

A review of on-road vehicle emission inventory

Daoyuan Yang^{1,2}, Yonglin Zhang^{1,2,*}, Rui Wu^{1,2}, Yuanyuan Song^{1,2}, and Honglei Xu^{1,2}

¹Transport Planning and Research Institute, Ministry of Transport, Beijing 100028, PR China

²Laboratory of Transport Pollution Control and Monitoring Technology, Beijing, 100028, PR China

Keywords: Vehicle emission inventory, Emission factor model, Traffic demand models, Machine learning, Traffic big data.

Abstract. The large increase in the on-road vehicle population in China has raised sustainability concerns regarding air pollution prevention, energy conservation, and climate change mitigation. Vehicle emission inventory is an irreplaceable tool to characterize the temporal and spatial distribution of the air pollutant and provide guidance to the policy makers with effective vehicle emission controls. This review paper reviewed two kinds of typical vehicle emission inventories. The top-down vehicle emission inventories is calculated based on the static datasets (e.g., vehicle population, vehicle kilometer traveled, and fuel consumption). These inventories could track historical emissions abatement progress and examine potential benefits from future regulations. The technological evolution in intelligent transportation systems have facilitated emission inventories to satisfy the increasing sophisticated management demand. The bottom-up link-level vehicle emission inventories are development based on the availability of the real-world traffic profiles. To simulate the temporal and spatial patterns with high-resolution, traffic demand model and machine learning methods are employed to elucidate traffic emissions.

1 Introduction

China has been embracing rapid motorization over the past two decades with a 14.1% annual average growth rate of vehicle population as result of the rapid social and economic growth^[1]. Vehicle ownership density is a widely used indicator to represent the motorization. The ownership of light-duty passenger vehicles (LDPVs) per thousand people increased from 6 in 2000 to 108 in 2016, an 18-fold increase over 16 years. Over the same time, the increased rates of ownership of LDPVs per thousand people in the European Union and the United States are 23%^[2] and 2%^[3], respectively (See Fig. 1).

*Corresponding author: 674057237@qq.com

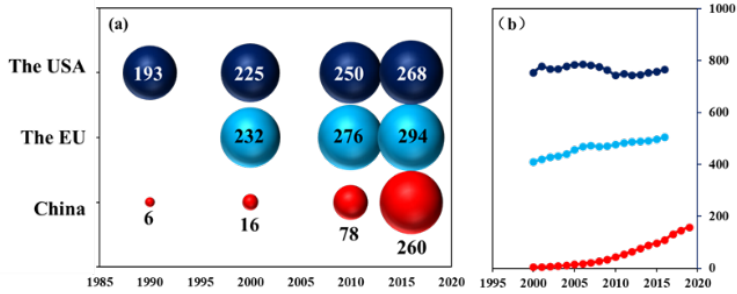


Fig. 1. The ownership of LDPVs per kilo people in China, the EU, and the USA, 2000-2016: (a) vehicle population, units in million, and (b) ownership of LDPVs per thousand people, units in veh per thousand people.

Note: The data markers in panel (a) represent the vehicle population in 1990, 2000, 2010, and 2016 for the USA; 2000, 2010, and 2016 for the EU; and 1990, 2000, 2010 and 2019 for China.

The rapid motorization poses substantial challenges for China concerning assuring energy security and alleviating global climate change^[4]. The petroleum consumption increased from 440 million tons in 2010 to 660 million tons in 2019, and the external dependence increased from 54% to 71% during the same period. On-road vehicles are a major driver for the surge of fuel demand accounting for approximately 48% of petroleum consumption. The Chinese government has announced its goals to peak CO₂ emissions before 2030 and achieve carbon neutrality by 2060. the Chinese government should make more efforts to focus on energy conservation and the CO₂ emission reductions in the transport sector. It is necessary for China within the time framework to make substantial progress to improve transportation system efficiency, lower energy consumption for fossil fuel powered vehicles and substantially increase the usage of low carbon on-road fuels

The dramatic increase in vehicle population has also triggered serious air pollution. Vehicle emissions has become the largest contribution to city air pollution (e.g. CO, NO_x, and PM_{2.5}). Many megacities in China are facing with similar environmental challenges. Beijing's annual concentration of PM_{2.5} in 2020 was 38 µg m⁻³. Although this value was reduced by 42% as opposed to that in 2013, it still exceeded the limit of China's national ambient air quality standard (35 µg m⁻³). The recent official source apportionment results of PM_{2.5} indicated that vehicle emissions were the largest local contributors to ambient for PM_{2.5} concentrations in five cities (Beijing, Shanghai, Guangzhou, Shenzhen, and Hangzhou) megacities in China^[5] (See Fig. 2).

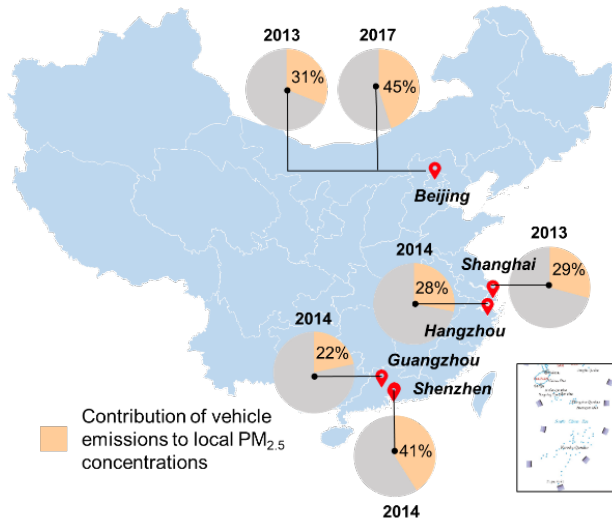


Fig. 2. The official source apportionment results of PM_{2.5} pollutions for megacities in China.

The air pollutants from on-road vehicles have adverse health impacts^[6]. The emission from heavy duty trucks have been confirmed human carcinogen. The European Environment Agency (EEA) reported that 94% of the monitoring sites whose NO₂ concentration are exceeded the limits are traffic-related sites. The EEA are further estimated that around 70,000 premature deaths are due to the NO₂ exposure. Jerrett et al.^[7] found out significant positive correlation between the traffic-related NO₂ concentrations mortality in California.

Vehicle emission inventory is an irreplaceable tool to characterized the spatiotemporal distribution of the air pollutant and provide guidance to the policy makers with effective vehicle emission controls. This review paper aims to summarize the basic methodology and necessary data to development the two kinds of typical vehicle emission inventories (i.e. top-down and the bottom-up vehicle emission inventories). Based on the review, this paper further has a discussion on the comparison of the two kinds of inventories and the future trends of the development of vehicle emission inventories.

2 Research on vehicle emission inventory

2.1 Top-down vehicle emission inventory

The top-down vehicle emission inventory is calculated based on registered population by vehicle category, as Eq. 1 illustrates.

$$E_{p,y} = \sum_{vc} \sum_f \sum_{va} (VP_{vc,f,y,va} \times VKT_{vc,f,y,va} \times EF_{vc,f,y,va,p} \times 10^{-6}) \quad (1)$$

where $E_{p,y}$ is the total emissions of pollutant p during the certain period y , units in t; $VP_{vc,f,y,va}$ is the vehicle population defined by vehicle category vc , fuel type f , and vehicle age va in the certain time period y , units in veh; $VKT_{vc,f,y,va}$ is the fleet-average vehicle kilometers traveled of vehicle category vc , fuel type f , and vehicle age va in the certain time period y , units in km veh⁻¹; $EF_{vc,f,y,va,p}$ is the fleet-average emission factors of pollutant p of vehicle category vc , fuel type f , and vehicle age va in the certain time period y , units in g km⁻¹;

As the Eq. 1 illustrates, the top-down emission inventory is calculated based on annual statistics data, which is commonly used to characterize the total emission trends in a large

area and a long period. There are relatively more sophisticated emission inventories overseas. The Long-range Transboundary Air Pollution (LRTAP) released by the EEA is developed based on the emission inventories reported by its member states. The on-road emission inventories of LRTAP are mainly calculated based on the fuel consumption and vehicle registration data. The National Emission Inventory (NEI) reported by the US Environment Protection Agency (EPA) employs the similar approach to established the national and county inventories.

Since the methodology to establish the emission inventories is relatively simple and the vehicle population and the fleet-average VKT are commonly from the statistical data, the researchers who established the top-down vehicle emission inventory often attempt to track historical emissions abatement progress and examine potential benefits from future regulations. Wu et al.^[8] established the top-down vehicle emission inventory in China to illustrate vehicle emission trends during 1998-2013 and the spatial patterns of the emissions at the provincial resolution. Based on the province-level emission inventory, China's first fifteen-year efforts in controlling vehicles emissions was assessed. Wu et al.^[9] further evaluated the vehicle emission reductions by designing comprehensive control scenarios based on the Wu's datasets and methodology, and provided detailed policy roadmaps and technical options related to these future emission reductions for governmental stakeholders. Both these researches employed a localized emission model named Emission Factor Model for the Beijing Vehicle Fleet Version (EMBEV 2.0). The EMBEV model was developed based on thousands of in-lab dynamometer tests and hundreds of on-road tests. Now, the EMBEV methodology and key parameters have been essentially referred to by China's National Emission Inventory Guidebook. Zheng et al.^[10] built a GHG emission inventory to predict vehicular GHG emissions on provincial basis based on the statistic fuel consumption, and proposed an integrated policy to peak GHG emissions of 90% provinces and whole China by 2030.

Although previous top-down emission inventory could convincingly support the policy makers with vehicle emission mitigation strategies, some major limitations have not been addressed. The top-down emission inventory is developed based on the registered data lacking temporal and spatial associations with real-world traffic patterns. Recently, the growing awareness of urban sustainability across the world has increased the needs for dynamically managing the road transportation systems within the cities and communities, which has spurred a research focus on developing bottom-up high-resolution road transportation emission inventories, including emission inventories evaluated at street levels. We will have a detail discussion on this kind of inventories in the next section.

2.2 Bottom-up vehicle emission inventory

There are three main motivations for preparing high-resolution road emission inventories. First, sustainability challenges are more significant in populous urban areas and traffic hotspots where vehicle usage is more extensive than rural or remote locations. Second, bottom-up emission inventories are useful to address local land use and transportation planning policies, because they are more representative of actual vehicle usage than conventional approaches based on macro-scale and static profiles of vehicle registration or fuel consumption. Third, traffic management systems have been adopted by many municipal governments (e.g., congestion fee and low emission zone programs in London, UK; license control and traffic restrictions in Beijing, China) and these systems require fine-grained tools to assess policy efficacy.

It should be noted that the bottom-up vehicle emission inventory referred in this paper is developed based on the real-world traffic datasets with the high temporal resolution of hour and the high spatial resolution of road segment level (~500 m in the urban core area). For

each road link, the hourly emission is the product of traffic volume, link length and speed-dependent factors (See Eq. 2)

$$E_{h,j,l} = \sum_c EF_{c,j}(v) \times TV_{c,h,l} \times L_l \tag{2}$$

where $E_{h,j,l}$ is the total emission of pollutant j on road link l at hour h , units in $g\ h^{-1}$; $EF_{c,j}(v)$ is the average emission factor of pollutant j for vehicle category c at speed v , units in $g\ km^{-1}$; $TV_{c,h,l}$ is the traffic volume of vehicle category c on road link l at hour h , units in $veh\ h^{-1}$; L_l is the length of road link l , units in km.

However, limited by the sparseness of the real-world traffic profiles, many current researches could not reach the high temporal and spatial resolution as mentioned above. Some bottom-up inventories established using the artificial allocation methods based real-world traffic profiles or spatial surrogates (e.g., population density, road length density) that could be the input of the air quality simulation models are included in this section.

Due to the technological difficulties in data mining, traffic data availability is a significant challenge in characterizing real-world spatial and temporal distributions of vehicle emissions.^[11] As high-resolution emissions are essentially required by air quality simulations, other accessible spatial surrogates are used to artificially allocate total vehicle emissions into fine spatial cells. Population density and/or road length density are two typical varieties of spatial indicators to allocate vehicle emissions by assuming linear relationships between vehicle emissions and spatial surrogates.

Zheng et al.^[12] introduced allocation weights according to the road type (e.g., highways, national roads, provincial roads, and county roads) to assign the county-level emissions onto $0.05^\circ \times 0.05^\circ$ grids based on the China Digital Road-network Map (CDRM). Zheng et al.^[13] proposed an allocation method to develop the vehicle emission inventory in the Pearl River delta region based on the standard road length instead of actual road length which uses road length density as the spatial indicators to allocate the vehicle emissions by further considering the influence from the road type and the distinctions between urban and rural areas. Further comparison was made between this method and the traditional method assuming the liner relationship between the emission intensity and the population density by comparing the modeled pollutant concentrations (e.g., ozone) with the observed ones. The air quality simulation results showed that Zheng’s method could improve the accuracy of model predictions for fine-resolution modeling application. However, such top-down allocations are often questioned with respect to the accurate representation of real-world traffic activity.

In the US, annual averaged daily traffic (AADT) data are reported annually by the Highway Administration, which are used to establish high-resolution vehicle emission inventories from city to national scales. Gately et al.^[14, 15] reported high-resolution road CO₂ emission inventories for Massachusetts and the US. The results indicate that the top-down approach based on macro-scale parameters may lead to deviations in the central districts that are as large as 500%. Open-access traffic count data, such as the AADT data, usually provide annual averaged characteristics and lack finer temporal resolution.

However, first, the AADT data use the traffic profiles of some road links that are reported in “Sample Panel”, a select portion of a given roadway system, to represent “Full Extent” of the systems. Second, some empirical assumptions and adjustments of VKT are often used to downscale state-level or national-level AADT profiles to county-level traffic patterns. Both of these factors could result in the estimated spatial variations in traffic volumes may not represent real-world patterns. Furthermore, the AADT datasets are updated per year according to the submission from all states annually, the collection and process of the traffic data prevent the establishing the real-time inventory to reflect the temporal patterns. Therefore, the AADT datasets could support the analysis of seasonal or day-of-week variations, but are limited to calculate road emissions in hourly or even finer temporal

resolutions. To improve the temporal resolution, McDonald et al.^[16] used 70 weigh-in-motion detectors in California to resolve diurnal and weekly variations for road CO₂ emissions, but still could only distinguish gasoline and diesel fleets in a coarsely aggregated way.

The technological evolution in intelligent transportation systems have facilitated emission inventories. Gately et al.^[17] applied GPS-informed speed data from mobile phones and vehicles to map emission fluxes from vehicles in Boston. In addition to such trajectory data, open-accessed traffic congestion indexes could be also derived from navigation companies or municipal government agencies to dynamically estimate road speeds. For traffic volume and mix, radio-frequency identification (RFID) and traffic camera are capably of reporting detailed vehicle counts by license plate number⁸. These real-world traffic datasets be useful to elucidate temporal and spatial variations in traffic emissions.

To capture the traffic volume of each road segments and avoid the uncertainty due to simple empirical assumptions, transportation demand modeling has been utilized to assist the development of emission inventories. Transportation demand modeling includes the traffic equilibrium model and the traffic speed-flow model. Xie et al.^[18] developed an integrated the microscopic traffic simulation model (i.e., PARAMICS) and the traffic emission model (i.e., MOVES from US EPA) to calculated the link-level emissions on a well-calibrated road network in Greenville, South Carolina to evaluate the fuel consumption impacts. Abou-Senna et al.^[19] presented an approach to capture the environmental impacts of vehicular operations on a 10-mile stretch of Interstate 4 (I-4), an urban limited-access highway in Orlando, Florida by using the traffic demand model (i.e., VISSIM) to analyze the average speed and traffic volumes of the highway. Zhou et al.^[20] allocated the vehicle emission into the grids in the urban area of Beijing by a transportation simulation platform based on a transportation demand model (i.e., TransCAD), and evaluated the emission reduction from the vehicle emission controls during the 2008 Olympics. The results suggested that reasonable traffic system improvement strategies along with vehicle technology improvements can contribute to controlling total vehicle emissions. Jing et al.^[21] presented a bottom-up methodology based on the near-real-time traffic data on road segments collected by the manual camera. The localized speed-flow model was employed to resolve the relationship between traffic volume and road segment speed, and further to develop a vehicle emission inventory with high temporal–spatial resolution for the Beijing urban area on a typical weekday. Yang et al.^[22] expanded the research domain the improved the temporal-spatial resolution compared with Jing’s research, employed the localized speed-flow model and resolved the congestion maps to established large-scale, real-world traffic datasets of the entire municipal area of Beijing and the emission contribution from the nonlocal freight was first estimated.

Such transportation simulation methods are often time-consuming, which has motivated us to explore more efficient data-driven method to map traffic flows in an entire network. That is to utilize machine learning methods to analyze the spatial distribution of traffic characteristics (e.g., volume, speed) by relating to some physical land-use features (e.g., population, infrastructure). Compared with traditional parametric methods (e.g., linear regression), machine learning methods are more attractive tools to perform supervised learning tasks on complex datasets by avoiding a prior rigid assumption about the nature form of the model. Random forests (RF) are often implemented in prediction analyses because of their increased accuracy and resistance to multi-collinearity and complex interaction problems as compared to linear regression^[23, 24]. Random forest models have been widely used in predicting the spatial distribution of pollutant concentrations, showing better accuracy than traditional land-use regression models.

The data-driven methods are primarily employed to predict the temporal and spatial distribution of the air pollutant concentrations by establishing the relationship between observed concentrations and land-use and economic variables within the research domain. Hoek et al.^[25] reviewed 25 land-use regression studies and identified population density, land

use, physical geography and climate as significant predictor variables Brokamp et al.^[26, 27] compared the capability of land use models based on regression (LUR) and random forest (LURF) in predicting the PM_{2.5} concentration and its elements, and the results showed that LURF models were more accurate and precise than LUR models for most elements and could be used for more accurate exposure assessment. The LURF model performed well with an overall cross-validated R² of 0.9 in the seven county area surrounding the Cincinnati, OH, area and could facilitate high-resolution assessment of both long-term and acute PM_{2.5} exposures in order to quantify their associations with related health outcomes.

Artificial neural networks (ANN) could be suitable for simulating traffic profiles when the relationship between input and output are not clear by considering the implicit layers between them. Fu and Rilett^[28] presented an ANN based method for estimating route travel times between individual locations in an urban traffic network during different time periods of the day peak and off-peak. The computational results showed that the ANN-based route travel time estimation model is appropriate, with respect to accuracy and speed, for use in real applications. Ghanim and Abu-Lebdeh^[29] developed a real-time traffic signal control integrating traffic signal timing optimization using ANN and genetic algorithms (GA) modeling. The simulation results showed that the proposed control system can reduce transit vehicle delay and improve schedule adherence. The reductions in delay and schedule adherence are statistically significant.

Support vector machine is a pattern classifier constructed from a unique learning algorithm that extracts training vectors that lie closest to the class boundary. The learning algorithm uses these vectors to construct a decision boundary that optimally separates the data and widely used in the prediction of the traffic accidents. Yuan et al.^[30] trained the simulated incident data from an arterial network in California and showed that SVMs offers a lower misclassification rate, higher correct detection rate, lower false alarm rate and slightly faster detection time in arterial incident detection. Sun et al.^[31] collected the crash data and the corresponding traffic flow detector data on expressways in Shanghai and employed an SVM model to predict the likelihood of crashes based on the important and significant variables from the traffic flow 5-10 minutes before the crash occurred using RF. The results showed that the crash prediction model can obtain a satisfactory prediction performance for crashes with the accuracy of the crash prediction model can be as high as 78.0%.

Gaussian Processes (GPs)^[32] are another machine learning method that could offer flexibility in finding a suitable parametric form for a complex dataset without prior experience. Previous research showed that GPs are easier to use in comparison to alternatives like neural networks and can offer some practical advantages over SVMs. Xie et al.^[33] evaluated the GPs and SVMs performance on short-term traffic flow forecasting based on different sets of traffic volume data collected from highways in Seattle, Washington. The comparative results showed that because the GPs is formulated in a full Bayesian framework, it could allow for explicit probabilistic interpretation of forecasting outputs and give the GPs an advantage over SVMs to model and forecast traffic flow. Liu et al.^[34] proposed dynamic congestion model based on GPs that can effectively characterize both the dynamics and the uncertainty of congestion conditions to minimize the collective travel time of all vehicles in the system. The model are validated in two Asian cities and showed the routing algorithm could generate significant faster routes and achieve near-optimal performance. GPs are flexible non-parametric Bayesian models that have been successfully applied to model and predict with state-of-the-art results various traffic related phenomena such as traffic congestion and traffic volumes.

The data-driven approach is both scientific and adaptable to process real-time traffic big data, making it an ideal tool for characterize temporal the and spatial distribution of traffic flows.

3 Conclusion and perspectives

Over the past two decades, China has experienced rapid growth in vehicle population. Emission inventories are an important tool for environmental and climate managements. We review the basic methodology and necessary data to development the typical vehicle emission inventories and their application. For top-down emission inventories based on static data, they do not account for actual traffic activity, notably inter-city transportation, may have a significant bias. In some megacities, the mismatch can be more significant due to special municipal traffic managements, and would further underestimate traffic contribution to ambient pollutant concentrations. For example, city dwellers in Shanghai may choose drive a car issued with a license plate from nearby provinces (e.g., Jiangsu, Zhejiang) to save the cost of the license plate auction in Shanghai. In Beijing^[9], freight companies may use trucks registered in other provinces as result of more stringent emission regulations and urban restriction policy for heavy-duty trucks in Beijing, and non-local trucks can emit more black carbon than their local counterparts by a multiple fold.

On the other hand, temporal and spatial patterns of air pollutant emissions from on-road vehicles are of increasingly substantial interest because of the associated potential public health impacts.^[35] Intake fractions of vehicular pollutants are greatest in areas with high vehicle usage and population density.^[36] High-resolution vehicle emission inventories for air pollutants will be valuable to evaluate potential health benefits from vehicle emission control measures in particular for traffic spots. Local emission measurement data for air pollutants are needed to estimate and validate emission factors.^[22, 37, 38] For example, Carslaw et al.^[39] reported substantial discrepancies between average NO_x emission factors estimated from a nationwide model in UK and from local remote sensing data in London. Spatial and temporal heterogeneities of road emission inventory estimates for air pollutants will be greater than those for CO₂.

We suggest collaborative and continuous efforts in China to collect traffic data alongside the development of ITS facilities to improve high-resolution emission inventory technologies (e.g., traffic demand models, data-driven methods). Different elements of ITS infrastructure are owned and operated by different stakeholders. For example, local floating car data (e.g., probe taxis) are collected by local municipal governments, while inter-city highway traffic is monitored by Ministry of Transport. Developing a high-resolution road emission inventory requires collaboration between different ITS facility operators. A long-term effort to develop high-resolution road emission inventories would be useful to understand long-term ambient pollutant concentration trends and the drivers (economic development, land use change, urbanization) of emission changes^[15]. As illustrated here, the growing availability of data from advanced traffic control systems coupled with a growing ability to process these data using big data techniques will provide Chinese policy makers with increasingly sophisticated high-resolution emission inventories in the future.

This study was supported by the FAW-Volkswagen China Environmental Protection Foundation automobile environmental protection innovation leading plan and Laboratory of Transport Pollution Control and Monitoring Technology.

Reference

1. China Daily. The population of private passenger cars has reached 200 million.(in Chinese) [EB/OL]. http://paper.people.com.cn/rmrb/html/2020-01/08/nw.D110000renmrb_20200108_5-04.htm
2. Bureau of Transportation Statistics U.S. Number of U.S Aircraft, Vehicles, Vessels, and Other Conveyances [M]. U.S. Bureau of Transportation Statistics.

3. European Environment Agency E. Vehicle Ownership and Truck Intensity [M]. European Environment Agency.
4. Uherek E, Halenka T, Borken-Kleefeld J, et al. Transport impacts on atmosphere and climate: Land transport [J]. *Atmospheric Environment*, 2010, 44(37): 4772-816.
5. Chian News. Nine city released the results of source apportionment. (in Chinese) [EB/OL]. <http://www.chinanews.com/gn/2015/04-01/7176182.shtml>
6. Puntoni R, Ceppi M, Reggiardo G, et al. Occupational exposure to carbon black and risk of bladder cancer [J]. *The Lancet*, 2001, 358(9281): 562.
7. Jerrett M, Burnett R T, Beckerman B, et al. Spatial Analysis of Air Pollution and Mortality in California [J]. *American Journal of Respiratory and Critical Care Medicine*, 2013, 188(5): 593-9.
8. Wu X, Wu Y, Zhang S, et al. Assessment of vehicle emission programs in China during 1998-2013: Achievement, challenges and implications [J]. *Environ Pollut*, 2016, 214(556-67).
9. Wu Y, Zhang S, Hao J, et al. On-road vehicle emissions and their control in China: A review and outlook [J]. *Sci Total Environ*, 2017, 574(332-49).
10. Zheng B, Zhang Q, Borken-Kleefeld J, et al. How will greenhouse gas emissions from motor vehicles be constrained in China around 2030? [J]. *Applied Energy*, 2015, 156(230-40).
11. Pan L, Yao E, Yang Y. Impact analysis of traffic-related air pollution based on real-time traffic and basic meteorological information [J]. *Journal of Environmental Management*, 2016, 183(510-20).
12. Zheng B, Huo H, Zhang Q, et al. High-resolution mapping of vehicle emissions in China in 2008 [J]. *Atmospheric Chemistry and Physics*, 2014, 14(18): 9787-805.
13. Zheng J, Che W, Wang X, et al. Road-Network-Based Spatial Allocation of On-Road Mobile Source Emissions in the Pearl River Delta Region, China, and Comparisons with Population-Based Approach [J]. *J Air Waste Manage Assoc*, 2009, 59(12): 1405-16.
14. Gately C K, Hutyra L R, Wing I S, et al. A bottom up approach to on-road CO₂ emissions estimates: improved spatial accuracy and applications for regional planning [J]. *Environ Sci Technol*, 2013, 47(5): 2423-30.
15. Gately C K, Hutyra L R, Sue Wing I. Cities, traffic, and CO₂: A multidecadal assessment of trends, drivers, and scaling relationships [J]. *Proc Natl Acad Sci U S A*, 2015, 112(16): 4999-5004.
16. C. M B, C. M Z, W. M E, et al. High - resolution mapping of motor vehicle carbon dioxide emissions [J]. *Journal of Geophysical Research: Atmospheres*, 2014, 119(9): 5283-98.
17. Gately C K, Hutyra L R, Peterson S, et al. Urban emissions hotspots: Quantifying vehicle congestion and air pollution using mobile phone GPS data [J]. *Environmental Pollution*, 2017, 229(496-504).
18. Xie Y, Chowdhury M, Bhavsar P, et al. An integrated modeling approach for facilitating emission estimations of alternative fueled vehicles [J]. *Transportation Research Part D: Transport and Environment*, 2012, 17(1): 15-20.
19. Abou-Senna H, Radwan E, Westerlund K, et al. Using a traffic simulation model (VISSIM) with an emissions model (MOVES) to predict emissions from vehicles on a limited-access highway [J]. *J Air Waste Manage Assoc*, 2013, 63(7): 819-31.

20. Zhou Y, Wu Y, Yang L, et al. The impact of transportation control measures on emission reductions during the 2008 Olympic Games in Beijing, China [J]. *Atmospheric Environment*, 2010, 44(3): 285-93.
21. Jing B Y, Wu L, Mao H J, et al. Development of a high temporal-spatial resolution vehicle emission inventory based on NRT traffic data and its impact on air pollution in Beijing – Part 1: Development and evaluation of vehicle emission inventory [J]. *Atmospheric Chemistry and Physics Discussions*, 2015, 15(19): 26711-44.
22. Yang D, Zhang S, Niu T, et al. High-resolution mapping of vehicle emissions of atmospheric pollutants based on large-scale, real-world traffic datasets [J]. *Atmos Chem Phys*, 2019, 19(13): 8831-43.
23. James G, Witten D, Hastie T, et al. *An Introduction to Statistical Learning: With Applications in R* [M]. 2013.
24. Liaw A, Wiener M. Classification and Regression by RandomForest [J]. *Forest*, 2001, 23(
25. Hoek G, Beelen R, De Hoogh K, et al. A review of land-use regression models to assess spatial variation of outdoor air pollution [J]. *Atmospheric Environment*, 2008, 42(33): 7561-78.
26. Brokamp C, Jandarov R, Rao M B, et al. Exposure assessment models for elemental components of particulate matter in an urban environment: A comparison of regression and random forest approaches [J]. *Atmospheric Environment*, 2017, 151(1-11).
27. Brokamp C, Jandarov R, Hossain M, et al. Predicting Daily Urban Fine Particulate Matter Concentrations Using a Random Forest Model [J]. *Environmental Science & Technology*, 2018, 52(7): 4173-9.
28. Fu L, Rilett L R. Estimation of time dependent, stochastic route travel times using artificial neural networks. [J]. *Transportation Planning and Technology*, 2000, 24(1): 25-48.
29. Ghanim M, Abulebdeh G. Real-Time Dynamic Transit Signal Priority Optimization for Coordinated Traffic Networks Using Genetic Algorithms and Artificial Neural Networks [J]. *Journal of Intelligent Transportation Systems*, 2015, 19(4): 327-38.
30. Yuan F, Cheu R L. Incident detection using support vector machines [J]. *Transportation Research Part C: Emerging Technologies*, 2003, 11(3): 309-28.
31. Sun J, Sun J. Real-time crash prediction on urban expressways: identification of key variables and a hybrid support vector machine model [J]. *Intelligent Transport Systems*, 2016, 10(5): 331-7.
32. Williams C K I, Rasmussen C E. Gaussian Processes for Machine Learning [J]. 2005, 2(
33. Xie Y, Zhao K, Sun Y, et al. Gaussian Processes for Short-Term Traffic Volume Forecasting [J]. *Transportation Research Record*, 2010, 2165(1): 69-78.
34. Liu S, Yue Y, Krishnan R. Adaptive Collective Routing Using Gaussian Process Dynamic Congestion Models [M]. 2013.
35. Marshall J D, Mckone T E, Deakin E, et al. Inhalation of motor vehicle emissions: effects of urban population and land area [J]. *Atmospheric Environment*, 2005, 39(2): 283-95.
36. Evans J S, Wolff S K, Phonboon K, et al. Exposure efficiency: an idea whose time has come? [J]. *Chemosphere*, 2002, 49(9): 1075-91.
37. Carslaw D C, Beevers S D, Tate J E, et al. Recent evidence concerning higher NO_x emissions from passenger cars and light duty vehicles [J]. *Atmospheric Environment*, 2011, 45(39): 7053-63.

38. Zhang S, Wu Y, Huang R, et al. High-resolution simulation of link-level vehicle emissions and concentrations for air pollutants in a traffic-populated eastern Asian city [J]. *Atmospheric Chemistry and Physics*, 2016, 16(15): 9965-81.
39. Carslaw D C, Beevers S D. Estimations of road vehicle primary NO₂ exhaust emission fractions using monitoring data in London [J]. *Atmospheric Environment*, 2005, 39(1): 167-77.