

Building-scale functional mixing reveals a decarbonization window in cities

Xin Huang

xhuang@whu.edu.cn

School of Remote Sensing and Information Engineering, Wuhan University <https://orcid.org/0000-0002-5625-0338>

Jiayi Li

School of Remote Sensing and Information Engineering, Wuhan University

Zhen Zhang

School of Remote Sensing and Information Engineering, Wuhan University

Liangpei Zhang

Wuhan University <https://orcid.org/0009-0005-4894-2552>

Xiaofeng Pan

Shenzhen Ecological and Environmental Monitoring Center of Guangdong Province

Song Ma

Shenzhen Ecological and Environmental Monitoring Center of Guangdong Province

Xiangyun Xiong

Shenzhen Ecological and Environmental Monitoring Center of Guangdong Province

Article

Keywords:

Posted Date: July 21st, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-9910345/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Additional Declarations: There is **NO** Competing Interest.

Building-scale functional mixing reveals a decarbonization window in cities

Jiayi Li^{1,†}, Zhen Zhang^{1,†}, Xiaofeng Pan², Song Ma², Xiangyun Xiong², Liangpei Zhang^{3,*} and Xin Huang^{1,*}

¹School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430079, China.

²Shenzhen Ecological and Environmental Monitoring Center of Guangdong Province, Shenzhen 518049, China.

³State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing (LIESMARS), Wuhan University, Wuhan 430079, China.

[†]These authors contributed equally: Jiayi Li, Zhen Zhang.

*e-mail: xhuang@whu.edu.cn (Xin Huang); zlp62@whu.edu.cn (Liangpei Zhang).

Abstract

Mixed-use development is promoted for urban decarbonization, but conventional metrics cannot resolve the building-level functions shaping energy use and travel. We develop a nationwide functional atlas of approximately 200 million buildings in China by integrating sub-meter imagery, semantic information and vector data. ODIAC-derived emissions reveal that neighborhood-scale mixing is associated with steadily lower emissions, whereas building-scale mixing shows a three-stage relationship. At low levels, scattered functions may

raise emissions; at moderate levels (0.14–0.72), increasing mixing reduces emissions; beyond this range, benefits level off or reverse. Shifting under-mixed areas into the moderate range could conservatively mitigate about 100 Mt C yr⁻¹. Analysis of 31 international cities reveals a similar pattern, although transition points vary. Thus, more mixing is not always better: planners should prioritize low-mixing areas, avoid excessive mixing where benefits have saturated, and adopt locally tailored building-scale targets.

Introduction

Urban operational emissions, driven by building operations, mobility and other land-based urban activities, sit at the center of the global mitigation challenge^{1,2}. Cities account for ~80% of final energy use and a comparable share of energy-related greenhouse-gas emissions^{3,4}. Here “operational carbon” denotes territorial, land-based carbon emissions from the Open-source Data Inventory for Anthropogenic CO₂ (ODIAC) within urban areas, reported in carbon-mass units, rather than embodied emissions or consumption-based carbon footprints. China provides an unusually stringent stress test for identifying actionable pathways, because it has been the world’s largest annual greenhouse-gas emitter since the mid-2000s while simultaneously undergoing the most extensive urbanization project in history, adding more floor area over the past two decades than any other country^{5,6}. Under its “dual-carbon” strategy,

China is pivoting from rapid greenfield expansion to renovation and optimization of the existing urban stock, compressing the timetable for near-term, stock-focused abatement⁷. As expansion slows, outcomes hinge increasingly on how buildings are used and co-located, what functions are mixed, at what intensities, and where within the urban fabric—yet the quantitative rules that link mixed use to urban operational carbon remain unsettled.

Accumulating evidence establishes a link between urban form and carbon outcomes^{8,9}. Compact and mixed-use patterns generally reduce emissions by reshaping daily mobility demand, shortening travel distances, enabling shared infrastructure, and balancing diurnal loads, whereas fragmented or single-function layouts increase carbon intensity^{10,11}. Emerging evidence further suggests that these benefits are bounded and context-dependent, with non-linear responses and saturation thresholds^{12,13}, underscoring the need for quantitative, rule-based guidance rather than heuristic prescriptions. Taken together, these mechanisms place mixed use at the land-use–mobility–energy nexus, yet they are seldom quantified at building resolution, because most analyses rely on coarse-scale city-level studies or proxies such as land-use categories and night-time lights^{14,16,17}. Such approaches often lack the resolution to isolate individual buildings as the fundamental units of energy consumption and activity^{15,18}. Moreover, "mixed use" is frequently measured at a single spatial resolution, despite scale-sensitive mechanisms that can shift from substitution effects to frictions and capacity penalties as functions interweave more finely. Systematic tests of this critical scale dependency remain rare^{19,20}.

This gap calls for building-resolved evidence on how functional mixing translates into operational carbon, at the spatial scale where energy use and mobility demand are organized. Mixed use can generate proximity-driven substitution benefits, yet it can also impose coordination frictions²¹ and capacity constraints^{22,23}, implying that carbon benefits may operate only within a bounded intensity range rather than increasing monotonically. Critically, these mechanisms are scale-sensitive: neighborhood-scale mixing reflects coarse functional balance, whereas building-scale mixing captures fine-grained co-location and vertical interweaving of uses. We therefore test whether functional mixing exhibits a quantifiable “decarbonization window” with identifiable thresholds, and whether its location and strength shift systematically across scales and urban contexts.



Fig. 1. Multiscale functional labeling and scale-dependent mixing-carbon relationships. Schematic 3D visualizations created in *Cities: Skylines* illustrate the multiscale labeling framework and key concepts used in this study. **a.** A stylized urban area with four nested spatial scales: **Scale 1** (building), **Scale 2** (block/functional zone), **Scale 3** (2 km × 2 km grid cell), and **Scale 4** (10 km × 10 km cross-grid context). Example blocks/functional zones (B1–B8) are shaded by relative operational-carbon intensity

(darker tones indicate higher intensity), and route markers illustrate accessibility and cross-area interactions. Blocks/functional zones are labeled by dominant functions: Residential (B1), Industrial (B2), Transport (B3), Education (B4), Healthcare (B5), Administrative (B6), Public open space (B7), and Commercial (B8). **b.** Example grid cells (G1–G6) with increasing degrees of functional mixing, quantified by BFMI. **c–f.** Illustrative building-level mixed-use configurations highlighted in (a): (c) dormitories within a school (education–residential), (d) ground-floor commerce with upper-floor residences (residential–commercial), (e) housing embedded in an industrial park (residential–industrial), and (f) shops in an airport (transport–commercial). The schematic emphasizes that the apparent carbon benefits of mixed use can be scale-dependent and that, at broader spatial extents, functional organization can also shape mobility-related carbon through cross-grid commuting (e.g., routes from residential units R1–R2 to employment centers).

To answer these questions, we construct the first nationwide, building-resolved functional atlas for China and quantify functional mixing consistently at both neighborhood and building scales. **Fig. 1** summarizes the multiscale study design and the two-tier functional labels. We introduce a multi-scalar Building Function Mixed Index (BFMI) to capture center–periphery structure, and use a spatial panel framework that accounts for spillovers and other confounders. To identify policy-relevant patterns beyond coarse correlations, we estimate the BFMI–carbon relationship in a Spatial Durbin panel with city and ring fixed effects, recovering non-linear marginal responses while accounting for spatial spillovers, and contrasting effects across neighborhood versus building resolution; we further separate movement versus morphology channels using two jobs–housing indicators (expected average commuting distance and Jensen–Shannon divergence). This design links land-use organization to territorial urban operational carbon through mobility constraints and non-mobility channels within a unified spatial framework. Empirically, we reveal a tri-phasic mixing–carbon response with a policy-relevant decarbonization window and a practical upper bound, translating mixed use from a qualitative planning ideal into an actionable target range for stock-

focused decarbonization. Finally, a comparative analysis of 31 global cities reproduces the characteristic tri-phasic structure, clarifying portable, capacity-conditioned boundary conditions for transferring building-scale mixed-use targets across contexts.

Results

A nationwide building-resolved functional atlas reveals China's mixed-use structure

We first establish a nationwide, building-resolved functional atlas that enables consistent measurement of mixed use across China's urban system. We harmonize a two-tier ontology aligned with China's planning land-use taxonomy according to *Urban Land Use Classification and Planning Construction Land Standards*²⁴, defining eight macro functional classes at both neighborhood (contextual) and building (intrinsic) scales, and further organizing them into 29 contextual–intrinsic combinations after screening for separability (**Supplementary Note 1, Supplementary Note 3 and Extended Data Fig.1**). To enable nationally consistent mapping, we then develop a multimodal POI–aerial imagery–vector fusion pipeline that integrates language-based POI semantics with sub-meter visual cues and vector context to assign two-tier functional labels at national scale (Methods). Applying this pipeline, we annotate ~200 million urban buildings across China (**Fig. 2a**). Independent manual validation of the final map using 3,024 interpreted buildings yields an overall accuracy of 90.77% (Cohen's $\kappa = 0.885$) at the neighborhood scale and 86.97% ($\kappa = 0.862$; macro-F1 = 0.861) at the building scale (**Extended Data Fig. 2**), providing substantially improved

national coverage and resolution relative to prior city-scale efforts.

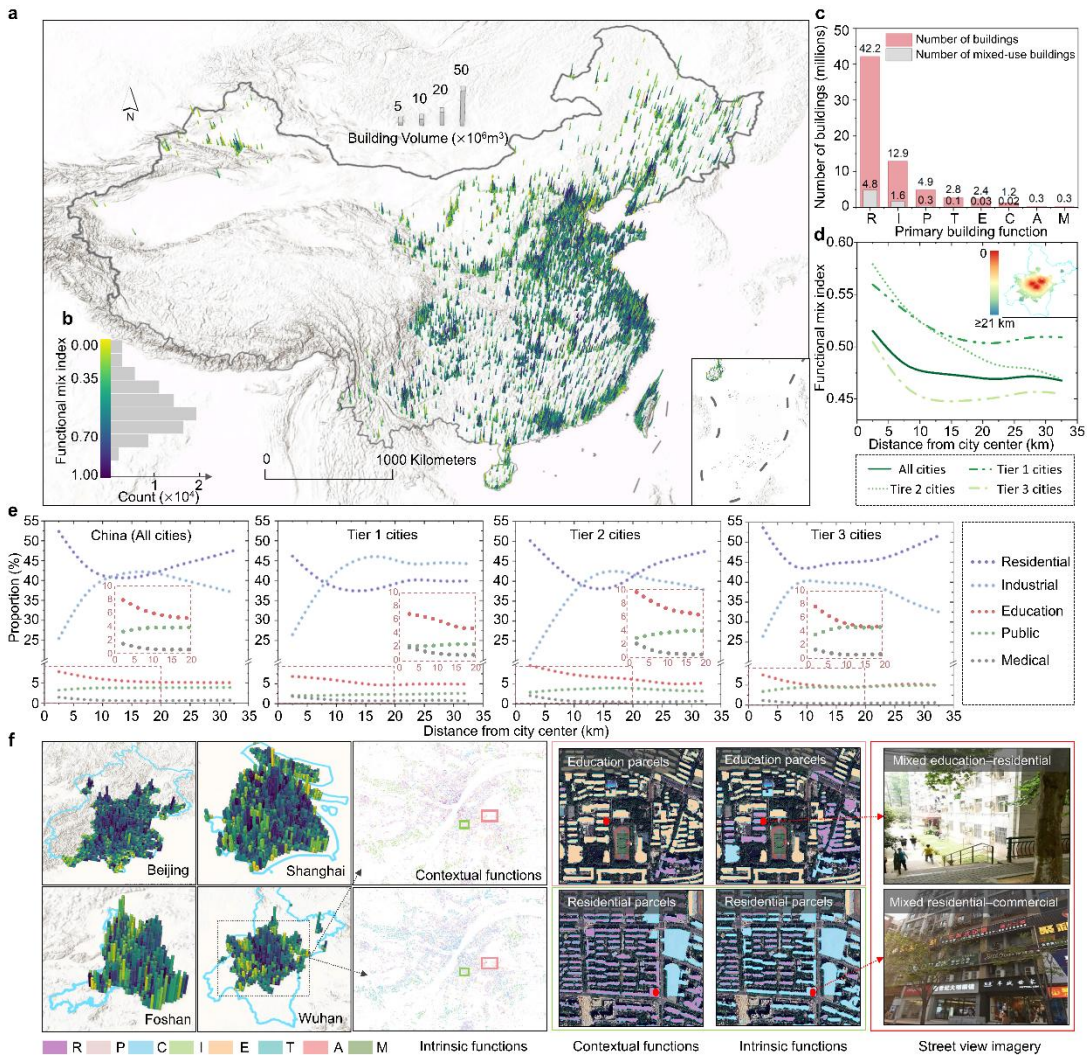


Fig. 2. Spatial patterns of building functions across Chinese cities. a. grid-scale spatial distribution of functional mix. Grid height represents total building volume, and color indicates the degree of functional mix. **b.** Frequency distribution of grids by functional mix level. **c.** statistical breakdown of building function categories across Chinese cities. **d.** Variations in functional mix relative to distance from the city center; the top-right inset illustrates the distance calculation for Wuhan, highlighting its polycentric structure. **e.** Distance-based gradients of proportions for key building functions. **f.** Representative patterns of functional distributions. **Left columns:** General patterns of functional mix in

four representative cities, characterizing high central diversity and low peripheral diversity. **Middle column:** Spatial configuration of functions in Wuhan, dominated by residential and commercial use in the center and industrial use in the periphery. **Right columns:** Zoomed-in views of the pink (education parcels) and green (residential parcels) boxes outlined in the middle column, demonstrating the accurate identification of mixed-use buildings, such as mixed education–residential and mixed residential–commercial structures.

The resulting atlas reveals a highly uneven functional composition of the building stock (**Fig. 2c**). Across built-up areas, Residential buildings dominate (62.98%), followed by Industrial (19.19%), while Commercial, Education, Public services, Transport, Administrative, and Medical constitute smaller but non-negligible shares. Mixed use is common at the neighborhood scale: in Residential-labeled neighborhoods, ~11.4% of buildings exhibit intrinsic mixed use (predominantly Residential–Commercial); in Industrial-labeled neighborhoods, ~12.4% are mixed (predominantly Industrial–Commercial), consistent with ground-floor retail embedded within housing blocks²⁵ and producer services co-located with manufacturing spaces²⁶.

We quantify functional mixing using BFMI (normalized Shannon entropy of building-volume shares) on 2-km grids (see multi-scale robust analysis in **Supplementary Methods 4** and **Supplementary Fig. S9**) and summarize it along center–periphery distance bands. Nationally, BFMI is unimodal (see **Fig. 2b**, modal bin 0.53–0.54) with an interquartile range of 0.38–0.58 (IQR = 0.202), indicating that moderate mixing is the prevailing urban condition, while very low- (BFMI $\lesssim$ 0.3) and very high-mixing pockets (BFMI $\gtrsim$ 0.65) form comparatively sparse tails (24.48%)

(**Fig. 2c**). Center–periphery profiles further reveal systematic gradients in both functional composition (**Fig. 2e**) and mixing intensity (**Fig. 2d**): residential shares decline from inner cores toward mid-rings and partially recover outward, while industrial shares show the opposite. Importantly, BFMI distance profiles differ by city tier: Tier-I (*i.e.*, urban built-up area > 3,000 km²) and Tier-III cities (*i.e.* urban built-up area < 800 km²) show high core diversity, mid-ring trough, modest outward rebound (minima at ~18 km and ~12 km, respectively), whereas Tier-II cities (*i.e.* urban built-up area 800-3,000 km²) display a more monotonic decline (**Fig. 2e**). The tier-wise heterogeneity highlights that “where mixing is scarce” is itself systemically structured rather than idiosyncratic.

Scale divergence: neighborhood mixing understates the non-linear costs of building-level mixing

We next quantify functional mixing-carbon relationship in a spatial panel framework with city and ring fixed effects (Methods). A central finding is that the mixing-carbon relationship is scale-divergent when mixing is computed from coarse (contextual) versus fine (contextual–intrinsic) functional labels. Using China-wide analyses, we find that neighborhood-scale mixing exhibits a largely monotonic reduction in urban operational carbon (**Fig. 3a**), which is consistent with a broad “substitution effect”, whereby balancing major urban functions (housing-jobs-services) supports localized activity and reduces long-distance commuting demands²⁷. In contrast, building-scale mixing reveals a distinctly non-linear pattern with two cost-sensitive tails (**Fig. 3b**).

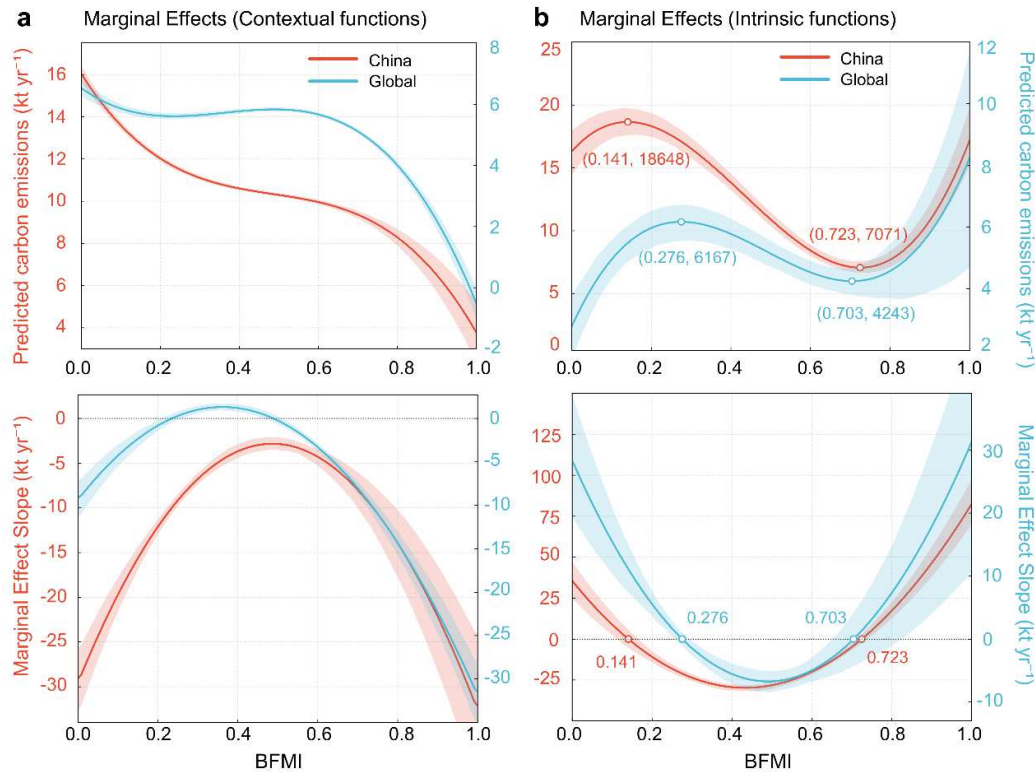


Fig. 3. Scale- and context-dependent marginal effects of functional mixing reveal a decarbonization window and an upper bound. The curves represent the average marginal effects (AMEs) evaluated at sample averages and their corresponding derivatives (slopes). Shaded areas indicate 95% confidence intervals (CIs) estimated via the delta method with standard errors clustered at the city level. **a.** The marginal effect curve (top panel) and its slope (bottom panel) for the building contextual function. **b.** The marginal effect curve (top panel) and its slope (bottom panel) for the building contextual-intrinsic function.

At very low building-scale mixing (0–0.141), introducing scattered fine-grained functions can impose “friction costs” (e.g., circulation complexity and operational inefficiencies²⁸) before a coherent internal service network emerges. More critically, we

identify a complexity threshold at $\text{BFMI} \approx 0.723$ (**Fig. 3b**): beyond this point, additional fine-grained (often vertical) interweaving can trigger diminishing returns and even rebounds, as operational inefficiencies and local congestion penalties begin to offset proximity benefits. This scale divergence implies that planning guidance based solely on neighborhood-scale land-use balance risks overstating the universal benefits of mixing and missing the upper-bound constraint that emerges only when functions are resolved at the building level. These turning points are robust to alternative functional-taxonomy granularities (**Supplementary Note 3; Supplementary Fig. S9**).

A tri-phasic mixing–carbon response delineates an effective “synergy window” and a practical upper bound

We next focus on the building-scale mixing–carbon relationship, allowing for non-linearity and spatial spillovers. The estimated marginal response of operational carbon to BFMI exhibits a clear tri-phasic structure (**Fig. 3b**): marginal effects are positive at low mixing ($\text{BFMI} < 0.141 \pm 0.027$; 95% CI 0.112–0.166), become significantly negative at moderate mixing, with the derivative of the marginal-effect curve reaching its most negative value around $\text{BFMI} \approx 0.437 \pm 0.013$ (95% CI 0.420–0.445), and turn positive again at very high mixing. The abatement window is bounded by derivative-based zero-crossings, while the steepest strengthening of the marginal-abatement response occurs around $\text{BFMI} \approx 0.437$. Derivative-based zero-crossings define an effective abatement range: $\text{BFMI} \in [0.141 \pm 0.027, 0.723 \pm 0.013]$ (95% CI bounds 0.112–0.166 and 0.714–0.739), within which increasing BFMI yields significantly negative marginal effects; beyond the upper threshold, benefits saturate or reverse.

Given the national BFMI distribution (median 0.491, IQR 0.202), these thresholds imply a practical rule: the largest gains come from lifting low-mix tails into the abatement window and stabilizing high-performing areas before the upper-bound regime, while avoiding unchecked escalation in already high-mix cores. To translate the response curve into policy-relevant magnitudes, we evaluate static counterfactual alignments (see **Supplementary Note 4** for details). An “ideal alignment” setting each grid to the upper turning point of the fitted response curve before saturation (BFMI $\approx$ 0.723) yields 310.4 Mt C yr⁻¹ of national abatement (95% CI, 267.9–355.1 Mt C yr⁻¹; 1,138.1 Mt CO₂ yr⁻¹ equivalent, 95% CI, 982.3–1,302.0 Mt CO₂ yr⁻¹), corresponding to 20.9% of urban operational carbon. A “cost-effective alignment” with minimal shifts (+0.30 for BFMI < 0.30, -0.25 for BFMI > 0.75) still yields 97.37 Mt C yr⁻¹ of abatement (95% CI, 90.74–104.17 Mt C yr⁻¹; 357.0 Mt CO₂ yr⁻¹ equivalent, 95% CI, 332.7–382.0 Mt CO₂ yr⁻¹), corresponding to 6.6% of urban operational carbon, with disproportionate contributions from Tier-III cities (**Supplementary Note 4**).

Channels, spillovers, and boundary conditions from global city comparisons

Composition matters beyond BFMI, and its carbon implications are strongly mediated by inter-area interactions. Conditional on BFMI and controls, total effects differ sharply across functions: per additional 1×10^5 m³ of volume, emissions increase by $+766 \pm 12.5$ t C yr⁻¹ (commercial), $+2,144 \pm 65.8$ t C yr⁻¹ (medical), $+278 \pm 3.3$ t C yr⁻¹ (industrial), and $+297 \pm 20.1$ t C yr⁻¹ (transport), while residential and education volumes are associated with reductions. Notably, positive-sector totals are dominated by indirect (spillover) effects (about 74% for industrial and 63% for

commercial), indicating that the carbon consequences of high-flux functions shape operational carbon through cross-grid propagation rather than purely local impacts.

To isolate the mobility-related constraints linking functional organization to system-wide circulation, we decompose jobs-housing separation into two complementary indicators: a distance channel (expected average commuting distance, EAD) and a morphology channel (Jensen-Shannon divergence, JSD, between residential and employment volume distributions). Entered jointly with BFMI, both channels are positively associated with operational carbon and are largely spillover-driven: +609.5 t C yr⁻¹ per +1 km in EAD and +248.4 t C yr⁻¹ per +0.1 in JSD (**Extended Data Table 2**). Heterogeneity across China's urban hierarchy is pronounced: EAD effects are largest in Tier-I cities, while JSD remains positive across tiers, suggesting that movement impedance and morphological polarization contribute differently across city types (**Extended Data Table 3**).

We next probe external validity by extending the analysis to 31 global cities spanning multiple continents and city-size classes (**Supplementary Methods 8** and **Supplementary Table S7**). At the contextual-intrinsic, building-resolved scale, global regressions recover a tri-phasic mixing-carbon response that is structurally consistent with China (**Fig. 3b**), including a mid-range synergy window and a high-mixing upturn with closely aligned upper-bound turning points (global BFMI ≈ 0.703 ; China ≈ 0.723). The main divergence emerges in the low-mixing regime: the first turning point occurs later in the global sample (≈ 0.276) than in China (≈ 0.141), indicating that early-stage frictions and the onset of the synergy window are more context-sensitive than the high-

mixing tail.

Discussion

Mixed use does not monotonically decarbonize cities, and its carbon implications are bounded, scale-dependent, and visible only at building resolution. This finding reconciles a central inconsistency in prior works^{9,16,17}, where coarse proxies and city-level summaries often imply uniformly beneficial effects. At neighborhood scale, mixing primarily captures aggregate functional balance and substitution, that is, how housing, jobs and services are distributed across districts, so its estimated benefits tend to look smooth and monotonic^{29,30}. At building scale, mixing reflects fine-grained co-location and vertical interweaving of uses, where operational frictions and local capacity limits can dominate³¹. The pronounced divergence across resolutions therefore indicates that “mixed use” is not a single mechanism but a family of mechanisms whose sign and magnitude depend on the scale at which functions are measured and organized.

At the contextual-intrinsic, building-resolved scale, mixing follows a tri-phasic (friction–synergy–saturation) response that delineates a policy-relevant decarbonization window and a practical upper bound. The low-mixing regime is the most context-sensitive: incremental diversity is not yet reliably decarbonizing and may even coincide with higher emissions, consistent with early frictions before local complementarities are established. At moderate mixing, proximity benefits dominate, yielding robust emission reductions, a ‘synergy window’, suggesting that coherent co-location can reduce network-wide circulation demands and improve shared-system

utilization²¹. At extreme mixing, benefits saturate or can reverse, indicating an upper-bound regime in which capacity saturation and congestion externalities offset proximity gains in high-throughput cores¹³. This bounded response converts mixed use from a directional heuristic (“more is better”) into an operational design problem: raise under-mixed areas into the synergy window, but treat already hyper-mixed centers as capacity-managed systems rather than candidates for further intensification.

The effectiveness of mixing is system-level. Its impacts can propagate through spatial spillovers, are conditioned by jobs–housing separation, and vary with urban context^{11,32,33}. Operational carbon is therefore shaped not only by the local functional composition of a place but also by inter-area interactions, implying that parcel-level prescriptions may not scale without accounting for network effects³⁴. Consistent with this, separating jobs–housing imbalance into a distance component and a morphological-polarization component suggests that mixing is most likely to decarbonize when it reduces system-wide circulation demand, rather than merely increasing local diversity in situ³⁵. Our cross-city comparison supports both generality and limits: China and the global sample share a mid-range synergy window and a broadly similar high-mixing upturn, while the low-mixing regime differs in the timing and steepness of the first turning point. China therefore does not overturn a broader urban regularity; instead, it makes the upper-bound regime most visible under stringent, high-throughput conditions. The portable implication is to treat mixed use as a bounded, capacity-aware target range, rather than maximizing mixing indiscriminately³⁶.

Our interpretation is subject to several limitations that motivate further work. First,

although independent manual validation indicates high map accuracy, residual category ambiguity can propagate into BFMI and the exact locations of inferred turning points. Uncertainty-aware labeling would further strengthen threshold robustness. Second, BFMI is computed on 2-km grids and cannot represent within-grid connectivity or temporal rhythms. Grid-size sensitivity analyses indicate that 2 km provides a stable compromise between diversity capture and micro-fragmentation noise (Supplementary Fig. S9; Supplementary Table S5), suggesting the upper-bound regime is unlikely to be a discretization artifact, while mobility and infrastructure data are needed to test the underlying mechanisms. Third, EAD and JSD capture complementary dimensions of jobs–housing separation but do not explicitly model mode choice or diurnal scheduling, which could refine channel attribution and heterogeneity. Finally, counterfactual abatement estimates are static and should be interpreted as indicative potentials; embedding implementation costs and behavioral adjustments in future scenario analyses would translate the identified target range into more actionable pathways. Together, these results establish a scale-aware, bounded mixing–carbon relationship that reveals a transferable decarbonization window and a density-conditioned upper bound, converting mixed-use planning from a qualitative ideal into a capacity-aware operational rule for urban decarbonization.

Methods

Study design and spatial units.

We quantify how functional organization shapes ODIAC-derived urban

operational carbon across a nested hierarchy: from individual buildings to neighborhood contexts, $2\text{ km} \times 2\text{ km}$ grids (see Robustness of the grid size in **Supplementary Methods 4**), and the city-wide urban system. The nested spatial units and the two-tier labeling logic are schematized in **Fig. 1a**. Building-scale characterization is enabled by sub-meter (0.5 m) remote-sensing imagery, allowing functions to be resolved at the level of individual structures and then consistently aggregated to neighborhood contexts and grids for computing BFMI, covariates, and grid-matched emissions. City-wide patterns and cross-grid interactions are evaluated over the entire urban grid system within each city's urban boundary, combining center-periphery profiling (distance-to-center bands) with spillover estimation through a spatial-weights structure defined across all grids; the same measurement framework is applied to 31 global cities to test boundary conditions. The baseline China analysis started from all $2\text{ km} \times 2\text{ km}$ built-up grid cells with valid building-function measurements and retained 91,111 regression-valid grid cells after matching BFMI, ODIAC-derived urban operational carbon, covariates, city identifiers and center-distance rings.

Data harmonization and two-tier functional labels

Crowd-sourced AOIs and POIs provide semantic cues but cannot be directly overlaid to label every building nationwide due to strong cross-city and within-city coverage bias³⁷, pervasive sparsity³⁸, and inconsistent quality/taxonomies across platforms³⁹. We therefore harmonize multi-source geodata to a unified schema (**Supplementary Methods 1**) and propose a two-tier functional representation that

separates the neighborhood-context setting from the intrinsic (building) use, explicitly accommodating mixed and fuzzy functional boundaries. Using rule-based decision logic that integrates AOI geometry, POI proximity, and priority-based conflict resolution (see **Supplementary Methods 2** for details), we extract a curated set of ~30 million high-confidence labeled buildings as supervision anchors, which enables scalable inference for the full national building stock (~200 million) in the subsequent model.

City centers are treated as stable anchors for ring construction. Centers are identified from statutory planning designations and corroborated by multi-source spatial evidence (e.g., lights and road topology), with detailed examples reported in **Supplementary Note 2**.

BHENet for nationwide functional inference

To obtain consistent contextual and intrinsic labels at national scale, we develop BHENet (**Fig. 4**), a multimodal fusion model that combines (i) POI text semantics, (ii) sub-meter (0.5 m) remote-sensing imagery, and (iii) vector context (roads and building footprints). For each target building, POIs within a 500 m radius of the centroid are encoded into text embeddings; an image patch centered on the building provides visual features; and vector layers capture accessibility and morphology. Features are fused using attention-based modules, followed by a two-head classifier that predicts contextual and intrinsic functions. Training, optimization, and hyperparameters are documented in the code repository and **Supplementary Methods 3**. To independently assess the final mapping accuracy, we used a manually interpreted validation set of

3,024 buildings based on high-resolution imagery, street-view or map evidence, and POI context. We computed overall accuracy, Cohen’s κ , precision, recall and F1 for both the neighborhood scale and the building scale.

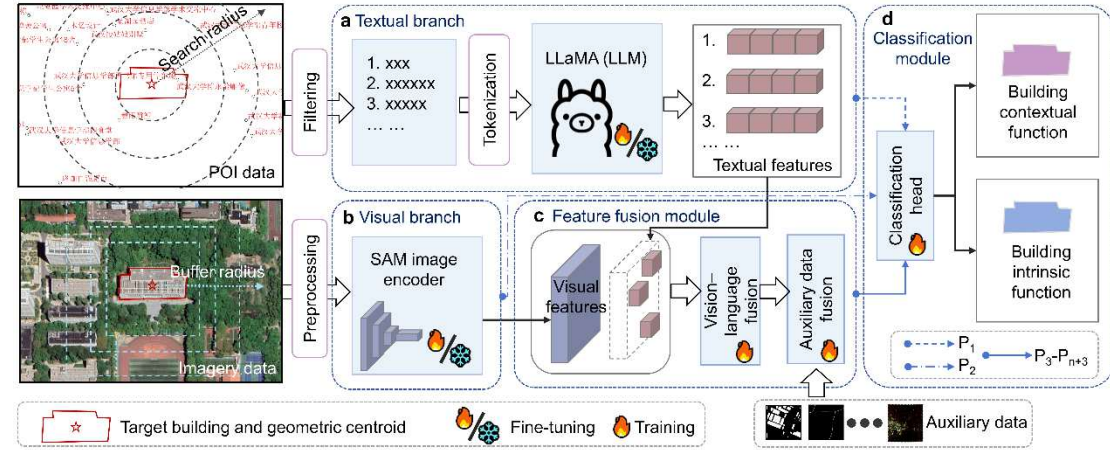


Fig. 4. The BHENet architecture for POI-aerial-vector fusion. The framework integrates multi-modal data to classify building functions through four main components. **a.** Textual branch. Points of interest (POIs) within a 500 m search radius are retrieved; the two nearest POIs for each category are retained and tokenized to generate text embeddings using the large language model LLaMA⁴⁰. **b.** Visual branch. A 0.2km×0.2km aerial imagery patch centered on the target building is normalized and encoded using the SAM image encoder⁴¹. **c.** Feature fusion module. This module performs spatial alignment and integrates vision-language features. An auxiliary branch embeds road networks and building footprints, which are subsequently fused with the multi-modal features. **d.** Classification head. The final module predicts the building contextual or intrinsic function.

Mixing and jobs-housing metrics

Functional mixing is quantified on 2-km grids using the Building Function Mixed Index (BFMI), defined as the normalized Shannon entropy of building-volume shares

by function. Building volume is computed as footprint area $\times$ height and aggregated by function within each grid cell. Let V_i be the total volume of function i and $p_i = V_i / \sum_{i=1}^N V_i$. Then,

$$\text{BFMI} = - \sum_{i=1}^N p_i \log_2 p_i / \log_2 N$$

where N denotes number of function categories. Sector-level BFMI is the arithmetic mean of cell values within each annular sector. We compute BFMI under both contextual and intrinsic labels to test scale divergence between neighborhood mixing and building mixing (**Supplementary Methods 4**).

To characterize mobility-related mechanisms within daily activity spaces, we compute two complementary jobs–housing separation indicators for each grid cell using a 10 km moving window. Let R_k and J_k denote residential (“housing”, refers to residential) and employment (“jobs”, including industrial, commercial, education, administrative, transport, public services, and medical) building volumes within the local window S_i of focal grid i , and let P_R and P_J be the corresponding locally normalized distributions. Local Expected average commuting distance (EAD) captures the distance channel as:

$$\text{EAD}_i = \sum_{k \in S_i} \sum_{l \in S_i} d_{k,l} \cdot p_R(k) \cdot p_J(l) \cdot \frac{1}{d_{\max}} \in [0,1]$$

where $d_{k,l}$ is the distance between grid centroids and $d_{\max} = 10\text{km}$. Jensen–Shannon divergence (JSD) captures morphological polarization between housing and jobs within the window as

$$\text{JSD}_i = \frac{1}{2} D_{\text{KL}}(P_i \parallel M_i) + \frac{1}{2} D_{\text{KL}}(Q_i \parallel M_i), \text{ where } M_i = \frac{1}{2} (P_i + Q_i)$$

bounded in $[0,1]$. EAD isolates the separation-by-distance channel, whereas JSD

captures spatial mismatch in the shape of distributions even when mean distance is similar. Additional details and robustness checks for EAD and JSD are provided in **Supplementary Methods 5.**

Outcome variable

Our dependent variable is annual, grid-level urban operational carbon derived from ODIAC^{42,43}. We use “urban operational carbon” as a shorthand for ODIAC-derived territorial, land-based carbon emissions associated with fossil-fuel combustion and related industrial-process sources within the urban built-up mask. Specifically, we used the 2020 monthly layers from the ODIAC2024 1 km GeoTIFF land product. ODIAC2024 denotes the product release version, whereas 2020 denotes the emission year used in this study. The ODIAC land product includes emissions from fossil-fuel combustion, cement production and gas flaring, while international aviation and marine bunker emissions are excluded from the 1 km land rasters. Because the original GeoTIFF values are reported as metric tons of carbon per cell per month, we summed the 12 monthly layers for 2020 and aggregated the annual totals to the 2 km × 2 km analysis grid using area-weighted zonal summation within the urban built-up mask. Emissions are therefore expressed as t C yr⁻¹ per grid cell. Accordingly, the outcome should be interpreted as territorial urban operational carbon, rather than sector-specific transportation emissions, building-only operating emissions or demand-based electricity emissions. See **Supplementary Table S8** for data sources and preprocessing details.

Spatial econometric identification and counterfactuals

We estimate the BFMI-carbon relationship using a Spatial Durbin Model (SDM) with city fixed effects and ring fixed effects, capturing systematic center-periphery gradients while allowing spillovers through a spatial weight matrix W . To recover non-linearity and saturation, BFMI enters the model through an orthogonal polynomial basis (with spline-based alternatives used as robustness checks). We control for confounding using a parsimonious set of covariates capturing (i) functional scale and composition (function-specific building volumes), (ii) mobility structure (EAD and JSD), (iii) infrastructure and accessibility (road length/density), and (iv) background socioeconomic and physical conditions (population density, night-time lights, and terrain controls). Full variable definitions, units, and construction details are provided in **Extended Data Table 1-2** and **Supplementary Methods 6**. We report marginal-effect curves (e.g., average marginal effects) and their derivatives to identify regime shifts (thresholds), with uncertainty quantified via delta-method confidence intervals and city-clustered standard errors.

We then evaluate static counterfactual abatement potentials by shifting each grid cell's BFMI toward an empirically identified target range (an "ideal" alignment and a "cost-effective" piecewise alignment). Emission changes are computed by integrating the estimated marginal-effect curve along the BFMI shift and aggregating grid-level changes to city and national totals; corresponding uncertainty bounds follow the same inference framework (**Supplementary Methods 6 and 7**).

Data availability

An interactive web portal, including map-based visualization of the building-function atlas, BFMI layers and independent manual validation results, is available at <https://zhzhyd.github.io/china-mixed-use-building-map/>. The portal allows users to inspect nationwide building-function layers and manually interpreted validation samples used to assess final-map accuracy. The building-resolved functional atlas for China will be made available through Zenodo upon publication at <https://zenodo.org/records/20466219>. The repository will provide regional GeoPackage archives containing building geometries and essential functional attributes, together with checksum and manifest files for data verification. Input datasets are from a combination of open and licensed sources. OSM AOIs and vector layers are available under the OpenStreetMap license. Building footprints and gridded building height used to compute building volume are available from the cited sources in Methods. Administrative boundaries were obtained from an official national standard-map service. VIIRS night-time light composites, population and terrain covariates are publicly available and are fully itemized in Supplementary Table S8. POIs from Gaode Map were accessed via the provider's open platform/API and are subject to the provider's terms of use; redistribution may be restricted. High-resolution aerial imagery from Google Earth was used for feature extraction and visualization under its user license and is not redistributed; the processing workflow is provided to enable re-access by other researchers subject to the same terms. For the global comparative analysis, city lists and all city-specific data sources (building data, function signals and emissions

layers) are described in **Supplementary Table S8**; these data are subject to their respective third-party licenses and access conditions.

Code availability

The custom code used for data processing, BHENet model implementation, model training and inference, BFMI calculation, and spatial econometric analysis is available from the corresponding author upon reasonable request.

Reference:

1. Seto, K. C., Güneralp, B. & Hutyrá, L. R. Global forecasts of urban expansion to 2030 and direct impacts on biodiversity and carbon pools. *Proc. Natl Acad. Sci. USA* **109**, 16083–16088 (2012).
2. UN-Habitat. *Global Report on Human Settlements 2011: Cities and Climate Change* (Earthscan, London, 2011).
3. Shukla, P. R. *et al.* (eds) *Climate Change 2022: Mitigation of Climate Change* (Cambridge Univ. Press, Cambridge, 2022).
4. Estrada, F., Lupi, V., Botzen, W. J. W. & Tol, R. S. J. Urban and non-urban contributions to the social cost of carbon. *Nat. Commun.* **16**, 4193 (2025).
5. Liu, H., Evans, S., Zhang, Z., Song, W. & You, X. The Carbon Brief Profile: China. *Carbon Brief* <https://interactive.carbonbrief.org/the-carbon-brief-profile-china/> (2023).
6. PBL Netherlands Environmental Assessment Agency. China now no. 1 in CO₂

-
- emissions; USA in second position. Press release at <https://www.pbl.nl/en/news/2007/china-now-no.-1-in-co2-emissions-usa-in-second-position> (2007).
7. Central Committee of the Communist Party of China & State Council of the People's Republic of China. *Working Guidance for Carbon Dioxide Peaking and Carbon Neutrality in Full and Faithful Implementation of the New Development Philosophy*. National Development and Reform Commission https://en.ndrc.gov.cn/policies/202110/t20211024_1300725.html (2021).
8. Fang, C., Wang, S. & Li, G. Changing urban forms and carbon dioxide emissions in China: a case study of 30 provincial capital cities. *Appl. Energy* **158**, 519–531 (2015).
9. Fan, T., Ren, Y. & Chapman, A. Unveiling the carbon neutrality pathways of compact cities: a simulation-based scenario analysis from China. *Humanit. Soc. Sci. Commun.* **12**, 1205 (2025).
10. Tian, P. *et al.* Effects of 3D urban morphology on CO₂ emissions using machine learning: towards spatially tailored low-carbon strategies in Central Wuhan, China. *Urban Clim.* **57**, 102122 (2024).
11. Creutzig, F., Baiocchi, G., Bierkandt, R., Pichler, P.-P. & Seto, K. C. Global typology of urban energy use and potentials for an urbanization mitigation wedge. *Proc. Natl Acad. Sci. USA* **112**, 6283–6288 (2015).
12. Fragkias, M., Lobo, J., Strumsky, D. & Seto, K. C. Does size matter? Scaling of CO₂ emissions and US urban areas. *PLoS ONE* **8**, e64727 (2013).

-
- 511 13. Louf, R. & Barthelemy, M. How congestion shapes cities: from mobility patterns
512 to scaling. *Sci. Rep.* **4**, 5561 (2014).
- 513 14. Chen, Y. *et al.* A building-scale modeling framework for urban net-zero transitions
514 in Nanjing. *Nat. Commun.* **16**, 8954 (2025).
- 515 15. Yap, W., Wu, A. N., Miller, C. & Biljecki, F. Revealing building operating carbon
516 dynamics for multiple cities. *Nat. Sustain.* **8**, 1199–1210 (2025).
- 517 16. Wang, G. *et al.* The estimation of building carbon emission using nighttime light
518 images: a comparative study at various spatial scales. *Sustain. Cities Soc.* **101**,
519 105066 (2024).
- 520 17. Fang, G., Gao, Z., Tian, L. & Fu, M. What drives urban carbon emission efficiency?
521 Spatial analysis based on nighttime light data. *Appl. Energy* **312**, 118772 (2022).
- 522 18. Güneralp, B. *et al.* Global scenarios of urban density and its impacts on building
523 energy use through 2050. *Proc. Natl Acad. Sci. USA* **114**, 8945–8950 (2017).
- 524 19. Ghorbany, S. & Hu, M. A systematic framework to reduce urban embodied carbon
525 emissions using urban scale simulation. *npj Urban Sustain.* **5**, 7 (2025).
- 526 20. Zhao, J., Mo, B., Caros, N. S. & Zhao, J. Housing exchange framework to reduce
527 carbon emissions from commuting. *Nat. Sustain.* **8**, 1259–1269 (2025).
- 528 21. Bettencourt, L. M. A. The origins of scaling in cities. *Science* **340**, 1438–1441
529 (2013).
- 530 22. Batty, M. The size, scale, and shape of cities. *Science* **319**, 769–771 (2008).
- 531 23. Verbavatz, V. & Barthelemy, M. The growth equation of cities. *Nature* **587**, 397–
532 401 (2020).

-
- 533 24. Ministry of Housing and Urban–Rural Development of the People’s Republic of
534 China. *Urban Land Use Classification and Planning Construction Land Standards*
535 *(GB 50137–2011)* (China Architecture Publishing & Media Co., Ltd., Beijing,
536 2011).
- 537 25. Wang, M. *et al.* Assessing urban vitality in high-density cities: a spatial
538 accessibility approach using POI reviews and residential data. *Humanit. Soc. Sci.*
539 *Commun.* **12**, 1119 (2025).
- 540 26. Yuan, H., Liu, J., Li, X. & Zhong, S. The impact of industrial collaborative
541 agglomeration on total factor carbon emission efficiency in China. *Sci. Rep.* **13**,
542 12347 (2023).
- 543 27. Ewing, R. & Cervero, R. Travel and the built environment: a meta-analysis. *J. Am.*
544 *Plan. Assoc.* **76**, 265–294 (2010).
- 545 28. Ahlfeldt, G. M. & Pietrostefani, E. The compact city in empirical research: a
546 quantitative literature review. *SERC Discussion Paper* **215**. Spatial Economics
547 Research Centre <http://eprints.lse.ac.uk/83638/> (2017).
- 548 29. Stevens, M. R. Does compact development make people drive less? *J. Am. Plan.*
549 *Assoc.* **83**, 7–18 (2017).
- 550 30. Glaeser, E. L. & Kahn, M. E. The greenness of cities: carbon dioxide emissions
551 and urban development. *J. Urban Econ.* **67**, 404–418 (2010).
- 552 31. Batty, M. *The New Science of Cities* (MIT Press, Cambridge, MA, 2013).
- 553 32. Ahlfeldt, G. M., Redding, S. J., Sturm, D. M. & Wolf, N. The economics of density:
554 evidence from the Berlin Wall. *Econometrica* **83**, 2127–2189 (2015).

-
- 555 33. Chauvin, J. P., Glaeser, E. L., Ma, Y. & Tobio, K. What is different about
556 urbanization in rich and poor countries? Cities in Brazil, China, India and the
557 United States. *J. Urban Econ.* **98**, 17–49 (2017).
- 558 34. Jones, C. & Kammen, D. M. Spatial distribution of US household carbon
559 footprints reveals suburbanization undermines greenhouse gas benefits of urban
560 population density. *Environ. Sci. Technol.* **48**, 895–902 (2014).
- 561 35. Cervero, R. Jobs-housing balancing and regional mobility. *J. Am. Plan. Assoc.* **55**,
562 136–150 (1989).
- 563 36. Batty, M. *Inventing Future Cities* (MIT Press, Cambridge, MA, 2018).
- 564 37. Barrington-Leigh, C. & Millard-Ball, A. The world’s user-generated road map is
565 more than 80% complete. *PLoS ONE* **12**, e0180698 (2017).
- 566 38. Yuan, J., Zheng, Y. & Xie, X. Discovering regions of different functions in a city
567 using human mobility and POIs. In *Proc. 18th ACM SIGKDD Int. Conf. Knowl.*
568 *Discov. Data Min.* 186–194 (ACM, 2012).
- 569 39. Biljecki, F. & Chow, Y. S. Global building morphology indicators. *Comput.*
570 *Environ. Urban Syst.* **95**, 101809 (2022).
- 571 40. Touvron, H. *et al.* Llama: open and efficient foundation language models. Preprint
572 at <https://arxiv.org/abs/2302.13971> (2023).
- 573 41. Kirillov, A. *et al.* Segment anything. In *Proc. IEEE/CVF Int. Conf. Comput. Vis.*
574 4015–4026 (IEEE, 2023).
- 575 42. Oda, T., Maksyutov, S. & Andres, R. J. The Open-source Data Inventory for
576 Anthropogenic CO₂, version 2016 (ODIAC2016): a global monthly fossil fuel CO₂

gridded emissions data product for tracer transport simulations and surface flux
inversions. *Earth Syst. Sci. Data* **10**, 87–107 (2018).

43. Oda, T. & Maksyutov, S. ODIAC Fossil Fuel CO₂ Emissions Dataset
(ODIAC2024) (Center for Global Environmental Research, National Institute for
Environmental Studies, 2025); <https://doi.org/10.17595/20170411.001>.

Acknowledgements

This work was supported by the National Key Research and Development
Program of China (2024YFF1306102) and the National Natural Science Foundation of
China (42271328 and 42471391).

Author contributions

X.H., J.L. and Z.Z. conceived and designed the study. Z.Z., J.L. and X.H.
developed the BHENet architecture, conducted model training and inference, and
performed the Spatial Durbin Model (SDM) analysis. Z.Z., X.P., S.M. and X.X. were
responsible for raw data curation, preprocessing, and accuracy validation. X.H., L.Z.,
J.L. and Z.Z. optimized the experimental design and interpreted the results. J.L. and
Z.Z. wrote the original draft of the manuscript. X.H. and L.Z. supervised the project,
reviewed the findings, and revised the manuscript.

Competing interests

The authors declare no competing interests.

Correspondence and requests for materials should be addressed to Xin Huang
(xhuang@whu.edu.cn).

Extended Data Figure Legends

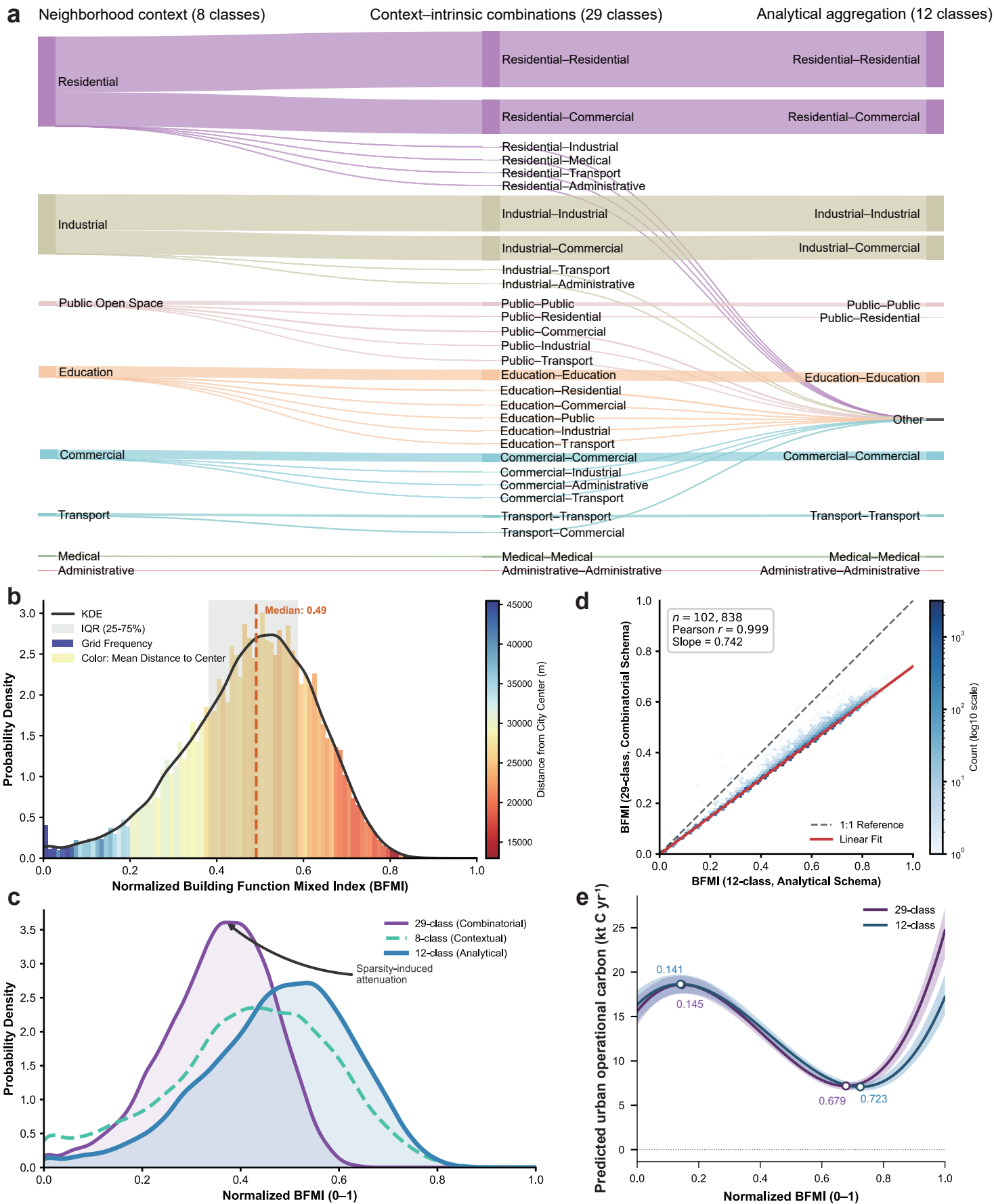
Extended Data Fig. 1: Taxonomic harmonization and statistical robustness of the building-scale functional mixing index. **a**, Sankey diagram illustrating the hierarchical mapping logic of the functional atlas. The taxonomy flows from the coarse-grained neighborhood context (left; 8 classes aligned with national land-use standards), to the fine-grained context–intrinsic combinations (middle; 29 types capturing specific mixed-use patterns), and finally aggregates into the robust analytical schema (right; 12 classes). Node colors represent the primary functional categories, and flow widths are proportional to total building volume. To ensure statistical stability, 18 sparse tail combinations are merged into the 'Other' category (1.184% of total volume). **b**, Distribution of the normalized BFMI (12-class analytical schema) across valid grid cells ($n = 102,838$). The distribution is unimodal (Median = 0.491, IQR = 0.202), with bar colors indicating the mean distance to the city center (Spectral scale), revealing a spatial gradient where mixing intensity correlates with urban centrality. **c**, Comparative probability density functions (KDE) for the three taxonomic schemas. The systematic leftward shift of the 29-class distribution (purple) relative to the 12-class baseline (blue) quantifies the sparsity bias (or denominator inflation): the normalization denominator ($\log_2 29 \approx 4.858$) increases significantly due to rare categories absent in most grids, artificially suppressing index values. **d**, Hexbin density scatter plot correlating the 12-

class and 29-class indices (color intensity represents grid count on a $\log_{10}$ scale). Despite the magnitude attenuation (regression slope ≈ 0.742), the exceptionally strong linear correlation (Pearson's $r > 0.999$) confirms that the aggregation strategy preserves the fundamental rank order and structural patterns of urban functional organization used in the main analysis. **e**, Comparison of fitted BFMI–carbon response curves for the 12-class and 29-class schemas, demonstrating that the functional form and turning points of the urban operational carbon response remain structurally consistent despite the scalar shift in BFMI values.

Extended Data Fig. 2 | Independent manual validation of the building-function

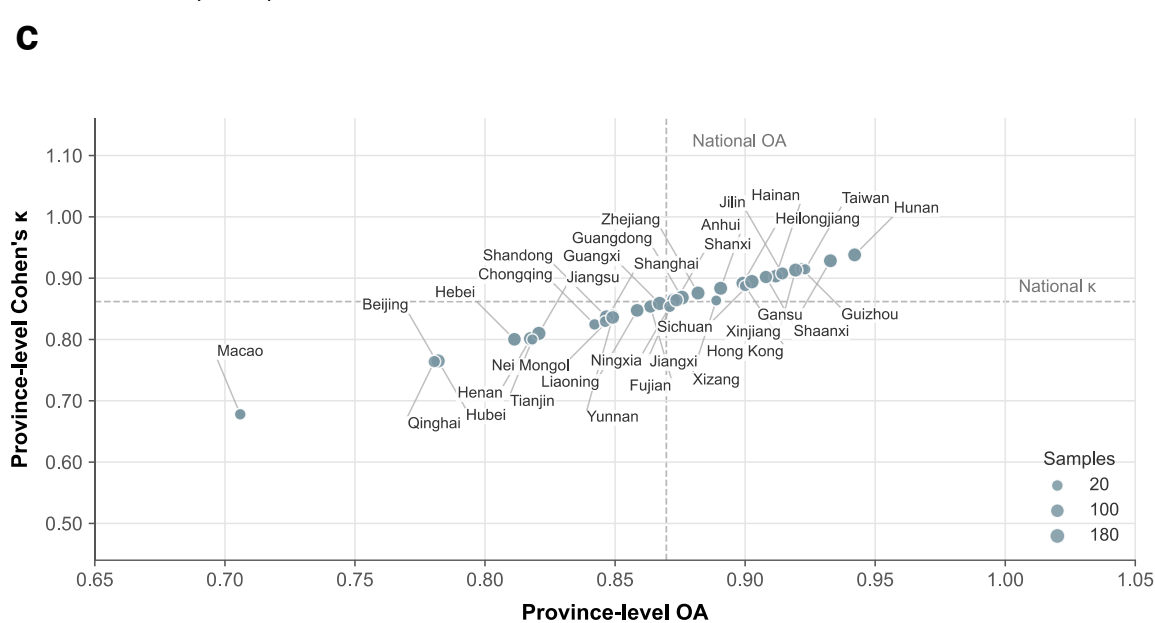
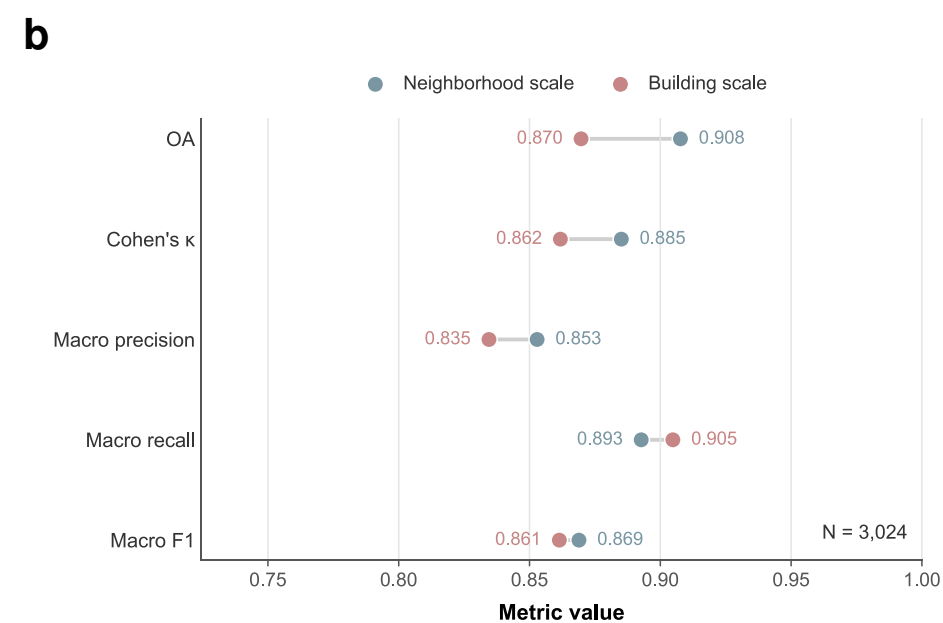
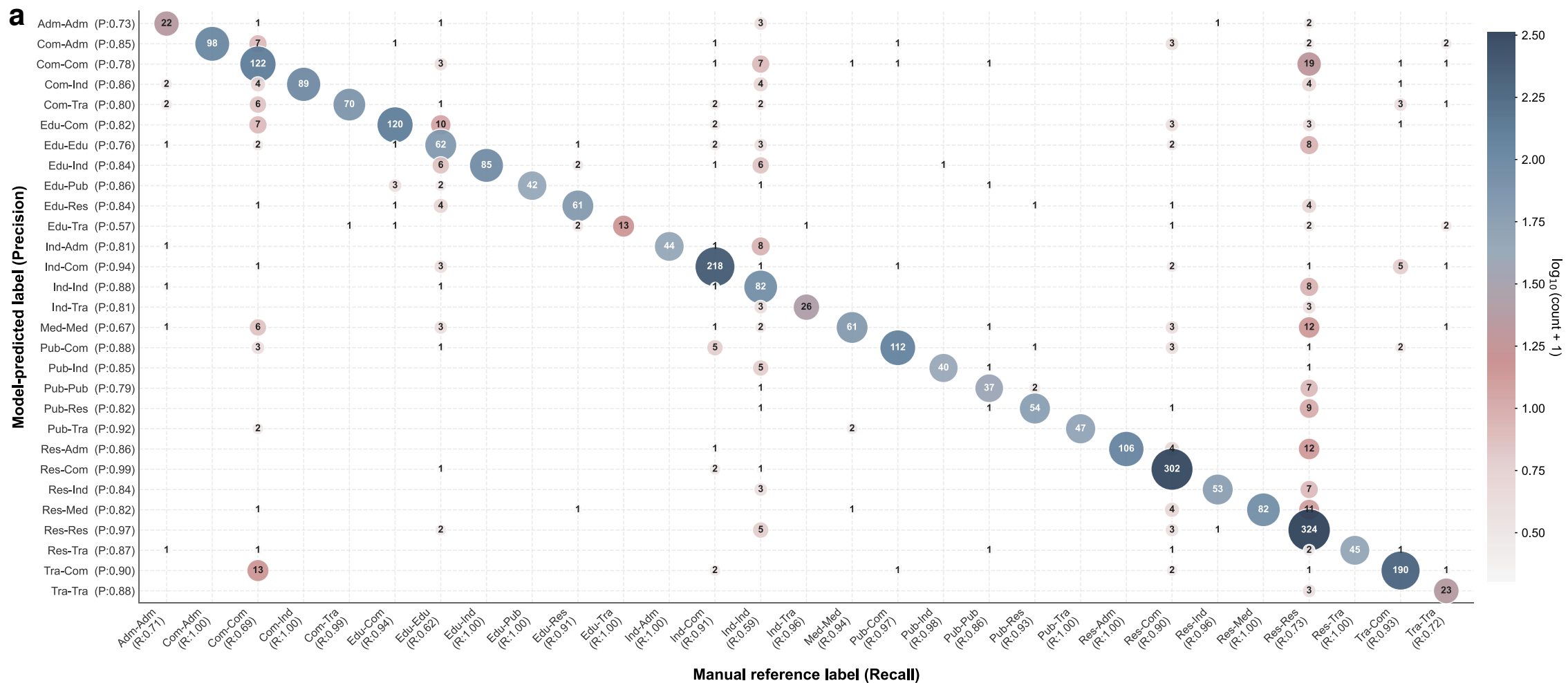
atlas. a, Bubble confusion matrix for the building-scale classification, comparing manually interpreted reference labels and model-predicted labels across the 29 retained context–intrinsic building-function categories ($N = 3,024$). The horizontal axis denotes manual reference labels, with class-level recall shown in parentheses, and the vertical axis denotes model-predicted labels, with class-level precision shown in parentheses. Bubble size and colour intensity represent the number of validation samples in each reference–prediction pair, shown on a log-transformed count scale. **b**, Summary validation metrics for the final mapped product at the neighborhood scale (8-class contextual dimension) and building scale (29-class context–intrinsic label system). Metrics include overall accuracy (OA), Cohen's κ , macro-precision, macro-recall and macro-F1. **c**, Province-level validation consistency for the building-scale classification. Each point represents one provincial-level region; point size indicates the number of manually interpreted validation samples, and dashed lines mark the national overall

639 accuracy and Cohen's κ estimated from the full building-scale validation set. Province-
640 level estimates are shown to characterize spatial consistency of the validation results
641 and should be interpreted with sample-size support.
642



Extended Data Fig. 1 | Taxonomic harmonization and statistical robustness of the building-scale functional mixing index.

a, Sankey diagram illustrating the hierarchical mapping logic of the functional atlas. The taxonomy flows from the coarse-grained **neighborhood context** (left; 8 classes aligned with national land-use standards), to the fine-grained **context-intrinsic combinations** (middle; 29 types capturing specific mixed-use patterns), and finally aggregates into the robust **analytical schema** (right; 12 classes). Node colors represent the primary functional categories, and flow widths are proportional to total building volume. To ensure statistical stability, 18 sparse tail combinations are merged into the 'Other' category (1.184% of total volume). **b**, Distribution of the normalized BFMI (12-class analytical schema) across valid grid cells ($n = 102,838$). The distribution is unimodal (Median = 0.491, IQR = 0.202), with bar colors indicating the mean distance to the city center (Spectral scale), revealing a spatial gradient where mixing intensity correlates with urban centrality. **c**, Comparative probability density functions (KDE) for the three taxonomic schemas. The systematic leftward shift of the 29-class distribution (purple) relative to the 12-class baseline (blue) quantifies the **sparsity bias** (or denominator inflation): the normalization denominator ($\log_{10} 29 \approx 4.858$) increases significantly due to rare categories absent in most grids, artificially suppressing index values. **d**, Hexbin density scatter plot correlating the 12-class and 29-class indices (color intensity represents grid count on a $\log_{10}$ scale). Despite the magnitude attenuation (regression slope ≈ 0.742), the exceptionally strong linear correlation (Pearson's $r > 0.999$) confirms that the aggregation strategy preserves the fundamental rank order and structural patterns of urban functional organization used in the main analysis. **e**, Comparison of fitted BFMI-carbon response curves for the 12-class and 29-class schemas, demonstrating that the functional form and turning points of the urban operational carbon response remain structurally consistent despite the scalar shift in BFMI values.



Extended Data Fig. 2 | Independent manual validation of the building-function atlas.

a, Bubble confusion matrix for the building-scale classification, comparing manually interpreted reference labels and model-predicted labels across the 29 retained context–intrinsic building-function categories ($N = 3,024$). The horizontal axis denotes manual reference labels, with class-level recall shown in parentheses, and the vertical axis denotes model-predicted labels, with class-level precision shown in parentheses. Bubble size and colour intensity represent the number of validation samples in each reference–prediction pair, shown on a log-transformed count scale. **b**, Summary validation metrics for the final mapped product at the neighborhood scale (8-class contextual dimension) and building scale (29-class context–intrinsic label system). Metrics include overall accuracy (OA), Cohen's κ , macro-precision, macro-recall and macro-F1. **c**, Province-level validation consistency for the building-scale classification. Each point represents one provincial-level region; point size indicates the number of manually interpreted validation samples, and dashed lines mark the national overall accuracy and Cohen's κ estimated from the full building-scale validation set. Province-level estimates are shown to characterize spatial consistency of the validation results and should be interpreted with sample-size support.

Extended Data Table 1: Variable definitions, construction protocols, and model specifications.

Category	Variable	Description and Construction Logic	Unit	Data Source	Model Role
Outcome	Urban operational carbon	Annual grid-level urban operational carbon derived from ODIAC. It represents ODIAC-derived territorial, land-based carbon emissions within the urban built-up mask. The 2020 monthly ODIAC2024 1 km GeoTIFF land layers were summed to annual carbon mass and aggregated to 2 km × 2 km grids using area-weighted zonal summation.	t C yr ⁻¹	ODIAC2024, 2020 monthly layers	Dependent Variable (y)
Core Regressor	BFMI	Building Function Mixed Index. Normalized Shannon entropy of volume-weighted functional shares. Computed using the 12-class analytical schema (building-scale) or 8-class contextual schema (neighborhood-scale). Entered as orthogonal polynomial basis (cubic Legendre) to capture non-linearities.	Index [0, 1]	BHENet Functional Atlas	Main Regressor (X)
Jobs–Housing	EAD	Expected Average Commuting Distance. Localized metric capturing the expected physical separation between housing and employment within a 10 km moving window centered on the grid cell. Distances are normalized by the 10 km moving-window radius for model estimation; main-text effects are rescaled to a +1 km change.	Normalized distance index [0, 1]	Functional Atlas (incl. Geometry)	Control (Z)
	JSD	Jensen–Shannon Divergence. Localized metric capturing the morphological dissimilarity between residential and employment distributions within the same 10 km moving window. Represents spatial polarization independent of distance.	Index [0, 1]	Functional Atlas	Control (Z)
Functional Volume	Vol _R	Residential Volume. Total volume of buildings classified as Residential at the contextual (neighborhood) scale.	10 ⁴ m ³	Functional Atlas	Control (Z)
	Vol _I	Industrial Volume. Total volume of buildings classified as Industrial at the	10 ⁴ m ³	Functional Atlas	Control (Z)

Category	Variable	Description and Construction Logic	Unit	Data Source	Model Role
		contextual scale.			
	Vol_C	Commercial Volume. Total volume of buildings classified as Commercial at the contextual scale.	10^4 m^3	Functional Atlas	Control (Z)
	Vol_E	Education Volume. Total volume of buildings classified as Education at the contextual scale.	10^4 m^3	Functional Atlas	Control (Z)
	Vol_A	Administrative Volume. Total volume of buildings classified as Administrative at the contextual scale.	10^4 m^3	Functional Atlas	Control (Z)
	Vol_M	Medical Volume. Total volume of buildings classified as Medical at the contextual scale.	10^4 m^3	Functional Atlas	Control (Z)
	Vol_T	Transport Volume. Total volume of buildings classified as Transport at the contextual scale.	10^4 m^3	Functional Atlas	Control (Z)
	Vol_P	Public Open Space Volume. Total volume of buildings classified as Public Open Space at the contextual scale.	10^4 m^3	Functional Atlas	Control (Z)
Physical & Activity	$Road_{Len}$	Road Network Density. Total length of all road segments within the grid cell, capturing transport infrastructure intensity.	km	OpenStreetMap (OSM)	Control (Z)
	Pop_{Den}	Population Density. Grid-level residential population count, derived from the high-resolution 7th National Census gridded dataset using ensemble learning.	per km^2	7th Census Grid	Control (Z)
	NTL	Night-Time Light Intensity. Mean radiance of stable night-time lights, proxying overall socio-economic activity and urbanization level.	$\text{nW cm}^{-2} \text{ sr}^{-1}$	VIIRS (VNP46A4)	Control (Z)
Topography	Elev	Elevation. Mean elevation above sea level within the grid, controlling for terrain-induced transport energy penalties.	m	Copernicus DEM	Control (Z)
	Slope	Slope. Mean terrain slope, capturing ruggedness constraints on mobility and development.	Degree ($^\circ$)	Derived from DEM	Control (Z)
	Relief	Local Relief. Standard deviation of elevation within the grid, representing topographic complexity.	m	Derived from DEM	Control (Z)
Fixed Effects	$City_{FE}$	City Fixed Effects. Dummy variables for each city to absorb time-invariant, city-specific unobservables (e.g., energy mix,	Binary (0/1)	Administrative Boundaries	Fixed Effect (α_{city})

Category	Variable	Description and Construction Logic	Unit	Data Source	Model Role
		industrial base).			
	Ring _{FE}	Ring Fixed Effects. Dummy variables for concentric distance bands (5 km intervals) from the city center(s), absorbing systematic center–periphery gradients.	Binary (0/1)	City Center Identification	Fixed Effect (α_{ring})

Note:

Data Sources: Functional volumes and BFMI are derived from the BHENet-generated national building atlas (**Supplementary Methods 1–3**). For full details on external datasets (ODIAC, 7th Census Grid, VIIRS, Copernicus DEM), including versioning, licensing, and citations, please refer to the Data Dictionary in **Supplementary Table S8**.

Construction Details: Continuous covariates were standardized during model estimation to improve numerical stability. For interpretation, the marginal effects reported in Extended Data Tables 2–3 are expressed in the physical or index units listed here after rescaling, except for the BFMI Legendre basis and the normalized EAD index in Extended Data Table 2. Model Role: y , dependent variable; X , explanatory variable of interest (including spatial lags WX); Z , control covariates (including spatial lags WZ); Fixed Effects, dummy variables absorbing unobserved heterogeneity. EAD/JSD: Calculated using a 10 km moving window radius to capture localized jobs–housing balance (see **Supplementary Methods 5**).

Extended Data Table 2: Spatial Durbin Model estimates of direct, indirect and total marginal effects.

Variable	Direct Effect	Indirect Effect (Spillover)	Total Effect
Core Regressor			
BFMI (Legendre P_1)	-986.0 (157.1) ***	-4,012.5 (415.2) ***	-4,998.5 (420.5) ***
BFMI (Legendre P_2)	48.8 (137.0)	3,848.7 (349.6) ***	3,897.5 (355.1) ***
BFMI (Legendre P_3)	735.1 (153.8) ***	5,314.2 (437.6) ***	6,049.3 (450.9) ***
Jobs–Housing Metrics			
EAD	-801.5 (463.2) *	6,896.7 (564.0) ***	6,095.2 (305.7) ***
JSD	-306.6 (839.6)	2,790.8 (893.3) **	2,484.1 (258.4) ***
Controls: Functional Volume			
Vol _R (Residential)	-0.78 (0.20) ***	-6.93 (0.32) ***	-7.71 (0.27) ***
Vol _I (Industrial)	7.21 (0.22) ***	20.62 (0.38) ***	27.83 (0.33) ***
Vol _C (Commercial)	28.11 (0.64) ***	48.50 (1.38) ***	76.62 (1.25) ***
Vol _E (Education)	3.28 (0.59) ***	-14.87 (1.27) ***	-11.59 (1.23) ***
Vol _P (Public Open Space)	8.76 (1.03) ***	-2.20 (1.80)	6.57 (1.59) ***
Vol _M (Medical)	63.51 (2.26) ***	150.92 (6.47) ***	214.43 (6.58) ***
Vol _A (Administrative)	13.51 (2.15) ***	-5.59 (4.71)	7.92 (4.59) *
Vol _T (Transport)	6.13 (0.99) ***	23.53 (2.12) ***	29.66 (2.06) ***
Controls: Physical & Activity			
Road _{Len}	-15.58 (4.22) ***	-8.28 (5.64)	-23.86 (4.01) ***
Pop _{Den}	0.47 (0.02) ***	1.14 (0.04) ***	1.62 (0.03) ***
NTL	25.59 (0.36) ***	34.77 (0.50) ***	60.36 (0.33) ***
Controls: Topography			
Elev	-6.66 (1.67) ***	5.82 (1.67) ***	-0.84 (0.12) ***
Slope	130.2 (18.2) ***	-355.2 (20.8) ***	-225.0 (10.3) ***
Relief	702.7 (134.8) ***	448.5 (147.4) ***	1,151.2 (59.4) ***
Model Diagnostics			
Observations (N)	91,111		
Pseudo R_2	0.7833		
Spatial ρ	0.103***		
Log-Likelihood	-940,816.7		

Note: Estimates are derived from the baseline Spatial Durbin Model (SDM) utilizing the 12-class volume-weighted BFMI on a 2 km × 2 km grid system across China; the final analytical sample comprises $N = 91,111$ grid cells, following the exclusion of cells with missing covariates, zero population, or no spatial neighbors (isolates) required for spatial lag construction. The dependent variable is annual grid-level urban operational carbon derived from ODIAC, reported in carbon-

mass units (t C yr⁻¹ per grid cell). The model includes city fixed effects and ring fixed effects to control for time-invariant and center–periphery unobserved heterogeneity. **Core Regressor:** BFMI enters the model as an orthogonal cubic Legendre polynomial basis (P_1, P_2, P_3) to capture non-linear responses; reported coefficients correspond to these transformed basis functions rather than the raw index. **Effects:** Coefficients are decomposed into Direct (local feedback), Indirect (spatial spillover), and Total marginal effects. Unless otherwise stated, effects are reported as changes in urban operational carbon (t C yr⁻¹) per unit increase in each variable as defined in **Extended Data Table 1**. **Model Diagnostics:** The spatial autoregressive coefficient ρ is estimated at 0.103 ($P < 0.01$), indicating significant positive spatial dependence in emissions, validating the spatial specification. The Log-Likelihood is -940,816.7, consistent with the large sample size. **Units:** Functional volumes are measured in 10⁴ m³; road network length in km; population density in people km⁻²; NTL in nW cm⁻² sr⁻¹; elevation/relief in m; slope in degrees. EAD coefficients in this table correspond to the normalized 10 km moving-window distance index; main-text EAD effects are rescaled to a +1 km change. **Inference:** Standard errors (in parentheses) are clustered at the city level to account for within-city correlation. Significance levels: * $P < 0.1$, ** $P < 0.05$, *** $P < 0.01$. See **Extended Data Table 1** for detailed variable definitions.

Extended Data Table 3: Heterogeneity of jobs–housing separation effects on operational carbon across city tiers.

City Stratification	Sample Size (N)	EAD Total Effect (t C yr ⁻¹ per +1 km)	JSD Total Effect (t C yr ⁻¹ per +0.1 unit)
Tier 1 Cities	9,714	2,116.6 ± 465.8*** [1,203.7, 3,029.5]	1,320.9 ± 271.3*** [789.1, 1,852.7]
Tier 2 Cities	20,149	-9.0 ± 115.9 ^{ns} [-236.0, 218.1]	329.8 ± 82.2*** [168.7, 490.9]
Tier 3 Cities	61,248	567.2 ± 41.0*** [486.8, 647.5]	227.1 ± 36.5*** [155.6, 298.6]

Note: Values represent the total marginal effect on annual grid-level urban operational carbon, reported in tonnes of carbon per year (t C yr⁻¹), derived from the Spatial Durbin Model stratified by city tier. Data are presented as Estimate ± Standard Error [95% Confidence Interval]. Sample Definitions: Tier 1: Urban built-up area > 3,000 km² (e.g., Beijing, Shanghai); Tier 2: Built-up area 800–3,000 km²; Tier 3: Built-up area < 800 km². Regressors: EAD Total Effect indicates the change in urban operational carbon associated with a 1 km increase in local Expected Average Commuting Distance. JSD Total Effect indicates the change in urban operational carbon associated with a 0.1 unit increase in Jensen-Shannon Divergence (morphological mismatch). Significance: *** $P < 0.01$, ** $P < 0.05$, * $P < 0.10$, ^{ns} Not Significant ($P \geq 0.10$). Interpretation: The impact of spatial separation varies by urban scale. In Tier 1 cities, both movement impedance (EAD) and morphological polarization (JSD) are associated with stronger carbon penalties, reflecting acute congestion and structural imbalance. In Tier 2 cities, the EAD effect is statistically insignificant, likely reflecting high internal heterogeneity (e.g., polycentric manufacturing hubs vs. monocentric capitals) where distance does not linearly translate to carbon flux, whereas morphological mismatch (JSD) remains a significant driver. Tier 3 cities show moderate but highly consistent penalties across both channels.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryInformation.pdf](#)
- [SupplementaryInformationrevised17841333452.pdf](#)