Federated Machine Learning-Based Urban Attributes Mapping using Multi-Source Urban Data

Junxian Yu [1], Xiana Chen [1], Zhaoyue Cai [1], Wei Tu [1,2,3*], Mo Su [4], Jianghai Liao [5], Tsz Nam Chan [5]

[1] Shenzhen Key Laboratory of Spatial Information Smart Sensing and Services, and Department of Urban Informatics, School of Architecture and Urban Planning, Shenzhen University, 3688 Nanhai Avenue, Nanshan District, Shenzhen, Guangdong, China - (yujunxian2019, chenxiana2019, 2021104012)@email.szu.edu.cn, tuwei@szu.edu.cn
[2] Key Laboratory for Geo-Environmental Monitoring of Great Bay Area, Ministry of Nature Resource, Shenzhen University, Shenzhen, China - tuwei@szu.edu.cn
[3] State Key Laboratory of Subtropical Building and Urban Science, Shenzhen, China - tuwei@szu.edu.cn
[4] Shenzhen Urban Planning and Land Resource Research Center, Shenzhen, China - 2014102050019@whu.edu.cn
[5] College of Computer Science and Software Engineering, Shenzhen University, 3688 Nanhai Avenue, Nanshan District, Shenzhen, Guangdong, China - (liaojh, edisonchan)@szu.edu.cn

Keywords: Multi-Source Urban Data, Hetero Federated Learning, Urban Attribute Mapping, Data Sharing.

Abstract

The complementary nature of multi-source urban data (including remote sensing images, mobile phone data, and other spatial datasets) underscores the critical importance of data fusion in spatial attribute mapping. However, exponentially growing data coupled with decentralized data management has rendered traditional centralized data analytics and learning models increasingly inefficient at scale. To address these challenges, we propose an alternative urban attribute mapping framework based on federated learning which resolves data silo issues in multi-source collaborative modeling. Unlike traditional direct data sharing, the presented framework shares encrypted handcrafted features derived from multi-source data between dataholders. Therefore, it preserves the privacy of original urban data. With popular machine learning methods, the framework enables accurate inference of diverse urban attributes such as population, economic development, urban mobility, land use, and air quality. Experimental results from a case study in Shenzhen demonstrate that the presented framework successfully facilitates multi-source collaborative mapping solely through the exchange of model parameters and structures. The framework exhibits exceptional performance across five urban attribute mapping tasks: non-resident population, GDP, taxi travels, land use mix, and PM2.5 concentration. These results validate the effectiveness of multi-source data collaboration in data-rich environments. The primary contribution of this research lies in the development of a distributed multi-source federated mapping framework for urban attributes, offering an alternative solution to overcome urban data silos while establishing practical foundations for expanding urban mapping and cross-regional applications.

1. Introduction

Urban attributes comprise the fundamental components that characterize cities' socioeconomic, physical, and functional structures, including population distribution, economic development, land use patterns, human mobility, etc (Fan et al., 2023; Cao et al., 2025a; Fang et al., 2024). The mapping of urban attributes faces many challenges, such as unavailable high-resolution urban population data and unclear land use mix information in some areas. Traditionally, urban attribute mapping relies on survey data collected by professional personnel, which is time-consuming and cannot be updated in a timely manner. For example, the population surveys are conducted every ten years in China, therefore the urban population density data does not keep pace with urban development, especially for rapidly changing Chinese cities.

Inferring urban attributes from multi-source urban data with machine learning is an alternative approach. Different urban data sources contain diverse and valuable urban information. For example, remote sensing imagery (Wang et al., 2024) and street view imagery (Wu et al., 2023) reveal physical structures, while building footprints (Li et al., 2023) and points-of-interest (POI) data (Bai et al., 2023) represent facility distributions, and vehicle trajectory data (Huang et al., 2025) reflects human activity patterns. However, acquiring such comprehensive data presents significant challenges. High-resolution remote sensing imagery incurs substantial licensing fees per scene, while street view imagery is constrained by API call limits and stringent commercial usage terms (Danish et al., 2025). POI data, though seemingly accessible, faces similar API restrictions and platform-exclusive licensing, complicating cross-source integration (Sun et al., 2023). Human mobility data, including mobile positioning and vehicle trajectories (Brauer et al., 2024), is heavily restricted by privacy regulations and data protection laws. Despite these challenges, multi-source urban data provides a comprehensive capture of multidimensional information across urban areas, offering unprecedented insights into complex dynamics of urban environments. Recently, these data have demonstrated good performance in data-driven urban studies, including population density estimation, traffic flow prediction and PM2. 5 prediction (Chen et al., 2022; Feng et al., 2024; Putra et al., 2021).

However, data-driven urban attribute mapping confronts multiple challenges. First, these diverse datasets are usually owned by different governmental sectors, organizations, or companies. Traditionally, raw urban data is exchanged according to strict laws or regulations, which impose heavy barriers to the wide usage of these datasets. Therefore, they are isolated in "data silos", making it difficult to achieve information exchange and sharing, particularly as privacy concerns and regulatory restrictions continue to intensify. Consequently, alternative approaches are needed to bridge these data silos and mitigate the challenges associated with centralized data collection and processing.

Second, there are diverse urban attribute mapping tasks, such as population portraits, environment mapping, and land use mapping. They are usually solved by different machine

learning-based models or deep neural networks. For example, Zhao et al. (2021) mapped population distribution based on XGBoost using land use, POIs, subway routes, road networks, remote sensing imagery, building data, and Anjuke real estate data. Cao et al. (2022) combined remote sensing imagery and road networks to construct a multi-view graph neural network for estimating and mapping economic development. Cao et al. (2025b) proposed SemiGTX, a graph-based semi-supervised framework integrating street view imagery, POI, and mobility data for multi-sector economic mapping. In the past decades, tens, or even hundreds of urban models have been proposed to infer different attributes. Although these models extend the frontier of urban attributes mapping, they propose diverse features and model structures, therefore requiring a universal attribute inference model.

Figure 1. Urban data silo.

Recent years have witnessed remarkable advancements in distributed machine learning, which leverages multiple computing nodes to train and deploy machine learning algorithms through data or model parallelism. This paradigm offers novel solutions for multi-source spatiotemporal data integration under privacy constraints. Federated learning, built upon distributed frameworks, enables heterogeneous data sources to collaboratively develop high-performance machine learning models without disclosing raw data (Kairouz et al., 2021). By distributing training tasks to local devices where data originates, it eliminates data transmission while preserving privacy. Concurrently, participating data holders collectively enhance model accuracy and robustness through shared training insights, demonstrating significant potential to resolve the longstanding dilemma between "data silos" and "privacy preservation" (Pandya et al., 2023).

Currently, federated learning has achieved significant progress in enabling multi-source data collaboration. Salehi et al. (2022) locally fused LiDAR scans, GPS trajectories, and camera images on client devices. The server then distributed the optimized fused model back to clients. Gao et al. (2022) integrated power grid operations, socioeconomic statistics, and energy consumption data to achieve cross-provincial carbon emission predictions without raw data sharing. Fan et al. (2024) leveraged joint user data from four banks to train a risk control model, enabling financial risk assessment, thereby improving the identification of potential fraud and financial risks. These studies demonstrate the strong potential of federated learning in urban data analysis and urban computing by effectively integrating multi-source data from enterprises and government agencies.

To tackle the challenges of data silos and sharing in urban computing, this study leverages federated learning for urban attributes mapping. This approach maintains diverse urban data sources locally, while exchanging only model parameters to the server, instead of accessing raw data. It enables multiple stakeholders to collaboratively achieve urban attributes mapping while preserving data privacy though with additional computational and communication costs. The server aggregates these parameters into a global model and redistributes the global model for local updates. Through iterative optimization, the resulting federated model realizes urban attribute mapping (non-resident population, GDP, taxi travels, land use mix, and PM2.5 concentration) while maintaining data locality. We also develop a comprehensive urban attribute inference validation framework that encompasses multiple core urban dimensions: population, economy, transport, land use, and environment. The framework overcomes the constraints of 'data silos' that typically exist among urban management departments, research institutions, and commercial enterprises, thereby enabling integrated modeling of urban entities and advancing data-driven smart city applications.

2. Study Area and Datasets

2.1 Study Area

This study was conducted in Shenzhen, Guangdong Province, China. As one of China's Special Economic Zones, Shenzhen serves as a global innovation hub for technology industries and possesses abundant urban data resources, including satellite images, street-view imagery, POIs, building footprints, road networks, transit stations, and mobile signaling data. These diverse datasets provide rich materials for inferring multidimensional urban spatial elements. Shenzhen's rapid development and high-density urban characteristics make it an ideal case study for examining complex socioeconomic phenomena. Therefore, taking Shenzhen as the study area not only offers valuable insights for smart data and smart city management but also facilitates the application of the proposed methodology to other urban environments.

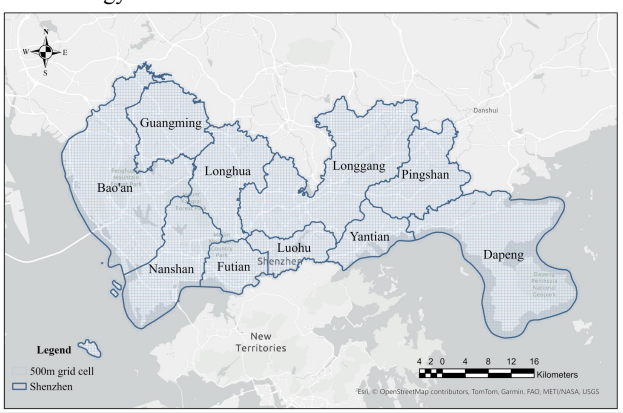

Figure 2. Study area.

2.2 Data

We employed a 500-meter grid system as the fundamental spatial unit (as illustrated in Figure 2), with each cell assigned a unique identifier. Several urban data sources were used, including the ZY-3 remote sensing imagery, street-view images, POIs, building footprints, road networks, transit stations, and mobile signaling data. Remote sensing imagery was obtained from the China Centre for Resources Satellite Data and Application. Street-view images were provided by Tencent. POIs, road networks, and transit stations were sourced from Baidu Maps and OpenStreetMap to capture the built environmental characteristics. Building footprints were obtained from the Bureau of Natural Resources and Urban Planning and

were produced by cadastral surveying. Mobile signal data was sourced from mobile communication operators to capture the changing population. These datasets are difficult to integrate due to their fragmented collection and management across various government agencies and enterprises.

To evaluate the performance of the models in urban attribute mapping, we utilized ground truth data across five key urban attributes, as shown in Table 1. These ground truth datasets were spatially aligned with the multi-source data, enabling supervised learning and validating the federated framework's urban attribute inference capability.

Attribute	Data
Population	Non-Resident Population
Economy	GDP
Transport	Taxi Travels
Land use	Land Use Mix Index
Environment	PM2.5 Concentration

Table 1. Ground truth data.

3. Methods

This study focuses on aligning sample data across massive urban datasets from diverse government agencies, enterprises, and open platforms, establishing a heterogeneous federated learning framework. Each participant trains machine learning models on local data and uploads parameters to a central server for aggregation. The server integrates parameters to produce and update a global model. It distributes updated global models back to nodes, iterating until convergence. The final federated model enables urban attribute mapping across population, economy, transport, land use, and environment. The methodological framework is illustrated in Figure 3.

3.1 Multi-source Data Feature Engineering

We orchestrate collaboration among multiple heterogeneous data custodians to model complex urban dynamics while maintaining strict data sovereignty. Each participant manages distinct domain-specific data assets with unique characteristics. To ensure data privacy protection, all participating entities perform data preprocessing locally before engaging in the federated learning process. Multi-source urban data features obtained through feature engineering are presented in Table 2. Five data sources are involved in this study.

- **Source A:** Maintains satellite imagery and street-view repositories. For remote sensing images, spectral features are extracted through mean and standard deviation calculations across three channels (6-dimensional features), while texture features are derived from GLCM metrics (12-dimensional features). Structural features are generated using Dense-SIFT (128-dimensional features). For street-view images, 19 semantic feature categories are extracted through Panoptic-DeepLab.

- **Source B:** Possesses POI distribution datasets with semantic richness (scientific research and education, catering, residential areas, shopping, large shopping plazas, tourism, leisure, medical, and finance), offering granular functional characterization of urban spaces. POI distributions are mapped to 500m grid cells through spatial aggregation, counting instances per category to generate 9-dimensional features for each grid.

- **Source C:** Maintains authoritative records of building footprints, road networks, and transit stations. Building footprints and road networks are processed to 500m grid resolution, with the following metrics extracted per cell: (a) For buildings: total area, count, and height characteristics (maximum, minimum, average). (b) For roads: classified length (L1/L2) and total areal coverage. Transport stations follow the same protocol as POI data.

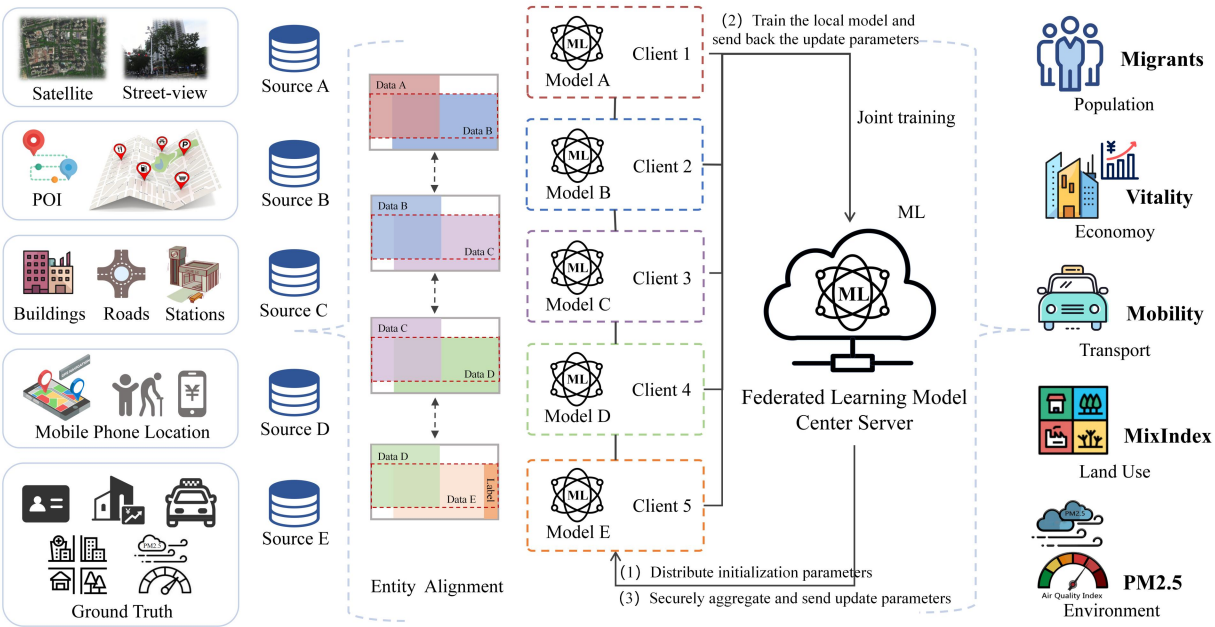

Figure 3. Overview of the workflow.

- **Source D:** Controls mobile phone signaling data with continuous temporal sampling and population-scale coverage, inherently containing sensitive personally identifiable information (PII) protected under privacy laws. For mobile signaling data, daily average trip counts per grid cell were calculated as both origin (o_grid) and destination (d_grid) using three weeks of March 2021 and two-week samples from April/May 2021. The analysis further disaggregated flows by age strata and prepaid/postpaid user categories.

- **Source E:** Holds various ground-truth datasets and selectively provides supervision signals including the number of non-resident population, GDP, taxi travels, land use mix, and PM2.5 concentration according to specific inferred requirements.

Data	Features	Dimension
Remote sensing image	spectral, texture, structural	146
Street view image	road, sidewalk, building, wall, fence, pole, traffic light, traffic sign, vegetation, terrain, sky, person, rider, car, truck, bus, train, motorcycle, bicycle	19
POI	scientific research and education, catering, residential areas, shopping, large shopping plazas, tourism, leisure, medical, finance	9
Buildings	area, quantity, height (max., min., average)	5
Road networks	length, area (L1, L2)	4
Transport stations	bus station, subway station, train station, motor station	4
Mobile signaling	age: 0-18, 19-44, 45-59, 60+	8
	Spending (CNY): 0-20, 20-50, 50-80, 80-200, 200+	10

Table 2. Multi-source data features.

All ground-truth datasets undergo geospatial processing: (1) Population: Non-resident population data are aggregated to 500m grids based on their spatial distribution, ensuring each individual is assigned to the grid where they are located; (2) Economy: 1km resolution GDP grid data obtained from the Institute of Geography, Chinese Academy of Sciences are resampled to 500m grids; (3) Transport: Average daily pick-offs and drop-offs per 500m grid are extracted from March and October 2021 GPS trajectories using spatial-temporal aggregation; (4) Land Use: 56 categories are reclassified into 9 superclasses (public, water, commercial, agricultural, green, industrial, residential, education-scientific, transport), with grid-level proportions and Shannon diversity index calculated; (5) Environment: 1km-resolution PM2.5 rasters (2021 annual average) are joined to 500m grids and converted to area-weighted values using zonal statistics.

Spatial harmonization occurs at the grid-cell level through secure cryptographic entity resolution techniques (detailed in Section 3.2), facilitating integrated modeling while preventing any exchange of raw data between sources.

3.2 Hetero Federated Learning

3.2.1 Entity Alignment

Hetero-federated learning enables collaborative model training where participants share the same sample space but have partially overlapping feature spaces. This setting is suitable for scenarios where multiple sources possess different urban data types but exhibit significant sample overlap. To align multi-source data samples while preserving privacy, we employ Private Set Intersection (PSI) based on secure cryptographic computations (Meadows, 1986).

To align sample IDs across sources $\{S_A, S_B, S_C, \ldots, S_N\}$ via a center server C, the protocol follows these key steps:

1. **Setup:** Each source generates a private key and shares its public key with C.
2. **Encrypted Submission:** Sources encrypt their samples using their private keys and send them to C.
3. **Aggregation & Decryption:** C aggregates encrypted values and requests partial decryptions from sources. Using threshold decryption, C verifies sample intersection.
4. **Result Distribution:** C broadcasts the aligned sample IDs to all sources for federated training.

Algorithm 1: Centralized Multi-Source PSI for Entity Alignment

Input:
- Sources $S_A, \ldots, S_N$ with sample sets $\mathcal{X}_A, \ldots, \mathcal{X}_N$
- Central server $\mathcal{C}$

Output: Aligned sample IDs $\mathcal{X}^* = \bigcap_{i=1}^{N} \mathcal{X}_i$;

Server Setup: foreach *source* S_i **do**
  S_i generates private key $s_i \in \mathbb{Z}_p$ // `Generate private key`
  $S_i \to \mathcal{C}$: $S_i = s_i \cdot G$ // `Public key`

Sample Submission: foreach *source* S_i **do**
  foreach *sample* $x \in \mathcal{X}_i$ **do**
    Compute $H(x) = \texttt{HashToPoint}(x)$ // `Hash the sample`
    $C_i(x) = s_i \cdot H(x)$ // `Encrypt the sample`
    $S_i \to \mathcal{C}$: $\{C_i(x)\}_{x \in \mathcal{X}_i}$ // `Send encrypted sample to server`

Intersection Computation: $\mathcal{C}$ selects reference set $\mathcal{X}_A$;
foreach $x \in \mathcal{X}_A$ **do**
  $C_{\text{joint}}(x) = \sum_{i=1}^{N} C_i(x)$ // `Aggregate the encrypted values`
  foreach *source* S_i **do**
    $\mathcal{C} \to S_i$: Request $K_i(x) = s_i^{-1} \cdot C_{\text{joint}}(x)$ // `Decrypt the aggregated value`
  Recover $H'(x) = \left(\sum_{i=1}^{N} K_i(x)\right)^{-1} \cdot C_{\text{joint}}(x)$;
  if $H'(x) == H(x)$ **then**
    Add x to $\mathcal{X}^*$ // `If match, add sample to intersection`

Result Distribution: foreach *source* S_i **do**
  $\mathcal{C} \to S_i$: Broadcast $\mathcal{X}^*$ // `Send aligned sample IDs to sources`

Figure 4. The pseudocode for the Entity Alignment.

3.2.2 Machine Learning based Attribute Prediction

Machine learning is a powerful tool for extracting insights from data by training models to recognize patterns and make predictions. It involves feeding data into algorithms that learn to map input features to output labels. XGBoost, as a popular machine learning model, stands out with its advanced Gradient Boosted Decision Trees (GBDT) algorithm, enabling efficient processing of structured data. It builds multiple decision trees sequentially to model complex feature relationships, learning from labeled datasets where input features (e.g., physical attributes, functional patterns) map to target values (e.g., demographics, economic indicators). For prediction, XGBoost

processes new data by traversing its ensemble of trees, combining their outputs to generate accurate and interpretable results. Considering these good characteristics, XGBoost is adopted for attribute inferences from features.

3.2.3 Joint Training

The HeteroSecureBoost model is an adaptation of the GBDT algorithm for secure multi-source federated learning, where participants collaboratively train a model without exposing their raw data (Chen et al., 2021). The core principle involves using gradient boosting in a federated setting where central sever coordinates the learning process while passive sources contribute encrypted features without revealing raw data. As shown in Figure 5, the model is built iteratively through boosted tree ensembles. For each tree, first-order gradients and second-order Hessians are calculated locally, encrypted, and then aggregated in feature histograms that enable the coordinator to determine optimal splits without accessing raw data. These splits maximize information gain while preserving privacy, and the resulting ensemble progressively improves prediction accuracy across boosting rounds. Below, we elaborate on the implementation with mathematical formulations.

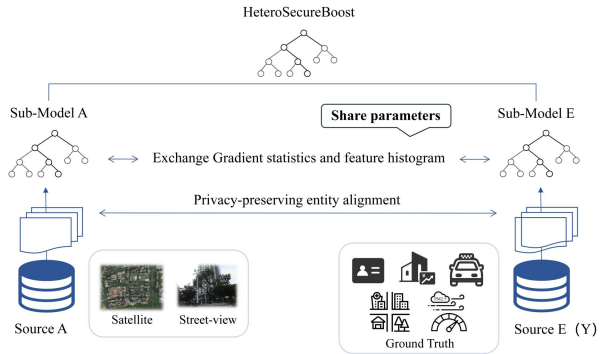

Figure 5. Federated SecureBoost.

The joint training process is iteratively implemented through T boosting rounds. In each iteration t, Source E first computes predictions $\hat{y}_i^{(t-1)}$ and encrypted gradients for all samples:

$$g_i=\hat{y}_i^{(t-1)}-y_i;\mathrm{h}_i=\hat{y}_i^{(t-1)}\left(1-\hat{y}_i^{(t-1)}\right), \quad (1)$$

where g_i = gradient of the loss function with respect to the predicted value for sample i
h_i = second-order derivative (Hessian) of the loss function with respect to the predicted value for sample i
$\hat{y}_i^{(t-1)}$ = predicted value for sample i at iteration $t-1$
y_i = true label for sample i

These encrypted values $Enc(g_i)$ and $Enc(\mathrm{h}_i)$ are distributed to passive sources $\{P_A, P_B, P_C, \ldots, P_N\}$, who construct feature histograms:

$$Enc(G_{j,k})=\sum_{i\in Bin_k} Enc(g_i), \quad (2)$$

$$Enc(H_{j,k})=\sum_{i\in Bin_k} Enc(\mathrm{h}_i), \quad (3)$$

where $Enc(G_{j,k})$ = encrypted sum of gradients for feature j in bin k
$Enc(H_{j,k})$ = encrypted sum of Hessians for feature j in bin k
Bin_k = set of samples in the k^{th} bin
$Enc(\bullet)$ = encryption function

Source E decrypts and aggregates these histograms to determine the optimal split through gain maximization:

$$Gain=\frac{1}{2}\left[\frac{G_L^2}{H_L+\lambda}+\frac{G_R^2}{H_R+\lambda}-\frac{(G_L+G_R)^2}{H_L+H_R+\lambda}\right]-\gamma, \quad (4)$$

where $Gain$ = split criterion to maximize
G_L, G_R = sum of gradients for left and right child nodes
H_L, H_R = sum of Hessians for left and right child nodes
λ = regularization parameter
γ = complexity control parameter

If the optimal feature belongs to a passive source, its encrypted split threshold $Enc(s^*)$ is retrieved. The model then updates additively:

$$\hat{y}_i^{(t)}=\hat{y}_i^{(t-1)}+\eta f_t(x_i), \quad (5)$$

where $\hat{y}_i^{(t)}$ = updated predicted value for sample i at iteration t
η = learning rate
$f_t(x_i)$ = output of the weak learner (tree) at iteration t for sample i

This process repeats until completion of T rounds, yielding the final ensemble model:

$$F(x)=\sum_{t=1}^{T}\eta f_t(x_i), \quad (6)$$

where $F(x)$ = final ensemble model prediction for input x
T = total number of boosting rounds
$f_t(x)$ = weak learner (tree) at round l
η = learning rate (may vary per round if subscripted)

3.3 Urban Attributes Mapping

The study trains five specialized models to predict distinct urban attributes. Each model targets a specific prediction task: (1) non-resident population distribution (persons/grid), (2) GDP estimation (10,000 CNY/km^2), (3) taxi travels (trips/grid), (4) land use mix index (entropy score 0-2), and (5) PM2.5 concentration (μg/m^3).

All prediction results were visualized as continuous heatmaps in 500m grid cells using the WGS84 coordinate system. Each urban attribute was displayed with distinct color schemes. These standardized raster layers enabled integrated visualization and pixel-level comparative analysis in GIS platforms, while maintaining precise spatial alignment across all five urban dimensions for comprehensive pattern recognition.

4. Experiments and Analysis

4.1 Experimental Setup

This experiment sets up a two-source federated learning scenario, deploying the federated learning environment using the open-source FATE framework (Liu et al., 2021) on two virtual machines. Source Guest serves as the central server with mobile signaling data. These data were processed to obtain 18-

dimensional human mobility features and prediction labels for any target spatial element. Source Host contains remote sensing images, street-view images, POIs, buildings, road networks, and transit stations. These data were processed to obtain 187-dimensional physical spatial features. The modeling process is initiated by the Source Guest, and joint training is conducted between the two data-holding participants, ensuring data privacy protection.

The evaluation framework employs a 50-50 data partitioning strategy, where half of the dataset with ground truth labels is further split into a 4:1 ratio for model training and testing. The trained models are then applied to predict the remaining half of the dataset. This design serves dual purposes: (1) The initial 50% split allows conventional performance validation using test metrics, Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and Coefficient of Determination (R^2) , while (2) the prediction on the second half simulates real-world deployment scenarios where models must map urban attributes in data-scarce regions.

4.2 Overall Results

Table 3 presents the centralized non-federated learning (Non-FL) performance across all five urban attributes, while Table 4 provides the corresponding federated learning (FL) results.

Urban Attributes	MAE	RMSE	MAPE	R^2
Non-Resident	0.352	0.539	3.565	0.727
GDP	0.411	0.561	2.770	0.702
Taxi Travels	0.184	0.418	0.714	0.785
Land Use Mix	0.502	0.639	1.891	0.554
PM2.5 Concentration	0.458	0.647	1.163	0.590
Average	0.381	0.561	2.021	0.672

Table 3. Results of Non-FL model.

Urban Attributes	MAE	RMSE	MAPE	R^2
Non-Resident	0.431	0.433	2.536	0.623
GDP	0.525	0.699	2.200	0.519
Taxi Travels	0.219	0.538	0.838	0.646
Land Use Mix	0.634	0.795	1.640	0.380
PM2.5 Concentration	0.631	0.834	1.695	0.318
Average	0.488	0.660	1.782	0.497

Table 4. Results of FL model.

The Non-FL approach demonstrates varying levels of predictive performance across the urban attributes. The model performs particularly well for Taxi Travels prediction with the highest R^2 and lowest error metrics. Non-Resident population and GDP also show good performance with R^2 values above 0.7. Land Use Mix and PM2.5 Concentration show moderate performance. The FL framework achieves 53.9-85.7% of the Non-FL model's predictive accuracy (R^2) across urban attributes, highlighting a trade-off between maintaining data privacy and predictive performance. Additionally, this privacy preservation comes at a substantial computational cost, with FL training times extending from seconds in centralized approaches to hours due to encrypted communication overhead. While the Non-FL baseline demonstrates superior R^2 metrics (average 35.2% higher R^2), the FL approach achieves better MAPE performance (11.8% lower than Non-FL).

HeteroSecureBoost demonstrates distinct performance patterns between dynamic and static urban attributes mapping. For certain dynamic attributes (non-resident, taxi travels), FL maintains 83.8% of Non-FL's explanatory power (R^2 difference < 0.14) with a 5.75% lower MAPE, showcasing the feasibility of decentralized learning in urban attributes mapping. However, significant performance degradation occurs in static attributes mapping (land use mix and PM2.5), where FL shows only 61.3% R^2 retention and 16.2% higher MAPE compared to Non-FL benchmarks. This discrepancy may stem from the inherent challenges of modeling static attributes in a federated setting, such as data distribution heterogeneity or model sensitivity to these features.

4.3 Urban Attribute Spatial Pattern Analysis

The spatial visualization of urban attributes presented in Figure 6 was generated using a composite mapping approach that integrates both ground truth data and model prediction values (shown within dashed boundaries).

The non-resident population mapping accurately captures high-density areas in central urban districts, while underestimating population in peripheral areas of Dapeng, northeastern Longgang and coastal Bao'an regions. GDP predictions correctly identify high-value concentrations in Nanshan, Futian and southern Luohu districts, but overestimate economic activity in southwestern Bao'an, northern Guangming, and eastern Pingshan. Taxi travel mapping effectively represents high-volume areas in central Nanshan, Futian and southwestern Luohu, while overestimating travel density in southern Longhua, parts of Longgang and Dapeng. Land use mix predictions underestimate diversity in northern Bao'an and central Longgang, while slightly overestimating coastal Yantian and Dapeng. PM2.5 concentration mapping reflects high pollution levels in Guangming, Bao'an and Longhua districts, but underestimates pollution in northern Longgang and southeastern Nanshan regions.

Shenzhen's urban attributes exhibit varied spatial distribution patterns across the city. The northwestern districts of Nanshan and Futian show high concentrations across all three attributes (non-resident population, GDP, and taxi trips), establishing them as the city's core areas. Luohu District, as the original central business district, demonstrates medium-to-high values in GDP (Fig. 6b) and taxi trip volumes (Fig. 6c), reflecting its status as Shenzhen's traditional urban center. Longgang District displays more dispersed patterns with several sub-centers of economic activity, particularly evident in the GDP distribution. Bao'an and Longhua districts show emerging hotspots in taxi trip volumes (Fig. 6c). The peripheral districts of Pingshan, Dapeng, and Yantian exhibit relatively lower values across most indicators, with more scattered distribution patterns.

The land use mix index (Fig. 6d) shows widespread medium to high values across the city, with particularly concentrated high-value areas forming in Bao'an, Longhua, and Longgang districts, indicating high land function diversity in these regions. Guangming and northern Nanshan also display relatively high land use mixture, while eastern Dapeng and Yantian areas show comparatively lower values. PM2.5 concentration (Fig. 6e) displays a clear northwest-southeast gradient, with the highest concentrations in the northwestern districts of Bao'an, Guangming, and Longhua; while the eastern and southeastern districts exhibit lower PM2.5 levels. This distribution pattern reflects the higher industrial activity and population density in the northwestern urban areas, whereas the southeastern region maintains better environmental quality, demonstrating the spatial relationship between urban development and environmental conditions.

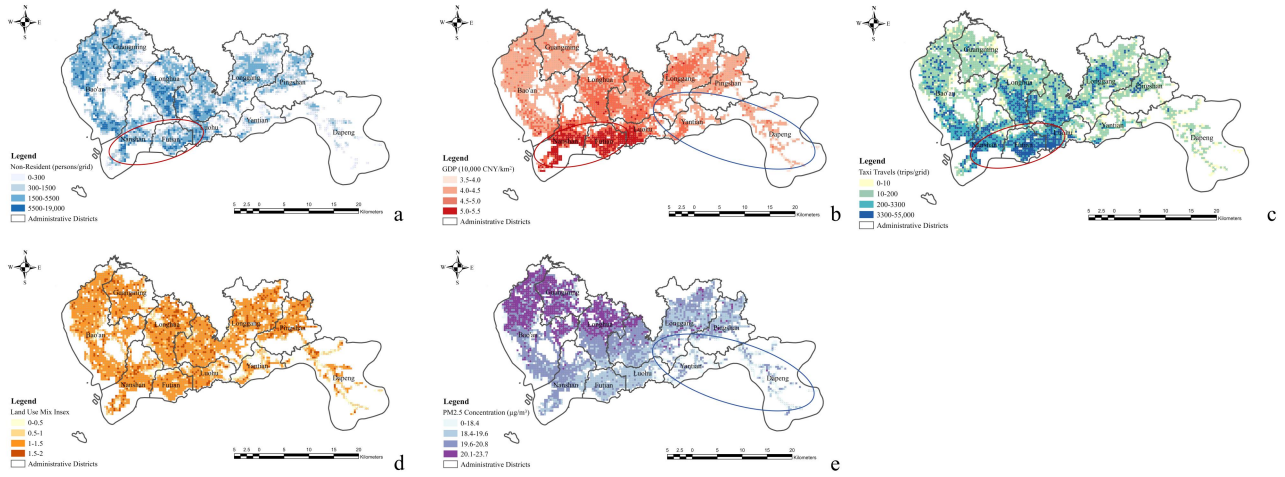

Figure 6. Spatial distribution of urban attributes.

Beyond individual patterns, significant spatial correlations exist between different urban attributes. High GDP areas (Fig. 6b) strongly correlate with high taxi trip volumes (Fig. 6c), particularly evident in Nanshan and Futian districts. This relationship stems from economic prosperity generating increased mobility demand, while enhanced transportation accessibility simultaneously attracts business investment - creating a positive feedback loop where economic activities stimulate transportation needs and vice versa. Similarly, the co-location of non-resident population (Fig. 6a) and economic activities (Fig. 6b) demonstrates how commercial centers naturally attract temporary populations for employment, shopping, and services, which in turn support and amplify economic vitality in these zones.

Comparing land-use mix index (Fig. 6d) with PM2.5 concentration (Fig. 6e) reveals a positive correlation. Areas with higher land-use diversity exhibit elevated PM2.5 levels, while homogeneous zones show lower pollution. Northern Bao'an and Guangming districts exemplify this pattern, with high land-use mix indices aligning with significantly higher PM2.5 levels. Similarly, the eastern districts (Yantian and Dapeng) display lower economic indicator values (Fig. 6b) but significantly better environmental metrics (Fig. 6e). These patterns demonstrate that urban development intensity and environmental quality exhibit clear spatial trade-offs across Shenzhen's urban fabric.

5. Conclusion and Discussion

The unprecedented multi-source urban data (including remote sensing images, mobile phone data, and other urban datasets) provides an alternative approach to mapping urban attributes. This study presents a federated learning-based urban attribute mapping framework that resolves issues of data isolation and sharing. It shares the encrypted handcrafted features derived from private data sourced from multi-source dataholders. It predicts diverse urban attributes such as population, economic development, urban mobility, land use, and air quality using popular machine learning techniques. The experiment in Shenzhen demonstrates that the framework successfully enables multi-source collaborative mapping by sharing only model parameters and structures, rather than raw urban data. The framework demonstrates exceptional performance in five urban attribute mapping tasks: non-resident population, GDP, taxi travels, land use mix, and PM2.5 concentration. The study provides an alternative solution to overcome urban data silos and lays the groundwork for expanding urban mapping and cross-regional applications.

However, several limitations should be acknowledged. The presented federated learning approach introduces additional computational and communication overhead through encrypted gradient exchanges, which may constrain scalability in resource-limited environments. The framework's performance under varying federation scenarios—such as different numbers of participating data holders or imbalanced data distributions—remains to be thoroughly evaluated. Additionally, the current validation is based on Shenzhen's data-rich urban environment, and the framework's adaptability to cities with sparser data sources or less reliable connectivity requires further investigation. Future work should focus on optimizing communication efficiency to reduce network overhead, developing robust aggregation mechanisms to handle data imbalance across participants, and extending the framework's applicability to diverse urban contexts with varying data availability. Additionally, integrating deep neural networks could enable the development of a universal urban foundation model.

Acknowledgement

This study is supported and funded by Shenzhen Science and Technology Program (JCYJ20220818100200001); Science and Technology R&D Program of Shenzhen, China (20220810135520002); the National Natural Science Foundation of China (No. 42471496; 42101472); The Center for Scientific Research and Development in Higher Education Institutes, MOE (2024HT013); The Innovation Team of the Department of Education of Guangdong Province (2024KCXTD013).

References

Bai, L., Huang, W., Zhang, X., Du, S., Cong, G., Wang, H., Liu, B., 2023. Geographic mapping with unsupervised multi-modal representation learning from VHR images and POIs. *ISPRS Journal of Photogrammetry and Remote Sensing*, 201, 193-208. doi.org/10.1016/j.isprsjprs.2023.05.006.

Brauer, A., Mäkinen, V., Ruotsalainen, L., Oksanen, J., 2024. Time will not tell: Temporal approaches for privacy-preserving trajectory publishing. *Computers, Environment and Urban*

Systems, 112, 102154. doi.org/10.1016/j.compenvurbsys.2024.102154.

Cao, J., Cao, X., Tu, W., Tan, X., Wang, T., Chen, G., Zhang, X., Li, Q., 2025. Nighttime light imagery or mobile phone footprints: Which better reflects urban socio-economics at the grid level? A case study in the Pearl River Delta, China. Computers, Environment and Urban Systems, 116, 102220. doi.org/10.1016/j.compenvurbsys.2024.102220.

Cao, J., Wang, X., Chen, J., Tu, W., Li, Z., Yang, X., Zhang, X., Li, Q., 2025. Urban representation learning for fine-grained economic mapping: A semi-supervised graph-based approach. ISPRS Journal of Photogrammetry and Remote Sensing, 226, 317-331. doi.org/10.1016/j.isprsjprs.2025.05.007.

Chen, G., Su, Y., Zhang, X., Hu, A., Chen, G., Feng, S., Xiang, J., Zhang, J., Zheng, Y., 2022. A Cross-City Federated Transfer Learning Framework: A Case Study on Urban Region Profiling. *arXiv:2206.00007*.

Cheng, K., Fan, T., Jin, Y., Liu, Y., Chen, T., Papadopoulos, D., Yang, Q., 2021. Secureboost: A lossless federated learning framework. *IEEE intelligent systems*, 36(6), 87-98. doi.org/10.1109/MIS.2021.3082561.

Danish, M., Labib, S. M., Ricker, B., Helbich, M., 2025. A citizen science toolkit to collect human perceptions of urban environments using open street view images. *Computers, Environment and Urban Systems*, 116, 102207. doi.org/10.1016/j.compenvurbsys.2024.102207.

Fan, M., Zhang, Z., Li, Z., Sun, G., Yu, H., Kang, J., Guizani, M., 2024. SecureVFL: privacy-preserving multi-party vertical federated learning based on blockchain and RSS. *Digital Communications and Networks*. doi.org/10.1016/j.dcan.2024.07.008.

Fan, Z., Zhang, F., Loo, B. P., Ratti, C., 2023. Urban visual intelligence: Uncovering hidden city profiles with street view images. *Proceedings of the National Academy of Sciences*, 120(27), e2220417120. doi.org/10.1073/pnas.222041712.

Fang, B., Li, M., Huang, Z., Yue, Y., Tu, W., Guo, R., 2024. Revealing multi-scale spatial synergy of mega-city region from a human mobility perspective. *Geo-Spatial Information Science*, 1-16. doi.org/10.1080/10095020.2024.2379060.

Feng, J., Du, C., Mu, Q., 2024. Traffic Flow Prediction Based on Federated Learning and Spatio-Temporal Graph Neural Networks. *ISPRS International Journal of Geo-Information*, 13(6), 210. doi.org/10.3390/ijgi13060210.

Gao, Y., Yan, D., Liu, N., Luo, S., Zhou, Z., Wang, Y., Gao, B., 2022, November. A High-Frequency Monitoring Method for Urban Carbon Emissions Based on Vertical Federated Deep Learning and Multi-Source Heterogeneous Data Fusion Methods. *In Proceedings of the 5th International Conference on Information Technologies and Electrical Engineering* (pp. 232-237). doi.org/10.1145/3582935.3582972.

Huang, Q., Cui, H., Xiang, L., 2025. Classification of urban road functional structure by integrating physical and behavioral features. *ISPRS Journal of Photogrammetry and Remote Sensing*, 220, 753-769. doi.org/10.1016/j.isprsjprs.2025.01.018.

Kairouz, P., McMahan, H. B., Avent, B., Bellet, A., Bennis, M., Bhagoji, A. N., Zhao, S., 2021. Advances and open problems in federated learning. *Foundations and trends® in machine learning*, 14(1–2), 1-210. doi.org/10.1561/220000008.

Li, Y., Huang, W., Cong, G., Wang, H., Wang, Z., 2023, August. Urban region representation learning with openstreetmap building footprints. *In Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining* (pp. 1363-1373). doi.org/10.1145/3580305.35995.

Liu, Y., Fan, T., Chen, T., Xu, Q., Yang, Q., 2021. Fate: An industrial grade platform for collaborative learning with data protection. *Journal of Machine Learning Research*, 22(226), 1-6.

Meadows, C., 1986. A more efficient cryptographic matchmaking protocol for use in the absence of a continuously available third party. *In 1986 IEEE Symposium on Security and Privacy* (pp. 134-134). doi.ieeecomputersociety.org/10.1109/SP.1986.10022.

Pandya, S., Srivastava, G., Jhaveri, R., Babu, M. R., Bhattacharya, S., Maddikunta, P. K. R., Gadekallu, T. R., 2023. Federated learning for smart cities: A comprehensive survey. *Sustainable Energy Technologies and Assessments*, 55, 102987. doi.org/10.1016/j.seta.2022.102987.

Putra, K. T., Chen, H. C., Prayitno, Ogiela, M. R., Chou, C. L., Weng, C. E., Shae, Z. Y., 2021. Federated compressed learning edge computing framework with ensuring data privacy for PM2.5 prediction in smart city sensing applications. *Sensors*, 21(13), 4586. doi.org/10.3390/s21134586.

Salehi, B., Gu, J., Roy, D., Chowdhury, K., 2022, May. Flash: Federated learning for automated selection of high-band mmwave sectors. *In IEEE INFOCOM 2022-IEEE Conference on Computer Communications* (pp. 1719-1728). doi.org/10.1109/INFOCOM48880.2022.979686.

Sun, K., Hu, Y., Ma, Y., Zhou, R. Z., Zhu, Y., 2023. Conflating point of interest (POI) data: A systematic review of matching methods. *Computers, Environment and Urban Systems*, 103, 101977. doi.org/10.1016/j.compenvurbsys.2023.101977.

Wang, Q., Wang, S., Zheng, Y., Lin, H., Zhang, X., Zhao, J., Walker, J., 2024. Deep hybrid model with satellite imagery: How to combine demand modeling and computer vision for travel behavior analysis?. *Transportation Research Part B: Methodological*, 179, 102869. doi.org/10.1016/j.trb.2023.102869.

Wu, M., Huang, Q., Gao, S., Zhang, Z., 2023. Mixed land use measurement and mapping with street view images and spatial context-aware prompts via zero-shot multimodal learning. *International Journal of Applied Earth Observation and Geoinformation*, 125, 103591. doi.org/10.1016/j.jag.2023.103591.

Zhao, X., Xia, N., Xu, Y., Huang, X., Li, M., 2021. Mapping population distribution based on XGBoost using multisource data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 14, 11567-11580. doi.org/10.1109/JSTARS.2021.3125197.