

Exploring Spatiotemporal Pattern and Agglomeration of Road CO₂ Emissions in Guangdong, China

Xingdong Deng ^{a,b}

Email: dengxingdong@gzpi.com.cn

Wangyang Chen ^{a,b}

Email: e0403826@u.nus.edu

Qingya Zhou ^{a,b,*}

Email: qy_zhou@yeah.net

Yuming Zheng ^{a,b}

Email: 779076427@qq.com

Hongbao Li ^{a,b}

Email: lihongbao@gzpi.com.cn

Shunyi Liao ^{a,b}

Email: liaoshunyi@gzpi.com.cn

Filip Biljecki ^{c,d}

Email: filip@nus.edu.sg

* Corresponding Author

^a Guangzhou Urban Planning and Design Survey Research Institute, Guangzhou 510060, China

^b Guangdong Enterprise Key Laboratory for Urban Sensing, Monitoring and Early Warning, Guangzhou 510030, China

^c Department of Architecture, National University of Singapore, Singapore

^d Department of Real Estate, National University of Singapore, Singapore

November 23, 2022

Exploring Spatiotemporal Pattern and Agglomeration of Road CO2 Emissions in Guangdong, China

Abstract

Road transport is a prominent source of carbon emissions. However, fine-grained regional estimations on road carbon dioxide (CO₂) emissions are still lacking. This study estimates road CO₂ emissions in Guangdong Province, China, at high spatiotemporal resolution, with a bottom-up framework leveraging massive vehicle trajectory data. We unveil the spatiotemporal pattern of regional road CO₂ emissions and highlight the contrasts among cities. The Greater Bay Area (GBA) is found to produce 76% of the total emissions, wherein Guangzhou emits the most while Shenzhen has the highest emission intensity. Emission agglomeration is still an under-explored field, which we advance in this paper. We propose Quantile-based Hierarchical DBSCAN (QH-DBSCAN) to explore road CO₂ emission agglomeration in GBA. Our method is the first one to identify the specific location and scope of emission hotspots. Emission hotspots exhibit significant concentration on major urban centers. Considering emission characteristics from multiple perspectives, we derive six emission categories, including four emission zones and two emission connectors. The density-based property of our method results in spatially contiguous regions with similar emission patterns. Accordingly, we divide policy zones and propose targeted strategies for road carbon reduction. The study provides new technologies and insights to achieve regional sustainable development.

Keywords: Road transport, Carbon emission, Spatiotemporal distribution, Vehicle trajectory, GBA, Emission agglomeration

1. Introduction

Carbon dioxide (CO₂) emissions pose a great threat to the global environment as a culprit of global warming. The pattern of carbon emission and its relationship with different aspects of human development is an enduring topic to discuss (Shindell et al., 2008; Zhang et al., 2014; Zhang and Da, 2015; Huang et al., 2018; Shan et al., 2021). Among the multitudinous sectors of human activities, transport sector accounts for more than 20% of the global carbon emissions. (Yan et al., 2017; Van Fan et al., 2018). The share continues to climb up (IEA, 2019; Mohsin et al., 2019), rendering transport a challenging sector in carbon emission mitigation (Yang et al., 2015; Batur et al., 2019). Among all the transport modes, road transport is the closest to our

Abbreviations: GBA, Guangdong-Hong Kong-Macao Greater Bay Area ; HDFV, Heavy-Duty Freight Vehicle; QH-DBSCAN, Quantile-Based Hierarchical Density-Based Spatial Clustering Of Applications With Noise; O-D, Origin-Destination.

10 daily lives and contributes to 82% of the total carbon emission in the sector (IEA, 2020). In this
11 regard, estimating road carbon emissions and exploring the patterns are essential for relevant
12 carbon reduction policy making.

13 To estimate road CO₂ emissions, various kinds of data sources have been leveraged, but
14 each faces different deficiencies. One group of studies use collective data, such as statistical
15 yearbooks (Lin and Li, 2020; Zhou et al., 2018; Cai et al., 2018). Such method usually covers
16 a broad geographic scale, but suffers from coarse resolution in both space and time. Some
17 scholars seek to ameliorate the estimation granularity by using survey (Pérez-Martínez et al.,
18 2020; McQueen et al., 2020; Sobrino and Arce, 2021; Patiño-Aroca et al., 2022), sensors (Liu
19 and Zimmerman, 2021), smartphones (Manzoni et al., 2010) or video surveillance data (Li et al.,
20 2019b). Although these methods can obtain fine-grained results, they are only applicable at a
21 limited geographic scale. Balancing the trade-off between both data fineness and wide spatial
22 coverage, vehicle trajectory is considered to be a suitable instrument to estimate road emissions
23 (Kan et al., 2018; Xia et al., 2020; Sun et al., 2015; Zhao et al., 2017). However, existing vehicle
24 trajectory datasets often only contain a subset of vehicle types, e.g. taxis (Zhao et al., 2017),
25 ride-hailing (Sui et al., 2019), and light-duty vehicles (Li et al., 2019a). Emission patterns
26 obtained with these datasets are characterized by low sampling rate and thus may considerably
27 deviate from the actual pattern of the entire fleet, which is more valuable for policy-makers to
28 understand the status quo and design emission mitigation strategies. Solving the problem calls
29 for a trajectory dataset with higher sampling rate and more vehicle types. Correspondingly, a
30 framework for estimating road CO₂ emissions with massive vehicle trajectory dataset is also
31 required.

32 Besides deficiencies in emission estimation, limited attention is paid to emission agglom-
33 eration. As an effective strategy to catalyze the development of economy, urbanization and
34 industry (Fujita and Thisse, 1996; Malmberg and Maskell, 1997; Fang and Yu, 2017), agglom-
35 eration also brings some negative externalities, of which an important aspect is the convergence
36 of carbon emissions (Yu et al., 2020; Wang et al., 2019). In China, 1% of the land contributes
37 to 70% national CO₂ emissions, and the regions with high level of urban agglomerations are
38 also those emitting the most (Wang et al., 2014). In our context, we define the concept of
39 emission agglomeration as the phenomenon that massive emissions are concentrated in certain

40 contiguous territory. To better spatialize and visualize emission agglomeration, we introduce
41 the “emission hotspot”, which indicates the spatially continuous region with high emissions per
42 unit area. Although regional emission agglomeration has been verified statistically with city-
43 level or county-level emissions, studies on identifying emission hotspots is extremely lacking.
44 The hotspots of CO₂ emissions may be prone to worse environmental problems and could also
45 be the crux for overall emission reduction. Thus, how to determine the location and scope of
46 emission hotspots becomes crucial. However, the literature falls short on relevant methods,
47 which is another gap that this study intends to fill.

48 In this paper, we conduct a regional study in Guangdong Province, which has the largest
49 population and economy in China, and accordingly produces intensive emissions. Guangdong-
50 Hong Kong-Macao Greater Bay Area (GBA) denotes the most developed region in Guangdong
51 and is primary to the overall development of China (Hui et al., 2020). According to the *Out-*
52 *line Development Plan for the Guangdong-Hong Kong-Macao Greater Bay Area* published by
53 Chinese State Council, low-carbon development is among the major objectives of GBA. GBA
54 maintains great economic, population and urban agglomeration (Chen et al., 2017, 2020; Yu,
55 2019), making it a favorable region to study mesoscale emission agglomeration (Zhou et al.,
56 2022a). Although some insights have been given in spatial (Lin and Li, 2020; Chen et al., 2017)
57 and temporal (Zhou et al., 2018; Chen et al., 2017) patterns of carbon emissions in Guangdong
58 or GBA, the results are either coarse in granularity or focus on total emissions instead of road
59 emissions. Fine-grained regional patterns of road transportation CO₂ emission and agglomera-
60 tions in Guangdong and GBA is still under-explored. Exploring regional patterns of road CO₂
61 emission is the prerequisite GBA to formulate strategies for sustainable development.

62 Taking into account the research opportunities and deficiencies described so far, we aim to
63 bridge the following gaps and present the following contributions to the field: (1) We implement
64 a bottom-up framework for estimating road CO₂ emissions based on a vehicle trajectory dataset
65 with sampling rates surpassing the counterparts in the literature, and develop approaches to han-
66 dle inherent data deficiencies. (2) We present fine-grained spatiotemporal patterns of regional
67 road CO₂ emissions in Guangdong Province where similar analysis is lacking. (3) We propose
68 an approach to identify hierarchical emission hotspots, which uncover the spatiotemporal emis-
69 sion agglomeration patterns in GBA. (4) We categorize our study area into emission zones with

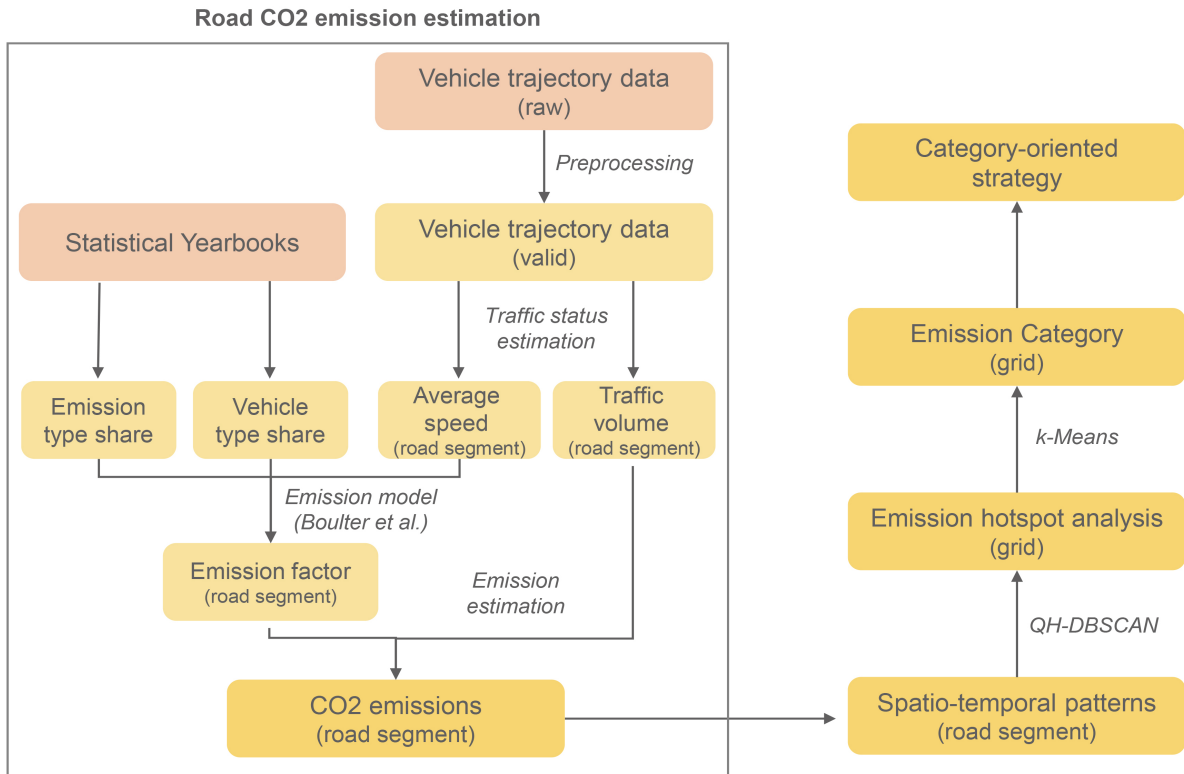


Figure 1: Analytical framework.

70 distinct emission patterns and propose targeted emission reduction strategies.

71 The analytical framework of this study is designed as illustrated in Figure 1. Our study
 72 estimates CO₂ emissions at road segment level at the first stage. To achieve such an advance-
 73 ment, we harness a massive vehicle trajectory dataset that contains 12.9 million records in a
 74 day, encompassing all vehicle types. Its sampling rate exceeds most counterparts used hith-
 75 erto. We firstly introduce the dataset and elucidate the estimation process. Then we analyze the
 76 spatial distribution and the hourly variation of road CO₂ emissions in Guangdong, highlighting
 77 the disparities among cities. Next, a density-based clustering algorithm is proposed to detect
 78 hierarchical emission hotspots. Their temporal variations are also explored. Finally, we inte-
 79 grate all facets of emission features and divide our study area into different emission categories.
 80 Based on the emission characteristics of each category, targeted carbon reduction strategies are
 81 proposed for the sustainable development in GBA and Guangdong.

82 2. Methodology

83 2.1. Study area

84 Our study area (Figure 2) includes all the twenty-one prefectural-Level cities in Guangzhou
 85 Province. We highlight the nine GBA cities (Guangzhou, Shenzhen, Foshan, Dongguan, Zhuhai,

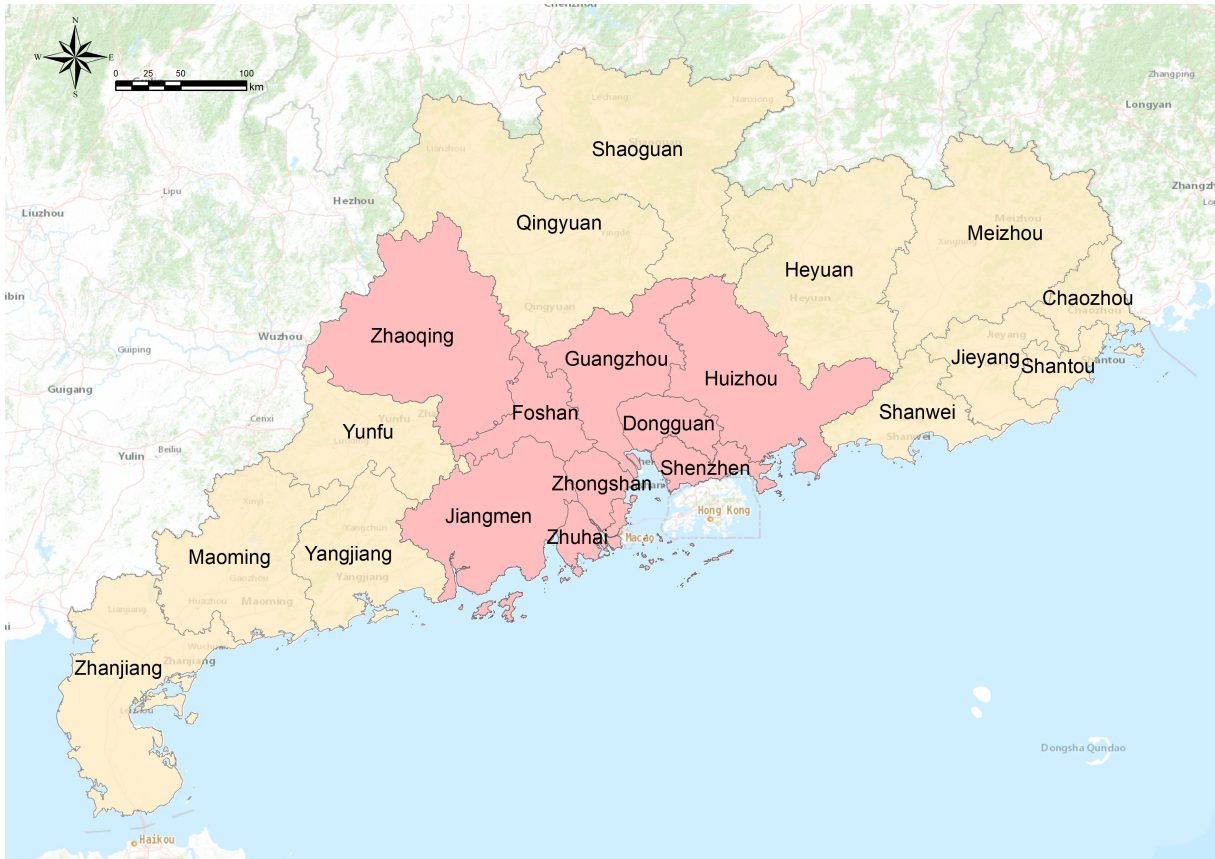


Figure 2: Study area: Guangdong Province (colored area) and GBA (pink area). Source of the basemap: (c) Esri.

86 Huizhou, Zhongshan, Jiangmen and Zhaoqing), considering their paramount role in the region.
 87 Hong Kong and Macao, which also belong to GBA but are not part of Guangdong Province, are
 88 not included in this study due to their absence in the trajectory dataset. Thus, we refer to the
 89 nine aforementioned and contiguous cities when we refer to GBA in this paper. The socioeco-
 90 nomic indicators of the corresponding cities are displayed in Supplementary Table S1 according
 91 to the Statistical Yearbook of Guangdong Province (2021) ([Guangdong Province Statistical Bu-](#)
 92 [reau, 2021](#)). GBA accommodates 62% of the population and contributes to 81% of the GDP
 93 with only 30% of land in Guangdong. It is four times as densely populated and three times as
 94 economically developed as the rest areas in Guangdong, showing significant agglomerations.
 95 To put that in context for international readers, these values are comparable with the population
 96 of Germany, economy of Brazil, and area of Croatia. Considering that GBA also has denser
 97 road networks and heavier traffic ([Hui et al., 2020](#)), in this study, we estimate the road CO2
 98 emissions and delineate the patterns in the entire province, but focus on GBA when exploring
 99 emission hotspot patterns.

Table 1: Summary statistics of the trajectory data.

Vehicle type	Raw trajectories	Valid trajectories	Samples	Total travel distance (10 ⁴ km)	Sampling rate (%)
HDFV	527,396	520,924	137,963,458	505.1	100
Non-HDFV	12,371,644	10,623,090	385,648,577	1022.4	40
Total	12,899,040	11,144,014	523,612,035	1527.5	–

100 2.2. Data description and preprocessing

101 In this study, we introduce a new vehicle trajectory dataset to estimate road CO₂ emissions.
102 The dataset is provided by PalmGo, a company offering driving navigation services in China.
103 The dataset contains 12.9 million navigation trajectories of individual vehicles in a day (Monday
104 19 October 2020) in Guangdong (Table 1). The dataset covers our study area well and is verified
105 representative to the general traffic condition (Supplementary Note S1). A trajectory sample is
106 shown in Supplementary Table S3. Each trajectory records the location and time of departure
107 and arrival. The route is represented by a series of passed road segments (*SegmentIDs*). Every
108 time the vehicle drives onto a new road segment, the road segment ID, direction, and time
109 of entry would be appended to the record. By matching trajectory records with road network
110 geodata provided by PalmGo, we can reconstruct the trajectories. Vehicles are divided by two
111 types: heavy-duty freight vehicle (HDFV) and the others (non-HDFV). Since all HDFVs are
112 obliged to install Global Positioning System (GPS) devices by law in China, the sampling rate
113 of HDFV fleets is 100%. For non-HDFVs, the overall sampling rate is 40% according to the
114 company. Despite covering the entire fleet with favorable sampling rate, the dataset is subject
115 to one major issue. All the time-related fields of trajectory records are only accurate to minute,
116 which causes difficulty in computing vehicular instantaneous status (speed and acceleration).
117 To exclude abnormal records that may root from device errors, trajectories less than 500 m or
118 with single sample are filtered out. As a result, we obtain 11.1 million valid trajectories.

119 Meshing is a common approach to regularize and integrate anisotropic spatial attributes
120 (road networks) to isotropic ones (grids). In our case, CO₂ emissions would be measured at
121 road segment level as the foundation. To facilitate emission hotspot analysis, the emissions
122 would be further allocated to regular grids. Instead of defining our own grids, we use the 1
123 km-resolution grids of WorldPop (WorldPop, 2018; Tatem, 2017) to integrate the population
124 information and make our results attachable to other relevant initiatives using the same grid

125 system (Chen et al., 2021).

126 2.3. Traffic status estimation

127 To estimate road CO2 emissions with trajectories, the first step is to calculate the traffic
128 status on road. Some studies are able to compute the instantaneous speeds and accelerations
129 of individual vehicles at different positions based on second-level trajectory data (Böhm et al.,
130 2022; Deng et al., 2020). Since our dataset is only accurate to minute, we propose the following
131 strategies to estimate the traffic status with minimal deviation. Instead of distinguishing indi-
132 vidual vehicles, we emphasize the average traffic status per road segment and update the status
133 every hour. Besides, we only calculate speeds and avoid computing accelerations which require
134 better fineness in time. Traffic volume per road segment per hour is gleaned by counting the
135 number of passed trajectories. The volume of HDFVs and non-HDFVs is counted separately,
136 but they are summed up to estimate the speed. The average speed on a road segment at an hour
137 is calculated by dividing the aggregated travel distance by the total time cost (Equation 1).

$$s_{i,t} = \frac{|U(i,t)| \cdot l_i}{\sum_{u \in U(i,t)} T_{u,i}} \quad (1)$$

138 where $s_{i,t}$ denotes the average speed on road segment i at hour t , $U(i,t)$ is the set of vehicle
139 trajectories that pass road segment i at hour t , and $|\cdot|$ represents the size of the set. In our case,
140 $|U(i,t)|$ or $q_{i,t}$ denotes the traffic volume on road segment i at hour t . l_i is the length of road
141 segment i , and $T_{u,i}$ is the passing time of vehicle u through road segment i .

142 Another challenge is how to estimate actual passing times on road segments ($T_{u,i}$) with
143 minute-level trajectories. Considering the way the trajectory is recorded (Supplementary Table
144 S3), simply regarding time differences between consecutive road segments as passing times
145 becomes infeasible because of two issues. First, if a vehicle passes through more than one road
146 segment at the same minute, the time difference would be zero, resulting in the inability to
147 calculate speed. Second, even with a non-zero time difference, the actual passing time would
148 still be out of calibration. For example, a time difference of 1 minute could correspond to
149 any actual passing time between 1 and 120 seconds. Out of our control, since 93.2% of the
150 road segments are shorter than 500 m, 95.4% of the time differences are no larger than 1 min,
151 rendering it a pervasive problem. To alleviate the problem, we amend the time differences based

152 on following strategies. We introduce the minimal time cost (T_{min}^i) for each road segment, which
 153 is calculated with the maximum speed limit. The maximum speed is determined by the road
 154 class referring to the speed limit regulation (100 km/h for highways, 80 km/h for provincial and
 155 urban expressways, 60 km/h for arterial roads, 40 km/h for ramps and secondary roads, and 20
 156 km/h for branch roads). $T_{u,i}$ smaller than T_{min}^i is replaced by T_{min}^i . Furthermore, we mitigate
 157 the uncertainty by accumulating the length and time difference of continuous segments until
 158 the added-up time difference exceeds a minimum threshold (τ). Then we calculate the average
 159 speed based on the cumulative length and time, and reallocate the cumulative passing time to
 160 each road segment proportional to the segment length. After grid searching with τ from 2 min
 161 to 7 min, τ is finally set to 5 min as the trade-off between estimation accuracy and variance.
 162 Mathematically, the speed error caused by time uncertainty is constrained within a range of
 163 $\pm 20\%$ (-1/5 to 1/5).

164 2.4. Road CO2 emission estimation

Table 2: Weighted average parameters to calculate emission factors for HDFV and non-HDFV. (Parameter $a - k$ are the coefficients of the polynomial fitting function between speeds and emission factors provided by (Boulter et al., 2009).)

Vehicle type	a	b	c	d	e	f	g	k
HDFV	5237.8	1003.4	-20.164	0.1603	1.28e-3	-1.67e-5	3.28e-8	1
Non-HDFV	2735.7	104.2	-0.468	0.0098	-3.43e-5	2.39e-7	-4.09e-10	1

165 With the segment-level traffic status per hour, we estimate the road CO2 emissions with
 166 a speed-based microscopic emission model (Boulter et al., 2009), which fits our data charac-
 167 teristics and emission standard system well, and has been widely applied in various scenarios
 168 (Carslaw et al., 2010; Lomas et al., 2010; Hicks et al., 2021). The model determines the emis-
 169 sion factor f by speed and parameters varying with vehicle attributes including the vehicle type,
 170 fuel type and emission standard (Equation 2). Due to lack of these vehicle attributes, we follow
 171 the approach commonly used in the literature (Pla et al., 2021; Zhou et al., 2022b), by assum-
 172 ing that the on-road shares of vehicle types, fuel types, and emission standards are consistent
 173 across all road segments, and estimating CO2 emissions with weighted average emission fac-
 174 tors. Non-HDFV consists of all the vehicle types except for HDFV, including large passenger
 175 vehicle, medium passenger vehicle, small passenger vehicle, mini passenger vehicle, medium

176 freight vehicle, light freight vehicle, and mini freight vehicle. The shares of vehicle types in
177 non-HDFVs are obtained from the vehicle possession structure in 2020 ([Guangdong Province](#)
178 [Statistical Bureau, 2021](#)). The shares of fuel types and emission standards refer to the national
179 level in 2018 ([MEE, 2019](#)). Despite some time inconsistencies, it is the official data closest
180 in time to our trajectories. CO2 emissions from new energy vehicles are ignored in this study,
181 given that new energy vehicles only account for 1.75% of the total vehicles in China in 2020
182 ([MEE, 2021](#)). Details in shares of vehicle types, fuel types, and emission standards could be
183 found in Supplementary Table S4-S6, while the parameters to calculate CO2 emission fac-
184 tors for each combination are available Supplementary Table S7. With all of this information,
185 we obtain the weighted average emission factors for HDFVs and non-HDFVs (Table 2), re-
186 spectively. Accordingly, road segment-level CO2 emissions are estimated with the following
187 formulas (Equation 3, 4, and 5):

$$f_{i,t} = k \cdot (a + bs_{i,t} + cs_{i,t}^2 + ds_{i,t}^3 + es_{i,t}^4 + fs_{i,t}^5 + gs_{i,t}^6) / s_{i,t} \quad (2)$$

188

$$E_{i,t}^{non} = \frac{1}{\beta_{non}} \times (f_{i,t}^{non} \times l_i \times q_{i,t}^{non}) \quad (3)$$

189

$$E_{i,t}^{hdfv} = \frac{1}{\beta_{hdfv}} \times (f_{i,t}^{hdfv} \times l_i \times q_{i,t}^{hdfv}) \quad (4)$$

190

$$E_{i,t} = E_{i,t}^{non} + E_{i,t}^{hdfv} \quad (5)$$

191 where $a-f$ are the weighted average parameters depending on vehicle type (Table 2), $E_{i,t}^{hdfv}$ and
192 $E_{i,t}^{non}$ denote CO2 emissions at segment i and hour t from HDFVs and non-HDFVs respectively,
193 $E_{i,t}$ is the total CO2 emissions, $f_{i,t}^{hdfv}$ and $f_{i,t}^{non}$ are the average CO2 emission factors (g/km), and
194 $q_{i,t}^{hdfv}$ and $q_{i,t}^{non}$ indicate traffic volumes. The sampling rate of HDFVs β_{hdfv} and non-HDFVs is
195 1.0 and 0.4 respectively.

196 2.5. Emission hotspot detection (QH-DBSCAN)

197 Considering that emissions are continuously distributed in space, we deem density-based
198 clustering algorithms an appropriate option to capture the hotspots. Density-based spatial clus-

Table 3: Description of emission hotspots at different levels.

Level	Color	Quantile	Description
High	Red	90%	Clustered grids with emissions higher than the 90% quantile
Middle	Orange	75%	Clustered grids with emissions higher than the 75% quantile
Low	Green	50%	Clustered grids with emissions higher than the 50% quantile

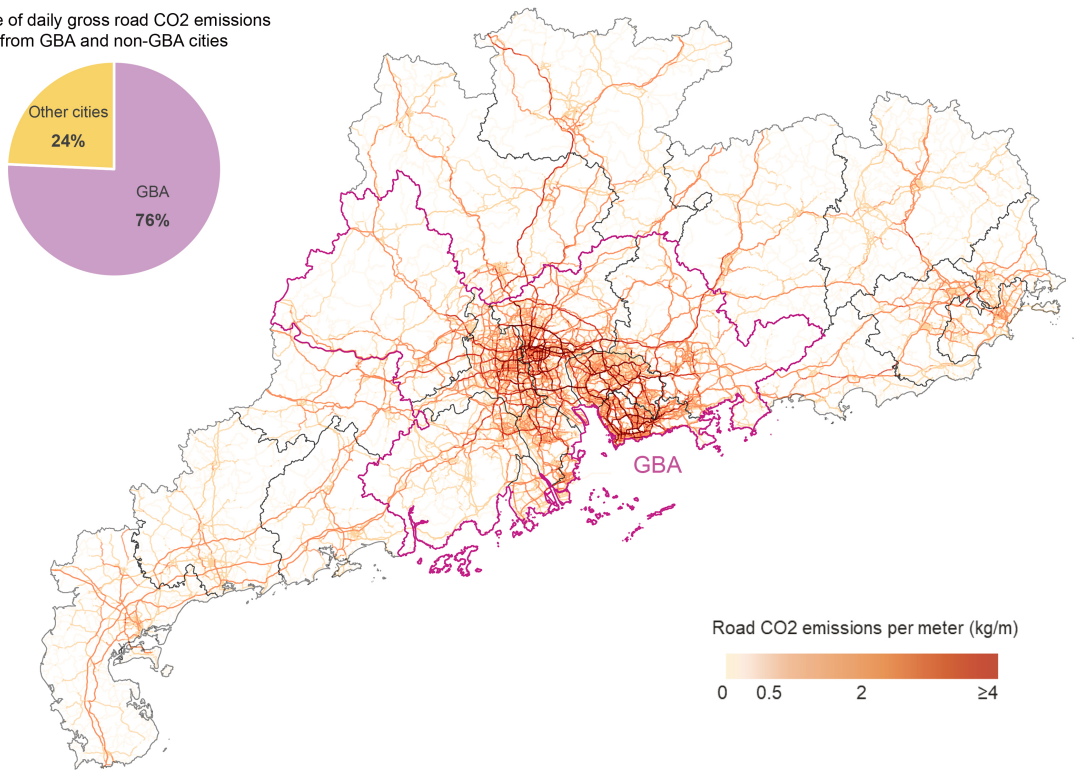
199 tering of applications with noise (DBSCAN) (Ester et al., 1996) is widely used to detect clusters
 200 based on the proximity in the feature space. To decipher the hierarchy of emission hotspots, on
 201 the ground of DBSCAN, we propose a method called Quantile-based Hierarchical DBSCAN
 202 (QH-DBSCAN) to recognize hierarchical emission hotspots, which refer to spatially clustered
 203 high-emission areas. First, road-level CO₂ emissions are allocated onto WorldPop grids. We
 204 use quantiles to identify high-emission grids. To explore a hierarchy of emission hotspots, we
 205 prepare a list of quantiles and repeat emission hotspot detection with each. For each quantile,
 206 the corresponding emission hotspots are identified by feeding the grids with emissions over the
 207 quantile into a DBSCAN model. We fully appraise the uncertainty hidden behind the method.
 208 The list of quantiles is determined through exploratory experiment on the relationship between
 209 the quantiles and the covered percentages of total emissions, while the hyper-parameters of DB-
 210 SCAN are determined by sensitivity analysis. Both experiments would be elaborated in Sec-
 211 tion 3.3. The selected quantiles with the associated levels of emission hotspots are displayed in
 212 Table 3. To clarify, the 90%-quantile grids are equivalent to the top 10% emitting grids, and so
 213 on. QH-DBSCAN has two major merits. First, it could detect hierarchical emission hotspots
 214 and present the results with intuitive maps. Second, as a density-based clustering method, it
 215 leads to results with spatial contiguity, which is more useful for potential policy making.

216 3. Results

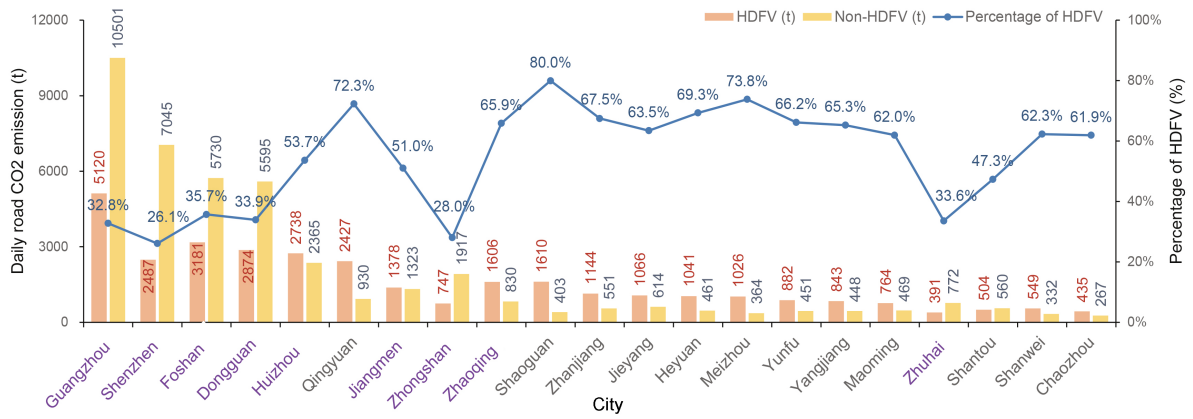
217 3.1. Spatial pattern of road CO₂ emissions in Guangdong.

218 Based on the vehicle trajectory dataset and the proposed emission estimation method, segment-
 219 level CO₂ emissions per hour are obtained. The spatial distribution of daily gross road CO₂
 220 emission intensity is illustrated in Figure 3(a). Emission intensity denotes CO₂ emissions per
 221 meter for each road segment. Road CO₂ emissions in Guangdong Province show significant
 222 spatial agglomeration. GBA accounts for 76% of the total emissions, while non-GBA cities

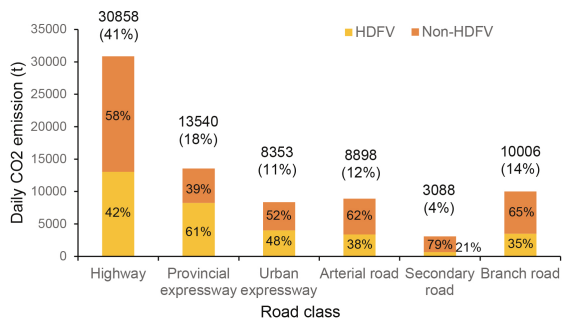
Share of daily gross road CO2 emissions from GBA and non-GBA cities



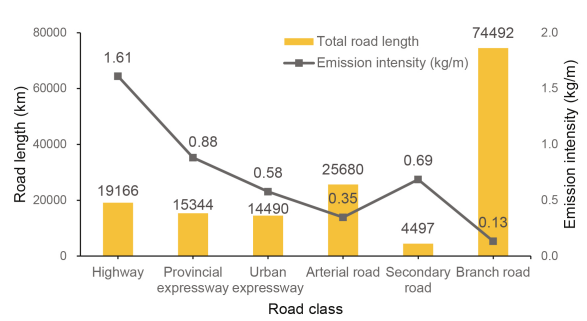
(a)



(b)



(c)



(d)

Figure 3: Daily gross road CO2 emission pattern in Guangdong Province. (a) Daily gross CO2 emission intensity per road segment in Guangdong. (b) Daily gross road CO2 emissions per city and vehicle type (GBA cities in purple). (c) Daily gross road CO2 emissions per road class and vehicle type. (d) Emission intensity and total road length per road class.

223 generate considerably less emissions on a broader land. Comparing the daily gross emissions
224 among the cities (Fig 3(b)), the capital city, Guangzhou, appears to be the topmost emitter,
225 producing 15621t CO₂ and contributing to 21% of the total. It is followed by the other ma-
226 jor city, Shenzhen, with 9532t of road CO₂ emissions. Although the total emission is lower
227 than Guangzhou, Shenzhen is the most densely emitting city, with an average emission inten-
228 sity (1.01 kg/m) twice as much as the provincial average (0.49 kg/m). Foshan and Dongguan
229 are the other two emission-intensive cities with a total emission close to the one of Shenzhen.
230 Their emission intensities are between Shenzhen and Guangzhou. These four cities are the ma-
231 jor industrial cities in Guangdong. For the rest cities, both the total emissions and emission
232 intensity drop dramatically. Zhuhai is the only GBA city at the bottom of the emission ranking.
233 Qingyuan produces the most road CO₂ among the non-GBA cities.

234 GBA cities appear to have higher percentages of road CO₂ emissions from non-HDFVs.
235 Comparing the shares of emissions from different vehicle types across the cities, we find that
236 GBA cities tend to have less proportions of HDFV emissions (generally around 30%), while
237 those of most non-GBA cities are over 60%. GBA cities possess 71.7% of the non-HDFVs
238 in Guangdong ([Guangdong Province Statistical Bureau, 2021](#)), but they generate 86.0% of the
239 provincial non-HDFV emissions by our estimation. To discover whether non-HDFVs in GBA
240 cities tend to emit more in average, we calculate and display the road CO₂ emissions per non-
241 HDFV possession across the cities in Supplementary Table S8. The results are affirmative.
242 The average emissions per non-HDFV possession for GBA cities are 2.24 kg, while those for
243 non-GBA cities are only 0.92 kg. GBA cities occupy the top 8 places in this indicator. In
244 comparison, regarding emissions per HDFV possession, the mean of non-GBA cities is 16.02
245 kg, which is a bit larger than that of GBA cities (11.14 kg). The difference is not as significant as
246 that of non-HDFV. Our findings suggest that non-HDFVs are used in a more emission-intensive
247 way in GBA than outside. It may be due to longer travel distances, more frequent vehicle travel,
248 and severer congestion in GBA. However, we also cannot rule out the possibility that drivers in
249 non-GBA cities use navigation less due to a less complicated traffic system and smaller active
250 area, which leads to undersampling of trajectories in these regions.

251 The distribution of CO₂ emissions by road class is also heterogeneous (Figure 3(c)). High-
252 way generates the most emissions (41% of the total), with 42% from HDFVs. The emission

253 intensity on highways is also the highest with 1.61 kg/m in average (Figure 3(d)). Provincial
254 expressways as the other major regional connector, emit the second most CO₂ (18% of the
255 total). It has the highest share of HDFV emissions (61%), indicating its vital role in regional
256 freight transportation. The rest road classes mainly serves local transportation. Although their
257 emission intensities are generally more moderate, they cover 77% of the total road length and
258 accordingly contribute to 50% of the total emissions. In particular, the branch roads exhibit
259 substantially minor emission intensity (0.13 kg/m), but the massive length makes them generate
260 14% of the total emissions.

261 3.2. Temporal pattern of road CO₂ emissions

262 The hourly variations of road CO₂ emissions during the day are revealed in Figure 4. Given
263 that the total emissions of GBA and non-GBA cities differ considerably, we further compare
264 their temporal fluctuation on emissions (Figure 4(a), 4(b)). In general, the city-level road CO₂
265 emissions share a “three-stage” pattern, including an ascending stage (4:00-10:00), a plateau
266 stage (10:00-19:00) and a descending stage (19:00-4:00 the next day). During the plateau stage,
267 the emissions experience a slight decrease at noon (11:00-13:00). An obvious trough is ob-
268 served at dawn (around 4:00). To quantify the emission variance of a city in a day, we define
269 the variation rate as the maximum hourly emissions divided by the minimum. Variation rates
270 of GBA cities (ranging from 2.7 for Zhaoqing to 9.6 for Zhongshan) surpass non-GBA cities
271 (ranging from 1.8 for Shaoguan to 6.8 for Shantou). Guangzhou is the top emitter at all hours,
272 with a variation rate of 4.7. The other three emission-intensive cities, Shenzhen, Foshan and
273 Dongguan, possess larger variation rates, which are 6.7, 6.8, and 5.3 respectively.

274 To provide more insights in GBA, we compare the temporal patterns by vehicle types. The
275 hourly variations of road CO₂ emissions from HDFVs and non-HDFVs for GBA cities are
276 demonstrated in Figure 4(c) and 4(d), respectively. Non-HDFV emissions follow a very similar
277 “three-stage” pattern with the total emissions. In comparison, HDFV emissions fluctuate less
278 during the day. Guangzhou is the city with the most stable and intensive HDFV emissions dur-
279 ing the day, manifesting its dominant position in freight transportation in GBA. There are two
280 obvious valleys of HDFV emissions at 6:00-9:00 and 17:00-20:00 in Guangdong, Shenzhen,
281 Dongguan and Foshan. The declines may be attributed to HDFV traffic control during the peak
282 time. In these four major cities, HDFVs show some degree of off-peak travel, and the pattern

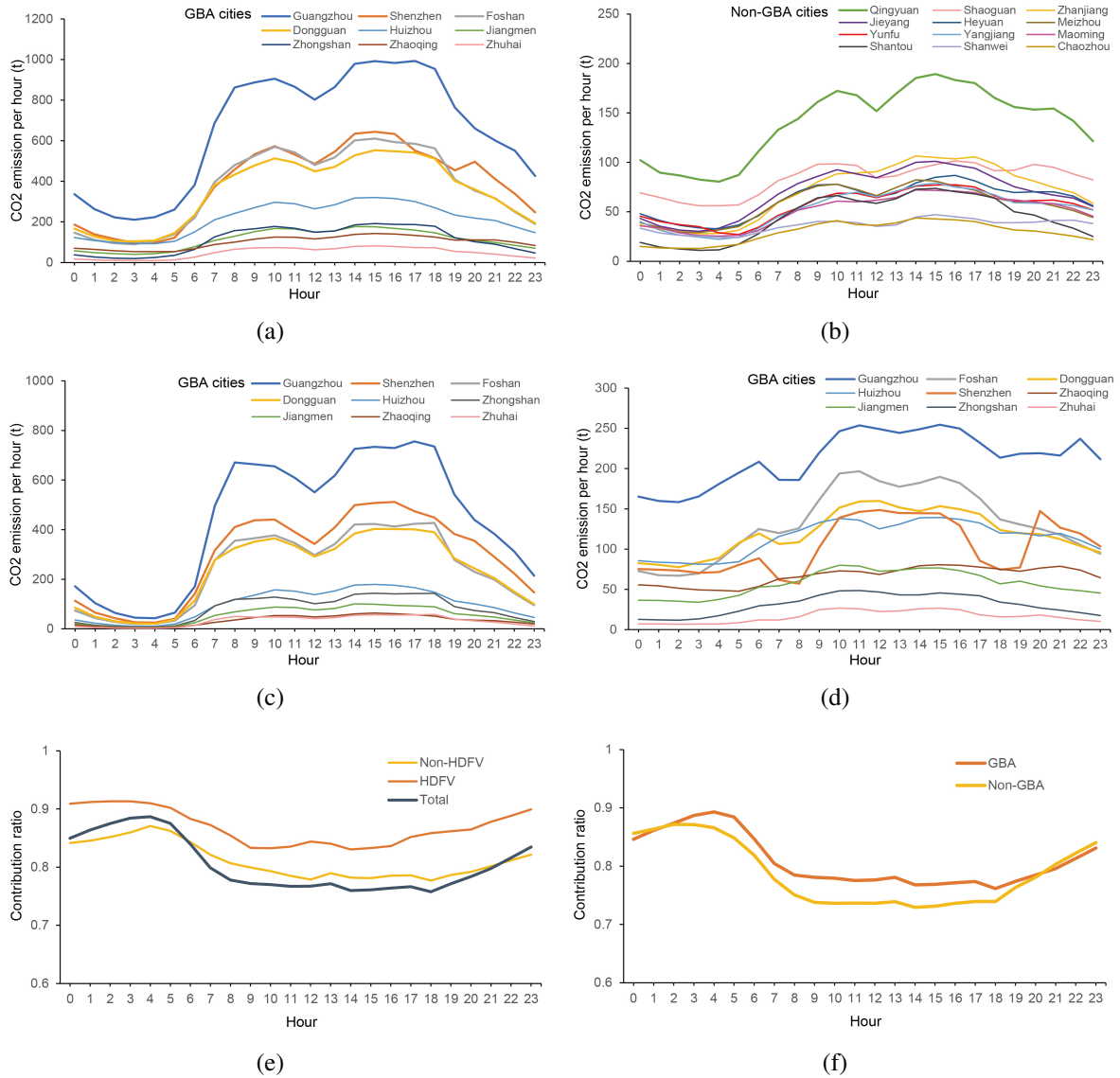


Figure 4: Temporal patterns of road CO₂ emissions in Guangdong. **(a)** Hourly variations of road CO₂ emissions for GBA cities. **(b)** Hourly variations of road CO₂ emissions for non-GBA cities. **(c)** Hourly variations of non-HDFV CO₂ emissions for GBA cities. **(d)** Hourly variation of HDFV CO₂ emissions for GBA cities. **(e)** Hourly variations of contribution ratio for GBA and non-GBA cities. **(f)** Hourly variations of contribution ratio per vehicle type in Guangdong.

283 is the most significant in Shenzhen. However, there are still apparent overlaps of high emis-
 284 sions from HDFVs and non-HDFVs during the period between the morning and evening peaks,
 285 rendering a space for relevant policy making.

286 The study also outlines the temporal pattern of road CO₂ emission agglomeration. We use
 287 the contribution ratio of the top 10% emitting road segments to the total emissions to quantify
 288 the level of emission hotspots in a region. In Guangdong, the top 10% emitting roads produce
 289 77.6% of the total emissions, presenting significant emission agglomeration. The phenomenon
 290 is also reported in other city-level road emission studies (Böhm et al., 2022; Chen et al., 2022).
 291 The contribution ratio of GBA cities (78.6%) exceeds the provincial average, while non-GBA

292 cities (75.9%) is below. The hourly variation of the contribution ratio of GBA and non-GBA
 293 cities is plotted in Figure 4(e). For both groups of cities, the contribution ratio culminates near
 294 90% around the valley hour of emission (4:00-5:00). It indicates that, although the total emis-
 295 sion shrinks at dawn, emissions are actually more concentrated on few roads. In terms of vehicle
 296 types (Figure 4(f)), HDFV emissions are more concentrated than non-HDFV emissions at all
 297 time. Nevertheless, the emission agglomeration of non-HDFV emissions is also nonnegligible.

298 3.3. Emission hotspot analysis

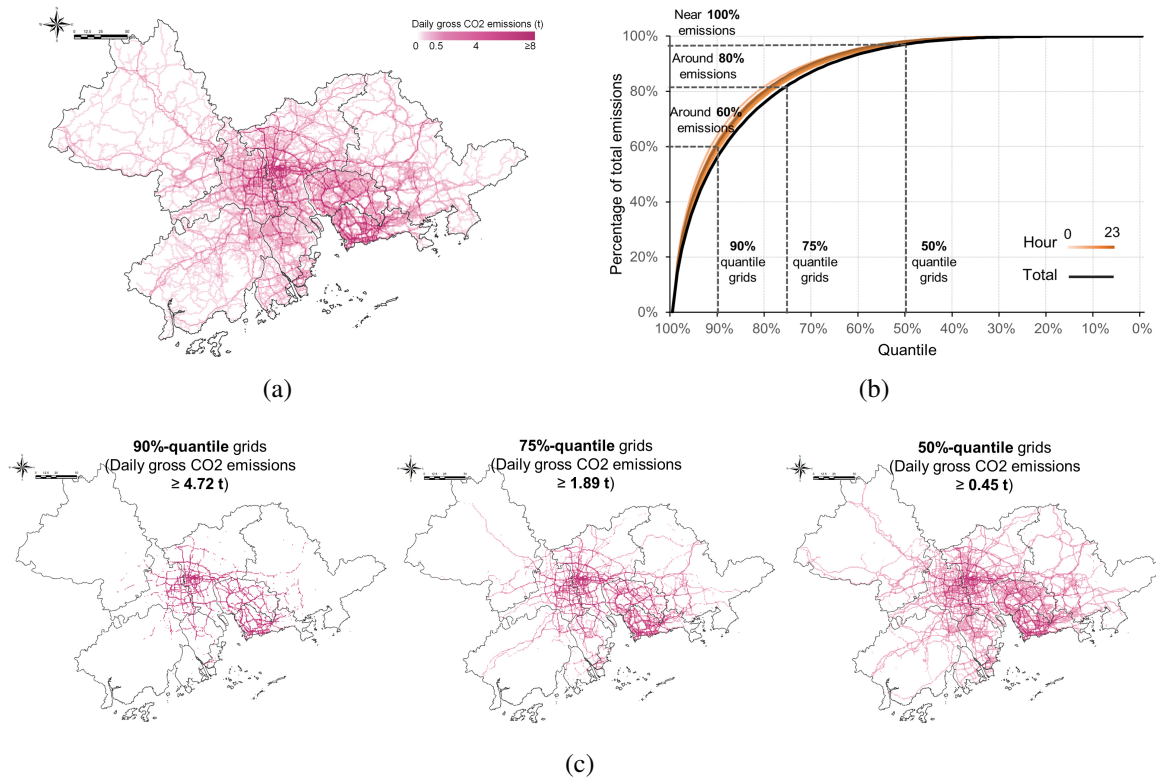


Figure 5: Process of quantile selection for emission hotspot and analysis. (a) Spatial distribution of grid-level daily gross road CO2 emissions in GBA. (b) Curves on relationships between quantiles and the covered percentages of total emissions. (c) Spatial distributions of grids associated with the three selected quantiles.

299 In the preceding sections, we have presented the spatiotemporal pattern of road CO2 emis-
 300 sions in Guangdong. The findings suggest that GBA contributes to 76% of the total emissions
 301 with only 30% of the land. GBA is also found to exhibit more significant emission agglom-
 302 eration gauged by contribution ratios. Thus, in the following analysis, we focus on GBA to
 303 conduct emission hotspot analysis. To regularize the road-level emissions spatially, we lattice
 304 the GBA area into 70,750 1km-resolution (0.01°) grids and aggregate the emissions within each
 305 grid, resulting in the grid-level map of daily gross emissions in GBA (Figure 5(a)).

306 Before implementing QH-DBSCAN method to identify emission hotspots, we need to de-
 307 termine the list of quantiles. The relationship between quantiles and the percentages of the cov-
 308 ered daily gross CO₂ emissions is illustrated in Figure 5(b) (the black curve). As the quantile
 309 decreases, the percentage increases dramatically at first. The 90%-quantile grids have covered
 310 around 60% of the total emissions. After that, the slope becomes flatter gradually and near all
 311 the emissions are covered with the minimum amount of grids at the 50% quantile. Emissions at
 312 different hours also follow a similar trend. Given the observation, we select 90% and 50% quan-
 313 tiles as the upper and lower limits respectively. To introduce some continuity and hierarchy, the
 314 75% quantile is added, which is associated with a percentage in the middle (80%). As a result,
 315 we confirm three quantiles, namely 90%, 75% and 50%. The spatial distribution of the grids
 316 fulfilling each quantile is shown in Figure 5(c). The 90%-quantile grids primarily distribute at
 317 the downtown of Guangzhou and Shenzhen and along the regional connectors among the four
 318 major cities. More areas with backbone roads are included for the 75%-quantile, and most built
 319 areas in GBA are involved for the 50%-quantile.

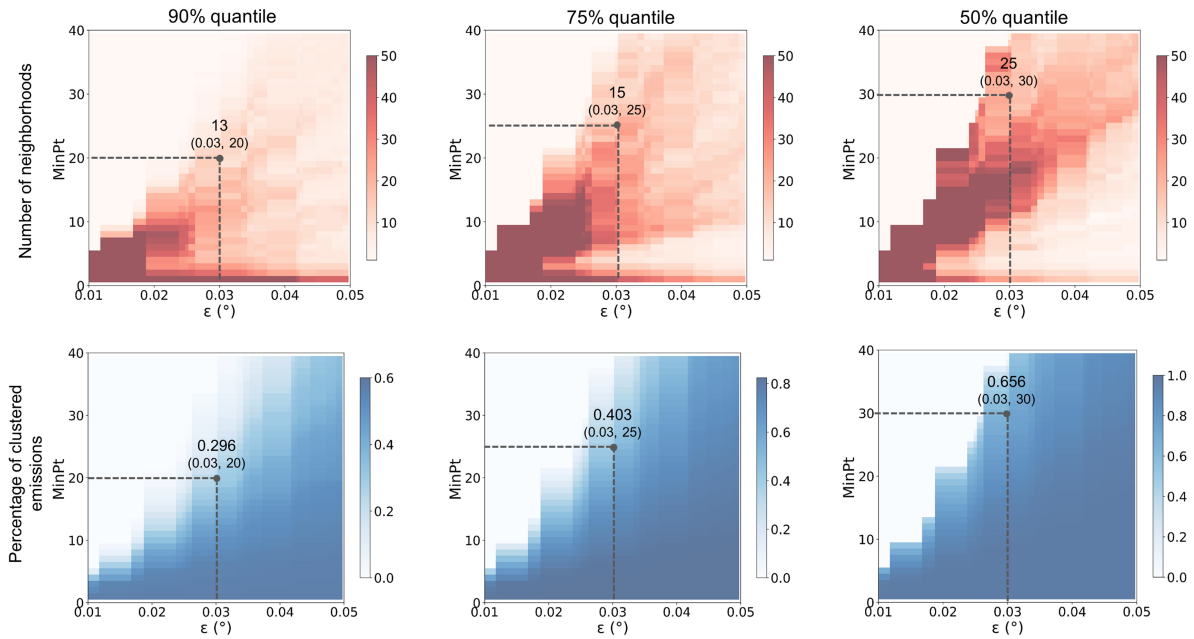


Figure 6: The process of determining hyper-parameters of DBSCAN: the maximum neighborhood-searching distance(ϵ) and the minimum sample in a neighborhood ($MinPt$).

320 Have the quantiles determined, we perform DBSCAN with grids with emissions over each
 321 quantile respectively. To minimize the uncertainty caused by parameter selection, we design a
 322 sensitivity analysis on the two main hyper-parameters of DBSCAN, the maximum neighborhood-
 323 searching distance(ϵ) and the minimum sample in a neighborhood ($MinPt$). The number of

neighborhoods and the percentage of clustered emissions, are the two metrics that we use to evaluate the rationality of clustering results. Based on daily gross emissions, we visualize how these two metrics change with different combinations of parameters (Figure 6). In general, the number of neighborhoods decreases the fastest when either ϵ or $MinPt$ increments individually. When the hyper-parameters grow synchronously, the rate of decrease slows down, and the smaller the quantile, the more obvious the decrease is. For the percentage of clustered emissions, the trend is simpler. Larger ϵ induces more clustered emissions, while $MinPt$ is the opposite. We note that extreme values of either metric can lead to unwanted clustering results. A too large number of neighborhoods could result in small and dispersive emission hotspots, while a too small one makes emission hotspots over-merged. Similarly, the percentage of clustered emissions reflects the severity of the conditions for identifying emission hotspots, which cannot be too loose nor too tight. Therefore, our strategy is to select a parameter couple that yields moderate values of both metrics at the same time. ϵ determines how stringent we define the proximity between grids, so it should be constrained by the size of the grids. Given that our grids are in 0.01° resolution, the available range of ϵ is from 0.01° to 0.05° . Finally, striking a balance among all the factors aforementioned, the hyper-parameters are settled for each quantile: 90% quantile ($\epsilon = 0.03^\circ$, $MinPt = 20$), 75% quantile ($\epsilon = 0.03^\circ$, $MinPt = 25$) and 50% quantile ($\epsilon = 0.03^\circ$, $MinPt = 30$).

With all the prerequisites set, we obtain the hierarchical emission hotspots of daily gross road CO₂ emissions in GBA (Figure 7(a)). Metrics of different levels of emission hotspots are presented in Figure 7(b). Most hotspots locate in the core area of GBA (the zoomed-in area). Statistically, the 90%-quantile hotspots emit 29.6% of the total emissions with only 5% of land in GBA. The 75%-quantile hotspots cover twice the area of the 90%-quantile hotspots, contributing to 40.3% emissions. 50%-quantile hotspots aggregate 65.7% of emissions on 23% of land. For each quantile, we plot the location and scope of every single emission hotspot with the ranking of total emissions (Figure 7(c)). The 90%-quantile hotspots include two primary components, namely the Guangzhou-Foshan metropolis and Shenzhen center, exhibiting huge emission volume that far exceeds other hotspots scattering at the major industry and transportation nodes in Guangzhou, Shenzhen, Foshan and Dongguan. The results suggest that road CO₂ emissions in Guangzhou and Foshan have formed a unified high-level emission hotspot, reflect-

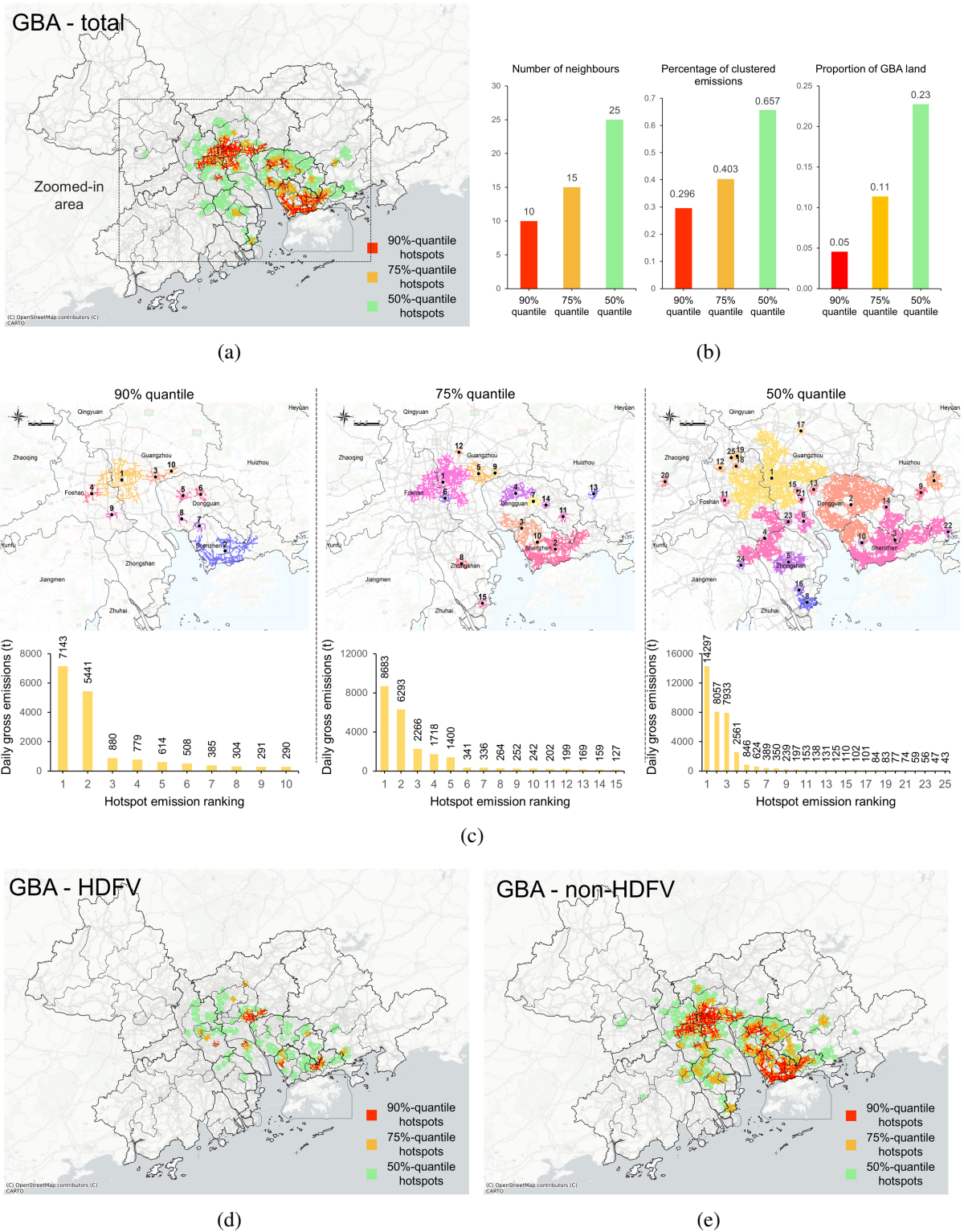


Figure 7: Maps and statistics of hierarchical emission hotspots in GBA. **(a)** Hierarchical emission hotspots of daily gross road CO₂ emissions in GBA. **(b)** Metrics of different levels of emission hotspots. **(c)** Spatial distribution and emission rankings of different levels of emission hotspots. **(d)** Hierarchical emission hotspots of daily CO₂ emissions from HDFVs in GBA. **(e)** Hierarchical emission hotspots of daily CO₂ emissions from non-HDFVs in GBA. Source of the basemap: (c) OpenStreetMap contributors.

354 ing the promotion of integrated urbanization of the two cities (Zhang et al., 2021b). For the
355 75%-quantile hotspots, the two primary hotspots retain and absorb broader area around. Some
356 adjacent 90%-quantile hotspots are merged into 75%-quantile hotspots, and glean second-tier
357 emissions. 90%-quantile hotspot 5 and 6 are integrated into 75%-quantile hotspot 4 in Dong-
358 guan downtown. 90%-quantile hotspot 7 and 8 become 75%-quantile hotspot 3 at Shenzhen-
359 Dongguan junction area. Meanwhile, some new hotspots pop up in central area of other small
360 cities, including Zhongshan (75%-quantile hotspot 8), Huizhou (75%-quantile hotspot 13) and
361 Zhuhai (75%-quantile 15). The top 50%-quantile hotspots (1 - 4) are characterized by huge foot-
362 prints. Since the 50%-quantile hotspots are obtained with nearly all the grids with emissions,
363 they reflect the spatially contiguous road CO₂ emitting zones in GBA. One observation is that
364 all the top hotspots cross city boundaries. It indicates that regional integration development
365 in GBA is promoting huge spatially continuous road emission zones. Cross-border emission
366 hotspots have been constituted between Guangzhou and Foshan, Dongguan and Shenzhen, and
367 Foshan, Zhongshan and Jiangmen, while the contiguous discharge in the east and south direc-
368 tions of Guangzhou has not yet formed. Even so, the current top 50%-hotspots have gathered
369 the majority of road CO₂ emissions. Although a large amount of small hotspots exists around
370 the top hotspots, they only contribute to a tiny proportion.

371 The disparity of agglomeration patterns between HDFV and non-HDFV emissions is high-
372 lighted (Figure 7(d), 7(e)). HDFV CO₂ emissions are significantly less clustered than the coun-
373 terpart. Even the 50%-quantile hotspots only account for 32.8% of the total HDFV emissions.
374 The 90%-quantile HDFV hotspots mainly refer to Huangpu District in east Guangzhou, Long-
375 gang District in north Shenzhen, and Baoan District in west Shenzhen. These areas are the
376 principal industrial zones in GBA. Comparatively, emission agglomeration of non-HDFV is far
377 more considerable. The hierarchical pattern resembles that of the total emissions. Non-HDFV
378 emissions are even more concentrated than the total emissions, with the three levels of hotspots
379 covering 42.5%, 61.6% and 78.9% of the total non-HDFV emissions respectively.

380 With the hyper parameters consistent, the temporal variation of hierarchical emission hotspots
381 in GBA is analyzed (Supplementary Figure S2). The patterns from 7:00 to 0:00 the next day
382 are analogous and are basically consistent with that of the daily gross emissions. From 1:00,
383 all levels of emission hotspots start to shrink in scale visibly and become the smallest at 4:00.

384 After that, emission hotspots resume spreading and finally return the full state at around 7:00.
385 As a scale metric of emission hotspots, the percentage of clustered emissions records the same
386 process of variation in a quantitative way (Supplementary Figure S2(b)). These curves have
387 a similar trend with the daily gross emissions (Figure 4(a)). However, they carry different in-
388 formation. QH-DBSCAN reflects the relative level of emission agglomeration regardless of
389 the absolute emissions. Thus, their remarkable consistency indicates that, in GBA, the rise of
390 absolute emissions would promote the level of emission agglomeration.

391 *3.4. Category of emission patterns*

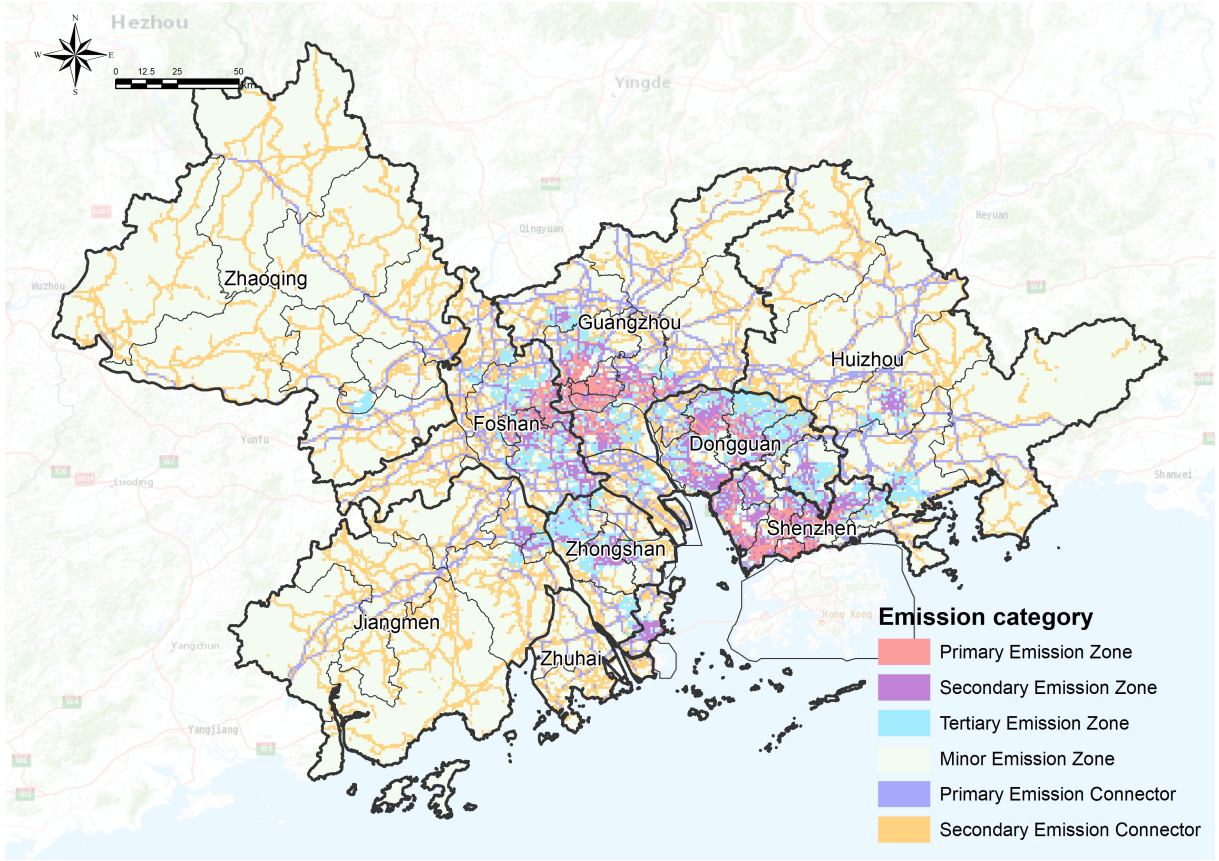
392 The preceding analysis has unfolded the road CO₂ emission patterns in GBA from the per-
393 spective of space, time, vehicle type and agglomeration. To gain an all-round insight from
394 those sides, we attempt to categorize the grid-level emission patterns. The attributes used for
395 the categorization consist of two portions. The first part indicates the overall emission quan-
396 tities, including daily gross emissions, daily gross HDFV emissions, daily gross non-HDFV
397 emissions, emission intensity and the percentage of HDFV emissions. The second one contains
398 the cases where the grids belong to different levels of emission hotspots at different times. We
399 use hotspot level rather than the absolute emissions to guide the generation of spatially continu-
400 ous emission categories, which is conducive to policy making. Since emission hotspots do not
401 vary drastically in consecutive hours, we use typical hours to represent the entire day. Focusing
402 on representative hours alleviates the multicollinearity among attributes and is instrumental to
403 interpret the results. Specifically, we choose five hours, that is the emission valley hour (4:00),
404 the two emission peak hours (8:00 and 18:00), and two transitional hours (12:00 and 0:00). The
405 hierarchical emission hotspots at each selected hour are converted into three dummy variables,
406 indicating whether the grid belongs to each level of hotspot. As a result, the process gives rise
407 to 15 dummies. In summary, the attribute set contains 20 explanatory variables. We employ
408 K-Means clustering algorithm (Hartigan and Wong, 1979) to categorize the emission patterns.
409 All the non-dummy attributes are normalized before entering the model. We experiment among
410 Z-score, min-max scaling and inverse hyperbolic sine function, and find that the last one fits the
411 best because it is capable of dealing with zero-inflated attribute sets as ours. To determine the
412 optimal hyper-parameter k (number of clusters), we conduct assessments with multiple methods
413 including elbow method (Bholowalia and Kumar, 2014), Silhouette score (Rousseeuw, 1987),

414 and gap statistics (Tibshirani et al., 2001). After testing with 100 random seeds, the optimal k
415 is set to be 6. Consequently, we obtain the spatial distribution of the six emission categories
416 in GBA (Figure 8(a)). There are two main geographical forms for the emission categories.
417 Among the six categories, four are in zone form, whilst two are in linear form. According to
418 the geographical form and the emission intensity, we define the zone categories as Primary,
419 Secondary, Tertiary and Minor Emission Zone, and name the linear categories as Primary and
420 Secondary Emission Connector. The differences in emission metrics among the categories are
421 demonstrated in Table 4.

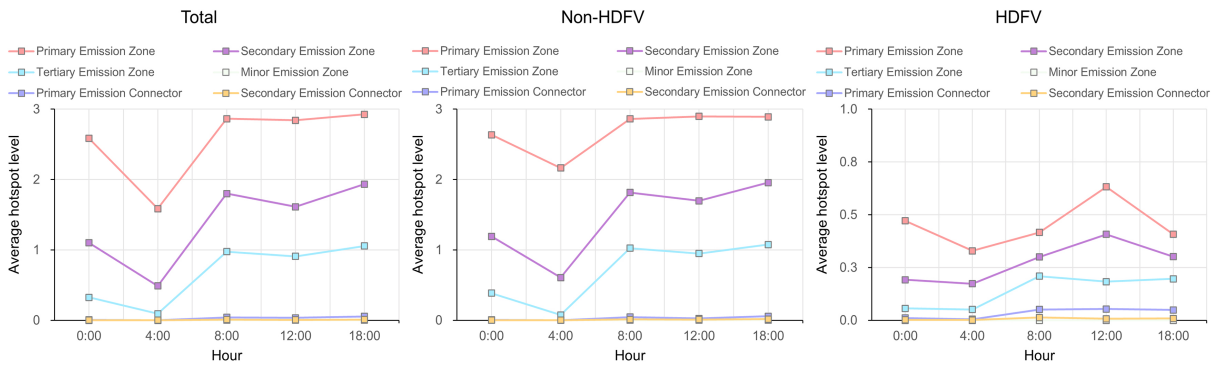
Table 4: Differences in road CO2 emissions metrics among emission categories.

Category	Population (million)	Daily gross emissions (t)	Per capita emissions (kg/person)	Emission intensity (kg/m)	Non-HDFV emissions (t)	HDFV emissions (t)	HDFV percentage (%)
Primary Emission Zone	16.46	17,002	1.03	1.63	13,850	3152	18.5
Secondary Emission Zone	16.29	12,808	0.79	0.80	8897	3911	30.5
Tertiary Emission Zone	12.46	5333	0.43	0.34	3385	1948	36.5
Minor Emission Zone	11.22	344	0.03	0.04	283	61	17.7
Primary Emission Connector	5.73	15,890	2.77	1.03	7704	8186	51.5
Secondary Emission Connector	9.88	4864	0.49	0.24	1761	3103	63.8

422 The Primary, Secondary and Tertiary Emission Zone overlap with the urban area in GBA.
423 The Primary Emission Zone includes the Guangzhou-Foshan metropolis and the main cities of
424 Shenzhen and Dongguan, which are the most developed area in GBA, accommodating 16.46
425 million residents according to the WorldPop (WorldPop, 2018). It not only generates the largest
426 total emissions (17002 t), but also possesses the topmost emission intensity (1.63 kg/m) and
427 per capita emission (1.03 kg/person). In Primary Emission Zone, non-HDFVs are responsible
428 for the majority of emissions, with only 18.5% of the total emissions from HDFVs. The Sec-
429 ondary Emission Zone points to the satellite towns around Guangzhou, Shenzhen, Dongguan
430 and Foshan, and also the major cities of Zhongshan, Huizhou, Jiangmen and Zhuhai. The total
431 population is close to the Primary Emission Zone, but less road CO2 emissions are produced
432 (12808 t) because of the lower per capita emissions (0.79 kg/person) and emission intensity
433 (0.80 kg/m). The Tertiary Emission Zone contains more outer suburbs around the Primary and
434 Secondary Emission Zone. With 12.46 million residents in the zone, the zone products 5333t
435 road CO2 emissions per day. The per capita emissions (0.43 kg/person) and emission intensity
436 (0.34 kg/m) decline further. In the meantime, HDFVs account for a higher proportion of the
437 total emissions (36.5%), which nearly doubles that of the Primary Emission Zone. The find-
438 ings unveil that population size may not be the decisive factor of road CO2 emissions. In fact,



(a)



(b)

Figure 8: Emission Category in GBA. (a) Spatial distribution of road CO₂ emission categories in GBA. (b) Temporal variations of average emission hotspot level. Source of the basemap: (c) Esri.

439 some zones emit surplus emissions because of a higher unit average emission instead of exces-
440 sive population. Besides, the emission addition is mainly triggered by non-HDFVs, implying
441 a possibly more emission-intensive mode of mobility in those high-emission zones. The broad
442 rural and natural area in GBA is mostly demarcated into the Minor Emission Zone. Road CO₂
443 emissions in this zone are comparatively negligible.

444 By contrast with zones, the rest two emission categories are characterized by a spatially
445 linear distribution along the road, especially the regional connector (highway and provincial
446 expressway). The Primary Emission Connector mainly consists of highways that radiate from
447 the emission-intensive zones. These connectors produce 15890 t road CO₂ a day, which is
448 second only to the volume of Primary Emission Zone. This emission category involves the least
449 residents and has the highest per capita emissions (2.77 kg/person). Most emissions are related
450 to inter-city long-distance travel or freight transportation. HDFVs contribute to 51.5% of the
451 total emissions, which is considerably larger than those of the emission zones. The Secondary
452 Emission Connector refers to the provincial expressways that primarily serve intra-city medium-
453 distance mobility. The daily gross emissions and the emission intensity are much lower than the
454 Primary Emission Connector. However, the HDFV percentage (63.8%) is the highest among
455 all the categories. On emission connectors, HDFVs appear to be the dominant emitter, with
456 55.4% of the daily gross HDFV emissions concentrating on Primary and Secondary Emission
457 Connector.

458 To further interpret the differences in the emission hotspot level among emission categories
459 across the time, we conduct the following analysis. For grids included in multiple levels of emis-
460 sion hotspots, the highest level is taken. For each grid, a numerical score is used to represent
461 the highest hotspot level (90%-quantile hotspot— 3, 75%-quantile hotspot— 2, 50%-quantile
462 hotspot — 1, no hotspot — 0). For each category, its average emission hotspot level is reflected
463 by the mean score across the grids. In this way, we obtain the temporal variations of the av-
464 erage hotspot level in Figure 8(b). Regarding the total emissions, only the Primary, Secondary
465 and Tertiary Emission Zone exhibit significant emission agglomeration. Grids in the Primary
466 Emission Zone mostly belong to 90%-quantile hotspots from 8:00 to 18:00, while one level of
467 downgrade occurs at 4:00. Similar trend is also observed for the Secondary and Tertiary Emis-
468 sion Zone with lower overall levels. The average hotspot level of Non-HDFV emissions has

469 analogous characteristics with the total emissions. In contrast, HDFV emission agglomeration
 470 is less considerable across all the categories, with average hotspot levels lower than 1 during
 471 the day.

472 4. Discussion

473 4.1. Validation of emission estimation

Table 5: Comparison of estimation results between this study and the literature.

Study area	Road type	Study	Year	Data Source	Approach	CO2 Emissions (Mt/year)
Guangdong Province	All	Jia et al. (2018)	2011	Statistical Yearbooks	Bottom-up	84.5
		Jia et al. (2018)	2012	Statistical Yearbooks	Bottom-up	94.1
		Guo et al. (2014)	2012	Statistical Yearbooks	Top-down	56.82
		Jia et al. (2018)	2013	Statistical Yearbooks	Bottom-up	101.9
		Jia et al. (2018)	2014	Statistical Yearbooks	Bottom-up	109.4
		Jia et al. (2018)	2015	Statistical Yearbooks	Bottom-up	112.5
		Crippa et al. (2021)	2018	Energy Balances	Top-down	90.9
		Guan et al. (2021)	2019	Statistical Yearbooks	Top-down	61.17
		Xu et al. (2021)	2019	Statistical Yearbooks	Bottom-up	50.1
		Gao et al. (2022)	2020	China's continuous emissions monitoring systems (CEMS)	Top-down	70.15
Guangdong Province	Highway	Liu et al. (2022)	2021	EDGAR transport emissions	Top-down	80.3
		This study	2020	Vehicle trajectories	Bottom-up	27.28
		Li et al. (2022)	2021	Highway toll data	Bottom-up	5.59
Futian and Nanshan, Shenzhen	All	This study	2020	Vehicle trajectories	Bottom-up	11.26
		Zhou et al. (2022b)	2017	Vehicle trajectories	Bottom-up	0.03
		This study	2020	Vehicle trajectories	Bottom-up	0.84

474 Emission estimation is always vulnerable to the method and data source used. To validate
 475 our results and discover the pros and cons, we compare our results with other studies using
 476 different datasets and methods (Table 5). We make necessary conversions on our results to
 477 guarantee the consistency of study scenario (year, study area, road type and emission type) with
 478 previous studies and ensure the comparability. It turns out that road CO2 emission estimates
 479 vary significantly across studies. Top-down and bottom-up approach can lead to very different
 480 results using similar data source. Besides, data source also has a profound impact. Estimates
 481 based on trajectory data are commonly lower than those based on collective data. As the first
 482 study that uses vehicle trajectories to estimate daily gross road CO2 emissions in Guangdong,
 483 our estimates are considerably lower than those obtained from collective data such as China
 484 Statistical Yearbook and Energy Balances. However, it cannot be assumed that the sampling rate
 485 of our dataset is overrated, because our study yields higher emission estimates than other studies
 486 using similar dataset such as highway toll data ([Li et al., 2022](#)) or vehicle trajectories ([Zhou
 487 et al., 2022b](#)) at the same time. [Li et al. \(2022\)](#) harness full-sample O-D pairs between highway

488 tolls and route simulations rather than real trajectories to estimate highway CO₂ emissions.
489 Their lower estimates may root from overlooking accidents and detours that often happen in
490 real life. [Zhou et al. \(2022b\)](#) use real trajectories in 2017 with limited samples. Considering
491 the three-year time difference and the fact that our dataset contains over 16 times as many
492 samples as theirs in the same area, the difference in results is generally acceptable. Even though
493 most inputs are fixed, uncertainty of estimates can also be introduced by different empirical
494 schemes adopted to determine the emission factor, vehicle kilometer traveled, fuel consumption,
495 etc.. Therefore, in general, the estimates of this study lie in an appropriate range, and they
496 compensate the lack of road carbon emission estimation in Guangdong using individual level
497 activity data.

498 *4.2. Policy implications*

499 Based on the findings above, we intend to recommend some targeted measures to mitigate
500 road CO₂ emissions in our study area, especially in GBA.

501 Several universal strategies can be widely applied. We outline the necessity of periodically
502 supervising fine-grained regional road carbon emissions with intelligent systems and all sorts
503 of big data approaches ([Wang et al., 2022a](#)). Our study suggests that special attention should be
504 given to the core area of Guangzhou-Foshan metropolis and Shenzhen where emissions are the
505 most intensive and clustered. The continuously updated database lays the foundation for down-
506 stream applications, such as emission pattern analysis and emission reduction policy making.
507 This study enriches the technical toolkits for both the supervision side and the application side,
508 by proposing methods to handle trajectory data with deficiency in time granularity, identify
509 emission hotspots, and categorize emission zones. Besides, our research confirms that raising
510 vehicle emission standards and promoting new energy vehicles is instrumental to reduce over-
511 all road carbon emissions ([Zhang et al., 2022](#)). HDFVs are found to be associated with much
512 more intensive emissions per vehicle possession than non-HDFVs (Supplementary Table S8).
513 Thus, HDFVs, especially those with higher accumulated mileage ([Wang et al., 2022b](#)), should
514 be prioritized for emission standard upgrade or electrification.

515 Furthermore, considering that resources are limited and extensive policies may cause nega-
516 tive externalities, emission reduction policies should be adaptive to local emission patterns ([Cai](#)
517 [et al., 2018](#)). To achieve that, we introduce policy zones where targeted measures are proposed,

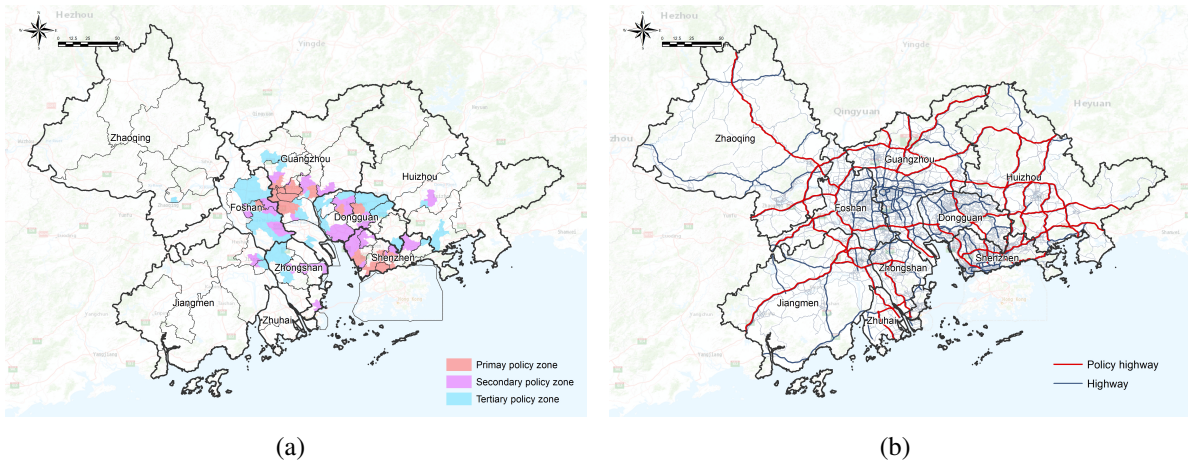


Figure 9: Maps for policy implications. **(a)** Spatial distribution of emission policy zones **(b)** Spatial distribution of policy highways.

518 based on the spatial distribution of emission categories. Our methodology generates spatially
 519 continuous regions with similar emission patterns, which is inductive to defining policy zones.
 520 To further facilitate policy implementation, we use street (the smallest administrative division
 521 in China) as the unit to divide policy zones. The emission category of a street is determined
 522 by the emission category of the grids covering the largest area within the street. In line with
 523 the three major emission zones (Figure 8(a), Table 4), we define the Primary, Secondary and
 524 Tertiary Policy Zone (Figure 9(a)). The spatial distribution of emission zones and policy zones
 525 are similar. Primary Emission Zone is characterized by the highest emission intensity but low
 526 proportions from HDFVs. The measures should accordingly focus on mitigating emissions
 527 from daily travel. Considering that areas within Primary Emission Zone are equipped with
 528 comparatively more advanced public transportation resources in GBA, we suggest encouraging
 529 sustainability-oriented travel mode choice from two directions. First, more people should be
 530 guided to use the public transits by improving the competitiveness of the system. For exam-
 531 ple, the transportation department could optimize the bus lines and subway operations within
 532 this zone to achieve more volume and better accessibility. Second, government could toll extra
 533 emission taxes on personal vehicles driving in the Primary Policy Zone. Under our estimation
 534 regarding road CO₂ emission reduction, converting residents in the Primary Policy Zone to
 535 public transits is five times more efficient than converting the same number of random people
 536 in GBA. The Secondary Policy Zone indicates the streets with most of their area categorized
 537 as Secondary Emission Zone whose emission intensity halved but proportions from HDFVs
 538 doubled compared with the Primary Emission Zone. Although some level of off-peak HDFV

539 delivery has been observed (Figure 4(c), 4(d)), there are still apparent overlaps of high emis-
540 sions from HDFVs and non-HDFVs during the period between the morning and evening peaks.
541 Therefore, we recommend enacting incentive policies such as midnight toll discount to guide
542 HDFVs to reschedule their itinerary and operate during low-emission hours (22:00 - 6:00). Pro-
543 moting off-peak delivery could alleviate the emission surplus caused by traffic congestion at
544 peak hours (Chen et al., 2022). Last, Tertiary Policy Zone mainly indicates the marginal streets
545 around the major urban areas. Although both the local emission volume and intensity are much
546 less, the nonnegligible population size and the geographic adjacency with the major urban ar-
547 eas may mean heavy external traffics. We would recommend more detailed future studies to
548 uncover the pattern of emissions from traffics between these streets and the major urban ar-
549 eas. Some measures might be necessary to reduce emission surplus from long-distance vehicle
550 travels.

551 The emission category, Primary Emission Connector, mainly consists of the inter-city high-
552 ways with intensive emissions and a high proportion from HDFVs. We highlight these highways
553 as the “policy highway”, which should be prioritized for carbon reduction policies (Figure 9(b)).
554 We advocate promoting electrified HDFVs to run the transportation routes on these highways.
555 Some policy and infrastructure facilities could be provided on these highways for electrified
556 HDFVs, such as more convenient and affordable charging services and lower passing tolls.
557 Also, we note that many of the policy highways are those radiating from Guangzhou and Shen-
558 zhen. Thus, another strategy could be strengthening the construction and integration of regional
559 freight railway infrastructure around these two cities, to divert corresponding transportation
560 needs from roads to railways.

561 4.3. *Limitations and future studies*

562 The findings of this study have to be seen in light of the following limitations. The ve-
563 hicle trajectory data used in our study is at minute-level granularity, which may still induce
564 uncertainty in traffic status and CO₂ emission estimation, although some corrections are im-
565 plemented. Besides, due to limited data availability, this analysis is conducted with data in a
566 single day. Panel data should be helpful to examine the emission fluctuation in a longer time
567 span. Since the vehicle type, fuel type and emission standard of each trajectory are unavailable,
568 these factors are considered to be constant across all the road segments. We suggest using more

569 advanced data sources and local emission model to validate our results. Besides, our results
570 identify a considerably higher proportion of HDFV CO₂ emissions in non-GBA cities, which
571 might be partly attributed to the possibility that drivers in these cities count less on navigation
572 during daily travel. It means the possible spatial heterogeneity in the sampling rates, which may
573 influence trajectory-based emission estimation but remains under-explored in the literature.

574 This study obtains emission category zones in GBA and accordingly divides policy zones
575 with targeted carbon reduction strategies. We are aware that policy making is a rigorous and
576 complicated process where multiple aspects should be considered. Due to space limitations, this
577 study focuses more on emission pattern analysis, and thus does not involve other perspectives
578 such as demography, land use, finance, and so on. A more detailed depiction on the demo-
579 graphic structure in each emission category zone may help to understand the status quo better
580 and formulate more solid strategies (Deng et al., 2021; Zhang et al., 2021a). Other indicators
581 regarding urban form (Shi et al., 2020; Wang et al., 2017; Chen et al., 2021; Biljecki and Chow,
582 2022) and built environment (Cao and Yang, 2017) could also be taken into consideration. Fur-
583 thermore, distinguishing the emission patterns of different vehicle usages and trip purposes
584 (Zhao et al., 2017) is also instrumental to deepen the understanding of road emissions. Our
585 follow-up research would be intended to remedy these deficiencies.

586 **5. Conclusions**

587 Based on massive vehicle trajectories, this study demonstrates the spatiotemporal pattern
588 of road CO₂ emissions in Guangdong. Emissions are estimated per road segment and per
589 hour, with both broad spatial coverage and fine granularity. Overall, GBA produces 76% of
590 the total emissions with only 30% of land in Guangdong. Guangzhou is the primary emitter
591 in Guangdong, producing 15621 t CO₂ and contributing to 21% of the provincial emissions.
592 Shenzhen has the highest average emission intensity (1.01 kg/m). Most GBA cities rank high in
593 road CO₂ emissions except for Zhuhai. Compared with GBA cities, non-GBA cities have twice
594 the percentage of HDFV emissions. Regarding the road class, highway appears to generate the
595 most emissions, accounting for 41% of the total. Road classes serving local transportation also
596 contribute to 50% of the total emissions. Temporally, we discover a “three-stage” pattern for all
597 the cities, while the trend of non-GBA cities is flatter. We use the contribution ratio of the top
598 10% emitting road segments to quantify the level of emission agglomeration. Top 10% emitting

599 roads discharge 77.6% of the total emissions in Guangzhou. The contribution ratio culminates
600 at the emission valley hour. Besides, GBA cities exhibit more significant contribution ratio.

601 We propose QH-DBSCAN to detect hierarchical emission hotspots with emissions projected
602 onto 1 km grids in GBA, identifying the location and scope of emission hotspots. Then we
603 demonstrate a comprehensive sensitivity analysis to examine the influence of hyper-parameter
604 selection for the method. The results show that the 90%-quantile hotspots emit 29.6% of the
605 daily gross road CO₂ emissions with only 5% of land in GBA. The associated regions refer to
606 the Guangzhou-Foshan metropolis and Shenzhen center. The 50%-quantile emission hotspots
607 cover most urban areas in GBA, presenting cross-border integration development between ad-
608 jacent cities represented by Guangzhou and Foshan, and Shenzhen and Dongguan. HDFV CO₂
609 emissions show less significant agglomeration than the non-HDFV counterpart. The temporal
610 variation of emission hotspots is highly synchronized with the daily gross emissions. Differ-
611 ent interactions among the three levels of emission hotspots are found in high-emission and
612 low-emission periods.

613 We aggregate all aspects of emission features and derive six emission categories, including
614 four emission zones and two emission connectors. The Primary Emission Zone refers to the
615 most developed urban area in GBA, which produces the largest total emissions (30% of GBA
616 total). It has the topmost emission intensity but the lowest percentage of HDFV emissions
617 in emission zones. The Secondary and Tertiary Emission Zone emit less intensively but have
618 a higher percentage of HDFV emissions. The Primary Emission Connector mainly contains
619 the inter-city highways radiating from the emission-intensive zones, with the highest emission
620 intensity and over 50% of emissions contributed by HDFVs. On the ground of the differences of
621 emission patterns among emission categories, we propose policy zones and recommend targeted
622 strategies for carbon reduction.

623 In summary, the study demonstrate a bottom-up road CO₂ estimation method based on a
624 massive vehicle trajectory dataset. The results unveil the spatiotemporal pattern of road CO₂
625 emission in Guangdong and emission agglomeration in GBA. This study expands the toolkit of
626 regional emission studies and offers insightful findings that support the carbon reduction policy
627 making and sustainable development in Guangdong and especially GBA.

628 **Funding**

629 This research was supported by Guangdong Enterprise Key Laboratory for Urban Sensing,
630 Monitoring and Early Warning (No.2020B121202019) and The Science and Technology Foun-
631 dation of Guangzhou Urban Planning & Design Survey Research Institute (RDI2220205141).

632 **CRedit authorship contribution statement**

633 **Xingdong Deng:** Conceptualization, Methodology, Funding acquisition, Writing - Ori-
634 ginal draft preparation. **Wangyang Chen:** Writing- Original draft preparation, Writing - Review
635 & Editing, Software. **Qingya Zhou:** Writing - Review & Editing, Investigation, Validation,
636 Supervision. **Yuming Zheng:** Writing- Original draft preparation, Formal analysis, Visualiza-
637 tion. **Hongbao Li:** Writing - Review & Editing, Supervision. **Shunyi Liao:** Resources, Project
638 administration. **Filip Biljecki:** Conceptualization, Writing - Review & Editing.

639 **Declarations of interest**

640 None

641 **Acknowledgments**

642 We sincerely appreciate the kindly help from the editors and reviewers to improve this study.
643 We also acknowledge the data sources and the open-source packages used in this study.

644 **References**

- 645 Batur, İ., Bayram, I.S., Koc, M., 2019. Impact assessment of supply-side and demand-side policies on energy
646 consumption and co2 emissions from urban passenger transportation: The case of istanbul. *Journal of Cleaner*
647 *Production* 219, 391–410.
- 648 Bholowalia, P., Kumar, A., 2014. Ebk-means: A clustering technique based on elbow method and k-means in wsn.
649 *International Journal of Computer Applications* 105.
- 650 Biljecki, F., Chow, Y.S., 2022. Global Building Morphology Indicators. *Computers, Environment and Urban*
651 *Systems* 95, 101809.
- 652 Böhm, M., Nanni, M., Pappalardo, L., 2022. Gross polluters and vehicle emissions reduction. *Nature Sustainability*
653 , 1–9.
- 654 Boulter, P., Barlow, T., McCrae, I., Latham, S., 2009. Emission factors 2009: Final summary report. TRL Published
655 Project Report .

656 Cai, B., Guo, H., Cao, L., Guan, D., Bai, H., 2018. Local strategies for china's carbon mitigation: An investigation
657 of chinese city-level co2 emissions. *Journal of Cleaner Production* 178, 890–902.

658 Cao, X., Yang, W., 2017. Examining the effects of the built environment and residential self-selection on commut-
659 ing trips and the related co2 emissions: An empirical study in guangzhou, china. *Transportation Research Part*
660 *D: Transport and Environment* 52, 480–494.

661 Carslaw, D.C., Goodman, P.S., Lai, F.C., Carsten, O.M., 2010. Comprehensive analysis of the carbon impacts of
662 vehicle intelligent speed control. *Atmospheric Environment* 44, 2674–2680.

663 Chen, L., Xu, L., Yang, Z., 2017. Accounting carbon emission changes under regional industrial transfer in an
664 urban agglomeration in china's pearl river delta. *Journal of Cleaner Production* 167, 110–119.

665 Chen, W., Wu, A.N., Biljecki, F., 2021. Classification of urban morphology with deep learning: Application on
666 urban vitality. *Computers, Environment and Urban Systems* 90, 101706.

667 Chen, X., Jiang, L., Xia, Y., Wang, L., Ye, J., Hou, T., Zhang, Y., Li, M., Li, Z., Song, Z., et al., 2022. Quanti-
668 fying on-road vehicle emissions during traffic congestion using updated emission factors of light-duty gasoline
669 vehicles and real-world traffic monitoring big data. *Science of the Total Environment* 847, 157581.

670 Chen, Z., Li, Y., Wang, P., 2020. Transportation accessibility and regional growth in the greater bay area of china.
671 *Transportation Research Part D: Transport and Environment* 86, 102453.

672 Crippa, M., Guizzardi, D., Muntean, M., Schaaf, E., Lo Vullo, E., Solazzo, E., Monforti-Ferrario, F., Olivier, J.,
673 Vignati, E.E., 2021. v6. 0 global greenhouse gas emissions. European Commission, Joint Research Centre
674 (JRC): Vienna, Austria .

675 Deng, F., Lv, Z., Qi, L., Wang, X., Shi, M., Liu, H., 2020. A big data approach to improving the vehicle emission
676 inventory in china. *Nature communications* 11, 1–12.

677 Deng, X., Liu, Y., Gao, F., Liao, S., Zhou, F., Cai, G., 2021. Spatial distribution and mechanism of urban occupation
678 mixture in guangzhou: An optimized geodetector-based index to compare individual and interactive effects.
679 *ISPRS International Journal of Geo-Information* 10, 659.

680 Ester, M., Kriegel, H.P., Sander, J., Xu, X., et al., 1996. A density-based algorithm for discovering clusters in large
681 spatial databases with noise., in: *kdd*, pp. 226–231.

682 Fang, C., Yu, D., 2017. Urban agglomeration: An evolving concept of an emerging phenomenon. *Landscape and*
683 *urban planning* 162, 126–136.

684 Fujita, M., Thisse, J.F., 1996. Economics of agglomeration. *Journal of the Japanese and international economies*
685 10, 339–378.

686 Gao, Y., Zhang, L., Huang, A., Kou, W., Bo, X., Cai, B., Qu, J., 2022. Unveiling the spatial and sectoral char-
687 acteristics of a high-resolution emission inventory of co2 and air pollutants in china. *Science of The Total*
688 *Environment* 847, 157623.

689 Guan, Y., Shan, Y., Huang, Q., Chen, H., Wang, D., Hubacek, K., 2021. Assessment to china's recent emission
690 pattern shifts. *Earth's Future* 9, e2021EF002241.

691 Guangdong Province Statistical Bureau, 2021. *Guangdong Statistical Yearbook: 2021*. China Statistics Publishers.

692 Guo, B., Geng, Y., Franke, B., Hao, H., Liu, Y., Chiu, A., 2014. Uncovering china's transport co2 emission patterns

693 at the regional level. *Energy Policy* 74, 134–146.

694 Hartigan, J.A., Wong, M.A., 1979. Algorithm as 136: A k-means clustering algorithm. *Journal of the royal*
695 *statistical society. series c (applied statistics)* 28, 100–108.

696 Hicks, W., Beevers, S., Tremper, A.H., Stewart, G., Priestman, M., Kelly, F.J., Lanoisellé, M., Lowry, D., Green,
697 D.C., 2021. Quantification of non-exhaust particulate matter traffic emissions and the impact of covid-19 lock-
698 down at london marylebone road. *Atmosphere* 12, 190.

699 Huang, L., Krigsvoll, G., Johansen, F., Liu, Y., Zhang, X., 2018. Carbon emission of global construction sector.
700 *Renewable and Sustainable Energy Reviews* 81, 1906–1916.

701 Hui, E.C., Li, X., Chen, T., Lang, W., 2020. Deciphering the spatial structure of china’s megacity region: A new
702 bay area—the guangdong-hong kong-macao greater bay area in the making. *Cities* 105, 102168.

703 IEA, G.E., 2019. Co2 emissions from fuel combustion. International Energy Agency, Paris .

704 IEA, G.E., 2020. Co2 emissions from fuel combustion. International Energy Agency, Paris .

705 Jia, T., Li, Q., Shi, W., 2018. Estimation and analysis of emissions from on-road vehicles in mainland china for
706 the period 2011–2015. *Atmospheric Environment* 191, 500–512.

707 Kan, Z., Tang, L., Kwan, M.P., Ren, C., Liu, D., Pei, T., Liu, Y., Deng, M., Li, Q., 2018. Fine-grained analysis on
708 fuel-consumption and emission from vehicles trace. *Journal of cleaner production* 203, 340–352.

709 Li, T., Wu, J., Dang, A., Liao, L., Xu, M., 2019a. Emission pattern mining based on taxi trajectory data in beijing.
710 *Journal of cleaner production* 206, 688–700.

711 Li, Y., Wu, Q., Zhang, Y., Huang, G., Jin, S., Fang, S., 2022. Mapping highway mobile carbon source emissions
712 using traffic flow big data: A case study of guangdong province, china. *Frontiers in Energy Research* , 496.

713 Li, Y., Zheng, J., Dong, S., Wen, X., Jin, X., Zhang, L., Peng, X., 2019b. Temporal variations of local traffic co2
714 emissions and its relationship with co2 flux in beijing, china. *Transportation Research Part D: Transport and*
715 *Environment* 67, 1–15.

716 Lin, B., Li, Z., 2020. Spatial analysis of mainland cities’ carbon emissions of and around guangdong-hong kong-
717 macao greater bay area. *Sustainable Cities and Society* 61, 102299.

718 Liu, B., Zimmerman, N., 2021. Fleet-based vehicle emission factors using low-cost sensors: Case study in parking
719 garages. *Transportation Research Part D: Transport and Environment* 91, 102635.

720 Liu, Z., Deng, Z., Zhu, B., Ciais, P., Davis, S.J., Tan, J., Andrew, R.M., Boucher, O., Arous, S.B., Canadell, J.G.,
721 et al., 2022. Global patterns of daily co2 emissions reductions in the first year of covid-19. *Nature Geoscience*
722 15, 615–620.

723 Lomas, K.J., Bell, M., Firth, S., Gaston, K.J., Goodman, P., Leake, J.R., Namdeo, A., Rylatt, M., Allinson, D.,
724 Davies, Z.G., et al., 2010. 4m: Measurement, modelling, mapping and management—the carbon footprint of uk
725 cities .

726 Malmberg, A., Maskell, P., 1997. Towards an explanation of regional specialization and industry agglomeration.
727 *European planning studies* 5, 25–41.

728 Manzoni, V., Maniloff, D., Kloeckl, K., Ratti, C., 2010. Transportation mode identification and real-time co2
729 emission estimation using smartphones. SENSEable City Lab, Massachusetts Institute of Technology, nd .

730 McQueen, M., MacArthur, J., Cherry, C., 2020. The e-bike potential: Estimating regional e-bike impacts on
731 greenhouse gas emissions. *Transportation Research Part D: Transport and Environment* 87, 102482.

732 MEE, 2019. *China Mobile Source Environmental Management Annual Report (2019)*. Ministry of Ecology and
733 Environment of the People's Republic of China.

734 MEE, 2021. *China Mobile Source Environmental Management Annual Report (2021)*. Ministry of Ecology and
735 Environment of the People's Republic of China.

736 Mohsin, M., Abbas, Q., Zhang, J., Ikram, M., Iqbal, N., 2019. Integrated effect of energy consumption, economic
737 development, and population growth on co₂ based environmental degradation: a case of transport sector. *En-
738 vironmental Science and Pollution Research* 26, 32824–32835.

739 Patiño-Aroca, M., Parra, A., Borge, R., 2022. On-road vehicle emission inventory and its spatial and temporal
740 distribution in the city of guayaquil, ecuador. *Science of The Total Environment* 848, 157664.

741 Pérez-Martínez, P., Miranda, R., Andrade, M., 2020. Freight road transport analysis in the metro são paulo:
742 Logistical activities and co₂ emissions. *Transportation Research Part A: Policy and Practice* 137, 16–33.

743 Pla, M.A.M., Lorenzo-Sáez, E., Luzuriaga, J.E., Prats, S.M., Moreno-Pérez, J.A., Urchueguía, J.F., Oliver-
744 Villanueva, J.V., Lemus, L.G., 2021. From traffic data to ghg emissions: A novel bottom-up methodology
745 and its application to valencia city. *Sustainable Cities and Society* 66, 102643.

746 Rousseeuw, P.J., 1987. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *Journal
747 of computational and applied mathematics* 20, 53–65.

748 Shan, Y., Fang, S., Cai, B., Zhou, Y., Li, D., Feng, K., Hubacek, K., 2021. Chinese cities exhibit varying degrees
749 of decoupling of economic growth and co₂ emissions between 2005 and 2015. *One Earth* 4, 124–134.

750 Shi, K., Xu, T., Li, Y., Chen, Z., Gong, W., Wu, J., Yu, B., 2020. Effects of urban forms on co₂ emissions in china
751 from a multi-perspective analysis. *Journal of environmental management* 262, 110300.

752 Shindell, D.T., Levy, H., Schwarzkopf, M.D., Horowitz, L.W., Lamarque, J.F., Faluvegi, G., 2008. Multimodel
753 projections of climate change from short-lived emissions due to human activities. *Journal of Geophysical
754 Research: Atmospheres* 113.

755 Sobrino, N., Arce, R., 2021. Understanding per-trip commuting co₂ emissions: A case study of the technical
756 university of madrid. *Transportation Research Part D: Transport and Environment* 96, 102895.

757 Sui, Y., Zhang, H., Song, X., Shao, F., Yu, X., Shibasaki, R., Sun, R., Yuan, M., Wang, C., Li, S., et al., 2019.
758 Gps data in urban online ride-hailing: A comparative analysis on fuel consumption and emissions. *Journal of
759 Cleaner Production* 227, 495–505.

760 Sun, Z., Hao, P., Ban, X.J., Yang, D., 2015. Trajectory-based vehicle energy/emissions estimation for signalized
761 arterials using mobile sensing data. *Transportation Research Part D: Transport and Environment* 34, 27–40.

762 Tatem, A.J., 2017. Worldpop, open data for spatial demography. *Scientific data* 4, 1–4.

763 Tibshirani, R., Walther, G., Hastie, T., 2001. Estimating the number of clusters in a data set via the gap statistic.
764 *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 63, 411–423.

765 Van Fan, Y., Perry, S., Klemeš, J.J., Lee, C.T., 2018. A review on air emissions assessment: Transportation. *Journal
766 of cleaner production* 194, 673–684.

- 767 Wang, F., Wang, G., Liu, J., Chen, H., 2019. How does urbanization affect carbon emission intensity under a
768 hierarchical nesting structure? empirical research on the china yangtze river delta urban agglomeration. *Envi-
769 ronmental Science and Pollution Research* 26, 31770–31785.
- 770 Wang, J., Cai, B., Zhang, L., Cao, D., Liu, L., Zhou, Y., Zhang, Z., Xue, W., 2014. High resolution carbon
771 dioxide emission gridded data for china derived from point sources. *Environmental science & technology* 48,
772 7085–7093.
- 773 Wang, L., Chen, X., Xia, Y., Jiang, L., Ye, J., Hou, T., Wang, L., Zhang, Y., Li, M., Li, Z., et al., 2022a. Operational
774 data-driven intelligent modelling and visualization system for real-world, on-road vehicle emissions—a case
775 study in hangzhou city, china. *Sustainability* 14, 5434.
- 776 Wang, M., Madden, M., Liu, X., 2017. Exploring the relationship between urban forms and co₂ emissions in 104
777 chinese cities. *Journal of Urban Planning and Development* 143, 04017014.
- 778 Wang, Y., Xing, Z., Zhang, H., Wang, Y., Du, K., 2022b. On-road mileage-based emission factors of gaseous
779 pollutants from bi-fuel taxi fleets in china: The influence of fuel type, vehicle speed, and accumulated mileage.
780 *Science of The Total Environment* 819, 151999.
- 781 WorldPop, 2018. Global 1km population. doi:[10.5258/SOTON/WP00647](https://doi.org/10.5258/SOTON/WP00647).
- 782 Xia, C., Xiang, M., Fang, K., Li, Y., Ye, Y., Shi, Z., Liu, J., 2020. Spatial-temporal distribution of carbon emissions
783 by daily travel and its response to urban form: A case study of hangzhou, china. *Journal of Cleaner Production*
784 257, 120797.
- 785 Xu, Y., Liu, Z., Xue, W., Yan, G., Shi, X., Zhao, D., Zhang, Y., Lei, Y., Wang, J., 2021. Identification of on-
786 road vehicle co₂ emission pattern in china: A study based on a high-resolution emission inventory. *Resources,
787 Conservation and Recycling* 175, 105891.
- 788 Yan, D., Lei, Y., Li, L., Song, W., 2017. Carbon emission efficiency and spatial clustering analyses in china's
789 thermal power industry: Evidence from the provincial level. *Journal of Cleaner Production* 156, 518–527.
- 790 Yang, W., Li, T., Cao, X., 2015. Examining the impacts of socio-economic factors, urban form and transportation
791 development on co₂ emissions from transportation in china: a panel data analysis of china's provinces. *Habitat
792 International* 49, 212–220.
- 793 Yu, H., 2019. The guangdong-hong kong-macau greater bay area in the making: development plan and challenges.
794 *Cambridge Review of International Affairs* , 1–29.
- 795 Yu, X., Wu, Z., Zheng, H., Li, M., Tan, T., 2020. How urban agglomeration improve the emission efficiency? a
796 spatial econometric analysis of the yangtze river delta urban agglomeration in china. *Journal of environmental
797 management* 260, 110061.
- 798 Zhang, J., Jia, R., Yang, H., Dong, K., 2022. Does electric vehicle promotion in the public sector contribute to
799 urban transport carbon emissions reduction? *Transport Policy* 125, 151–163.
- 800 Zhang, X., Gao, F., Liao, S., Zhou, F., Cai, G., Li, S., 2021a. Portraying citizens' occupations and assessing urban
801 occupation mixture with mobile phone data: A novel spatiotemporal analytical framework. *ISPRS International
802 Journal of Geo-Information* 10, 392.
- 803 Zhang, X., Shen, J., Gao, X., 2021b. Towards a comprehensive understanding of intercity cooperation in china's

804 city-regionalization: A comparative study of shenzhen-hong kong and guangzhou-foshan city groups. *Land*
805 *Use Policy* 103, 105339.

806 Zhang, Y.J., Da, Y.B., 2015. The decomposition of energy-related carbon emission and its decoupling with eco-
807 nomic growth in china. *Renewable and Sustainable Energy Reviews* 41, 1255–1266.

808 Zhang, Y.J., Liu, Z., Zhang, H., Tan, T.D., 2014. The impact of economic growth, industrial structure and urban-
809 ization on carbon emission intensity in china. *Natural hazards* 73, 579–595.

810 Zhao, P., Kwan, M.P., Qin, K., 2017. Uncovering the spatiotemporal patterns of co2 emissions by taxis based on
811 individuals' daily travel. *Journal of Transport Geography* 62, 122–135.

812 Zhou, X., Bai, L., Bai, J., Tian, Y., Li, W., 2022a. Scenario prediction and critical factors of co2 emissions in the
813 pearl river delta: A regional imbalanced development perspective. *Urban Climate* 44, 101226.

814 Zhou, X., Wang, H., Huang, Z., Bao, Y., Zhou, G., Liu, Y., 2022b. Identifying spatiotemporal characteristics and
815 driving factors for road traffic co2 emissions. *Science of The Total Environment* 834, 155270.

816 Zhou, Y., Shan, Y., Liu, G., Guan, D., 2018. Emissions and low-carbon development in guangdong-hong kong-
817 macao greater bay area cities and their surroundings. *Applied energy* 228, 1683–1692.