https://doi.org/10.5194/essd-2026-345
Preprint. Discussion started: 17 June 2026

Open Access Earth System Science Data Discussions

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

Hanyu Yin[1], Fan Zhang[1,*], Yuqing Wang[1], Lun Wu[1], and Yu Liu[1]

[1]Institute of Remote Sensing and Geographic Information System, Peking University, Beijing, 100871, China

Correspondence: Fan Zhang (fanzhanggis@pku.edu.cn)

Abstract. Building facade materials are closely associated with urban microclimates, energy consumption, and carbon emissions, yet existing urban building datasets typically capture only building footprints and heights, leaving semantic material information unavailable at the individual-building level. Here we present BMAT, a footprint-level building facade material dataset covering 22.09 million buildings across 73 major cities worldwide. BMAT links individual building footprints with facade material labels inferred from over 147 million temporally stamped street-view images collected between 2007 and 2025, where historical imagery is available. We developed an automated inference pipeline based on a fine-tuned Vision-Language Model trained on 39,405 manually annotated samples, achieving an F1-score of 0.91 against held-out test data. Dataset reliability was further assessed through independent cross-validation against 47,434 Overture Maps building records with known material attributes, yielding an overall accuracy of 0.80. BMAT reveals spatially distinct material signatures associated with climate, geography, and regional building traditions. In cities with sufficient repeated imagery, the temporally stamped records further provide exploratory evidence of facade material transitions, including increasing glass facade adoption in selected rapidly urbanizing regions. By bridging building geometry and semantic surface properties, BMAT supports studies in building energy modelling, embodied carbon accounting, urban microclimate analysis, and urban hazard risk assessment. The dataset is openly available at https://doi.org/10.6084/m9.figshare.31569370 (Yin, 2026).

1 Introduction

Rapid global urbanization is accelerating, with nearly 68% of the world's population projected to reside in urban areas by 2050 (United Nations Population Division, 2014). This unprecedented expansion intensifies environmental challenges, including urban heat islands (Manoli et al., 2019; Hsu et al., 2021), escalating building energy consumption (Güneralp et al., 2017; Wang et al., 2023), and rising carbon emissions (Yap et al., 2025; Gurney et al., 2021). Facade materials form the critical interface between buildings and their surrounding environment. Their thermal and radiative properties directly affect building envelope performance and urban microclimates (Wang et al., 2024a; Liu and Lei, 2025), with implications for building energy consumption, embodied carbon estimation (Liang et al., 2025a), and disaster risk assessment (Zamanialaei et al., 2025). Beyond environmental performance, the choice of facade materials also reflects the construction technologies (Xie et al., 2025; Huang and Hsu, 2003), aesthetic preferences (Xue et al., 2025), and urban development processes of different eras (Naik et al., 2017), making them a tangible record of how cities have changed over time. However, the lack of large-scale, multi-temporal

https://doi.org/10.5194/essd-2026-345
Preprint. Discussion started: 17 June 2026

facade material datasets with fine-grained classification leaves a critical gap for data-driven urban sustainability and resilience research.

Despite rapid progress in large-scale building mapping, no existing dataset provides systematic facade material information at the individual building level. Recent advances in remote sensing and deep learning have enabled the extraction of building footprints and heights at global scale, producing widely used datasets such as 3D-GloBFP (Che et al., 2024), GlobalBuildingAtlas (Zhu et al., 2025), Google Open Buildings (Sirko et al., 2023), and Microsoft Building Footprints (Microsoft, 2024). While these resources have greatly advanced urban morphology research, they are limited to geometric attributes—an inherent constraint of the nadir satellite perspective, which captures rooftops but cannot observe vertical facades. Open geospatial platforms such as OpenStreetMap and Overture Maps do include building material tags, but coverage is extremely sparse, with the vast majority of buildings worldwide lacking any material annotation (Herfort et al., 2023; Biljecki et al., 2023). Moreover, these records are typically static snapshots lacking temporal metadata, making them unable to reflect material changes from renovations or redevelopment (Zhang et al., 2025). As a result, no existing dataset provides ready-to-use, footprint-level facade material attributes across a broad set of major cities, while also preserving temporal references for multi-year analysis.

Street-view imagery has emerged as a promising alternative for facade material identification. Unlike nadir satellite observations, street-view images capture the urban environment from a horizontal perspective, clearly revealing facade materials, textures, and architectural details (Wang et al., 2024b; Chen et al., 2025). Importantly, major platforms such as Google Street View (GSV) and Baidu Street View (BSV) now provide dense, repeated coverage across cities worldwide, offering both broad spatial extent and a multi-temporal archive (Biljecki and Ito, 2021). Recent studies have demonstrated the feasibility of combining street-view images with deep learning for building material classification, progressing from end-to-end CNN and Transformer classifiers applied to whole images (Raghu et al., 2023; Sun et al., 2025), to patch-based strategies that better handle mixed-material facades (Chen et al., 2025), and more recently to fine-tuned Vision-Language Models that achieve superior accuracy (Liang et al., 2025b). However, existing efforts have focused primarily on methodological development, releasing annotated image samples (Raghu et al., 2023; Wang et al., 2024b) or inference toolkits (Liang et al., 2025b) rather than producing a ready-to-use, building-level facade material dataset across major cities.

In summary, existing building datasets derived from remote sensing capture only geometric attributes, while street-view-based material studies remain at the methodological stage. No publicly available dataset yet links facade material attributes to individual building footprints across a broad set of major cities. To address these gaps, this study presents BMAT, a building-level facade material dataset derived from over 147 million multi-temporal street-view images covering 73 major cities worldwide—37 in mainland China (via Baidu Street View, 2013–2023) and 36 in other countries and regions (via Google Street View, 2007–2025). These cities encompass national capitals, provincial centers, and economically significant metropolises with large populations, providing broad geographic representativeness across diverse urban contexts worldwide. By linking material attributes to individual building footprints and retaining the acquisition year of each street-view observation, BMAT moves urban material analysis beyond assumption-based or material-blind approaches. It supports applications in building energy simulation, embodied carbon accounting, urban microclimate analysis, and urban hazard risk assessment, while also providing a basis for exploratory longitudinal analysis in cities and periods with sufficient records. The main contributions are:

https://doi.org/10.5194/essd-2026-345
Preprint. Discussion started: 17 June 2026


Open Access Earth System Science Data Discussions

(1) Broad multi-city coverage: the dataset spans 73 major cities across diverse climate zones and development contexts, providing broad geographic coverage for building-level facade material mapping and enabling cross-regional comparative analysis;

(2) Building-level granularity: material attributes are mapped to 22.09 million individual building footprints, providing the spatial precision required for downstream urban analytics that aggregate-level data cannot support;

(3) Temporally stamped records: the dataset preserves acquisition-year information from available street-view imagery between 2007 and 2025, supporting exploratory analysis of facade material transitions in cities and periods with sufficient repeated observations.

2 Data Source

2.1 City administrative boundaries

This study uses the GADM database v4.1 to obtain administrative boundaries of major cities worldwide. GADM provides multi-level administrative division vector data for countries and regions worldwide (https://gadm.org/data.html). It should be noted that for mainland Chinese cities, the official administrative boundaries are typically much larger than the actual built-up areas. Therefore, this study focuses on the main urban areas where street view data coverage is available, rather than the entire administrative region, to ensure targeted and efficient data collection.

2.2 Building footprint and height data

Building footprint data were acquired from the Overture Maps Foundation global dataset. This repository integrates multiple authoritative sources—including OpenStreetMap, Esri Community Maps, Microsoft Buildings, and Google Open Buildings, among others. It comprises over 2.6 billion individual building vector features worldwide, referenced to the WGS84 coordinate system and attributed with geometric boundaries, heights, and classifications. The dataset's multi-source fusion strategy significantly improves coverage in developing regions and emerging cities compared to single-source alternatives. For buildings lacking height records in Overture Maps, missing values were supplemented using the GlobalBuildingAtlas (Zhu et al., 2025) dataset (Supplementary Figure S1).

2.3 Road network data

Road network data were retrieved from OpenStreetMap (OSM) using the OSMnx Python library. Using the city administrative boundaries defined previously as spatial filters, we queried the network with the type parameter set to "drive service". This configuration captures not only public drivable streets but also service roads (e.g., alleyways, access aisles), ensuring comprehensive coverage of potential street view sampling locations.

https://doi.org/10.5194/essd-2026-345
Preprint. Discussion started: 17 June 2026

Open Access Earth System Science Data Discussions

2.4 Street-view images

Street-view images were collected from two platforms. Google Street View (GSV), launched in 2007, provides regularly updated panoramic imagery with historical archives dating back over a decade in many cities; GSV images were collected for 36 cities outside mainland China, covering available observations between 2007 and 2025. As GSV is not available in mainland China, we used Baidu Street View (BSV), which has been operational since 2013; BSV images were collected for 37 cities, covering the period 2013–2023. The 73 cities were selected based on economic significance and population size, primarily encompassing national capitals, provincial centers, and major metropolises across diverse geographic regions and climatic zones worldwide. In total, approximately 133 million GSV images and 14 million BSV images were acquired, yielding around 147 million street-view images across all available timestamps. After filtering out images where buildings are invisible, about 92 million images (63%) remain as valid samples for facade material inference.

3 Methods

The overall workflow comprises three main stages (Figure 1). In the first stage, street-view sampling points are determined based on the spatial relationship between building footprints and the road network, and multi-temporal building facade images are retrieved with calculated camera parameters. In the second stage, a lightweight classifier first filters out outlier images; the remaining valid images are then passed to a fine-tuned Vision-Language Model for material type inference. In the third stage, the trained model is applied to all valid building images across 73 cities to produce the final multi-temporal facade material dataset.

3.1 Data collection

3.1.1 Time-series building images retrieval

This study aims to acquire time-series street-view imagery at the individual building scale. We developed a geometric workflow linking building footprints to the road network. For each building, we computed the spatial relationship between its centroid and the nearest road segment to identify the optimal street-view sampling location, and calculated the corresponding viewing parameters (i.e., heading and pitch) to orient the camera towards the target facade (Figure 2(a)). This parameterized framework enabled automated, large-scale acquisition of building-centric images across entire cities.

Specifically, theoretical sampling points are determined by projecting a perpendicular line from the building's geometric centroid to the nearest road. Since these theoretical locations may not strictly coincide with the platform's recorded collection points, the available street view point nearest to the theoretical location is identified as the actual sampling point. Subsequently, camera parameters are computed based on the spatial relationship between this actual point and the building center: the heading parameter is defined as the clockwise angle from True North to the line connecting the actual sampling point to the building centroid, while the pitch parameter represents the angle between the ground and the line of sight directed at the building's

https://doi.org/10.5194/essd-2026-345
Preprint. Discussion started: 17 June 2026

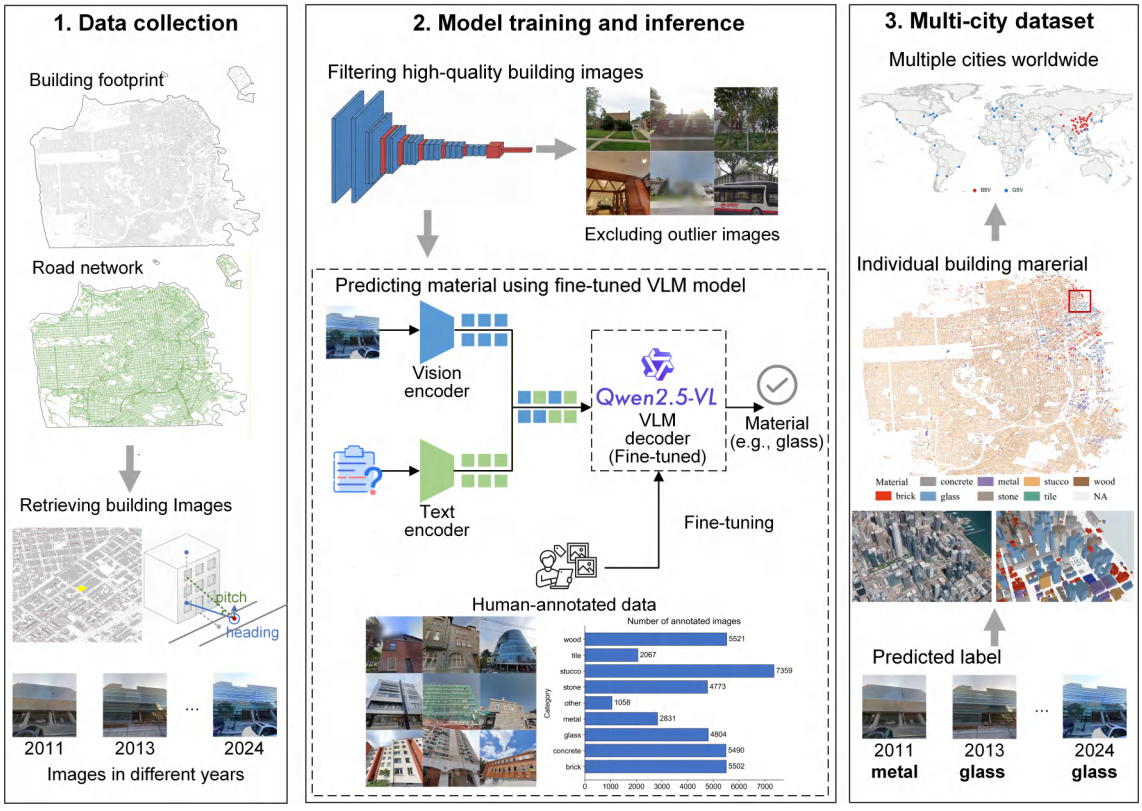


Figure 1. Overview of the dataset production workflow: (1) multi-temporal street-view image collection, (2) model training and material inference, and (3) multi-city dataset generation. Aerial images are from Google Earth Pro. Imagery © 2025 Landsat / Copernicus, Map data © 2025 Google.

vertical midpoint (half of the building height). For buildings lacking raw height information in Overture Maps, the pitch parameter is set to a default value of 30°.

3.1.2 Line-of-sight obstruction detection

In densely built-up areas, intervening structures may obstruct the view between street-view sampling points and target buildings. To systematically identify building visibility, we implemented a two-stage line-of-sight verification strategy (Figure 2(b)). First, a sight line is cast from the sampling point to the centroid of the target building footprint and tested for intersection with surrounding building polygons. If the centroid sight line is obstructed, a multi-point verification is triggered, in which sight lines are cast to all boundary vertices of the target footprint and each is tested independently. Buildings are classified as fully visible if the centroid sight line is unobstructed, or partially visible if at least one vertex sight line remains clear. Completely occluded buildings, where all sight lines are blocked, are excluded from further processing. Both fully and partially visible buildings are retained for material inference, as even partial facade exposure provides sufficient visual information for classi-

https://doi.org/10.5194/essd-2026-345
Preprint. Discussion started: 17 June 2026

Open Access Earth System Science Data Discussions

fication. This approach maximizes valid sample retention while ensuring that the target building is genuinely captured in the imagery.

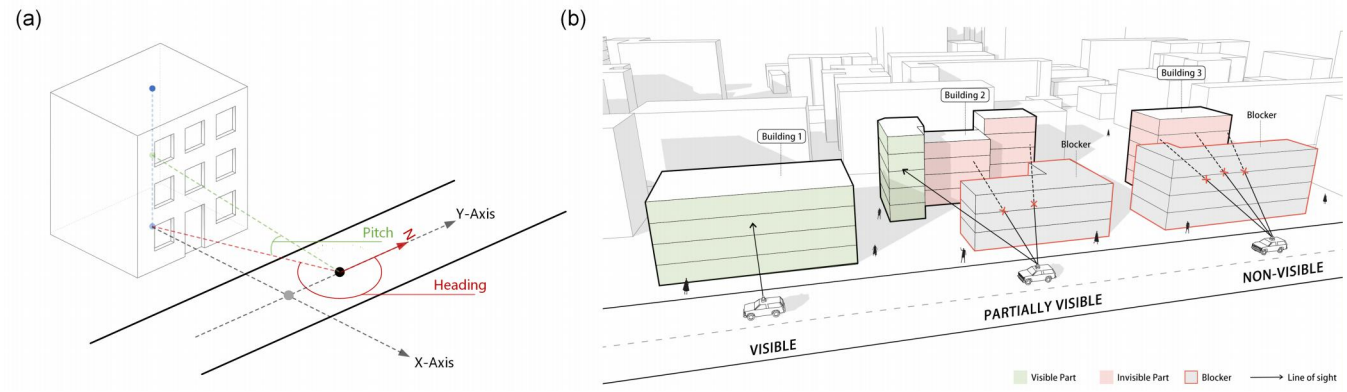


Figure 2. Workflow for facade image acquisition. (a) Sampling location and camera parameter determination; (b) Line-of-sight obstruction detection.

3.2 Model training and inference

3.2.1 Filtering high-quality building images

Street-view images frequently contain outlier samples unsuitable for material classification, such as overexposed scenes, vehicle-occluded views, indoor captures, or heavily mosaicked areas. To exclude these, we trained a binary classifier based on the MobileNetV2 architecture. A labeled dataset of 8,934 images (5,903 valid and 3,031 invalid) was constructed and split into training (80%) and validation (20%) sets. The model was fine-tuned from pre-trained ImageNet weights using the Adam optimizer (learning rate = 1e-4) for 100 epochs, achieving a validation accuracy of 88.75%. This classifier was then applied to the full image collection to retain only valid building facade images for subsequent inference.

3.2.2 Material inference using a fine-tuned Vision-Language Model

To support model training, we constructed the largest human-annotated dataset of building facade materials to date, comprising 39,405 images across nine material categories (Figure 3): brick, concrete, glass, metal, stone, stucco, tile, wood and other (primarily buildings under construction). The dataset was split into training and testing sets at a 7:3 ratio.

We adopted Qwen2.5-VL-7B-Instruct as the backbone model and implemented Supervised Fine-Tuning (SFT) using the LLaMA-Factory framework. To optimize computational efficiency, we applied LoRA (Low-Rank Adaptation) with 4-bit QLoRA quantization in bfloat16 precision. Training was conducted for 3 epochs with an initial learning rate of 5e-5. The prompt employed during fine-tuning was: "What is the facade material of the central building in this image? Please choose one of the following: [tile, brick, concrete, glass, metal, other, stone, stucco, wood] Only answer with the exact word from the list."

https://doi.org/10.5194/essd-2026-345
Preprint. Discussion started: 17 June 2026

Figure 3. Example images of the 9 human-annotated material types.

3.3 Developing multi-city building facade material dataset

The fine-tuned Vision-Language Model was applied to infer facade materials for all buildings with valid street-view images across 73 cities. For buildings captured at multiple timestamps, materials were inferred independently for each available year, enabling tracking of facade material transitions over time. By applying this approach across major cities worldwide, we finally produced a multi-temporal, individual-level building material dataset covering 73 cities (36 cities based on GSV and 37 cities based on BSV).

https://doi.org/10.5194/essd-2026-345
Preprint. Discussion started: 17 June 2026

Open Access Earth System Science Data Discussions

4 Results

4.1 Model performance

We evaluated several Vision-Language Models on the manually labeled test set (11,827 images) to select the most suitable backbone (Table 1). Among all models tested without fine-tuning, Qwen2.5-VL-7B-Instruct achieved the highest inference accuracy while also requiring the shortest inference time. We therefore selected it as the backbone and fine-tuned it on the training set (27,578 images). The fine-tuned model achieved an F1-score of 0.91, substantially outperforming both the original unfine-tuned VLMs and traditional vision backbones (e.g., ResNet, ViT) trained on the same data (Figure 4, Supplementary Table S1).

Table 1. Comprehensive performance comparison of different models.

Model	Fine-tuned	Infer Time (ms)	Accuracy	Precision	Recall	F1-Score
Traditional Vision Backbones						
VGG16	✓	1.26	0.83	0.82	0.80	0.81
ViT-B/16	✓	4.40	0.82	0.81	0.79	0.80
ResNet-50	✓	5.57	0.82	0.80	0.80	0.80
ResNet-101	✓	10.52	0.82	0.81	0.80	0.80
ShuffleNetV2	✓	5.57	0.79	0.78	0.76	0.77
EfficientNet-B0	✓	6.81	0.84	0.82	0.82	0.82
DenseNet-161	✓	18.37	0.84	0.83	0.82	0.82
MobileNetV3-Large	✓	5.54	0.82	0.81	0.80	0.80
Vision Language Models (VLMs)						
InternVL2.5-8B	×	185.55	0.69	0.62	0.53	0.54
MiniCPM2.6-8B	×	631.59	0.74	0.55	0.48	0.48
Qwen2-7B	×	276.06	0.70	0.51	0.52	0.51
mLLaMA3.2-11B	×	829.64	0.66	0.38	0.36	0.36
Qwen2.5-7B	×	221.48	0.73	0.76	0.65	0.66
Fine-tuned Qwen2.5-7B (Ours)	✓	**237.73**	**0.91**	**0.91**	**0.91**	**0.91**

To further validate model performance in real-world settings, we compared our inference results against Overture Maps building records that contain "facade material" attributes (Figure 5). Matching the most recent inferred materials with Overture Maps records yielded an accuracy of 0.72 (Figure 5(a)). However, Overture Maps records lack timestamps, meaning that a building's recorded material may correspond to an earlier state rather than its current facade. A direct comparison using only the most recent inference therefore likely underestimates true agreement. To account for this, we adopted a temporal matching strategy: a prediction was considered correct if the inferred material from any available year matched the Overture Maps record, yielding an improved accuracy of 0.80 (Figure 5(b)).

https://doi.org/10.5194/essd-2026-345
Preprint. Discussion started: 17 June 2026

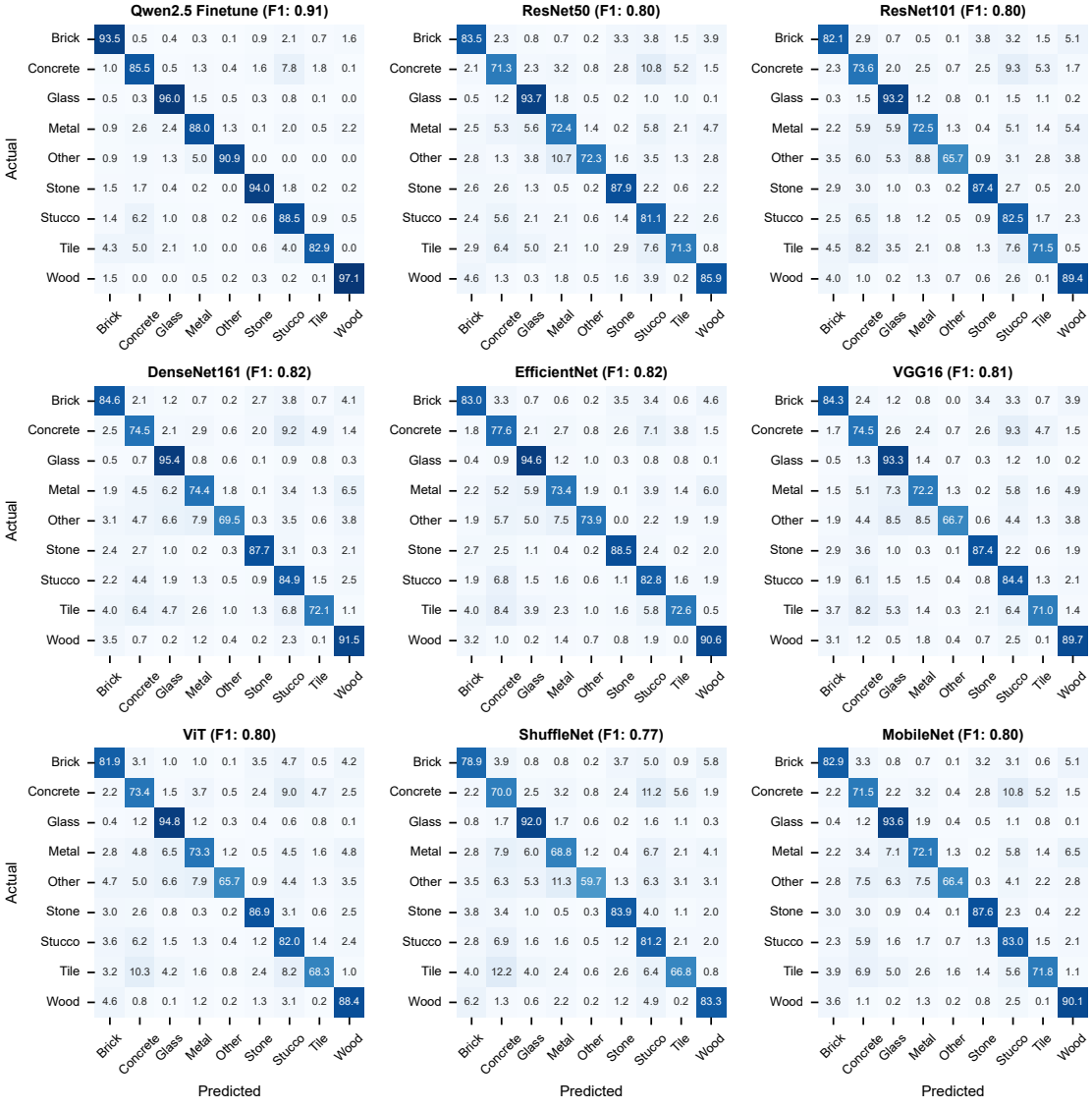


Figure 4. Comparison of classification performance between the fine-tuned Qwen2.5 model and traditional vision backbones on the test set.

It should be noted that both our model predictions and the Overture Maps records may contain errors. Figure 5(c) illustrates three representative scenarios: agreement between both sources, only our model correct, and only Overture Maps correct. Our model performs reliably in the majority of cases, with misclassifications concentrated in visually ambiguous scenarios, such as brick with indistinct textures being classified as concrete, stucco facades dominated by large glass windows being identified as glass, and wood and metal surfaces with similar visual appearance.

https://doi.org/10.5194/essd-2026-345
Preprint. Discussion started: 17 June 2026

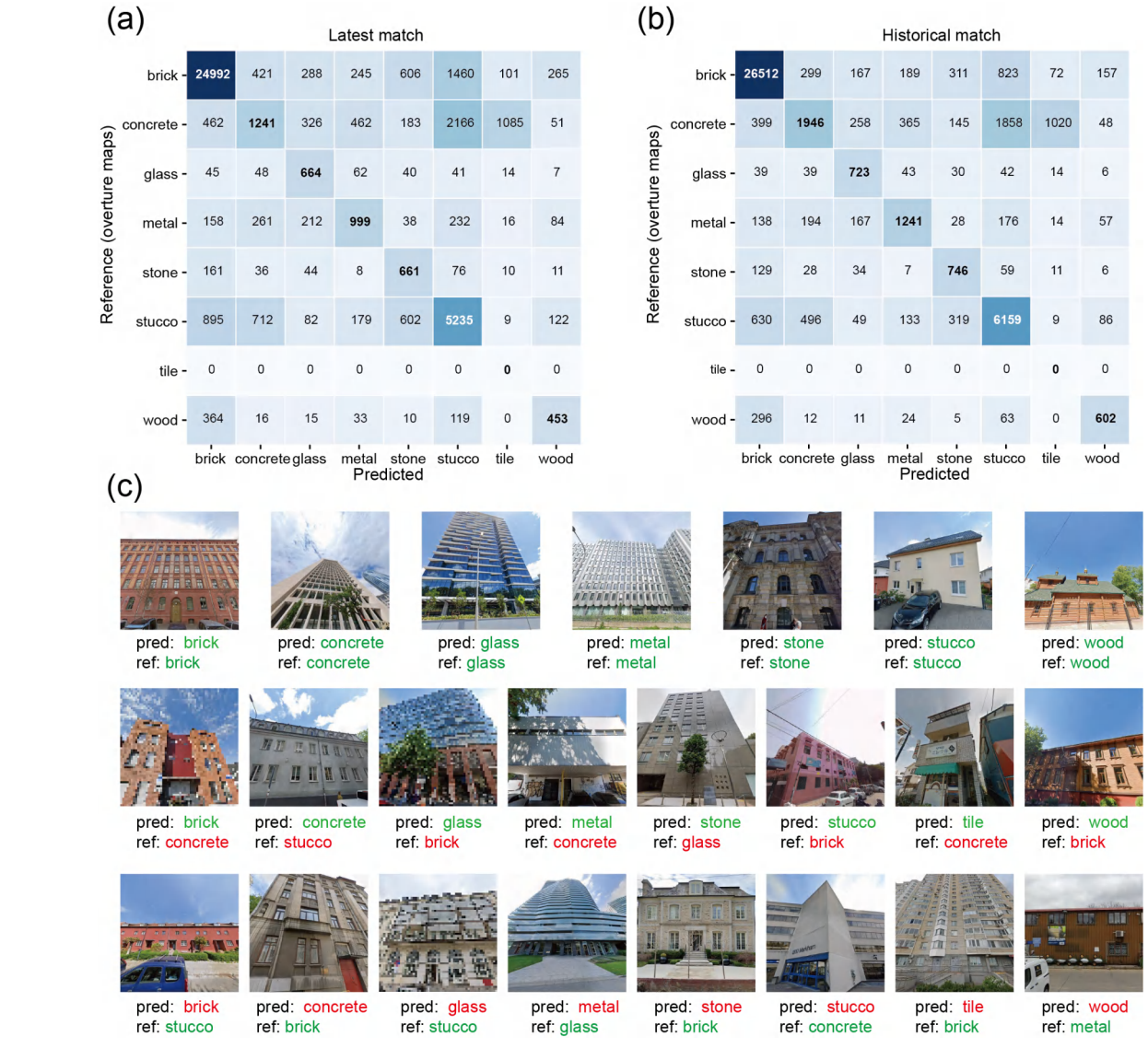


Figure 5. Comparison of model predictions with Overture Maps records for (a) latest and (b) historical material attributes, including (c) visual examples with green for correct and red for error.

4.2 Details of this dataset

Figure 6(a) presents the geographic distribution of the 73 cities covered in this dataset, which were selected based on economic significance and population size, primarily encompassing national capitals, provincial centers, and major metropolises across diverse geographic regions and climatic zones. The dataset classifies building facade materials into nine categories: brick, concrete, glass, metal, stone, stucco, tile, wood, and other. Material attributes are assigned to all buildings for which

https://doi.org/10.5194/essd-2026-345
Preprint. Discussion started: 17 June 2026

Open Access Earth System Science Data Discussions

valid street-view images are available; buildings without street-view coverage do not have material records. In total, the dataset encompasses 73 cities with 40.86 million buildings, of which 22.09 million buildings (54%) have material attributes assigned. Figure 6(b) illustrates the dataset at both city and local scales, showing the spatial distribution of facade material classifications across entire cities alongside zoomed-in views of representative neighborhoods in New York and San Francisco. A qualitative comparison with contemporaneous Google Earth Pro aerial imagery indicates that the dataset effectively distinguishes material types among adjacent buildings—for example, correctly differentiating glass-clad high-rises from brick-faced low-rise buildings in dense urban cores—demonstrating both the spatial resolution and high accuracy of the material classifications.

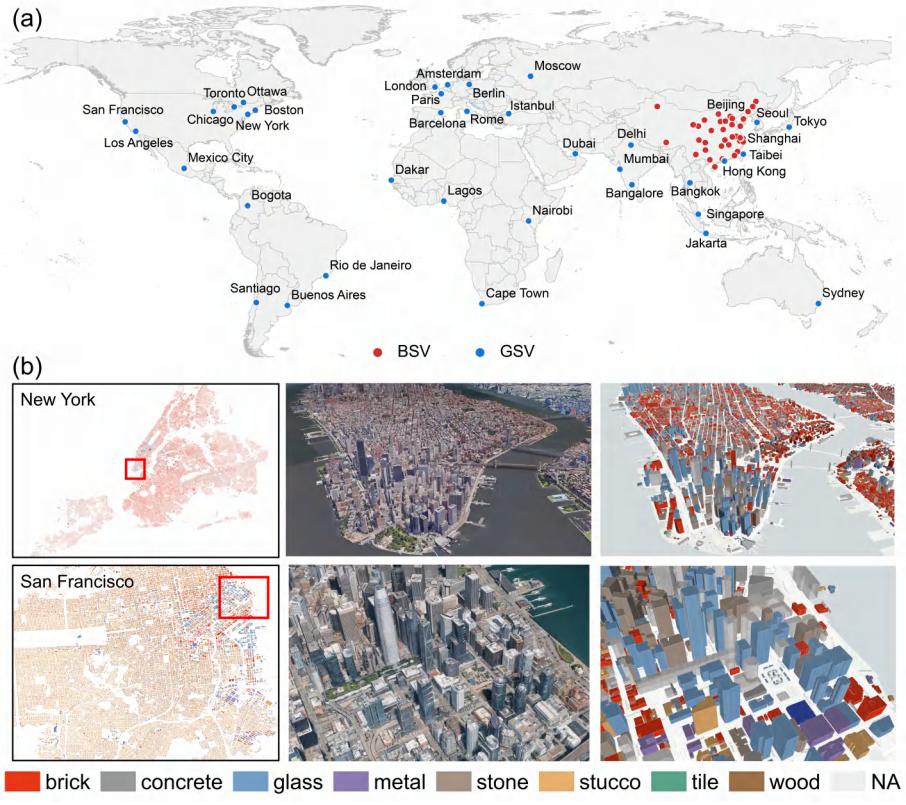


Figure 6. Building material dataset overview: (a) geographic distribution of 73 cities; (b) detailed comparison of Google Earth imagery and recorded material classifications for local regions. Aerial images are from Google Earth Pro. Imagery © 2025 Landsat / Copernicus, Map data © 2025 Google.

Figure 7 illustrates the distribution of building facade materials across cities, revealing distinct spatial patterns associated with climate and geography. Coastal cities (e.g., Singapore, Los Angeles, San Francisco, Tokyo, Mumbai) show a predominance of stucco facades, consistent with its suitability for humid maritime environments due to moisture resistance. Cities in mid-to-high latitude cold climate regions (e.g., Ottawa, Toronto, Amsterdam, Chicago, London) are characterised by a high proportion of brick, a material valued for its thermal mass and heat retention. Distinct regional patterns also emerge. Cities in

https://doi.org/10.5194/essd-2026-345
Preprint. Discussion started: 17 June 2026

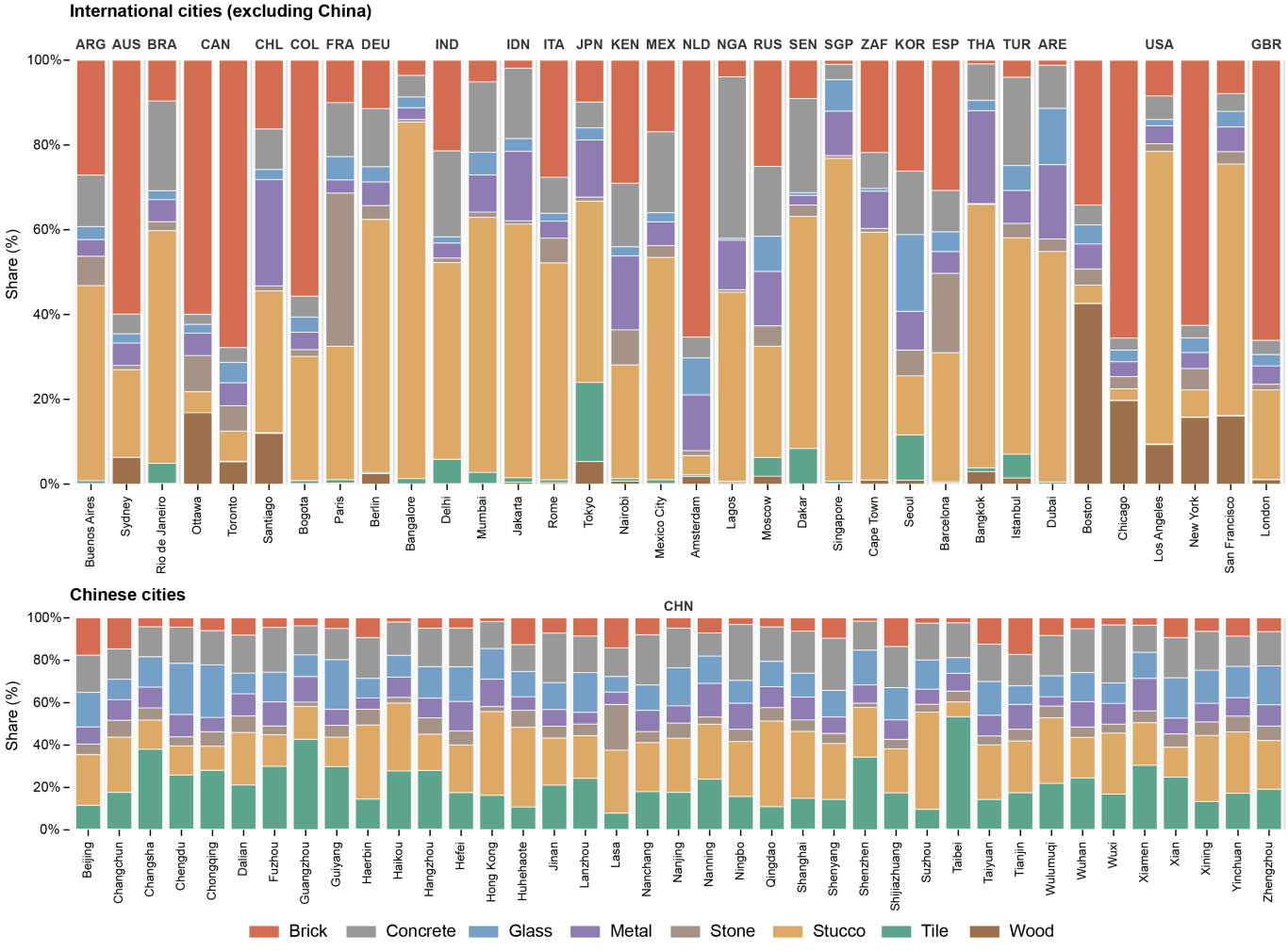


Figure 7. Distribution of building facade materials across the 73 studied cities.

Table 2 summarises the number of buildings with material information, the total building count, and the material coverage rate for each city. Coverage varies substantially across cities. Economically developed cities with relatively low building density—such as Sydney, Toronto, London, Boston, Chicago, New York, and San Francisco—exhibit notably higher coverage rates. In contrast, cities with high building density—such as Mumbai, Delhi, Nairobi, and many cities in mainland China—show considerably lower rates. This disparity is primarily attributable to densely packed residential areas, including informal settlements and high-density housing districts, where narrow streets and alleyways are inaccessible to street-view collection vehicles, resulting in gaps in image coverage (Supplementary Figure S3).

https://doi.org/10.5194/essd-2026-345
Preprint. Discussion started: 17 June 2026

Open Access Earth System Science Data Discussions

Table 2. Statistics of buildings and material information coverage across different cities in the dataset.

Country	City	w/ Mat.[1]	Total	%	Country	City	w/ Mat.[1]	Total	%
Argentina	Buenos Aires	325,364	494,210	65.84%	China	Beijing	308,555	809,109	38.14%
Australia	Sydney	997,068	1,197,990	83.23%	China	Changchun	74,336	432,388	17.19%
Brazil	Rio de Janeiro	774,478	1,467,922	52.76%	China	Changsha	72,507	150,244	48.26%
Canada	Ottawa	285,941	349,865	81.73%	China	Chengdu	92,899	216,643	42.88%
Canada	Toronto	391,850	419,217	93.47%	China	Chongqing	37,117	95,572	38.84%
Chile	Santiago	609,372	948,800	64.23%	China	Dalian	83,553	187,932	44.46%
Colombia	Bogota	1,054,336	2,226,860	47.35%	China	Fuzhou	40,855	95,964	42.57%
France	Paris	76,060	105,390	72.17%	China	Guangzhou	231,084	596,421	38.75%
Germany	Berlin	279,872	577,968	48.42%	China	Guiyang	26,440	96,587	27.37%
India	Bangalore	634,948	1,071,026	59.28%	China	Haerbin	58,893	246,935	23.85%
India	Delhi	362,766	1,212,355	29.92%	China	Haikou	29,197	82,703	35.30%
India	Mumbai	93,962	404,062	23.25%	China	Hangzhou	158,114	721,026	21.93%
Indonesia	Jakarta	1,056,179	1,912,973	55.21%	China	Hefei	65,424	183,558	35.64%
Italy	Rome	160,016	239,098	66.92%	China	Huhehaote	43,741	221,172	19.78%
Japan	Tokyo	2,196,946	2,581,836	85.09%	China	Jinan	50,462	100,604	50.16%
Kenya	Nairobi	228,988	724,700	31.60%	China	Lanzhou	15,565	36,181	43.02%
Mexico	Mexico City	948,183	1,593,166	59.52%	China	Lasa	21,936	96,256	22.79%
Netherlands	Amsterdam	141,070	200,275	70.44%	China	Nanchang	70,973	180,481	39.32%
Nigeria	Lagos	744,357	1,592,688	46.74%	China	Nanjing	112,001	581,126	19.27%
Russia	Moscow	110,222	260,956	42.24%	China	Nanning	60,565	185,928	32.57%
Senegal	Dakar	90,788	139,256	65.20%	China	Ningbo	108,548	503,086	21.58%
Singapore	Singapore	135,397	238,305	56.82%	China	Qingdao	74,315	167,654	44.33%
South Africa	Cape Town	865,241	1,460,930	59.23%	China	Shanghai	445,399	1,504,854	29.60%
South Korea	Seoul	279,308	352,282	79.29%	China	Shenyang	77,340	233,387	33.14%
Spain	Barcelona	61,196	74,984	81.61%	China	Shenzhen	194,831	540,962	36.02%
Thailand	Bangkok	768,713	1,768,080	43.48%	China	Shijiazhuang	62,466	225,788	27.67%
Turkey	Istanbul	784,412	1,151,729	68.11%	China	Suzhou	39,509	101,280	39.01%
UAE	Dubai	76,992	372,717	20.66%	China	Taiyuan	57,037	174,131	32.76%
UK	London	1,258,792	1,501,352	83.84%	China	Tianjin	128,095	444,915	28.79%
US	Boston	93,227	105,841	88.08%	China	Wulumuqi	13,395	32,969	40.63%
US	Chicago	741,837	832,709	89.09%	China	Wuhan	134,976	407,544	33.12%
US	Los Angeles	848,444	1,100,371	77.11%	China	Wuxi	87,361	261,714	33.38%
US	New York	930,542	1,091,533	85.25%	China	Xiamen	55,090	149,114	36.94%
US	San Francisco	146,030	163,766	89.17%	China	Xian	94,680	254,338	37.23%
China	Hong Kong	97,512	197,589	49.35%	China	Xining	16,560	43,713	37.88%
China	Taibei	90,269	110,653	81.58%	China	Yinchuan	20,380	59,655	34.16%
					China	Zhengzhou	91,260	324,292	28.14%

[1] "w/ Mat." refers to the number of buildings with identified material information. Region abbreviations: UAE = United Arab Emirates; UK = United Kingdom; US = United States.

4.3 Spatial patterns of building facade materials across cities

Figure 8 illustrates the spatial coverage of building material data inferred from GSV images. Figure 8(a) shows building counts and material coverage proportions at 1 km grid resolution across representative cities. In these developed cities, central areas generally exhibit higher proportions of buildings with material information than suburban zones, with most regions maintaining consistently high material data visibility. Figure 8(b) presents the proportion of buildings with material information (by count and by area) at the city level in this dataset. Most cities achieve count-based proportions exceeding 60%, while cities characterized by abundant high-density building areas (e.g., Delhi, Dubai, Hong Kong, Mumbai, and Nairobi) show lower count-based proportions. These lower proportions primarily reflect limited street view accessibility in dense neighborhoods

https://doi.org/10.5194/essd-2026-345
Preprint. Discussion started: 17 June 2026

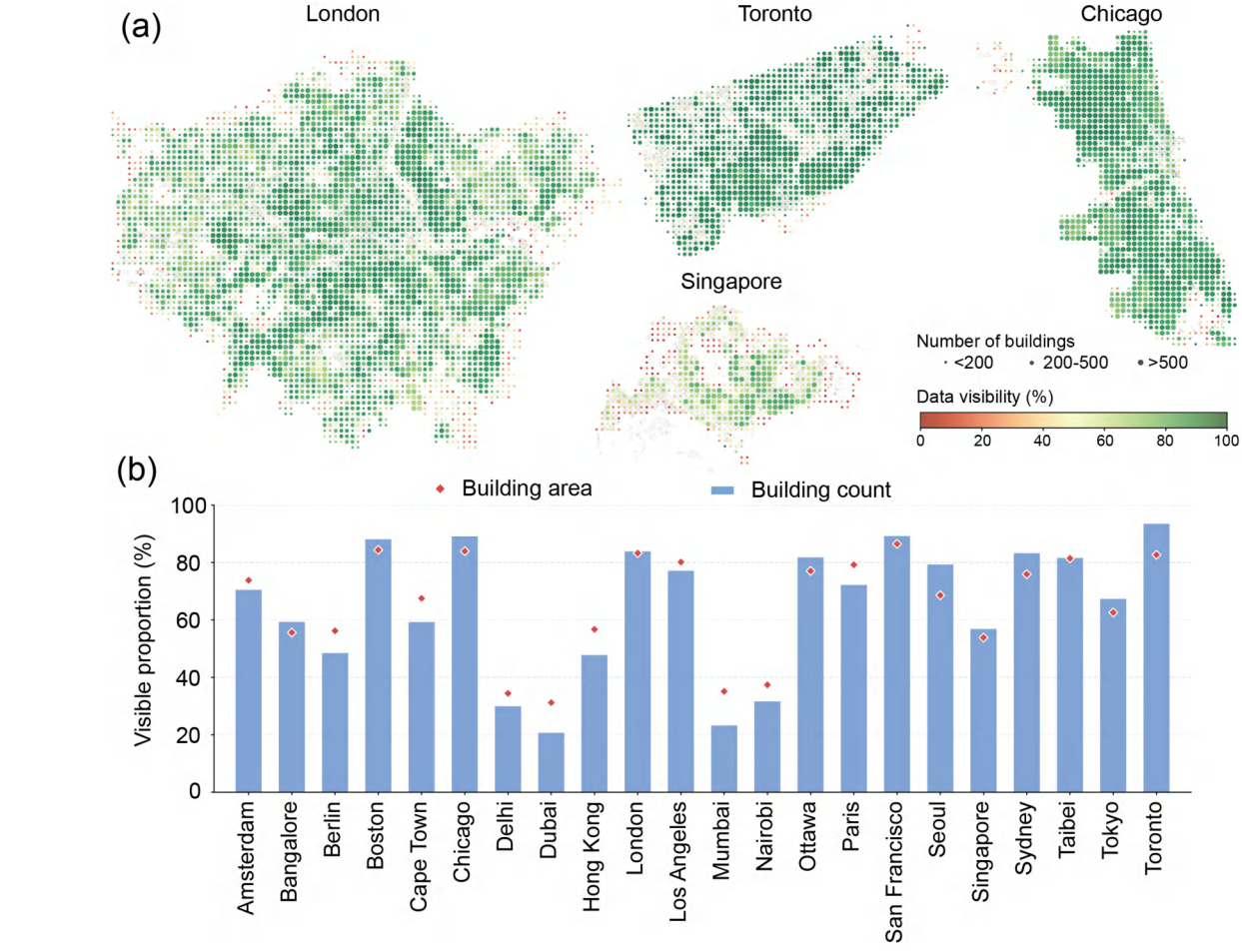
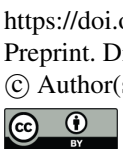


Figure 8. Spatial coverage of building material data inferred from GSV images: (a) Building counts and material coverage proportions at 1 km grid resolution in representative cities; (b) Proportion of buildings with material information by count (blue bars) and by area (red scatter) at the city level.

Figure 9 illustrates the spatial coverage of building material data inferred from BSV images. Figure 9(a) shows building counts and material coverage proportions at 2 km grid resolution in mainland Chinese cities. Central urban areas demonstrate notably higher material data visibility, while suburban zones show progressively lower proportions of buildings with material information, revealing distinct spatial patterns. Figure 9(b) and Supplementary Table S2 present the proportion of buildings with material information across buffer zones at increasing distances from city centers. These proportions generally decline with distance from urban centers. Some exceptions exist, such as Beijing, where the proportion within the 5 km buffer is lower than within the 10 km buffer, primarily due to street view image gaps in restricted areas like the Forbidden City. Overall, while

https://doi.org/10.5194/essd-2026-345
Preprint. Discussion started: 17 June 2026



Open Access Earth System Science Data Discussions

BSV-derived material data exhibit lower completeness than GSV-derived data, average proportions of buildings with material information within 10 km buffers reach approximately 40%.

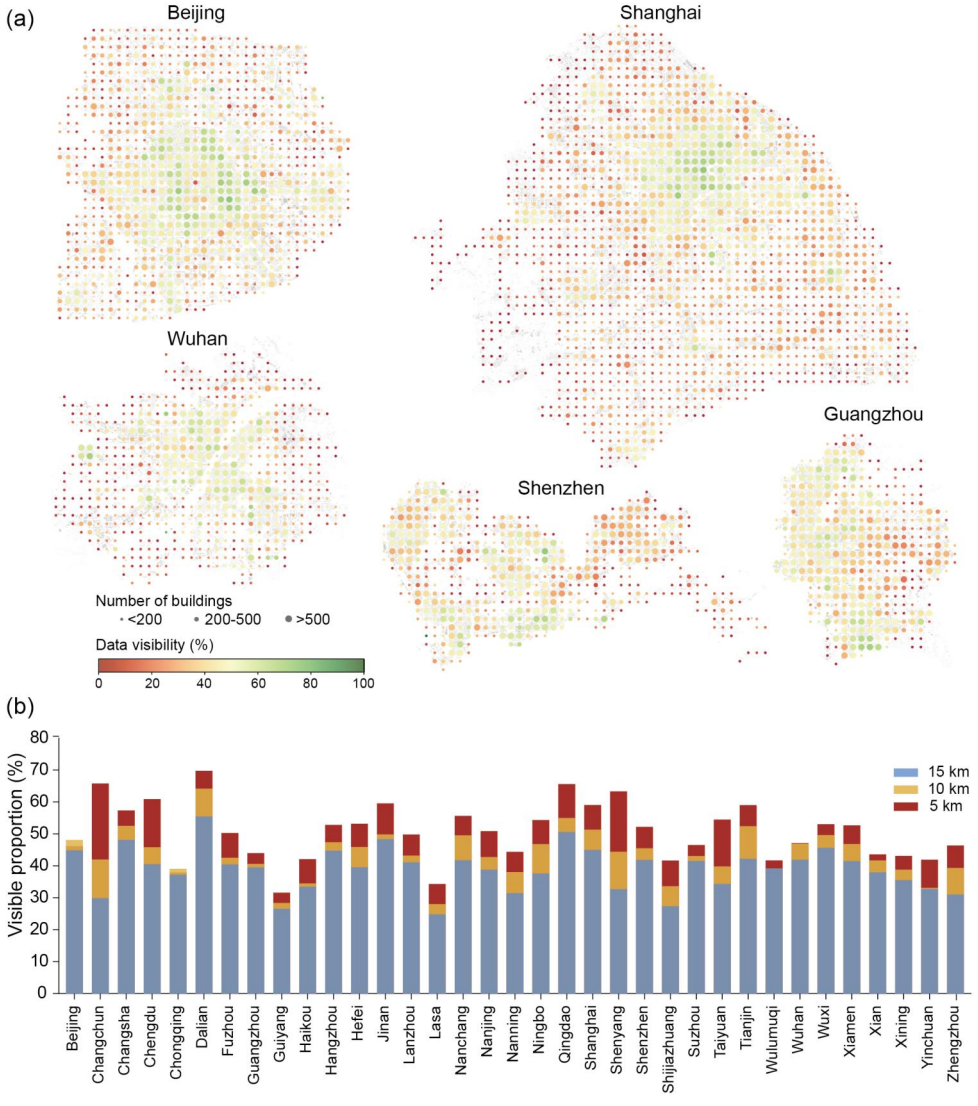


Figure 9. Spatial coverage of building material data inferred from BSV images: (a) Building counts and material coverage proportions at 2 km grid resolution in representative cities; (b) Proportions of visible building materials at varying distances from city centers.

To assess the spatial coverage quality of the dataset, we quantified building material diversity (Shannon index) and street-view visibility across concentric annular buffer zones at 3-km intervals extending to 30 km from each city center (Figure 10(a–b), Supplementary Figures S4-S5). Visibility rates generally decline with distance from city centers, reflecting sparser

https://doi.org/10.5194/essd-2026-345
Preprint. Discussion started: 17 June 2026

Open Access Earth System Science Data Discussions

street-view sampling in peripheral areas. The degree of decline varies across cities but valid material records are retained across the full 30-km range in all cases, confirming adequate spatial coverage from urban cores to suburban peripheries.

Material diversity patterns are spatially coherent and geographically interpretable (Figure 10(c), Supplementary Figures S6-S7). Chinese cities sustain consistently high diversity across all buffer rings, reflecting rapid and spatially diffuse urbanization since the 1990s in which construction occurred simultaneously across all urban zones. Cities such as Chicago, Sydney, and Toronto show monotonically declining diversity outward, consistent with the concentric zone model where centers accumulate historically diverse building stock while suburbs are dominated by uniform residential materials. More complex profiles are equally interpretable: Singapore's U-shaped curve reflects standardized public housing in inner zones and industrial facilities at the fringe, while Mexico City's outward-increasing diversity captures informal settlements built with locally variable materials at the urban edge. These geographically consistent patterns across diverse global cities confirm that the dataset captures meaningful material variation across urban gradients.

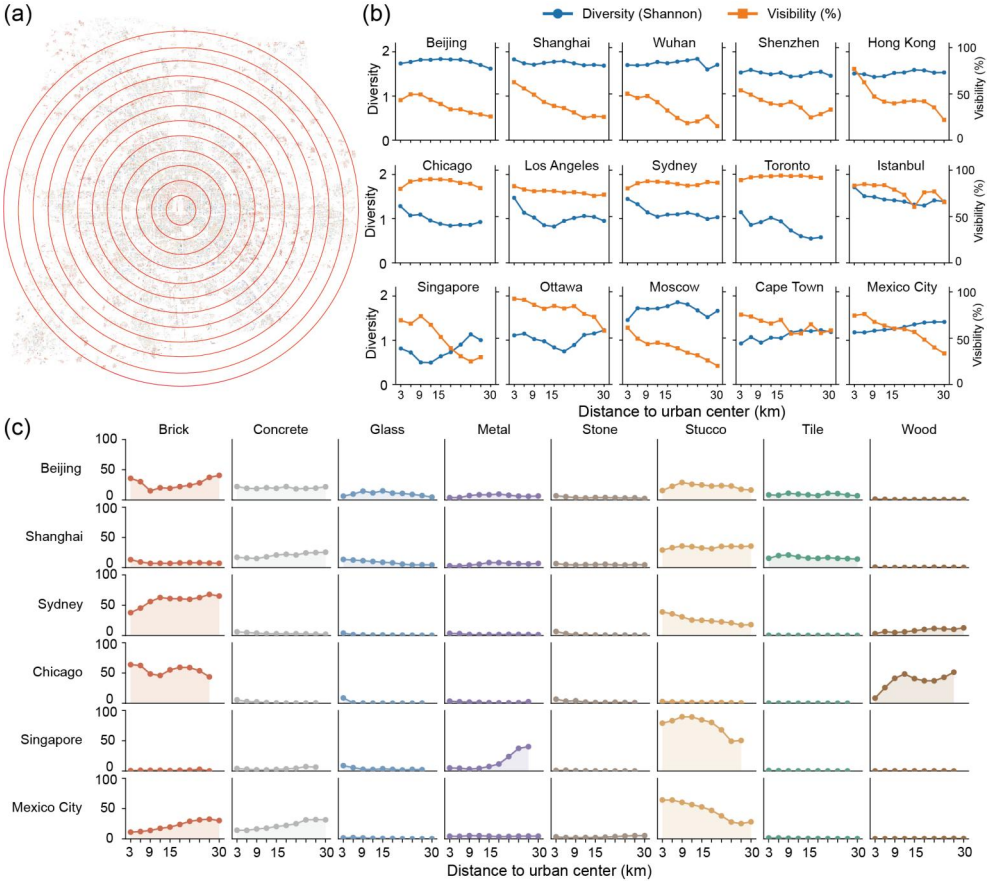


Figure 10. Spatial gradients of building material diversity and composition across different cities. (a) Illustration of equal-interval concentric buffers. (b) Shannon diversity trends and visibility rates across buffers. (c) Material proportions across buffer zones for some representative cities.

https://doi.org/10.5194/essd-2026-345
Preprint. Discussion started: 17 June 2026

4.4 Multi-year records of building facade materials

This dataset collects all available historical street-view images at the individual building level to derive year-specific material records. The temporal completeness of these records depends on street-view acquisition frequency, with GSV providing more frequent updates and more complete spatial coverage per cycle than BSV (Supplementary Figure S2).

GSV coverage spans 2007 to 2025 with frequent acquisition cycles, and the proportion of buildings with material records is relatively high across most cities, with many exceeding 40% coverage in many years (Figure 11). Cities ranked by overall coverage show that top-ranked cities such as Paris, San Francisco, and Boston maintain consistent records across nearly all years, while lower-ranked cities exhibit more intermittent patterns. BSV coverage starts from 2013 with less frequent acquisition intervals, and building coverage rates are generally lower than GSV (Figure 12). For most Chinese cities, material records are concentrated in several acquisition years rather than evenly distributed over time. Although BSV imagery has not been updated since 2023, the 2023 records still provide a recent snapshot of facade materials for most existing buildings in Chinese cities.

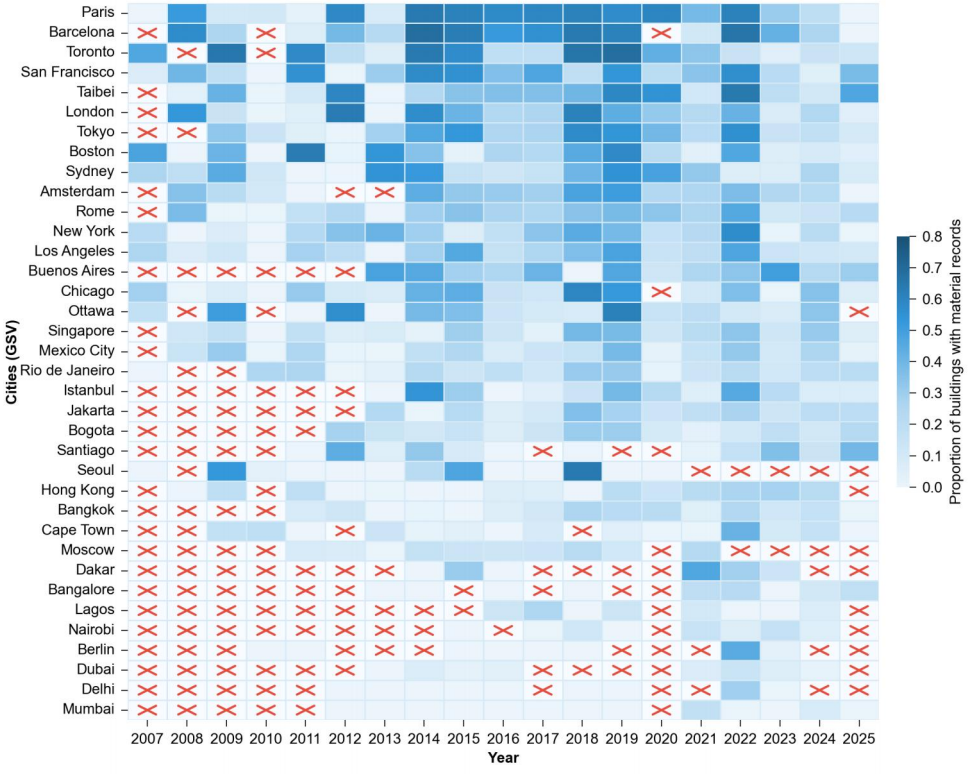


Figure 11. Annual proportion of buildings with material records derived from Google Street View (GSV) images across cities. Red crosses indicate years with no valid records.

https://doi.org/10.5194/essd-2026-345
Preprint. Discussion started: 17 June 2026

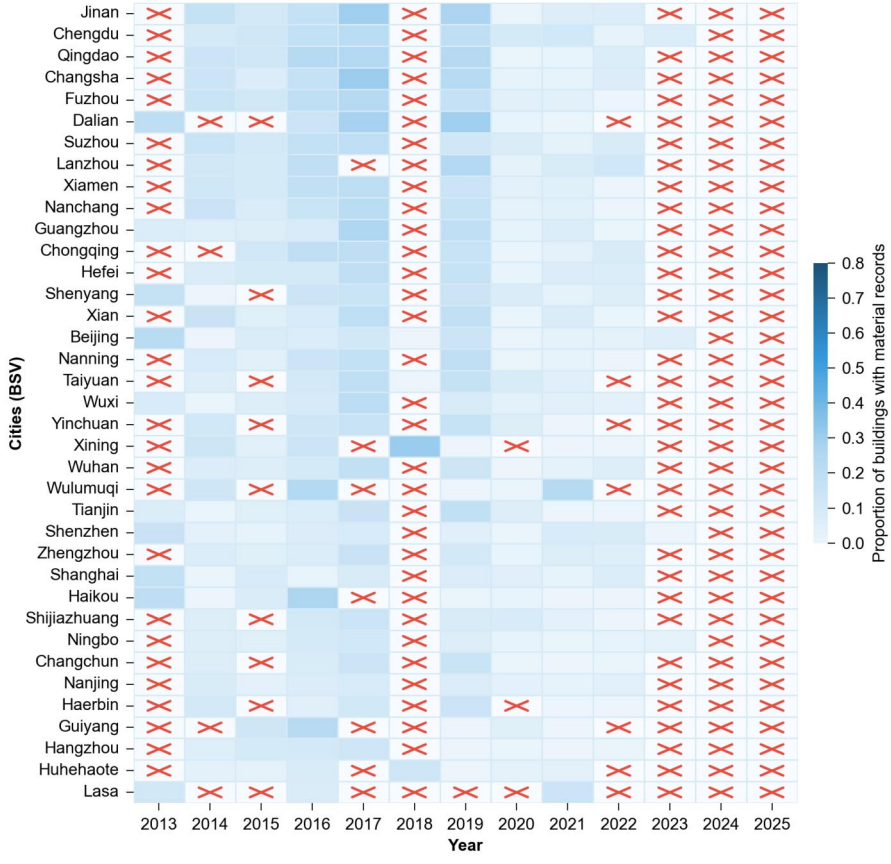


Figure 12. Annual proportion of buildings with material records derived from Baidu Street View (BSV) images across cities. Red crosses indicate years with no valid records.

To examine the potential use of BMAT's multi-year records, we tracked material transitions between 2015 and 2025 across five Asian cities (Figures 13(a)). The recorded transitions are broadly consistent with expected urban development patterns. Glass coverage increases across many cities, reflecting the broad trend toward modern cladding materials, further supported by building age analysis showing that glass and metal facades are concentrated in buildings constructed after 2000 (Supplementary Figures S8-S9). Chinese mainland cities show greater material turnover than Singapore and Tokyo, consistent with their faster pace of ongoing urbanization. At the individual-building level, paired street-view images show visually verifiable examples of apparent facade transitions, including brick-to-glass and concrete-to-tile changes on the same buildings across years (Figure 13(b)). These temporally coherent patterns suggest that the dataset's multi-temporal records can reflect real-world differences of urban renewal across cities.

https://doi.org/10.5194/essd-2026-345
Preprint. Discussion started: 17 June 2026

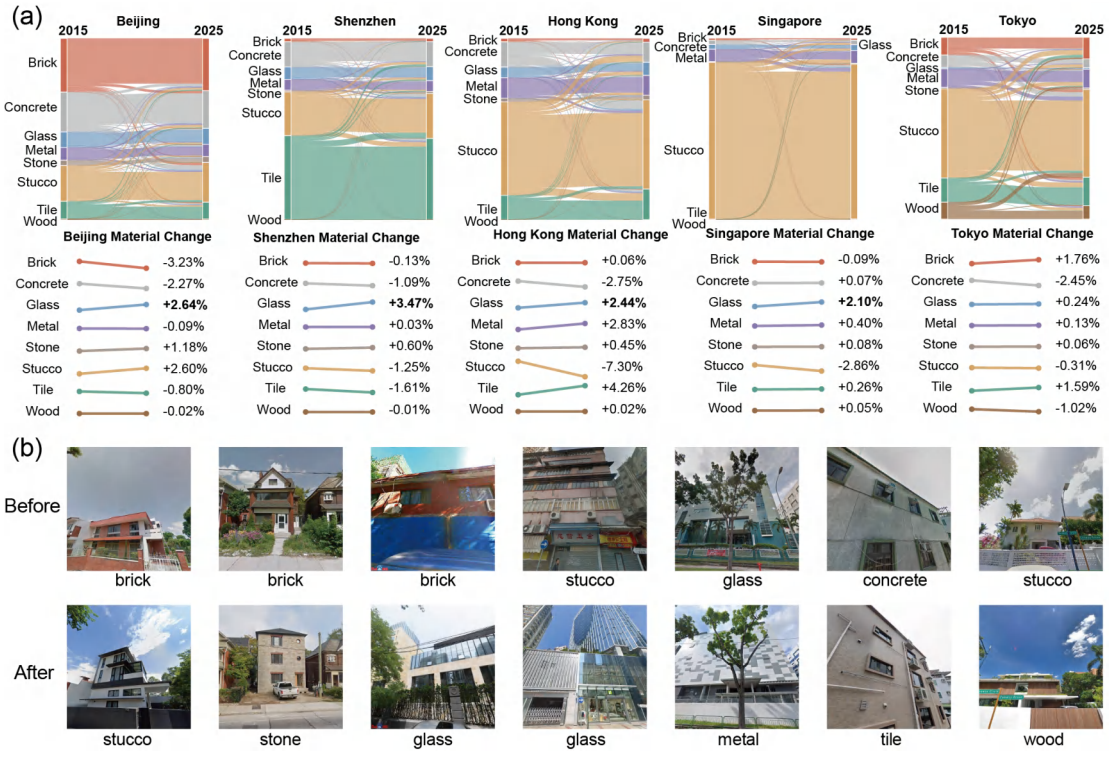


Figure 13. Building material transitions over time: (a) Flows and statistical shifts of material composition in case study cities; (b) Visual comparison of building facade before and after material changes.

5 Discussion

To the best of our knowledge, this study develops the first building material dataset that combines large-scale coverage with individual building-level precision. In addition to these spatial features, this dataset incorporates a temporal dimension, explicitly mapping time-series material attributes to specific building footprints. This dataset is derived from over 147 million street view building facade images. We utilized a fine-tuned Vision-Language Model (Qwen2.5-VL-7B-Instruct), which was trained on 39,405 human-labeled facade images—the largest manually annotated dataset for building material classification to date, notably exceeding the second-largest benchmark of 17,914 images (Chen et al., 2025). As a result, the model achieves an overall accuracy of 0.91, surpassing traditional deep learning models and ensuring high data reliability for downstream analysis.

This dataset enables diverse applications in built environment research and urban analytics. The individual building-level material information supports refined building stock estimation (Liang et al., 2025a; Fishman et al., 2024), facilitating more accurate assessments of urban material composition and construction patterns (Dai et al., 2024; Rajaratnam et al., 2025). Material attributes can be integrated into urban heat environment simulations to improve thermal modeling accuracy (Tarkhan et al., 2025; Wonorahardjo et al., 2022), and incorporated into building energy consumption estimation by accounting for

https://doi.org/10.5194/essd-2026-345
Preprint. Discussion started: 17 June 2026

envelope material properties (Liu and Abbasabadi, 2025; Liu et al., 2025). The dataset also enables specialized analyses such as photovoltaic potential assessment of glass facades (Dong et al., 2025; Zhao and Gou, 2023) and material-specific building vulnerability studies for hazards including floods (Balasbaneh et al., 2019; Xing et al., 2023) and fires (Zamanialaei et al., 2025; Kee et al., 2025). In addition, The year-specific material records can provide exploratory evidence of temporal material changes associated with urban renewal processes (Naik et al., 2017; Xue et al., 2025) and life cycle assessment of building material (Sharma et al., 2011; Bribián et al., 2011).

Beyond applications in the built environment, the dataset's metadata structure offers unique analytical opportunities. As material records derive from temporally-stamped street view images, the data inherently capture spatiotemporal sampling patterns of street view platforms (Biljecki and Ito, 2021; Kim and Jang, 2023). Temporal analysis can reveal geographic disparities in image update frequency (Wang et al., 2024c), identifying areas with outdated building records (Supplementary Figure S3) or systematic differences between platforms (e.g., Google Street View vs. Baidu Street View) (Wang et al., 2025). Spatially, the absence of material records may serve as a proxy for urban accessibility constraints, as street view vehicles cannot penetrate extremely narrow roads or informal settlements (Kamalipour and Dovey, 2019). This characteristic enables the indirect identification of underdeveloped urban areas, including informal housing and marginalized neighborhoods that lack coverage of street view images (Fan et al., 2025).

However, this dataset also has some limitations. First, because the data relies on street view images, it cannot capture every building in a city (Fan et al., 2025). Buildings located in gated communities or far away from main roads may be missing. Since building materials often show strong similarities within the same neighborhood (Fleischmann et al., 2020), future work could potentially address this by using data from nearby buildings to estimate these missing attributes. Second, the dataset currently records only the dominant facade material, which may simplify the representation of buildings composed of mixed materials (Wang et al., 2024b; Chen et al., 2025). It is worth noting that the time-series material information recorded in this study offers a potential way to address this. Frequent changes in predicted materials over time, for example, switching between brick and glass, might not be model classification errors. Instead, these variations likely reflect the fact that the building is composed of multiple materials, causing the model to generate inconsistent predictions across different years. Third, the dataset relies on a single street-level viewpoint. Therefore, it may not capture materials on the rear or side walls that may differ from the street-facing facade.

6 Conclusions

This study presents the first large-scale, footprint-level dataset of building facade materials, effectively bridging the gap between geometric building footprint data and material attributes. Derived from over 147 million street view images via a fine-tuned Vision-Language Model, the dataset covers 73 major cities with high classification accuracy. The dataset records material attributes for 22.09 million individual buildings across major cities worldwide. In addition, unlike conventional datasets that capture material information at a single time point, this dataset includes temporal records spanning multiple years, enabling the analysis of material changes and building renovation patterns over time.

https://doi.org/10.5194/essd-2026-345
Preprint. Discussion started: 17 June 2026

Open Access Earth System Science Data Discussions

This building material dataset provides new opportunities for built environment research and urban studies. The individual building-level material information supports more accurate assessments of building energy performance, embodied carbon, and disaster vulnerability, moving from assumption-based or material-blind approaches to empirical, material-specific assessments. Furthermore, the temporal material records allow for tracking of urban renewal processes and building stock transformations, providing essential data support for urban sustainability.

Code availability. The code used in this study is available at https://github.com/yhyJoy/BMAT

Data availability. The BMAT data is openly available at https://doi.org/10.6084/m9.figshare.31569370 (Yin, 2026)

Author contributions. **Hanyu Yin**: Conceptualization, Data curation, Investigation, Methodology, Software, Visualization, Writing (original draft preparation), Writing (review and editing). **Fan Zhang**: Funding acquisition, Project administration, Supervision, Writing (original draft preparation), Writing (review and editing). **Yuqing Wang**: Data curation, Formal analysis, Visualization, Writing (original draft preparation). **Lun Wu**: Resources, Supervision. **Yu Liu**: Resources, Supervision.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. This work was supported by the National Natural Science Foundation of China (Grant No. 42371468). This work was also supported by High-performance Computing Platform of Peking University.

https://doi.org/10.5194/essd-2026-345
Preprint. Discussion started: 17 June 2026

References

Balasbaneh, A. T., Marsono, A. K. B., and Gohari, A.: Sustainable materials selection based on flood damage assessment for a building using LCA and LCC, Journal of Cleaner Production, 222, 844–855, 2019.

Biljecki, F. and Ito, K.: Street view imagery in urban analytics and GIS: A review, Landscape and Urban Planning, 215, 104 217, 2021.

Biljecki, F., Chow, Y. S., and Lee, K.: Quality of crowdsourced geospatial building information: A global assessment of OpenStreetMap attributes, Building and Environment, 237, 110 295, 2023.

Bribián, I. Z., Capilla, A. V., and Usón, A. A.: Life cycle assessment of building materials: Comparative analysis of energy and environmental impacts and evaluation of the eco-efficiency improvement potential, Building and environment, 46, 1133–1140, 2011.

Che, Y., Li, X., Liu, X., Wang, Y., Liao, W., Zheng, X., Zhang, X., Xu, X., Shi, Q., Zhu, J., et al.: 3D-GloBFP: The first global three-dimensional building footprint dataset, Earth System Science Data Discussions, 2024, 1–28, 2024.

Chen, X., Ding, X., and Ye, Y.: Mapping sense of place as a measurable urban identity: Using street view images and machine learning to identify building façade materials, Environment and Planning B: Urban Analytics and City Science, 52, 965–984, 2025.

Dai, M., Jurczyk, J., Arbabi, H., Mao, R., Ward, W., Mayfield, M., Liu, G., and Tingley, D. D.: Component-level residential building material stock characterization using computer vision techniques, Environmental Science & Technology, 58, 3224–3234, 2024.

Dong, K., Yu, Q., Guo, Z., Xu, J., Tan, H., Zhang, H., and Yan, J.: Advancing building facade solar potential assessment through AIoT, GIS, and meteorology synergy, Advances in Applied Energy, 17, 100 212, 2025.

Fan, Z., Feng, C.-C., and Biljecki, F.: Coverage and bias of street view imagery in mapping the urban environment, Computers, Environment and Urban Systems, 117, 102 253, 2025.

Fishman, T., Mastrucci, A., Peled, Y., Saxe, S., and van Ruijven, B.: RASMI: Global ranges of building material intensities differentiated by region, structure, and function, Scientific Data, 11, 418, 2024.

Fleischmann, M., Feliciotti, A., Romice, O., and Porta, S.: Morphological tessellation as a way of partitioning space: Improving consistency in urban morphology at the plot scale, Computers, Environment and Urban Systems, 80, 101 441, 2020.

Güneralp, B., Zhou, Y., Ürge-Vorsatz, D., Gupta, M., Yu, S., Patel, P. L., Fragkias, M., Li, X., and Seto, K. C.: Global scenarios of urban density and its impacts on building energy use through 2050, Proceedings of the National Academy of Sciences, 114, 8945–8950, 2017.

Gurney, K. R., Liang, J., Roest, G., Song, Y., Mueller, K., and Lauvaux, T.: Under-reporting of greenhouse gas emissions in US cities, Nature communications, 12, 553, 2021.

Herfort, B., Lautenbach, S., Porto de Albuquerque, J., Anderson, J., and Zipf, A.: A spatio-temporal analysis investigating completeness and inequalities of global urban building data in OpenStreetMap, Nature Communications, 14, 3985, 2023.

Hsu, A., Sheriff, G., Chakraborty, T., and Manya, D.: Disproportionate exposure to urban heat island intensity across major US cities, Nature communications, 12, 2721, 2021.

Huang, S.-L. and Hsu, W.-L.: Materials flow analysis and emergy evaluation of Taipei's urban construction, Landscape and urban planning, 63, 61–74, 2003.

Kamalipour, H. and Dovey, K.: Mapping the visibility of informal settlements, Habitat International, 85, 63–75, 2019.

Kee, T., Qiu, B., Du, Q., and Sui, J.: An application of AHP-based fire vulnerability assessment for 20th century mountainous built heritage, npj Heritage Science, 13, 350, 2025.

Kim, J. and Jang, K. M.: An examination of the spatial coverage and temporal variability of Google Street View (GSV) images in small-and medium-sized cities: A people-based approach, Computers, Environment and Urban Systems, 102, 101 956, 2023.

https://doi.org/10.5194/essd-2026-345
Preprint. Discussion started: 17 June 2026

Liang, H., Zhang, B., Qian, X., Chen, Y., Luo, Y., Huang, Y., Sun, L., Zhou, C., Long, J., and Dong, L.: Building material stocks and embodied carbon dataset for two urban agglomerations in China from 2000 to 2020, Scientific Data, 12, 930, 2025a.

Liang, X., Xie, J., Zhao, T., Stouffs, R., and Biljecki, F.: OpenFACADES: An Open Framework for Architectural Caption and Attribute Data Enrichment via Street View Imagery, arXiv preprint arXiv:2504.02866, 2025b.

Liu, Y. and Abbasabadi, N.: Enhancing urban building energy models with Vision Transformers: A Case study in material classification from Google street view, Energy and Buildings, 333, 115 457, 2025.

Liu, Y. and Lei, D.: Painting humid cities cool: Cooling technologies, Nature Energy, pp. 1–2, 2025.

Liu, Z.-A., Li, Y., Hou, J., Tian, L., Wang, S., Hu, W., and Zhang, L.: Impact of exterior envelope thermal performance on energy demand and optimization strategies for university teaching-office buildings, Scientific Reports, 15, 15 171, 2025.

Manoli, G., Fatichi, S., Schläpfer, M., Yu, K., Crowther, T. W., Meili, N., Burlando, P., Katul, G. G., and Bou-Zeid, E.: Magnitude of urban heat islands largely explained by climate and population, Nature, 573, 55–60, 2019.

Microsoft: Microsoft Building Footprints, https://github.com/microsoft/GlobalMLBuildingFootprints, accessed December 29, 2025, 2024.

Naik, N., Kominers, S. D., Raskar, R., Glaeser, E. L., and Hidalgo, C. A.: Computer vision uncovers predictors of physical urban change, Proceedings of the National Academy of Sciences, 114, 7571–7576, 2017.

Raghu, D., Bucher, M. J. J., and De Wolf, C.: Towards a 'resource cadastre'for a circular economy–urban-scale building material detection using street view imagery and computer vision, Resources, Conservation and Recycling, 198, 107 140, 2023.

Rajaratnam, D., Stewart, R. A., Liu, T., and Vieira, A. S.: A novel computer vision-based approach for autonomous building facade material stock estimation, Resources, Conservation and Recycling, 219, 108 311, 2025.

Sharma, A., Saxena, A., Sethi, M., Shree, V., et al.: Life cycle assessment of buildings: a review, Renewable and Sustainable Energy Reviews, 15, 871–875, 2011.

Sirko, W., Brempong, E. A., Marcos, J. T., Annkah, A., Korme, A., Hassen, M. A., Sapkota, K., Shekel, T., Diack, A., Nevo, S., et al.: High-resolution building and road detection from sentinel-2, arXiv preprint arXiv:2310.11622, 2023.

Sun, K., Li, Q., Liu, Q., Song, J., Dai, M., Qian, X., Gummidi, S. R. B., Yu, B., Creutzig, F., and Liu, G.: Urban fabric decoded: High-precision building material identification via deep learning and remote sensing, Environmental Science and Ecotechnology, 24, 100 538, 2025.

Tarkhan, N., Klimenka, M., Fang, K., Duarte, F., Ratti, C., and Reinhart, C.: Mapping facade materials utilizing zero-shot segmentation for applications in urban microclimate research, Scientific Reports, 15, 5492, 2025.

United Nations Population Division: World Urbanization Prospects: The 2014 Revision, Tech. rep., United Nations Department of Economic and Social Affairs/Population Division, https://esa.un.org/unpd/wup/, accessed December 29, 2025, 2014.

Wang, C., Song, J., Shi, D., Reyna, J. L., Horsey, H., Feron, S., Zhou, Y., Ouyang, Z., Li, Y., and Jackson, R. B.: Impacts of climate change, population growth, and power sector decarbonization on urban building energy use, Nature Communications, 14, 6434, 2023.

Wang, G., Ryu, K., Dong, Z., Hu, Y., Ke, Y., Dong, Z., and Long, Y.: Micro/nanofabrication of heat management materials for energy-efficient building facades, Microsystems & Nanoengineering, 10, 115, 2024a.

Wang, L., Zhang, T., He, J., and Kada, M.: Cross-platform complementarity: Assessing the data quality and availability of Google Street View and Baidu Street View, Transactions in Urban Data, Science, and Technology, 4, 22–47, 2025.

Wang, S., Park, S., Park, S., and Kim, J.: Building façade datasets for analyzing building characteristics using deep learning, Data in Brief, 57, 110 885, 2024b.

https://doi.org/10.5194/essd-2026-345
Preprint. Discussion started: 17 June 2026

Wang, Z., Ito, K., and Biljecki, F.: Assessing the equity and evolution of urban visual perceptual quality with time series street view imagery, Cities, 145, 104 704, 2024c.

Wonorahardjo, S., Sutjahja, I. M., Mardiyati, Y., Andoni, H., Achsani, R. A., Steven, S., Thomas, D., Tunçbilek, E., Arıcı, M., Rahmah, N., et al.: Effect of different building façade systems on thermal comfort and urban heat island phenomenon: An experimental analysis, Building and Environment, 217, 109 063, 2022.

Xie, J., Li, M., Wu, J., Zhang, X., and Zhang, J.: Semantic segmentation of building façade materials and colors for urban conservation, npj Heritage Science, 13, 378, 2025.

Xing, Z., Yang, S., Zan, X., Dong, X., Yao, Y., Liu, Z., and Zhang, X.: Flood vulnerability assessment of urban buildings based on integrating high-resolution remote sensing and street view images, Sustainable Cities and Society, 92, 104 467, 2023.

Xue, X., Tian, Z., Yang, Y., Wang, J., and Cao, S.-J.: Sustaining the local color of a global city, Nature Cities, pp. 1–13, 2025.

Yap, W., Wu, A. N., Miller, C., and Biljecki, F.: Revealing building operating carbon dynamics for multiple cities, Nature Sustainability, 8, 1199–1210, 2025.

Yin, H.: BMAT: A Global Dataset of Building Material Attributes and Temporal Records at the Footprint Level, https://doi.org/10.6084/m9.figshare.31569370, dataset, 2026.

Zamanialaei, M., San Martin, D., Theodori, M., Purnomo, D. M. J., Tohidi, A., Lautenberger, C., Qin, Y., Trouvé, A., and Gollner, M.: Fire risk to structures in California's Wildland-Urban Interface, Nature Communications, 16, 8041, 2025.

Zhang, C., Hu, M., Sprecher, B., Sacchi, R., Yang, X., Yang, S., Verhagen, T. J., Zhang, C., Steubing, B., and Tukker, A.: Revealing the interplay between decarbonisation, circularity, and cost-effectiveness in building energy renovation, Nature Communications, 16, 7153, 2025.

Zhao, K. and Gou, Z.: Influence of urban morphology on facade solar potential in mixed-use neighborhoods: Block prototypes and design benchmark, Energy and Buildings, 297, 113 446, 2023.

Zhu, X. X., Chen, S., Zhang, F., Shi, Y., and Wang, Y.: GlobalBuildingAtlas: an open global and complete dataset of building polygons, heights and LoD1 3D models, Earth System Science Data Discussions, 2025, 1–31, 2025.