Supplementary Materials for
BMAT: A footprint-level building facade material dataset for 73 major cities worldwide

This PDF file includes:

Supplementary Text

Figures S1 to S9

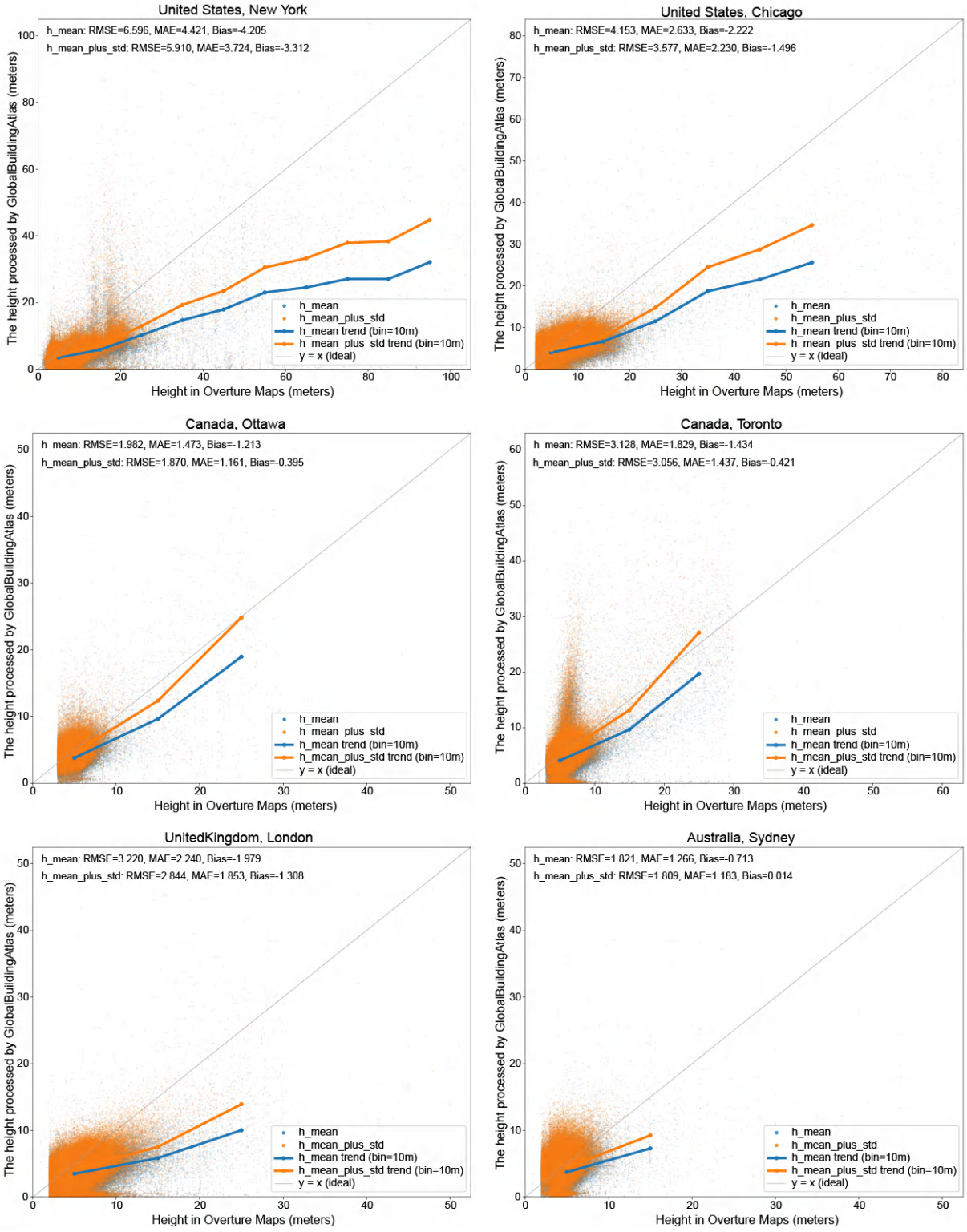
Tables S1 to S2

Contents

S1 Supplementing missing building heights for each building

Building footprints in Overture Maps lack height records for a substantial portion of buildings, particularly in developing regions. To support downstream applications such as energy modeling, disaster risk assessment, and urban carbon estimation using this dataset, we supplemented these missing values using the Global Building Atlas (GBA) (Zhu et al., 2025), a global raster dataset at 3 m $\times$ 3 m resolution that provides grid-cell-level building height estimates.

Specifically, we evaluated two aggregation strategies in cities where Overture Maps height records are relatively complete, using those records as reference values (Supplementary Figure. S1). Assigning the mean height (μ) of all GBA cells within a building footprint consistently underestimated building height, likely because boundary cells partially covering non-building surfaces such as roads record lower values. To compensate, we instead assigned $\mu + \sigma$, which yielded closer agreement with Overture Maps reference values across all tested cities. We therefore retained original height records where Overture Maps provides them, and applied the $\mu + \sigma$ estimator to supplement height information for all remaining footprints.



Supplementary Figure. S1. Comparison of two raster-to-vector height assignment methods against Overture Maps reference values across six cities. Each point represents one building. Trend lines (10 m bins) show that the mean-plus-standard-deviation estimator (orange) outperforms the mean alone (blue) across all cities, with consistently lower RMSE and reduced negative bias.

S2 Model classification results comparison

Table S1 compares the classification performance of the fine-tuned Qwen2.5 model against traditional vision backbones on the test set. The results demonstrate that the fine-tuned Qwen2.5 achieves superior performance.

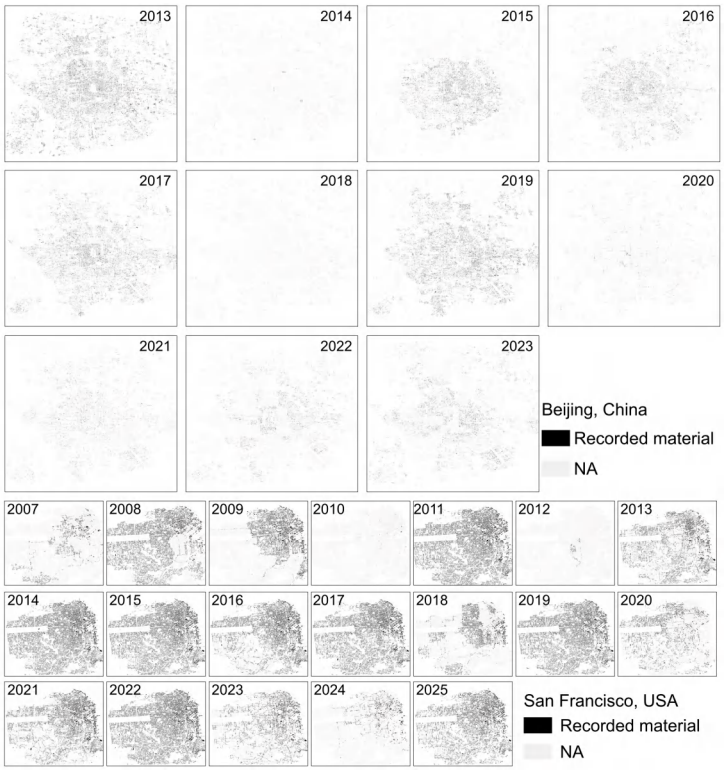
Supplementary Table S1. Class-wise performance comparison with Traditional Models

Category	Metric	VGG16	ViT	Res50	Shuff	Res101	EffB0	Dense	MbNet	Ours
Overall	Accuracy	0.83	0.82	0.82	0.79	0.82	0.84	0.84	0.82	**0.91**
	Precision	0.82	0.81	0.80	0.78	0.81	0.82	0.83	0.81	**0.91**
	Recall	0.80	0.79	0.80	0.76	0.80	0.82	0.82	0.80	**0.91**
	F1-Score	0.81	0.80	0.80	0.77	0.80	0.82	0.82	0.80	**0.91**
Brick	Precision	0.87	0.82	0.84	0.80	0.84	0.86	0.85	0.85	**0.92**
	Recall	0.84	0.82	0.84	0.79	0.82	0.83	0.85	0.83	**0.94**
	F1-Score	0.86	0.82	0.84	0.79	0.83	0.85	0.85	0.84	**0.93**
Concrete	Precision	0.77	0.76	0.78	0.71	0.75	0.77	0.81	0.77	**0.86**
	Recall	0.75	0.73	0.71	0.70	0.74	0.78	0.74	0.72	**0.85**
	F1-Score	0.76	0.75	0.75	0.70	0.75	0.77	0.78	0.74	**0.86**
Glass	Precision	0.85	0.88	0.86	0.87	0.88	0.88	0.86	0.86	**0.95**
	Recall	0.93	0.95	0.94	0.92	0.93	0.95	0.95	0.94	**0.96**
	F1-Score	0.89	0.91	0.90	0.89	0.90	0.91	0.91	0.90	**0.95**
Metal	Precision	0.78	0.77	0.74	0.73	0.79	0.78	0.79	0.76	**0.88**
	Recall	0.72	0.73	0.72	0.69	0.72	0.73	0.74	0.72	**0.88**
	F1-Score	0.75	0.75	0.73	0.71	0.76	0.76	0.77	0.74	**0.88**
Other	Precision	0.81	0.83	0.77	0.76	0.76	0.74	0.78	0.79	**0.89**
	Recall	0.67	0.66	0.72	0.60	0.66	0.74	0.69	0.66	**0.91**
	F1-Score	0.73	0.73	0.75	0.67	0.70	0.74	0.74	0.72	**0.90**
Stone	Precision	0.89	0.88	0.87	0.87	0.89	0.89	0.92	0.89	**0.95**
	Recall	0.87	0.87	0.88	0.84	0.87	0.89	0.88	0.88	**0.94**
	F1-Score	0.88	0.87	0.88	0.86	0.88	0.89	0.90	0.88	**0.95**
Stucco	Precision	0.82	0.81	0.80	0.77	0.82	0.85	0.83	0.81	**0.89**
	Recall	0.84	0.82	0.81	0.81	0.83	0.83	0.85	0.83	**0.89**
	F1-Score	0.83	0.81	0.80	0.79	0.82	0.84	0.84	0.82	**0.89**
Tile	Precision	0.73	0.73	0.68	0.66	0.69	0.76	0.74	0.73	**0.88**
	Recall	0.71	0.68	0.71	0.67	0.71	0.73	0.72	0.72	**0.83**
	F1-Score	0.72	0.70	0.70	0.66	0.70	0.74	0.73	0.72	**0.85**
Wood	Precision	0.87	0.85	0.86	0.84	0.85	0.87	0.86	0.85	**0.96**
	Recall	0.90	0.88	0.86	0.83	0.89	0.91	0.91	0.90	**0.97**
	F1-Score	0.88	0.87	0.86	0.84	0.87	0.89	0.88	0.88	**0.97**

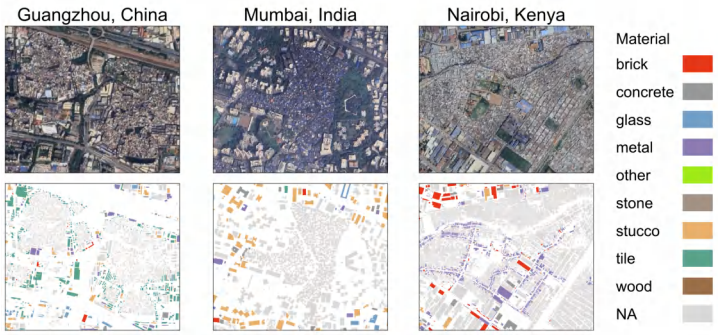
Note: Abbreviations: VGG16: VGG16 (Simonyan and Zisserman, 2014); ViT: Vision Transformer (ViT-B/16) (Dosovitskiy, 2020); Res50: ResNet-50 (He et al., 2016); Shuff: ShuffleNetV2 (Ma et al., 2018); Res101: ResNet-101 (He et al., 2016); EffB0: EfficientNet-B0 (Tan and Le, 2019); Dense: DenseNet-161 (Huang et al., 2017); MbNet: MobileNetV3-Large (Howard et al., 2019); **Ours**: Our fine-tuned Qwen2.5-7B model.

S3 Spatiotemporal coverage of buildings with material attributes

Supplementary Figure. S2 illustrates spatiotemporal sampling characteristics of street view platforms. GSV demonstrates more frequent updates and complete spatial coverage per cycle, while BSV shows longer intervals and more fragmented coverage. Both platforms exhibit systematic gaps in densely built areas with narrow streets inaccessible to collection vehicles, such as informal settlements and historic alleyways (Supplementary Figure. S3).



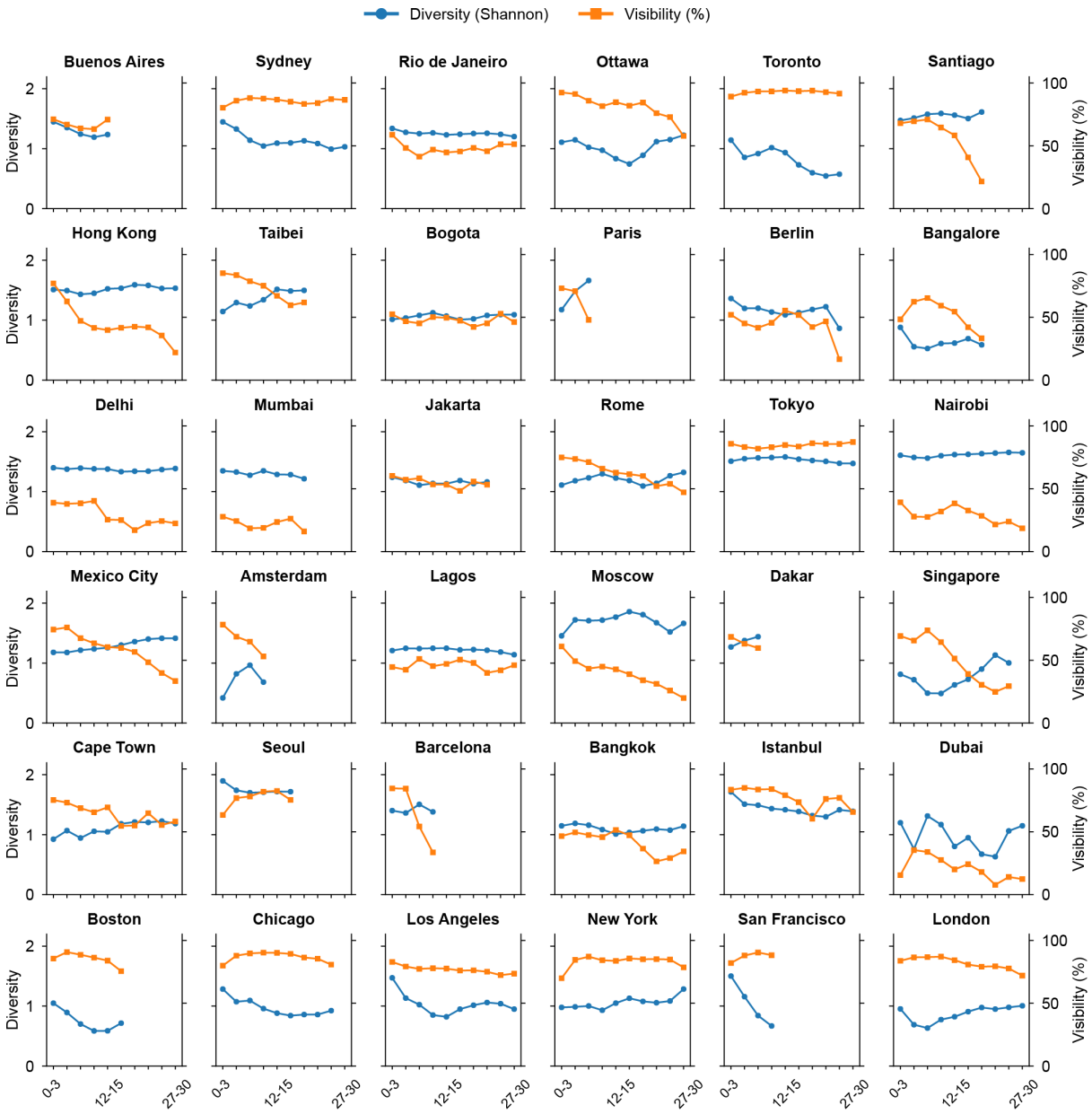
Supplementary Figure. S2. Spatiotemporal availability of building material records: Reflecting sampling differences between BSV and GSV through the case studies of Beijing and San Francisco.



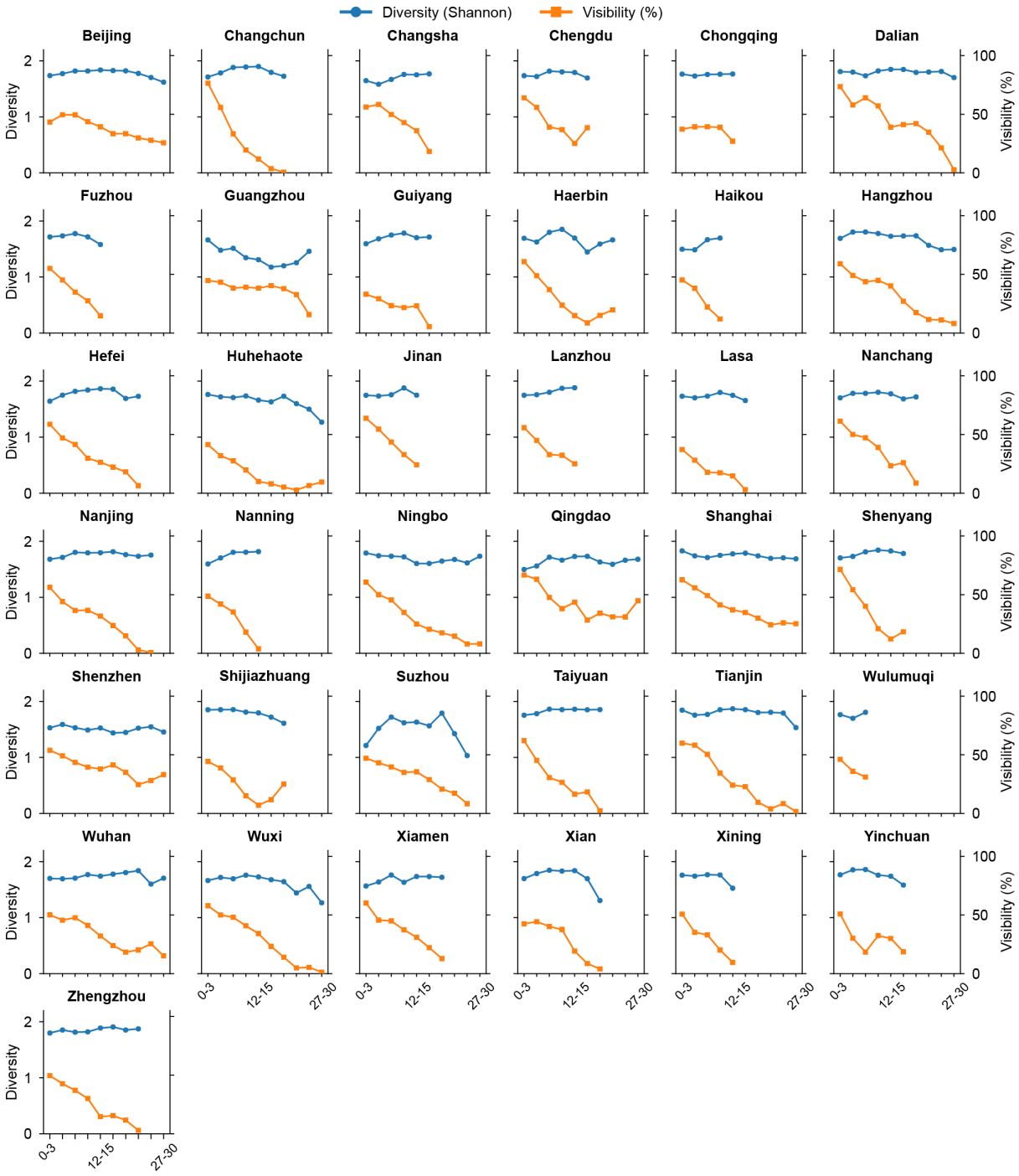
Supplementary Figure. S3. Missing building material attributes in high-density areas. Imagery © 2025 Airbus, CNES/Airbus, Maxar Technologies.

S4 Spatial heterogeneity of building facade materials across different cities

Supplementary Figures. S4-S5 present the proportion of buildings with material attributes and material diversity (entropy values) across different buffer zones from city centers. Material diversity varies distinctly among cities: some show minimal variation (e.g., Beijing, Shanghai, Delhi, Mumbai, Tokyo, Nairobi, Lagos, Dakar—predominantly Asian and African cities), some exhibit declining trends (e.g., Sydney, Ottawa, Chicago, Los Angeles, San Francisco—predominantly cities in the Americas and Oceania), while others display increasing diversity (e.g., Paris, Rome, Moscow, Amsterdam, London—predominantly European cities).



Supplementary Figure. S4. Trends in Shannon diversity and visibility rates across buffer zones in representative GSV-derived cities.



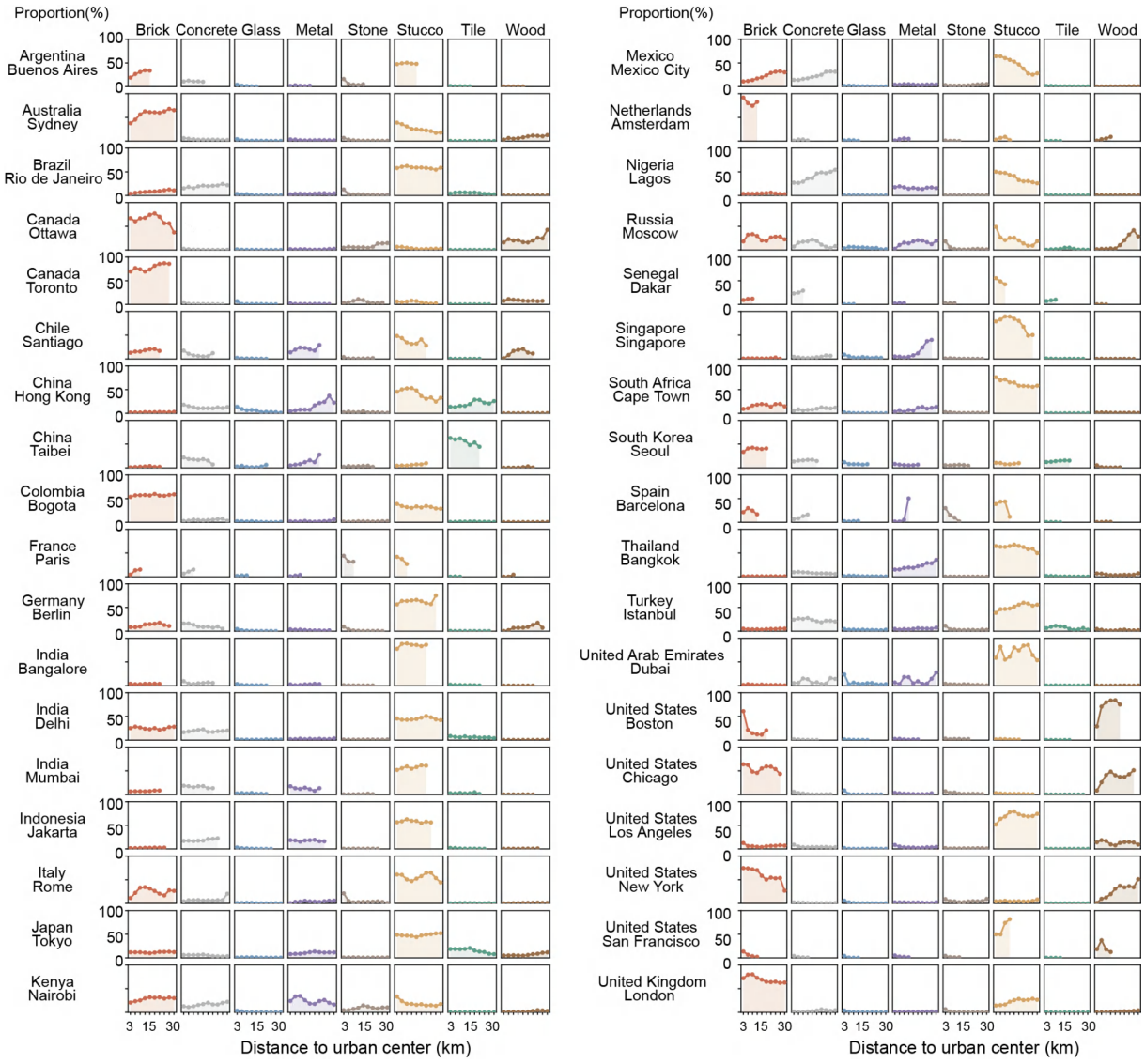
Supplementary Figure. S5. Trends in Shannon diversity and visibility rates across buffer zones in BSV-derived cities.

Supplementary Table S2. Statistics of recorded building materials in different buffer zones (China)

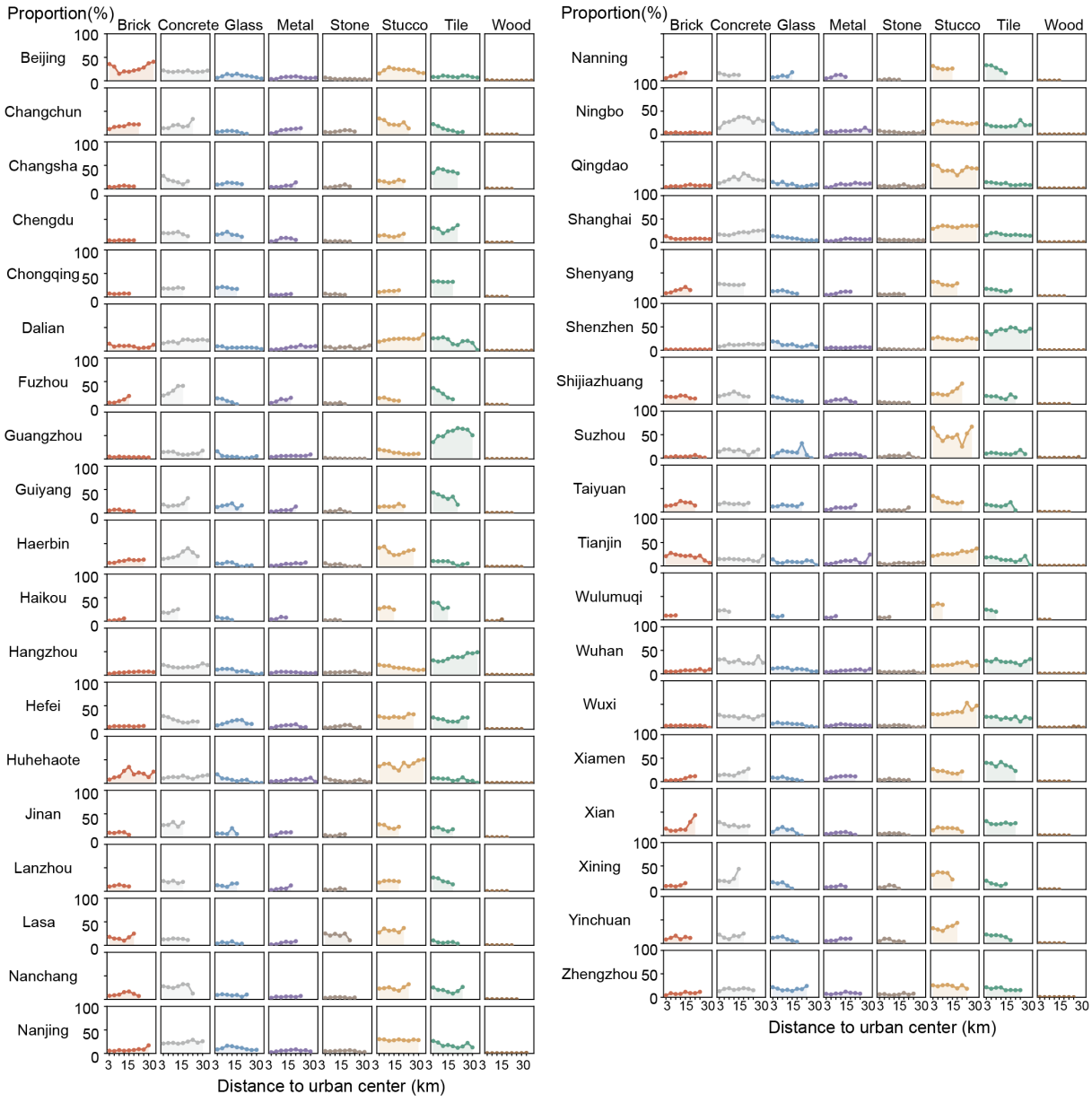
City	w/ Mat.[1]	0–5 km		0–10 km		0–15 km	
		Count	Prop(%)	Count	Prop(%)	Count	Prop(%)
Beijing	308,555	34,479	46.1	93,444	48.1	154,636	44.9
Changchun	74,336	21,266	65.8	52,621	42.0	67,000	29.9
Changsha	72,507	17,230	57.4	47,887	52.5	66,368	48.1
Chengdu	92,899	22,044	60.9	60,510	45.8	85,174	40.5
Chongqing	37,117	9,446	37.9	25,100	39.1	34,506	37.2
Dalian	83,553	12,545	69.8	33,929	64.2	50,973	55.5
Fuzhou	40,855	19,280	50.3	35,665	42.5	37,929	40.5
Guangzhou	231,084	19,430	44.0	72,270	40.6	136,540	39.5
Guiyang	26,440	10,741	31.6	18,864	28.3	24,531	26.5
Haerbin	58,893	9,023	54.5	30,963	40.2	46,291	28.3
Haikou	29,197	17,415	42.1	26,849	34.5	27,135	33.5
Hangzhou	158,114	17,416	52.8	48,602	47.4	88,384	44.8
Hefei	65,424	20,692	53.2	42,135	45.9	54,133	39.6
Huhehaote	43,741	9,432	36.3	28,359	28.7	37,348	23.1
Jinan	50,462	20,471	59.6	45,205	49.8	47,353	48.4
Lanzhou	15,565	7,673	49.8	12,532	43.2	14,446	41.1
Lasa	21,936	12,163	34.3	16,741	28.0	20,132	24.8
Nanchang	70,973	15,698	55.6	42,676	49.5	59,451	41.8
Nanjing	112,001	25,227	50.9	49,897	42.7	82,241	38.8
Nanning	60,565	21,891	44.4	51,599	38.0	55,542	31.4
Ningbo	108,548	13,063	54.3	42,113	46.8	69,052	37.6
Qingdao	74,315	16,059	65.6	33,978	55.0	46,490	50.6
Shanghai	445,399	29,255	59.0	87,593	51.3	151,928	45.0
Shenyang	77,340	22,087	63.3	56,099	44.4	69,890	32.7
Shenzhen	194,831	10,866	52.2	29,830	45.5	55,838	41.9
Shijiazhuang	62,466	22,245	41.7	50,749	33.6	55,379	27.4
Suzhou	39,509	18,400	46.5	27,127	43.1	34,077	41.5
Taiyuan	57,037	19,964	54.5	41,024	39.8	49,627	34.3
Tianjin	128,095	25,586	59.0	73,186	52.4	99,347	42.2
Wulumuqi	13,395	9,990	41.7	12,589	39.2	12,589	39.2
Wuhan	134,976	18,863	47.1	57,197	46.9	86,387	41.9
Wuxi	87,361	21,535	53.0	54,189	49.6	71,248	45.6
Xiamen	55,090	13,457	52.7	30,851	46.8	44,758	41.5
Xian	94,680	29,832	43.6	66,739	41.7	85,951	37.9
Xining	16,560	9,308	43.1	14,381	38.8	15,104	35.5
Yinchuan	20,380	11,825	41.9	16,098	33.1	18,634	32.8
Zhengzhou	91,260	17,023	46.4	49,615	39.3	70,959	31.0

[1] "w/ Mat." refers to the number of buildings with identified material information.

Supplementary Figures. S6-S7 illustrate the proportional distribution of facade materials across urban buffer zones. Chinese cities exhibit greater facade material diversity, with stucco and tile as predominant materials, while other cities display more homogeneous compositions dominated by brick or stucco. Notably, glass facade proportion consistently declines with distance from urban centers across all cities, reflecting the concentration of modern commercial architecture and glass curtain wall buildings in urban centers.



Supplementary Figure. S6. Variations in building material proportions with increasing distance from the urban center across GSV-derived cities.

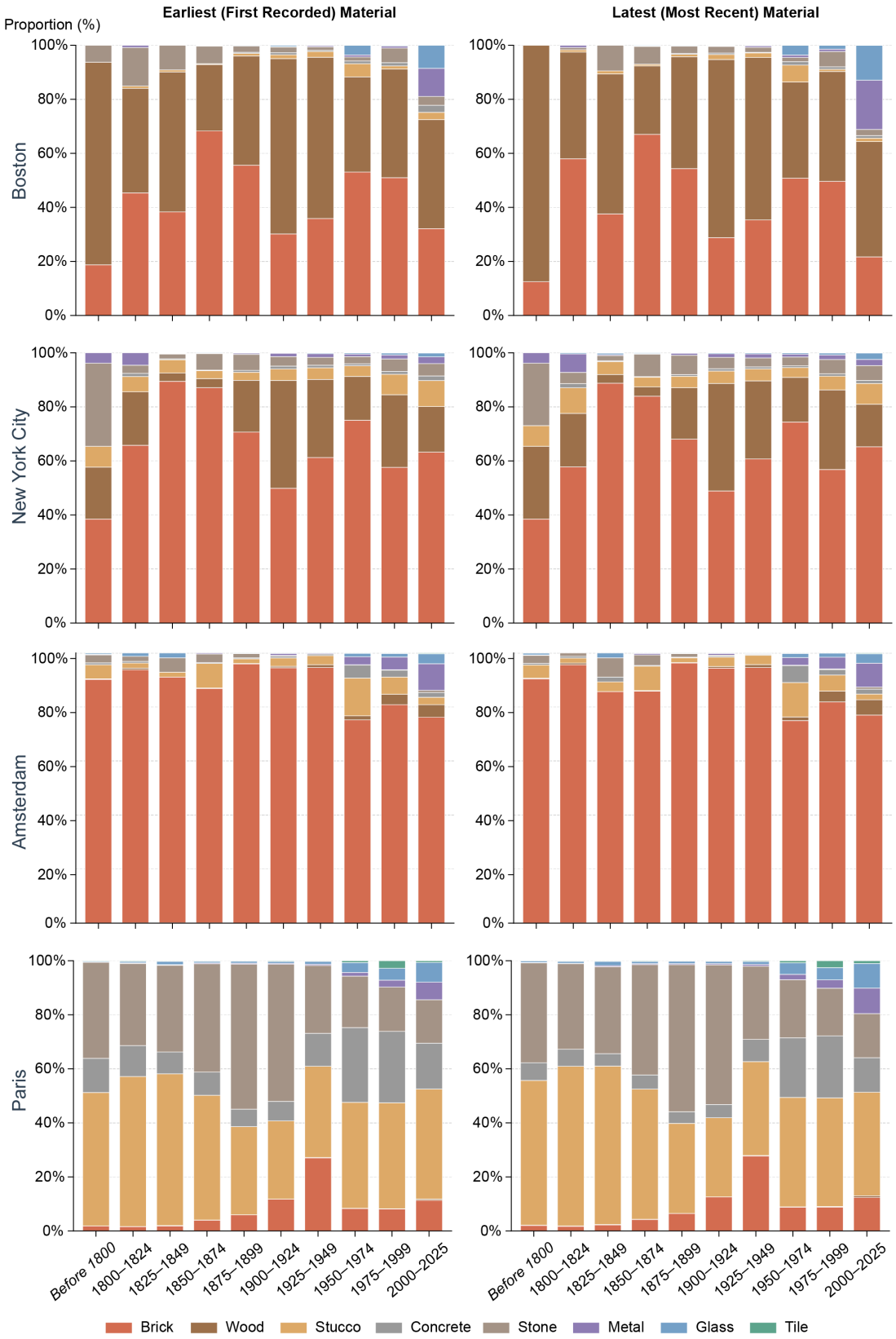


Supplementary Figure. S7. Variations in building material proportions with increasing distance from the urban center across BSV-derived cities.

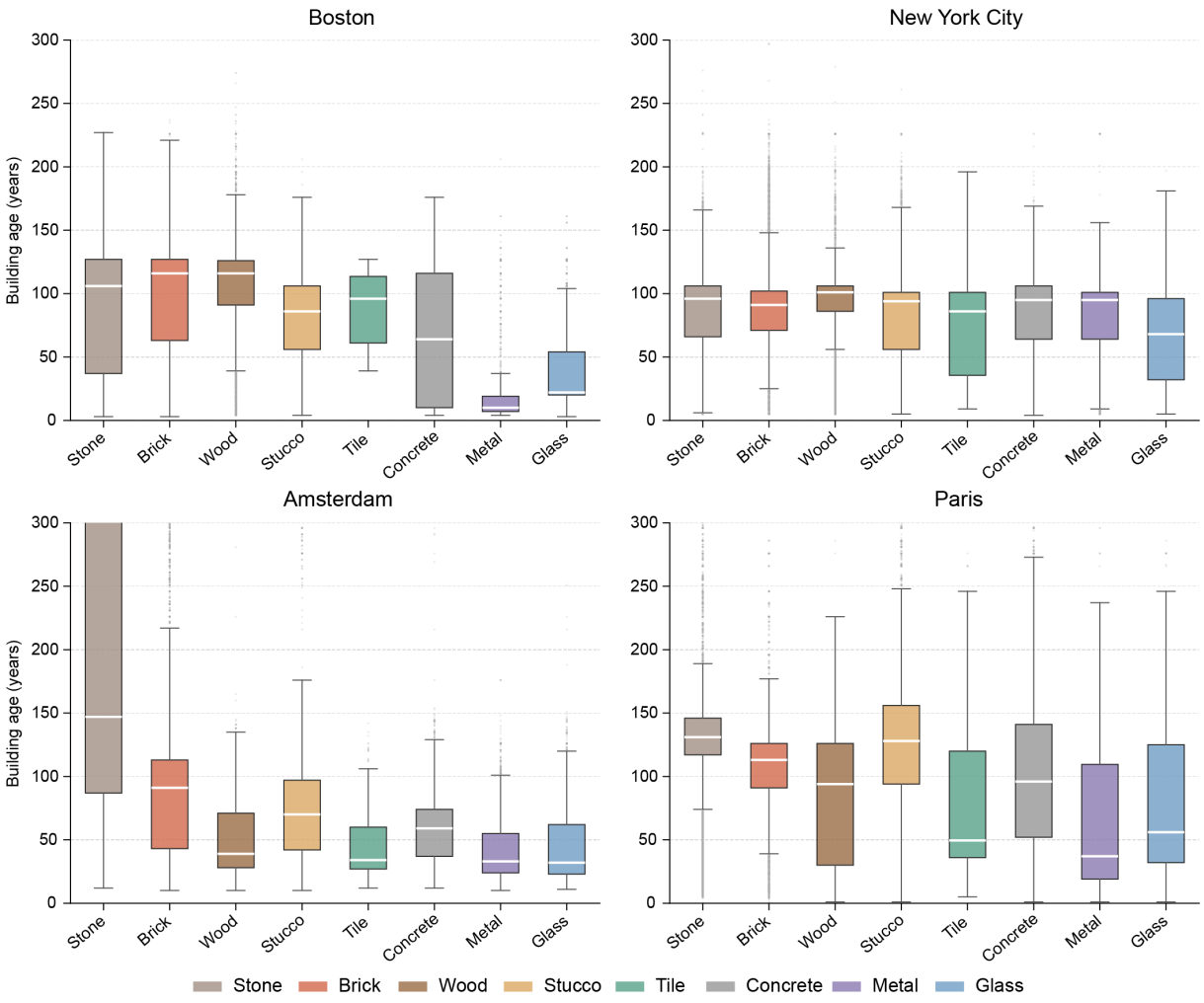
S5 Building material composition across construction eras

Supplementary Figure. S8 illustrates how building facade material composition varies with construction era across four cities. Before 1950, facades in each city were dominated by a single material — brick in Boston, New York, and Amsterdam, and stucco in Paris — with little variation across buildings. After 1950, a wider range of materials began to appear, and by 2000, glass and metal had become increasingly common, driven by the rise of modern commercial towers and industrial buildings. Notably, when comparing the earliest and latest recorded materials for buildings constructed after 2000, glass proportion is higher in the latest records, suggesting that some of these buildings have since been renovated with modern glass facade. This points to a growing trend of facade renewal in the most recent construction period.

Supplementary Figure. S9 presents the distribution of building age for each facade material type across different cities. In all cities, modern materials such as glass and metal correspond to younger buildings, with median ages typically below 50 years, reflecting their widespread adoption in recent decades alongside the rise of commercial and industrial construction. Tile and concrete buildings also tend to be relatively young, consistent with mid-to-late 20th century construction practices. In contrast, traditional materials such as brick and stone are found predominantly in older buildings, with median ages commonly exceeding 100 years, particularly in Amsterdam and Paris where historic building stock remains largely intact. These differences highlight the close relationship between facade material and construction era, as dominant building materials have shifted over time in response to changing construction technologies and urban development demands.



Supplementary Figure. S8. Building facade material composition by construction era. Left column: earliest recorded facade material per building; right column: most recent recorded material.



Supplementary Figure. S9. Building age distribution by facade material across different cities.

References

Dosovitskiy, A. (2020). An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929*.

He, K., Zhang, X., Ren, S., and Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778.

Howard, A., Sandler, M., Chu, G., Chen, L.-C., Chen, B., Tan, M., Wang, W., Zhu, Y., Pang, R., Vasudevan, V., et al. (2019). Searching for mobilenetv3. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 1314–1324.

Huang, G., Liu, Z., Van Der Maaten, L., and Weinberger, K. Q. (2017). Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4700–4708.

Ma, N., Zhang, X., Zheng, H.-T., and Sun, J. (2018). Shufflenet v2: Practical guidelines for efficient cnn architecture design. In *Proceedings of the European conference on computer vision (ECCV)*, pages 116–131.

Simonyan, K. and Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.

Tan, M. and Le, Q. (2019). Efficientnet: Rethinking model scaling for convolutional neural networks. In *International conference on machine learning*, pages 6105–6114. PMLR.

Zhu, X. X., Chen, S., Zhang, F., Shi, Y., and Wang, Y. (2025). Globalbuildingatlas: an open global and complete dataset of building polygons, heights and lod1 3d models. *Earth System Science Data Discussions*, 2025:1–31.