

Supplementary Information

Daily patterns of global urban populations reveal how cities breathe

Andrew Zimmer, Daniel Adams, Lucas Poudyal and Marie Urban

Supplementary Validation

Temporal orientation in Helsinki

The Helsinki dynamic population dataset provides 24 modeled weekday distributions on 250-m cells using mobile-phone activity, land use and building data; it is a time-resolved comparator rather than direct ground truth.¹ We reprojected and aggregated LandScan Mosaic day and night to the aligned 250-m grid without changing their totals, then compared each state with every Helsinki hour (Fig. S1). We defined daytime as the mean from 07:00 to 18:59 and nighttime as the mean from 19:00 to 06:59. These equal 12-hour windows avoid selecting favorable single hours but are not universal definitions of day and night.

On the 250-m grid, the Helsinki composites correlated with LandScan Mosaic at $r = 0.550$ for day and $r = 0.678$ for night. After both datasets were aggregated to common 1-km cells and normalized, agreement increased to $r = 0.782$ and $r = 0.913$, respectively. The hourly curves also show the expected ordering: agreement with LandScan Mosaic night is highest overnight, whereas agreement with LandScan Mosaic day is highest during working hours. Because both products are modeled, this comparison supports temporal orientation rather than occupancy accuracy.

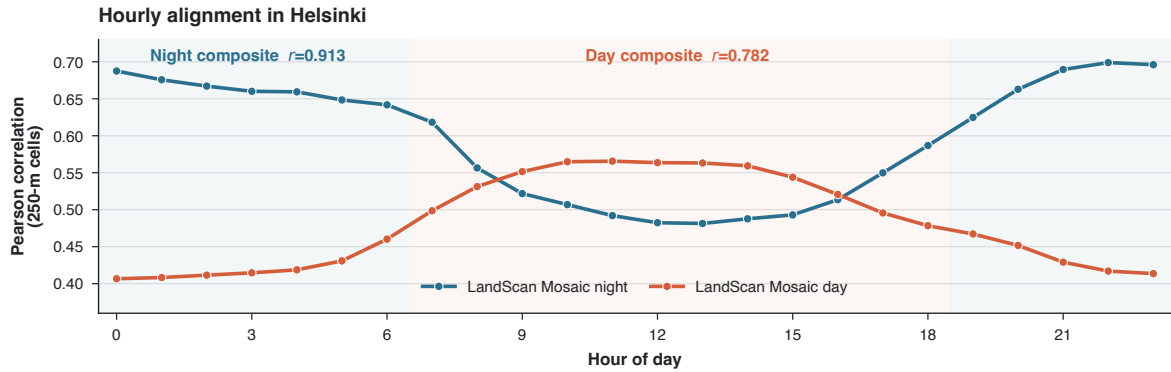


Figure S1: Temporal orientation in Helsinki. Correlation of LandScan Mosaic day and night with hourly weekday distributions on 250-m cells. Shading marks the defined 12-hour night (19:00–06:59) and day (07:00–18:59) windows; annotations report composite correlations on a common 1-km grid.

Official residential-grid comparisons

We separately compared LandScan Mosaic night with three official residential sources: the 2015 Helsinki population register on 250-m cells, Brazil’s 2022 Census on 200-m urban cells and Japan’s 2020 census population grid on 1-km mesh cells.^{2–4} We clipped Helsinki and five urban centers in each of Brazil and Japan to their GUEs. Helsinki and Brazilian counts were aggregated to aligned 1-km cells; Japan was already at 1 km. LandScan Mosaic night was aggregated to the same cells, and each pair was normalized within its location.

Across the 11 comparisons, Pearson r ranged from 0.829 to 0.945, Spearman ρ from 0.828 to 0.952 and JS divergence from 0.024 to 0.075 (Fig. S2; Table S1). The agreement supports the broad within-center geography of LandScan Mosaic night but is not a formal error estimate: the references describe registered residence rather than all nighttime presence, use different years and may share census controls with LandScan Mosaic.

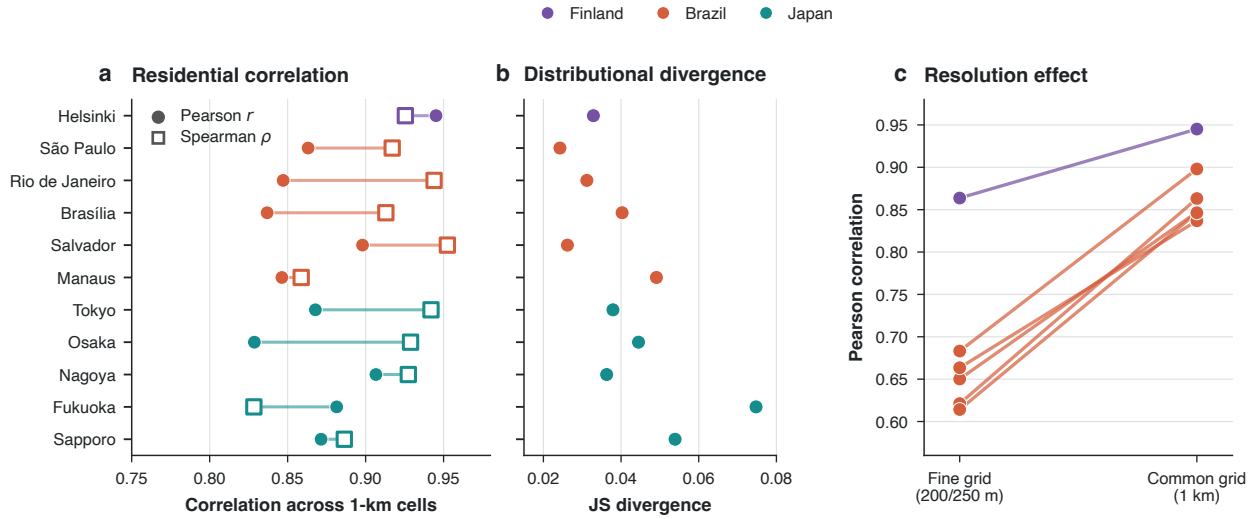


Figure S2: Agreement with official residential grids. a,b, Correlation and JS divergence on a common 1-km grid; circles show Pearson r , squares show Spearman ρ and color denotes country. Lower JS values indicate closer distributions. c, Pearson correlation at the original resolution and after aggregation to 1 km where both were available. Metrics compare normalized within-location shares.

Table S1: Agreement between LandScan Mosaic night and official residential grids. Metrics compare normalized population shares at the original resolution and after aggregation to common 1-km cells. Original-resolution values are omitted where the reference grid was already 1 km.

Location	Native resolution	Native r	Native ρ	1-km r	1-km ρ	1-km JS
Helsinki	250 m	0.864	0.843	0.945	0.925	0.033
São Paulo	200 m	0.621	0.799	0.863	0.917	0.024
Rio de Janeiro	200 m	0.650	0.871	0.847	0.944	0.031
Brasília	200 m	0.663	0.846	0.837	0.913	0.040
Salvador	200 m	0.683	0.856	0.898	0.952	0.026
Manaus	200 m	0.614	0.703	0.846	0.859	0.049
Tokyo	1 km	—	—	0.868	0.942	0.038
Osaka	1 km	—	—	0.829	0.929	0.044
Nagoya	1 km	—	—	0.907	0.927	0.036
Fukuoka	1 km	—	—	0.881	0.828	0.075
Sapporo	1 km	—	—	0.871	0.886	0.054

Supplementary Robustness

Urban-boundary concepts

Different urban boundaries describe different spatial units: jurisdictions, contiguous built settlements, population-defined centers or functional regions connected by commuting and accessibility.^{5–7} We selected GUE because it matches the 2025 population year, includes small named settlements and does not use gridded population counts to define the outcome analyzed here. GUE is not the only population-independent physical boundary; Global Urban Boundaries, for example, use impervious surface and rules that join nearby built areas and fill internal gaps.⁸ Population-thresholded urban centers and wider functional regions remain useful for different questions (Table S2).

Table S2: Urban-boundary concepts relevant to the analysis. Population dependence refers to whether population contributes to delineating the unit, not whether population is later summarized within it.

Boundary	Delineation concept	Uses population?	Principal strength	Main difference from GUE
GUE ^{9,10}	Built form, nighttime lights and place labels	No	Current global physical settlements, including small centers	Primary unit used here
GHS Urban Centre ^{5,11}	Contiguous built-up and population-density thresholds	Yes	Harmonized Degree of Urbanisation definition	Population partly determines membership
Global Urban Boundaries ⁸	Impervious surface, density and morphological closure	No	Fine-resolution physical urban extents over time	Earlier epoch and different rules for infill and fragmentation
GHS Functional Urban Area ⁵	Urban centre plus commuting zone	Yes	Captures labor-market relationships beyond the core	Wider functional rather than physical settlement
Multi-Tier City–Region (1 h) ⁷	Population-tiered centers linked by travel-time catchments	Yes	Potential access to employment and services across several urban tiers	Primary regional system containing secondary centers and hinterland
Administrative city	Statutory jurisdiction	Varies	Direct connection to governance and policy	Definitions and extents are not globally harmonized

Boundary and resolution sensitivity

We first tested the exact position of the urban edge by moving each GUE boundary inward and outward by 250 m in a geographically and size-stratified panel of 60 centers. Every center retained its daytime-increase, approximately balanced or daytime-decrease category. Spearman rank correlations with the original GUE were 0.989 and 0.999 for total balance and 0.995 and 0.998 for native-cell JS after the inward and outward adjustments, respectively.

We also aggregated the native 3-arc-second population cells to 9, 18 and 36 arcseconds—approximately 270 m, 540 m and 1.08 km at the equator—in the same panel. Median JS declined from 0.179 at native resolution to 0.138, 0.098 and 0.065, as expected when neighboring gains and losses are combined. Center ordering remained strong: correlations with native-resolution JS were 0.981, 0.945 and 0.910, respectively. Fine-redistribution magnitude is therefore resolution dependent, while its comparative ordering is substantially retained.

Alternative urban boundaries

We re-extracted LandScan Mosaic day and night within the complete public boundaries of GHS Urban Centres, GHS Functional Urban Areas and one-hour Multi-Tier City–Regions (MTCR).^{7,12–14} GHS Urban Centres combine built-up and population-density thresholds; Functional Urban Areas add modeled commuting zones. MTCR instead joins population-tiered centers through one-hour travel-time catchments that represent potential access to employment and services, not observed commuting. We combined its 48,990 patches into 4,210 primary regions before extraction so that JS was calculated within each complete region. Applying the same threshold of at least 1,000 nighttime residents retained 12,931 Urban Centres, 9,027 Functional Urban Areas and 4,207 MTCR regions.

To separate differences in the boundaries from differences in the places represented, we also matched the same centers across all four products. GUEs and GHS Urban Centres were linked when each contained the other’s reference point within the same country; Functional Urban Areas used their published Urban Centre links; and each MTCR used its Tier-1 center. This produced 2,193 unique matches without using LandScan population. The matched set spans 160 countries (828 centers in Asia, 569 in Africa, 487 in the Americas, 290 in Europe and 19 in Oceania) but favors larger functional centers, so it tests sensitivity among comparable places rather than estimating global prevalence.

Across the matched size classes, median JS differed by at most 0.018 among the four boundary definitions and showed the same rise toward larger centers after the two smallest classes (Fig. S3f). MTCR total changes remained smaller because its wider, multi-center regions contain more offsetting gains and losses (Fig. S3e). Internal redistribution is visible under all four definitions; total balance depends more strongly on whether the unit is a settled footprint or the broader functional system around it.

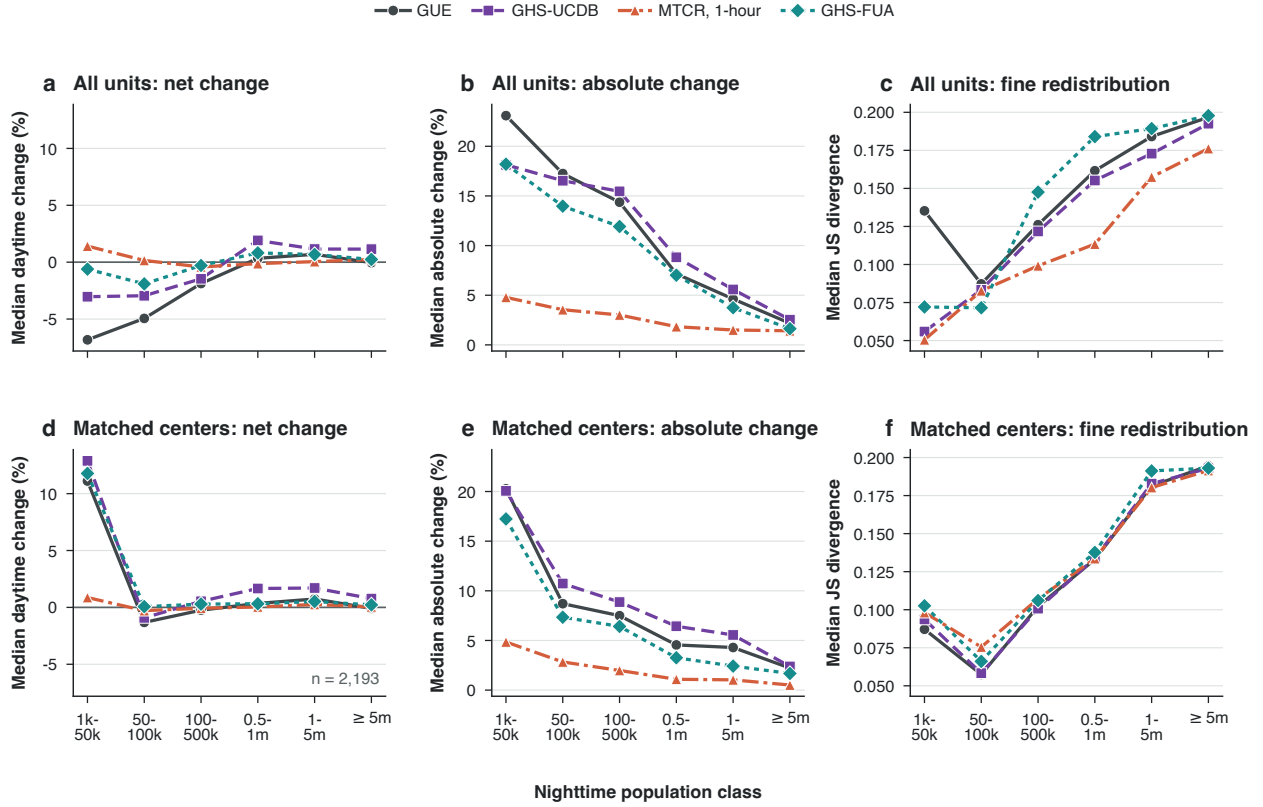


Figure S3: Sensitivity to urban delineation. Median net change, absolute change and native-cell JS by nighttime population class. **a–c**, All eligible units under each definition. **d–f**, The same 2,193 matched centers represented in all four products. Column scales are fixed across rows; wider MTCR catchments damp total changes, while matched JS profiles remain similar.

Spatial resolution and the population gradient

Fig. 2b asks the same day–night question at two spatial scales. Native-cell JS measures change among approximately 90-m cells across each GUE. Because median equivalent radius rises from 1.22 km in the smallest population class to 20.68 km in the largest, this comparison encompasses progressively more land, cells and local variation. The 32-zone measure instead uses the same number and relative arrangement of zones in every center, although each zone is physically larger in a larger GUE. Median native-cell JS falls from 0.135 to 0.087 between the first two classes and then rises to 0.197; median 32-zone JS declines from 0.048 to 0.008 (Fig. S4a,b). The measures remain strongly correlated across centers ($\rho = 0.902$), but describe change at different spatial scales.

The regression models described in the main Methods clarify this contrast. After accounting for region, population was positively associated with native-cell JS ($\beta = 0.095$) but negatively associated with 32-zone JS ($\beta = -0.142$). After also accounting for polygon area, both associations were negative (-0.446 and -0.507 , respectively); using valid-cell count in place of area produced similar estimates (-0.455 and -0.528 ; Fig. S4c). These associations show that the increasing physical extent of larger centers produces the apparent rise in fine-cell redistribution, whereas broad redistribution still declines after accounting for either area or cell count.

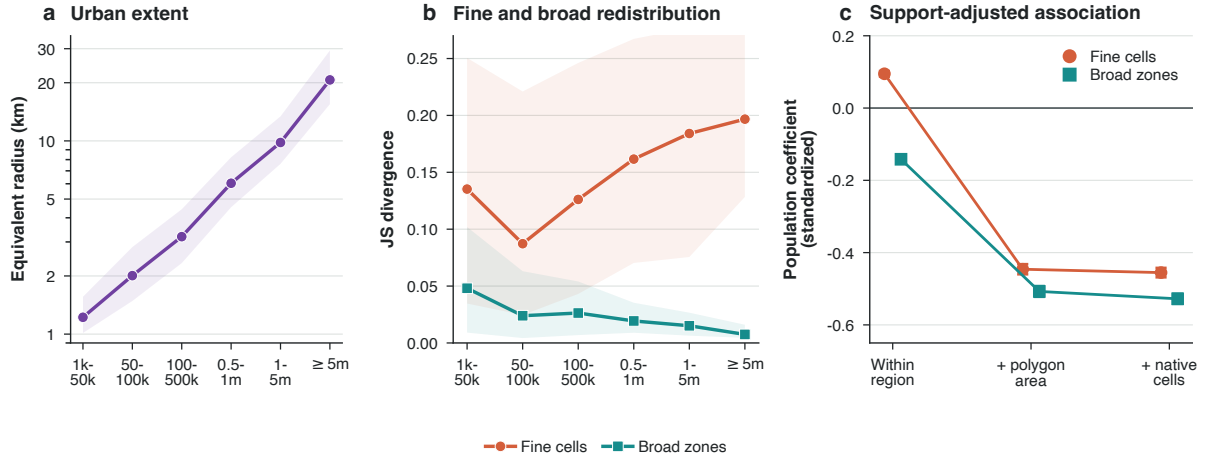


Figure S4: Population scale and measurement resolution. **a,b**, Median equivalent radius and JS divergence by nighttime population class; shading is the interquartile range. Broad redistribution uses the same 32 relative zones in every center. **c**, Standardized population coefficients after accounting for region, polygon area or valid-cell count.

Sensitivity to LandScan Mosaic confidence

LandScan Mosaic is the only major global gridded population product that supplies a formal cell-level confidence surface with its population estimates.¹⁵ It propagates uncertainty in building use, floor count and occupancy through Monte Carlo simulation. Cell confidence is calculated from the spread between the simulated 5th and 95th percentiles, $1 - (P_{95} - P_5)/(P_{95} + P_5)$; values near one indicate a narrower interval and therefore higher confidence. The same confidence surface accompanies the day and night grids.

For each GUE, we averaged cell confidence after weighting cells by their combined day and night population, $D_i + N_i$, where D_i and N_i are the two population counts. This gives equal treatment to the day and night states and produced a score for 49,701 centers. We recalculated the two main population-size profiles for the full sample and after progressively retaining only the highest-confidence 90% through 10% (Fig. S5); score cutoffs and class counts are in the Source Data. Both relationships persisted: similar day and night totals became more common with population, and native-cell JS rose after the two smallest classes. Retaining the highest-confidence 80%, for example, left 39,761 centers and reproduced the full sample's non-monotonic JS profile. The most restrictive subsets are less stable because their class sizes and geographic composition change.

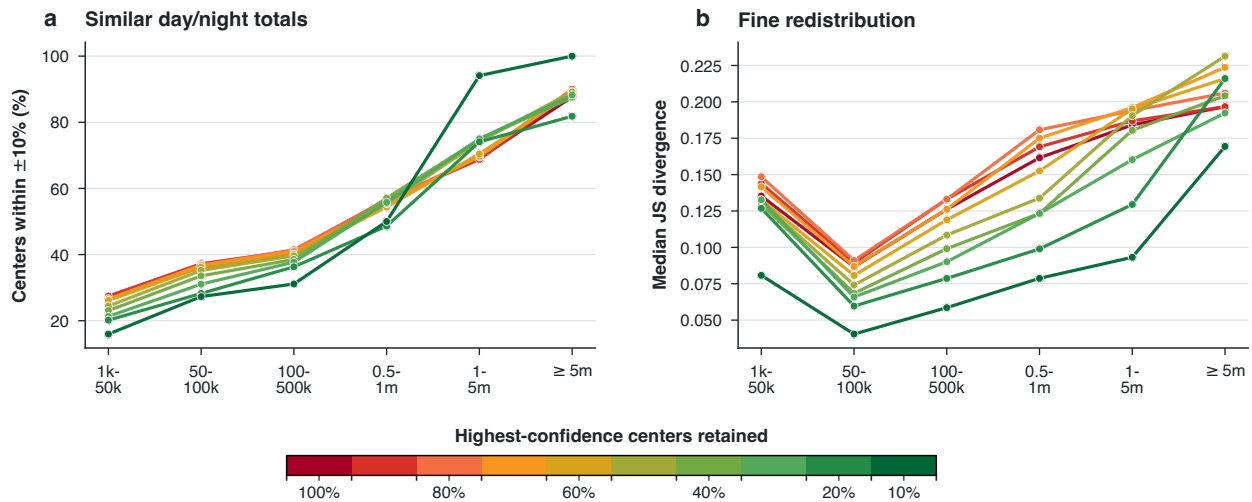


Figure S5: Robustness across LandScan Mosaic confidence. Profiles after progressively retaining the highest-confidence centers. **a**, Share with daytime and nighttime totals within 10%. **b**, Median native-cell JS. Red includes all 49,701 covered centers; green retains the highest-confidence 10%; intermediate lines show successive deciles. Lines are descriptive summaries within population classes, not confidence intervals.

References

1. Bergroth, C., Järv, O., Tenkanen, H., Manninen, M. & Toivonen, T. A 24-hour population distribution dataset based on mobile phone data from Helsinki metropolitan area, Finland. *Sci. Data* **9**, 39 (2022).
2. Helsinki Region Environmental Services Authority (HSY). Population grid of Helsinki metropolitan area. Helsinki Region Infoshare [Dataset], 2015 release (2015). <https://hri.fi/data/en/dataset/vaestotietoruudukko>. Accessed 26 August 2026.
3. Brazilian Institute of Geography and Statistics (IBGE). Censo Demográfico 2022: Grade Estatística. IBGE [Dataset] (2025). <https://mapasinterativos.ibge.gov.br/grade2022/>. Accessed 26 August 2026.
4. Ministry of Land, Infrastructure, Transport and Tourism (MLIT). 1-km mesh future population projections. National Land Numerical Information [Dataset] (2024). <https://nlftp.mlit.go.jp/ksj/gml/datalist/KsjTmplt-mesh1000r6.html>. Accessed 26 August 2026.
5. Dijkstra, L. *et al.* Applying the Degree of Urbanisation to the globe: A new harmonised definition reveals a different picture of global urbanisation. *J. Urban Econ.* **125**, 103312 (2021).
6. Van Migerode, C., Poorthuis, A. & Derudder, B. Proxy datasets to identify city boundaries. *Nat. Cities* **3**, 582–590 (2026).
7. Cattaneo, A. *et al.* Worldwide delineation of multi-tier city–regions. *Nat. Cities* **1**, 469–479 (2024).
8. Li, X. *et al.* Mapping global urban boundaries from the global artificial impervious area (GAIA) data. *Environ. Res. Lett.* **15**, 094044 (2020).
9. Liu, Z., Huang, S., Fang, C., Guan, L. & Liu, M. Global urban and rural settlement dataset from 2000 to 2020. *Sci. Data* **11**, 1359 (2024).
10. Liu, Z., Huang, S., Fang, C., Guan, L. & Liu, M. Global Urban and Rural Settlements (GURS) and Global Urban Envelopes (GUE) Dataset. Zenodo [Dataset], version 4 (2026). <https://doi.org/10.5281/zenodo.19630155>.
11. Melchiorri, M. *et al.* The multi-temporal and multi-dimensional Global Urban Centre Database to delineate and analyse world cities. *Sci. Data* **11**, 82 (2024).
12. European Commission, Joint Research Centre. GHS urban centre database 2015, multitemporal and multidimensional attributes. European Commission, Joint Research Centre [Dataset] (2019). <https://doi.org/10.2905/53473144-B88C-44BC-B4A3-4583ED1F547E>.
13. European Commission, Joint Research Centre. GHS functional urban areas 2015. European Commission, Joint Research Centre [Dataset] (2019). <https://doi.org/10.2905/347F0337-F2DA-4592-87B3-E25975EC2C95>.
14. Cattaneo, A. *et al.* Worldwide delineation of multi-tier city–regions. Zenodo [Dataset] (2024). <https://doi.org/10.5281/zenodo.11187634>.
15. Adams, D. S. *et al.* LandScan Mosaic enables high-resolution gridded population estimates with explicit uncertainty. *Sci. Rep.* **15**, 44493 (2025).