

Article

Nonlinear Impact Analysis of Urban Road Traffic Carbon Emissions Based on the Integration of Gasoline and Electric Vehicles

Dongcheng Xie¹, Xingzi Shi¹, Kai Li¹, Jinwei Li¹ and Gen Li^{2,*} 

¹ College of Transportation and Logistics, Southwest Jiaotong University, Chengdu 611756, China; orangedc@my.swjtu.edu.cn (D.X.); 2024200770@my.swjtu.edu.cn (X.S.); 2023200744@my.swjtu.edu.cn (K.L.); 2023211424@my.swjtu.edu.cn (J.L.)
² College of Automobile and Traffic Engineering, Nanjing Forestry University, Nanjing 210037, China
* Correspondence: ligen@njfu.edu.cn

Abstract: With the rapid proliferation of electric vehicles (EVs) in China, the landscape of transportation carbon emissions has undergone significant changes. However, research on the impact of the built environment on the carbon emissions of mixed traffic from gasoline and electric vehicles remains sparse. This paper focuses on urban traffic scenarios with a mix of gasoline and electric vehicles, analyzing the spatiotemporal distribution of carbon emissions from both types of vehicles and their nonlinear association with the built environment. Utilizing trajectory data from gasoline-powered and electric taxis in Chengdu, China, we establish segment-level carbon emission estimation models based on the vehicle-specific power of gasoline vehicles and the equivalent energy consumption of electric vehicles. Subsequently, we employ the XGBoost algorithm and SHapley Additive ExPlanation (SHAP) to analyze the nonlinear relationships between 13 built environment variables and vehicle carbon emissions. This paper reveals that most built environment variables exhibit nonlinear relationships with traffic carbon emissions, with five factors—population density, road density, residential density, metro accessibility, and the number of parking lots—having a significant impact on road carbon emissions. Finally, we discuss the carbon reduction benefits of EV adoption and propose policy recommendations for low-carbon initiatives in the transportation field.



Academic Editor: Antonio Caggiano

Received: 26 November 2024

Revised: 20 January 2025

Accepted: 26 January 2025

Published: 4 February 2025

Citation: Xie, D.; Shi, X.; Li, K.; Li, J.; Li, G. Nonlinear Impact Analysis of Urban Road Traffic Carbon Emissions Based on the Integration of Gasoline and Electric Vehicles. *Buildings* **2025**, *15*, 488. <https://doi.org/10.3390/buildings15030488>

Copyright: © 2025 by the authors. 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 (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: built environment; traffic carbon emissions; electric vehicle; nonlinear effects; XGBoost

1. Introduction

Climate change represents one of the greatest challenges to global sustainable development, with controlling carbon dioxide (CO₂) emissions recognized as a central strategy to address this issue [1,2]. China has pledged to achieve peak carbon emissions by 2030 and carbon neutrality by 2060 [3]. Approximately 24% of energy-related emissions are attributed to the transportation sector, with road transport accounting for 74% of the sector's total emissions [4–6]. In China, where road transportation emissions remain persistently high, reducing emissions has become a critical focus for both policy and research [7].

To address climate change and promote “green development”, China has identified the promotion of new energy vehicles (NEVs) as a pivotal measure under its carbon peaking and carbon neutrality strategies [8]. In 2024, China's NEV fleet reached 10 million vehicles, reflecting an annual growth rate of 59.25%, with battery electric vehicles (BEVs)

accounting for 82% of the total. While BEVs achieve zero emissions during operation, the emissions generated during electricity production cannot be overlooked [9]. In this context, the large-scale adoption of BEVs contributes to carbon reduction. However, China's coal-dominated power generation mix raises concerns that supplying electricity for BEVs could contribute to air pollution concerns, posing potential threats to air quality and public health [10,11]. Consequently, it is crucial to study the coupling effects of vehicle electrification (VE) and power generation mix (PGM) to evaluate the impact of electrification trends on carbon and pollutant emissions. Investigating the carbon emissions of electric vehicles on the consumption side, particularly during transportation, not only provides a more comprehensive understanding of energy use and spatial distribution but also enables a more accurate quantification of carbon reduction benefits.

Research on carbon emissions from single-gasoline vehicles has extensively examined spatiotemporal distribution and influencing factors. As electric vehicles (EVs) proliferate in urban traffic, there is a pressing need to quantify their carbon footprint and elucidate the coexistence of road carbon emissions from both EVs and internal combustion engine vehicles. This analysis is crucial for understanding how urban spatial configuration affects carbon emissions and guiding urban land use and infrastructure planning. Despite its importance, this area of research remains largely unexplored.

Traditional studies have often assumed a linear or log-linear relationship between the built environment and carbon emissions, typically employing linear models to explore this relationship in the context of transportation-related emissions [12–14]. However, recent findings suggest that linear assumptions may fail to fully capture the complexities of real-world scenarios, thereby compromising the accuracy of carbon emission predictions [15]. To address this limitation, researchers have begun exploring nonlinear relationships and threshold effects to uncover the intricate characteristics of how built environment factors influence carbon emissions. For instance, high levels of mixed land use, population density, employment density, and public transit route density are generally effective in reducing private vehicle usage. However, once these factors exceed certain thresholds, their positive effects may diminish rapidly [16–18].

Recent studies have highlighted the importance of non-linear relationships in the context of urban environments and carbon emissions. For example, Yang et al. [19] employed a gradient boosting decision tree (GBDT) method to investigate the nonlinear effects of multi-scale built environments on CO₂ emissions from commuting, revealing complex and non-linear relationships between built environment factors and carbon emissions. Similarly, Wu and Li [20] utilized GBDT to examine the nonlinear impacts of urban form factors on transportation carbon emissions in 282 Chinese cities. The study highlights significant threshold effects, particularly in polycentricity and urban dispersion, providing valuable guidance for urban planning to reduce carbon emissions through optimized spatial strategies. These studies underscore the need to move beyond linear assumptions to better understand and predict the impacts of urban development on carbon emissions.

Despite these advances, research on transportation carbon emissions remains limited by the difficulty of fully incorporating built environment factors. Although recent studies have introduced machine learning techniques such as gradient boosting decision trees and random forests to enhance analysis [21–23], spatial heterogeneity in carbon emissions—such as differences observed in urban canyons, high-density areas, and open spaces—remains inadequately addressed. Recent advances in machine learning offer promising tools to address these challenges. Song et al. [24] proposed a tree-based MCST-tree model integrating multi-source data for high-resolution air pollution mapping. Similarly, Song et al. [25] introduced the Deep-MAPS framework, which leverages urban big data and machine learning to perform spatial inference of PM_{2.5} concentrations with fine

granularity. However, while these studies have made significant strides in addressing the spatial heterogeneity of air pollution, they primarily focus on PM_{2.5} concentrations and do not directly address the specific challenges of carbon emissions in complex urban environments. Consequently, developing a comprehensive model that accounts for the spatial variability of transportation carbon emissions is critical. Such a model would provide robust scientific support for crafting more targeted and effective environmental policies.

This paper explores the nonlinear relationship between the built environment and transportation carbon emissions, focusing on an urban road traffic scenario where gasoline vehicles and electric vehicles coexist. Therefore, this paper addresses two key questions: (1) How do urban road carbon emissions evolve over time and space with the coexistence of gasoline and electric vehicles? (2) How do built environment variables impact transportation carbon emissions from gasoline and electric vehicles, and what spatial adjustments effectively reduce traffic emissions?

Leveraging GPS data from both gasoline and electric taxis in Chengdu, this paper adopts the XGBoost regression model along with the SHAP method to enhance the interpretation of complex nonlinear dynamics. By examining the spatiotemporal nonlinear effects of built environment factors on transportation carbon emissions, this paper enriches existing studies and provides deeper insights into global and local nonlinear interactions. The results indicate that the built environment's impact on carbon emissions varies significantly across temporal and spatial scales, showing distinct threshold effects. These findings provide critical evidence for developing targeted carbon reduction policies tailored to regional and temporal needs.

The structure of this paper is as follows. Section 2 reviews the literature on the relationship between road transportation carbon emissions and the built environment. Section 3 describes the data and variables included in the model. Section 4 outlines the research methodology. Section 5 presents the detailed results. Finally, Section 6 summarizes the key findings, proposes carbon reduction policy recommendations, and discusses directions for future research.

2. Literature Review

2.1. Calculation of Road Traffic Emissions

Road transportation emissions are estimated using top-down or bottom-up methods, selected based on geographic scale, data granularity, and availability [26]. Top-down methods are effective for large-scale analyses, such as at national, provincial, or county levels [2], relying on total energy consumption and emission factors. For example, Alam et al. [27] calculated Ireland's national road emissions using the 2016 emission inventory, while Singh et al. [28] evaluated CO₂ trends in India's road transport sector through fuel consumption data. In Tianjin, Sun et al. [29] developed a vehicle emission inventory using detailed local datasets. De Nunzio et al. [30] further introduced a framework to estimate road traffic emissions using macroscopic traffic and road data.

Bottom-up methods, utilizing detailed data such as vehicle type, travel distance, and fuel consumption per unit, provide a more precise characterization of carbon emissions from mobile sources [31]. For instance, Luo et al. [32] analyzed the spatial distribution of taxi energy consumption and emissions in Shanghai using GPS data. Kan et al. [33] proposed a fine-grained microscopic model to calculate vehicle fuel consumption and emissions. Liu et al. [34] reconstructed vehicle emission profiles by integrating taxi GPS trajectories with license plate recognition data. Pla et al. [2] developed a bottom-up method to quantify greenhouse gas emissions from urban road traffic.

Previous studies have proposed various methods for calculating road transportation emissions, but limitations remain in terms of spatiotemporal resolution, particularly at the

road level [2]. High-resolution road-level emission data are crucial for advancing sustainable urban development. Onboard GPS data, which accurately captures vehicle driving states, has gradually become a focal point of research. The vehicle specific power (VSP) model, based on regression analysis of measured emission data, establishes the relationship between driving parameters and average emissions. It has emerged as a key tool for studying road emissions, demonstrating high accuracy in fuel consumption estimation for public vehicles [35] and light-duty taxis [36]. In recent years, researchers have applied the VSP model to analyze urban travel activities and associated emissions. For example, Zhang et al. [37] used taxi trajectory data to estimate fuel consumption and carbon emissions, visualizing the spatial distribution of emissions within traffic networks. Xia et al. [38] developed a method to calculate daily travel emissions using taxi records, exploring the relationship between urban morphology and carbon emissions. Similarly, Chen et al. [39] analyzed the temporal characteristics of multimodal commuters, investigating the potential impact of low-carbon travel on emission reductions.

However, current GPS-based emission studies primarily focus on emissions at the administrative scale or within traffic networks. Research on the spatial distribution of emissions from daily travel sources at finer resolutions remains limited, highlighting the need for more detailed and localized analyses.

2.2. Driving Factors Analysis on Road Traffic Emissions

Transportation emissions have been extensively studied in relation to the built environment, with a particular focus on the role of density. Zahabi et al. [40] found that a 10% increase in residential density can lead to a 2.5% reduction in carbon emissions. Similarly, Wu et al. [22] observed that both high population density and high employment density contribute to reducing carbon emissions, with employment density having a more significant impact. However, the relationship between density and emissions is not straightforward. Some studies suggest that increased population density can lead to higher energy consumption and emissions from transportation [12,41].

In areas with diverse land use, optimizing land allocation reduces travel distances, which in turn decreases car dependency and related carbon emissions. A study found that increasing land use diversity by 10% can result in a 2.5% reduction in transportation-related carbon emissions [42]. Another study observed that combining different land uses boosts subway usage, thereby lowering emissions [43]. It was highlighted that a greater concentration of workplaces relative to residential areas can shorten commute distances and cut down greenhouse gas outputs [44]. Another study noted that the influence of land use mix on carbon emissions is notable only when the index surpasses 0.4 [22]. Furthermore, studies suggested that if land use diversity exceeded a certain threshold, it might lead to congestion and parking issues, impacting emissions [45]. This suggested a potentially nonlinear relationship between land use diversity and transportation carbon emissions.

With increasing distance from the city center, automobile use tended to rise, exacerbating environmental pollution [46]. However, researchers found that the distance from city center was negatively correlated with the proportion of people using cars at work. On the whole, there was a complex nonlinear relationship between traffic carbon emissions and the distance from city center and transportation carbon emissions [22,23].

2.3. Research Method Development in Built Environment and Carbon Emissions Studies

The relationship between the built environment and carbon emissions has been widely researched, with initial studies primarily using linear models to assume a direct and proportional link between characteristics like density, diversity, and design, and car-

bon emissions. For example, Ewing and Cervero [47] used linear regression to demonstrate that higher urban density and diversity were associated with lower vehicle miles traveled (VMTs) and reduced carbon emissions. Li et al. [48] employed linear regression to examine how 13 factors impact traffic CO₂ emissions, highlighting economic scale, population density, transportation structure, and energy consumption as critical influences. However, these linear models were limited by their simplicity, inability to capture threshold effects, and lack of interaction effects among variables.

Recognizing these limitations, researchers have increasingly turned to nonlinear models to more accurately capture the complex dynamics between the built environment and carbon emissions. Nonlinear models, such as machine learning techniques like random forests and gradient boosting machines, offer a more nuanced approach. These models can account for threshold effects and interactions, providing a better fit to the data and enhancing predictive accuracy. He et al. [49] utilized the random forest method to examine the influence of key variables on the three travel patterns of short-distance car users, exploring the nonlinear correlations and interactions among different variables. Wu et al. [50] constructed a tree-boosting algorithm based on GPS taxi trip data from Dalian City to investigate the spatiotemporal heterogeneity of taxi travel carbon emissions and their nonlinear relationship with the built environment.

Despite the advantages of nonlinear models, current research still faces challenges. Nonlinear models require larger and more detailed datasets to accurately estimate complex relationships, which can be difficult to obtain. Additionally, these models are more complex and harder to interpret than linear models, making it difficult for policymakers and practitioners to understand and apply the results. Many nonlinear studies are also context-specific and may not be generalizable to other regions or settings.

The above review highlights several limitations of existing research. Firstly, most studies used linear models to analyze the relationship between the built environment and transportation carbon emissions, overlooking potential nonlinearities. Secondly, traditional machine learning research had not adequately considered the spatial heterogeneity of the built environment, limiting its effectiveness for spatial policy decisions. Additionally, commonly used PDP interpreters face limitations due to independence assumptions. To address these gaps, this paper introduces an optimized gradient boosting decision tree model and compares its performance with existing machine learning models. Using SHAP interpreters, we explore the complex impacts of built environment factors on traffic carbon emissions across different spatial and temporal dimensions, reveal threshold effects and simultaneously explain spatiotemporal heterogeneity and nonlinear relationships with the built environment, which greatly expands the depth of existing literature.

3. Study Area and Data Sources

3.1. Study Area

Chengdu is the provincial capital of Sichuan Province, the central city in Southwest China, and an international comprehensive transportation hub city. Its urban function is the “Western Economic and Technological Innovation Center, Western Center for Foreign Exchanges, and National Advanced Manufacturing Base”. By the end of 2023, the permanent resident population in Chengdu had reached 21.403 million, among which the urban permanent resident population was 17.229 million, and the urbanization rate was 80.5% [51]. In 2023, Chengdu’s regional gross domestic product (GDP) reached CNY 2.20747 trillion. By the end of 2023, the number of automobiles in Chengdu had reached 6.7406 million, ranking first in China. The study area is shown in Figure 1.

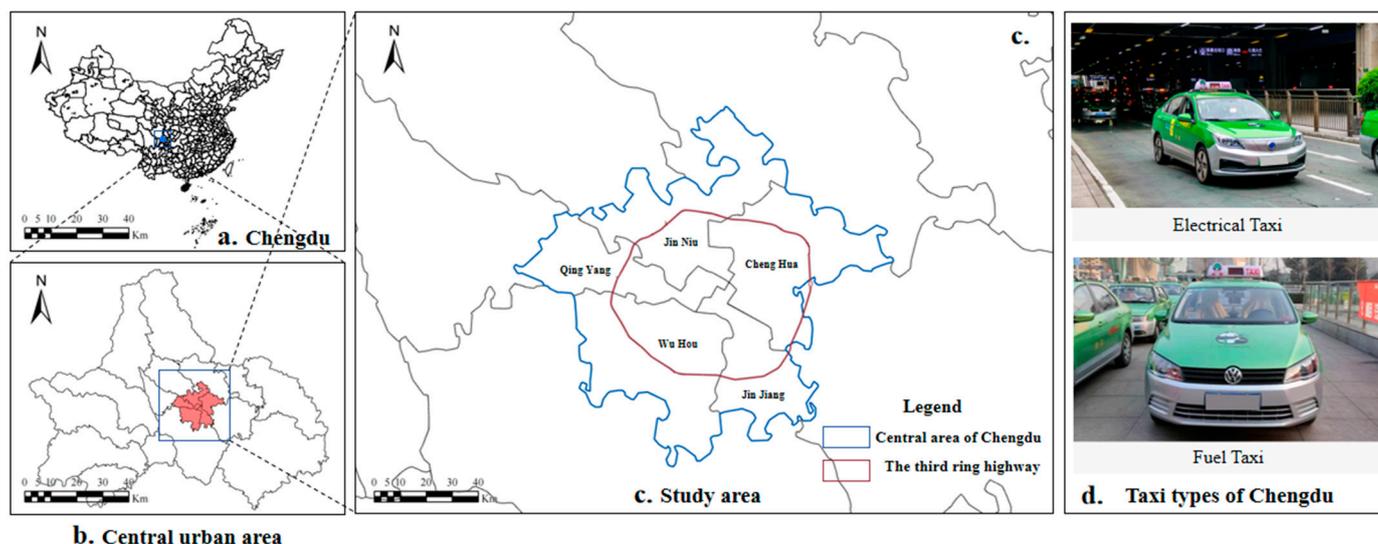


Figure 1. Study area.

Figure 1a,b show the geographical location and scope of Sichuan Province and Chengdu City; Figure 1c shows the central area of Chengdu; Figure 1d shows the taxi types of Chengdu.

The study area comprises the five central urban districts of Chengdu: Chenghua, Jinniu, Qingyang, Wuhou, and Jinjiang. These districts cover an area of 424.06 km², with a population of 6.5 million, accounting for 30% of the city's total population. The central urban area constitutes 3% of Chengdu's total area but contributes 35% of the city's GDP, amounting to CNY 764.164 billion. Economically, this region is the most developed, with high-end services predominating and focusing on three high-tech industries: electronic information, biomedicine, and digital economy.

To accurately analyze the taxi carbon emissions of the study area and investigate the relationship of urban infrastructure, we divided the study area into 500 m × 500 m grids. The 500 m grid division is widely used in studies of urban spatial regression [50,52,53]. The smaller grid enriches the spatial difference [54] in each region, which helps the regression model to better capture the difference features. This division resulted in 2096 grids, enabling us to find carbon emission characteristics across the entire city more accurately.

3.2. Data Resource and Processing

3.2.1. Taxi GPS Data

In this paper, the GPS data of taxis across the whole area of Chengdu from 30th March to 5th April 2022, a total of seven days, are selected to analyze the road carbon emission situation in the central area of Chengdu. There are approximately 12,000 taxis in Chengdu, generating about 9.05 million pieces of trajectory data daily. Taxi GPS data are collected by the on-board positioning equipment and sent to the data center regularly (the data sampling interval is 30 to 60 s). The data includes nine fields, namely generation time, license plate number, longitude and latitude, driving speed, direction angle, operating status, and data availability. Table 1 is a sample of the original taxi GPS data.

Table 1. Taxi GPS data sample.

Date	Time	Vehicle_ID	Lon	Lat	Speed	Angle	Status	Avail
2022 0601	51,919	川 ADU**1	104.031197	30.639927	49.0	103	1	1
2022 0601	51,935	川 ADT0**2	104.109182	30.686768	0.3	105	1	1
2022 0601	52,235	川 ADT3**6	103.882307	30.814181	23.0	76	1	1

The dataset spans seven days and covers many urban areas, providing comprehensive spatio-temporal characterization. As shown in Table 1, each Vehicle_ID is unique, starting with “川|A” for Chengdu, Sichuan Province, followed by five characters for gasoline taxis and six for electric taxis. The Time column represents GPS timestamps. Lat and Lon indicate vehicle coordinates. Angles is the azimuth describing the direction of the taxi. Speed denotes the instantaneous velocity of the taxi. Status indicates the taxi’s operation mode, with 0 for empty and 1 for occupied. Avail reflects the device status, with 0 for offline due to malfunction and 1 for normal operation.

Before conducting data analysis, it is essential to process the raw taxi GPS data. The data processing involves data organization, cleaning, and trajectory extraction. During the data organization phase, since time segments store the original trajectory data, we first merge the data by date, select the complete daily trajectory data for each vehicle based on the license plate number, and store them in groups sorted by time order. In the data cleaning phase, we remove data with a device status of 0 and blank fields. Additionally, trajectory points with speeds exceeding 120 km/h will be removed by the speed limits of urban roads in China [34,55]. Error filtering is applied to each vehicle’s order trajectory data. If a taxi order mainly consists of location data with a speed of 0 and a constant position, then the order will be deleted. Furthermore, data for single-order trips with distances less than 300 m or greater than 350 km will be filtered out [33]. After the completion of data preprocessing, taxi trajectory data for different time periods were obtained as 260,010, 491,358, 197,888, 447,221, 256,652, and 466,871 records, respectively.

3.2.2. Built Environment Data

This paper acquires land development data for the central urban area of Chengdu in April 2022. According to the Chinese national standard “Classification of Land Use Status (GB T 2010–2017)” [56], various types of land are categorized and merged, ultimately retaining four types of construction land: public management and service land, commercial land, residential land, and transportation land.

Population data are obtained through the WorldPop project of the University of Southampton (<https://www.worldpop.org/>, accessed on 1 December 2024). The WorldPop project is an open, high-resolution geospatial dataset from the University of Southampton that generates information on population distribution and demographics using various statistical and simulation methods. This dataset is widely used by scholars worldwide for scientific research [57,58]. This paper obtained the 2022 China 100 m·100 m grid population spatial dataset corrected by the United Nations.

The urban road network is a ground road structure composed of roads of different grades, functions, and locations distributed across various urban areas with a certain density and form. The density of the road network also reflects the convenience and accessibility of travel in different urban areas and is one of the important bases for guiding and formulating urban traffic regulations. The road network data in this paper comes from

the OpenStreetMap. OpenStreetMap is an open-source map that can be downloaded from the web.

The POI data in this paper comes from Amap, one of China’s largest online map applications. The Chengdu POI data for April 2022 includes 14 major categories: shopping services, medical services, residential communities, and others. To improve the sample’s representativeness, we screen and classify the POIs based on their actual operating hours, effectiveness, and relevant regulations [59]. Among the 14 major categories, the life service category mainly includes facilities such as public toilets, which have a weak correlation with travel and thus were removed from the study.

In most Chinese cities, there is a lack of statistical data that can quantify the economic level of smaller-scale areas. Since housing prices reflect residents’ income to a certain extent [60], they are used in this paper to reflect residents’ economic levels. The housing price data in this paper comes from the Chinese real estate intermediary website Beike, and the Chengdu housing price data for April 2022 are crawled.

This paper obtain data on subway and bus stations in Chengdu in April 2022 through Amap. Considering the different impacts of buses and subways on taxi travel, the two are processed separately. Existing studies generally believe that the impact of the subway has a range of effects. The “Urban Rail Transit Line Planning and Design Guidelines” issued by the Ministry of Housing and Urban–Rural Development in 2015 defines the area within 500–800 m around urban rail transit stations (a 15-min walkable distance) as the traffic impact zone of urban rail transit stations. In domestic and international studies, the impact radius of rail transit stations often adopts an 800 m radius. Therefore, this paper constructs a buffer zone with an 800 m radius around each subway station as the impact range of the subway station. Figure 2 shows the spatial distribution and heatmap of the above data in the study area.

3.2.3. Variables

Numerous studies have demonstrated the significant role of the built environment in effective urban planning, with a scientific distribution of buildings capable of reducing traffic-related carbon emissions [61]. This paper employs a nonlinear model to analyze the relationship between traffic carbon dioxide emissions and the built environment. The dependent variable is the carbon dioxide emissions of vehicles in the study area grid. The explanatory variables of the built environment mainly include five dimensions: facility density, facility diversity, transportation service level, road network coverage level, and social population. These dimensions are based on the 5D theory [62]. It specifically includes 13 factors, including land type, population, road density, public transportation accessibility, and residential and work facility coverage density [22,38,63,64]. Table 2 provides a descriptive analysis of each variable.

Table 2. Built environment variable definition and statistics.

Variable	Variable Description	Mean	S.D.	Min	Max
Density					
Work poi	Locations related to work or business activities	569.02	721.10	0.00	7933.09
Live poi	Locations that are primarily residential or related to living	36.52	52.43	0.00	376.24
Government land	Area designated for government buildings or public institutions	0.20	0.32	0.00	1.00
Business land	Area designated for business or commercial use	0.09	0.22	0.00	1.00
Live land	Area designated for residential use	0.43	0.38	0.00	1.00

Table 2. Cont.

Variable	Variable Description	Mean	S.D.	Min	Max
Work land	Area designated for industrial	0.11	0.27	0.00	1.00
Diversity					
HHI	A measure of concentration or diversity of land use or business types in an area	1.55	0.56	0.00	2.33
Transportation service					
Parking density	The number of parking spaces per unit of area.	39.57	47.36	0.00	1129.17
Bus stop density	The number of bus stops per unit of area.	8.22	6.45	0.00	40.02
Subway buffer zone ratio	The proportion of area within a certain distance	39.57	47.36	0.00	1129.17
Road network coverage					
Road diversity	The variety of road types or road functions in an area	12.01	7.96	0.00	88.89
Social population					
Population density	The number of people living per unit of area	14,123.86	15,820.99	40.00	98,216.00
Average house price	The average price of houses in the area.	18,084.17	7928.91	5700.00	65,121.81

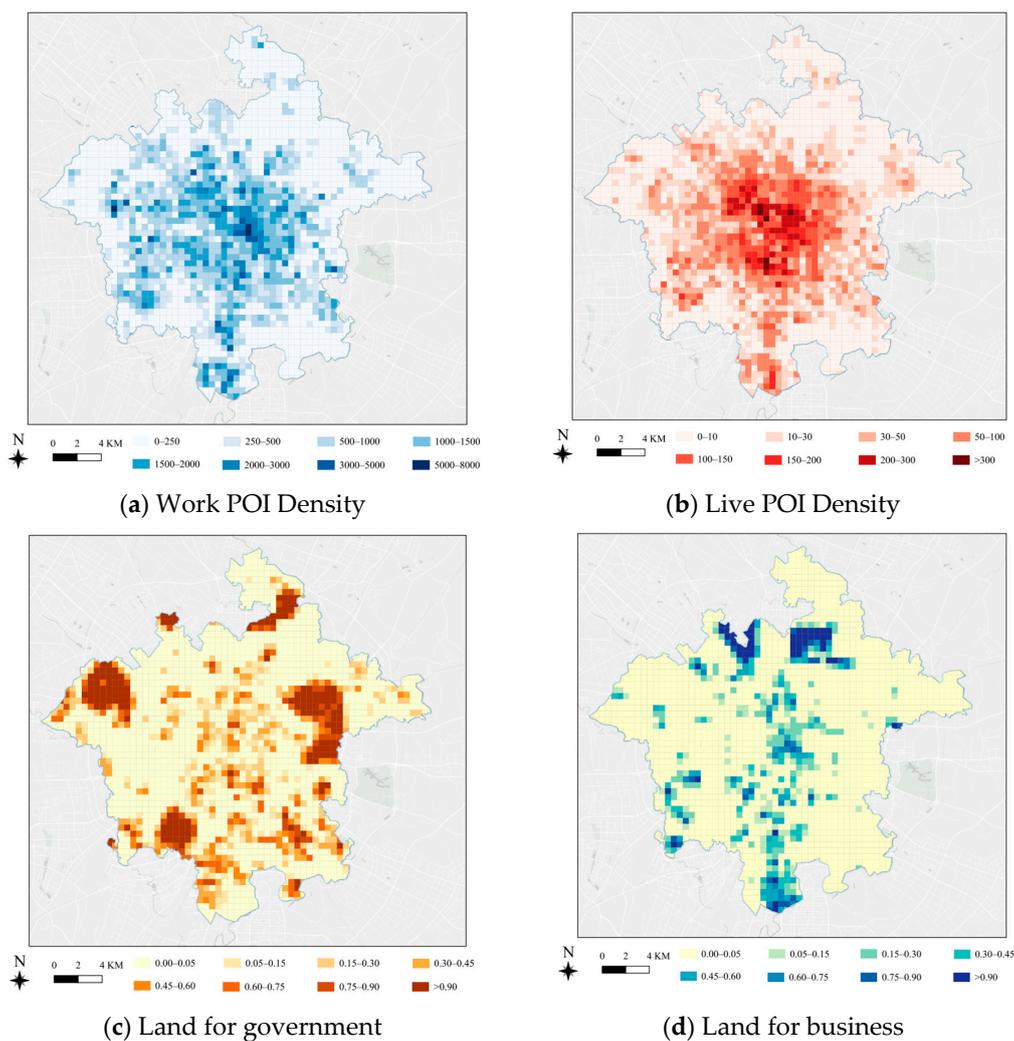


Figure 2. Cont.

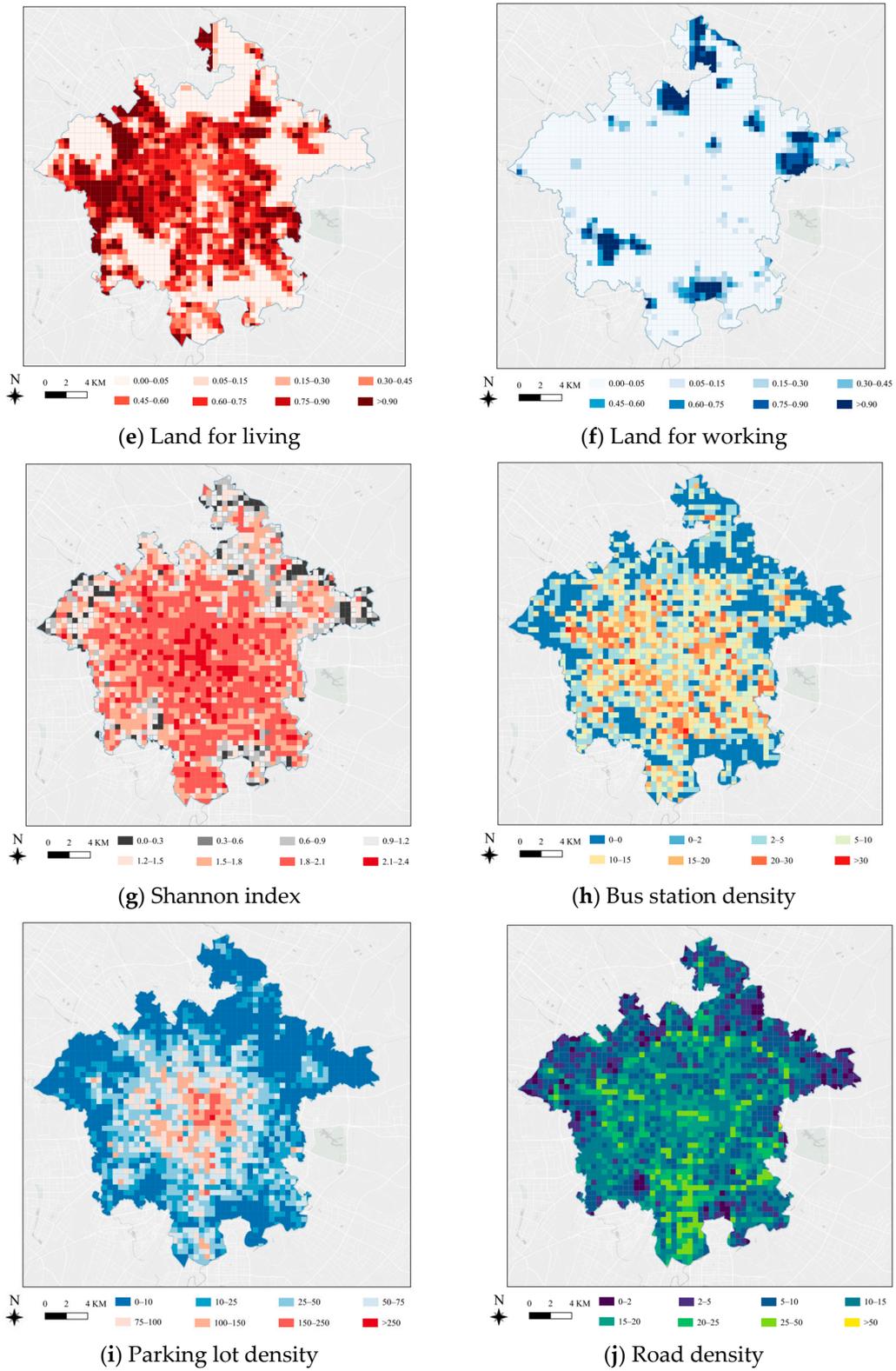


Figure 2. Cont.

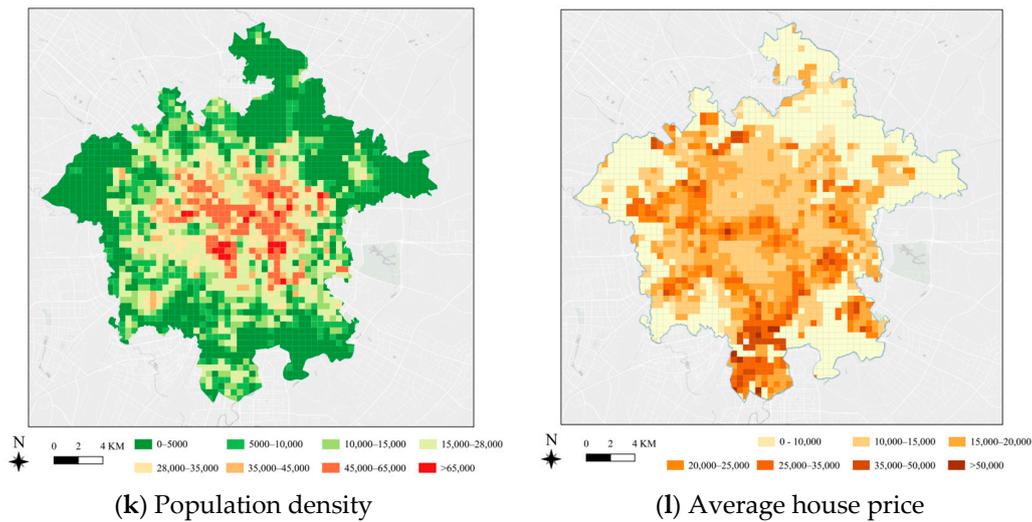


Figure 2. The spatial distribution and heatmap of the built environment data.

4. Methodology

4.1. Study Framework

Figure 3 illustrates the research framework, which consists of three main steps:

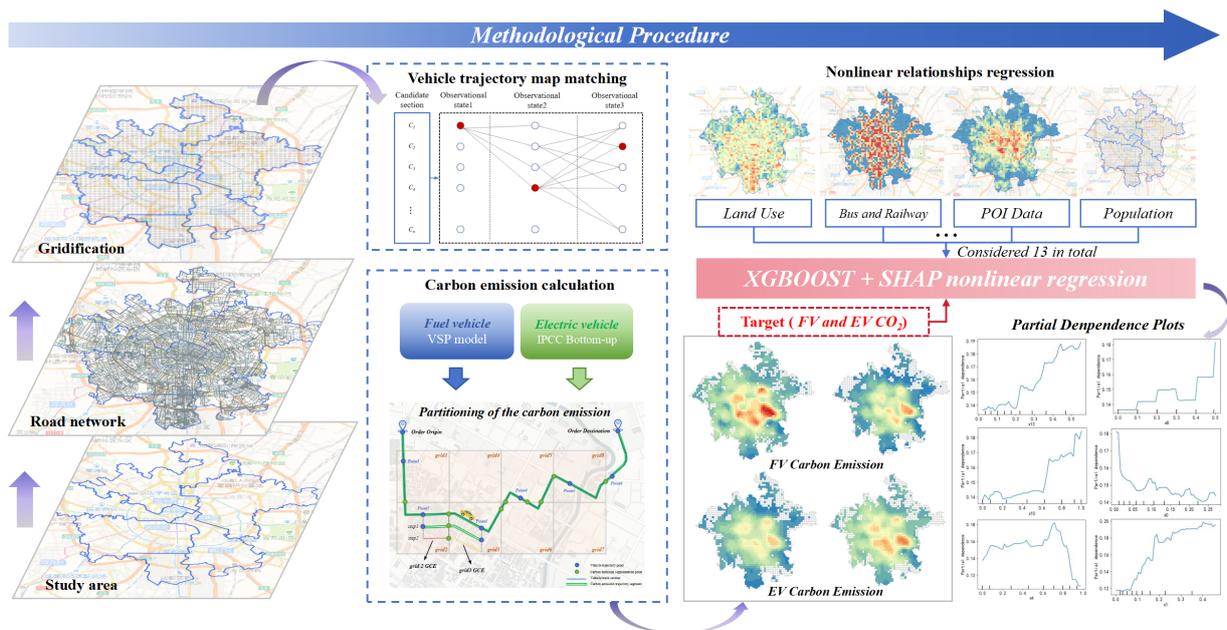


Figure 3. Study methodology.

(1) Trajectory Matching

This paper converts discrete trajectory points into continuous road sequences, as calculating the distance between two adjacent points alone may underestimate the mileage of motor vehicles. This matching process transforms point-to-point linear connections into actual road-based travel distances.

(2) Road Traffic Carbon Emission Accounting

This paper establishes a carbon emission accounting model for electric and traditional gasoline vehicles based on trajectory data is established. For gasoline taxis, the VSP model calculates traffic carbon emissions for each road segment using vehicle trajectories and emission factors. For electric taxis, this paper follows the Chinese national standard “GBT37340-2019 Calculation Method for Energy Consumption of Electric Vehicles” [65],

electricity consumption is first converted into equivalent fuel consumption, and then carbon emissions are estimated by the IPCC model. After calculating the carbon emissions of both vehicles, the study area is divided into several grids for grid allocation of road traffic carbon emissions and analysis of the spatiotemporal characteristics of carbon emissions.

(3) Driving Factor Analysis

This paper examines 13 key built environment factors, and the impact of each factor on road traffic carbon emissions is analyzed using the XGBoost algorithm, with the SHAP model used for nonlinear interpretation and analysis of intrinsic characteristics and relationships from PDP plots. Finally, the spatiotemporal heterogeneity of the built environment's impact on urban traffic emissions is analyzed by comparing algorithms such as random forests.

4.2. Map Matching for Taxi Trajectory Data

Some taxi positioning devices are unstable, leading to resulting in trajectory point loss and drift. Additionally, due to urban multipath propagation loss, especially in areas with high road network density and complex urban overpasses, the raw taxi trajectory may significantly deviate from the actual roads, resulting in a large discrepancy between the connected distance of trajectory points and the actual travel distance [34]. The sampling interval of onboard devices is often more than 30 s, leading to a sparse distribution of trajectory points. In urban areas with dense road networks, multiple potential paths may emerge, necessitating accurate identification of the actual travel path. Figure 4 illustrates the phenomena of trajectory data drift and multipath selection. As shown in Figure 4, at points A and B, there are two possible travel paths, and incorrect path selection can lead to errors in travel distance calculations. For instance, Chengdu's Second Ring Road is an elevated expressway with arterial roads constructed beneath it, overlapping with the overpass. Surrounding commercial and residential buildings interfere with positioning signals, resulting in extensive trajectory point loss and drift when vehicles traverse this area. These conditions prevent accurate travel path identification.

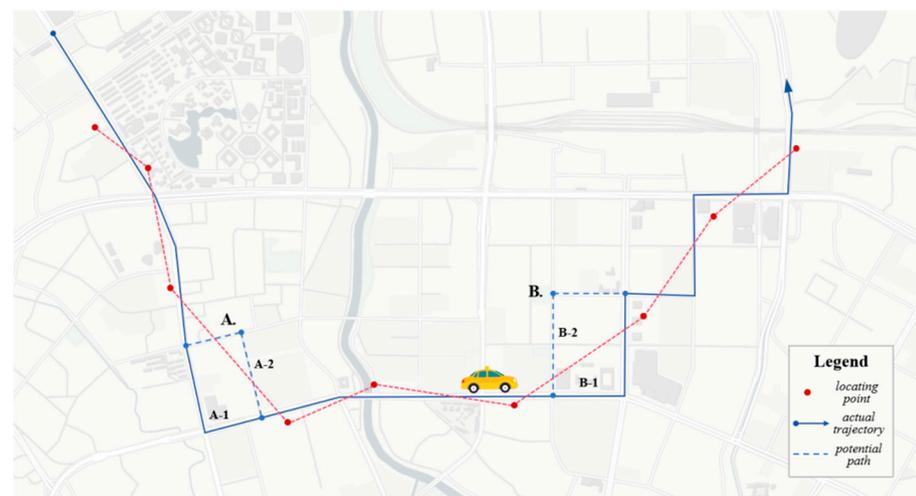


Figure 4. Map matching.

In response to the issues above, this paper employs a map-matching algorithm for identifying the true taxi trajectory. After map matching, the raw taxi location line can be converted into a real road connection, so as to extract the accurate driving distance and reduce the error of carbon emission calculation.

Map matching algorithms encompass mapping, geometric similarity and distribution probability [66–69]. This paper selects a hidden Markov model (HMM)-based probabilistic

map-matching algorithm. The hidden Markov model is a statistical model that describes a Markov process with hidden, unknown parameters. In practical applications, the internal states of certain systems and processes are not fully observable, and HMM is used to describe the transition of these states and the generation process of observations. HMM is widely applied in natural language processing, speech recognition, and time series analysis. The concept of HMM is that the current state of a system depends solely on the state of the previous moment, forming a Markov chain with hidden states that are not directly observable but can be inferred from a set of generated observation data. HMM consists of three basic elements: states, observations, and three types of probabilities (observation probabilities, state transition probabilities, and initial state probabilities). The set of states comprises all hidden states; the set of observations represents the observable manifestations of each state; observation probability is the likelihood of observing a specific observation given a particular state; state transition probability is the likelihood of the system transitioning from one specific state to another; and initial state probability refers to the probabilities of the system being in various states at the beginning.

In the context of map-matching, HMM functions as a decoding problem, determining true travel segments from GPS trajectory points. Figure 5 presents the Viterbi algorithm-based framework for hidden Markov map matching:

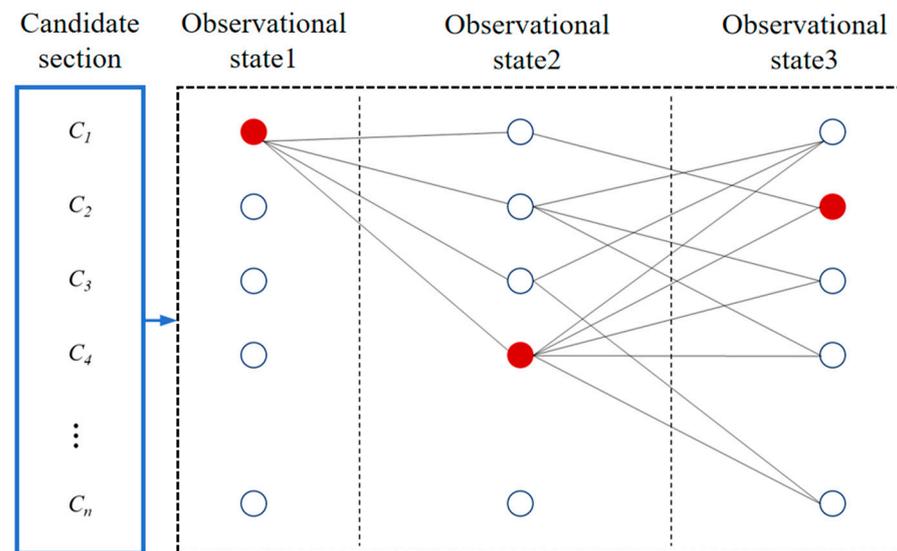


Figure 5. HMM map matching principle.

(1) Hidden States

Hidden states refer to the potential locations of a moving object. These states are not directly observable but are inferred based on the observed states. Each hidden state corresponds to a specific location on the road network. In this paper, the projection point of the observed position onto a road segment is considered as the hidden states (i.e., the candidate point).

(2) Observation States

Observation states are the positions directly observed through trajectory data, which are the vehicle's latitude and longitude information, as shown in Figure 6. Although observation states can be directly obtained through positioning devices, they are often inaccurate due to multipath effects, signal attenuation, and environmental factors. In the HMM system, observation states are used to infer the most probable hidden state, representing the moving object's true location.

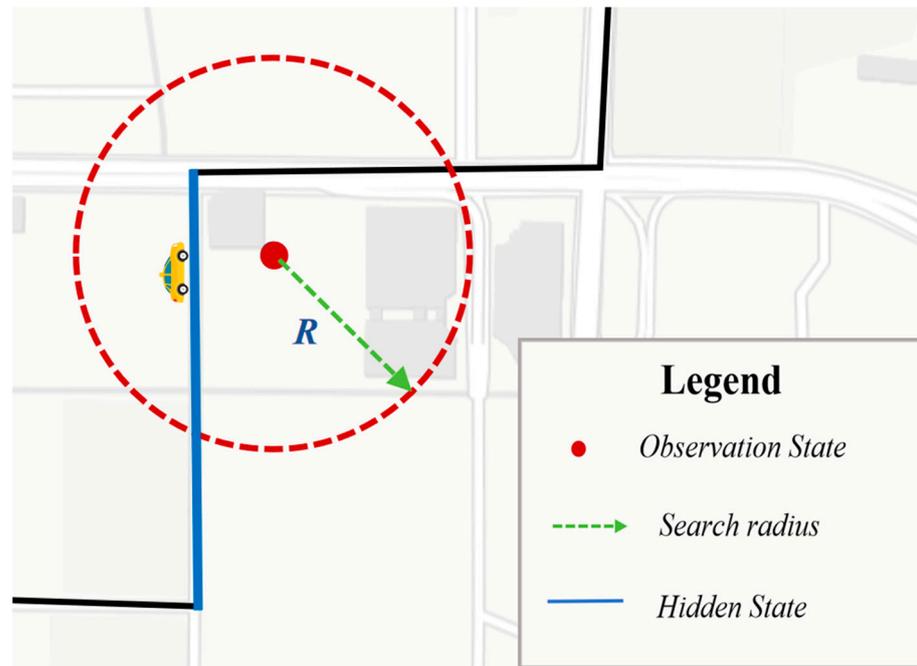


Figure 6. Hidden states and observation states.

(3) Observation Probability Matrix

The observation probability matrix maps the relationship between hidden states and observation states. The elements in the matrix represent the probability of observing the current GPS position, given the assumption that the moving object is in a specific hidden state.

The observation probability is calculated based on a normal distribution, as shown in Equation (1). In the equation, $B(k)$ represents the observation probability, P_i represents the observation point, P_i^c represents the hidden state (mean of the distribution), D_E represents the distance from the observation state to candidate road segments, and σ represents the standard deviation.

$$B(k) = P(O_k = P_i | H_k = P_i^c) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{D_E(P_i - P_i^c)^2}{2\sigma^2}} \quad (1)$$

(4) Transition Probability Matrix

Road segment transition probabilities constitute a key parameter reflecting hidden state dependencies and transition patterns in the HMM framework. The transition probability a_{ij} indicates the likelihood of moving from state i to state j in the subsequent time step. The configuration of transition probabilities must consider both the topological relationships inherent in the road network and the similarity between the time differences of adjacent data points and the corresponding travel times between road segments. The calculation formula is as follows:

$$P(S_j, S_k) = \begin{cases} \frac{1}{|t_{i,i+1} - T_{j,k}|} & T_{j,k} \neq t_{i,i+1} \\ 1 & T_{j,k} = t_{i,i+1} \end{cases} \quad (2)$$

$T_{j,k}$ represents the average travel time from road segment j to road segment k , $t_{i,i+1}$ represents the time interval between the i th observation point and the $(i + 1)$ th observation point. Using this approach, the HMM can estimate the most likely sequence of road segments based on the GPS trajectory points and the statistical model of the observation

probabilities. This method helps to correct for inaccuracies in GPS data and provides a more reliable representation of the vehicle's actual path.

4.3. Vehicle Trajectory-Based Emission Model

4.3.1. Gasoline Vehicle Carbon Emissions Calculation Method

The calculation of carbon emissions of gasoline vehicles is based on the vehicle specific power (VSP). This model first inputs vehicle speed, acceleration, road slope and other data to estimate the fuel consumption of vehicles. After obtaining the fuel consumption, the carbon emission is calculated by superimposing the carbon emission factor of the gasoline [35,62]. The optimized VSP carbon emission model is shown in Equation (3).

$$VSP = v \times [1.1a + 9.8a \times \text{grade} (\%) + 0.132] + 0.000302 \times v^3 \quad (3)$$

v is the speed of the vehicle, m/s. The grade is a dimensionless parameter, taking 0 in the application. a is the instantaneous acceleration, m/s^2 .

Due to difficulties in inputting some data (e.g., grade, acceleration), many scholars have optimized the carbon emission model based on VSP and established fuel consumption correction factors for different speed ranges. Therefore, this paper builds on these findings [70,71]. The instantaneous fuel consumption calculation incorporates the vehicle fuel consumption correction factor for each speed range. Then base on the carbon emission factor corresponding to each type of gasoline, we calculate the carbon emission of the vehicle.

Step1: Fuel consumption calculation. Fuel consumption of each section is determined as shown:

$$f_{i,l} = ER_0 \times NFCR_l \times T_{i,1} \quad (4)$$

In the equation, $f_{i,l}$ represents the actual fuel consumption of the segment l of the taxi order i ; ER_0 denotes the average fuel consumption rate of the taxi when VSP is 0, set to 0.274; $NFCR_l$ indicates the normalized fuel consumption rate for the average speed range, which is shown in Table 3; $T_{i,1}$ represents the travel time (seconds) of the taxi order i in the trajectory interval l .

Table 3. NFCR values for the different velocity intervals.

Speed (km/h)	NFCR	Speed (km/h)	NFCR	Speed (km/h)	NFCR
0–2	1.085	28–30	2.187	56–58	2.756
2–4	1.259	30–32	2.251	58–60	2.810
4–6	1.311	32–34	2.329	60–62	2.865
6–8	1.476	34–36	2.338	62–64	2.914
8–10	1.573	36–38	2.361	64–66	2.956
10–12	1.646	38–40	2.395	66–68	3.049
12–14	1.730	40–42	2.441	68–70	3.136
14–16	1.807	42–44	2.470	70–72	3.289
16–18	1.841	44–46	2.538	72–74	3.334
18–20	1.923	46–48	2.566	74–76	3.370
20–22	1.997	48–50	2.581	76–78	3.410
22–24	2.045	50–52	2.596	78–80	3.439
24–26	2.092	52–54	2.680	above80	3.551
26–28	2.163	54–56	2.716		

Step2: Carbon emissions calculation

$$CE_{i,l} = f_{i,l} \times EF_k \quad (5)$$

$CE_{i,l}$ represents the carbon emissions of the trajectory segment l of the order i ; EF_k is the carbon emission factor for 92# gasoline. According to the national standard GB 19578-2021 “Fuel Consumption Limits for Passenger Cars” [72] in China, which specifies the conversion relationship between CO₂ emissions and fuel consumption, 1 L of gasoline emits 2.21 kg of carbon dioxide. Therefore, EF_k equals 2.21 kg/L [38].

Step3: Calculate the total carbon emissions of the taxi order i

The total carbon emissions of the taxi order i from the starting point $(x_{o,i}, y_{o,i})$ to the ending point $(x_{d,i}, y_{d,i})$ is the sum of the carbon emissions of each trajectory segment l , which can be expressed as shown in Equation (6):

$$CE_i = \sum_{l=1}^n CE_{i,l} \quad (6)$$

where CE_i is the total carbon emissions for the taxi order i , and $CE_{i,l}$ is the carbon emissions of the segment l in the taxi order i . n represents the total number of trajectory segments in the taxi order i .

4.3.2. Carbon Emission Calculation of Electric Taxis

Regarding the calculation method for electric vehicle carbon emissions, according to the national standard “GBT37340-2019 Calculation Method for Energy Consumption of Electric Vehicles” [65], a carbon dioxide emission conversion method is used. This method first converts the vehicle’s electrical consumption into equivalent fuel consumption before calculating carbon emissions.

Step1: Calculation of the equivalent fuel consumption (FC_{CO_2})

$$\text{Calculate the equivalent fuel consumption } (FC_{CO_2}) \text{ according to the following : } FC_{CO_2} = E \times F_{CO_2} \quad (7)$$

where:

FC_{CO_2} is the equivalent fuel consumption in L/100 km;

E is the vehicle’s electric energy consumption in kW·h/100 km;

F_{CO_2} is the carbon dioxide conversion factor in L/(kW·h).

Step2: Calculation of the carbon dioxide conversion factor F_{CO_2} The carbon dioxide conversion factor F_{CO_2} is calculated according to the following:

$$F_{CO_2} = \frac{T_E \times T_C \times \phi}{T_F \times t_M \times i_{ch} \times (1 - i_{tr})} \quad (8)$$

where:

F_{CO_2} is the carbon dioxide conversion factor, in units of L/(kW·h);

E is the vehicle’s electric energy consumption, in units of kW·h/100 km;

T_E is the coal consumption per unit electricity generated by thermal power plants, in units of kg/(kW·h);

T_C is the carbon dioxide emission factor of coal gasoline, in units of kgCO₂/kg coal;

ϕ is the percentage of electricity generated by thermal power plants (%);

T_F is the carbon dioxide emission factor of the gasoline used in power generation, in units of kgCO₂/kWh;

t_M is the conversion factor between gasoline coal and standard coal;

i_{ch} is the charging efficiency (%);

i_{tr} is the transmission loss rate (%).

The value of F_{CO_2} obtained from this calculation is 0.31 L/(kW·h), based on the parameter values provided in Table 4.

Table 4. Parameter values.

ID	Variable	Variable Description	Value
1	F_E	Gasoline energy factor	92 # gasoline:0.1161
2	i_{ch}	Charge efficiency	100%
3	i_{tr}	Line loss rate	6.34%
4	φ	The proportion of thermal power generation	75.2%
5	s_{ge}	Power supply efficiency	38.63%
6	r_P	Refinery efficiency	92.8%
7	t_P	Delivery and filling efficiency	95%
8	T_E	Standard coal consumption for thermal power supply	0.318
9	T_C	Carbon dioxide emission factor of gasoline coal	3.09
10	t_M	Discount coefficient of gasoline coal and standard coal	1.07
11	T_F	Carbon dioxide emission factor of the gasoline	92 # gasoline: 2.38 kg/L

The density of 92# gasoline is referenced from GB 17930 [73], taken as 720 kg/m³; the average lower calorific value of gasoline is referenced from GB/T 2589 [74], taken as 43,070.

Step3: Carbon emissions calculation base on the equivalent fuel consumption

According to the IPCC carbon emission calculation method, based on GB/T37340 “Calculation Method for Energy Consumption of Electric Vehicles” [65] and GB 19578-2021 “Fuel Consumption Limits for Passenger Cars.” [72], the carbon emission of electric vehicles is calculated according to the actual energy consumption distance, as shown in following.

$$F_{j,b} = D_{j,b} \times FC_{CO_2} \times EF_k \quad (9)$$

where $F_{j,b}$ represents the total carbon emission of the order j of the electric taxi b from the starting point to the destination; $D_{j,b}$ represents the order j travel distance of the electric taxi b .

Given that there are no significant differences in the models and engine of electric taxis, the same emission coefficient and electricity consumption values are used to calculate the carbon emissions. The WLTC electricity consumption for the standard electric taxi (Geely Dihao new energy edition) in Chengdu is 11.9 kWh/100 km. The equivalent carbon dioxide emission factor for fuel consumption is 3.689 L/100 km, and the carbon emission factor is 8.1158 kg/100 km.

According to the different carbon emission calculation methods for gasoline vehicles and electric vehicles, within a complete trajectory of an order, each pair of trajectory positioning points is considered a carbon emission unit. The carbon emission of each unit is allocated to the road section [33,56,75] or area grid [32,34]. In this paper, we divide the study area into numerous grids with 500-500 m, and then the CO₂ emissions of trajectory unit are allocated to each grid. Finally, the grid carbon emissions (GCE) are calculated by summing up all the carbon emissions unit on the grid. The principle of grid allocation of taxi carbon emissions is illustrated in the Figure 7.

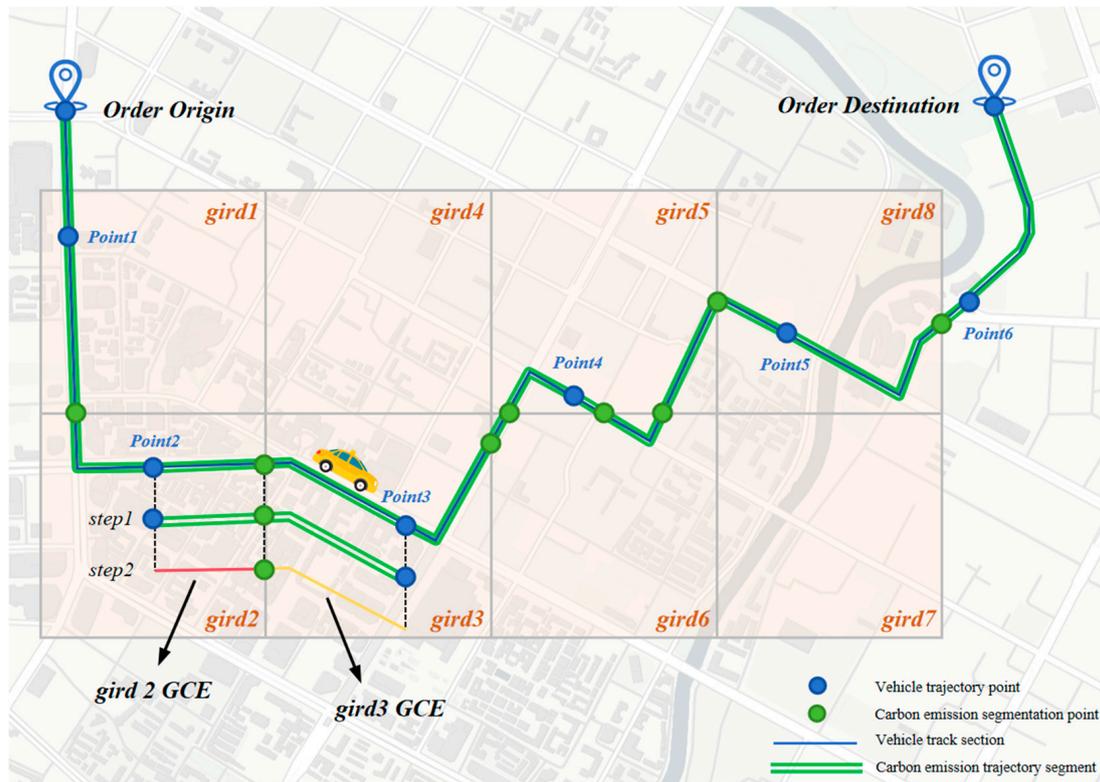


Figure 7. Taxi trajectory carbon emission grid allocation.

4.4. Non-Linear Geographic Regression Model

4.4.1. XGboost

XGBoost (eXtreme gradient boosting) represents an efficient implementation of the gradient boosting algorithm and serves as an enhancement of the gradient boosting decision tree (GBDT) algorithm. XGBoost's improvements are specifically manifested in its parallel processing design, regularization modules, and loss function optimization. Compared to GBDT, the regularization modules reduce the likelihood of overfitting, while the optimized loss functions yield superior predictive performance. Due to its parallel processing architecture, XGBoost also demonstrates enhanced computational efficiency, enabling rapid processing of large datasets and meeting the performance requirements of this modeling task.

The XGBoost computation method involves calculating the value of each data sample on every regression tree, and the final prediction for the sample is obtained by summing up the values from each tree. The calculation formula for the XGBoost model is shown in Equation (9).

$$\hat{y}_i = \sum_{k=1}^K f_k(x_i) \quad (10)$$

Here \hat{y}_i is the predicted value of the model, x_i represents the eigenvector of the sample i , and is the expression of the tree k .

The objective function of the XGBoost model is shown in Equation (11).

$$Obj = \sum_{i=1}^n L(y_i, \hat{y}_i) + \sum_{k=1}^K \Omega(f_k) \quad (11)$$

Here \hat{y}_i is the predicted value of the sample i , n represents the number of samples, y_i is the observed value of the sample i , k represents the number of trees, and f_k is the

expression of the tree k ; $L(y_i, \hat{y}_i)$ represents the loss function, $\Omega(f_k)$ is the regularization function.

The XGBoost algorithm decomposes the objective function using a Taylor expansion, relying on the first and second derivatives of the loss function at each data point. This approach allows XGBoost to train in a parallel manner, significantly improving the efficiency of the training process. Its form is shown in Equation (12).

$$Obj^{(t)} = \sum_{i=1}^n L \left[\left(y_i, \hat{y}_i^{(t-1)} \right) + g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i) \right] + \Omega(f_t) + \text{const} \quad (12)$$

g_i is the first partial derivative of the loss function L , and h_i is the second partial derivative of the loss function L . Just know the $\Omega(f_t)$, to get the result.

$\Omega(f_t)$ is a regular term of the model. The regularization term in the XGBoost model is implemented by adding the complexity of the tree as a regularization term to the objective function and pruning the tree in the later stages of model training to control the complexity of the model. This can result in a simpler final trained model, effectively preventing overfitting. The formula for the regularization term is shown in Equation (13):

$$\Omega(f_t) = \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T \omega_j^2 \quad (13)$$

In the expression, f_t is the expression for the tree t , T represents the number of leaf nodes in the tree t , ω_j represents the score on the leaf node j , γ and λ is the penalty factor. The larger their values, the greater the penalty for the complexity of the tree. Incorporating the regularization term into the objective function yields the final objective function as shown in Equation (14).

$$\begin{aligned} Obj^{(t)} &= \sum_{i=1}^n [g_i \omega_{q(x_i)} + \frac{1}{2} h_i \omega_{q(x_i)}^2] + \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T \omega_j^2 \\ &= \sum_{j=1}^T \left[\left(\sum_{i \in I_j} g_i \right) \omega_j + \frac{1}{2} \left(\sum_{i \in I_j} h_i + \lambda \right) \omega_j^2 \right] + \gamma T \end{aligned} \quad (14)$$

According to the objective function formula, I_j is the leaf node j corresponding to the sample set. The optimal solution obtained is follows:

$$\omega = - \frac{\sum_{i \in I_j} g_i}{\sum_{i \in I_j} h_i + \lambda}, \quad Obj = - \frac{1}{2} \sum_{j=1}^T \frac{\left(\sum_{i \in I_j} g_i \right)^2}{\sum_{i \in I_j} h_i + \lambda} + \gamma T \quad (15)$$

At the same time, XGBoost adopts a segmentation search algorithm that can sense the sparse features to adapt to the high-dimensional sparse features of variables. This algorithm can learn the sparsity of data and conduct parallelization learning, which enables XGBoost to reduce memory and obtain better training effects when computing.

4.4.2. Model Interpretation Using SHAP

Understanding the influence of various built environment factors and their combined effects is crucial for informing practical policy decisions. In this study, we utilize the SHAP

(Shapley additive explanations) model, an innovative additive explanation method rooted in game theory. SHAP values are computed as follows:

$$\varphi_i = \sum_{D \subseteq Q \setminus \{i\}} \frac{|D|!(Q - |D| - 1)!}{Q!} (P_{D \cup \{i\}}(\chi_{D \cup \{i\}}) - P_D(\chi_D)) \quad (16)$$

The SHAP scatter dependence plot shows how each independent variable affects traffic carbon emissions, accounting for the influence of other variables. Analyzing these plots helps determine whether these variables control carbon emissions within certain ranges.

5. Result

5.1. Spatiotemporal Analysis of Taxi Operations

In April 2022, an average of 12,078 taxis operated daily in Chengdu, with 4429 (36.67%) being traditional gasoline-powered vehicles and 7649 (63.33%) being electric vehicles (EVs). The traditional gasoline-powered taxis are primarily FAW-Volkswagen Jetta sedans with 1.5 L engines, while the electric taxis are exclusively EVs, primarily consisting of Geely Emgrand and Dongfeng Fukang E600 models.

Operational performance analysis shows that gasoline-powered taxis operate for an average of 17.75 h daily, EVs operate for 19.64 h. The average daily mileage per gasoline-powered taxi is 152.44 km, while for EVs, it is slightly higher at 160.23 km. Additionally, the average daily order count for gasoline-powered taxis is 11 trips, whereas EVs average 10.6 trips. When analyzing average trip distances, gasoline-powered taxis cover 13.97 km per order, while EVs average 14.99 km.

In terms of revenue, EV taxis generate an average daily income of approximately CNY 370, compared to CNY 338.8 for gasoline-powered taxis. EV taxis attract more passengers due to better conditions, smoother operation, and greater ride comfort, resulting in higher order counts and daily revenue.

As shown in Figure 8, an analysis of speed at various times of the day reveals that there is little difference between the average speeds of gasoline-powered and electric taxis during off-peak hours, with the speeds being 17.83 km/h and 16.66 km/h, respectively. Also, during peak hours, gasoline-powered taxis average 15.43 km/h, whereas electric vehicles (EVs) average 16.37 km/h. The maximum observed speed for gasoline-powered taxis within urban areas is 52.13 km/h, slightly lower than the 55.35 km/h maximum speed for EV taxis.

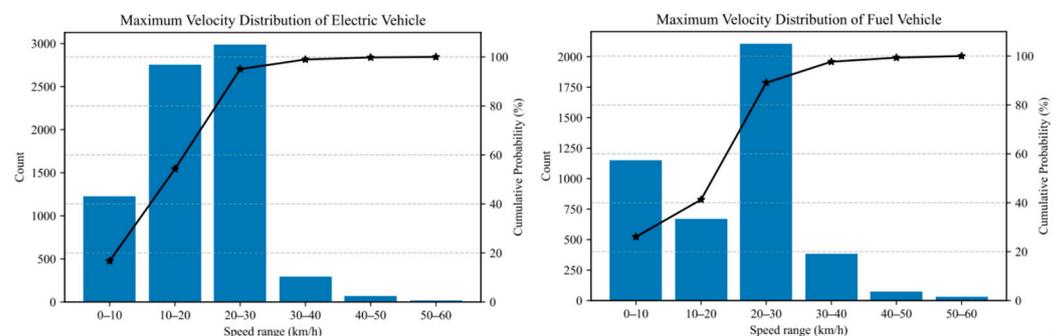


Figure 8. Description of taxi speed ranges.

Additionally, as shown in Figure 9, an analysis of the data sampling frequency for taxi GPS positioning reveals that electric taxis generate approximately 2000 data points per vehicle per day, with an average sampling interval of 32 s. In contrast, traditional gasoline-powered taxis have a stable data generation rate of around 2000 points per vehicle per

day, with a slightly longer average sampling interval of 40 s. The average interval between consecutive trajectory records for taxis is approximately 30 s.

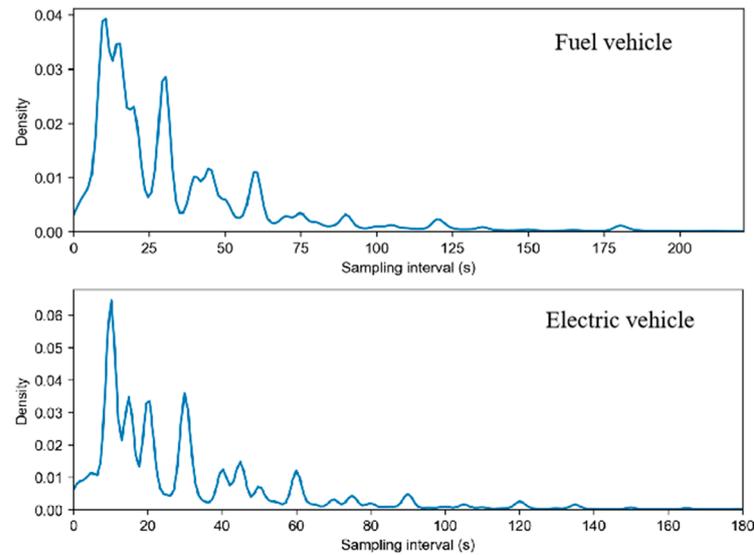


Figure 9. Statistical results of data sampling intervals.

Based on the aforementioned findings, electric taxis in Chengdu demonstrate potential advantages in operational efficiency and revenue generation compared to traditional gasoline-powered vehicles. This study recommends accelerating Chengdu's taxi electrification process and increasing financial subsidies to reduce operators' vehicle acquisition costs, with the goal of achieving complete electrification by 2025. This initiative would effectively leverage the advantages of electric taxis, enhance road traffic efficiency, facilitate low-carbon transition in the transportation sector, and promote high-quality urban transportation development.

5.2. Analysis of Carbon Emission Results

5.2.1. Time Distribution for Carbon Emission

Using the VSP model, the total carbon emissions on one typical workday in 2022 are estimated and measured, results are shown in Figure 10. The results of carbon emissions for two types of vehicles at different time periods show significant temporal variation, with multiple peaks and troughs. There is a carbon emission trough in the early morning (4:00–6:00), a peak in the morning (8:00–10:00), and another emission peak at night (21:00–23:00).

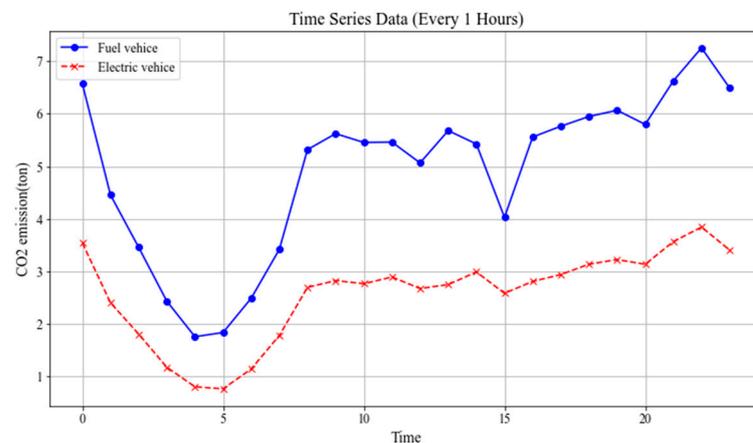


Figure 10. Time-varying diagram of carbon emissions.

The morning peak in emission is strongly correlated with commuting patterns, and the variation in carbon emissions from gasoline-powered taxis is more significant compared to that of new energy vehicles. The emission peak at night is related to leisure and entertainment activities, as Chengdu has many recreational parks and a well-developed night market. This period is the peak time for returning home after entertainment, and with public transportation services like buses and subways having ended, the demand for taxis increases, leading to a carbon emission peak.

Overall, vehicle carbon emissions exhibit significant temporal variation, with gasoline-powered taxi emissions significantly higher than those of new energy vehicles. Based on this finding, it is recommended to prioritize electric vehicle deployment during morning and evening peak hours when transport demand is concentrated, thereby reducing per-vehicle carbon emission intensity. This differentiated dispatch model enhances vehicle utilization efficiency, simultaneously reducing costs while maximizing emission reductions.

5.2.2. Spatial Distribution for Carbon Emission

Based on the model's calculations, Table 5 presents a statistical analysis of the total carbon emissions from gasoline and electric vehicles at various times of the day. When examining total carbon emissions across different periods, weekday morning peak hours show the highest emissions, followed by weekends, with weekday off-peak hours having the lowest emissions. This pattern aligns with real-world observations. The high carbon emissions during morning rush hours are driven by increased commuting demand. Similarly, weekends see elevated emissions due to leisure and entertainment activities that boost travel demand. In contrast, off-peak hours have lower emissions because of reduced travel demand.

Table 5. Statistics of total carbon emissions.

Period of Time	Electric Vehicle (kg)	Gasoline Vehicle (kg)	Total (kg)
Weekend	5292.57	10,540.71	15,833.28
Weekday Morning Peak	5458.74	10,806.50	16,265.24
Weekday Off-peak	5466.23	9273.13	14,739.36

In analyzing the differences in carbon emissions between gasoline and electric vehicles across various time periods, it is clear that gasoline vehicles remain the dominant mode of transportation, with their quantity being approximately twice those of electric vehicles. Additionally, the lower carbon emissions from gasoline vehicles during off-peak hours are likely due to decreased travel demand and favorable road conditions.

Spatial distribution maps of carbon emissions across different times of the day indicate that taxi emissions are primarily concentrated in central urban areas. Carbon emissions generally decrease from the city center towards peripheral areas, with a strong correlation to major and arterial roads. Specifically, high-emission zones are concentrated along the first, second, and third ring roads, which serve as major traffic arteries. The density of the road network is a key factor affecting the spatial distribution of carbon emissions.

(1) Weekend carbon emissions (14:00–16:00)

As shown in Figure 11, the carbon emissions from electric taxis are mainly concentrated in the city center, showing a pattern of high emissions in central areas and lower emissions towards the outskirts. This indicates that electric taxis contribute less to carbon emissions overall. In contrast, gasoline-powered taxis exhibit a pronounced concentration of emissions in the city center and along major traffic corridors, with high emissions

particularly along Tianfu Avenue, South Third Ring Road, and Shudu Avenue. Gasoline-powered taxis contribute significantly more carbon emissions than electric taxis.

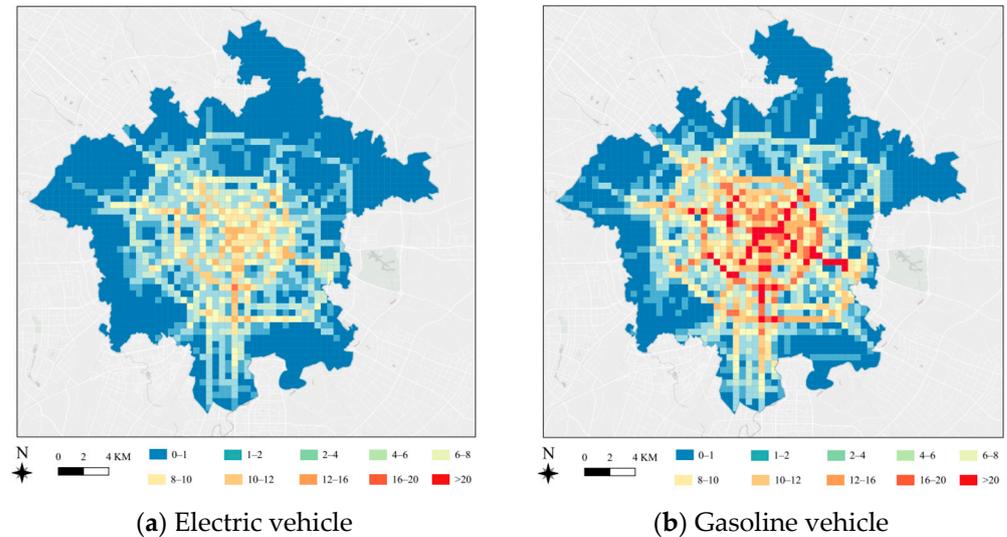


Figure 11. Spatial distribution of carbon emissions on weekends.

(2) Weekday morning peak carbon emissions (8:00–10:00)

Figure 12 is the spatial distribution of carbon emissions during weekday morning peak. As shown in Figure 12, carbon emissions patterns on weekday mornings are similar to those on weekends, with a strong correlation to the main road network. Carbon emissions are especially high along Tianfu Avenue, Shudu Avenue, and the elevated Second Ring Road.

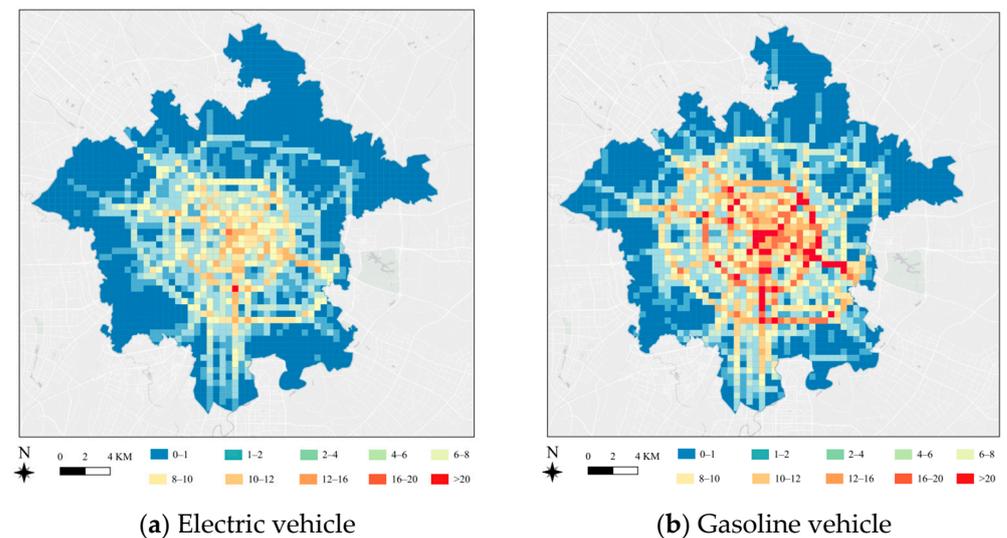


Figure 12. Spatial distribution of carbon emissions during weekday morning peak.

(3) Weekday off-peak carbon emissions (14:00–16:00)

Figure 13 is the spatial distribution of carbon emissions during weekday off-peak hours. As shown in Figure 13, compared to the weekday morning peak, off-peak emissions exhibit a different spatial pattern, with lower emissions in the northern and higher in the southern areas. Notable areas of high emissions include the major connecting roads, such as Tianfu Avenue and Shudu Avenue, as well as popular downtown destinations like Chunxi Road and Taikoo Li. This pattern reflects the travel demand for leisure and recreational activities among residents.

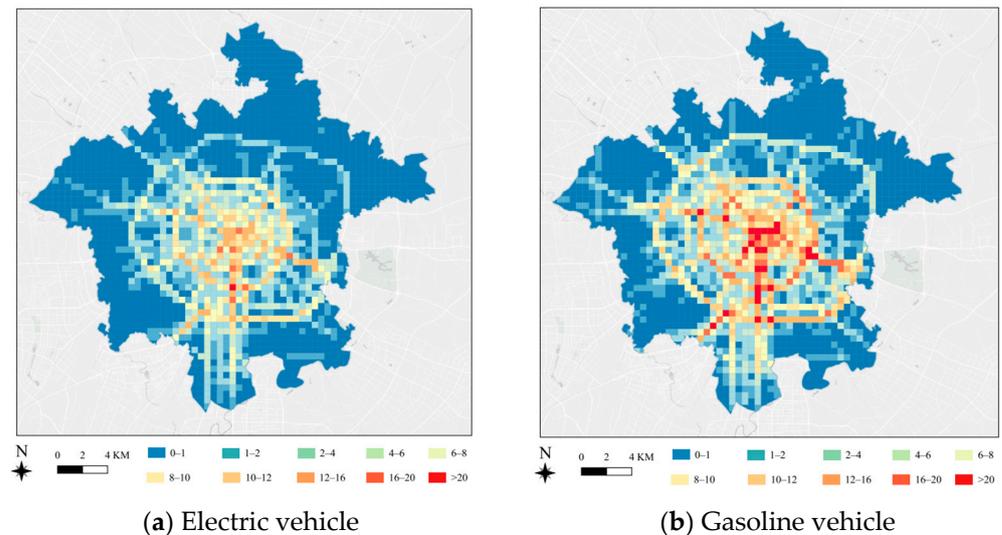


Figure 13. Spatial distribution of carbon emissions during weekday off-peak hours.

Analysis of spatial distribution patterns of carbon emissions between electric and gasoline-powered taxis reveals that gasoline-powered taxi emissions are highly concentrated along major urban arteries such as Tianfu Avenue and the Third Ring Road, as well as popular commercial districts like Chunxi Road and Taikoo Li. Electric vehicles demonstrate notably lower emission intensity in these areas. This indicates that in high-traffic, low-speed road segments, gasoline-powered vehicles generate higher carbon emissions due to frequent idling and low-speed operations.

For these hotspot areas, it is recommended that taxi platforms implement electric taxi voucher programs to increase the proportion of electric taxi operations in popular areas, thereby reducing inefficient operational modes of gasoline-powered vehicles such as empty cruising and idling.

5.3. Spatial Correlation Analysis of Carbon Emissions

5.3.1. Model Selection

Based on 5D elements, 13 variables representing the urban built environment were selected, and XGBoost, GBDT, and RF models are used to construct carbon emission prediction models for different periods in order to analyze the heterogeneity of factors affecting taxi emissions in different regions.

In this paper, we use mean squared error (MSE), root mean square error (RMSE), and R-squared (R^2) as metrics to assess the predictive performance of various models. We evaluate the performance of the gradient boosted regression trees (GBRT) model against random forest (RF) and XGBoost models. As shown in Table 6, the XGBoost model outperforms GBDT and RF, evidenced by its lower MSE and RMSE values and a higher R^2 . Therefore, we selected the XGBoost model for subsequent analysis.

Table 6. Comparison of Results from Multiple Machine Learning Models.

Model	RMSE	MSE	R^2
XGBoost	40.27	1621.62	0.6387
Random Forest	41.25	1701.63	0.6209
Gradient Boosting	40.85	1668.34	0.6283

To ensure the model's generalizability and prevent overfitting, we apply a grid search method and five-fold cross-validation to determine the optimal hyperparameters. In each

iteration, the original training samples are randomly divided into five subsets, with four subsets (80% of the data) used for training and the remaining subset (20%) used for testing.

The five key hyperparameters are: maximum tree depth (`max_depth`), the number of estimators (`n_estimators`), subsample ratio, learning rate, and column sample ratio per tree (`colsample_bytree`). The `max_depth` determines the maximum depth of trees and controls model complexity. Higher values can capture more complex relationships but risk overfitting, while lower values offer better generalization but may lead to underfitting. The `n_estimators` determines the number of trees. While more trees generally improve model performance, there are diminishing returns and increased training time. Too few trees may result in underfitting. The `colsample_bytree` determines the proportion of features randomly sampled for each tree, while the `subsample` determines the proportion of training instances randomly sampled for each tree. Lower values introduce randomness to reduce overfitting, but excessively low values might ignore important features. The `learning_rate` is the step size shrinkage used to prevent overfitting. Lower values require more iterations but typically yield better generalization performance, though extremely low values can lead to slow model training.

Based on the function and significance of XGBoost hyperparameters, and balancing performance with generalization capability, the parameter thresholds are set as follows: learning rate (0.001, 0.01, 0.05, 0.1), tree depth (3, 5, 7, 10), number of trees (100, 300, 500), subsample ratio (0.8, 0.9, 1.0), and column sample ratio per tree (0.8, 0.9, 1.0). The optimal hyperparameter combination obtained through grid search is shown in the Table 7.

Table 7. Hyperparameter selection for XGBoost model.

Max_Depth	N_Estimators	Colsample_Bytree	Learning_Rate	Subsample
5	500	0.9	0.01	0.80

The XGBoost model, retrained with these optimized parameters, exhibits a gradually decreasing RMSE during iterations, indicating continuous improvement in the training process, as shown in Figure 14. In the final prediction, the model achieves an R^2 of approximately 69.25%, demonstrating good fit and accurate target value predictions. Figure 15 compares the actual values with predicted values, showing close alignment and further validating the model's effectiveness.

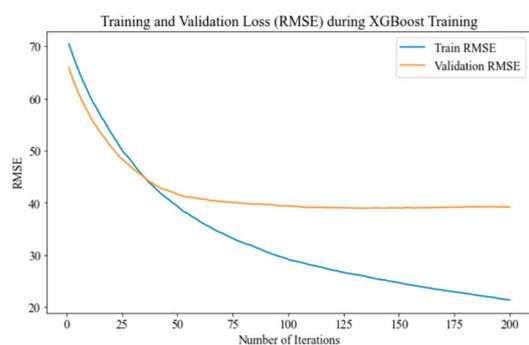


Figure 14. XGBoost training loss.

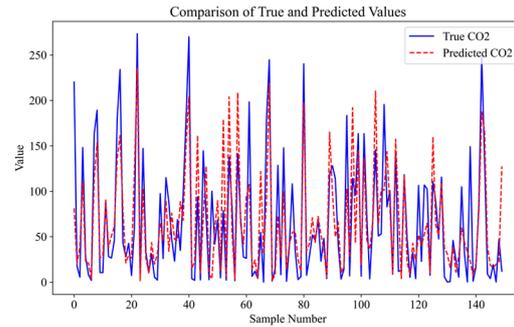


Figure 15. XGBoost prediction.

5.3.2. Variable Analysis

To clarify how each building environment variable affects carbon emission prediction, the characteristic importance diagram is generated as shown in Figure 16.

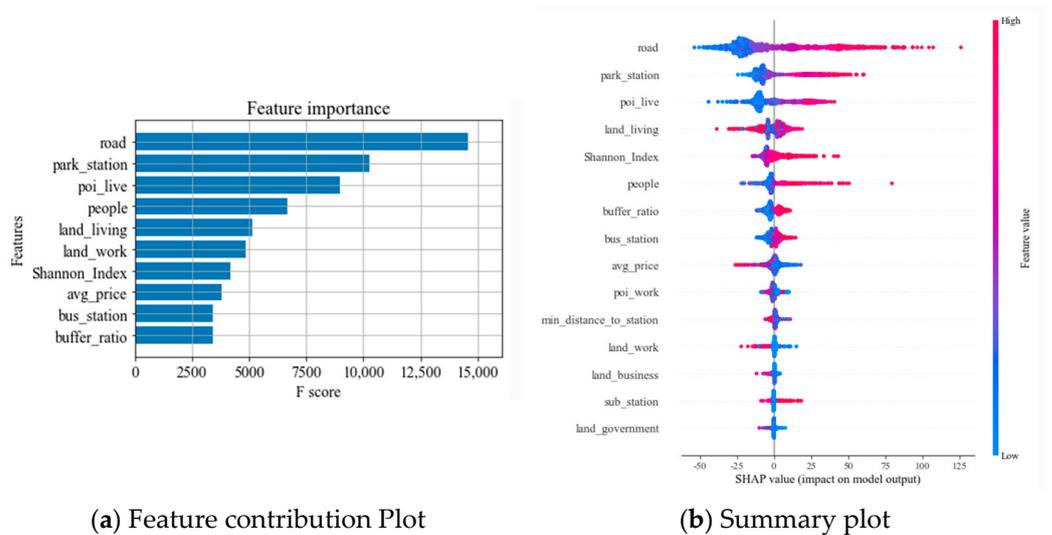


Figure 16. Feature importance plot.

Analysis of both the feature contribution plot and SHAP plot indicates that the features “road”, “park_station”, and “poi_live” are the most influential in the model’s predictions with high F-scores. Furthermore, the SHAP plot shows substantial variation in their impact values, which underscores their substantial contribution to the model’s output. Collectively, these features are essential for accurate model predictions and warrant further detailed examination and analysis. The force plot provides a detailed view of each feature’s specific contribution to the model’s prediction, as shown in Figure 17. In this visualization, red sections indicate features that positively influence the predicted value (i.e., positively correlated), while blue sections represent features that decrease the predicted value (i.e., negatively correlated). In this case, “park_station”, “poi_live”, and “road” contribute notably to the increase in predicted values, with “park_station” having the greatest impact. Conversely, features that reduce the predicted value have a lesser effect. Together, these features elevate the final predicted value to 175.47, higher than the baseline. This plot enables a straightforward understanding of each feature’s importance and directional influence in a specific prediction.



Figure 17. Force plot.

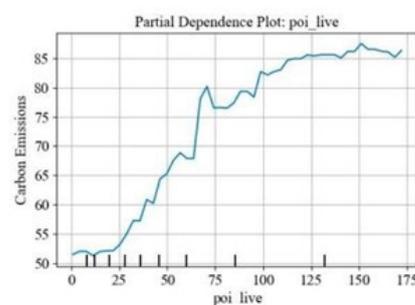
5.3.3. Nonlinear Impacts

The layout of urban roads, building density, and public transportation accessibility significantly influence residents’ travel choices and energy demand, impacting local carbon emissions. For instance, high-density development areas attract larger crowds and generate greater travel demand, which may result in traffic congestion and higher carbon emissions. By utilizing SHAP and XGBoost model, we can capture how these spatial heterogeneities affect carbon emissions, providing scientific insights for further analysis of localized carbon emission characteristics. This paper offers crucial support for optimizing public transportation planning and green building design, aiding in the effective reduction of carbon footprints. In this paper, five key variables are selected for visualization and analysis, with the dependence plot clearly demonstrating the impact of individual features on the model’s predictive outcomes.

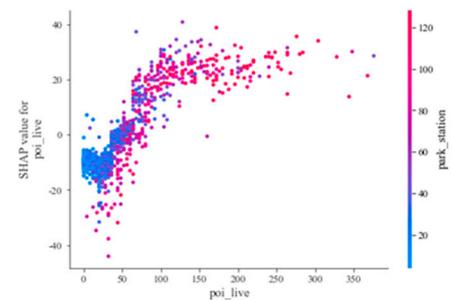
(1) POI live density

Overall, as residential POI density increases, taxi carbon emissions rise correspondingly. In areas where residential density is below 25 units/km², carbon emissions remain low and increase gradually, likely due to low demand for taxis in these less dense areas, as shown in Figure 18. When residential density exceeds 25 units/km², emissions significantly increase as taxi demand rises with higher population density. At densities around 100 units/km², emissions stabilize, likely due to road capacity limits and the diversion of travel demand to mass transit options like subways and buses, resulting in a saturation of taxi demand and a peak in carbon emissions.

Analysis reveals that a residential POI density of 25/km² represents a critical threshold. Above this value, travel demand increases significantly, leading to markedly higher taxi carbon emission intensity, indicating that taxi-dominated transportation systems become unsuitable. For high-density communities in main urban districts such as Chenghua, Jinniu, and Wuhou, it is essential to optimize public transit layout and enhance high-capacity public transportation services. This involves accelerating the development of a multi-level, integrated public transit system to encourage modal shift from taxis to public transportation, thereby alleviating congestion and carbon emissions.

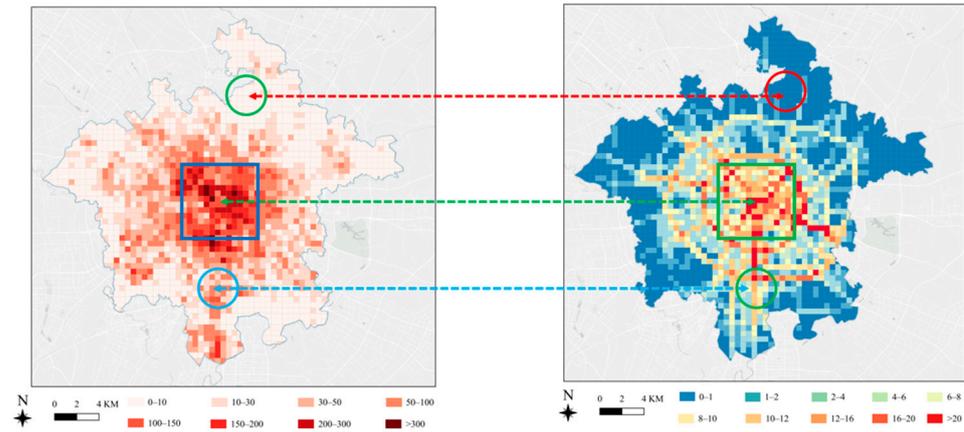


(a) Continuous dependence plot



(b) Discrete dependence plot

Figure 18. Cont.



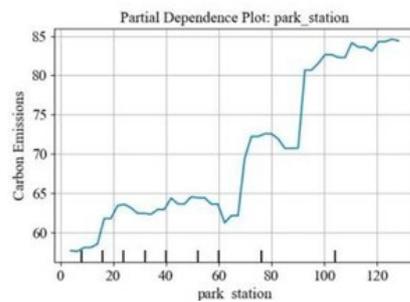
(c) Spatial distribution of Poi_live

Figure 18. Non-linear Impact analysis of Poi_live.

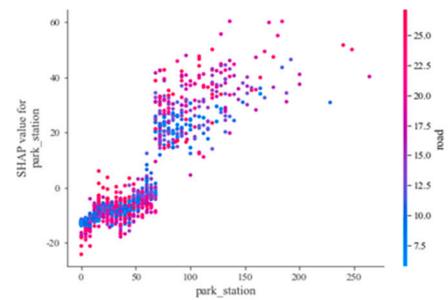
(2) Parking lot density

An increase in parking availability leads to a rise in taxi carbon emissions, which has a noticeable threshold effect. When parking lot density reaches 64 lots/km² and 88 lots/km², carbon emissions experience significant jumps, as shown in Figure 19. This indicates that within certain density ranges, increased parking supply induces greater taxi travel demand, resulting in rapid carbon emission growth. Urban parking policies should carefully consider this non-linear effect and rationally control parking supply in key areas to prevent uncontrolled carbon emissions.

It is found that the density of parking lots decreases from the city center to the outskirts, indicating a strong correlation between parking lot density and urban development intensity. High-density parking areas also mean higher development and stronger travel attraction, resulting in increased taxi demand and thus higher CO₂ emissions. When parking lot density exceeds 100 per km², the attraction of developed areas to private cars increases, thereby suppressing the growth in taxi demand, resulting in stabilized carbon emissions.

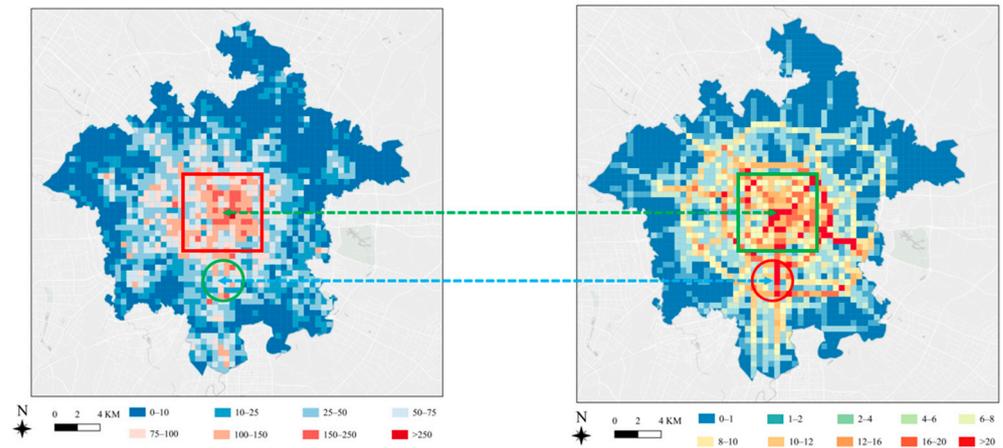


(a) Continuous dependence Plot



(b) Discrete dependence plot

Figure 19. Cont.



(c) Spatial distribution of parking lot density

Figure 19. Non-linear impact analysis of parking lot density.

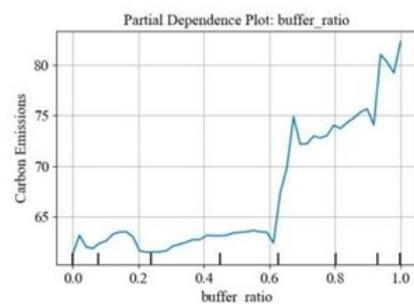
Future strategies should strengthen parking demand management in central urban areas, such as moderate parking fee adjustments, to encourage adoption of green transportation modes like buses and metros.

(3) Subway buffer zone ratio

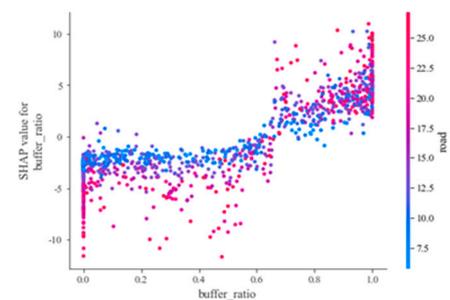
In areas where the subway buffer zone ratio is below 0.6, carbon emissions are relatively stable and low, as shown in Figure 20. This stability may be attributed to the fact that these areas are farther from subway stations, leading residents to rely on other transportation modes rather than taxis for their journeys. Consequently, this results in lower demand and activity for taxis and hence lower carbon emissions. However, when the buffer zone ratio reaches 0.6, a significant increase in emissions occurs, likely indicating strong passenger dependence on taxis for metro connectivity.

In addition, the large buffer zone around the subway station is usually prone to congestion, where taxis driving at a low speed will increase fuel consumption and carbon emissions. As the subway buffer zone ratio increases, carbon emissions increase, which illustrates this trend.

This research recommends improving slow-traffic systems around metro stations to provide more green connection options. Actively promoting flexible “micro-mobility” solutions like shared bicycles and electric scooters can create seamless green transportation networks connecting with metro systems, reducing excessive taxi dependence.

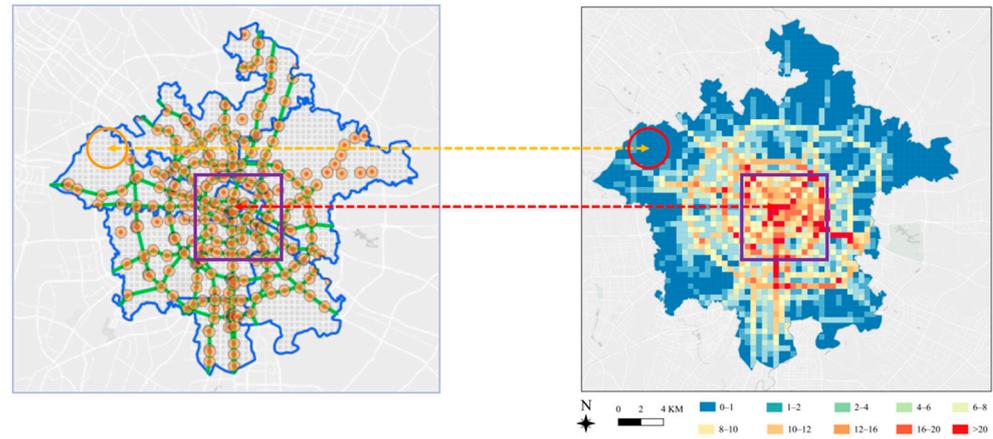


(a) Continuous dependence plot



(b) Discrete dependence plot

Figure 20. Cont.

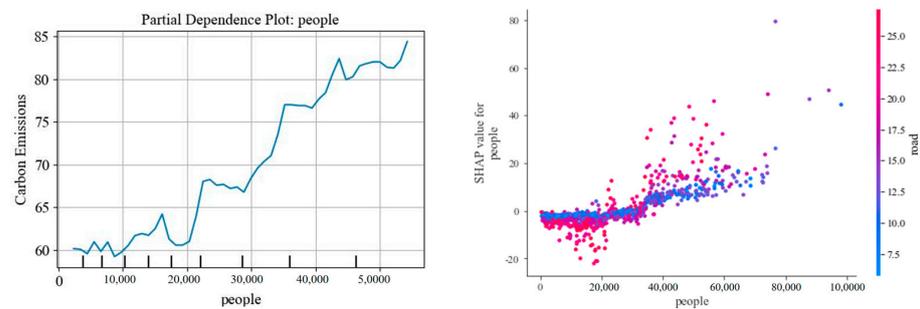


(c) Spatial distribution of subway buffer zone ratio

Figure 20. Non-linear impact analysis of subway buffer zone ratio.

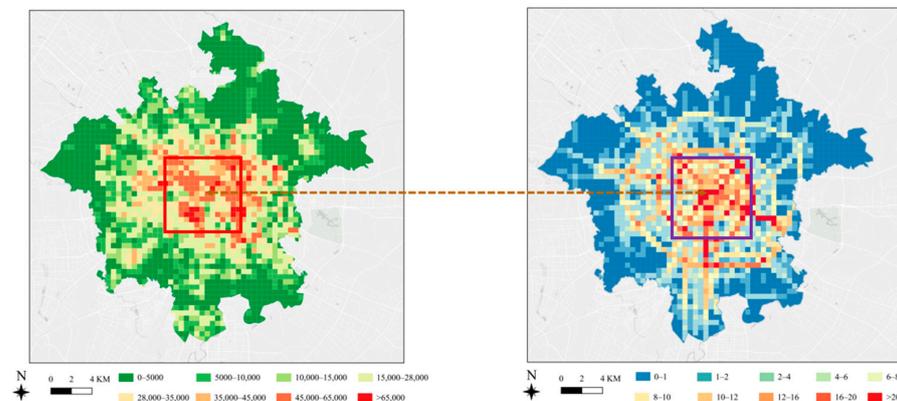
(4) Population density

Increased population density is closely linked to higher carbon emissions, as it elevates travel demand and subsequently raises carbon emissions. Overall, emissions increase with population density, indicating that areas with higher population density have more frequent taxi usage, raising trip frequency and driving distances, which causes higher carbon emissions. Notably, when population density surpasses 20,000, the rate of increase in emissions accelerates, suggesting that carbon emissions grow faster beyond a certain population density threshold, as shown in Figure 21. At densities above 50,000, emission growth tends to stabilize, potentially due to the marginal effect of carbon reduction measures or the intervention of alternative transportation modes. Based on the analysis, this study proposes the following policy recommendations:



(a) Continuous dependence plot

(b) Discrete dependence plot



(c) Spatial distribution of subway population density

Figure 21. Non-linear impact analysis of population density.

Focus on high-density areas (population density above 20,000) for priority emission reduction measures, implementing targeted interventions such as accelerating taxi electrification and optimizing vehicle dispatch efficiency through big data.

Additionally, improving public transit services in high-density areas, enhancing bus and metro system coverage and service quality, and optimizing first/last-mile connections with transit hubs can effectively share taxi travel demand. Beyond “supply-side” interventions, policymakers should consider “demand-side” transportation management measures in densely populated areas, such as encouraging carpooling to increase occupancy rates.

(5) Road network density

The density of road networks, particularly main roads and expressways, significantly increases carbon emissions. Higher road density improves regional accessibility, thus inducing more traffic demand, which results in greater carbon emissions, Figure 22c intuitively reflects this conclusion. Based on this, we recommend scientific road network planning to avoid excessive expansion. In central urban areas with already high road network density, focus should be on improving traffic efficiency and optimizing existing network resources rather than continuous expansion.

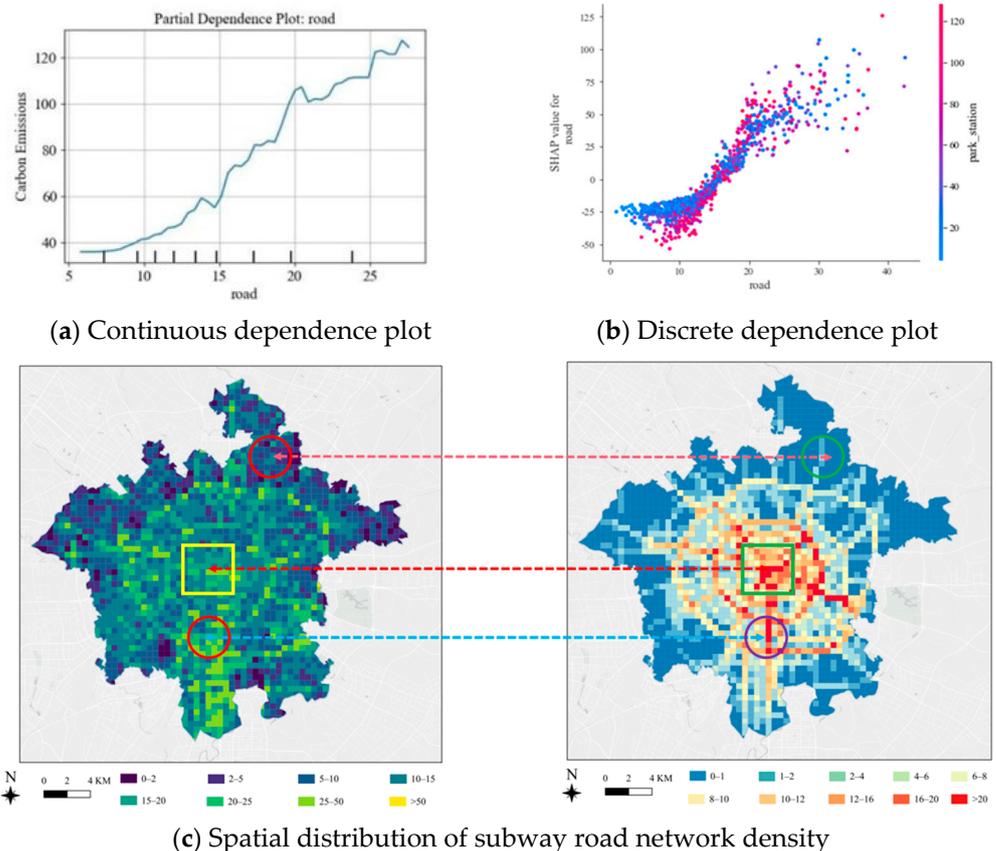


Figure 22. Non-linear impact analysis of road network density.

Specifically, carbon emissions increase as road density rises, but different density ranges show varying rates of increase: at lower densities (5–10 km/km²), the rise in emissions is moderate; however, starting from 10 km/km², the emission growth rate sharply accelerates, likely due to a sudden increase in traffic activity beyond a certain road density threshold. At high densities close to 20 km/km², the increase in emissions exhibits some fluctuations, potentially due to factors like road congestion and reduced vehicle speeds. To mitigate this situation, vehicular management policies can be implemented to reduce taxi proportions, thereby decreasing carbon emissions. It is recommended to establish ded-

icated bus lanes or reversible lanes on major and secondary arterial roads with high emission intensity, while prioritizing bus passage through traffic signal systems to enhance transportation efficiency and reduce emissions.

Table 8 clearly presents various critical thresholds based on research findings and their associated policy recommendations.

Table 8. Policy recommendations from key findings.

Key Findings	Policy Recommendations
Residential POI density: 25/km ²	Optimize public transportation layout
Parking facility density: 64/km ²	Rationally control parking supply in key areas
Metro buffer zone coverage ratio > 0.6	Strengthen parking demand management in central urban districts
Population density > 20,000/km ²	Improve slow-traffic systems around metro stations
Road network density > 10 km/km ²	Prioritize emission reduction measures
	Scientific road network planning to avoid excessive expansion

5.4. Emission Reduction Benefits

By the end of 2020, Chengdu had 5.98 million motor vehicles, and the average daily travel volume in the central urban area reached 27.9 million trips. While this has brought tremendous pressure on urban traffic operations, it has also result in a continuous increase in urban transportation carbon emissions and a worsening of regional air pollution.

According to data from the Chengdu Environmental Protection Science Research Institute, the contribution rate of motor vehicle emissions to carbon dioxide (CO₂) among local emission sources in Chengdu was 31% between 2019 and 2020. The transportation emissions have gradually become the fastest-growing source of carbon emissions in the city.

The Chengdu Green and Low-Carbon Development Report (2022) indicates that the annual market penetration rate of new energy vehicles (NEVs) in Chengdu has reached 31%. According to a document from the Chengdu Environmental Protection Bureau, Chengdu's cruising taxis will achieve full electrification by 2025. Additionally, the Special Plan for Electric Vehicle Charging and Swapping Infrastructure in Chengdu (2023–2025), issued by the Chengdu Economic and Information Technology Bureau, states that under stable growth conditions, logistics delivery vehicles, buses, and cruising taxis in Chengdu's central urban area will essentially transition to new energy vehicles by 2025. Taxis in Chengdu are predicted to number between 15,000 and 17,000 by 2025.

Furthermore, the granular carbon emission calculation methodology established in this study enables precise tracking of emission reductions achieved through electrification initiatives. This methodology has significant implications for policy refinement, transparent emission reduction reporting, and quantifiable demonstration of environmental benefits to stakeholders. After full electrification, compared to the current 1:2 ratio of gasoline to electric vehicles, it is conservatively estimated that the daily reduction in emissions will reach 99.04 tons, assuming an average of 16,000 operating vehicles per day and an average daily mileage of 180 km per vehicle. The environmental benefits will be significant.

6. Discussion and Conclusions

Lowering carbon emissions from road traffic is crucial for mitigating global warming. Using taxis trajectory data, this paper proposes a bottom-up CO₂ emission calculation method for traditional fuel vehicles and electric vehicles at the road level and analyzes their spatiotemporal variation characteristics. Building on this, this paper employs

XGBoost and SHAP explanation models to analyze the impact of the built environment on road carbon emissions.

This paper successfully identifies the spatiotemporal characteristics and emission hotspots of road carbon emissions in Chengdu. CO₂ emissions rise between 06:00 and 09:00 and peak between 15:00 and 18:00. Emission hotspots are primarily concentrated on major urban roads, such as Tianfu Avenue, the Second Ring Road, and the Third Ring Road. This paper examines how elements of the built environment influence road carbon emissions at the grid level and explains the relationships between various influencing factors. The results indicate that workplaces, main roads, residences, and bus stations have significantly positive impact coefficients, while population density has a significantly negative impact coefficient. Moreover, access to subway stations has a bidirectional impact on road carbon emissions.

This paper highlights the significant impact of urban spatial layout and transportation infrastructure on the spatiotemporal distribution of road carbon emissions. Specifically, the clustering of commercial and residential areas leads to a marked increase in carbon emissions. By optimizing regional characteristics and improving the diversity and coordination of internal facilities, carbon emissions can be effectively reduced. Additionally, this paper reveals that carbon emissions are notably higher on major roads and ring roads, a result primarily attributed to the combination of concentric and radial road networks in Chengdu.

For future urban land use and spatial development policies in Chengdu, the following recommendations are proposed: First, optimize the distribution of population by encouraging the relocation of certain facilities and industries from the city center to new urban districts, such as government offices, universities, and high-tech industries. These new districts, located in the eastern and southern parts of Chengdu, have been developed in recent years and are equipped with residential, commercial, and public service infrastructure, along with comprehensive public transportation services, including bus and rail systems. Connecting these newly developed districts to the city center via major roads and ring roads can not only reduce road carbon emissions in the central urban area but also foster rapid development in the new districts, promote a balanced industrial distribution, and narrow regional disparities. Second, in the central urban areas with high population density and development intensity, it is recommended to reduce traffic-related carbon emissions by improving road traffic efficiency, creating street parks and recreational green spaces, building cycling and walking greenways, developing shared slow transport systems, expanding rail transit, and optimizing the balance between public and private transportation. These measures can help lower carbon emissions during peak hours and mitigate traffic congestion.

Current research mainly focuses on the spatiotemporal distribution of urban traffic carbon emissions and the influence of the built environment, using data from travel surveys, ride-hailing, and taxi trajectories [70,71]. However, with the widespread adoption of electric vehicles in China, studies have yet to fully address the carbon emission distribution patterns and influencing factors of electric vehicles. This paper integrates multi-day trajectory data from both conventional gasoline taxis and electric taxis to investigate their respective carbon emission patterns. It quantitatively estimates the carbon reduction benefits of electric vehicles and analyzes the impact of the built environment on the emissions from both vehicle types. Existing carbon emission estimation methods, such as IVE and VSP [68,75], are highly sensitive to input data, and fine-grained data significantly improves accuracy. In this paper, vehicle emissions are estimated based on trajectory data and HHM road matching, minimizing errors from location deviations and incorrect linkages, ensuring the reliability and precision of the built environment impact analysis.

This paper can be further refined. For example, multi-source data such as street cameras and sensors can be used to improve the accuracy of model results. More comprehensive GPS trajectory data and vehicle type information would deepen the analysis of carbon emission characteristics. In addition, other greenhouse gases can also be analyzed in depth. Future research could concentrate on addressing these issues.

Author Contributions: Conceptualization, G.L.; Data curation, X.S.; Funding acquisition, G.L.; Investigation, K.L.; Methodology, D.X., X.S., K.L., J.L. and G.L.; Project administration, G.L.; Software, D.X., X.S. and J.L.; Visualization, D.X. and J.L.; Writing—original draft, D.X., X.S. and K.L.; Writing—review & editing, D.X. and J.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no funding.

Data Availability Statement: The data that support the findings of this study are available on request from the corresponding author, upon reasonable request.

Acknowledgments: The authors would like to thank the editors and reviewers for their valuable comments and suggestions on this paper.

Conflicts of Interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. Benzie, M.; Carter, T.R.; Carlsen, H.; Taylor, R. Cross-border climate change impacts: Implications for the European Union. *Reg. Environ. Change* **2019**, *19*, 763–776. [CrossRef]
2. Pla, M.A.M.; Lorenzo-Sáez, E.; Luzuriaga, J.E.; Prats, S.M.; Moreno-Pérez, J.A.; Urchueguía, J.F.; Oliver-Villanueva, J.V.; Lemus, L.G. From traffic data to GHG emissions: A novel bottom-up methodology and its application to Valencia city. *Sustain. Cities Soc.* **2021**, *66*, 102643.
3. Shi, X.; Zheng, Y.; Lei, Y.; Xue, W.; Yan, G.; Liu, X.; Cai, B.; Tong, D.; Wang, J. Air quality benefits of achieving carbon neutrality in China. *Sci. Total Environ.* **2021**, *795*, 148784. [CrossRef] [PubMed]
4. Solaymani, S. CO₂ emissions patterns in 7 top carbon emitter economies: The case of transport sector. *Energy* **2019**, *168*, 989–1001. [CrossRef]
5. Zhou, X.; Zhao, R.; Cheng, L.; Min, X. Impact of policy incentives on electric vehicles development: A system dynamics-based evolutionary game theoretical analysis. *Clean Technol. Environ.* **2019**, *21*, 1039–1053. [CrossRef]
6. Ahmed, Z.; Ali, S.; Saud, S.; Shahzad, S.J.H. Transport CO₂ emissions, drivers, and mitigation: An empirical investigation in India. *Air Qual. Atmos. Health* **2020**, *13*, 1367–1374. [CrossRef]
7. Liu, Y.; Yuan, Y.; Guan, H.; Sun, X.; Huang, C. Technology and threshold: An empirical study of road passenger transport emissions. *Res. Transp. Bus. Manag.* **2021**, *38*, 100487. [CrossRef]
8. State Council. New Energy Vehicle Industry Development Plan (2021–2035). 2020. Available online: http://www.gov.cn/zhengce/content/2020-11/02/content_5556716.htm (accessed on 21 December 2021).
9. Tu, R.; Gai, Y.; Farooq, B.; Posen, D.; Hatzopoulou, M. Electric vehicle charging optimization to minimize marginal greenhouse gas emissions from power generation. *Appl. Energy* **2020**, *277*, 115517. [CrossRef]
10. Frey, H. Trends in onroad transportation energy and emissions. *J. Air Waste Manag. Assoc.* **2018**, *68*, 514–563. [CrossRef]
11. Peng, W.; Yang, J.; Lu, X.; Mauzerall, D. Potential co-benefits of electrification for air quality, health, and CO₂ mitigation in 2030 China. *Appl. Energy* **2018**, *218*, 511–519. [CrossRef]
12. Sun, D.J.; Zhang, K.; Shen, S. Analyzing spatiotemporal traffic line source emissions based on massive Didi online car-hailing service data. *Transp. Res. Part D Transp. Environ.* **2018**, *62*, 699–714. [CrossRef]
13. Cao, X.; Yang, W. Examining the effects of the built environment and residential self-selection on commuting trips and the related CO₂ emissions: An empirical study in Guangzhou, China. *Transp. Res. Part D Transp. Environ.* **2017**, *52*, 480–494. [CrossRef]
14. Choi, K.; Zhang, M. The net effects of the built environment on household vehicle emissions: A case study of Austin, TX. *Transp. Res. Part D Transp. Environ.* **2017**, *50*, 254–268. [CrossRef]
15. Alfaaseh, L.; Tu, R.; Farooq, B.; Hatzopoulou, M. Greenhouse gas emission prediction on road network using deep sequence learning. *Transp. Res. Part D Transp. Environ.* **2020**, *88*, 102593. [CrossRef]
16. Tao, T.; Wang, J.; Cao, X. Exploring the non-linear associations between spatial attributes and walking distance to transit. *J. Transp. Geogr.* **2020**, *82*, 102560. [CrossRef]

17. Zhang, W.; Zhao, Y.; Cao, X.J.; Lu, D.; Chai, Y. Nonlinear effect of accessibility on car ownership in Beijing: Pedestrian-scale neighborhood planning. *Transp. Res. Part D Transp. Environ.* **2020**, *86*, 102445. [[CrossRef](#)]
18. Liu, J.; Wang, B.; Xiao, L. Non-linear associations between built environment and active travel for working and shopping: An extreme gradient boosting approach. *J. Transp. Geogr.* **2021**, *92*, 103034. [[CrossRef](#)]
19. Yang, W. The nonlinear effects of multi-scale built environments on CO₂ emissions from commuting. *Transp. Res. Part D Transp. Environ.* **2023**, *118*, 103736. [[CrossRef](#)]
20. Wu, J.; Li, C. Illustrating the nonlinear effects of urban form factors on transportation carbon emissions based on gradient boosting decision trees. *Sci. Total Environ.* **2024**, *929*, 172547. [[CrossRef](#)]
21. Yang, W.; Zhou, S. Using decision tree analysis to identify the determinants of residents' CO₂ emissions from different types of trips: A case study of Guangzhou, China. *J. Clean. Prod.* **2020**, *277*, 124071. [[CrossRef](#)]
22. Wu, X.; Tao, T.; Cao, J.; Fan, Y.; Ramaswami, A. Examining threshold effects of built environment elements on travel-related carbon-dioxide emissions. *Transp. Res. Part D Transp. Environ.* **2019**, *75*, 1–12. [[CrossRef](#)]
23. Gao, J.; Ma, S.; Peng, B.; Zuo, J.; Du, H. Exploring the nonlinear and asymmetric influences of built environment on CO₂ emission of ride-hailing trips. *Environ. Impact Assess. Rev.* **2021**, *92*, 106691. [[CrossRef](#)]
24. Song, J.; Fan, H.; Gao, M.; Xu, Y.; Ran, M.; Liu, X.; Guo, Y. Toward high-performance map-recovery of air pollution using machine learning. *ACS EST Eng.* **2022**, *3*, 73–85. [[CrossRef](#)]
25. Song, J.; Han, K.; Stettler, M.E. Deep-MAPS: Machine-learning-based mobile air pollution sensing. *IEEE Internet Things J.* **2020**, *8*, 7649–7660. [[CrossRef](#)]
26. Colville, R.N.; Hutchinson, E.J.; Mindell, J.S.; Warren, R.F. The transport sector as a source of air pollution. *Atmos. Environ.* **2001**, *35*, 1537–1565. [[CrossRef](#)]
27. Alam, M.S.; Duffy, P.; Hyde, B.; McNabola, A. Downscaling national road transport emission to street level: A case study in Dublin, Ireland. *J. Clean. Prod.* **2018**, *183*, 797–809. [[CrossRef](#)]
28. Singh, N.; Mishra, T.; Banerjee, R. Greenhouse gas emissions in India's road transport sector. In *Climate Change Signals and Response*; Springer: Singapore, 2019; pp. 197–209.
29. Sun, S.; Sun, L.; Liu, G.; Zou, C.; Wang, Y.; Wu, L.; Mao, H. Developing a vehicle emission inventory with high temporal-spatial resolution in Tianjin, China. *Sci. Total Environ.* **2021**, *776*, 145873. [[CrossRef](#)]
30. De Nunzio, G.; Laraki, M.; Thibault, L. Road traffic dynamic pollutant emissions estimation: From macroscopic road information to microscopic environmental impact. *Atmosphere* **2021**, *12*, 53. [[CrossRef](#)]
31. Zhang, L.; Long, R.; Chen, H.; Geng, J. A review of China's road traffic carbon emissions. *J. Clean. Prod.* **2019**, *207*, 569–581. [[CrossRef](#)]
32. Luo, X.; Dong, L.; Dou, Y.; Zhang, N.; Ren, J.; Li, Y.; Sun, L.; Yao, S. Analysis on spatial-temporal features of taxis' emissions from big data informed travel patterns: A case of Shanghai, China. *J. Clean. Prod.* **2017**, *142*, 926–935. [[CrossRef](#)]
33. Kan, Z.; Tang, L.; Kwan, M.P.; Ren, C.; Liu, D.; Pei, T.; Liu, Y.; Deng, M.; Li, Q. Fine-grained analysis on gasoline-consumption and emission from vehicles trace. *J. Clean. Prod.* **2018**, *203*, 340–352. [[CrossRef](#)]
34. Liu, J.; Han, K.; Chen, X.M.; Ong, G.P. Spatial-temporal inference of urban traffic emissions based on taxi trajectories and multi-source urban data. *Transp. Res. Part C Emerg.* **2019**, *106*, 145–165. [[CrossRef](#)]
35. Duarte, G.O.; Gonçalves, G.A.; Baptista, P.C.; Farias, T.L. Establishing bonds between vehicle certification data and real-world vehicle fuel consumption—A vehicle specific power approach. *Energy Convers. Manag.* **2015**, *92*, 251–265. [[CrossRef](#)]
36. Chong, H.S.; Park, Y.; Kwon, S.; Hong, Y. Analysis of real driving gaseous emissions from light-duty diesel vehicles. *Transp. Res. Part D Transp. Environ.* **2018**, *65*, 485–499. [[CrossRef](#)]
37. Zhang, J.L.; Chen, F.; Wang, Z.J.; Wang, R.; Shi, S.W. Spatiotemporal patterns of carbon emissions and taxi travel using GPS data in Beijing. *Energies* **2018**, *11*, 500. [[CrossRef](#)]
38. Xia, C.Y.; Xiang, M.; Fang, K.; Li, Y.; Ye, Y.; Shi, Z.; Liu, J. Spatial-temporal distribution of carbon emissions by daily travel and its response to urban form: A case study of Hangzhou, China. *J. Clean. Prod.* **2020**, *257*, 120797. [[CrossRef](#)]
39. Chen, F.; Yin, Z.; Ye, Y.; Sun, D. Taxi hailing choice behavior and economic benefit analysis of emission reduction based on multi-mode travel big data. *Transp. Policy* **2020**, *97*, 73–84. [[CrossRef](#)]
40. Zahabi SA, H.; Miranda-Moreno, L.; Patterson, Z.; Barla, P.; Harding, C. Transportation greenhouse gas emissions and its relationship with urban form, transit accessibility, and emerging green technologies: A Montreal case study. *Procedia Soc. Behav. Sci.* **2012**, *54*, 966–978. [[CrossRef](#)]
41. Cárdenas-Mamani, U.; Kahhat, R.; Vázquez-Rowe, I. District-level analysis for household-related energy consumption and greenhouse gas emissions: A case study in Lima, Peru. *Sustain. Cities Soc.* **2021**, *77*, 103572. [[CrossRef](#)]
42. Hong, J.; Goodchild, A. Land use policies and transport emissions: Modeling the impact of trip speed, vehicle characteristics, and residential location. *Transp. Res. Part D Transp. Environ.* **2014**, *26*, 47–51. [[CrossRef](#)]
43. Jun, M.J.; Choi, K.; Jeong, J.E.; Kwon, K.H.; Kim, H.J. Land use characteristics of subway catchment areas and their influence on subway ridership in Seoul. *J. Transp. Geogr.* **2015**, *48*, 30. [[CrossRef](#)]

44. Zhao, P. The impact of the built environment on individual workers' commuting behavior in Beijing. *Int. J. Sustain. Transp.* **2013**, *7*, 389–415. [[CrossRef](#)]
45. Raman, R.; Roy, U.K. Taxonomy of urban mixed land use planning. *Land Use Policy* **2019**, *88*, 104102. [[CrossRef](#)]
46. Ding, C.; Wang, Y.; Yang, J.; Liu, C.; Lin, Y. Spatial heterogeneous impact of built environment on household auto ownership levels: Evidence from analysis at traffic analysis zone scales. *Transp. Lett.* **2016**, *8*, 26–34. [[CrossRef](#)]
47. Ewing, R.; Cervero, R. Travel and the built environment: A meta-analysis. *J. Am. Plann. Assoc.* **2010**, *76*, 265–294. [[CrossRef](#)]
48. Li, Y.; Li, T.; Lu, S. Forecast of urban traffic carbon emission and analysis of influencing factors. *Energy Effic.* **2021**, *14*, 84. [[CrossRef](#)]
49. He, M.; Pu, L.; Liu, Y.; Shi, Z.; He, C.; Lei, J. Research on Nonlinear Associations and Interactions for Short-Distance Travel Mode Choice of Car Users. *J. Adv. Transp.* **2022**, *2022*, 8598320. [[CrossRef](#)]
50. Wu, J.; Jia, P.; Feng, T.; Li, H.; Kuang, H. Spatiotemporal analysis of built environment restrained traffic carbon emissions and policy implications. *Transp. Res. Part D Transp. Environ.* **2023**, *121*, 103839. [[CrossRef](#)]
51. Chengdu Municipal Bureau of Statistics, *Chengdu Statistical Yearbook on National Economic and Social Development*; China Statistics Press: Beijing, China, 2023.
52. Zhou, X.; Wang, H.; Huang, Z.; Bao, Y.; Zhou, G.; Liu, Y. Identifying spatiotemporal characteristics and driving factors for road traffic CO₂ emissions. *Sci. Total Environ.* **2022**, *834*, 155270. [[CrossRef](#)]
53. Zhao, C.; Tang, J.; Zeng, Y.; Li, Z.; Gao, F. Understanding the spatio-temporally heterogeneous effects of built environment on urban travel emissions. *J. Transp. Geogr.* **2023**, *112*, 103689. [[CrossRef](#)]
54. McDonald, B.C.; McBride, Z.C.; Martin, E.W.; Harley, R.A. High-resolution mapping of motor vehicle carbon dioxide emissions. *JGR Atmos.* **2014**, *119*, 5283–5298. [[CrossRef](#)]
55. Nyhan, M.; Sobolevsky, S.; Kang, C.; Robinson, P.; Corti, A.; Szell, M.; Ratti, C. Predicting vehicular emissions in high spatial resolution using pervasively measured transportation data and microscopic emissions model. *Atmos. Environ.* **2016**, *140*, 352–363. [[CrossRef](#)]
56. GB/T 21010-2017; Classification of Land Use Status. Standardization Administration of China: Beijing, China, 2017.
57. Wu, C.; Ye, X.; Ren, F.; Du, Q. Check-in behaviour and spatio-temporal vibrancy: An exploratory analysis in Shenzhen, China. *Cities* **2018**, *77*, 104–116. [[CrossRef](#)]
58. Taubenbock, H.; Weigand, M.; Esch, T.; Staab, J.; Wurm, M.; Mast, J.; Dech, S. A new ranking of the world's largest cities—Do administrative units obscure morphological realities. *Remote Sens. Environ.* **2019**, *232*, 111353. [[CrossRef](#)]
59. Zhao, P.; Kwan, M.-P.; Qin, K. Uncovering the spatiotemporal patterns of CO₂ emissions by taxis based on individuals' daily travel. *Transp. Geogr.* **2017**, *62*, 122–135. [[CrossRef](#)]
60. Zhang, C.; Jia, S.; Yang, R. Housing affordability and housing vacancy in China: The role of income inequality. *J. Hous. Econ.* **2016**, *33*, 4–14. [[CrossRef](#)]
61. Shen, Y.S.; Lin, Y.C.; Cui, S.; Li, Y.; Zhai, X. Crucial factors of the built environment for mitigating carbon emissions. *Sci. Total Environ.* **2022**, *806*, 150864. [[CrossRef](#)]
62. Cervero, R.; Sarmiento, O.L.; Jacoby, E.; Gomez, L.F.; Neiman, A. Influences of built environments on walking and cycling: Lessons from Bogotá. *Int. J. Sustain. Transp.* **2009**, *3*, 203–226. [[CrossRef](#)]
63. Song, S.; Diao, M.; Feng, C.C. Individual transport emissions and the built environment: A structural equation modelling approach. *Transp. Res. Part A Policy Pract.* **2016**, *92*, 206–219. [[CrossRef](#)]
64. Ding, C.; Cao, X.; Wang, Y. Synergistic effects of the built environment and commuting programs on commute mode choice. *Transp. Res. Part A Policy Pract.* **2018**, *118*, 104–118. [[CrossRef](#)]
65. GB/T 37340-2019; Calculation Method for Energy Consumption of Electric Vehicles. Standardization Administration of China: Beijing, China, 2019.
66. Greenfeld, J.S. Matching GPS observations to locations on a digital map. In Proceedings of the Transportation Research Board 81st Annual Meeting, Washington, DC, USA, 12–16 January 2020.
67. Yang, C.; Gidofalvi, G. Fast map matching, an algorithm integrating hidden Markov model with precomputation. *Int. J. Geogr. Inf. Sci.* **2018**, *32*, 547–570. [[CrossRef](#)]
68. Velaga, N.R.; Quddus, M.A.; Bristow, A.L. Developing an enhanced weight-based topological map-matching algorithm for intelligent transport systems. *Transp. Res. Part C Emerg. Technol.* **2009**, *17*, 672–683. [[CrossRef](#)]
69. He, Z.; Zheng, L.; Chen, P.; Guan, W. Mapping to cells: A simple method to extract traffic dynamics from probe vehicle data. *Comput. Aided Civ. Infrastruct. Eng.* **2017**, *32*, 252–267. [[CrossRef](#)]
70. Zhao, T. On-Road Fuel Consumption Algorithm Based on Floating Car Data for Light-Duty Vehicles. Master's Thesis, School of Traffic and Transportation, Beijing Jiaotong University, Beijing, China, 2009.
71. Zhang, J.; Zhao, Y.; Xue, W.; Li, J. Vehicle routing problem with fuel consumption and carbon emission. *Int. J. Prod. Econ.* **2015**, *170*, 234–242. [[CrossRef](#)]
72. GB 19578-2021; Fuel Consumption Limits for Passenger Cars. Standardization Administration of China: Beijing, China, 2021.

73. GB 17930-2016; Standard for Gasoline. Standardization Administration of China: Beijing, China, 2016.
74. GB/T 2589-2020; General Rules for Calculation of Comprehensive Energy Consumption. Standardization Administration of China: Beijing, China, 2020.
75. Cheng, S.; Lu, F.; Peng, P. A high-resolution emissions inventory and its spatiotemporal pattern variations for heavy-duty diesel trucks in Beijing, China. *J. Clean. Prod.* **2020**, *250*, 119445. [[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.