Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Urban by Decree: Morphological Reality versus Administrative Boundaries Across 12 Million Nepali Buildings

Diwash Mainali

diwash006@gmail.com

Research Article

Keywords: urban morphology, administrative boundaries, percolation, urban scaling, MAUP, Nepal

Posted Date: November 12th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11463872/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Urban by Decree: Morphological Reality versus Administrative Boundaries Across 12 Million Nepali Buildings

Anonymous Author[1*]

[1]Author details withheld for double-anonymous review.

Corresponding author(s). E-mail(s): anonymous@example.org;

Abstract

Nepal's 2015–2017 restructuring reclassified roughly two-thirds of its citizens as urban overnight, on legal status rather than built form. We audit that decree against morphology using a civic geodatabase of 12.16 million building footprints across all 753 local-level governments, delineating dense cores by percolation of a 50 m building lattice, anchored against the Global Human Settlement Layer degree-of-urbanisation grid and 2021 census ward tables. Decree and reality disagree in both directions: about a third of legally urban units hold no dense core dominant over their municipal territory, and about a third of legally rural units contain one. The external anchor supports the fine-grained classifier, and converting the audit to people bounds the stakes: 31.1% of the population (27–46% under external reclassification) lives in a misclassified unit, and the morphological classes track independent census dwelling density far better than the legal label does. The disagreement extends into inference: administrative road-scaling exponents separate along the official Terai-to-mountain gradient whereas morphological exponents do not, and the cross-system mountain-bearing divergences are the only contrasts surviving Holm correction within a confirmatory family fixed before the final analysis. Administrative boundaries are not a neutral container: they distort both who counts as urban and what scaling laws say.

Keywords: urban morphology, administrative boundaries, percolation, urban scaling, MAUP, Nepal

1 Introduction: Who Counts as Urban?

When a state declares a territory a "city," two things happen at once: a legal status is created—with fiscal transfers, service mandates, and a place in the national urban aggregate—and an implicit empirical claim is made, that somewhere inside that line there exists an urban *thing*: a dense, connected, built-up concentration of human activity. The first is an act of decree; the second a claim about morphology. Much of comparative urbanism assumes the two travel together. They need not.

Nowhere is the gap more consequential than in Nepal. Under the 2015 Constitution and the 2017 Local Government Operation Act, the country reconstituted itself into 753 local-level governments (LLGs), of which 293 were designated urban and 460 rural (*Gaunpalika*) (Agergaard, Subedi, & Brøgger, 2022); our gazetteer reconciliation matches 291 urban and 438 rural units to a legal status and leaves 24 unresolved (§ 3). The share of the population living in "urban" municipalities jumped to roughly two-thirds almost overnight—a reclassification driven by administrative redrawing rather than any observable change in the built environment (National Statistics Office, Government of Nepal, 2023). We test that claim directly: the 2021 census join confirms it in level (66.3% of the matched population lives in legally urban units) while overturning it in composition (only 63.9% in morphologically urban ones; 31.1% of people live under a mismatched classification). Whether the decree describes cities or merely names them is the question this paper answers with buildings.

Nepal is an extreme test case. Physiographically it traverses three belts in roughly 200 kilometres—Terai floodplain, Middle Hills, High Himalaya; administratively, the 2015–2017 restructuring recast every local map at once, and some 30 million people are now enumerated through these units. The gap between decree and form is therefore not a cartographic curiosity but a live fact about fiscal allocation and service entitlement.

We ask two linked questions. *First*, how much of "urban Nepal" is morphologically rural, and where are emergent high-density settlements still classified as administratively rural? *Second*, does the choice between administrative and morphological units change what urban science concludes—specifically, the scaling laws we fit to settlement systems? We address both from a single national ground truth: 12,162,781 building footprints, 834,790 road segments, and 501,197 points of interest (POIs) across 6,743 constitutional wards, in a reproducible DuckDB pipeline.

We contribute two findings and one artifact. (1) A *de facto* definition of the urban operationalised from building morphology by spatial percolation, cross-tabulated against *de jure* status for all 753 LLGs: 38.5% of legally urban units carry no dense core that qualifies against their territory and 32.2% of legally rural units contain one; the association is far above a random-label null but only weak-to-moderate (χ^2_{Yates}=59.85, p=1.0×10^{-14}; Cramér's V=0.29). Against the external GHSL degree-of-urbanisation anchor the legal map's misalignment is 28.7%, and our classifier agrees with the anchor at least as well as the legal map ($\kappa = 0.40$ versus 0.38; paired difference not significant), with concordance degrading to 0.03 on coarser lattices even at each lattice's own optimal radius. The census join converts the audit to people: 31.1% of the population (27–46% under external reclassification) lives in a misclassified unit, and morphological classes track independent census dwelling density ($\rho = 0.86$) far better than the legal label ($\rho = 0.44$). (2) The disagreement extends from classification into *inference*:

administrative road-scaling exponents diverge along the belt gradient—sub-linear Terai, near-isometric Hills, strongly super-linear Mountains—whereas morphological exponents show no significant belt divergence anywhere; the cross-system mountain-bearing divergences are the *only* members of our declared confirmatory family to survive Holm correction ($p = 0.044$). At the mountain periphery, the boundary definition silently determines whether a foundational regularity of urban science appears to hold. (3) An aggregate-only, coordinate-free replication package with an offline self-test, re-runnable end-to-end without redistributing sensitive geometry.

Answering with buildings rather than census categories is the methodological core: the measured quantity—presence, position, and density of roofed structures—is exactly what the administrative claim asserts. The corpus shares the provenance and biases of volunteer geographic information (Goodchild, 2007; Haklay, 2010), whose building-layer completeness is systematically lower and more uneven across the Global South than in the mapped core (Hecht, Kunze, & Hahmann, 2013; Herfort, Lautenbach, Porto de Albuquerque, Anderson, & Zipf, 2021, 2023); we bound that bias against an independent global built-up-area product (§ 5.7) and treat the residual divergence as the object of study.

2 Theory: Three Literatures and Their Intersection

The political geography of demarcation. A first literature treats the urban boundary as a political object. Agergaard et al. (2022) document how Nepal's restructuring demarcated municipal boundaries for administrative and fiscal reasons bearing only loose relation to settlement reality; the "planetary urbanization" program argues that the urban/rural distinction is an analytical fiction imposed on a continuous process (Brenner, 2013; Brenner & Schmid, 2015); and the South Asian "subaltern urbanisation" school (Denis & Zérah, 2017; Roy, 2011) shows autonomous market towns growing *outside* state recognition while rural hinterlands are amalgamated into cities—precisely the two error modes we quantify as "urban-in-rural" and "urban by decree." What this literature lacks is a national, exhaustively enumerated, reproducible *measurement* of the decree–reality gap; Nepal's fresh, complete set of 753 LLG boundaries supplies it. The stake generalises because administrative definitions of "urban" remain the norm worldwide and cross-national commensurability is a recognised open problem: UN harmonisation re-bases urban aggregates onto population-defined units, and DEGURBA exists precisely because administrative boundaries distort comparisons.

Morphological cities, percolation, and the degree of urbanisation. A second literature defines cities by physical form, now measurable from above (Seto, Güneralp, & Hutyra, 2012). The conceptual root for turning built surface into a *city* is *percolation*: a settlement exists as a connected whole once its built patches link above a critical proximity—formalised for cities by Arcaute et al. (2015) and extended to regional scale by Fleischmann, Feliciotti, Romice, and Porta (2022), with long cellular antecedents (Batty & Xie, 1994) and a fractal warrant for reading extent, density, and connectivity as one object (Batty, 2008). National and global building-footprint layers have made morphometric delineation routine (Jochem et al., 2021; Jochem & Tatem, 2021); in parallel, the European Commission / UN "Degree of Urbanisation"

(DEGURBA) method classifies 1 km grid cells into cities, towns, and rural areas by density and contiguous built-up extent, decoupling the "urban" from administrative borders (Dijkstra et al., 2021; United Nations Statistical Commission, 2020). Our index is a building-footprint realisation of that logic for a data-scarce national context: we percolate a fine building lattice rather than a coarse population grid, treating link radius, density floor, and area share as explicit, swept decisions.

Urban scaling laws and the unit-of-analysis problem. A third literature supplies the quantitative stake: socio-economic outputs scale super-linearly and infrastructure sub-linearly with population, $Y \propto N^{\beta}$, with canonical exponents $\beta \approx 0.85$ for road length and $\beta \approx 1.15$ for interactive activity (Bettencourt, 2013; Bettencourt, Lobo, Strumsky, & West, 2007; Rozenfeld et al., 2008). These laws are asserted as *properties of cities as functional organisms*, not of the map units a statistician happens to have drawn; what counts as "the city" is itself contested (Batty & Ferguson, 2011; Louf & Barthelemy, 2014), and fitted exponents shift systematically with delineation practice (Cottineau, Hatna, Arcaute, & Batty, 2017). The modifiable areal unit problem (MAUP) has warned for four decades that the aggregation zone alters regression outcomes (Fotheringham & Wong, 1991; Openshaw, 1984). We reframe boundary choice as part of the estimand—an allometric exponent is defined only relative to a unit, and a unit splicing an urban core to agrarian hinterland is a different economic object from the core. If administrative units merely added noise, belt-specific exponents would scatter around the morphological value; instead they separate into belt-dependent regimes, the signature of territorial confounding rather than sampling error.

The three literatures converge on one testable proposition: if administrative boundaries are politically authored containers rather than functionally measured organisms, then (a) a substantial, systematically patterned share of legal units will disagree with morphological evidence in both directions, and (b) regularities asserted to hold "of cities" will be unstable across zone choices in ways that vanish under functional delineation. We test (a) with the mismatch matrix and its permutation null, and (b) with belt-stratified allometric fits on both unit systems simultaneously. Neither test has, to our knowledge, been run exhaustively on a national system of decreed local units before.

3 Data: A National Civic Geodatabase

Since 2017 Nepal's local map is a two-level mosaic: 753 LLGs nested within 77 districts and 7 provinces, subdivided into 6,743 constitutional wards (our spatial crosswalk dissolves the extract's 34,529 ward-tier polygon fragments to the same architecture). Unit scale spans more than two orders of magnitude, from a 6.6 km^2 metropolitan core to mountain municipalities exceeding 2,400 km^2, and the settlement fabric is bimodal: dense nucleations embedded in a diffuse agrarian dwelling matrix—so a single "urban" label requires an explicit connectivity and density definition that administrative classification never supplies.

Analysis runs against a single DuckDB file of a volunteer civic-mapping extract. `buildings` holds 12,162,781 footprints (polygon geometry, tile key, 6-decimal centroid); `roads` 834,790 polylines; `pois` 501,197 points with a 191-category taxonomy and administrative attributes; a `boundaries` table holds 58,459 polygon records (one row

per geometry part of each multipolygon) across administrative levels 2–9, of which the 34,529 level-9 rows are ward-tier fragments. Table 1 reports the canonical counts; every statistic is re-derived from scratch by the replication harness, with computed tables in a second database leaving the raw extract read-only.

Legal status is the *de jure* axis. From 774 geoBoundaries ADM3 polygons (Runfola et al., 2020), twenty-one conservation or special-zone polygons are flagged out of the universe, leaving the constitutional **753 LLGs**. Status follows the name suffix (*Mahanagarpalika*/*Upa-*/*Nagarpalika* = urban; *Gaunpalika* = rural); a conservative, injective, district-constrained rematch restores two spelling-variant metropolitan units (Pokhara Lekhnath, Janakpurdham), leaving 729 LLGs with a legal status (291 urban, 438 rural) and **24** documented-unresolved—listed with districts in Appendix A rather than hand-curated, since a fuzzy cross-district match could silently mis-assign.

Two extraction properties are enforced in code. First, an *extraction-zoom* confound: buildings and roads were decoded at tile zoom 14 (≈0.53 m), so they carry all continuous geometry; POIs were cached at zoom 5 (longitude pitch ≈271 m), so we *never* compute sub-ward point patterns or building-to-POI Euclidean distances—POIs enter only through containment (point-in-polygon; lattice-cell identity against percolation clusters, where a core spans many square kilometres so a ±271 m displacement almost always resolves to the same cluster). Second, a redistribution guard: the replication package exports only LLG-level aggregates and per-cell statistics on 250 m, 500 m, and 1 km UTM index lattices—never coordinate arrays or geometries—asserted on every run by an ethics coordinate-scanner (a leak heuristic, not a proof; § 6.3).

4 Methods

4.1 Grid morphology and the percolated urban index

We rasterise footprints onto lattices at 250 m, 500 m, and 1,000 m, summing building counts and area per cell; the reconstruction is exact, per-cell totals summing to the national 12,162,781 at all three resolutions (§ 5.6). Each occupied cell is attributed to the LLG containing its centroid; coverage is 99.4–99.5%, the residual protected and unincorporated land (inside one of the 753 LLG polygons: 99.15%). Percolation itself runs on a finer 50 m building-centroid lattice.

A city is a *connected* built mass, so we apply bond-percolation (connected-component) analysis to the occupied lattice, linking built cells whose centroids lie within radius r. Sweeping r gives the percolation curve (Figure 2): the largest-cluster fraction rises sharply near a critical radius—the connectivity transition used to delineate cities operationally (Arcaute et al., 2015). We freeze $r_c = 500$ m at the steepest rise of the reference-zone (UTM 45N) curve, where the largest-cluster fraction is 0.789. Nationally (both UTM zones, 50 m lattice) percolation yields 2,870 clusters (minimum 4 cells; the reference zone alone gives 1,162); only 197 clear the dense-core threshold. An LLG is *morphologically urban* when a cluster within it holds at least 1,000 buildings and at least 4 cells *and* that core occupies at least $\tau = 0.10$ of the LLG's *total land area*—a bivariate rule of absolute floor plus relative territorial share.

Three disclosures. "Urban by decree" does *not* mean core absence: all 291 legally urban units contain a ≥1,000-building core (mean dense-core area 16.8 km^2, falling to

11.4 km^2 among the 112 flagged units); they fail because the core occupies under 10% of a territory drawn far beyond the settlement—territorial *dilution*. Freezing past the transition admits giants: 13 of the 197 cores exceed 100,000 buildings and hold 11.0 M of the 11.9 M buildings attributed to any core, the largest spanning 5.73 M structures and 22,898 km of road; excluding all $N > 100{,}000$ clusters leaves the pooled morphological road exponent essentially unchanged (1.112→1.100) while R^2 falls 0.85→0.64 (§ 5.7). And lattice quantisation bounds contact distance to ±50 m of nominal r_c, absorbed by the 250 m-spaced sweep.

4.2 The mismatch matrix

Cross-tabulating legal status against morphological reality over 729 LLGs yields a 2×2 plus an unresolved row, with cells named **True Urban**, **Urban by Decree** (legally urban, no core clearing the τ test), **Urban-in-Rural** (legally rural, qualifying core—the subaltern settlement), and **True Rural**. We report counts, buildings, footprint and land areas, error rates, χ^2 with Cramér's V, and Cohen's κ with bootstrap intervals—in a 2×2, V equals the Matthews coefficient in absolute value, so only κ is carried forward.

4.3 Urban scaling: administrative versus morphological units

We estimate Bettencourt scaling (Bettencourt, 2013; Bettencourt et al., 2007) as $\ln Y = \ln Y_0 + \beta \ln N + \varepsilon$, with N the building count and Y either road-network length (infrastructure, expected $\beta < 1$) or commercial POI count (activity, expected $\beta > 1$), fitted by OLS and by Huber robust regression (Huber, 1981) on *administrative* units (LLG aggregates) and *morphological* units (unique dense-core clusters). To expose the MAUP directly we stratify both fits by belt (Terai, Hills, Mountains) and test cross-belt equality of the road exponent for *all three* belt pairs on both systems—never a selected contrast—adding two-one-sided equivalence tests (TOST) at prespecified margins (±0.10–±0.30), since failing to reject a difference is not establishing equality, and a cross-system difference-in-differences (DiD) testing the two systems against each other. The contrast between the belt-spreads is the core theoretical result.

4.4 Operationalisations

Only units with ≥200 buildings enter the fits—a non-binding floor administratively (all 753 clear it; smallest 446 buildings; the road fit uses $n = 742$, dropping 11 units without mapped street-class length) and morphologically (all 197 cores clear it; smallest 1,002 buildings; fits use the 163 cores with street segments and the 141 with a commercial POI, zero outputs excluded). Road output is total geodesic length of street-class segments, attributed whole to the unit containing its *midpoint*, projected into that midpoint's UTM zone (44N/45N): a road whose midpoint falls outside a core contributes nothing to it, while a highway bisecting a vast mountain polygon contributes its whole length—so we test rather than concede the shortcut (geometric clipping moves the mountain exponent by 0.014 and leaves the gradient intact, § 5.6). Activity output is commercial POI counts matching a fixed pattern list, entering only through containment (§ 3). Belts are centroid-latitude proxies (Terai below 27.2°N,

Hills below 28.4°N, Mountains above; the four far-western plain districts—Banke, Bardiya, Kailali, Kanchanpur—forced to Terai); because physiography runs NW–SE while the cut is horizontal, the misclassification is *systematic*, so the claim rests on the administrative-versus-morphological *contrast*, never the *level* of a single belt cell. We additionally rebuild the official CBS crosswalk privately (LGOA-2017 district-to-zone table under the 2021 census three-belt convention, gated against published belt population shares—Terai 53.66%, Hill 40.25%, Mountain 6.09%) and release only aggregate outcomes: every cell, contrast, and DiD is refitted under both definitions (§§ 4.5d, 5.7).

4.5 Sensitivity controls and external validation battery

Internal controls (§ 5.6): link-radius sweep $r_c \in \{250, 500, 750, 1000\}$ m; threshold sweep $\tau \in \{0.05, 0.10, 0.20\}$; absolute-floor sweep $\{500, 1000, 2000\}$ buildings; a random-label null permuting legal status across the 729 units 5,000 times with marginals preserved (291 legal urban, 320 morphological—both structural, only their pairing stochastic), with a district-block variant (within-district shuffling) as the conservative null; a grid-MAUP rebuild of the full matrix at 250 m and 1 km; and the clipping control.

An external battery anchors, converts, bounds, and corrects (scripts 12–19, run against the frozen databases and exporting only aggregate tables). *(a) Anchor*: the GHSL-SMOD R2023A 1 km grid (Pesaresi et al., 2024; Schiavina, Melchiorri, & Pesaresi, 2023) (md5-pinned at fetch) flags each LLG externally urban under two readings—DEGURBA-conformant (contains an urban-centre or dense-cluster cell) and a τ-analogue ($\geq$10% of land in such cells)—cross-tabulated with both classifications (unweighted κ), and arbitrates the lattice question by scoring the classifier rebuilt at 50/250/1,000 m. *(b) Census join*: NSO 2021 ward population tables name-matched to the 753 LLGs (exact, district-disambiguated, then fuzzy with an audit trail; unmatched units excluded, never guessed), yielding population-weighted misclassification and the legal/morphological urban *population* shares. *(c) Completeness*: GHS-BUILT-S R2023A built-up area (Pesaresi & Politis, 2023) as a mapping-intensity covariate; footprint-to-benchmark ratio, rank concordance, and the matrix recomputed within terciles of benchmark built density. *(d) Belts*: all scaling cells and contrasts refitted under the official district-to-zone classification with i.i.d. and district-clustered standard errors. *(e) Spatial diagnostics*: Moran's I on k-nearest-neighbour belt-residual weights; a district-block permutation null for the matrix; a district-block bootstrap of the DiD so the two systems' belt-gap covariance enters empirically. *(f) Delineation*: per-cluster diagnostics for all cores plus a re-delineation trimming every core below the first quartile of its 5×5 local building density and refitting on the fragments. *(g) Multiplicity*: Holm step-down within the declared confirmatory family—6 administrative belt-pair contrasts (both definitions) + 12 cross-system DiDs (both definitions $\times$ both variance schemes) = 18 tests—plus a declared 12-test delineation-sensitivity family, fixed before the final analysis rather than externally pre-registered; the POI axis stays exploratory. *(h) Second-round checks*: paired bootstrap of the anchor-kappa difference; the kappa triangle and census-density concordance bounding the anchor's externality; the population band under both readings; the lattice re-sweep at each pitch's own optimal radius. *(i) Tool transparency*: the analysis code and parts of the manuscript

text were drafted with a large-language-model coding assistant; every reported statistic is independently recomputed from the frozen databases and all AI output was reviewed and verified by the author.

5 Results

5.1 The decree–reality gap is large and systematic

Table 2 presents the matrix. Of 729 classified LLGs, 179 are True Urban and 297 True Rural, but the off-diagonal cells are far from residual: **112** legally urban municipalities contain no dense core clearing the $\tau = 0.10$ test—*urban by decree*, a false-positive rate of $112/291 = 38.5\%$—and **141** legally rural municipalities host one—*urban-in-rural*, $141/438 = 32.2\%$. Combined misalignment is $(112 + 141)/729 = 34.7\%$, a point on a band spanning 33.5–43.3% over the τ sweep and 33.6–36.8% once the unresolved 24 are adjudicated at the extremes. The mass distribution is telling: the 112 urban-by-decree municipalities hold 1.73 M buildings spread across 23,184 km^2, while the 141 urban-in-rural settlements pack 1.85 M into 9,955 km^2—denser, connected, legally invisible.[1]

Table 1 Canonical ground-truth counts (DuckDB).

quantity	value
buildings	12,162,781
roads	834,790
POIs	501,197
constitutional LLGs	753
legal urban	291
legal rural	438
unresolved	24
protected excluded	21

The association is real but weak: $\chi^2_{\text{Yates}} = 59.85$, $p = 1.0 \times 10^{-14}$, Cramér's $V = 0.287$ (uncorrected $\chi^2 = 61.03$, $p = 5.6 \times 10^{-15}$); $\kappa = 0.288$ (bootstrap 95% CI 0.22–0.36).[2] Marginals preserved, a random-label permutation expects urban-by-decree 163.2 (s.d. 6.5) and urban-in-rural 192.2 off-diagonals (simulated means 163.4/192.4)—*above* the observed 112/141; the district-block null lowers expectations to 138.3/167.3, still above. Decree and morphology are *more* concordant than chance, but only weakly-to-moderately, leaving a third of units in systematic—not random—disagreement. Figure 1 locates the gap: the two maps agree along Middle-Hills ridge towns and the Terai corridor but diverge in the interior hills and far-west.

Medians place the off-diagonal cells well past the threshold: core share of land has median 0.150 among True Urban but 0.067 among urban-by-decree units (True Rural:

[1] The four cells account for 11,809,829 of the national 12,162,781 buildings; with the 249,874 buildings of the 24 unresolved units the reconciled total is 12,059,703, leaving 103,078 (0.85%) outside any LLG polygon (§ 3).
[2] Yates-corrected and uncorrected statistics agree substantively.

Table 2 Legal-decree × morphological-reality mismatch matrix. "Built footprint" is the summed polygon area of building footprints inside each cell; "Land area" is the municipal territory (polygonised LLG area), the τ-test denominator.

Cell	LLGs	Buildings	Footprint (km^2)	Land (km^2)
True Urban	179	5,809,053	378.770	20,230
Urban by Decree	112	1,730,725	98.500	23,184
Urban-in-Rural	141	1,854,909	101.060	9,955
True Rural	297	2,415,142	123.350	82,815
unresolved	24	249,874	13.830	6,899

0.050), with urban-in-rural at 0.131. Sitganga (Arghakhanchi), designated "urban," spans 611 km^2 with 2.2% of its land in any core; Thuli Bheri (Dolpa) reaches 0.2%; at the other pole Kathmandu packs 88.7% of its land into one core and Lalitpur 75.8%. The urban-in-rural exemplars are the subaltern market towns the literature predicts—Badhaiyatal (28,819 buildings), Khajura (28,637), Baijanath (27,048), Janaki (26,312), Buddhashanti (25,860)—each holding more buildings than the median legal-urban unit (19,945). In aggregate, 23% of the legal-urban building stock (1.73 M) sits in diluted units and 43% of the legal-rural stock (1.85 M) in rural units carrying a qualifying core.

5.2 Percolation defines the urban threshold

Figure 2 shows the curve: the largest-cluster fraction climbs gradually through $r = 150$–400 m, then steps up steeply around $r_c = 500$ m (fraction 0.789). Below, the surface fragments into tens of thousands of micro-clusters (71,339 in the reference zone at $r = 50$ m); above, counts collapse toward congealed metropolitan systems (1,162 at $r = 500$ m, 279 at $r = 1000$ m in the reference zone; 2,870 nationally at the frozen radius). We claim no critical-exponent physics for the freeze, only that the headline is stable across the swept band (§ 5.6): r_c is a documented operating threshold, not an estimated phase transition. The curve also describes Nepal's settlement system: counts thin gradually rather than congealing nationally even at $r = 1500$ m (built surface saturating at ≈2.07 M occupied 50 m cells)—a mosaic of intermediate regional systems, the ridge- and market-town pattern of a hill country, not a single primacy-driven conurbation.

5.3 Scaling regularity survives only on morphological units

Table 3 and Figure 3 report the fits. Administratively, pooled road scaling is super-linear (OLS $\beta = 1.104$, s.e. 0.033, $R^2 = 0.60$, $n = 742$), above the near-isometric value theory predicts, falling to plausible only under robust downweighting (Huber $\beta = 0.928$); commercial POIs are strongly super-linear ($\beta = 1.372$, $n = 753$). On morphological units the fit is far tighter and belt-stable—though *not* closer to theory on sign: road scaling is near-isometric to slightly super-linear (OLS $\beta = 1.112$, s.e. 0.037, $R^2 = 0.85$, $n = 163$; Huber $\beta = 1.084$, s.e. 0.030); the morphological advantage is the exponent's low dispersion and cross-belt *stability*, not fidelity to a canonical value. Commercial activity on cores settles at a moderate but unambiguously super-linear level (OLS

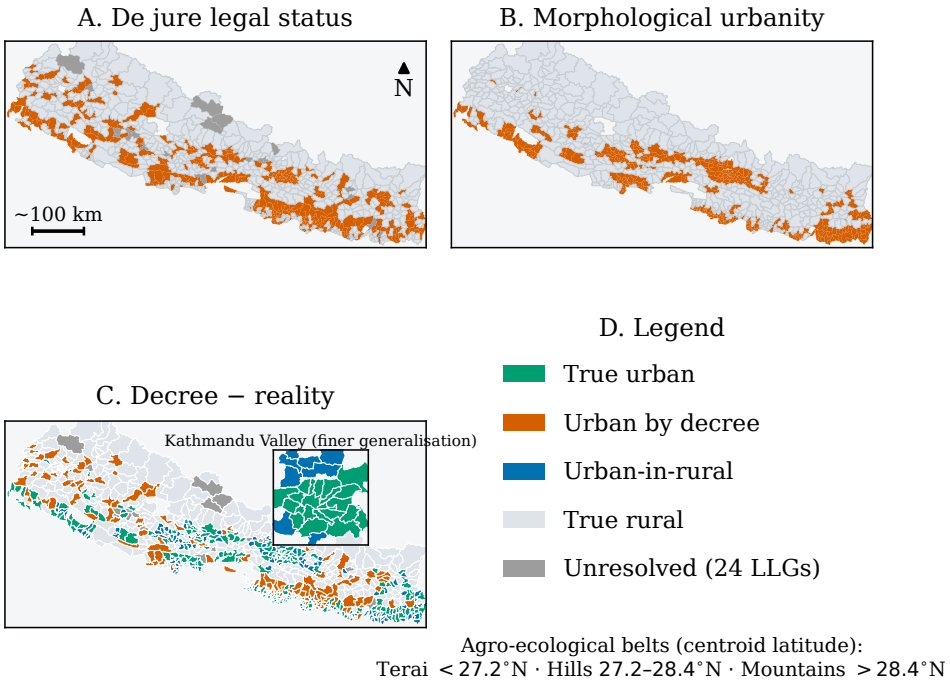


Fig. 1 National legal versus morphological urbanity, laid out as a 2×2 grid; categories use the Okabe–Ito colourblind-safe palette. (A) *de jure* legal status of the 753 local-level governments (vermillion = legally urban). (B) *de facto* morphological urbanity from percolated building cores ($r_c = 500\,\text{m}$, $\tau = 0.10$; vermillion = morphologically urban, pale = rural, grey = unresolved). (C) the four-cell decree−reality mismatch, with a Kathmandu Valley inset drawn at finer generalisation; (D) the shared legend, which adds the grey unresolved (24 LLG) class. The divergence between panels A and B is the decree–reality gap. Aggregated administrative polygons only, generalised to $\approx 0.004°$ ($\approx 400\,\text{m}$) nationally; no building, road, or POI geometry is drawn.

$\beta = 1.218$, $n = 141$; Huber 1.252). We frame the slightly super-linear road exponent as an *open empirical anomaly* in building-count-based infrastructure scaling—candidate mechanisms (leapfrog arterial sprawl; vertical densification, since N counts buildings not floor area) are suggestive but untested. That the *order* of exponents (activity above infrastructure) survives both unit choices while *magnitudes* move with the choice is the MAUP pathology: qualitative theory is robust to delineation, quantitative calibration is not, and policy thresholds built on calibrated values inherit the boundary's politics.

5.4 The MAUP headline: administrative belt divergence, morphological belt consistency

The central evidence stratifies the road exponent by belt (Table 4; all pairs on both systems in Table 5A). Administrative exponents rise monotonically: Terai $\beta = 0.895$

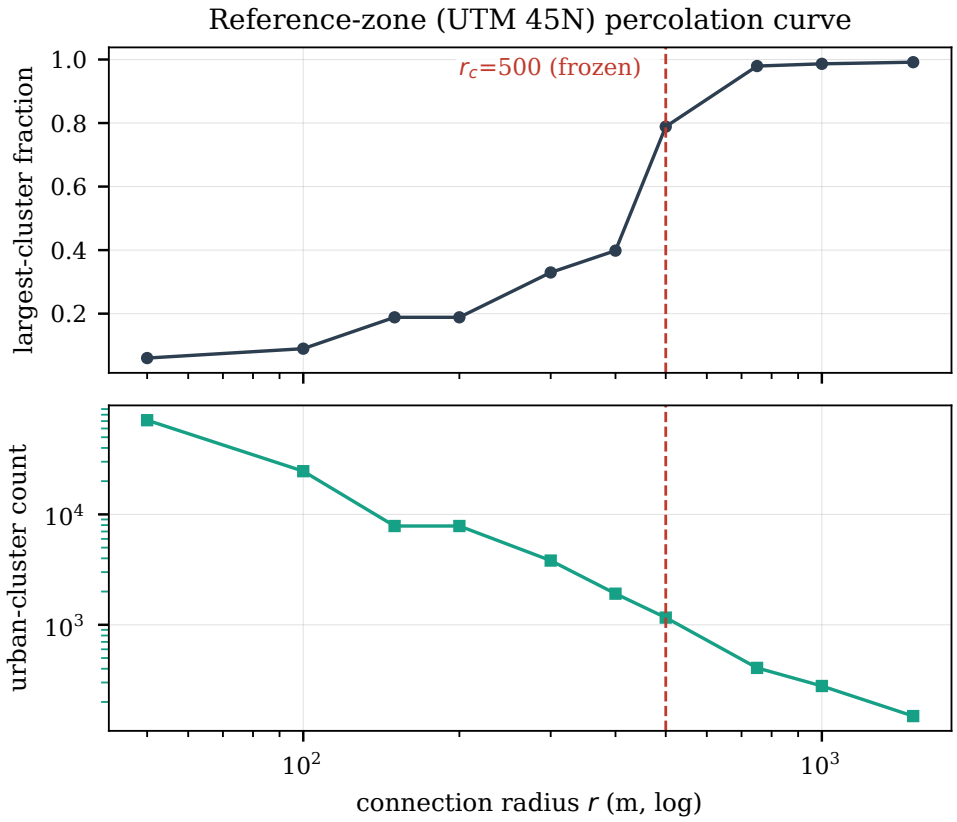


Fig. 2 Reference-zone (UTM 45N) percolation curve, shown as two stacked panels sharing one radius axis: largest-cluster fraction (top) and urban-cluster count (bottom) against link radius r; cluster counts here are for UTM 45N alone (the national two-zone count at r_c is 2,870, § 4.1). The steep rise near $r_c = 500$ m marks the connectivity transition used to delineate morphological urban cores.

(s.e. 0.032, $n = 265$), Hills 0.963 (0.041, $n = 324$), Mountains 1.391 (0.142, $n = 153$)—spanning the pivotal $\beta = 1$ and flipping its policy interpretation with the zone. Wald tests locate the divergence at the mountain endpoint: Terai–Hills $\Delta\beta = 0.068$, $p = 0.19$; Hills–Mountains $\Delta\beta = 0.428$, $z = 2.90$, $p = 0.004$; Terai–Mountains $\Delta\beta = 0.496$, $z = 3.42$, $p = 0.001$. Morphological exponents are flat: 1.117 (0.045, $n = 73$), 1.101 (0.062, $n = 46$), 1.054 (0.107, $n = 44$); no pair differs ($p = 0.84$, 0.70, 0.58), and the Terai–Hills contrast is statistically equivalent at ± 0.15 ($p_{\mathrm{TOST}} = 0.039$) and ± 0.20 (0.008) though not ± 0.10 (0.13); endpoint pairs reach equivalence only at ± 0.30 ($p \approx 0.02$), so we claim equivalence at the heartland and bounded smallness, not equivalence, at the endpoints. Symmetry stated plainly: the administrative Terai–Hills pair also clears ± 0.20 ($p \approx 0.006$), failing only the ± 0.15 margin the morphological pair passes—heartland equivalence is shared, not system-distinguishing. The R^2 gain (0.60→0.85) corroborates but is not the claim. Under the official 77-district belts the pattern *sharpens*: administrative $\beta = 0.961/0.916/1.540$ (n 276/348/118), the

Table 3 Urban-scaling regressions ($\ln Y = a + \beta \ln N$) for administrative LLGs and morphological dense-core percolation clusters (197 national clusters spanning both UTM zones; because zero-valued outputs are dropped from the log–log fits, the road fit uses the 163 clusters carrying at least one mapped street segment and the POI fit the 141 carrying at least one commercial POI). OLS and Huber (robust) on road length and commercial POI counts. On the morphological axes the OLS and Huber R^2 coincide to three decimals (0.851, 0.756)—not because the fits are identical but because the dense-core distribution carries no extreme-leverage outlier, so robust reweighting barely moves them (road-length β 1.112→1.084); on the administrative axes, dominated by a few very large units, the two estimators separate as expected (road R^2 0.601 vs. 0.586).

Unit	Output	Model	n	β	s.e.	R^2
Admin. LLGs	Road length	OLS	742	1.104	0.033	0.601
Admin. LLGs	Road length	Huber	742	0.928	0.020	0.586
Admin. LLGs	Comm. POIs	OLS	753	1.372	0.051	0.489
Admin. LLGs	Comm. POIs	Huber	753	1.375	0.051	0.489
Morph. clusters	Road length	OLS	163	1.112	0.037	0.851
Morph. clusters	Road length	Huber	163	1.084	0.030	0.851
Morph. clusters	Comm. POIs	OLS	141	1.218	0.059	0.756
Morph. clusters	Comm. POIs	Huber	141	1.252	0.061	0.756

Table 4 Belt-stratified OLS β by agro-ecological belt with small-sample Student-t 95% CIs in brackets; n = units with $N \geq 200$ and non-zero output. The administrative exponents rise monotonically from Terai to Mountains, significantly at the mountain endpoint on the i.i.d. Wald test (Hills–Mountains $p = 0.004$; Terai–Mountains $p = 0.001$; district-clustered and Holm-corrected status in § 5.7), whereas the morphological exponents are statistically indistinguishable across belts (all $\beta \approx 1.05$–1.12, no pair significant). The commercial-POI companion is demoted to Appendix Table B2: that axis tracks volunteer mapping intensity and must not be quoted as a scaling result.

Road length (β, ln**km**)			
Belt	Terai	Hills	Mountains
Admin β	0.895 [0.833, 0.957]	0.963 [0.882, 1.044]	1.391 [1.111, 1.672]
n	265	324	153
Morph β	1.117 [1.028, 1.206]	1.101 [0.977, 1.226]	1.054 [0.838, 1.269]
n	73	46	44

heartland pair inverting and falling away ($\Delta\beta = -0.045$, clustered $p = 0.60$) while both mountain-bearing pairs strengthen ($\Delta\beta = 0.623$, $p = 0.022$; 0.579, $p = 0.028$); the morphological small mountain cell moves 1.054→0.884 (n 44→28) exactly as its size warns, still with no significant pair (clustered $p \geq 0.237$).

The variation is mechanical: Terai units blend town cores with agricultural hinterlands while mountain municipalities are vast sparse slivers in which road length co-varies with territory rather than population—the same regression fits different economic objects by belt. The commercial axis behaves likewise but is a weaker instrument (it tracks mapping intensity; § 6), so its belt table sits in Appendix B.

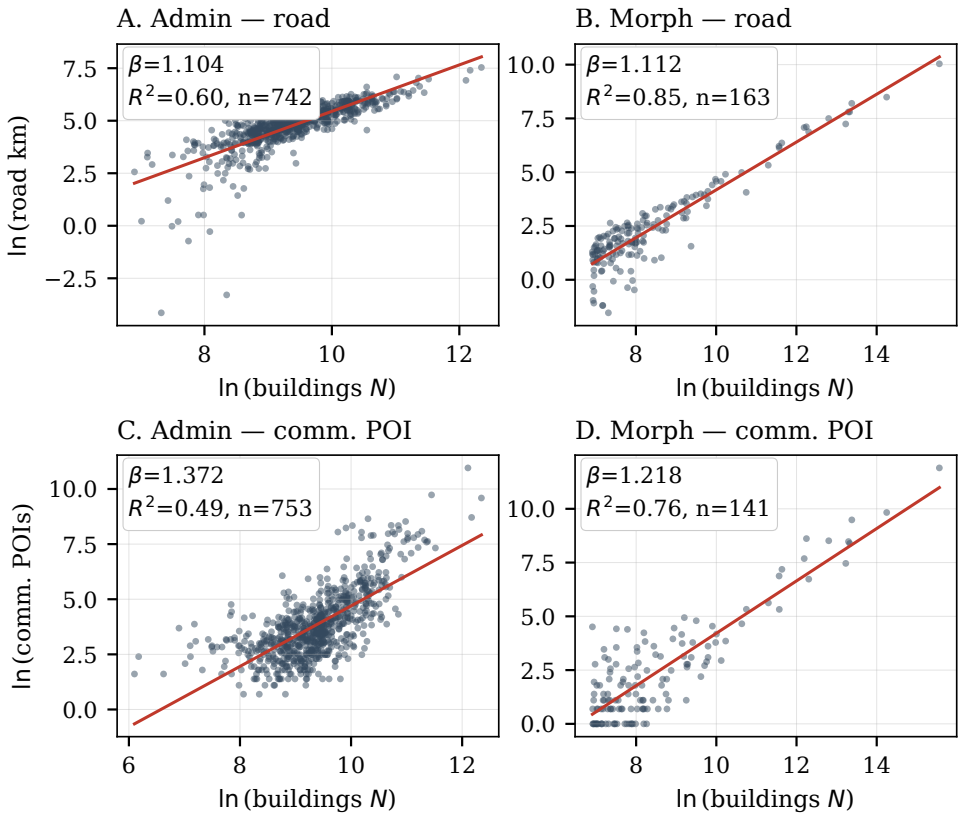


Fig. 3 Urban scaling on administrative (left) versus morphological (right) units, for both scaling axes: road length (top row) and commercial POIs (bottom row). Lines are OLS fits of $\ln Y = \ln Y_0 + \beta \ln N$; boxes report β, R^2, and n (full belt-stratified coefficients in Table 4). The morphological road fit is markedly tighter (R^2 0.60→0.85), while both systems' commercial axes are super-linear—though the POI axis additionally tracks volunteer mapping intensity (§ 6), so the substantive argument rests on the road axis.

5.5 Cross-system interaction, matched controls, and spatial dependence

"Significant versus not significant" is not a test of the difference between systems, so we test it directly for all three pairs (Table 5A). Terai–Hills: $D = 0.084$, s.e. 0.092, $p = 0.36$—no difference. The gradient-bearing pairs differ: Hills–Mountains $D = 0.476$, $z = 2.47$, $p = 0.013$; Terai–Mountains $D = 0.560$, $z = 3.01$, $p = 0.003$. The district-block bootstrap (§ 4.5e) finds small *negative* covariance between the systems' belt gaps ($\widehat{\text{Cov}}$ −0.007 to −0.046), so bootstrap standard errors exceed the quadrature value, by up to ∼46% (official Hills–Mountains 0.284 → 0.415); bootstrap p-values are the most conservative reported (proxy 0.48/0.11/0.066; official 0.61/0.039/0.046). With district-clustered administrative errors the proxy divergence attenuates but survives (Terai–Mountains $z = 2.06$, $p = 0.039$; Hills–Mountains $p = 0.088$; Terai–Hills null

Table 5 Cross-system interaction test and selection-matched control for the belt road exponent, recomputed live from the frozen replication database (`scripts/09` in the package). **Panel A**: every belt pair is reported, so no contrast is cherry-picked. The administrative gradient is concentrated at the mountain endpoint (Hills–Mountains and Terai–Mountains significant; Terai–Hills not), whereas no morphological pair differs and the Terai–Hills morphological difference is statistically equivalent within ±0.15 (TOST). The difference-in-differences is significant only for the pairs that carry the administrative gradient. **Panel B**: applying the paper's own dense-core screen to *administrative* units does not reproduce the morphological flat profile: the screened Terai–Mountains gap stays large, and on the Terai–Hills pair (Panel A row 1) the screened administrative gap even reverses sign—so the administrative divergence is not a rural-dilution selection artefact, though its direction under screening is unstable. The compact screen is omitted for lack of mountain units ($n < 10$), and the absolute-floor ≥1000 screen is omitted because it is non-binding—every unit clearing the $\tau = 0.10$ rule already carries ≥4,557 buildings, so its row is a byte-for-byte duplicate of the dense-core rule. n = units with non-zero road and building counts passing the screen.

A. Cross-system interaction (road $\Delta\beta$, all belt pairs)

Belt pair	Admin $\Delta\beta$ (s.e.)	p_{admin}	Morph $\Delta\beta$ (s.e.)	p_{morph}	DiD p
Terai – Hills	0.068 (0.052)	0.190	−0.016 (0.076)	0.835	0.363
Hills – Mountains	0.428 (0.148)	0.004	−0.048 (0.123)	0.699	0.013
Terai – Mountains	0.496 (0.145)	0.001	−0.064 (0.116)	0.583	0.003

B. Selection-matched control (administrative road β, screened, Terai–Mountains)

Screen	Terai β (n)	Mountains β (n)	$\Delta\beta_{M-T}$	p
Unscreened	0.895 (265)	1.391 (153)	+0.496	0.001
Dense-core rule	0.964 (163)	1.286 (15)	+0.322	0.118
Core ≥4 cells	0.895 (265)	1.291 (141)	+0.396	0.009
Morphological ref	1.117 (73)	1.054 (44)	−0.064	0.583

$p = 0.45$) and under official belts strengthens (Hills–Mountains $D = 0.855$, $p = 0.015$; Terai–Mountains $D = 0.781$, $p = 0.022$; Terai–Hills null $p = 0.50$). Holm step-down in the 18-test confirmatory family then decides what may be claimed: *only* the two official-belt clustered DiDs survive ($p_{\text{Holm}} = 0.044$ both). Everything else fails—all six proxy-belt tests ($p_{\text{Holm}} \geq 0.12$, including the formerly headline proxy administrative Terai–Mountains contrast at nominal $p = 0.043$), the block-bootstrap DiDs (0.118), and the official administrative mountain contrasts (nominal 0.022–0.028, $p_{\text{Holm}} = 0.065$). The headline is re-anchored to this: the unit systems separate along the gradient *under the official belts, with clustered inference and family-wise correction*; all other variants are suggestive (full ledger, Appendix Table B3).

Two threats are controlled. *Selection*: applying the paper's own dense-core screen (plus variants) to *administrative* units and refitting (Table 5B; Appendix B) shows screening reorganises rather than flattens the profile—the mountain-end excess survives every screen in *magnitude* ($\Delta\beta$ +0.32 to +0.40) but loses *statistical* resolution as the cell thins to 15 units ($p = 0.12$; DiD $p = 0.10$), and the Terai–Hills ordering reverses to −0.153 ($p = 0.003$) under screening, a fragility on which we rest no claim. *Spatial dependence*: district-clustered standard errors ($n = 742$, 73 districts) leave point estimates identical and inflate belt s.e. by 47/65/70% (Terai/Hills/Mountains); Terai–Mountains survives ($p = 0.043$), Hills–Mountains turns marginal ($p = 0.087$), Terai–Hills null ($p = 0.41$). Residual autocorrelation is large administratively (Moran's $I = 0.26$–0.28 pooled, $p \leq 0.0002$; Terai ≈0.47), small morphologically ($I = 0.18$, $p = 0.0008$ proxy; $I = 0.07$, $p = 0.065$ official; § 5.7).

5.6 Sensitivity and null controls

Figure 4 collects the sweeps. (i) *Link radius*: urban-by-decree counts $\{128, 112, 109, 109\}$ over $r_c \in \{250, 500, 750, 1000\}$ m—stable across a 4× range. (ii) *Threshold*: urban-by-decree swings $\{25, 112, 233\}$ over $\tau \in \{0.05, 0.10, 0.20\}$, urban-in-rural oppositely $\{291, 141, 11\}$; because the cells move oppositely, *total* misalignment spans only 33.5–43.3%. (iii) *Absolute floor*: over $\{500, 1000, 2000\}$ buildings, urban-by-decree moves only $\{111, 112, 114\}$, every other cell by at most six units. (iv) *Grid MAUP*: rebuilding the entire matrix at 250 m and 1 km, per-LLG totals are rank-consistent (Spearman $\rho = 0.999$, 0.998, 0.998) *but* headline misalignment moves with the lattice: 34.7% at 50 m, 52.0% at 250 m, 56.7% at 1 km—coarse cells smear dispersed dwellings into qualifying patches, inflating urban-in-rural 141→413 while urban-by-decree collapses toward zero. The gap's existence is lattice-robust; its magnitude and dominant error mode are construction-dependent—exactly the thesis that the unit choice, legal *or* lattice, is part of the estimand. We freeze the 50 m lattice and the external anchor arbitrates (§ 5.7): agreement $\kappa = 0.40$ at 50 m but 0.09 and 0.03 at the coarser pitches, so the inflated coarse headlines are grid smearing, and the legal-versus-external misalignment (28.7%) bands the quoted headline. Two further controls: geometric clipping of roads moves levels not slopes (mountain β 1.391→1.377; contrast 0.496→0.492; median clipped-to-midpoint ratio 0.67–0.76; Appendix B); and the building-share alternative to τ is degenerate—every cell's median unit places >96% of its buildings inside some core and none of the 112 falls below a 10% share, so the classifier's discriminating content is territorial dominance (Appendix B).

On multiplicity, the district-clustered proxy Terai–Mountains contrast (nominal $p = 0.043$) does *not* survive correction ($p_{\text{Holm}} = 0.129$); the claim carrying the scaling argument is the official-belt cross-system DiD ($p_{\text{Holm}} = 0.044$; §§ 5.5, 5.7). All other p-values are descriptive.

5.7 External benchmarks anchor the audit

Degree-of-urbanisation anchor. Against reading A, the legal map agrees with GHSL-SMOD at $\kappa = 0.38$ and our morphological classifier at $\kappa = 0.40$; the paired difference on the same 729 units is not distinguishable from zero ($\Delta\kappa = +0.015$, 95% CI −0.07 to +0.10, two-sided $p = 0.733$), so the defensible claim is *at least as well as*, not better. Legal misalignment against the anchor is 28.7% (131 legally-urban units with no external city; 78 legally-rural holding one), the same order as the internal 34.7%; under reading B agreement falls ($\kappa = 0.16/0.26$; 36.2% misalignment), since SMOD's classes fit functional urban areas, not land-dominance rules. The anchor thus puts the gap in a 28.7–36.2% band containing our headline. Its provenance is disclosed: the instruments agree with SMOD at 0.38–0.40 but with each other only at $\kappa = 0.288$, and SMOD is itself satellite built-up area—so $\kappa = 0.40$ bounds, not establishes, external concordance. The form-independent datum is the census: morphological core share ranks with census dwelling density at $\rho = 0.86$ against the legal label's 0.44—on a variable no satellite product touches, morphology tracks people far better than the

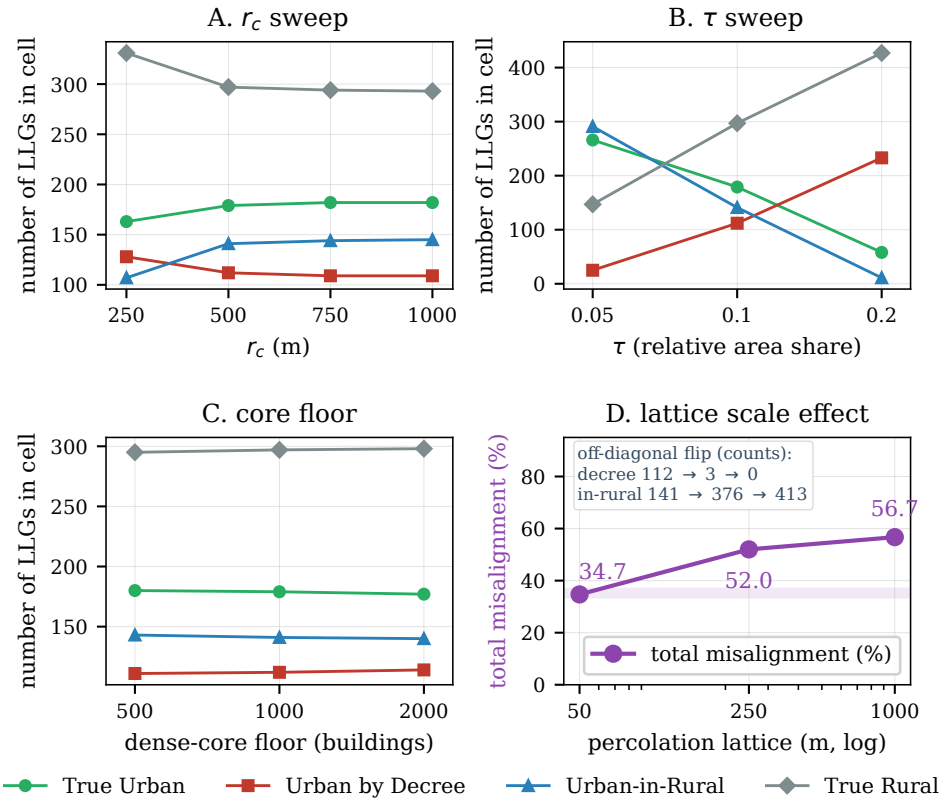


Fig. 4 Sensitivity controls. Panels A–C: the four mismatch cells under the link-radius (r_c), relative-threshold (τ), and absolute-core-floor sweeps. Panel D: headline misalignment from rebuilding the full matrix on the 50/250/1,000 m lattices (the grid-MAUP panel; these *classification* lattices are distinct from the 250 m–1 km attribution grids of § 4.1); points carry one-decimal values (34.7/52.0/56.7%), the shaded band is the ±unresolved-24 bound at 50 m; the composition flip (urban-by-decree 112→3→0 against urban-in-rural 141→376→413) is annotated in panel D. Panels are full matrix rebuilds at the frozen rules, not re-thresholdings. Stability holds for the geometric rules (A, C) and the τ-total (B); panel D is a lattice *scale effect*, reported honestly in § 5.6.

decree[3]. Lattice arbitration: rebuilding the classifier at each pitch and re-sweeping *each lattice at its own best radius* (500 m is sub-pixel at a 1 km pitch) still degrades monotonically—best attainable κ 0.40 at 50 m, 0.09 at 250 m, 0.03 at 1 km—a genuine loss of resolution, not a connectivity artefact; the 50 m point is what the benchmark *supports as the operationalisation*, not a scale it *settles* as uniquely true.

Census join: the audit in people. The matched tables cover 28.42 M of the 2021 census population. Population-weighted, **31.1%** of people live in a misclassified unit

[3]Metric asymmetry, controlled: 0.86 uses a continuous regressor, 0.44 a binary label. Against the same census dwelling density the *binary* morphological class ranks at $\rho = 0.70$ ($n = 727$; script 19, `anchor_provenance.csv`)—attenuated by dichotomisation yet still far above the legal 0.44: the gap survives the scale-matched comparison.

(bounded **27.1–46.2%** under the SMOD readings): 16.8% (4.75 M) under urban-by-decree, 14.3% (4.06 M) under urban-in-rural. The two-thirds claim is validated in level and overturned in composition: 66.3% of the matched population lives in legally urban units (within 0.4 pp of the claim) but only 63.9% in morphologically urban ones, the 2.4 pp net of two ≈15 pp opposing flows. Population-weighted anchor agreement exceeds unweighted ($\kappa = 0.43/0.44$), so misclassification is not a small-unit artefact.

Completeness. Mapped footprint totals 715.5 km^2 against GHS-BUILT-S's 1,100.7 km^2 inside LLG boundaries (ratio 0.65; per-unit median 0.63, IQR 0.51–0.75; rank concordance $\rho = 0.91$; log-log slope 0.85, R^2 0.83): under-mapping is substantial but *close to uniform*, not selectively urban. Terciles of benchmark built density keep both error cells live in the best-mapped tercile (40.9%; 15 urban-by-decree, 86 urban-in-rural) while the least-built reads 16.9% (41 urban-by-decree, *zero* urban-in-rural)—the subaltern exemplars sit where the independent product also measures dense built-up area, and none appears where over-mapped ribbon could manufacture one. Ratio-of-logs exponents are insensitive to a common scale.

Belts and delineation. Under the official belts (Appendix Table B1), the administrative mountain exponent rises (1.391→1.540) and the morphological small cell falls (0.884, $n = 28$—at that n the minimum detectable cell-pair $\Delta\beta$ (two-sided $\alpha = 0.05$, 80% power) is ≈ 0.54 (Δ SE 0.194–0.196), the same order as the administrative mountain divergences (0.579, 0.623): bounded smallness, not equivalence): cell *levels* are belt-sensitive, the system *contrast* is not (no morphological pair significant, clustered $p \geq 0.237$; DiD strengthens, clustered $p = 0.015/0.022$). Per-cluster diagnostics: the 13 giants hold 11.00 M of 11.86 M core-attributed buildings (92.8%); the largest (5.73 M buildings) spans a 413 × 229 km box at 4.3% fill, covering **327 LLGs across 35 districts**—a corridor, not a settlement. The density-minimum trim dissolves the giants (largest fragment retaining 14.8% of the mega-corridor's buildings; 75,907 fragments, 7,582 above the fit floor) and re-anchors the morphological mountain cell at 1.030 ($n = 277$); the profile *remains* flat (official belts 1.130/1.097/1.030, clustered $p \geq 0.086$; the single nominal exception—proxy Hills–Terai $\Delta\beta = 0.074$, $p_{\text{Holm}} = 0.045$ in its sensitivity family—sits an order of magnitude below the administrative divergences). Morphology's flatness is no corridor artefact. Nor is it fragment pseudo-replication: thinning the trimmed fit to its largest fragment per core × district (487 of 7,582 fitting fragments) leaves the official-belt profile flat (β 1.152/1.125/1.204, n 100/71/27; all paired clustered $p \geq 0.575$; script 17, `trim_belt_contrasts.csv`).

Spatial ledger. The district-block permutation of morphological labels leaves observed off-diagonals (112, 141) far below within-district null means (138.0, 167.0; s.d. 4.9), $p < 1/5{,}000$; the Holm ledger (Appendix Table B3) confirms: of 18 confirmatory tests exactly two reject—the official-belt clustered DiDs at both mountain-bearing pairs ($p_{\text{Holm}} = 0.044$). Every headline claim is worded to that anchor.

6 Discussion

6.1 Decree, reality, and what the gap means

Two readings carry opposite policy valences. Measurement error—a noisy administrative map that better cadastre would fix—versus structure: the restructuring deliberately

distributed urban status across agrarian territories for fiscal equalisation and representation (Agergaard et al., 2022), so the "error" is the design. Our evidence favours the structural reading: the gap is *large* (34.7%), beyond what boundary generalisation would plausibly produce[4]; *systematic*, $V \approx 0.29$ against a null expecting *more* off-diagonal mass than observed; and *asymmetric in density* (1.85 M buildings in 9,955 km^2 rural-urban cores against 1.73 M stretched over 23,184 km^2 decreed towns)—subaltern urbanisation recognised late or never (Denis & Zérah, 2017; Roy, 2011), and municipalisation decreed early and not yet built.

The implication is composition, not net level: decreed-urban units hold 62.0% (7.54 M) of national footprints against morphological urbanity's 63.0% (7.66 M); 291 legally urban units against 320 morphologically—the decree map slightly *under*states urbanity in total. What a near-balanced net conceals are the gross flows: two-fifths of legal-urban units (1.73 M buildings, 14.2% of the national stock) with no qualifying core, a third of legal-rural units (1.85 M, 15.3%) harbouring one, and nearly nine million people mislabelled in each direction. The two-thirds figure is right in level and wrong in composition. DEGURBA-style methods (Dijkstra et al., 2021; United Nations Statistical Commission, 2020) exist to absorb exactly this; we quantify its magnitude for a complete national system where previous estimates rested on census-town lists.

The fiscal stakes follow: transfers blend population, land area, and cost factors, and unit category enters some grant components (Agergaard et al., 2022; Government of Nepal, 2017), while urban status attaches service and planning mandates premised on agglomeration economies whose core our measurement dilutes below the qualifying threshold in 112 of 291 urban units; conversely the 141 urban-in-rural municipalities operate de facto towns of 20,000-plus buildings on rural-category transfers without urban planning authority. The morphological map is the missing feedback: where governance is premature, and where it is overdue. The dilution judgment is sharpest at fine resolution (at 1 km DEGURBA grain these units register urban) but not scale-arbitrary: the anchor validates the fine lattice ($\kappa = 0.40$ against 0.03 at 1 km). The claim is not that the 112 towns are non-urban in any absolute sense but that statutory status outruns the territorially dominant core.

6.2 Why the unit of analysis changes the law

Nepal's administrative map cuts across ecological regions, so fitting scaling laws on its containers imports the territorial confounding MAUP warns of (Fotheringham & Wong, 1991; Openshaw, 1984): the administrative road exponent spans $\beta = 1$ (0.895/0.963/1.391 proxy; 0.961/0.916/1.540 official), while percolated clusters give a flat profile under *every* delineation tried (1.05–1.12 frozen cores; 0.88–1.11 official belts; 1.03–1.13 trimmed settlement-scale; § 5.7), no pair significant anywhere, heartland equivalence within ±0.15. Systems diverge only at the mountain-bearing pairs, under official belts with clustered inference and Holm correction ($p_{\text{Holm}} = 0.044$)—no other scaling contrast survives. The morphologically delineated aggregate, not the decreed municipality, is the unit at which belt structure vanishes, as scaling theorists assume

[4] Boundary geometry cannot manufacture the gap: the matrix is computed on the as-published geoBoundaries ADM3 polygons (SHA-256 pinned at ingest; the ≈400 m generalisation is display-only), and boundary shape enters only through municipal land area—whose sensitivity the τ-sweep spans ($\tau = 0.05$–0.20: mismatched share 33.5–43.3% of resolved units, containing the headline; `sensitivity_tau_sweep.csv`).

of the functional city (Bettencourt, 2013) and morphologists demonstrated on street networks (Arcaute et al., 2015). We stress what this is *not*: recovery of the canonical $\beta \approx 0.85$ (the morphological Huber exponent is 1.08, s.e. 0.03; mechanisms untested), nor proof that morphology is "truer"—it too is a choice—but the choice under which the asserted theoretical object is actually the unit of regression. A law fitted on administratively blended containers mixes agglomerate physics with territorial geometry under belt-varying weights; the belt-stable morphological exponent is the first honest candidate for comparison against canonical values, in Nepal and in every system sharing this blending.

6.3 Limitations, generality, and next steps

Four boundaries on the claims. *Extraction*: the morphological axis inherits the volunteer geometry—midpoint road attribution (bounded empirically: clipping moves the exponent 0.014), ±50 m link quantisation, latitude-proxy belts whose leakage moves cell *levels* (mountain cell 1.054→0.884 frozen, 1.030 trimmed) without reversing the system *contrast*. *Reconstruction*: the legal axis leaves 24 units unresolved (Appendix A); adjudicating them at either extreme moves misalignment only 33.6–36.8%. *Proxy and provenance*: the POI axis tracks mapping intensity and stays exploratory; and the external benchmarks are not independent of one another or of us—SMOD shares built-form provenance with our footprints, hence $\kappa = 0.40$ is an *upper* bound on external concordance, motivating the form-independent census-density $\rho = 0.86$, and the ethics scanner is a leak heuristic, not a proof. *Corpus*: volunteer footprints are densest where mappers are; the benchmark bounds the direction (under-mapping real, roughly uniform, ratio 0.65), so measured urbanity is a lower bound and 34.7% sits inside the 28.7–36.2% external band.

The protocol and the warning travel: wherever urban categories are decreed rather than measured—statutory-town definitions across South Asia and Africa, reclassification waves in middle-income federations, national aggregates feeding World Urbanisation Prospects-style estimates. Building-morphology audits are becoming comparative practice (Prieto-Curiel, Patino, & Anderson, 2023), but not yet joined to a complete universe of decreed legal units. The procedure ports compactly: percolate at a swept radius; demand an absolute core and relative share; cross-tabulate against legal units; stratify scaling fits; permute under fixed marginals—aggregate-only by construction, wherever volunteer buildings exist. Next: difference successive footprint vintages to date the wedge, and recompute service-access indicators on morphological units.

7 Conclusion

Administrative boundaries are instruments of governance, not containers of urban reality. Scoring all 753 Nepali LLGs against percolated building morphology, decree and reality disagree in both directions and by no marginal amount: 38.5% of legally urban units contain no dense core covering ≥10% of their municipal land, 32.2% of legally rural units contain one—34.7% misalignment on the 50 m lattice, which the GHSL-SMOD anchor places in a 28.7–36.2% band and validates at that lattice ($\kappa = 0.40$, collapsing to 0.03 at 1 km), and the 2021 census join fixes at 31.1% of the population

(band 27–46%). Concordance is far above chance ($p = 1.0 \times 10^{-14}$) but only weak-to-moderate ($V \approx 0.29$). The consequence extends into inference: administrative road exponents diverge at the mountain end (β 0.90/0.96/1.39 proxy; 0.96/0.92/1.54 official) while morphological exponents split nowhere (β 1.03–1.13 under every delineation); the cross-system difference survives only at the mountain-bearing pairs under official belts with clustered inference and Holm correction ($p_{\text{Holm}} = 0.044$—the paper's only surviving scaling family). Every state's urban statistics rest on legally drawn boundaries of uneven morphological fidelity; Nepal's fresh, complete, suddenly reclassified map makes the gap measurable end-to-end, and makes visible what comparative urbanism risks whenever it takes the administrative map as its unit: a distorted census of who lives in cities, and a boundary-dependent physics of how cities work. Urban policy indicators should be computed on morphologically delineated units (or at minimum reported on both), and scaling studies in data-scarce regions should treat boundary choice as a first-order identification decision.

Three numbers carry away the case: about a third of units—34.7% on the validated lattice, 28.7–36.2% against the external grid—disagree with an independent city definition in both directions (112 diluted decreed cities, 141 unrecognised towns); 31.1% of the population, some nine million people (27–46% band), live under a mismatched label; and of 18 confirmatory tests—declared before the final analysis, not externally pre-registered—exactly two survive family-wise correction ($p_{\text{Holm}} = 0.044$), the official-belt mountain-bearing divergences that decide, silently, whether a foundational scaling regularity appears to hold. All recomputable end-to-end from the aggregate tables released with this paper.

Appendix A The 24 unresolved local-level governments

These 24 geoBoundaries ADM3 polygons are real LLG-shaped units whose names did not match the municipality codebook under our conservative, injective, district-constrained rematch. We refuse a fuzzy assignment that could silently mis-join a same-named unit across districts; all 24 are excluded from the mismatch denominators ($N = 729$) and from all χ^2 and scaling fits. Table A1 lists them verbatim for audit; the District column reproduces the source string exactly, including legacy VDC-level labels, which we deliberately do not silently recode.

Note. District codes follow the standard 77-district enumeration. The two Nawalparasi entries carry "—" because their source polygons predate the district's East/West division; the correct code cannot be assigned without a fuzzy cross-district join, which we deliberately avoid.

Appendix B Displaced robustness analyses and validation detail

Selection-matched dense-core screens. Applying the dense-core rule ($N \geq 200$, a percolation core cell, $\geq 10\%$ of unit land in core) plus variants to *administrative* units: the screened Terai–Mountains gap is $+0.322$ under the dense-core rule (Terai $\beta = 0.964$, $n = 163$; Mountains $\beta = 1.286$, $n = 15$—too few for power, $p = 0.12$) and $+0.396$, $p = 0.009$ under the weaker core-cell screen retaining 141 of 153 mountain units ($\beta = 1.291$). The Terai–Hills contrast, non-significant unscreened ($+0.068$, $p = 0.19$), reverses under the dense-core screen ($\Delta\beta = -0.153$, $z = -3.01$, $p = 0.003$). Neither administrative Terai–Hills gap differs from its morphological counterpart (DiD $p = 0.36$, 0.13); the screened Terai–Mountains DiD is suggestive ($p = 0.10$). The stable signature across screens: administrative units retain a mountain-end excess *in magnitude* where

Table A1 The 24 unresolved LLG polygons (excluded from classification).

Name	District	Code	Name	District	Code
Pandav Gupha	Baijura	66	Meghang	Nuwakot	36
Kanda	Bajhang	67	Netrawati	Nuwakot	36
Bhimsen	Gorkha	41	Parbati Kunda	Rasuwa	74
Dhurkot	Gulmi	51	Duikholi	Rolpa	49
Lamidanda	Khotang	19	Sukidaha	Rolpa	49
Narphu	Manang	75	Suwarnabati	Rolpa	49
Neshyang	Manang	75	Tribeni	Rukum_E	77
Dalome	Mustang	76	Dhorchaur	Salyan	64
Binayee Tribeni	Nawalparasi	—	Kumakhmalika	Salyan	64
Bungdikali	Nawalparasi	—	Aurahi	Saptari	27
Belhi Chapena	Saptari	27	Dudhkaushika	Solukhumbu	17
Barah	Sunsari	11	Budhiganga	Sunsari	11

morphological cores show no belt structure; the heartland ordering is screen-dependent and carries no claim, and the excess loses *statistical* resolution once the mountain cell thins to 15—a disclosed power limitation.

Geometric-clipping re-attribution. Re-deriving per-LLG street length by true clipping against municipal polygons and refitting: mountain exponent 1.391 → 1.377, Terai–Mountains contrast 0.496 → 0.492 (per-belt shift ≤ 0.014); clipping lowers absolute length roughly proportionally (median ratio 0.67–0.76), moving intercepts, not slopes.

Building-share criterion. Flagged units are systematically large (mean territory 207 km^2 urban-by-decree; 113 true-urban; 71 urban-in-rural), so we test the building-share alternative: it is degenerate—median share >96% in *every* cell (by status: 0.999 urban, 0.992 rural; by cell: 1.0 true-urban and urban-in-rural, 0.98 urban-by-decree, 0.97 true-rural) and zero of the 112 below 10%, since 500 m percolation absorbs nearly every occupied cell into the 197 clusters. The classifier's content is territorial dominance, and the title's "by decree" claims statutory status outruns the *territorially dominant* core—not that the towns lack urban form.

Table B1 Every belt-stratified road and administrative-POI scaling cell under both belt definitions (script 15; shipped as `official_belt_scaling.csv` and `trim_belt_scaling.csv`). Cells are OLS β with district-clustered standard errors in brackets; n below counts fitting units—cores in the untrimmed rows and *fragments* in the trimmed rows, which are spatially adjacent pieces of delineated cores rather than independent municipal units (§ 5.7); morphological frozen-core rows carry i.i.d. standard errors because their clusters are spatially dispersed across districts (§ 5.5). The trimmed rows re-delineate morphological units at the density-minimum first quartile (script 17; § 5.7), dissolving the 13 mega-corridors into settlement-scale fragments. Administrative mountain-bearing contrasts under the official belts: Hills–Mountains $\Delta\beta = 0.623$, clustered $p = 0.022$; Terai–Mountains $\Delta\beta = 0.579$, $p = 0.028$; no morphological pair is significant under either definition or delineation (clustered $p \geq 0.086$ throughout). POI coefficients are reported for completeness only; the axis is mapping-contaminated (§ 5.4) and must not be quoted as a scaling result.

Definition	Series	Terai	Hills	Mountains
Panel A: Road-length β				
Proxy	Admin	0.895 (0.046)	0.963 (0.068)	1.391 (0.241)
	n	265	324	153
Official	Admin	0.961 (0.039)	0.916 (0.076)	1.540 (0.261)
	n	276	348	118
Proxy	Morph (frozen)	1.117 (0.045)	1.101 (0.062)	1.054 (0.107)
	n	73	46	44
Official	Morph (frozen)	1.086 (0.036)	1.115 (0.059)	0.884 (0.212)
	n	81	54	28
Proxy	Morph (trimmed)	1.091 (0.018)	1.165 (0.025)	1.068 (0.066)
	n	3,725	2,462	298
Official	Morph (trimmed)	1.130 (0.020)	1.097 (0.022)	1.030 (0.055)
	n	4,244	1,907	277
Panel B: Administrative commercial-POI β (mapping-contaminated; descriptive only)				
Proxy	Admin	1.528 (0.168)	2.087 (0.177)	0.581 (0.154)
	n	265	325	163
Official	Admin	1.583 (0.154)	1.997 (0.212)	0.231 (0.106)
	n	276	349	128

Table B2 Appendix companion to Table 4 (§ 5.4): the same belt-stratification conventions applied to commercial-POI counts. Reported here in the appendix rather than in the body because the axis tracks volunteer POI mapping intensity as much as underlying activity—on decreed units the profile is dominated by that contamination (the near-non-physical sub-unity mountain exponent) and none of these coefficients enters the declared confirmatory family. Reported for completeness only; not quotable as a scaling result.

Commercial POI count (β)			
Belt	Terai	Hills	Mountains
Admin β	1.528 [1.352, 1.705]	2.087 [1.909, 2.266]	0.581 [0.387, 0.775]
n	265	325	163
Morph β	1.254 [1.104, 1.404]	1.213 [0.977, 1.449]	1.134 [0.869, 1.398]
n	56	45	40

Table B3 Holm step-down correction within each pre-declared family (script 18; shipped as `holm_multiplicity.csv`). Families `admin_contrast_*` and `did_*` are the confirmatory family of § 4.5g (18 tests); `trim_*` is the declared delineation-sensitivity family (12 tests). Only the two official-belt district-clustered cross-system DiDs reject at 0.05 after correction in the confirmatory family; within the sensitivity family the single rejection (proxy Hills–Terai) is reported in § 5.7 as a bounded exception, not a headline. All other p-values in the paper are descriptive context.

Family	Belt pair	p_{raw}	p_{Holm}	Decision
Confirmatory family				
admin contrast, proxy	Terai–Hills	0.409	0.409	not rejected
admin contrast, proxy	Hills–Mountains	0.087	0.174	not rejected
admin contrast, proxy	Terai–Mountains	0.043	0.129	not rejected
DiD clustered, proxy	Terai–Hills	0.455	0.455	not rejected
DiD clustered, proxy	Hills–Mountains	0.088	0.176	not rejected
DiD clustered, proxy	Terai–Mountains	0.039	0.117	not rejected
admin contrast, official	Terai–Hills	0.602	0.602	not rejected
admin contrast, official	Hills–Mountains	0.022	0.065	not rejected
admin contrast, official	Terai–Mountains	0.028	0.065	not rejected
DiD clustered, official	Terai–Hills	0.502	0.502	not rejected
DiD clustered, official	**Hills–Mountains**	**0.015**	**0.044**	**reject**
DiD clustered, official	**Terai–Mountains**	**0.022**	**0.044**	**reject**
DiD block-bootstrap, proxy	Terai–Hills	0.480	0.480	not rejected
DiD block-bootstrap, proxy	Hills–Mountains	0.110	0.219	not rejected
DiD block-bootstrap, proxy	Terai–Mountains	0.066	0.198	not rejected
DiD block-bootstrap, official	Terai–Hills	0.615	0.615	not rejected
DiD block-bootstrap, official	Hills–Mountains	0.039	0.118	not rejected
DiD block-bootstrap, official	Terai–Mountains	0.046	0.118	not rejected
Delineation-sensitivity family				
post-trim contrasts, official	Hills–Terai	0.267	0.515	not rejected
post-trim contrasts, official	Mountains–Hills	0.257	0.515	not rejected
post-trim contrasts, official	Mountains–Terai	0.086	0.258	not rejected
post-trim contrasts, proxy	Hills–Terai	0.015	0.045	reject
post-trim contrasts, proxy	Mountains–Hills	0.166	0.332	not rejected
post-trim contrasts, proxy	Mountains–Terai	0.733	0.733	not rejected
pre-trim cores, official	Hills–Terai	0.540	0.711	not rejected
pre-trim cores, official	Mountains–Hills	0.237	0.711	not rejected
pre-trim cores, official	Mountains–Terai	0.299	0.711	not rejected
pre-trim cores, proxy	Hills–Terai	0.768	1.000	not rejected
pre-trim cores, proxy	Mountains–Hills	0.579	1.000	not rejected
pre-trim cores, proxy	Mountains–Terai	0.462	1.000	not rejected

References

Agergaard, J., Subedi, B.P., Brøgger, D. (2022). Political geographies of urban demarcation: Learning from Nepal's state-restructuring process. *Political Geography*, *98*, 102704, https://doi.org/10.1016/j.polgeo.2022.102704

Arcaute, E., Hatna, E., Ferguson, P., Youn, H., Johansson, A., Batty, M. (2015). Constructing cities, deconstructing scaling laws. *Journal of the Royal Society Interface*, *12*(102), 20141355, https://doi.org/10.1098/rsif.2014.1355

Batty, M. (2008). The size, scale, and shape of cities. *Science*, *319*(5864), 769–771, https://doi.org/10.1126/science.1151419

Batty, M., & Ferguson, P. (2011). Defining city size. *Environment and Planning B: Planning and Design*, *38*(5), 753–756, https://doi.org/10.1068/b3805ed

Batty, M., & Xie, Y. (1994). From cells to cities. *Environment and Planning B: Planning and Design*, *21*(7), s31–s48, https://doi.org/10.1068/b210s31

Bettencourt, L.M.A. (2013). The origins of scaling in cities. *Science*, *340*(6139), 1438–1441, https://doi.org/10.1126/science.1233424

Bettencourt, L.M.A., Lobo, J., Strumsky, D., West, G.B. (2007). Growth, innovation, scaling, and the pace of life in cities. *Proceedings of the National Academy of Sciences*, *104*(17), 7301–7306, https://doi.org/10.1073/pnas.0610147104

Brenner, N. (Ed.). (2013). *Implosions/explosions: Towards a study of planetary urbanization.* Berlin: Jovis.

Brenner, N., & Schmid, C. (2015). Towards a new epistemology of the urban? *City*, *19*(2–3), 151–182, https://doi.org/10.1080/13604813.2015.1014712

Cottineau, C., Hatna, E., Arcaute, E., Batty, M. (2017). Diverse cities or the systematic paradox of Urban Scaling Laws. *Computers, Environment and Urban Systems*, *63*, 80–94, https://doi.org/10.1016/j.compenvurbsys.2016.04.006

Denis, É., & Zérah, M.-H. (Eds.). (2017). *Subaltern urbanisation in India: An introduction to the dynamics of ordinary towns.* New Delhi: Springer India.

Dijkstra, L., Florczyk, A.J., Freire, S., Kemper, T., Melchiorri, M., Pesaresi, M., Schiavina, M. (2021). Applying the Degree of Urbanisation to the globe: A new harmonised definition reveals a different picture of global urbanisation. *Journal of Urban Economics*, *125*, 103312, https://doi.org/10.1016/j.jue.2020.103312

Fleischmann, M., Feliciotti, A., Romice, O., Porta, S. (2022). Methodological foundation of a numerical taxonomy of urban form. *Environment and Planning B: Urban Analytics and City Science*, *49*(4), 1283–1299, https://doi.org/10.1177/23998083211059835

Fotheringham, A.S., & Wong, D.W.S. (1991). The modifiable areal unit problem in multivariate statistical analysis. *Environment and Planning A*, *23*(7), 1025–1044, https://doi.org/10.1068/a231025

Goodchild, M.F. (2007). Citizens as sensors: the world of volunteered geography. *GeoJournal*, *69*(4), 211–221, https://doi.org/10.1007/s10708-007-9111-y

Government of Nepal (2017). *Inter-Governmental Fiscal Arrangement Act, 2074 (2017).* Ministry of Federal Affairs and General Administration, Government of Nepal, Kathmandu. (English translation as published in the Nepal Gazette)

Haklay, M. (2010). How good is volunteered geographical information? a comparative study of OpenStreetMap and Ordnance Survey datasets. *Environment and Planning B: Planning and Design*, *37*(4), 682–703, https://doi.org/10.1068/b35097

Hecht, R., Kunze, C., Hahmann, S. (2013). Measuring completeness of building footprints in OpenStreetMap over space and time. *ISPRS International Journal of Geo-Information*, *2*(4), 1066–1091, https://doi.org/10.3390/ijgi2041066

Herfort, B., Lautenbach, S., Porto de Albuquerque, J., Anderson, J., Zipf, A. (2021). The evolution of humanitarian mapping within the OpenStreetMap community. *Scientific Reports*, *11*(1), 3037, https://doi.org/10.1038/s41598-021-82404-z

Herfort, B., Lautenbach, S., Porto de Albuquerque, J., Anderson, J., Zipf, A. (2023). A spatio-temporal analysis investigating completeness and inequalities of global urban building data in OpenStreetMap. *Nature Communications*, *14*(1), 3985, https://doi.org/10.1038/s41467-023-39698-6

Huber, P.J. (1981). *Robust statistics*. New York: Wiley.

Jochem, W.C., Leasure, D.R., Pannell, O., Chamberlain, H.R., Jones, P., Tatem, A.J. (2021). Classifying settlement types from multi-scale spatial patterns of building footprints. *Environment and Planning B: Urban Analytics and City Science*, *48*(5), 1161–1179, https://doi.org/10.1177/2399808320921208

Jochem, W.C., & Tatem, A.J. (2021). Tools for mapping multi-scale settlement patterns of building footprints: An introduction to the R package foot. *PLOS ONE*, *16*(2), e0247535, https://doi.org/10.1371/journal.pone.0247535

Louf, R., & Barthelemy, M. (2014). Scaling: lost in the smog. *Environment and Planning B: Planning and Design*, *41*(5), 767–769, https://doi.org/10.1068/b4105c

National Statistics Office, Government of Nepal (2023). *National population and housing census 2021: National report* (Tech. Rep.). Kathmandu: National Statistics Office.

Openshaw, S. (1984). *The modifiable areal unit problem*. Norwich: Geo Books.

Pesaresi, M., & Politis, P. (2023). *GHS-BUILT-S R2023A — GHS built-up surface grid, derived from Sentinel-2 composite and Landsat, multitemporal (1975–2030)*. Dataset, European Commission, Joint Research Centre. https://doi.org/10.2905/9F06F36F-4B11-47EC-ABB0-4F8B7B1D72EA.

Pesaresi, M., Schiavina, M., Politis, P., Freire, S., Krasnodębska, K., Uhl, J.H., ... Kemper, T. (2024). Advances on the Global Human Settlement Layer by joint assessment of Earth Observation and population survey data. *International Journal of Digital Earth*, *17*(1), 2390454, https://doi.org/10.1080/17538947.2024.2390454

Prieto-Curiel, R., Patino, J.E., Anderson, B. (2023). Scaling of the morphology of African cities. *Proceedings of the National Academy of Sciences*, *120*(9), e2214254120, https://doi.org/10.1073/pnas.2214254120

Roy, A. (2011). Slumdog cities: Rethinking subaltern urbanism. *International Journal of Urban and Regional Research*, *35*(1), 223–238, https://doi.org/10.1111/j.1468-2427.2010.00981.x

Rozenfeld, N.D., Rybski, D., de Andrade, J.S., Batty, M., Stanley, H.E., Makse, H.A. (2008). Laws of population growth. *Proceedings of the National Academy of*

Sciences, *105*(48), 18702–18706, https://doi.org/10.1073/pnas.0808151105

Runfola, D., Anderson, A., Baier, H., Crittenden, M., Dowker, E., Fuhrig, S., ... Hobbs, L. (2020). geoBoundaries: A global database of political administrative boundaries. *PLOS ONE*, *15*(4), e0231866, https://doi.org/10.1371/journal.pone.0231866

Schiavina, M., Melchiorri, M., Pesaresi, M. (2023). *GHS-SMOD R2023A — GHS settlement layers, application of the Degree of Urbanisation methodology (stage I) to GHS-POP R2023A and GHS-BUILT-S R2023A, multitemporal (1975–2030).* Dataset, European Commission, Joint Research Centre. https://doi.org/10.2905/A0DF7A6F-49DE-46EA-9BDE-563437A6E2BA.

Seto, K.C., Güneralp, B., Hutyra, L.R. (2012). Global forecasts of urban expansion to 2030 and direct impacts on biodiversity and carbon pools. *Proceedings of the National Academy of Sciences*, *109*(38), 15231–15236, https://doi.org/10.1073/pnas.1117144109

United Nations Statistical Commission (2020). *A recommendation on the method to delineate cities, urban and rural areas for statistical and spatial purposes: Applying the Degree of Urbanisation* (Tech. Rep.). United Nations, UN Statistical Commission (51st session).