Atmospheric Environment 222 (2020) 117136

Contents lists available at ScienceDirect

Atmospheric Environment

journal homepage: http://www.elsevier.com/locate/atmosenv

High spatial and temporal resolution vehicular emissions in south-east Brazil with traffic data from real-time GPS and travel demand models

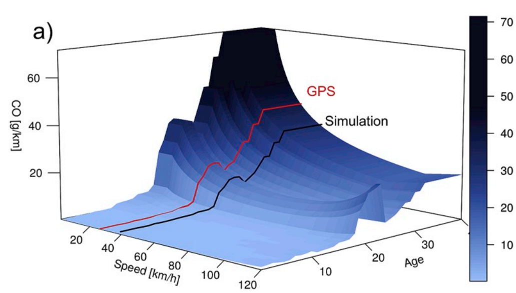
Sergio Ibarra-Espinosa [a,b,*], Rita Yuri Ynoue [a], Karl Ropkins [c], Xuelei Zhang [b], Edmilson Dias de Freitas [a]

[a] *Departamento de Ciências Atmosféricas, Instituto de Astronomia, Geofísica e Ciências Atmosféricas, Universidade de São Paulo, Brazil*
[b] *Key Laboratory of Wetland Ecology and Environment, Northeast Institute of Geography and Agroecology, Chinese Academy of Sciences, Changchun, 130102, China*
[c] *Institute for Transport Studies, University of Leeds, UK*

HIGHLIGHTS

- A comprehensive regional scale vehicular emissions inventory was developed for South-East Brazil.
- The vehicular activity comes from real-time GPS recordings and travel demand models.
- The emission factors are based on Brazilian measurements and were scaled to incorporate speed variation.
- The peak of emissions of CO and HC occurs at 20 years, for NOX at 5 years and PM2.5, at 17 years of vehicle use.

GRAPHICAL ABSTRACT



ARTICLE INFO

Keywords:
VEIN
Vehicular emissions
Travel demand
Brazil
GPS

ABSTRACT

Vehicular emissions are one of the most important source of pollution in urban centers, impacting air quality with a deleterious effect on human health and ecosystems. Air quality managers rely on emissions inventories to characterize pollution and sources. In this study we predicted vehicular emissions, using three sources of traffic data: 1) travel demand model outputs consisting of traffic simulations of light-duty vehicles, trucks, 2) and urban buses, and 3) a massive data set of real-time GPS coordinates of light-duty vehicles and trucks. The study area comprises the metropolitan areas of São Paulo, Santos, Vale de Paraíba, Sorocaba, and Campinas, which have a population of more than 30 million inhabitant. Once we generated hourly traffic flows, we used the Vehicular Emissions INventory Model (VEIN) to predict fuel consumption and emissions. Emissions using travel demand model for the metropolitan area of São Paulo are CO 177406 t/y, NO_X 73554 t/y, NMHC 33999 t/y and $PM_{2.5}$ 2281 t/y. The emissions using GPS data were higher than using travel demand outputs, because GPS average speeds were lower, producing higher emission factors.

1. Introduction

In urban centers, vehicles are the most important source of air pollution (Molina and Molina, 2004). The biggest urban agglomeration in South America is located in Southeast of Brazil, conformed by the Metropolitan Area of São Paulo (MASP) located near to other 4 metropolitan areas: Baixada Santista (hereafter identified as Santos), Vale do Paraíba (in this work it will be identified by the name of its main city São

* Corresponding author. Departamento de Ciências Atmosféricas, Instituto de Astronomia, Geofísica e Ciências Atmosféricas, Universidade de São Paulo, Brazil.
E-mail address: sergio.ibarra@usp.br (S. Ibarra-Espinosa).

https://doi.org/10.1016/j.atmosenv.2019.117136
Received 13 May 2019; Received in revised form 6 November 2019; Accepted 9 November 2019
Available online 15 November 2019

José Dos Campos - SJDC), Sorocaba and Campinas, as shown on Fig. 1. Altogether these 5 metropolitan areas accounted for more than 30 million of inhabitants in 2016 according to the Brazilian Institute of Geography and Statistics (IBGE, 2016). Air quality simulations conducted in these regions confirms that the main source of air pollution is road transportation (Andrade et al., 2015; Vara-Vela et al., 2016; Hoshyaripour et al., 2016). An emissions inventory is a spatial and temporal characterization of the release of the mass of pollutants into the atmosphere which serves as a decision making tool and provides inputs for air quality modeling (CETESB, 2015). Therefore, improving the accuracy of emission inventorying activities would help to improve the assessment of environmental policies, evaluate the emissions impacts on the climate, ecosystem and population exposure.

A key element to develop high temporal and spatial resolution bottom-up vehicular emissions inventories is the availability of traffic data. Some studies use macro (Corvalán et al., 2002) and micro (Panis et al., 2006; Nyhan et al., 2016) traffic simulations to estimate vehicular emissions. Others use traffic counts to allocate or produce traffic flow to later estimate emissions (Kinnee et al., 2004). However, most of studies are often limited to the cruder top-down approaches based on aggregated traffic fleet information (Gómez et al., 2018; Rafee et al., 2017). Obviously, transportation emissions based on more representative traffic flow estimates and emission factors would improve air quality management and modeling.

Recently, vehicular activity has been generated using data from Internet transportation companies. For instance, GPS datasets have been used to improve travel demands models (TDM) (Gately et al., 2017) to estimate emissions. Another approach consisted transform categorical congestion maps speed and traffic flow estimates (Yang et al., 2019). While these approaches are promising they still rely on secondary data as the congestion maps and not the primary source of data. Also, it is not assured that the driving cycle associated with the categorization of the level of congestion informed by the Internet companies is representative elsewhere.

In this study we estimated road transportation emissions using three sources of traffic data in developing bottom-up vehicular emissions inventories: 1) 120 million GPS observations of light-duty Vehicles (LDV) and trucks from the company Maplink (http://transito.maplink.global/), 2) travel demand model simulations of light-duty and trucks from the Company of Traffic and Engineering of São Paulo (CET,

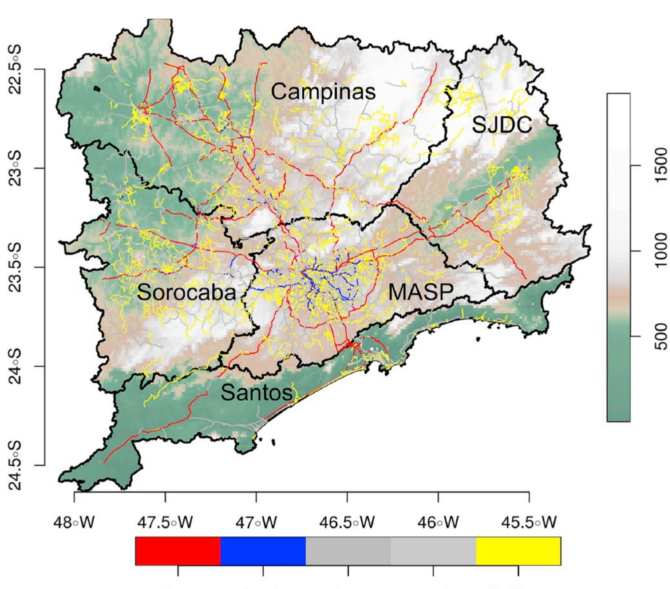


Fig. 1. Area of study: Metropolitan Areas (delimited by black lines), topography (shaded) and road network (colored lines).

http://cetsp.com.br/) and 3) travel demand model simulations for Urban Buses (UB) from the Secretary of Transport and Mobility (Sptrans, http://www.sptrans.com.br/). In contrast to other studies, we make direct use of primary GPS location data in this study and provide further details of the methods to extract traffic flow and speed information in an accompanying paper (Ibarra-Espinosa et al., 2019).

To work with this new type of data we improved the Vehicular Emissions INventory Model (VEIN) (Ibarra-Espinosa et al., 2018), which is an R package available at https://CRAN.R-project.org/package=vein (Ibarra-Espinosa and Schuchvein, 2019). VEIN predicts hot and cold exhaust, evaporative, wear and resuspension emissions, and also provides further emission speciation for chemical mechanisms and spatial geo-processing tasks associated with handling of gridded emissions. We used the emission factors from the official environmental agency of São Paulo (CETESB, 2015), however, as these factors are constant by type of vehicle and year they entered into the market, we generated speed-dependent functions based on these values. Our objective is to estimate bottom-up vehicular emissions using traffic flow from real-time GPS and traffic demand models outputs, to produce emissions maps and compare results.

2. Material and methods

2.1. Study area

The study area is delimited by the coordinates −47.856 to −45.484 degrees of longitude and −24.488 to −22.467 degrees of latitude. The area is located at Southeast Brazil and covers the metropolitan areas of São Paulo (MASP), Santos, Sao José do Campos (SJDC), Sorocaba and Campinas. The area of study is shown in Fig. 1. The political delimitation of each metropolitan area are smaller than the regions presented, with exception of MASP. In other words, we added more municipalities into each metropolitan area because we have data for the whole area, and not only for each geopolitical defined metropolitan area.

2.1.1. Road network

We used transport network data from OpenStreetMap (OpenStreetMap contributor, 2017) to add road layouts to Fig. 1. The OpenStreetMap data included several non-road transport networks (rail, pedestrian paths, etc). We excluded these by downloading the data with the QGIS software (QGIS Development Team, 2017) and filtering only streets with the following attributes: *motorway*, *trunks*, *primary*, *secondary*, *tertiary* and *residential* and the associated '_link' junction streets. The road network of Fig. (1) shows how the metropolitan areas are connected by motorways and trunk roads inside each region.

2.2. Traffic activity data

This study is focused in comparing emissions inventories using three sources of traffic data: GPS position of vehicles that feeds the Internet traffic site Maplink (http://transito.maplink.global/); Traffic simulations from Traffic Engineering Company (CET http://www.cetsp.com.br/) in São Paulo for Light Duty Vehicles (LDV) and Trucks; and Traffic simulation of Urban Buses (UB) from SPtrans (http://www.sptrans.com.br/).

2.2.1. Traffic simulations

The Traffic Engineering Company of São Paulo (CET) is the agency that performs the transport planning in São Paulo city. They have a specific department dedicated to 4-stage traffic simulations using data from an Origin-Destination-Survey (ODS) (Metro, 2017) which started in the decade of 1950 in São Paulo. A classic reference of a 4-stage modeling transport is presented in the work of Ortuzar and Willumsen (2002) (Ortuzar and Willumsen, 2002). The 4 stages of the traffic modeling includes characterization of trip attractions and productions by zone in some regions, distribution of these trips, preferred mode of transport for traveling and finally the allocation of the trips at each mode, in our case into the road network. The survey is made by Metro (http://www.metro.sp.gov.br/), which is the underground company in São Paulo, every 10 years with a smaller update after 5 years. The information gathered in the ODS is massive with the participation of thousand of commuters, helping to identify characteristics of the trips inside MASP. CET uses the information from the survey and performs the travel demand modelling. In this case, it is a macro or strategic travel demand simulation which represents the equilibrium between offer and supply of transportation at maximum load of the road network, that is, at the morning rush hour, which is from 08 to 09 local time.

The CET travel demand simulation provide initial and final identification number of node, volume of light-duty vehicles veh$\cdot$ h^{-1}, trucks, veh$\cdot$h^{-1}, number of lanes, capacity veh$\cdot$h^{-1}, free flow speed km$\cdot$ h^{-1}, peak speed km$\cdot$h^{-1}, type of street and travel-time for each link.

The Secretary of Transport and Mobility of São Paulo (SPTrans) is the agency responsible for public transportation by buses, and also uses origin-destination survey (Metro, 2017) to make traffic simulation for the rush hour.

The traffic simulations for light-duty vehicles, trucks and buses are shown on Fig. 2. Light-duty vehicles activity is more concentrated in urban motorways located near to the center of the city. There are streets where Light-duty vehicles traffic flows exceed more than 15,000 $veh\cdot h^{-1}$. Trucks are concentrated on motorways around the city where flows can exceed 7000 $veh\cdot h^{-1}$. As with light-duty vehicles, Bus activity is concentrated near the center of the city, but also has an east to southwest gradient. Bus flow can exceed 5000 per street $veh\cdot h^{-1}$. As the travel demand model only covers morning peak hour, we used expansion factors to extrapolate to the 168 h a week (Corvalán et al., 2002; Ibarra-Espinosa et al., 2018). The expansion factors were generated using hourly traffic counts from toll stations administrated by the Transportation Agency (ARTESP http://www.artesp.sp.gov.br) located near the city.

We obtained 124 million GPS observations of Cars and Trucks for the study area for the period between 2014 and 10-06 00:00 and 2014-10-12 23:00. This data set of 145 h was filled to sum 168 h of a week of data. We generated traffic flows for light-duty vehicles and trucks by associating average speed and data-base of traffic counts and speed recordings by type of street. The process used to extract traffic flow information from these GPS records is described in detail in an associated MethodsX paper (Ibarra-Espinosa et al., 2019).

Table 1 shows the a summary of the traffic data developed for this study. The GPS data includes passenger cars (PC) and Trucks. The second source of data is travel demand model simulation of light duty vehicles (LDV) and Trucks and the third is also TDM of buses. Therefore,

Table 1
Sources of traffic data used in this study.

Data	Method A	Method B
1) GPS: PC + Trucks	PC (GPS)	PC (TDM)
2) TDM: LDV + Trucks	Trucks (GPS)	Trucks (TDM)
3) TDM: Bus	Bus (TDM)	Bus (TDM)

TDM is travel demand model.

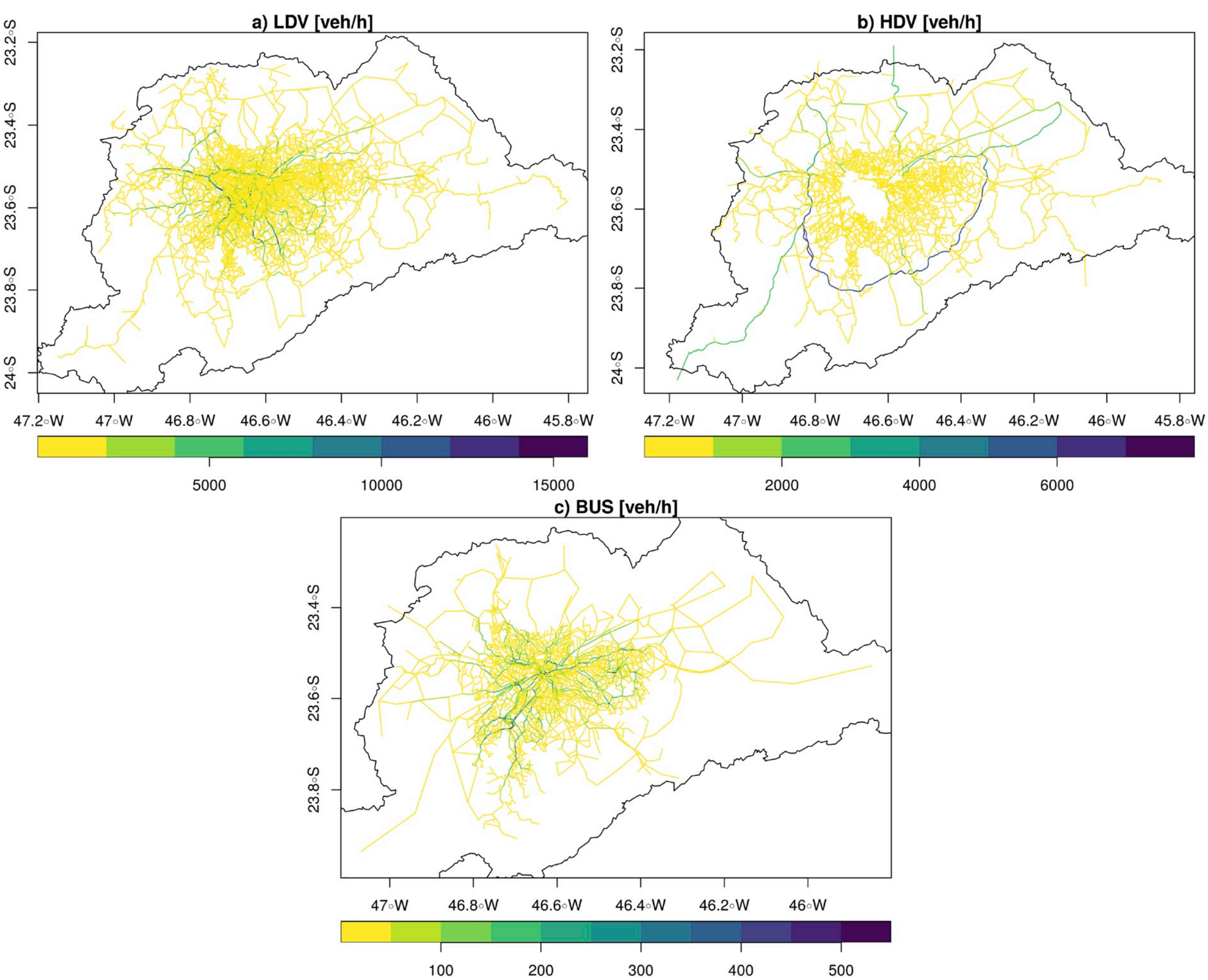


Fig. 2. CET Traffic simulation of (a) Light-duty vehicles and (b) Trucks, and (c) SPTrans simulation of buses from 08 to 09LT.

we aggregated the traffic data using two methods, the first method was named **A** which is formed by GPS recordings from light and trucks vehicles and TDM from buses. The second method was named **B** which is formed by TDM outputs from light vehicles, trucks and buses.

2.2.2. Vehicular composition

We considered 5 general types of vehicles: passenger cars (PC), light-commercial vehicles (LCV), trucks, buses and motorcycles (MC). The vehicular composition divides these categories by engine size and gross weight (Corvalán et al., 2002). As the input traffic data does not include specific type of vehicles, we used the vehicular composition from CETESB for São Paulo (CETESB, 2015). Cars from GPS recordings and light-duty vehicles from traffic simulation were split into 73.70% PC, 13.90% light-commercial vehicles and 12.40% motorcycles, the other percentages are shown on Table (2). In this way, we have a vehicular composition which is similar to the CETESB, however, we further identified the engine size of passenger cars using data from the using data of the National Association of Cars Manufacturers of Brazil (ANFAVEA (ANFAVEA, 2014),).

In Brazil there are mostly three types of automotive motors, 4-strokes, flex 4-strokes and diesel-compression. Standard 4-strokes engines uses gasohol, which is gasoline with 25% of ethanol (E25). Flex engines can use gasoline with any mixture of ethanol between 25% and 100%, so drivers are able to choose gasoline blends between E25 and E100. Compression engines use diesel with 5% of biodiesel (B5). We consider fleet statistics of engine size, weight and type of fuel used from the report of vehicular emissions inventory for São Paulo (CETESB, 2015). Passenger cars were divided by size of the engine with sales statistics of the National Association of Car Manufacturers (ANFAVEA, 2014) as: 38.14% $<=$ 1400 cc, 58.12% 1400 $< cc <=$ 2000 and 3.74% $>$ 2000 cc. Motorcycles were divided with sales statistics of the National Association of Motorcycles Manufacturers (ABRACICLO, 2014) as: 84.24% $<=$ 150 cc, 12.32% 150 $< cc <=$ 500 and 3.44% $>$ 500 cc. Each category is then distributed by age of vehicle, using the data from official inventory for São Paulo (CETESB, 2015). The categories Bus SP and ABus (Articulated Bus), shown on Table 2, are urban public transportation. According to SPtrans (http://www.sptrans.com.br/indicadores/) the average age of UB was 5.2 in 2014, therefore, we adjusted the max age of this category to 10 years of use so that the average age approximately 5.

2.3. Calibration of traffic

Vehicular emissions inventories must be validated. European Commission for emissions inventorying guidelines recommend that fuel consumption estimates be normalized by fuel sales in a region (Ntziachristos and Samaras). In the case of bottom-up approaches, the inventory is calibrated by factoring the traffic flow because the length of the road is constant. The fuel sales information comes from the Yearbook of Energy and Statistics for São Paulo State 2014 (de Energia Electrica, 2014) covering each municipality in São Paulo State. The total amount of fuel sold in 2014 for automotive use and for each metropolitan region is shown on Fig. 3. Most of fuel sales are in MASP because it is the bigger area with most inhabitants and number of vehicles with $4.65{\cdot}10^9$ litres of gasohol, $3.02{\cdot}10^9$ litres of ethanol and $3.14{\cdot}10^9$ litres of diesel. In MASP, the consumption of gasohol is clearly higher than the consumption of gasohol in other regions, while the consumption of the different type of fuels is more similar in the other regions. This could result in a higher emissions ratio of hydrocarbons versus nitrogen oxides in MASP, with implications in tropospheric ozone formation. Specifically, some studies have shown that the increase of gasohol over ethanol locally reduces O_3, while increasing NO_X and CO (Salvo and Geiger, 2014) and reducing hydrocarbon emissions, would reduce ozone formation (Alvim et al., 2017; Andrade et al., 1016).

The fuel sold for automotive use is not always used for that purpose. For example, a small fraction of electric generators use diesel or (less commonly) gasoline purchased from transport sector suppliers. Therefore, we expect that a fraction of the fuel sales were used for other uses. In order to account for such usage, we applied the assumption that 20% of B5 was used for other purposes than automotive. This means that the fuel consumption estimated from the traffic would be 80% B5 and 100% E25 and E100. There is no available information about diesel sold for electricity generation but we believe this value is reasonably conservative.

2.4. Emissions modeling

2.4.1. Aggregation of traffic flow

Hourly traffic flow information was extracted for light-duty vehicles and trucks from GPS records with a spatial coverage of all the metropolitan areas for this study. The traffic simulations from CET and SPtrans includes one morning rush traffic data for a typical working day of 2014, including volume of light-duty vehicles, trucks and buses for MASP. With this information we configured the traffic flows that cover all regions with two methods:

Method A: Traffic from GPS records in all regions. MASP with travel demand simulation of urban buses from SPtrans.

Table 2
Vehicular composition used in this study.

Veh	Fuel	Size	%	Veh	Fuel	Size	%
PC	E25	<1400 cc	14.21	MC	E25	<150 cc	74.04
		1400–2000 cc	21.65			150–500 cc	10.83
		>2000 cc	1.39			>500 cc	3.02
		<1400 cc	8.5			<150 cc	3.77
	FE25	1400–2000 cc	12.93		FE25	150–500 cc	0.55
		>2000 cc	0.83			>500 cc	0.15
		<1400 cc	14.48			<150 cc	6.42
	FE100	1400–2000 cc	22.07		FE100	150–500 cc	0.94
		>2000 cc	1.42			>500 cc	0.28
	E100	<1400 cc	0.93	Trucks	B5	3.8–6 t	8.38
		1400–2000 cc	1.42			6-10 t	25.5
		>2000 cc	0.09			10-15 t	15.28
LCV	E25	<3.8 t	39.13			15-40 t*	24.98
	FE25		15.21			15-40 t**	25.85
	FE100		25.9	Bus	B5	<15 t	9.07
	E100		1.18	Coach		>18 t	13.5
	B5		18.56	Bus SP		<15 t	17.23
				Bus SP		15-18 t	28.81
				ABus SP		15-18 t	14.07

Note: * Rigid Trucks <15 t and Articulated Trucks <40 t ** Rigid Trucks <15 t and Articulated Trucks >40 t.

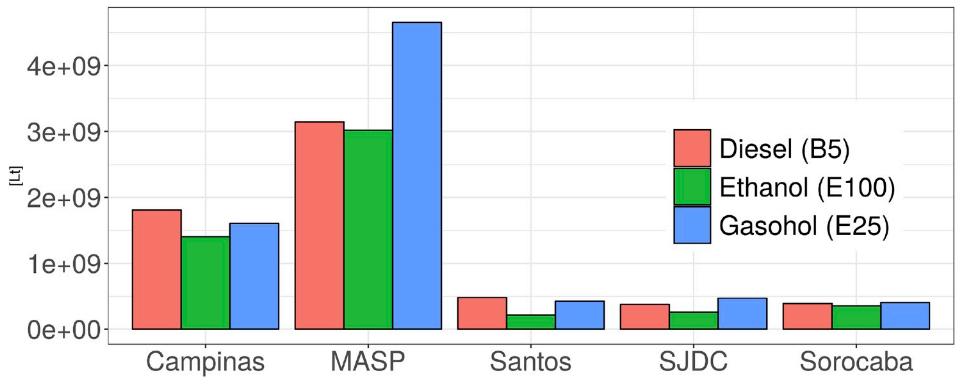


Fig. 3. Fuel sales for 2014 for metropolitan areas Campinas, MASP, Santos, SJDC and Sorocaba.

Method B: Traffic from GPS records except MASP. MASP with traffic simulation from CET and SPtrans of light-duty vehicles, trucks and buses.

We included travel demand simulation of urban buses in method A, because GPS data only considers light-duty vehicles and trucks, The proportion between buses from SPTrans and trucks from CET is 0.27. Therefore, including the bus flow from SPTrans into GPS data set, requires keeping the same proportion. This was accomplished by factoring the buses inside the GPS data-set by 0.17, which was obtained by multiplying 0.27 times the proportion of GPS trucks and buses. The method B uses GPS records for all regions with exception of MASP where travel demand simulations are used instead.

2.4.2. VEIN model

The Vehicular Emissions INventory (VEIN) model (Ibarra-Espinosa et al., 2018) is an R package that estimates bottom-up or top-down vehicular emissions. VEIN estimates hot and cold exhaust, evaporative, wear and re-suspension emissions. VEIN can use either morning rush hour data or hourly data for different hours as an input. Here, we used data from all hours when working with traffic information extracted from GPS records and morning rush hour information when working with traffic simulations. In this latter case where other hours were not output by the model, VEIN provides functions to extrapolate morning traffic and speeds to other hours. This data was divided according to estimated vehicular composition (type, engine size, weight and fuel) without considering the emission standards. So, for example, passenger cars using gasohol with engine lesser than 1400 cc is one category of vehicle composition. Then, age distribution was applied to each vehicular composition category. This allows to have the vehicular composition by age, for example: passenger cars using gasohol with engine lesser than 1400 cc with one year of use. The age distribution was based on statistics on vehicles in circulation from sources including licensing records (CETESB, 2015).

Emission Standards can then be assigned to these subcategories on the basis of age of vehicle. Emission factors available in VEIN are from: the official emissions inventory for São Paulo State developed by the Environmental Company (CETESB, 2015), the European emission guidelines (Ntziachristos and Samaras), and the International Vehicular Emissions model (Davis et al., 2005). Besides, the model has functions to transform constant emission factors into speed functions. Finally, VEIN also imports functions from the model eixport (Ibarra-Espinosa and Schuch, 2017) which exports emissions information in several different useful formats.

2.4.3. Emissions factors

VEIN model allows the user to choose between local emission factors, speed functions from Ntziachristos and Samaras (2016) (Ntziachristos and Samaras) or scaled factors to incorporate speed variation into local factors. In the case of São Paulo, the vehicular emissions inventory reports the emission factors of new vehicles that entered into the market each year (CETESB, 2015). These factors are the average emission certification values for: the Federal Test Procedure (FTP-75) for 4-strokes light-duty vehicles; the World Motorcycle Test Cycle for motorcycles; and the European Stationary Cycle for heavy-duty vehicles. The Federal Test Procedure consists in three phases: cold start, stabilized and hot start. It has a duration of 1877 s and a distance of 17789.56 m with an average speed of 34.2 $km{\cdot}h^{-1}$ (Barlow et al. 2009 Boulter). World Motorcycle Test Cycle consists in three parts of 600 s, the first part covers a distance of 4065 m in an average speed of 24.4 $km{\cdot}h^{-1}$, the second part 9111 m and 54.7 $km{\cdot}h^{-1}$, and the third part 15736 m and 94.4 $km{\cdot}h^{-1}$, which gives a weighted mean of 71.55 $km{\cdot}h^{-1}$ (Barlow et al. 2009 Boulter). The European Stationary Cycle is a cycle for testing the engine in a dynanometer at different combinations of engine speed, load and weight with different duration (Dieselnet, 2017).

Brazilian emission factors covers the following pollutants from exhaust and evaporative processes: CO, NO_x, PM, SO_2, HC, $NMHC$, *Aldehyde*, CO_2 and N_2O (CETESB, 2015). We assigned the vehicular composition with the categories of the European Emissions Guidelines (Ntziachristos and Samaras), including equivalence of emission standards (Ibarra-Espinosa et al., 2018). This allowed us to create an scaled emission factor between the local source of Brazilian emission factors and the speed dependent functions from Ntziachristos and Samaras (2016) (Ntziachristos and Samaras). As a result, we incorporated speed variation into the local emission factors. The scaling process consists in identify speed functions so that, at the speed of the driving cycles, it will return exactly the same value of the local emission factors (CETESB, 2015). European Stationary Cycle driving cycle does not have driven speed and instead it has motor speed, so we are assigning the speed of Federal Test Procedure-75, which is 34.2 $km{\cdot}h^{-1}$. In the case of World Motorcycle Test Cycle for motorcycles, the speed is 54.7 $km{\cdot}h^{-1}$ according to the reference report on driving cycles from the Transport and Research Laboratory (TRL) (Barlow et al. 2009 Boulter). This adjustment is done automatically by the function ef_ldv_scaled and ef_hdv_scaled in VEIN (Ibarra-Espinosa et al., 2018).

The European speed functions emission factors depend on the EURO emission standards (Ntziachristos and Samaras). Therefore, in order to select the right speed function, it is necessary do an equivalence between Brazilian and European Emission standards. Both emission standards are relatively similar but in Brazil they are delayed. Brazilian emission standards are named Proconve for light-duty vehicles and heavy-duty vehicles and Promot for motorcycles (de Medio Ambiente, 2011). Proconve L6 is equivalent to Euro V phase, L3 to Euro II, L2 to Euro I and L1 to Pre-Euro. In the case of heavy-duty vehicles, Proconve P7 is equivalent to Euro V, P5 to Euro III, P4 to Euro II, P3 to Euro I and the rest of the older fleet is equivalent to the Pre-Euro phase. Motorcycles are directly equivalent, Promot M3 to Euro III, M2 to Euro II, MI to Euro I. This equivalence was made based on the site http://www.transportpolicy.ne t/region/south-america/brazil, and already proposed for Brazil (Ibarra-Espinosa et al., 2018). In the absence of local data, we estimated deterioration on basis of accumulated mileage for CO, NO_X and HC emissions using the European emission guidelines method (Ntziachristos and Samaras).

The emission factors are numerous so we selected only some to show in Fig. 4. The highest emission factors of CO are due to older passenger cars at lwo speeds with 71.50 $g{\cdot}km^{-1}$. Newer vehicles emit less due to technological improvements made according to the stringent emission standards. Fig. 4 shows a) *CO*, c) NO_X, d) *HC*, and f) *PM* emission factors $g{\cdot}km^{-1}$ for passenger cars using gasohol and b) *CO*, d) NO_X, f) HC and h) PM emission factors for light trucks as function of speed and age of use. In all cases, when older vehicles are at lower speeds, the emissions factors are higher. The emission factors by age of use have incorporated the deterioration effect due to accumulated mileage. Hence, it is possible to see a peak of emissions around 20 years of use in panels a), c) and e). The emission factors by age of use have a general trend of lower certification emissions values for newer vehicles, despite that occasionally, new vehicles can have values a little higher than those of previous year. Brazilian provides emission factors for light-duty vehicles since 1976 and for heavy-duty vehicles, since 1999, but the for this study we are assuming that the oldest vehicle in circulation is 40 years of use, as shown on Fig. 4. This figure also show the average speed of GPS traffic data of 26.64 $km{\cdot}h^{-1}$ with red lines, and the speed of travel demand simulation of 39.58 $km{\cdot}h^{-1}$ with black lines. The average speeds shown as read and black lines of Fig. 4 covers the whole 168 h. As the average speed of GPS is lower than travel demand simulation, results in higher emission factors in most of cases, with exception of *PM* which does not vary with speed.

Evaporative emissions are due to the vaporization of fuel, driven by

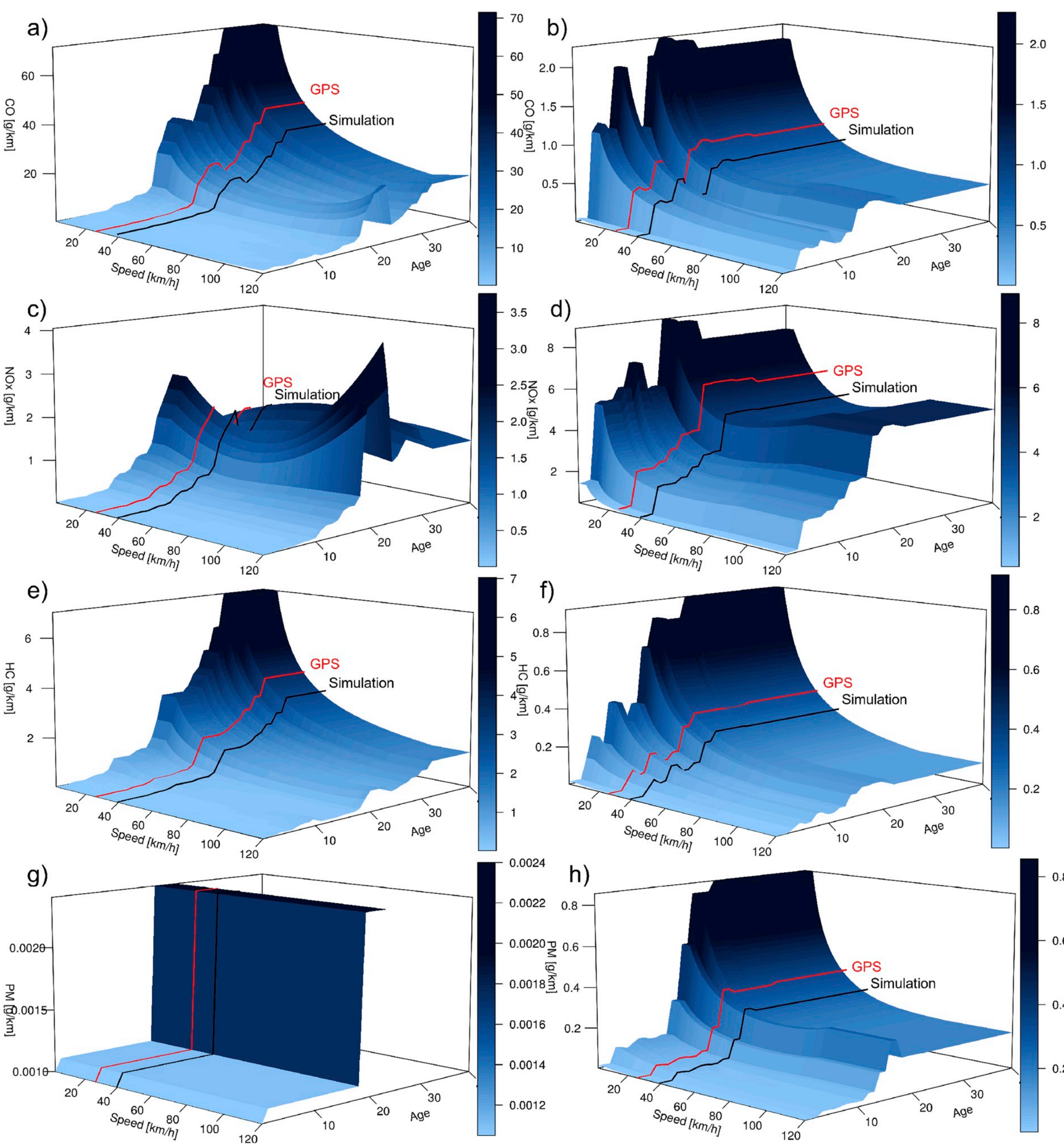


Fig. 4. Exhaust emission factors ($g{\cdot}km^{-1}$) depending on speed and age of use for Passenger Cars using Gasohol of a) *CO*, c) NO_X, e) *HC*, g) *PM*, and Light Trucks using Diesel of b) *CO*, d) NO_X, f) *HC* and h) *PM* .

variations in ambient temperatures. VEIN includes methods to estimate evaporative emissions based on the European Council emissions guidelines (Mellios and Ntziachristos, 2016). Also, VEIN includes CETESB emission factors, which are average values from emissions certification test Sealed Housing for Evaporative Determination (SHED) (CETESB, 2015). The standard consists in 1 g for only 2 h/test, while in China, Europe and USA the duration of the tests covers 24 h, 48 h or 72 h (Wu et al., 2017). Therefore, CETESB emission factors are presumably low. The emissions factors for diurnal ($g{\cdot}day^{-1}$), hot_soak ($g{\cdot}\ trip^{-1}$) and running_losses ($g{\cdot}trip^{-1}$) evaporative emissions, for 4-strokes PC and LCV and temperatures in ranges of and 20 °C and 35 °C are shown on Fig. 5. For comparison, it was added European emission factors (Mellios and Ntziachristos, 2016) showing that diurnal emissions are lower than European, hot-soak are relatively similar and running-losses are higher. As conclusion, CETESB emissions factors still provide insightful information. Vehicles older than 28 years produce higher emissions and, unsurprisingly, evaporative emissions increase with temperature.

To our knowledge, the only evaporative emission factors available in literature with Brazilian chemical composition in the fuel comes from CETESB. However, as our activity is traffic flow on streets, we had to transform the evaporative emission factors suitable. Hence, we converted each emission factor into g·km^{-1}. We divided the emission factors of diurnal evaporative ($g{\cdot}day^{-1}$) with the daily mileage ($km{\cdot}\ day^{-1}$) by age of use of each type of vehicle. We used the mileage functions inside VEIN that are based on odometer reading of more than 1.6 million vehicles from inspection and maintenance programs made in São Paulo (Bruni and Bales, 2013), these functions return the annual mileage which was divided by 365 to obtain the daily mileage. The other emission factors in $g{\cdot}trip^{-1}$ were converted to $g{\cdot}day^{-1}$ multiplying by 4.6 $trips{\cdot}day^{-1}$ and then we divided by the daily mileage $km{\cdot}day^{-1}$ by age of use of each type of vehicle. Therefore we used this value from the statistics of the (Bureau of Transportation Statistics, 2017), U.S. Department of Transportation. The sum of the three types of evaporative emission factors are shown in Fig. 5. To our knowledge, there are no statistics with the number of trips per day in São Paulo. Evaporative emissions factors, as shown on Fig. 5, have a clear pattern of higher emissions in older vehicles.

3. Results

3.1. Traffic calibration

Traffic data from all sources were calibrated according to local fuel sales. Table 3 shows the calibration factors for traffic by type of fuel and region. This indicates that, prior to calibration, GPS-based methods tended to overestimate traffic activity values for both gasohol (E25) and ethanol (E100) fueled vehicles and underestimate flows for diesel (B5) fueled vehicles, except at MASP. This region has more *motorways* and *trunks* than other regions, so may be more sensitive to the use of speed as proxy for other traffic properties, described in detail in an associated MethodsX paper (Ibarra-Espinosa et al., 2019). Simulation-based activity values were only available to MASP, but these tended to require less aggressive calibration.

3.2. Emissions inventories

Table 4 shows the emissions inventories for each region; CO_2* is expressed as 10^6t and Vehicles times kilometers (VKM)** as 10^{12}. In all regions, light vehicles are more responsible for emissions of CO, HC, NH_3 and CH_4, and heavy vehicles for PM, NO_X and SO_2. In the case of CO_2, the emissions are more equally distributed between these vehicles. Consistent with their VEIN profiles, lighter vehicles emits more CO, $NMHC$, NH_3 and CH_4, whilst heavier vehicles produce more PM, NO_X and SO_2, and CO_2 emissions are more evenly distributed across the vehicle fleet. The total emissions for MASP area are 3021.3 $t{\cdot}y^{-1}$ and 2281.1 $t{\cdot}y^{-1}$ of PM, 283322.6 $t{\cdot}y^{-1}$ and 177406.8 $t{\cdot}y^{-1}$ of CO, 89462.6 $t{\cdot}y^{-1}$ and 73554.6 $t{\cdot}y^{-1}$ of NO_X, 89485.4 $t{\cdot}y^{-1}$ and 33999.3 $t{\cdot}y^{-1}$ of $NMHC$, 2310.4 $t{\cdot}y^{-1}$ and 1670.6 $t{\cdot}y^{-1}$ of NH_3, 2019.8 $t{\cdot}y^{-1}$ and 2000.1 $t{\cdot}y^{-1}$ of SO_2, 6018.7 $t{\cdot}y^{-1}$ and 4156.9 $t{\cdot}y^{-1}$ of CH_4 and 21.72 $Mt{\cdot}y^{-1}$ and 21.05 $Mt{\cdot}y^{-1}$ of CO_2, from Method A (GPS) and Method B (travel demand simulation), respectively. The emissions from Method A are higher than Method B for all pollutants, however, the values for CO_2 are

Table 3
Calibration factors for traffic by type of fuel in each region and scenario.

Region	E25		E100		B5	
	Factor	%	Factor	%	Factor	%
MASP (Method A)	0.495	101.22	0.393	100.05	0.0124	84.14
Campinas (Method A)	0.2695	101.01	0.262	100.61	4.63	84.39
SJDC (Method A)	0.362	100.27	0.238	100.16	3.7	85.40
Santos (Method A)	0.265	100.09	0.164	100.73	5.15	85.61
Sorocaba (Method A)	0.181	100.50	0.199	100.55	3.5	85.56
MASP (Method B)	1.135	100.41	0.905	100.47	0.36	85.46

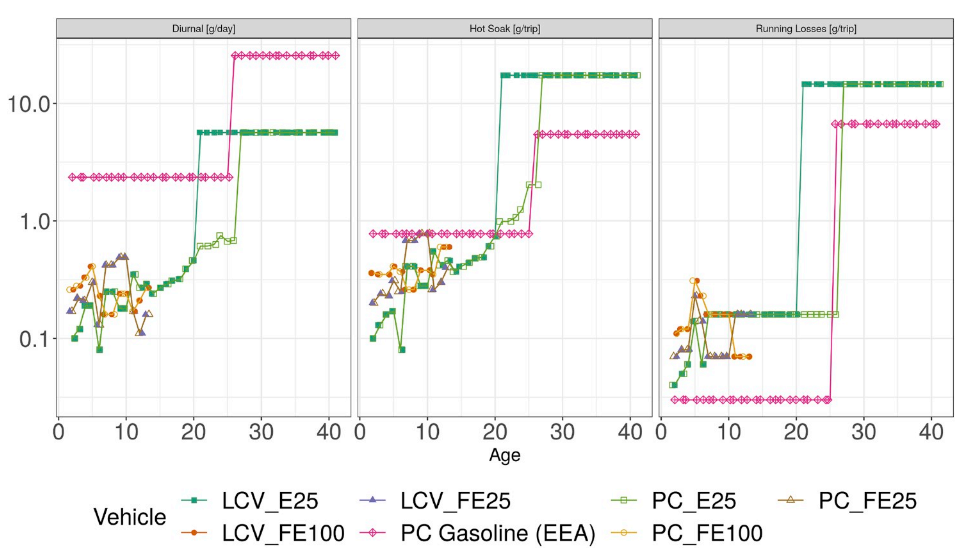


Fig. 5. Sum of Diurnal, Hot-Soak and Running Losses evaporative emission factors (g·km^{-1}) by age of use for an averaged temperature between 20 °C and 35 °C.

Table 4
Emissions ($t \cdot y^{-1}$) for each regions and scenario in 2014.

Region	*Veh*	*PM*	*CO*	NO_X	*NMHC*	NH_3	SO_2	CH_4	CO_2*	*VKM***
Camp Method A	PC	20.80	72743.70	7214.00	25056.00	741.80	174.50	1741.90	4.06	61.67
	LCV	52.50	7033.50	1752.50	4480.60	103.90	32.00	233.00	0.83	10.41
	HGV	2424.40	6660.30	42079.70	1830.30	24.40	942.20	23.10	3.98	9.78
	MC	0.00	11636.10	780.00	1151.50	8.70	10.30	224.30	0.28	10.47
Santos Method A	PC	5.40	17640.30	1868.00	6159.00	173.50	36.90	417.10	0.90	2.40
	LCV	13.70	1846.40	442.60	1134.60	25.20	7.30	56.00	0.20	0.43
	HGV	775.90	2052.50	13191.80	546.20	7.80	296.90	7.40	1.25	0.61
	MC	0.00	3275.00	222.30	297.80	2.20	2.70	57.50	0.07	0.46
SJDC Method A	PC	5.80	19148.50	1988.90	6630.60	187.30	40.60	449.00	0.99	2.73
	LCV	14.60	2025.70	474.40	1218.40	27.10	8.00	60.20	0.22	0.48
	HGV	513.30	1434.30	8965.40	393.50	5.20	202.00	4.90	0.85	0.39
	MC	0.00	3440.30	232.40	319.40	2.40	2.90	61.40	0.08	0.52
Soro Method A	PC	5.00	17708.90	1856.20	6061.70	186.00	43.80	433.60	1.00	1.97
	LCV	12.70	1834.00	444.00	1068.10	25.70	7.80	57.90	0.20	0.33
	HGV	556.80	1409.60	9254.60	364.10	5.60	208.40	5.30	0.88	0.36
	MC	0.00	3146.30	215.70	275.00	2.10	2.60	54.60	0.07	0.32
MASP Method A	PC	58.40	216870.40	19458.40	70080.00	1974.60	477.80	4689.70	11.40	493.06
	LCV	147.50	24418.70	4851.90	12983.20	281.60	92.60	628.50	2.43	101.06
	HGV	2555.70	7865.30	47567.40	2349.50	25.80	1071.60	24.40	4.49	70.24
	Bus	259.70	3364.00	15624.40	615.30	4.40	349.20	51.90	2.63	5.56
	MC	0.00	30804.20	1960.50	3457.40	24.00	28.60	624.20	0.77	105.36
MASP Method B	PC	53.70	127637.80	16307.70	30435.90	3422.40	278.20	3069.50	11.01	1932.57
	LCV	136.8	14021.70	3810.90	5692.70	339.30	178.50	433.9	2.63	356.35
	HGV	1582.2	6089.56	38455.53	1486.64	28.42	1125.89	17.05	4.92	122.25
	Bus	84.65	1319.95	5853.24	207.30	1.64	201.63	15.87	1.10	10.15
	MC	0.00	28622.60	1863.30	3382.80	25.20	11.90	651.40	0.76	336.43
MASP CETESB	PC	56	103858	11909	24302		302		7.16	
	LCV	10	18484	6118	4071		387		2.23	
	HGV	756	3440	20054	985		667		2.97	
	Bus	410	2842	15117	658		69		1.58	
	MC	69	34271	1134	4808		138		0.45	

practically the same.

In MASP, the aggregation of emissions from passenger cars, light-commercial vehicles and motorcycle account for 6.84% and 8.06% of *PM*, 96.04% and 95.16% of *CO*, 29.37% and 29.65% of NO_X, 96.69% and 93.96% of *NMHC*, 98.69% and 97.89% of NH_3, 29.66% and 23.88% of SO_2, 98.73% and 99.09% of CH_4 and 67.22% and 68.05% of CO_2, from GPS and travel demand simulation, respectively. The results present convergence between percentages of emissions using both methods, in agreement with other studies (D'Angiola et al., 2010a; D'Angiola et al., 2010b; de Souza et al., 2013; CETESB, 2015).

Camp is Campinas, SJDC is São José Dos Campos, MASP is Metropolitan Area of São Paulo. CO^{2*} are divided by 10^6, VKM are divided by. 10^12

Emissions were higher in MASP by comparison to other regions, reflecting trends in the activity data. For example, in the cases of *PM*, *CO*, NO_X, *NMHC*, CH_4 and CO_2, the highest emissions were all obtained for MASP using GPS methods, which estimated the lowest VKM values and speeds, as shown in Fig. 4. This is because lower speeds are associated with congestion and higher emissions in the VEIN source profiles. As this is highly likely to be representative of the situation in MASP by comparison to, e.g., Santos, this highlights the importance of incorporating the most representative measures of kinematic parameters into vehicular emissions inventorying methods.

The vehicular emissions inventories produced in this study also

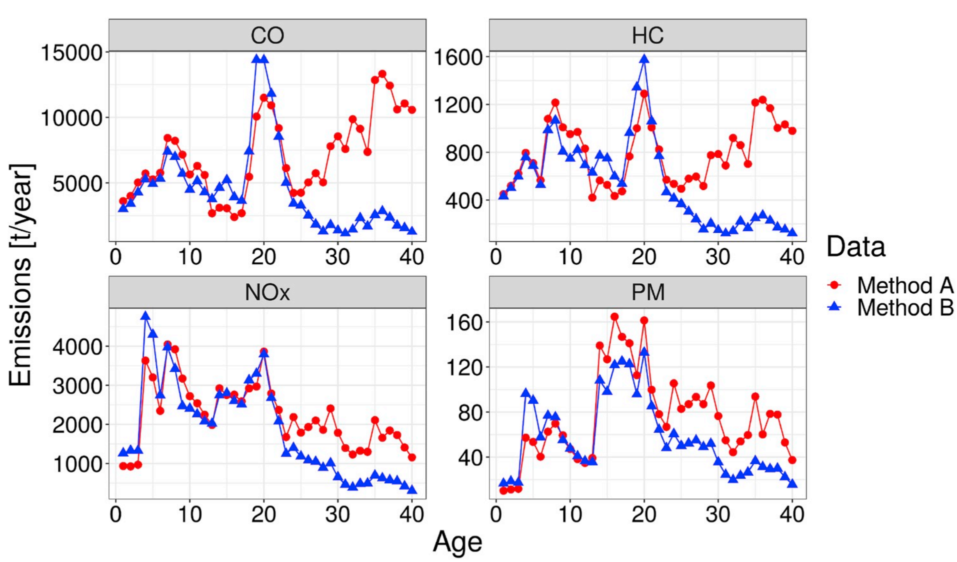


Fig. 6. Signature of vehicular emissions in $t \cdot y^{-1}$ by age of use and source of traffic in all metropolitan regions.

attempt to estimate the potential influence of vehicle age. With this in mind, we present figures showing the total amount of pollutants by age of vehicle. This allows us to propose a regional *signature* for each pollutant, which we define as the total amount of emissions by vehicle age, see e.g. Fig. 6. To understand these figures it is necessary to remember that emission factors are higher for older (by comparison to newer) vehicles as shown on Fig. 4 due to the deterioration of vehicles, and also due to the fleet distribution. Emissions of *CO* and *HC* shows three peaks, for vehicles newer than 10, at 20 and older than 30 years. The reason of the first peak is the rapid increase in fleet size between 2000 and 2011. The second peak is associated with 19–23 year old vehicles. According to the Manual of Air Pollution Control Program by Motor Vehicles (de Macedo et al., 2011), vehicles with catalytic systems entered the fleet with the Brazilian emissions standard Proconve L2, in 1992. Therefore, the second peak is due to vehicles of this era with deteriorated catalytic systems. The third peak is due to relatively small, but much older fleet with higher associated emission factors. Emissions of NO_x and *PM* shown on Fig. 6 exhibit a less similar trend with a higher central peak, that associates with newer vehicles.

The emission maps of *CO*, NO_X, *HC* and *PM* for traffic data from

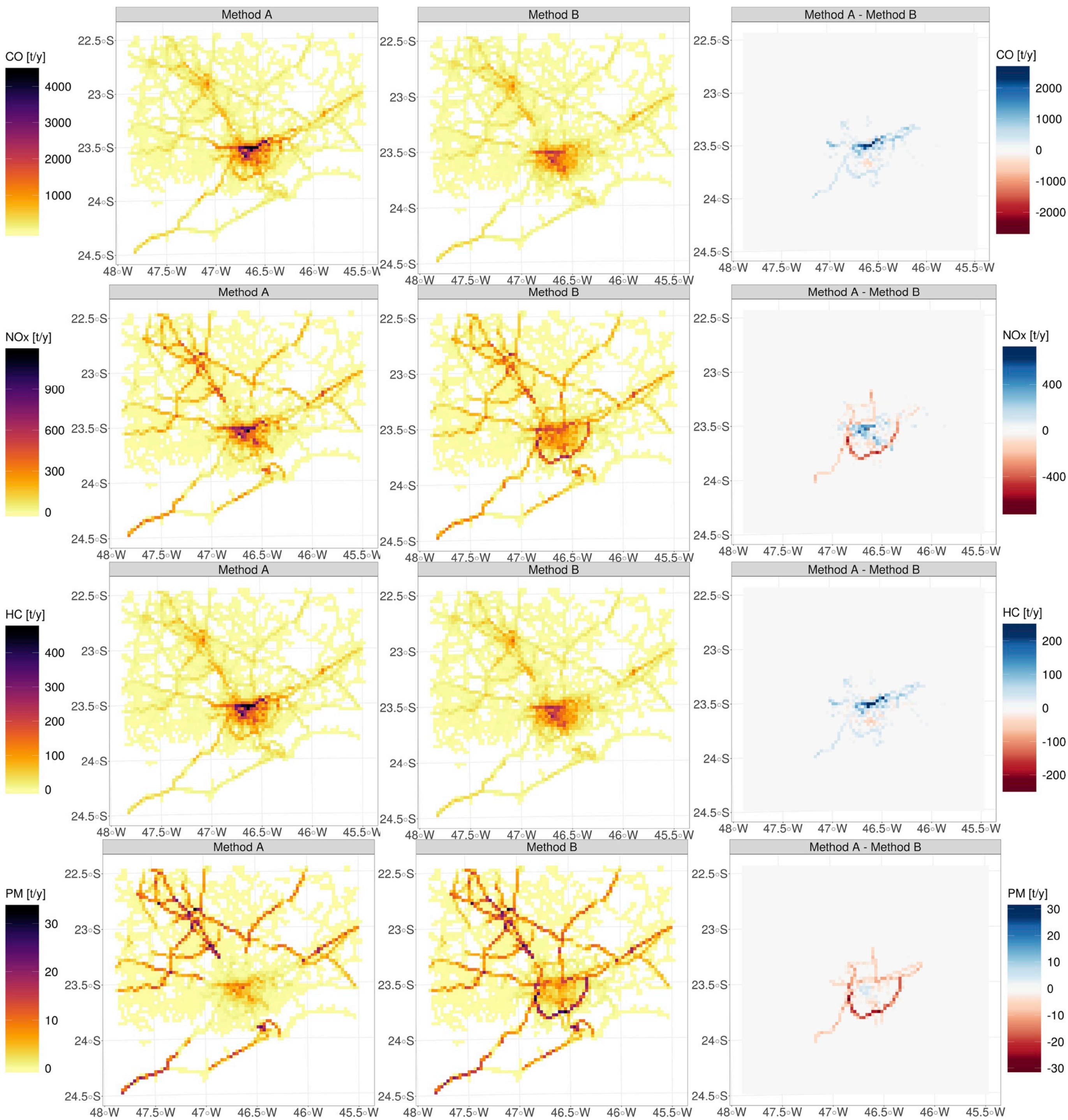


Fig. 7. Emission maps of *CO*, NO_X, *NMHC* and *PM*, in $t{\cdot}year^{-1}$, due traffic activity from scenarios Method A (left) and Method B (center) and difference between the two (right) in 2014.

Method A, and Method B and their difference are shown in Fig. 7. *CO* maps generated using Method A and Method B data look reasonably similar and their difference shows *CO* from Method A data is higher, specially in urban central motorways. In the case of NO_X, maps generated using travel demand model data clearly indicate a stronger contribution from the motorway ring road to the south of MASP. This is less pronounced for Method A data derived NO_x maps because speeds from *trunks* and *motorways* were used as a proxy for traffic flow. The difference of NO_X shows also the ring Method B produced more emissions than Method A into the motorway ring road, however, Method A produced more emissions towards the center of the city. The *HC* emissions seems to be similar with CO, with hot-spots in the center of the city and the difference also showing more contribution from Method A data. The emissions of PM follows a similar pattern to NO_x, with a clear contribution from Method B in the motorway ring. There is also an emissions peak in a street in the south west part of the region, at longitude 47.5 W and latitude 24.4 S, due to high truck activity in this area. We used satellite data to corroborate this finding.

The hourly emissions of *CO*, NO_X, *NMHC* and *PM*, for the 168 h of the study period are shown in Fig. 8, at the left the scenario with Method A and the right the scenario with Method B for MASP. Emissions estimated with Method A have a similar pattern for different pollutants, and, in the case of Method B, morning rush traffic flows were extrapolated using temporal factors of the same normalized traffic counts from the nearest toll stations from São Paulo. Emissions from GPS data presents profiles relatively constants for *CO* and *NMHC*, with lower emissions at night and early morning, and lower emissions on weekends. In the case of NO_x the profile is more marked, reflecting the contributions from Buses as estimated using the SPTrans traffic simulation. On contrast, emissions from the Method B have marked profiles in all emissions.

4. Discussion and uncertainty

This paper attempts to contribute to our understanding of air pollution by developing detailed bottom-up vehicular emissions inventories, using as its basis real-world GPS and simulation data on speeds and flows of different vehicle types, and regional fuel sales rescaling to account for provide a potentially more robust measure of vehicle activity. This also uses VEIN (Ibarra-Espinosa et al., 2018) to incorporate vehicle age adjusted emission factors for local fleets in Southeastern Brazil. Although we believe this provide a good measure of local traffic-related emissions, we also feel it is important to acknowledge that each stage of the inventorying process included difficulties, challenges, assumptions and uncertainty, and that these provide opportunities for future improvements.

Developing a bottom-up vehicular emissions inventory with data from massive GPS (multiple-vehicle type) recordings is, to our knowledge, a new approach and methods may require further refinement but this paper provides, at the very least, an exploratory analyses. The use of large GPS data sets as a source of traffic activity information seems an obvious step-forward by comparison to traditional used in traditional emissions inventorying projects. That said, it is important to acknowledge that the certainty of outputs is influenced by the quality and frequency of GPS, its degree of fleet coverage and its potential bias, for example, towards newer vehicles or commercial fleets as coverage of these is likely to be higher. As the traffic flow calculation is based on

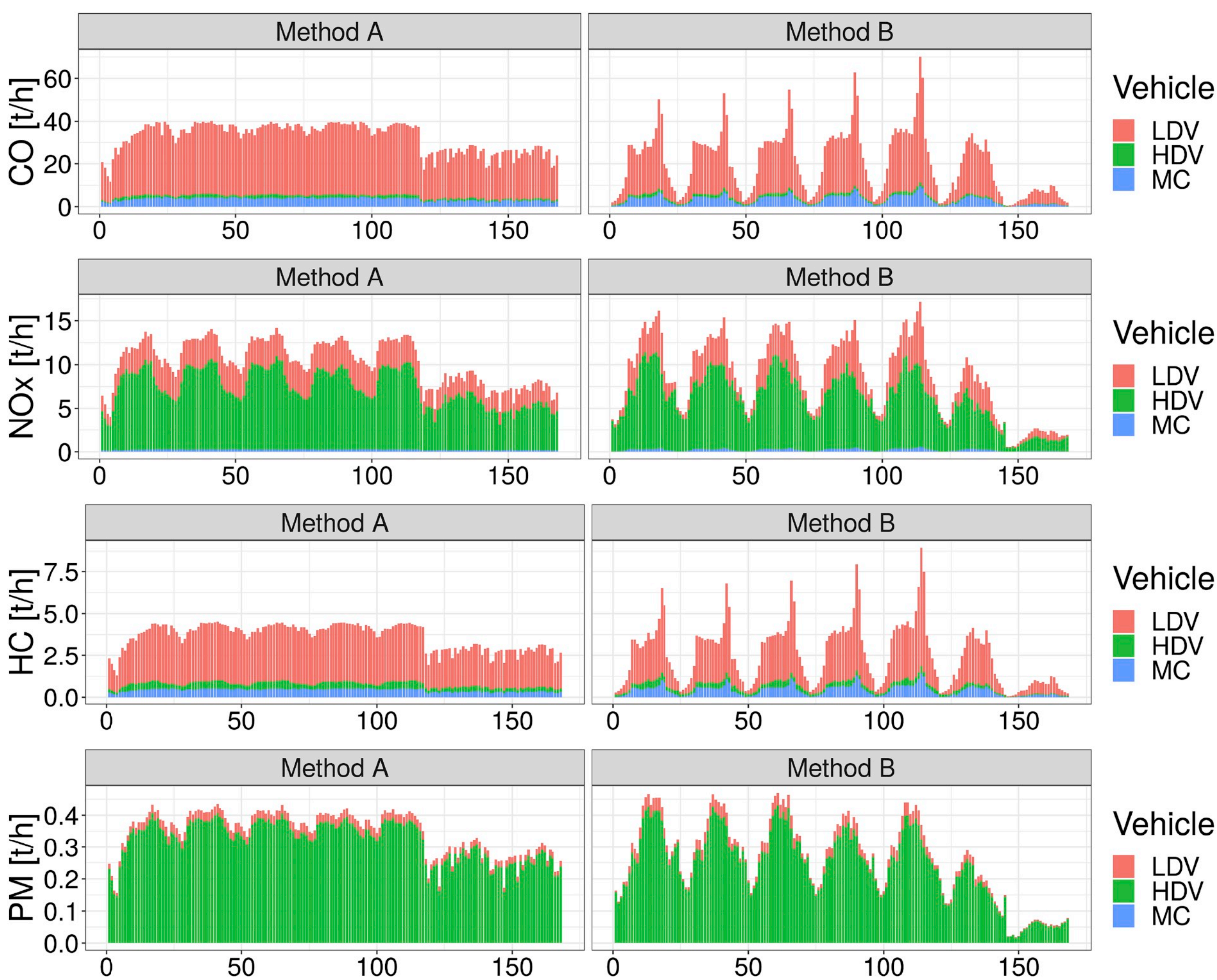


Fig. 8. Hourly emissions of *CO*, NO_X, *HC* and *PM*, in $t{\cdot}h^{-1}$ due traffic activity from scenarios Method A (left) and Method B in MASP (right).

average speed by vehicle, if the frequency of the GPS observation is higher the uncertainty on the speed calculation and hence the traffic flow is reduced and it is likely that newer fleets can count with GPS with the mentioned characteristics. Data sets and methods such as these that consider multiple vehicle types and attempt to re-scale measurements according to independent measures of fleet sizes provide an obvious step-forward by contrast to earlier GPS-based methods that extrapolate GPS data from one sub-group of drivers to all of the vehicle fleet independent of driver or vehicle type. Access to larger and more complete data sets with improve with time, as, hopefully, will our access to more detailed data, allowing us to move beyond the use of current compromises such as the use of speed as proxy for traffic flow.

The process of generating trucks flow from GPS recordings consisted in using the speed of trucks and when there is no available information, the average speed of all vehicles only for *motorways* and *trunks* type of streets. This process seems appropriate for light-duty vehicles but in the case of trucks could lead to assignment in the same streets of light-duty vehicles which could not necessary be truth. Circulation of trucks deliver goods and this determine the route they follow. Therefore, this traffic flow could not have an exact spatial representation. This could be reason why PM and NO_X emissions from GPS have different spatial distribution on Fig. 7 than the generated with travel demand models. This figure shows clearly a ring of NO_X and PM emissions surrounding the city of São Paulo in emission maps generated using travel demand simulation but not with GPS data. On the other hand, this data provided insights of trucks flow that hardly could be obtained from other sources. For example, trucks on the Santos Region are important because this city is a port, attracting and generating goods delivery using trucks. Traffic data from GPS in Santos identified streets with high number of trucks near the port and also, in the streets at south est of the study area. We corroborated this findings with Google image satellites (not shown here). Traffic speeds from GPS was used for generating traffic flow based on a set of traffic and speed measurements for MASP during 2012, (Ibarra-Espinosa, 2017). Unequivocally, updating and enhancing this information with new data sets could improve traffic flow generation. Therefore, traffic data from Internet GPS recordings is a promising new traffic data source for emissions inventories.

Where travel demand simulation was used, for example for CET light-duty vehicles and trucks morning rush hour SPtrans Urban buses, these incorporated uncertainties regarding the exact number of vehicles circulating. CET performs traffic counts to calibrate traffic simulations (Ibarra-Espinosa, 2017) in the mains routes at MASP. Both simulations include a limited number of local streets as virtual links located at the centroid of each zone of origin-destination survey. These links represent real traffic over virtual links, therefore, provides an associated element of uncertainty. However, most of traffic occurs in mains streets, such as *motorways*, *trunks* and *primary*, therefore the uncertainty of virtual links should be small for a emissions grid.

Traffic calibration using fuel sales resulted in consistent traffic flows regarding the origin of the data. This can be seen in the CO_2 emissions which resulted in 21 $10^6 \cdot y^{-1}$ in MASP using both scenarios. In the absence of complete GPS coverage, calibration according to fuel sales seems our most appropriate option, especially with regard to CO_2 emissions which depends most directly on fuel consumption. There is a source of uncertainty in fuel consumption of diesel, because sales statistics (de Energia Electrica, 2014) shows total diesel, without discriminating the fuel used only for automotive purpose. For instance, there are many electricity generators that consumes diesel. In addition, as trucks travel long distance for delivering goods, fuel bought for trucks could consumed outside the study area adding uncertainty. We believe that the using the criteria of calibrating the vehicles that consumes diesel so that it represents the 85% of diesel sales is conservative, but it is at the moment an assumption that we are applying until better information is available.

The emissions estimated were consistent in each region. The difference in MASP region for each pollutant presents differences associated with the speeds and amount of vehicles. As Table 4 shows, the VKM in MASP Method B is 2796 $10^{12} \cdot y^{-1}$, but in MASP Method A is 775 $10^{12} \cdot y^{-1}$, approximately 3.4 times lower. However, the emissions in MASP Method A are higher in pollutants emitted more for gasohol and ethanol vehicles. This emissions are due to lower speeds because emission factors are higher when lower the speeds as shown on Fig. 4, which speed variation comes from emission factors of European emission guidelines (Ntziachristos and Samaras). This highlights the importance of including kinematic parameters in the estimation of vehicular emissions. Regarding the emission factors, there are elements of uncertainty in the emission factors measurement. Brazilian emission factors of light-duty vehicles and motorcycle comes from dynanometer measurements (CETESB, 2015) and it has been discovered that these emissions are not representative of real world emissions (Pelkmans and Debal, 2006), in addition, car manufacturers incorporated software in the vehicles to cheat emission certification tests. In Brazil, the Institute of Natural Resources and Environment (IBAMA) fined Volkswagen with USD 15 10^6 after a tests using portable emissions measurements in Brazil made by the Environmental Company of São Paulo (IBAMA, 2017). However, the diesel fleet of light-duty vehicles is small and the impacts on vehicular estimation may not be significant. These examples all strongly suggest that Brazilian emission factors could be underestimated and highlight the need for a coordinated program of real world emissions measurements work in Brazil. In addition, there are uncertainty also in the deterioration factors used in this study, because they were calculated in Europe (Ntziachristos and Samaras) and they may not fully represent Brazilian vehicular deterioration. Lastly, the emission factors in (CETESB, 2015) are based on certification tests using the FTP-75 driving cycle. The factors are all-cycle values, independent of test phase, and it is known that, firstly, emissions vary significantly with driving conditions and environment, and, secondly, that cold start emissions can be very important, and they increase when the temperature decrease (Ntziachristos and Samaras).

There are other corrections that could improve the emission factors, such as the incorporation of percentage of load in trucks and the effect of the road gradient. This could be important in the streets that connect São Paulo and Santos, because the load could be different in each trip. In addition, the road gradient is likely to be significant because Santos is on the coast, whilst São Paulo is over 760 m above sea level. The streets with the highest gradients are located near Cubatão, an industrial district between Santos and São Paulo, and including this effect would change the atmospheric chemistry in simulations at the interface between both cities.

We compared our inventory of MASP simulation and GPS Sorocaba, Campinas, Santos and SJDC with the emissions of different cities in Latin America shown on Table (5). The first comparison is between the official emissions inventory for São Paulo, which uses a top-down approach with statistics of fleet and representative speeds. In this case, the CO emissions are quite similar, our estimate being only 10% higher, but differences for other pollutants are bigger, for example, our CO_2 estimate is 50% higher.

The comparison with the official inventories for the other cities shows that our emissions are higher in all cases. The difference in the emissions is that our traffic flow was calibrated with fuel sales and also due the influence if the speed in emission factor. The comparison between MASP and Rio de Janeiro and Buenos Aires shows the estimates in those cities are higher, which could be due to the fact the base years are 2010 and 2006 respectively and the circulating fleet had older technology allowing more emissions. Porto Alegre emissions were also relatively high but here the estimate is based on the 2004 fleet. The comparison with Santiago seems to give reasonable estimates despite that the fleet is older, however, the number of vehicles circulating in Santiago is smaller because of the size of the city and road network. The comparison with Manizales seems to be reasonable for most pollutants,

Table 5
Comparison with vehicular emissions inventories in different cities ($t \cdot y^{-1}$).

City	PM	CO	NO_X	NMHC	$CO_2 \cdot 10^6$	Reference
MASP, BR	2281	177406	73554	33999	21.05	[a]
MASP, BR	1484	162896	54334	34824	14.28	CETESB (2015)
RJ, BR	2917	583667	68400	69689		de Souza et al. (2013)
Santos, BR	795	24814	15724	8138	2.43	[a]
Santos, BR	182	13497	6157	2628	1.38	CETESB (2015)
Sorocaba, BR	574	24098	11770	7769	2.16	[a]
Sorocaba, BR	257	20203	8832	4070	1.9	CETESB (2015)
SJDC, BR	533	26048	11661	8562	2.13	[a]
SJDC, BR	271	27406	10116	5287	2.4	CETESB (2015)
Campinas, BR	2497	98073	51826	30871	9.14	[a]
Campinas, BR	377	34890	13851	7180	3.3	CETESB (2015)
Porto Alegre, BR	2350	195740	34111	23450		Teixeira et al. (2008)
Manizales, CO	800	43400	4900	9600	0.45	González et al. (2017)
Santiago, CL	1275	199884	40126	16668	7.3	Escobaret al. (2005)
Buenos Aires, AR	6370	569000	81900	69800	11.5	D'Angiola et al. (2010b)

[a] This study.

except for CO which our estimates are lower in most cities. The only city in which the HC emissions were significant higher than NO_X was Manizales. This seems to relate to evaporative emissions estimations, which here incorporate local survey data on both the distribution of trips and number of engine start-ups. The survey found that there were typically 5 trips per day for light vehicles. By contrast, we adopted a value of 4.5 taken from the literature. This small difference appears to have a large impact on evaporative emissions estimations, suggesting that any new evidence on this parameter could significantly enhance the robustness of the emissions inventory for São Paulo.

5. Conclusions

This study investigated vehicular emissions in Southeastern Brazil using two sources of traffic data 1): traffic flows using travel demand model outputs and 2) a massive data set of GPS location of cars and trucks. The increased availability of internet GPS traffic data provides us with the opportunity to use much more representative measures of vehicle activity when developing new emissions inventories. However, care must be taken when dealing with such data, especially in obtaining representative speeds, due to the low frequency of consecutive GPS readings of the same vehicle.

Arguably, at this stage traffic simulation provides the most straightforward and, perhaps even most reliable, option for sourcing traffic activity data, but this is a rapidly evolving field, and work such as this highlights the potential role, and associated challenges, for those seeking to make best use of local GPS data in such work.

The calibration factors for travel demand models in Table 3 are closer to one, meaning that the estimates of on-road fuel consumption derived from travel demand model outputs are more representative. Therefore, we encourage city planners to use travel demand model data when undertaking environmental assessments of city projects.

Declaration of competing interest

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

Acknowledgments

The authors would like to thank CETESB, CET, ARTESP, CAPES Financial code 001 (program by the Ministry of Education, Brazil) and the Fundação da Universidade de São Paulo. The authors thank the program Becas Chile from CONICYT (Ministry of Education, Chile), scholarship Resolución Exenta N°9198/214. The authors also thank the support provided by the Northeast Institute of Geography and Agroecology, Chinese Academy of Sciences. These entities provided invaluable support to this study.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.atmosenv.2019.117136.

References

ABRACICLO, 2014. Anuario 2014. http://www.abraciclo.com.br/.

D. Alvim, L. Gatti, S. Correa, J. Chiquetto, G. Santos, C. de Souza Rossatti, A. Pretto, J. Rozante, S. Figueroa, J. Pendharkar, P. Nobre, Determining vocs reactivity for ozone forming potential in the megacity of são paulo, Aerosol and Air Quality Research 18. doi:10.4209/aaqr.2017.10.0361.

M. Andrade, P. Kumar, E. Freitas, R. Ynoue, J. Martins, L. Martins, T. Nogueira, P. Perez-Martinez, R. Maura de Miranda, T. Albuquerque, F. Gonçalves, B. Oyama, Y. Zhang, Air Quality in the Megacity of São Paulo: Evolution over the Last 30 Years and Future Perspectives, Atmospheric Environment In Press. doi:10.1016/j.atmosenv.2017.03.051.

Andrade, M.d.F., Ynoue, R.Y., Freitas, E.D., Todesco, E., Vara Vela, A., Ibarra, S., Martins, L.D., Martins, J.A., Carvalho, V.S.B., 2015. Air quality forecasting system for southeastern Brazil. Frontiers in Environmental Science 3, 1–12. https://doi.org/10.3389/fenvs.2015.00009.

ANFAVEA, 2014. Estatisticas produção, vendas e exportação de autoveículos. http://www.anfavea.com.br/estatisticas.html.

Barlow, T., Latham, S., McCrae, I., Boulter, P., 2009. A Reference Book of Driving Cycles for Use in the Measurement of Road Vehicle Emissions. TRL Published Project Report.

Bruni, A.D.C., Bales, M.P., 2013. Curvas de intensidade de uso por tipo de veículo automotor da frota da cidade de são paulo. Tech. rep., CETESB 1. ISBN: 978-85-61405-61-8. http://veicular.cetesb.sp.gov.br/wp-content/uploads/sites/35/2013/12/curvas-intensidade-uso-veiculos-automotores-cidade-sao-paulo.pdf.

Bureau of Transportation Statistics, 2017. Daily Passenger Travel, Tech. Rep.. https://www.rita.dot.gov/bts/sites/rita.dot.gov.bts/files/publications/highlights_of_the_2001_national_household_travel_survey/html/section_02.html.

CETESB, 2015. Emissões veiculares no estado de são paulo 2014. http://veicular.cetesb.sp.gov.br/relatorios-e-publicacoes/.

Corvalán, R.M., Osses, M., Urrutia, C.M., 2002. Hot emission model for mobile sources: application to the metropolitan region of the city of santiago, Chile. J. Air Waste Manag. Assoc. 52 (2), 167–174.

Davis, N., Lents, J., Osses, M., Nikkila, N., Barth, M., 2005. Part 3: developing countries: development and application of an international vehicle emissions model, Transportation Research Record. Journal of the Transportation Research Board 8 (1939), 155–165.

de Energia Electrica, S., 2014. Anuario estatistico de energéticos por municipio no estado de são paulo 2014, Tech. rep. Goberno do Estado de São Paulo.

de Macedo, P.C., de Moraes Gomes, D., Pereira, E.A., Veloso, M.B., 2011. Air Pollution Control Program by Motor Vehicles, third ed., vol. 3. Brazilian Institute of Environment and Renewable Natural Resources (IBAMA), Brasília, Brazil.

de Medio Ambiente, M., 2011. Air pollution control program by motor vehicles; proconve/promot. In: http://www.ibama.gov.br/phocadownload/veiculosautomotoresmanual%20proconve%20promot_english.pdf, 2017-10-17.

de Souza, C.D.R., Silva, S.D., da Silva, M.A.V., Márcio de Almeida, D., Barboza, A.P., 2013. Inventory of conventional air pollutants emissions from road transportation for the state of rio de janeiro. Energy Policy 53, 125–135.

Dieselnet, 2017. European Stationary Cycle (ESC). ECOpoint Inc.. https://www.dieselnet.com/standards/cycles/esc.php

D'Angiola, A., Dawidowski, L.E., Gómez, D.R., Osses, M., 2010a. On-road traffic emissions in a megacity. Atmos. Environ. 44 (4), 483–493.

D'Angiola, A., Dawidowski, L.E., Gómez, D.R., Osses, M., 2010b. On-road traffic emissions in a megacity. Atmos. Environ. 44 (4), 483–493. https://doi.org/10.1016/j.atmosenv.2009.11.004. http://www.sciencedirect.com/science/article/pii/S1352231009009352.

J. Escobar, et al., Actualización del inventario de emisiones de contaminantes atmosféricos en la región metropolitana 2005, DICTUC, Pontificia Universidad Católica de Chile. Santiago: Commisión Nacional Del Medio Ambiente Región Metropolitana.

Gately, C.K., Hutyra, L.R., Peterson, S., Wing, I.S., 2017. Urban emissions hotspots: Quantifying vehicle congestion and air pollution using mobile phone gps data. Environ. Pollut. 229, 496–504.
Gómez, C., González, C., Osses, M., Aristizábal, B., 2018. Spatial and temporal disaggregation of the on-road vehicle emission inventory in a medium-sized andean city. comparison of gis-based top-down methodologies. Atmos. Environ. 179, 142–155.
González, C., Gómez, C., Rojas, N., Acevedo, H., Aristizábal, B., 2017. Relative impact of on-road vehicular and point-source industrial emissions of air pollutants in a medium-sized andean city. Atmos. Environ. 152, 279–289.
Hoshyaripour, G., Brasseur, G., Andrade, M.d.F., Gavidia-Calderón, M., Bouarar, I., Ynoue, R.Y., 2016. Prediction of ground-level ozone concentration in são paulo, Brazil: deterministic versus statistic models. Atmos. Environ. 145, 365–375.
IBAMA, 2017. Ibama confirma multa de R 50 milhoes por fraude em testes de emissao e exige recall da Volkswagen. http://www.ibama.gov.br/noticias/422-2017/1032-ibama-confirma-multa-de-r-50-milhoes-por-fraude-em-testes-de-emissao-e-exige-recall-da-volkswagen.
Ibarra-Espinosa, S., 2017. Traffic counts both ways cet 2012 sao paulo, Tech. rep., Medeley Data. 10-08-2017. https://doi.org/10.17632/rz2cymv6c2.1, 10.17632/rz2cymv6c2.1.
Ibarra-Espinosa, S., Schuch, D., 2017. eixport: Exporting emissions, r package version 0.2.10. https://github.com/ibarraespinosa/eixport.
Ibarra-Espinosa, S., Schuch, D., vein, 2019. Vehicular Emissions Inventories, r package version 0.7.4. https://atmoschem.github.io/vein/.
Ibarra-Espinosa, S., Ynoue, R., O'Sullivan, S., Pebesma, E., Andrade, M.D.F., Osses, M., 2018. Vein v0.2.2: an r package for bottom–up vehicular emissions inventories. Geosci. Model Dev. (GMD) 11 (6), 2209–2229. https://doi.org/10.5194/gmd-11-2209-2018. https://www.geosci-model-dev.net/11/2209/2018/.
Ibarra-Espinosa, S., Ynoue, R., Giannotti, M., Ropkins, K., de Freitas, E.D., 2019. Generating traffic flow and speed regional model data using internet gps vehicle records. Methods 6, 2065–2075. https://doi.org/10.1016/j.mex.2019.08.018.
IBGE, 2016. Ibge divulga as estimativas populacionais dos municípios em 2016. http://cgp.cfa.org.br/ibge-divulga-as-estimativas-populacionais-dos-municipios-em-2016/. (Accessed 9 August 2017).
Kinnee, E., Touma, J., Mason, R., Thurman, J., Beidler, A., Bailey, C., Cook, R., 2004. Allocation of onroad mobile emissions to road segments for air toxics modeling in an urban area. Transp. Res. D Transp. Environ. 9 (2), 139–150. https://doi.org/10.1016/j.trd.2003.09.003. http://www.sciencedirect.com/science/article/pii/S1361920903000877.
Mellios, G., Ntziachristos, L., 2016. Emep/eea Emission Inventory Guidebook; Gasoline Evaporation from Vehicles, Tech. Rep. European Environment Agency, Copenhagen.
Metro, 2017. Pesquisa origem e destino, 10-08-2017. https://transparencia.metrosp.com.br/dataset/pesquisa-origem-e-destino.
Molina, M.J., Molina, L.T., 2004. Megacities and atmospheric pollution. J. Air Waste Manag. Assoc. 54 (6), 644–680.
L. Ntziachristos, Z. Samaras, Emep/eea Emission Inventory Guidebook; Road Transport: Passenger Cars, Light Commercial Trucks, Heavy-Duty Vehicles Including Buses and Motorcycles, European Environment Agency, Copenhagen.
Nyhan, M., Sobolevsky, S., Kang, C., Robinson, P., Corti, A., Szell, M., Streets, D., Lu, Z., Britter, R., Barrett, S.R., et al., 2016. Predicting vehicular emissions in high spatial resolution using pervasively measured transportation data and microscopic emissions model. Atmos. Environ. 140, 352–363.
OpenStreetMap contributors, Planet dump. retrieved from. https://planet.osm.org. https://www.openstreetmap.org.
Ortuzar, J.d.D., Willumsen, L.G., 2002. Modelling Transport, vol. 3. Wiley.
Panis, L.I., Broekx, S., Liu, R., 2006. Modelling instantaneous traffic emission and the influence of traffic speed limits. Sci. Total Environ. 371 (1–3), 270–285.
Pelkmans, L., Debal, P., 2006. Comparison of on-road emissions with emissions measured on chassis dynamometer test cycles. Transp. Res. D Transp. Environ. 11 (4), 233–241.
QGIS Development Team, 2017. QGIS Geographic Information System. Open Source Geospatial Foundation. http://qgis.osgeo.org.
Rafee, S.A.A., Martins, L.D., Kawashima, A.B., Almeida, D.S., Morais, M.V.B., Souza, R.V.A., Oliveira, M.B.L., Souza, R.A.F., Medeiros, A.S.S., Urbina, V., Freitas, E.D., Martin, S.T., Martins, J.A., 2017. Mobile and stationary sources of air pollutants in the amazon rainforest: a numerical study with wrf-chem model. Atmos. Chem. Phys. 1–35. https://doi.org/10.5194/acp-2016-1190. http://www.atmos-chem-phys-discuss.net/acp-2016-1190/.
Salvo, A., Geiger, F.M., 2014. Reduction in local ozone levels in urban são paulo due to a shift from ethanol to gasoline use. Nat. Geosci. 7 (6), 450.
Teixeira, E.C., Feltes, S., Santana, E.R. R.d., 2008. Estudo das emissões de fontes móveis na região metropolitana de porto alegre, rio grande do sul. Quím. Nova 31 (2), 244–248.
Vara-Vela, A., Andrade, M.F., Kumar, P., Ynoue, R.Y., Muñoz, A.G., 2016. Impact of vehicular emissions on the formation of fine particles in the são paulo metropolitan area: a numerical study with the wrf-chem model. Atmos. Chem. Phys. 16, 777–797. https://doi.org/10.5194/acp-16-777-2016.
Wu, Y., Zhang, S., Hao, J., Liu, H., Wu, X., Hu, J., Walsh, M.P., Wallington, T.J., Zhang, K.M., Stevanovic, S., 2017. On-road vehicle emissions and their control in China: a review and outlook. Sci. Total Environ. 574, 332–349.
Yang, D., Zhang, S., Niu, T., Wang, Y., Xu, H., Zhang, K.M., Wu, Y., 2019. High-resolution mapping of vehicle emissions of atmospheric pollutants based on large-scale, real-world traffic datasets. Atmos. Chem. Phys. 19 (13), 8831–8843.