

Article

# Large Language Model-Derived Spatial Embedding for Revealing Electric Vehicle Energy Consumption: An Urban–Rural Case

Can Yin <sup>1</sup>, Lifu Jin <sup>1</sup> and Zhe Zhang <sup>2,\*</sup> 

<sup>1</sup> School of Management, Jiangsu University, Zhenjiang 212013, China; yincan@stmail.uj.edu.cn (C.Y.); jszjlf@uj.edu.cn (L.J.)

<sup>2</sup> School of Vehicle and Mobility, Tsinghua University, Beijing 100084, China

\* Correspondence: zhezhang@mail.tsinghua.edu.cn

## Abstract

Understanding spatial variation in electric vehicle energy consumption (EC) can support more targeted urban transport analysis, but observational data do not identify the causal effects of urban design. This study analyzes approximately 1700 private EVs operating in Shanghai from October to December 2020, integrating about 29 million high-resolution GPS records with multi-source built-environment data. Segment-level EC was aggregated to 500 m grids and communities. Community functional zones were derived by encoding POI sequences with Qwen2.5-1.5B-Instruct and clustering the resulting embeddings, while Random Forest models and SHAP were used to characterize model-attributed associations with spatial EC variation. At the grid level, mean speed and speed variability received approximately 84% of the native Random Forest importance, whereas the 5D-inspired indicators received the remaining 16%, led by destination accessibility, road density, and bus-stop density. At the community level, the Urban core zone had approximately 9% higher fleet-weighted EC than the Rural zone, while within-zone importance rankings differed between dense and sparse zones. These results describe conditional, location-level associations for the sampled fleet and season.

**Keywords:** electric vehicle; energy consumption; urban functional zones; large language model



Academic Editor: Wolfgang Kainz

Received: 16 June 2026

Revised: 18 August 2026

Accepted: 14 September 2026

Published: 15 September 2026

**Copyright:** © 2026 by the authors.

Published by MDPI on behalf of the International Society for

Photogrammetry and Remote Sensing.

Licensee MDPI, Basel, Switzerland.

This article is an open access article

distributed under the terms and

conditions of the [Creative Commons](#)

[Attribution \(CC BY\)](#) license.

## 1. Introduction

The decarbonization of urban transport increasingly relies on the diffusion of electric vehicles (EVs), with global electric-car sales exceeding 17 million in 2024 and China selling over 11 million electric cars, nearly two-thirds of the global total [1]. Shanghai provides a suitable empirical setting because it had registered 1.288 million new energy vehicles by the end of 2023 [2]. While EVs eliminate tailpipe emissions, their net climate benefit depends on the per-kilometer electricity demand, hereafter the EC, and the carbon intensity of the supplying grid. Accordingly, identifying where per-kilometer EC is systematically higher or lower can provide descriptive evidence for urban transport planning. The present study estimates spatial associations and model-attributed contributions in the observed data; it does not estimate the causal effect of changing urban form on EC.

A large body of literature has linked the urban built environment, typically captured by the 5D framework of density, diversity, design, destination accessibility and distance to transit, to vehicle kilometers traveled, modal share and emissions [3,4]. Empirical evidence

for EVs in particular, however, remains comparatively sparse. Existing EV energy studies have focused mainly on driving-cycle modeling, battery-state estimation or charging-load forecasting [5–7], while the spatial attribution of real-world EV energy use to the urban built environment remains under-explored. Two gaps stand out. The first concerns the identification of functional zones, where earlier work has largely relied on coarse categorical labels derived from administrative boundaries or shallow POI heuristics. Recent neural-embedding approaches such as Doc2Vec and Place2Vec capture richer semantics [8–11], but they were not designed for context-aware language models, and whether recent large language models pre-trained on web-scale Chinese text [12] can produce more discriminative functional zone embeddings remains an open question. The second gap concerns the lack of multi-scale evidence on EC. Most studies adopt a single spatial level, either a regular grid or an administrative boundary, which conflates fine-grained street-level signals with neighborhood-level processes [13,14].

More importantly, the per-kilometer energy demand of an EV is governed first and foremost by the speed regime under which it is driven, as demonstrated by sensitivity analyses of physics-based EV energy models [15,16]. However, these models are typically estimated under controlled conditions and do not directly reveal how the spatial configuration of the built environment is associated with real-world energy use across different urban contexts. In view of this, recent studies have increasingly adopted explainable machine learning frameworks such as SHAP (Shapley Additive exPlanations), which provide a unified framework for attributing model predictions to individual features [17,18]. Yet the application of such interpretability tools to multi-scale EV energy attribution in observational spatial data remains limited. This distinction is also consistent with EV emissions studies showing that use-phase CO<sub>2</sub> estimates depend strongly on local or marginal electricity-emission factors and on operating conditions such as temperature [19,20].

To better position our contribution within the broader literature, Table 1 summarizes five relevant streams of research and their remaining gaps, namely, (i) physics-based EV energy models; (ii) built-environment and travel-demand studies; (iii) urban functional-zone detection from POI and street networks; (iv) explainable ML and SHAP in transport energy; and (v) spatial econometric approaches.

To address these gaps, this study assembles a Shanghai-wide dataset that integrates real-world GPS records from approximately 1700 private EVs over a three-month period with about 920,000 cleaned POIs, the OpenStreetMap road network and the administrative boundaries. A multi-scale analytical design is adopted, in which a 500 m grid captures fine-grained built-environment variation and 7255 communities capture neighborhood-scale functional context. Random Forest models [17] with SHAP attribution [18] are estimated at each level. For the community-level analysis, a Qwen2.5-1.5B-Instruct embedding approach produces continuous functional zone representations from road-based POI sequences. The contributions of this study are threefold. First, a reproducible methodology is presented for transforming road-based POI sequences into large-language-model-based functional zone embeddings, which are then compared with discrete K-Means clusters. Second, multi-scale empirical evidence is provided on how the 5D built environment and the LLM-derived functional zones in Shanghai jointly explain the spatial variation in per-kilometer EC and the corresponding CO<sub>2</sub> emissions. The GIS and LLM components are used to characterize the spatial context of EC, not to estimate the grid carbon factor. Once EC has been estimated, the CO<sub>2</sub> value is obtained by a separate accounting conversion. Thus, the spatial modeling contribution concerns where observed per-kilometer EC differs and which included features the fitted model associates with that variation; it does not create an independent CO<sub>2</sub> model. Third, within the fitted Random Forest and the included feature set, the native feature-importance measure assigns approximately 84% of total importance

to the two speed descriptors and approximately 16% to the 5D-inspired built-environment indicators. These values are model-specific attribution shares rather than causal effects or proportions of energy use explained by urban form. This study therefore focuses on the conditional spatial association between the observed built environment and EC after accounting for the available speed descriptors.

**Table 1.** Structured literature table.

| Literature Stream                                           | Representative Studies | Typical Estimand/Objective                                                                                    | Typical Data                                                                            | Core Methods                                                                                                 | Core Gaps/Unresolved Issues                                                                                                                                                                                                                                                             |
|-------------------------------------------------------------|------------------------|---------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| (i) Physics-based EV energy models                          | [7,15,16]              | Instantaneous/segmental energy consumption (kWh/km) based on vehicle dynamics and powertrain efficiency       | On-board diagnostics (OBD), chassis dynamometer, or controlled driving cycles           | Physical equations (tractive force, regenerative braking, rolling resistance, aerodynamics)                  | Strong in vehicle physics but largely ignore spatial heterogeneity; rarely link macroscopic urban form to real-world energy use; not designed for city-wide spatial attribution.                                                                                                        |
| (ii) Built environment and travel demand                    | [3,4]                  | Vehicle kilometers traveled (VKT), modal split, or aggregate CO <sub>2</sub> emissions                        | Travel surveys, census, land use, POI, road networks                                    | Multivariate regression, structural equation modeling, discrete choice                                       | Usually focus on travel quantity (VKT) rather than per-km energy intensity; rarely validated with large-scale EV telemetry; often ignore within-city heterogeneity.                                                                                                                     |
| (iii) Urban functional-zone detection (POI/street networks) | [9–11]                 | Land-use/functional-zone classification (e.g., commercial, residential, mixed-use)                            | POI, street networks, remote sensing                                                    | Word2Vec, Doc2Vec, Place2Vec, or graph-based embeddings + clustering                                         | Mostly used for urban planning or land-use mapping; rarely linked to transport energy outcomes; conventional embeddings (Word2Vec/Doc2Vec) may not capture complex POI sequence semantics; recent LLMs remain underexplored.                                                            |
| (iv) Explainable ML and SHAP in transport/energy            | [17,18]                | Model interpretation/feature attribution (e.g., what drives predicted emissions or travel time)               | Traffic flow, weather, built environment, sensor data                                   | Random Forest, XGBoost, SHAP, permutation importance                                                         | SHAP attributions are often model-specific and correlational; rarely combined with LLM-derived spatial context; few studies apply it to per-km EV energy consumption at multiple spatial scales simultaneously.                                                                         |
| (v) Spatial econometric/Geographically weighted models      | [21,22]                | Spatial dependence and spatially varying relationships (e.g., land use—transport interactions)                | Zonal or point-based socio-economic, land-use, and transport data                       | Spatial lag/error models, Geographically Weighted Regression (GWR)                                           | Account for spatial autocorrelation but often rely on linearity assumptions; less suited for high-dimensional, non-linear interactions; typically, do not incorporate deep semantic embeddings of urban functions.                                                                      |
| This study                                                  |                        | Conditional spatial association between urban form/functional zones and observed per-km EV energy consumption | ~1700 private EVs, ~29 M GPS records, 920k POIs, OSM road network, community boundaries | Qwen2.5-1.5B-Instruct embeddings, K-Means clustering, Random Forest + SHAP, multi-scale (grid and community) | Directly links LLM-derived functional zone embeddings and 5D-inspired built-environment indicators to real-world EV telemetry; quantifies model-attributed importance conditional on observed speed descriptors; reveals zone-specific heterogeneity in built-environment associations. |

## 2. Materials and Methods

### 2.1. Overview Framework

Figure 1 presents the overall framework of this study. Four types of input data are used in this study, including the EV GPS records, the POI data, the OpenStreetMap road network, and the administrative boundaries. The GPS records are first processed into segment-level EC, and then aggregated to two parallel spatial scales, namely a 500 m grid for the built-environment analysis and a community cluster for the LLM-based functional zone analysis. The 5D built-environment indicators, the LLM-derived functional zone label and the segment-speed descriptors serve as the input features of the Random Forest models. The trained models are then interpreted using SHAP to describe how the fitted predictions allocate importance across the included built-environment and speed features. These attributions are conditional associations within the model and are not interpreted as causal effects.

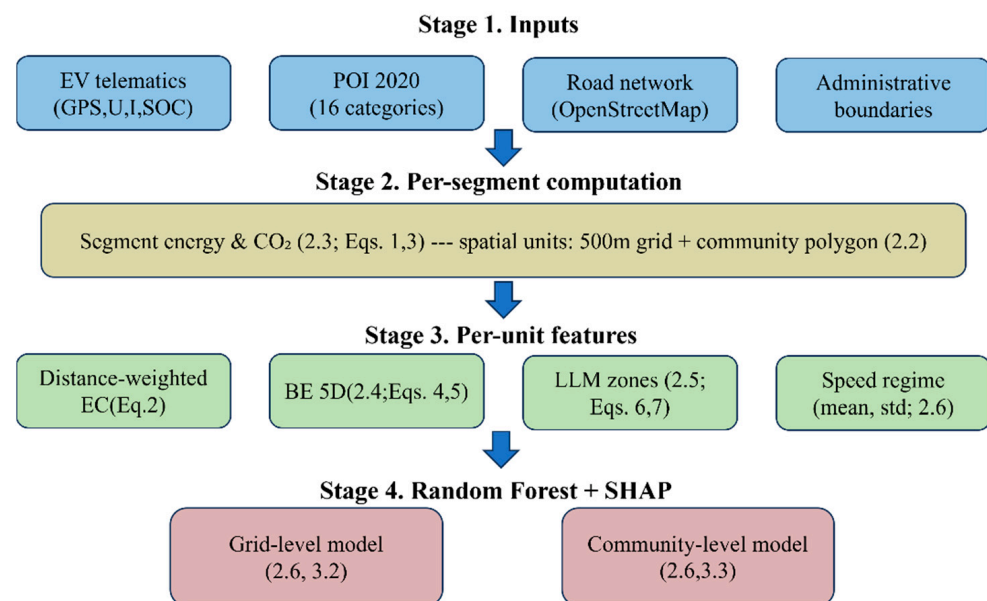


**Figure 1.** Overview framework of this study.

### 2.2. Study Area and Data

The study area is Shanghai, which comprises 17 administrative districts and approximately 8240 km<sup>2</sup> of land. This area is partitioned into 7255 communities, according to Chinese neighborhood committees and villages. The EV GPS records cover 1994 unique private electric vehicles operating in Shanghai between October and December 2020. All vehicles included in the dataset are battery electric sedans. The available vehicle records include several battery capacity specifications, including 35 kWh, 50.8 kWh and 55.5 kWh. Each record reports a timestamp, vehicle status, charge status, instantaneous speed, total voltage, total current, state of charge, and longitude and latitude. Figure 2 shows a sample of the analytical GPS points over the Shanghai district boundaries. The coverage is dense across the inner urban core and remains substantial across the outer ring.

The POI data comprise 919,599 POIs in 16 functional super-categories derived from the 2020 Amap POI dataset. During preprocessing, seven source categories were excluded because they mainly contained duplicate or redundant records, or POIs that occurred only during specific time periods, which could introduce redundancy or temporal inconsistency into the functional characterization. Representative excluded categories include place names and addresses, road furniture, incidents and events, and pass facilities. The remaining POIs were organized following the classification scheme adopted in previous urban functional-zone studies [9,10]. The OpenStreetMap road network of 114,170 road segments

and 12,753 bus stops provide a transit-density indicator. The administrative boundaries are obtained from open Chinese administrative datasets.

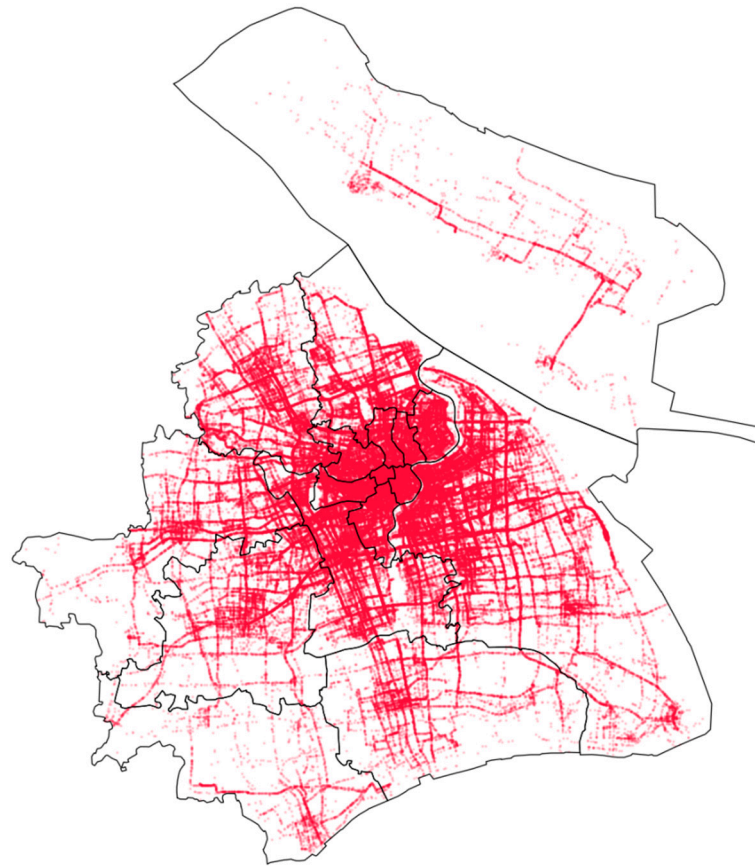


**Figure 2.** EV GPS points in Shanghai.

Two parallel spatial levels are adopted in this study. The first is the grid level, in which a regular 500 m grid over Shanghai produces 32,946 grid cells. Both square and hexagonal grids are commonly used as spatial aggregation units in built-environment analysis. In this study, we adopted a square grid because it allows cells of equal area to be generated straightforwardly within a regular coordinate-based framework and enables the spatial uniformity of the resulting grid to be visually inspected. Only the 20,035 cells (60.8%) that contain at least one analytical EV segment are retained for the analysis. The distance-weighted EC at each grid, defined as the sum of segment energies divided by the sum of segment distances, becomes the target of the grid-level model. Because many of these cells contain only a few segments, we further restrict the grid-level model to selected cells with at least 1000 analytical segments and at least 10 unique vehicles, which yields 3741 grids that together cover the majority of fleet activity. These thresholds are practical data quality control criteria rather than statistically optimized cutoff values, and sensitivity to alternative segment or vehicle-count thresholds was not tested. The second level is the community level, in which the 7255 communities provide a neighborhood scale. After filtering to mitigate the same kind of small-sample noise as the grid-level filter does, we obtain 4138 communities.

### 2.3. Segment-Level EC and CO<sub>2</sub> Accounting

The GPS records in the dataset have four charge statuses, labeled 1 through 4. A value of 1 indicates parking charging, 2 indicates charging while driving, 3 indicates a non-charging state, and 4 indicates that charging has been completed. GPS records are

retained when the charge status was 3 and the instantaneous speed is positive. A driving segment is defined as the interval between two consecutive retained records. Segments are excluded if their time gap falls outside 0 to 600 s, if their distance is below 50 m (attributable to GPS positioning noise), or if their distance exceeds 50 km (attributable to anomalous position jumps). Across the fleet, 29,375,856 raw segments were produced. Restricting to in-Shanghai segments belonging to usable vehicles yields 15,207,505 analytical segments, representing 7.8 million driven kilometers, 87.05 GWh of electricity and 47,704 tons of carbon dioxide. Table 2 summarizes the sample attrition process at each cleaning step. The majority of record loss occurred at the initial charge-status and speed filters, which removed stationary or charging records; subsequent filters removed segments with unrealistic time gaps or anomalous distances. The final analytical sample comprises 1706 vehicles and 15,207,505 segments, retaining approximately 51.77% of raw records after spatial restriction to Shanghai.

**Table 2.** Sample attrition across data cleaning and filtering steps.

| Filtering Step                                                 | Vehicles | Analytical Segments | Distance (km) | Energy (GWh) | Percent Retained (%) |
|----------------------------------------------------------------|----------|---------------------|---------------|--------------|----------------------|
| (1) Raw data                                                   | 1994     | 29,375,856          | 8,769,902     | 87.23        | 100                  |
| (2) After charge status = 3 and speed > 0                      | 1994     | 17,625,514          | 8,662,743     | 87.66        | 60.16                |
| (3) After segments time gap $\in [0, 600 \text{ s}]$           | 1974     | 17,184,926          | 8,451,286     | 87.45        | 58.50                |
| (4) After segment distance $\in [50 \text{ m}, 50 \text{ km}]$ | 1974     | 16,438,271          | 8,164,927     | 87.16        | 55.96                |
| (5) After filtering to in-Shanghai segments                    | 1706     | 15,207,505          | 7,800,845     | 87.05        | 51.77                |

Segment energy is computed by trapezoidal integration, as shown in Equation (1).

$$E_{i \rightarrow i+1} = \frac{\frac{1}{2}(U_i I_i + U_{i+1} I_{i+1}) \Delta t_{i \rightarrow i+1}}{3.6 \times 10^6} \quad (1)$$

where E is in kWh, U and I are the pack-level total voltage (V) and total current (A), respectively, and  $\Delta t$  (s) is the interval between consecutive records in seconds. Here, positive current denotes battery discharge and therefore produces positive segment EC, whereas the negative current denotes energy returned to the battery during regenerative braking and produces negative segment energy. Negative regenerative energy was retained when segment and spatial EC were aggregated.

The EC per segment is the ratio of segment energy to segment distance (kWh/km), and the EC per cell or per community used as the target of the spatial models is the distance-weighted aggregate defined in Equation (2).

$$\overline{EC} = \frac{\sum E_{seg}}{\sum d_{seg}} \quad (2)$$

Here, the sums run over all in-Shanghai analytical segments that fall inside a given grid cell or community. Segment distance  $d$  is computed using the haversine formula, and the segment carbon dioxide emission is obtained by multiplying the segment energy by the Shanghai 2020 electricity grid emission factor EF, as shown in Equation (3). This emission factor is sourced from the China Regional Power Grid Carbon Dioxide Emission Factors

2023 report published by the China Electricity Council. Although the report was published in 2023, it provides regional power grid emission factors for 2020.

$$CO_{2,i \rightarrow i+1} = E_{i \rightarrow i+1} \times EF \quad (3)$$

Because the emission factor  $EF$  is a spatial constant in this study, the per-kilometer  $CO_2$  emission at any grid cell or community is a scalar multiple of the per-kilometer EC at the same level. The grid-level and community-level Random Forest analyses in Section 3 therefore use EC as the single modeling target throughout, with the understanding that all SHAP attributions, urbanization gradients and zone-specific contributions reported below for EC translate directly into per-kilometer  $CO_2$  emissions by multiplying the EC magnitudes by  $EF = 0.548 \text{ kg } CO_2/kWh$ . The annual average emission factor yields approximate  $CO_2$  accounting estimates. Its constant application rescales totals but leaves the EC model, SHAP rankings, and spatial patterns unchanged. A time-varying emission factor could alter the magnitude and spatiotemporal distribution of the  $CO_2$  estimates if grid carbon intensity varies with travel time. However, temporally resolved emission factor data for the regional power grid during the study period were unavailable. We therefore used the annual average emission factor and acknowledged this limitation in the revised manuscript.

#### 2.4. Built-Environment Indicators Guided by the 5D Framework

In previous related studies, one or two indicators are computed per dimension of the 5D framework [4]. In this study, six observable indicators are constructed for each grid cell to operationalize the 5D framework. These indicators should not be regarded as comprehensive measures of all 5D dimensions. Specifically, intersection density is subject to the topological characteristics of OpenStreetMap data; bus stop density serves as an indirect indicator for transit accessibility; metro proximity is not included in the analysis; and destination accessibility is computed based on the Euclidean distance instead of the network distance. As shown in Table 3, the 5D dimensions can be related to per-kilometer EC through competing pathways, so a single directional hypothesis is not imposed for every indicator. A higher activity concentration and network friction may be associated with lower and more variable speeds, repeated acceleration, and greater auxiliary-energy use per kilometer. Conversely, compact and connected environments may shorten detours and support more continuous routing. Mean speed and speed variability may therefore lie partly on the pathway between the built environment and EC. Because the combined Random Forest conditions on these speed descriptors, its built-environment attributions are interpreted as direct conditional associations net of the observed speed summaries, rather than as total built-environment associations.

Density is represented by the POI density in counts per  $km^2$  and by the road length density in meters per  $km^2$ . Diversity is represented by the Shannon entropy over the 16 POI super-categories, defined in Equation (4),

$$H = -\sum_{i=1}^{16} p_i \log p_i \quad (4)$$

where  $p_i$  denotes the share of POIs in super-category  $i$  within the cell and the sum runs over the 16 retained POI super-categories. Design is represented by the road intersection density. Before calculating this indicator, the OpenStreetMap road network was topologically processed to identify shared road nodes and consolidate duplicate intersection records. Intersections were then defined as nodes of degree 3 or higher in the processed road-segment endpoint graph, and intersection density was calculated as the number of

identified intersections per km<sup>2</sup>. This preprocessing produced the intersection layer used in the subsequent Random Forest models.

**Table 3.** Conceptual pathways linking 5D-inspired indicators to per-kilometer EC.

| 5D-Inspired Indicator     | Competing Pathways to EC                                                                                                          | Role of Speed in Interpretation                                                               |
|---------------------------|-----------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------|
| POI density               | Higher EC: curbside friction, congestion, repeated acceleration; Lower EC: shorter local access and fewer detours.                | Possible mediator; attribution is conditional on mean speed and speed variability.            |
| Road length density       | Higher EC: more junctions and turning events; Lower EC: more route options and shorter circulation.                               | May summarize network operating conditions through local speed patterns.                      |
| POI entropy               | Higher EC: more local traffic interaction; Lower EC: multi-purpose trips and shorter trip chains.                                 | Interpreted as contextual association net of observed speed summaries.                        |
| Intersection density      | Higher EC: stopping and acceleration; Lower EC: greater connectivity and reduced detours. Measurement is limited by OSM topology. | Speed may mediate junction-related delay; attribution should be read cautiously.              |
| Destination accessibility | Higher EC: centrality, congestion, repeated acceleration; Lower EC: nearby destinations and shorter travel distances.             | Speed may lie on the accessibility-EC pathway; estimates are not total accessibility effects. |
| Bus-stop density          | Higher EC: indicator for dense central activity; Lower EC: indicator for transit-rich, less car-oriented areas.                   | Contextual indicator, not a direct transit-distance or causal transit effect.                 |

Note: The table summarizes competing pathways, not deterministic hypotheses. Because speed may be both physically related to EC and partly shaped by the built environment, model attributions are interpreted as speed-conditioned associations rather than total or causal effects.

Destination accessibility is measured as the distance-weighted count of POIs within 2 km of each grid cell center. Each POI is assigned an exponential weight of  $\exp\left(\frac{-d_{ij}}{1000}\right)$  as defined in Equation (5), where  $d_{ij}$  is its Euclidean distance from the cell center. Consequently, nearby POIs contribute more strongly than those farther away. Distance to a central business district is not used because Shanghai has a polycentric urban structure, making the selection of a single center arbitrary. Transit accessibility is represented by bus stop density within 2 km of each cell center.

$$A_i = \sum_{j:d_{ij} \leq 2000} \exp\left(\frac{-d_{ij}}{1000}\right) \quad (5)$$

## 2.5. Functional Zones via LLM Embedding and K-Means

Functional zones are derived through a four-step procedure that extends the POI-Doc2Vec approach proposed for urban functional zone identification [9,10] by replacing Doc2Vec with a recently developed decoder-only large language model [12]. In the first step, sampling points are interpolated every 50 m along all OpenStreetMap road

segments, and a 50 m grid with a 25 m offset pass deduplicates them to 497,912 sampling points. In the second step, a POI sequence is constructed for each sampling point by retrieving the POIs within a 200 m radius using a KD-tree spatial index, sorting them by Euclidean distance, and producing an ordered sequence of POI tokens in the form of a major category, mid category and minor category. Sequences with fewer than four tokens are excluded, leaving 238,272 sequences, of which 99.9% fall inside a community and which together cover 6626 out of 7255 communities. In the third step, the sequences are tokenized and encoded by Qwen2.5-1.5B-Instruct [12]. Each tokenized sequence of length  $T$  yields a stack of last-hidden-layer states  $\{h_1, \dots, h^T\}$  of dimension 1536, and mean pooling is adopted because POI sequences have no natural CLS token and variable lengths, making the equal-weight aggregation of all surrounding POI tokens a natural fit for capturing collective functional semantics. Max pooling tends to discard less frequent but functionally relevant categories, while attentive pooling introduces unnecessary complexity for this unsupervised task without guaranteeing stability. In preliminary experiments on 500 manually annotated communities, mean pooling achieved the highest clustering agreement with manual labels and the lowest variance across random seeds, outperforming both max pooling and a simplified attentive variant. Hence, we retain mean pooling as a computationally efficient and empirically robust choice for generating sequence-level embeddings. The sequence-level vector is then obtained by mean pooling as in Equation (6).

$$v_{seq} = \frac{1}{T} \sum_{t=1}^T h_t \quad (6)$$

The per-community vector is then obtained by concatenating the element-wise mean and standard deviation of all sequence vectors associated with that community, as in Equation (7), where  $S$  denotes the set of sequence vectors in the community and the superscript  $s$  indexes individual sequences. This doubles the dimensionality to 3072 and lets the clustering step exploit both the average semantic content of a community and its internal variation.

$$v_{com} = \left[ \text{mean}_{s \in S} \left( v_{seq}^s \right) ; \text{std}_{s \in S} \left( v_{seq}^s \right) \right] \quad (7)$$

In the fourth step, communities with fewer than three sampling points are excluded to avoid degenerate standard-deviation components, leaving 6330 communities. There is a multi-criteria selection of  $K$ , based on the silhouette score, the Calinski–Harabasz index, the Davies–Bouldin index and the adjusted Rand index across five random seeds over the range  $K \in [2, 15]$ . Nevertheless, qualitative inspection shows that the  $K = 3$  solution merges the urban core, dense residential areas and inner-suburban zones into one large cluster, masking intra-urban heterogeneity in built-environment intensity and road-network configuration critical to our energy-consumption analysis. This three-cluster scheme mixes functionally distinct neighborhoods with clearly different driving-speed and stop-and-go patterns. For subsequent descriptive comparisons, we adopt the finer-grained  $K = 5$ . It produces five spatially interpretable community archetypes and maintains high stability across random seeds (adjusted Rand index = 0.92 for five seeds).  $K = 5$  is therefore not the statistical optimum but a deliberate trade-off between within-cluster compactness and functional differentiation. The normalized mutual information between  $K = 5$  cluster labels and an external binary urban–rural reference labels is 0.168, suggesting the embeddings partly capture urban–rural divides while encoding richer functional semantics.

## 2.6. Random Forest Models with SHAP

Random Forest [17] regression models are fit at three levels that mirror the multi-scale design described above.

The grid-level model takes the six 5D built-environment indicators plus two speed descriptors of the local driving regime as inputs and the per-grid distance-weighted EC as the target. The speed descriptors are the mean and the standard deviation of the segment-level average speed within each grid, computed from the same in-Shanghai analytical segments that define the target. They are not built-environment indicators but rather summary statistics of the observed driving behavior in each grid, included so that the model can separate the contribution of the built environment from the contribution of the driving speed regime. The grid-level model is fit on the 3741 well-sampled grids, the optimal combination of hyperparameters selected through grid search.

The community-level zone-aggregated model is fit on the 4138 well-sampled communities. It takes the  $K = 5$  LLM-derived functional zone one-hot label together with the mean and the standard deviation of the segment-level average speed within each community, giving seven features, and predicts the per-community EC using this stated Random Forest specification. The same speed descriptors that appear in the grid-level model are included so that the model-attributed association of the functional-zone label is evaluated conditional on the observed speed descriptors. Because speed may be on the built-environment pathway, this quantity is not interpreted as the total functional-zone association.

To examine whether the relative importance of built-environment factors varies across functional zones, we additionally fit five zone-specific Random Forest models, one per zone. Each takes the six community-level 5D indicators (POI density, POI diversity, road density, intersection density, destination accessibility within 2 km and bus stop density within 2 km) plus mean speed and speed variability as predictors and the per-community EC as the target. The community-level 5D indicators are obtained by an area-weighted aggregation of the underlying grid-level indicators into the host community.

Each Random Forest yields its prediction by averaging across  $B = 500$  bootstrap trees as in Equation (8), which keeps the per-tree variance low while preserving the non-linear interactions captured inside each tree.

$$\hat{y}(x) = \frac{1}{B} \sum_{b=1}^B T_b(x) \quad (8)$$

Each model uses a 70/30 random train and test split. Performance is reported using MAE and RMSE. SHAP [18] is computed with the tree-based estimator on the entire test set and provides global feature-importance bars, beeswarm summaries and dependence plots for the top three features.

### 3. Results

#### 3.1. Electric-Vehicle Operation Profile

The 1706 usable private EVs traveled 7.8 million km between October and December 2020, consumed 87.1 GWh of electricity and emitted 47,704 tons of CO<sub>2</sub> under the Shanghai 2020 electricity grid emission factor. The median per-vehicle monthly distance is 1238 km, equivalent to approximately 41 km per day, which is consistent with private-car use in major Chinese cities. The median segment EC is 0.187 kWh/km, and the 5th to 95th percentile interval ranges from −0.13 to 0.76 kWh/km. The negative left tail around −0.5 kWh/km reflects regenerative-braking-dominated segments in which the battery is recharged during deceleration; negative EC segments account for 16.98% of analytical segments and were retained in both the descriptive distribution and the distance-weighted spatial aggregation because they represent net energy returned to the battery. By contrast, the long right tail beyond 1.5 kWh/km may arise from several distinct sources, including traction energy during repeated acceleration, low-speed operation over short distances, and auxiliary

HVAC loads. Because the traffic state and HVAC operation are not directly observed in the GPS records, the right tail cannot be uniquely attributed to congestion or cabin heating. The full distribution is shown in Figure 3. The hourly travel distribution exhibits the expected morning and evening peaks (07:00–09:00 and 17:00–19:00), and the in-Shanghai segment share has a median of 99.9% per vehicle.

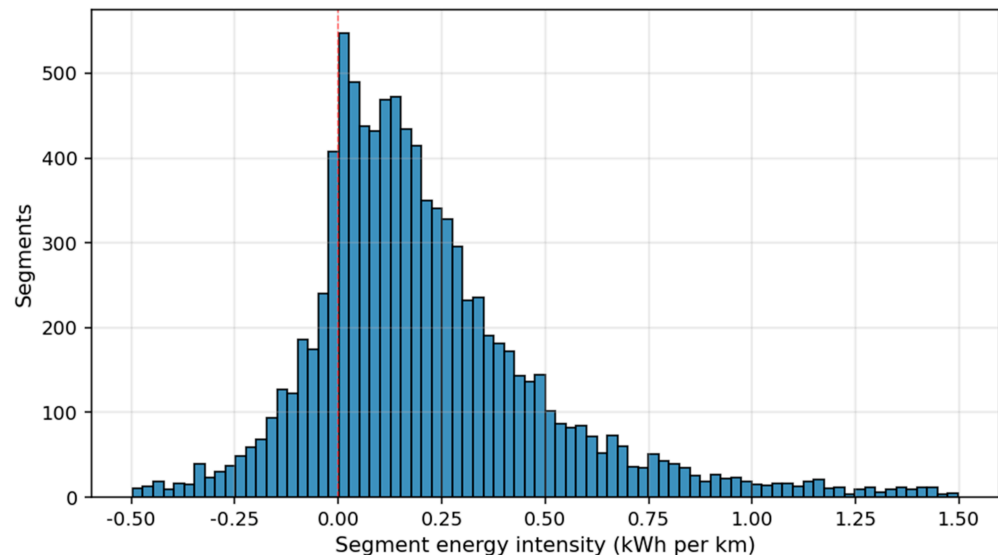


**Figure 3.** Distribution of segment EC (kWh per km) across the EV fleet.

### 3.2. Grid-Level EC

#### 3.2.1. Built-Environment 5D Factor Distribution

Across the 20,035 grids with EV records, the typical cell contains seven POIs at the median, a POI density of 28 per km<sup>2</sup>, a POI diversity (Shannon entropy) of 1.04, a road density of 5544 m/km<sup>2</sup>, 21 reachable bus stops within 2 km, and a destination-accessibility score of 296. The POI density and destination-accessibility distributions are strongly right-skewed, with maxima of 8108 per km<sup>2</sup> and 11,380, respectively, which reflects the multi-centric structure of Shanghai where commercial and service POIs concentrate inside the inner ring while sparse rural fringes dominate the outer area. Figure 4 shows the spatial distribution of the six 5D indicators over the 500 m grid. The maps reveal the multi-centric structure of Shanghai, with a high POI density, POI diversity and destination accessibility concentrated inside the inner ring and along secondary commercial corridors, while the road density and bus stop density extend across the broader urbanized area before tapering off toward the rural fringe.

#### 3.2.2. Attribution of Grid-Level EC

The grid-level Random Forest model is fit on the 3741 well-sampled grids, using the six 5D built-environment indicators plus the two segment-speed descriptors as input features and the per-grid distance-weighted EC as the target. On the held-out 30% test set, the model attains a coefficient of determination of 0.76, a mean absolute error of 0.0077 kWh/km, and a root mean squared error of 0.0106 kWh/km, while the training coefficient of determination is 0.92. The 11 features together explain about three-quarters of the spatial variance in grid-level EC, which is a substantial improvement over a baseline model that omits the speed descriptors and relies only on the 5D built-environment variables, which essentially fail to explain any spatial variance on the full set of grids with EV records. The native Random Forest feature importance attributes approximately 69% of the total to mean speed, 15% to speed variability, and the remaining 16% to the 5D built-environment indicators, led by destination accessibility, road density, road length

and bus stop density. The SHAP feature-importance bar in Figure 5 gives a consistent ranking, with mean speed as the dominant single contributor by mean absolute SHAP value, followed by speed variability and destination accessibility (2 km). Because native Random Forest importance and mean absolute SHAP are normalized differently, the SHAP chart is used for ranking consistency rather than as a second percentage decomposition. Nevertheless, mean speed is consistently identified as the most important predictor under both importance measures. The relative contributions reveal a coherent physical mechanism. The per-kilometer energy demand of an EV is governed first and foremost by the speed regime under which it is driven, since low-speed stop-and-go operation incurs frequent acceleration losses, while smoother cruising at intermediate speed operates near the powertrain's efficiency optimum [15,16]. A grid where the typical segment is slow or has highly variable speed therefore exhibits a systematically higher per-kilometer EC, while a grid with steady moderate-to-high speed sits near the efficient end of the EC distribution.

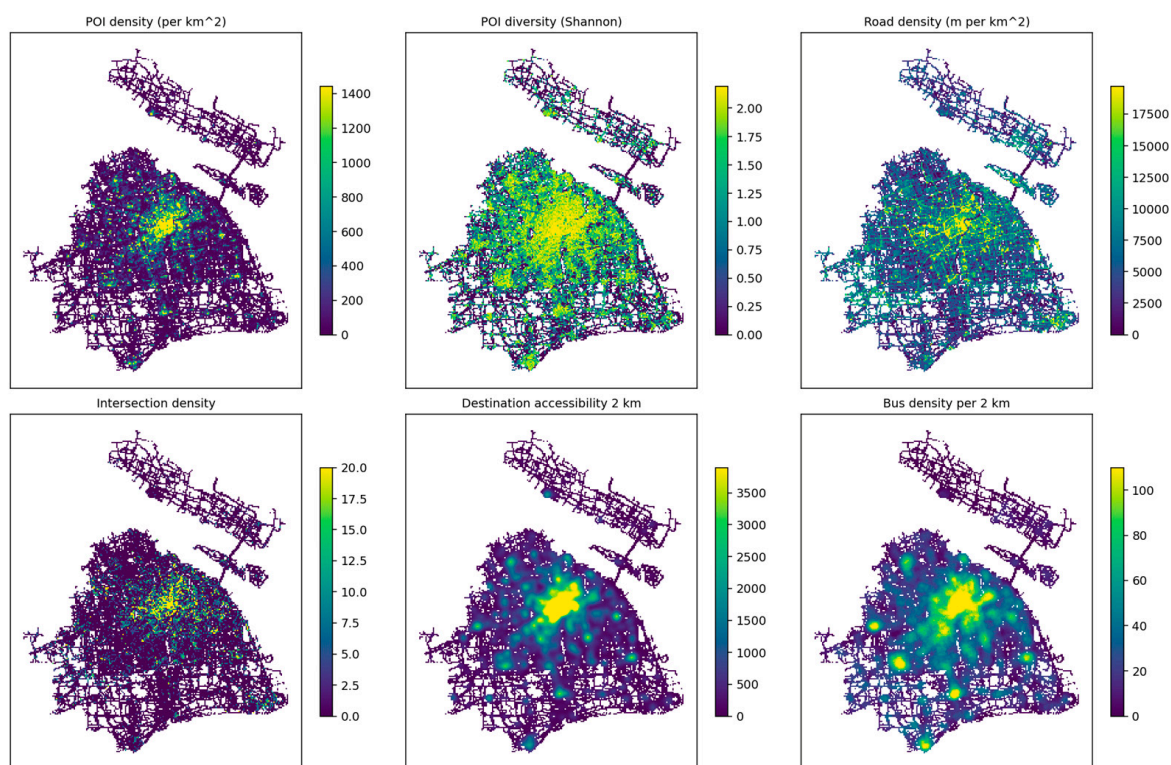


**Figure 4.** Spatial distribution of the six built-environment 5D indicators over the 500 m grid.

The SHAP beeswarm in Figure 6 and the dependence plots in Figure 7 describe associations within the fitted Random Forest. Low mean speed and high-speed variability are associated with positive SHAP values, whereas steady moderate-to-high speed is generally associated with negative SHAP values. The observed association between low or variable speed and higher EC is physically plausible because repeated acceleration increases traction-energy demand, while regenerative braking recovers only part of the kinetic energy lost during deceleration. Auxiliary loads may also raise per-kilometer EC during low-speed travel, but the present data do not separate these mechanisms. Among the 5D-inspired indicators, high destination accessibility and road density are associated with higher predicted EC over much of their observed ranges, while POI diversity, POI density, and bus-stop density show smaller and partly negative model-attributed patterns. These directions are consistent with several plausible mechanisms, including traffic concentration, repeated acceleration, route directness, and local trip chaining, but the observational design does not distinguish among them. The dependence plots show non-linearity in parts of the feature

ranges; however, neither weak displayed interactions nor the additive representation of SHAP establishes an independent marginal effect. Because the predictors are correlated, the magnitude and ordering of their SHAP values should be read as model-specific attribution rather than as causal or structural parameters.

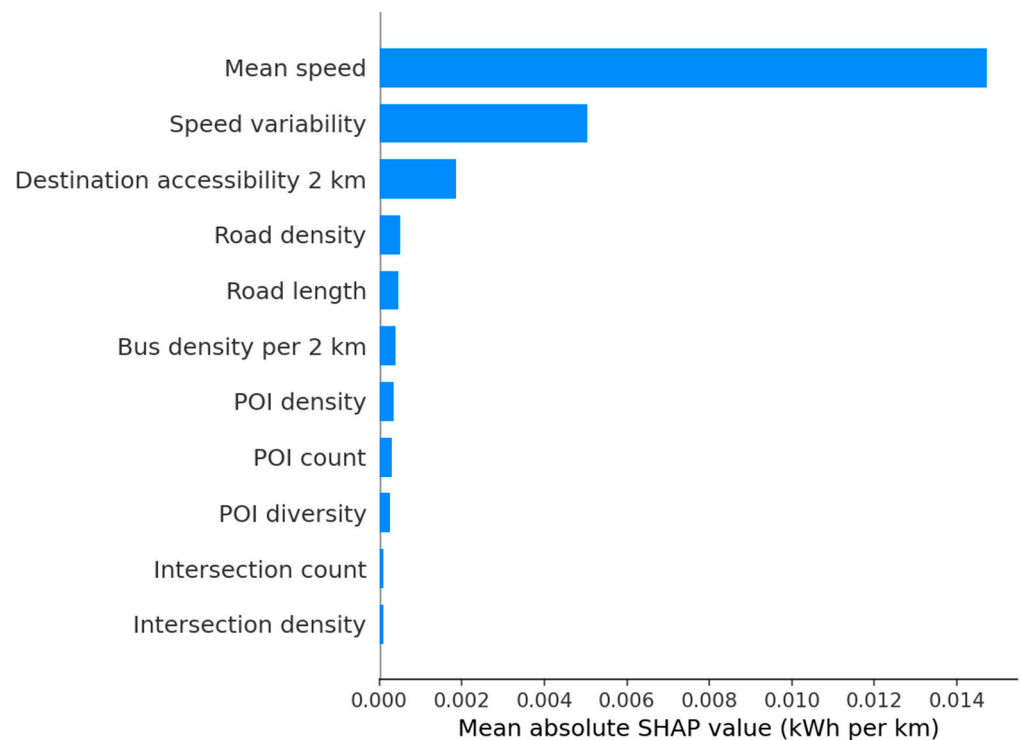


**Figure 5.** Grid-level model SHAP feature-importance bar chart.

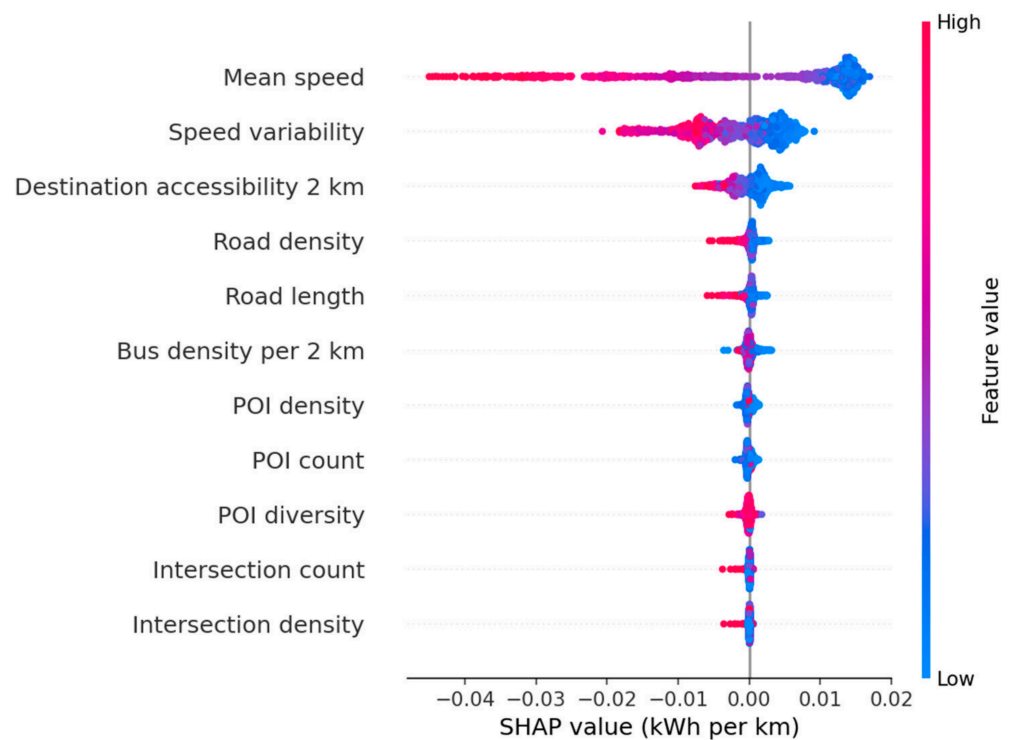


**Figure 6.** Grid-level model SHAP beeswarm summary.

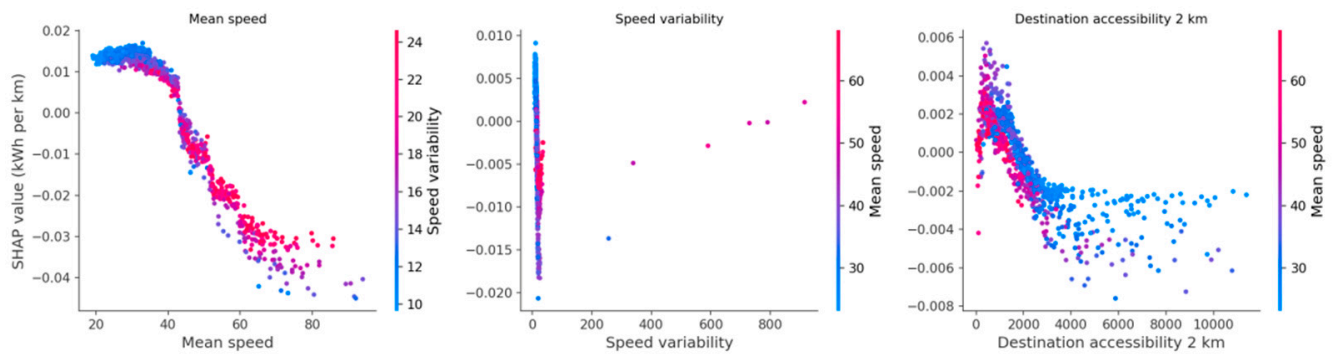


**Figure 7.** Grid-level model SHAP dependence plots for the top three features.

### 3.3. Community Functional Zone-Level EC

#### 3.3.1. Functional Zones from LLM Embeddings

The community-level K-Means clustering at  $K = 5$  produces five qualitatively distinct functional zones. These K-Means-derived community types are referred to as functional zones, that is, types that are empirically derived from POI composition rather than from statutory land-use classifications. A comparison of each cluster against POI density, POI diversity, destination accessibility and the host district reveal a clear urbanization gradient. The largest cluster ( $n = 2691$ , 43%) concentrates in the traditional inner districts of Xuhui, Jing'an, Putuo and Yangpu and exhibits the highest POI density (819 per  $\text{km}^2$ ), the highest destination accessibility (2401) and the lowest mean driving speed (31 km/h), and is labeled the Urban core zone (UCZ). A second dense cluster ( $n = 1255$ ) shares a similar inner-district footprint at a slightly lower density and is labeled the Dense urban residential zone (DUR). A larger outer cluster ( $n = 1278$ ) concentrates in Chongming, Qingpu and Songjiang, combines a moderate POI density with mixed uses, and is labeled the Suburban mixed-use zone (SMZ). A smaller outer cluster ( $n = 708$ ) covers Fengxian, Jiading and Jinshan with lower density and is labeled the Peri-urban residential zone (PRZ). The smallest cluster ( $n = 371$ ) sits at the lowest POI density (103 per  $\text{km}^2$ ), is heavily concentrated on Chongming and other agricultural districts, and is labeled the Rural zone (RZ). The normalized mutual information of 0.168 against the external urban-or-rural binary confirms that the embedding captures the urban-to-rural divide partially but is far from collapsing onto it, which indicates that the LLM zones encode richer functional information than a binary urban-or-rural categorization alone. Figure 8 displays the  $K = 5$  cluster assignment over the communities and visualizes the inner-ring versus outer-ring spatial structure.

#### 3.3.2. Among-Zone Variation in EC

The community-level Random Forest model is fit on the 4138 well-sampled communities, using the  $K = 5$  LLM-derived functional zone one-hot label together with mean speed and speed variability as predictors. On the held-out 30% test set, the model attains a coefficient of determination of 0.66, a mean absolute error of 0.0088 kWh/km and a training coefficient of determination of 0.78. The model therefore explains about two-thirds of the spatial variance in community-level EC, somewhat lower than the grid-level model above because the larger community averages over more heterogeneous driving conditions and therefore has a narrower energy-consumption distribution to predict. The community-level SHAP summary in Figure 9 reveals the same physical pattern as the grid-level model, where speed variability and mean speed dominate the bar chart, while the five LLM-derived functional zone indicators contribute small but consistent additional importance on top.

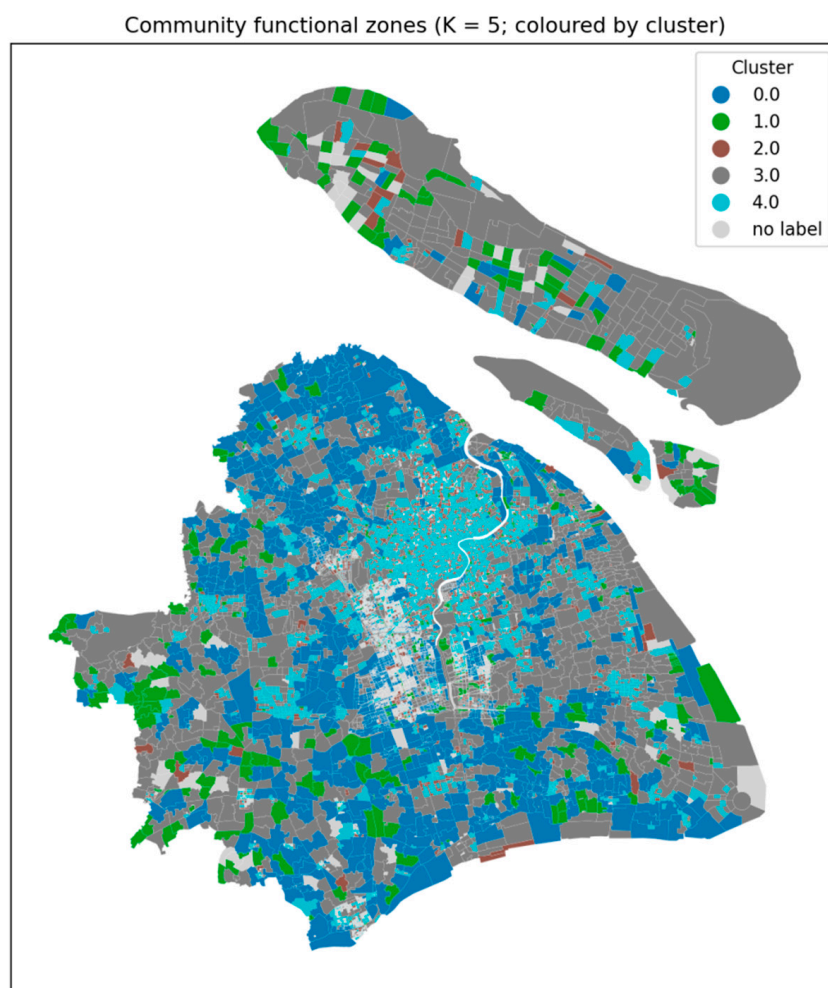


**Figure 8.** Community functional zones (K = 5) colored by cluster.

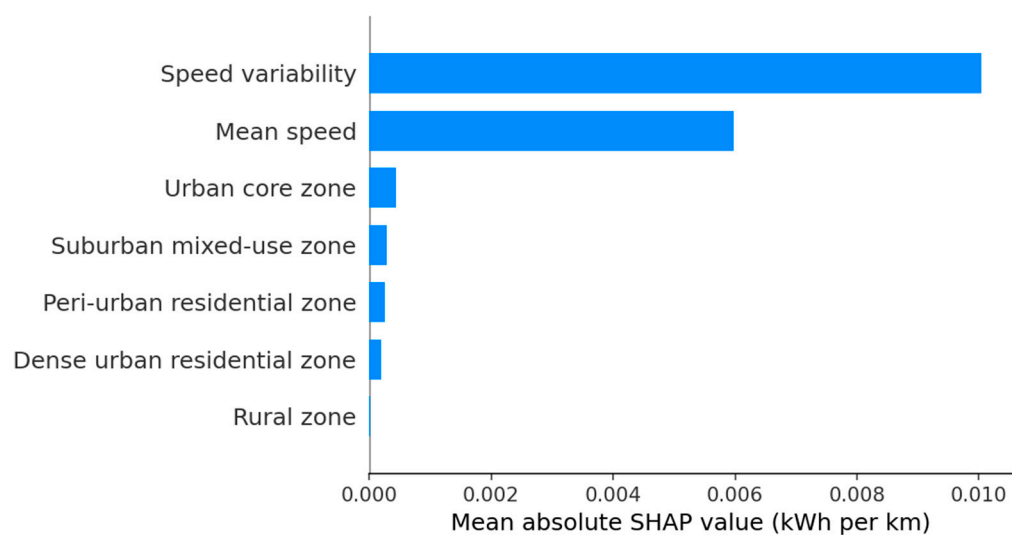


**Figure 9.** Community-level model SHAP feature-importance bar chart.

Aggregating EC by zone in a fleet-weighted manner, namely as the ratio of total segment energy to total segment distance within each zone, makes the urbanization gradient explicit. The fleet-weighted EC of the Urban core zone (0.219 kWh/km) and the Dense urban residential zone (0.219 kWh/km) is about 9% higher than that of the Rural zone (0.202 kWh/km), with the Suburban mixed-use and Peri-urban residential zones falling in between (0.209 kWh/km each). Among the five zone indicators in the Random Forest, the Urban core dummy contributes the largest mean absolute SHAP value, followed in descending order by Suburban mixed-use, Peri-urban residential, Dense urban residential and Rural. This ranking is consistent with the fleet-weighted gradient, since switching a community into or out of the Urban core zone produces the largest change in predicted EC, which makes the Urban core indicator the most informative zone signal in the model. The functional zone effect therefore aligns with the physical mechanism inferred at the grid level, where high destination accessibility and dense road networks (both concentrated in the urban core) push EC upward, while sparser peri-urban and rural communities pull EC downward.

### 3.3.3. Within-Zone Heterogeneity of Built-Environment Associations

To examine whether the relative importance of built-environment factors varies across zone types, we fit five independent Random Forest models, one per functional zone, each taking the six community-level 5D indicators (POI density, POI diversity, road density, intersection density, destination accessibility within 2 km and bus stop density within 2 km) plus mean speed and speed variability as predictors and the per-community EC as the target. The community-level 5D indicators are obtained by an area-weighted aggregation of the underlying grid-level indicators into the host community. The five zone-specific samples range from 250 communities in the Rural zone to 1651 in the Urban core zone. Because the within-zone EC dispersion is small relative to the among-zone gradient documented above, the absolute predictive performance of any single per-zone model is necessarily low. The analysis is therefore exploratory and focuses on relative SHAP rankings within each zone rather than absolute predictive performance.

The zone-specific SHAP rankings in Figure 10 reveal two descriptive patterns that are masked by the global community-level model. First, the dominant speed descriptor flips at the urbanization boundary. In the two dense zones (Urban core and Dense urban residential), speed variability outranks mean speed, since stop-and-go losses caused by signalized intersections and congestion dominate the per-kilometer energy account. In the three sparser zones (Suburban mixed-use, Peri-urban residential and Rural), mean speed outranks speed variability, since the average cruising speed on uncongested roads is the more informative summary of the local driving regime. Second, the top built-environment association flips along the same urbanization boundary. In the three denser zones, the highest-ranked built-environment factor is destination accessibility, consistent with a trip-generation channel through which accessible mixed-use cores attract shorter and stop-and-go trips. In the two sparser zones, the top built-environment factor is road density, consistent with a network-supply channel through which the local road density modulates routing choice and effective travel speed on otherwise lightly trafficked corridors. The heterogeneity analysis therefore complements the convergence finding above by showing that the same built-environment indicators operate through different mechanisms in different parts of the city.

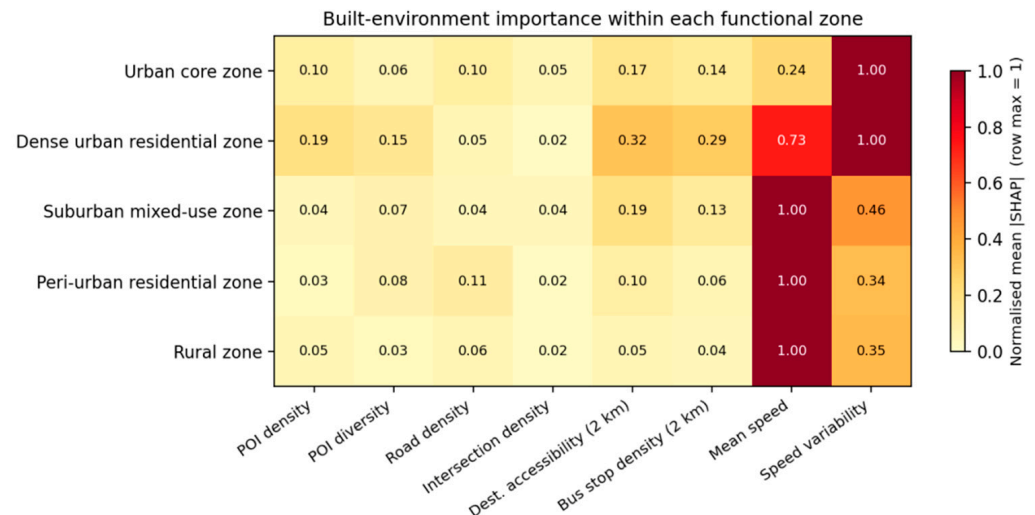


**Figure 10.** Per-zone SHAP rankings of built environment and speed features (normalized to row maximum).

## 4. Discussion

### 4.1. Conditional Built-Environment Associations Beyond the Observed Speed Descriptors

The grid-level Random Forest decomposition yields a clear and physically coherent picture of how the built environment is associated with per-kilometer EC. When the driving speed regime is excluded from the model, the 5D built-environment framework alone explains essentially none of the spatial variance in EC across the full set of grids. After restricting the analysis to well-sampled grids and adding two speed descriptors (mean speed and speed variability), the same model reaches a test coefficient of determination of 0.76, of which the two speed descriptors account for about 84% of the Random Forest feature importance and the 5D built-environment indicators account for the remaining 16%, led by destination accessibility, road density and bus stop density. The community-level analysis provides parallel evidence at a coarser spatial scale. The fleet-weighted EC of the Urban core zone is about 9% higher than that of the Rural zone, a between-zone gradient too large to be a sampling artefact and consistent with the upward pull of destination accessibility and road density identified at the grid level. The decomposition therefore supports a two-tiered interpretation of the spatial pattern of EV energy use, in which the speed regime sets the dominant first-order signal of EC by determining how often the powertrain operates near its efficiency optimum [15,16], while the built environment is associated with the spatial distribution of that regime by modulating how often a driver must accelerate, idle or stop. Within the fitted observational model, the speed descriptors and built-environment indicators receive distinguishable model-attributed importance. This separation is conditional on the included features and does not identify independent causal pathways or intervention effects. The implication is that built-environment interventions are not a small residual lever after speed-side optimization but a substantial parallel lever that operates on the same spatial pattern, and that overlooking either component would understate the policy space available for low-carbon EV adoption.

### 4.2. Multi-Scale Convergence

The dominant features at the grid and community levels point in the same direction and reinforce one another. Destination accessibility and road density at the grid level, together with the LLM-derived functional zone label at the community level whose urban core indicator carries the largest mean absolute SHAP value, are associated with higher predicted EC at the high-density end, while this association trends downward at the

low-density end, which corresponds to stop-and-go driving in dense mixed-use cores and smoother flow in lower-density areas. POI diversity and bus stop density exhibit small downward pulls on EC in both models, which suggests that neighborhoods with rich, transit-served destinations may reduce circuitous driving. The 9% fleet-weighted EC gradient between the Urban core and Rural zones quantifies the same upward push at the community scale and is consistent with the urban–rural EC gradient documented at the grid level, reinforcing the convergence finding.

The zone-specific analysis sharpens this picture by adding a heterogeneity dimension that is invisible to the global community-level model. The dominant built-environment association shifts at the urbanization boundary. In the dense zones (Urban core, Dense urban residential and Suburban mixed-use), destination accessibility is the top built-environment association with EC, consistent with a trip-generation channel through which accessible mixed-use cores attract shorter and stop-and-go trips. In the sparser zones (Peri-urban residential and Rural), road density takes over as the exploratory association, consistent with a network-supply channel through which the local road density modulates routing choice and effective travel speed on otherwise lightly trafficked corridors. The LLM embedding plays a complementary role to the 5D built-environment indicators. At the grid scale, the 5D variables enter the predictive model directly alongside the speed descriptors. At the community scale, the same built-environment signal is summarized more compactly as a single  $K = 5$  functional zone label derived from unsupervised K-Means clustering of the Qwen2.5-1.5B-Instruct representations [12]. This implementation is conceptually related to the broader urban neural-embedding literature [11] but uses a different model and aggregation procedure. The human-readable zone names (Urban core zone, Dense urban residential zone, Suburban mixed-use zone, Peri-urban residential zone and Rural zone) are obtained without any prior land-use annotation and align with the multi-centric structure of Shanghai shown in Figure 8. The combination of multi-scale convergence in the global signal and zone-specific heterogeneity in the model-attributed ranking supports a descriptive interpretation in which built-environment characteristics are systematically associated with the fitted spatial pattern of EC, conditional on the available speed descriptors and subject to the measurement and identification limitations discussed below.

#### *4.3. Planning Relevance and Hypotheses for Future Evaluation*

The results provide planning hypotheses and screening information rather than evaluated intervention effects. In the sampled data, locations characterized by a high destination accessibility or road density and by low or variable speeds tend to receive higher predicted EC or positive SHAP values. These spatial indicators may help identify areas in which additional traffic, vehicle, charging-demand, and infrastructure data should be collected before an intervention is designed. The differences in feature rankings across functional zones also suggest that a single city-wide interpretation may obscure local context; however, they do not establish that changing destination accessibility, road density, or transit supply would change EC by a corresponding amount. This study does not model charging demand or location choice, signal timing, bus-lane interventions, road-capacity changes, or time-varying electricity load and carbon intensity. It therefore cannot support direct prescriptions for fast or wireless charging, dedicated bus lanes, signal coordination, road expansion, or off-peak charging incentives. Evaluating such measures requires longitudinal or quasi-experimental designs that observe conditions before and after a clearly defined intervention.

#### 4.4. Limitations and Future Work

Several limitations should be considered when interpreting the results. First, the representation of the built environment remains incomplete. The six 5D-inspired indicators capture selected aspects of density, diversity, design, destination accessibility, and transit provision rather than a comprehensive operationalization of the full 5D framework. In particular, intersection density depends on the topological processing of OpenStreetMap road nodes and therefore represents street connectivity rather than the full Design dimension; bus-stop density reflects the concentration of surface-transit facilities rather than the literal distance to transit; metro accessibility is not included; and destination accessibility is calculated using Euclidean rather than network distance. In addition, potentially relevant vehicle characteristics (e.g., make-model, mass, battery capacity and health), driver characteristics (e.g., driver identity and driving style), route characteristics (e.g., gradient and road class), weather conditions (e.g., ambient temperature and precipitation), HVAC operation, and temporal covariates (e.g., month, day type, and hour-of-day) are unavailable in the present dataset. The two speed-based summary metrics cannot fully compensate for these missing predictors, and similar constraints also exist in physics-based EV energy modeling studies [15,16]. Where omitted factors exhibit systematic spatial variation, part of their influence may be reflected in the estimated associations with built-environment or functional-zone features. Future work could incorporate more complete multimodal accessibility measures and remote-sensing land-cover information, such as built-up, green, and barren land cover, to complement the present POI- and OSM-based characterization of destination functions and road-network context, together with richer vehicle and operating-condition data.

Second, the validation of both the derived EC estimates and the fitted models remains limited. Independent validation against odometer readings or manufacturer-reported energy consumption was not possible because these variables and vehicle-specific battery capacities were unavailable in the anonymized dataset. The derived EC values should therefore be interpreted as telemetry-based estimates rather than independently validated measurements. In addition, the minimum numbers of analytical segments and vehicles used to retain spatial units were adopted as practical data-quality criteria rather than statistically optimized thresholds, and sensitivity to alternative thresholds was not systematically evaluated. In addition, alternative spatial units, including 250 m and 1 km grids, hexagonal grids, road-buffer units, and origin/destination buffers, were not re-estimated. Therefore, the sensitivity of the reported spatial patterns and feature rankings to the modifiable areal unit problem (MAUP) remains untested. Model performance was assessed using a random train–test split, whereas spatial block validation, district-held-out validation, repeated splits, and uncertainty intervals for SHAP rankings were not implemented. Consequently, the reported predictive performance and zone-specific feature rankings should be interpreted as evidence within the present sample rather than fully validated measures of spatial transferability. Future studies should incorporate independent energy-use benchmarks where available and adopt spatially explicit, repeated, and sensitivity-based validation strategies.

Third, the LLM-derived functional-zone framework remains exploratory. The Qwen-based POI-sequence representation was not systematically benchmarked against alternative representations such as TF-IDF, Word2Vec, Doc2Vec, Place2Vec, sentence-transformer embeddings, or simple POI-composition vectors. Alternative pooling strategies and clustering solutions were also not comprehensively evaluated, and the resulting functional zones were not independently validated against authoritative land-use data. The identified community types should therefore be interpreted as empirically derived spatial archetypes rather than definitive land-use classes, and the present analysis does not establish the superiority of the Qwen-based representation. Systematic benchmarking of alternative representations

and pooling strategies, together with external validation of the resulting clusters, would provide stronger evidence for the robustness and transferability of the functional-zone framework.

Finally, this study is observational and is based on one private-EV fleet in Shanghai over October–December 2020, which limits temporal, fleet, and geographic generalizability. Seasonal HVAC demand, traffic conditions, and fleet composition may differ across periods and cities, while the use of an annual-average electricity emission factor does not capture short-term variation in grid carbon intensity. Moreover, mean speed and speed variability are operating variables directly related to EC but may themselves be partly shaped by the built environment. Conditioning the models on these variables may therefore account for part of the pathway through which the built environment is associated with EC. Consequently, the reported built-environment importance represents predictive associations conditional on the observed speed characteristics rather than the total association between the built environment and EC. The present observational design cannot separate speed-mediated and other pathways or establish the effects of specific planning interventions. Longer multi-season panels, cross-fleet and cross-city comparisons, temporally resolved electricity-emission factors, and longitudinal or quasi-experimental designs would provide stronger evidence for the transferability and causal interpretation of the observed relationships.

## 5. Conclusions

This study integrates telematics from 1706 private EVs and approximately 15 million analytical driving segments with POI, road, transit, grid, and community data in Shanghai. The multi-scale framework identifies systematic spatial variation in per-kilometer EC and uses Qwen2.5-1.5B-Instruct embeddings to derive five post hoc, interpretable community types. Within the fitted Random Forest and the included feature set, the two speed descriptors receive approximately 84% of the native feature importance, and the six 5D-inspired indicators receive approximately 16%. The Urban core zone has approximately 9% higher fleet-weighted EC than the Rural zone, and zone-specific models produce different feature rankings across dense and sparse areas. These findings are observational, location-level associations conditional on the available speed summaries. SHAP values are model attributions rather than causal effects, and the CO<sub>2</sub> estimate is an annual-average accounting conversion of EC rather than an independent modeling outcome. Random spatial splitting may overstate transferable performance, while unobserved vehicle, driver, route, weather, HVAC, and temporal factors may affect the reported attributions. The principal contribution is therefore the integration of EV telematics, built-environment indicators, and LLM-derived spatial representations for descriptive multi-scale analysis. Causal planning claims require spatially explicit validation, richer covariates, and longitudinal or quasi-experimental evidence.

**Author Contributions:** Conceptualization, Zhe Zhang; methodology, Zhe Zhang; software, Can Yin; validation, Can Yin and Lifu Jin; formal analysis, Can Yin and Zhe Zhang; investigation, Can Yin; resources, Zhe Zhang; data curation, Zhe Zhang; writing—original draft preparation, Can Yin; writing—review and editing, Zhe Zhang; visualization, Zhe Zhang; supervision, Zhe Zhang. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Data Availability Statement:** The data are available on request.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. International Energy Agency. *Global EV Outlook 2025*; IEA: Paris, France, 2025. Available online: <https://www.iea.org/reports/global-ev-outlook-2025> (accessed on 14 May 2025).
2. Shanghai Municipal People's Government. Shanghai Tops Global Cities in NEV Ownership. 2024. Available online: <https://english.shanghai.gov.cn/en-Latest-WhatsNew/20240305/53790527f40b46b083010a9690009184.html> (accessed on 5 March 2024).
3. Cervero, R.; Kockelman, K. Travel demand and the 3Ds: Density, diversity, and design. *Transp. Res. Part D Transp. Environ.* **1997**, *2*, 199–219. [\[CrossRef\]](#)
4. Ewing, R.; Cervero, R. Travel and the built environment: A meta-analysis. *J. Am. Plan. Assoc.* **2010**, *76*, 265–294. [\[CrossRef\]](#)
5. Fetene, G.M.; Kaplan, S.; Mabit, S.L.; Jensen, A.F.; Prato, C.G. Harnessing big data for estimating the energy consumption and driving range of electric vehicles. *Transp. Res. Part D Transp. Environ.* **2017**, *54*, 1–11. [\[CrossRef\]](#)
6. Wu, X.; Freese, D.; Cabrera, A.; Kitch, W.A. Electric vehicles' energy consumption measurement and estimation. *Transp. Res. Part D Transp. Environ.* **2015**, *34*, 52–67. [\[CrossRef\]](#)
7. Yuan, X.; Zhang, C.; Hong, G.; Huang, X.; Li, L. Method for evaluating the real-world driving energy consumptions of electric vehicles. *Energy* **2017**, *141*, 1955–1968. [\[CrossRef\]](#)
8. Le, Q.; Mikolov, T. Distributed representations of sentences and documents. In Proceedings of the International Conference on Machine Learning, Beijing, China, 22–24 June 2014.
9. Yao, Y.; Li, X.; Liu, X.; Liu, P.; Liang, Z.; Zhang, J.; Mai, K. Sensing spatial distribution of urban land use by integrating points-of-interest and Google Word2Vec model. *Int. J. Geogr. Inf. Sci.* **2017**, *31*, 825–848. [\[CrossRef\]](#)
10. Zhai, W.; Bai, X.; Shi, Y.; Han, Y.; Peng, Z.-R.; Gu, C. Beyond Word2vec: An approach for urban functional region extraction and identification by combining Place2vec and POIs. *Comput. Environ. Urban Syst.* **2019**, *74*, 1–12. [\[CrossRef\]](#)
11. Fan, C.; Yang, Y.; Mostafavi, A. Neural embeddings of urban big data reveal spatial structures in cities. *Humanit. Soc. Sci. Commun.* **2024**, *11*, 409. [\[CrossRef\]](#)
12. Yang, A.; Yang, B.; Zhang, B.; Hui, B.; Zheng, B.; Yu, B.; Li, C.; Liu, D.; Huang, F.; Wei, H.; et al. Qwen2.5 Technical Report. *arXiv* **2024**, arXiv:2412.15115. [\[CrossRef\]](#)
13. Huang, W.; Cui, L.; Chen, M.; Zhang, D.; Yao, Y. Estimating urban functional distributions with semantics preserved POI embedding. *Int. J. Geogr. Inf. Sci.* **2022**, *36*, 1905–1930. [\[CrossRef\]](#)
14. Liu, B.; Deng, Y.; Li, M.; Yang, J.; Liu, T. Classification schemes and identification methods for urban functional zone: A review of recent papers. *Appl. Sci.* **2021**, *11*, 9968. [\[CrossRef\]](#)
15. Asamer, J.; Graser, A.; Heilmann, B.; Ruthmair, M. Sensitivity analysis for energy demand estimation of electric vehicles. *Transp. Res. Part D Transp. Environ.* **2016**, *46*, 182–199. [\[CrossRef\]](#)
16. Fiori, C.; Ahn, K.; Rakha, H.A. Power-based electric vehicle energy consumption model: Model development and validation. *Appl. Energy* **2016**, *168*, 257–268. [\[CrossRef\]](#)
17. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [\[CrossRef\]](#)
18. Lundberg, S.M.; Lee, S.-I. A unified approach to interpreting model predictions. *Adv. Neural Inf. Process. Syst.* **2017**, *30*, 4768–4777.
19. Holland, S.P.; Mansur, E.T.; Muller, N.Z.; Yates, A.J. Are there environmental benefits from driving electric vehicles? The importance of local factors. *Am. Econ. Rev.* **2016**, *106*, 3700–3729. [\[CrossRef\]](#)
20. Zhong, Z.; Yu, Y.; Zhao, X. Revisiting electric vehicle life cycle greenhouse gas emissions in China: A marginal emission perspective. *iScience* **2023**, *26*, 106565. [\[CrossRef\]](#) [\[PubMed\]](#)
21. Xu, C.; Wang, Y.; Ding, W.; Liu, P. Modeling the Spatial Effects of Land-Use Patterns on Traffic Safety Using Geographically Weighted Poisson Regression. *Netw. Spat. Econ.* **2020**, *20*, 1015–1028. [\[CrossRef\]](#)
22. Zhang, W.; Guo, B.; Zhao, Y.; Wang, P. Relationship and spatiotemporal evolution of resource-based urban mixed land use and carbon emissions: A study of 11 cities in Shanxi province. *J. Urban Plan. Dev.* **2026**, *152*, 04026032. [\[CrossRef\]](#)

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.