# Physical and chemical properties of urban aerosols in São Paulo, Brazil: Links between composition and size distribution of submicron particles.

Djacinto Monteiro dos Santos[1], Luciana Varanda Rizzo[2], Samara Carbone[3], Patrick Schlag[1a], and Paulo Artaxo[1]

[1]Institute of Physics, University of São Paulo, São Paulo, Brazil
[2]Instituto de Ciências Ambientais, Químicas e Farmacêuticas, Universidade Federal de São Paulo, Diadema, São Paulo, Brazil
[3]Federal University of Uberlândia, Uberlândia, Minas Gerais, Brazil
[a]now at: Shimadzu Deutschland GmbH, 47269 Duisburg, Germany

**Correspondence:** Djacinto Monteiro dos Santos (djunior@if.usp.br)

**Abstract.** In this work, the relationships between size and composition of submicron particles ($PM_1$) were analyzed at an urban site in the Metropolitan Area of São Paulo (MASP), a megacity with about 21 million inhabitants. The measurements were carried out from $20^{th}$ December 2016 to $15^{th}$ March 2017. The chemical composition was measured with an Aerodyne Aerosol Chemical Speciation Monitor and size distribution with a TSI Scanning Mobility Particle Sizer 3082. $PM_1$ mass concentrations in the MASP had an average mass concentration of 11.4 $\mu$g m$^{-3}$. Organic aerosol (OA) dominated the $PM_1$ composition (56%), followed by sulfate (15%) and equivalent black carbon (eBC, 13%). Four OA classes were identified using Positive Matrix Factorization: oxygenated organic aerosol (OOA, 40% of OA), biomass burning organic aerosol (BBOA, 13%), and two hydrocarbon-like OA components (a typical HOA related to vehicular emissions (16%), and a second HOA (21%) representing a mix of anthropogenic sources). Particle number concentrations averaged 12100 $\pm$ 6900 cm$^{-3}$, dominated by the Aitken mode. Accumulation mode increased under relatively high $PM_1$ concentrations, suggesting an enhancement of secondary organic aerosol (SOA) production. Conversely, the contribution of nucleation mode particles was less dependent on $PM_1$ levels, coherent with vehicular emissions. The relationship between aerosol size modes and $PM_1$ composition was assessed by multilinear regression models. Secondary inorganic aerosols were partitioned between Aitken and accumulation modes, related to condensation particle growth processes. Submicron mass loading in accumulation mode was mostly associated with highly oxidized OOA and also traffic-related emissions. To the author's knowledge, this is the first work that uses the MLR methodology to estimate the chemical composition of the different aerosol size modes. The chemical composition with size-dependent PM provides innovative information on the properties of both primary and secondary organic aerosols, as well as inorganic aerosols in a complex urban environment. The results emphasize the relevance of vehicular emissions to the air quality at MASP and highlight the key role of secondary processes on the $PM_1$ ambient concentrations in the region since 56% of $PM_1$ mass loading was attributed to SOA and secondary inorganic aerosol.

# 1 Introduction

Air pollution has become a major problem in large urban conglomerates, particularly in megacities with a population exceeding 10 million inhabitants (Fuzzi et al., 2015). At these megacities, industrial and vehicular emissions are generally the main air pollution sources (Beirle et al., 2011; Zhu et al., 2012). The Metropolitan Area of São Paulo (MASP) is the largest South American megacity, with a population of about 21 million inhabitants (IBGE, 2020), an area of 7,947 km$^2$, comprising 39 municipalities and a fleet of 8 million vehicles. An important characteristic of the Brazilian vehicle fleet is that gasohol with 24% ethanol and pure ethanol are the main fuels, making Brazil, and especially the MASP, a unique case of large-scale biofuel usage worldwide. Successful public policies for controlling and monitoring industrial and vehicular pollutant sources have been implemented over the last decades in the MASP leading to positive impacts in reducing the emission of primary pollutants to the atmosphere (Andrade et al., 2017). However, many uncertainties are still found on the effect of emission reduction policies on the concentration of secondary pollutants, for example, ozone and secondary aerosols.

Aerosol particles have recognized effects on human health (Cohen et al., 2017) and critical role in the Earth's radiation budget (Boucher et al., 2013). However, the knowledge of the dynamics of physicochemical properties of submicron particles remains limited. The chemical composition and the size distribution of ambient aerosol are key parameters concerning its optical properties (Romano et al., 2019), ability to act as cloud condensation nuclei (CCN) (Che et al., 2016) and deposition efficiency into the human respiratory tract (Ching and Kajino, 2018). Furthermore, aerosol properties range widely in the atmosphere as a result of interaction between local and regional sources (Costabile et al., 2009) and atmospheric processing (Jimenez et al., 2009). Therefore, a current scientific effort is the characterization of the mechanisms of aerosol production and aging under ambient conditions, in particular the formation of secondary organic aerosols (SOA).

Source apportionment studies carried out at MASP (Castanho and Artaxo, 2001; Andrade et al., 2012; Pereira et al., 2017; de Miranda et al., 2018) have shown that vehicle traffic, resuspended soil dust and secondary processes are the main sources of PM$_{2.5}$ in the region. Although such studies have provided innovative knowledge about the main air pollution sources in MASP, most of them derived from filter-based offline chemical analyses. Low time resolution offline measurements are unable to describe the dynamic of the sources and processes of urban aerosols. To our knowledge, Brito et al. (2018) were the only ones to report high-resolution measurements of chemical composition of submicron particles recently in the MASP, so that conducting further studies in the region is fundamental for a better understanding of such dynamic processes.

Herein, near real-time PM$_1$ chemical composition was evaluated, including black carbon, inorganic aerosols and chemical classes of organic aerosol, using Positive Matrix Factorization. Furthermore, the relationships between aerosol chemical composition and particle number size distribution (PNSD) were investigated by using a multilinear regression (MLR) approach. The results presented here provide innovative insights on the association between sources and processes governing physical-chemical properties of atmospheric aerosol in a megacity largely impacted by traffic emissions and extensive biofuel usage.

## 2 Methodology

### 2.1 Sampling site and measurements

An aerosol and trace gas monitoring station was operated at one of the municipalities of the Metropolitan Area of São Paulo (MASP), at the Federal University of São Paulo (UNIFESP, 23º43'8"S 46º37'40"W, 769 m above mean sea level), in Diadema city (Figure 1). The municipality of Diadema is located 20 km southeast from São Paulo downtown in a region characterized by high population density and strong vehicular impact, besides emissions from industrial activities and proximity to preserved Atlantic forest areas. The combination of different anthropogenic and biogenic emissions results in complex physicochemical processes that promote the formation of SOA and ozone.

The measurements were carried out from $20^{th}$ December 2016 to $15^{th}$ March 2017, comprising 105 days of data acquisition. Trace gas analyzers Thermo 49i and Thermo 43i were used to monitor the concentration of trace gases $O_3$ and $SO_2$, respectively. $NO_2$ concentration was monitored using a Cavity Attenuated Phase Shift analyzer (CAPS, Aerodyne Research Inc.). Submicron particle number size distributions (PNSD) in the range 10-450 nm were measured every 2 minutes, using TSI-SMPS Model 3082, associated with a TSI-CPC Model 3772. Particle absorption coefficients were measured at 637 nm using a Thermo Scientific Model 5012 MAAP (Multi-Angle Absorption Photometer), compensated by a +5% factor and converted to equivalent black carbon (eBC) mass concentration assuming a mass absorption efficiency value of 6.6 $m^2$ $g^{-1}$ (Müller et al., 2011).

The chemical composition of non-refractory submicron aerosol was measured using a Quadrupole Aerosol Chemical Speciation Monitor (Q-ACSM, Aerodyne Research Inc.), described in detail by Ng et al. (2011a). Briefly, aerosol particles are sampled using a $PM_{2.5}$ cyclone inlet and focused through an aerodynamic lens forming a narrow particle beam, which is transmitted into the detection chamber under high vacuum onto a 600°C vaporizer. According to Ng et al. (2011a) the aerodynamic lens used in the Q-ACSM, has a high transmission efficiency in the aerodynamic diameter range 70–500 nm. Non-refractory material is flash-vaporized and ionized by electron impact at 70 eV. According to their mass-to-charge ratio (m/z), the fragments are analyzed in the quadrupole spectrometer. In this study, an averaging interval of 30 min was adopted. The collection efficiency (CE, defined as the ratio of the mass of particles detected by the instrument to the mass of particles introduced in the inlet) was calculated using the parametrization of Middlebrook et al. (2011) and resulted in an average of $0.45 \pm 0.01$. The transmission of the ACSM lens makes the instrument measure particles from 70 nm to about 900 nm, normally referred as $PM_1$. Detection limits of ACSM were calculated as three times the average noise level, using data of ambient air sampled through a HEPA filter. For 30 min of averaging time, the $3\sigma$ detection limits for chloride, ammonium, nitrate, organics, and sulfate are 0.03 $\mu$g m$^{-3}$, 0.36 $\mu$g m$^{-3}$, 0.03 $\mu$g m$^{-3}$, 0.57 $\mu$g m$^{-3}$, and 0.09 $\mu$g m$^{-3}$, respectively.

Ambient aerosols were sampled under dry conditions (RH<40%) using a nafion dryer, and concentrations were compensated for standard temperature and pressure conditions (1013.25 mbar; 273.15 K). Aerosol and trace gas measurements were averaged in periods of 30 min. Meteorological parameters (wind speed, wind direction, temperature, RH, solar radiation and precipitation) were provided by the Institute of Astronomy, Geophysics and Atmospheric Sciences of the University of São Paulo (IAG) meteorological station (23°39'02.61"S 46°37'18.55"W), 10 km North from the sampling site.

## 2.2 Identification of OA components with Positive Matrix Factorization (PMF)

Positive Matrix Factorization (PMF) is a statistical model that uses weighted least-square fitting for factor analysis (Paatero and Tapper, 1994; Paatero, 1997). It uses a bilinear factor analytic model defined in matrix notation as:

$$X = GF + E \tag{1}$$

where X denotes the matrix of the measured values, G and F are matrices computed by the model that represent the scores and loading, respectively, and E is the residual matrix, made up of the elements $e_{ij}$. For the ACSM and AMS (Aerosol Mass Spectrometer) data, the measured organic mass spectra are apportioned in terms of source/process-related components (Zhang et al., 2011). In this case, the columns j in X are the m/z's and each row i represents a single mass spectrum, G represents the time series and F the profile mass spectrum for the p factors computed by the algorithm. The model adjusts G and F using a least-squares algorithm that iteratively minimizes the quantity Q, defined as the sum of the squared residuals weighted by their respective uncertainties:

$$Q = \sum_{i=1}^{m} \sum_{j=1}^{n} \left( \frac{e_{ij}}{\sigma_{ij}} \right)^2 \tag{2}$$

where $\sigma_{ij}$ is the uncertainty for each element in the matrix X. An IGOR™-based source finder (Canonaco et al., 2013, Sofi) with a multilinear engine algorithm (Paatero, 1999, ME-2) was used to prepare the data, error estimates, execute the analysis and evaluate the results.

## 2.3 Multilinear regression (MLR) model

An analysis of the relationship between chemical composition and size distribution was performed using a multilinear regression (MLR) model. Previous studies have applied the MLR model to estimate aerosol mass scattering and extinction efficiencies, and source apportionment of optical properties (Ealo et al., 2018). A linear regression model describes the relationship between a dependent variable, y, and one or more independent variables, x. The dependent variable is also called the response variable and independent variables are also called explanatory or predictor variables. The MLR model is:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + ... + \beta_p x_{ip} + \epsilon_i, i = 1, ..., n \tag{3}$$

where, $y_i$ is the $i^{th}$ observation of the response variable, $\beta_j$ is the $j^{th}$ coefficient, $\beta_0$ is the constant term in the model, $x_{ij}$ is the $i^{th}$ observation on the $j^{th}$ predictor variable, j = 1, ..., p, and $\epsilon_i$ is the $i^{th}$ error term.

For the MLR model, the time series of $PM_1$ chemical constituents (i.e., eBC, inorganic species and OA PMF-derived chemical classes) were used as dependent variables, and the volume of particle size modes were taken as predictors. Initially, we use the volume of nucleation, Aitken and accumulation modes as independent variables in the model. However, the chemical mass

balance has been violated for nucleation mode, suggesting that more particles in the nucleation model are required by MLR than actual measurement. This effect is caused by the fact that the nucleation mode (local maximum at 22.5 nm) is out of the particle size range measured by the ACSM (75–650 nm). Taking this into account, we modified the model, using Aitken and accumulation modes as predictors. Since the nucleation mode represents only a small fraction of the aerosol volume (<0.5%), it is reasonable to explain the PM mass concentration based on these two modes. Volume size distribution was used in MLR calculations since it represents better the accumulation and Aitken modes than number size distribution.

For MLR analysis, intercept (constant term in the model) was set to 0, as it holds no physical significance for the mass attribution. Since the presence of multicollinearity (i.e. correlation between predictors) can adversely affect results in an MLR, the variance inflation factor (Hair et al., 2006, VIF) was determined. The VIF is a means of detecting multicollinearities between the independent variables of a model. If the VIF > 10, the multicollinearity is high and the variable should be excluded (Hair et al., 2006). A fit linear regression model was performed using fitlm function for Matlab 2015a by using standard least-squares fit.

## 3   Results and Discussion

### 3.1   Meteorology and trace gases

The campaign was carried out during the Southern Hemisphere spring/summer, with an average ($\pm$ standard deviation) air temperature of $23.0 \pm 3.5°C$, varying from 14 to 33°C. RH was $79.6 \pm 15.7\%$, on average (Table 1). The monthly accumulated precipitation ranged between 140 and 370 mm. The wind rose (Fig. 1) indicates dominant winds from the northeast, northwest and southeast along the measurement period.

Average mixing ratios of trace gases are shown in Table 1. The mid-day $O_3$ (averaged between 10h-18h LT) peak was 43.5 ppb, values comparable to the measurements at the Ibirapuera Park, one of the areas of MASP that often exceeds ozone air quality standards (Andrade et al., 2017). Diadema site is surrounded by green areas, an important contributor source of biogenic volatile organic compounds (BVOC), adding to VOC emissions from anthropogenic sources (Brito et al., 2015), in promoting the formation of ground-level ozone and SOA. In the MASP, the ozone levels are more sensitive to changes in VOC emissions rather than NOx (VOC-limited), being the only pollutant that shows an increasing trend for concentrations in the MASP over the years due to the large use of fossil fuel and biofuels (Andrade et al., 2017).

The average $NO_2$ concentration in this study was 13 ppb, significantly lower compared to São Paulo downtown (Andrade et al., 2017). According to emission inventories (CETESB, 2019), vehicles account for 64% of $NO_X$ emissions in the MASP, particularly diesel engines in HDVs (heavy-duty vehicles: mostly buses and trucks), which represent 44% of total $NO_X$ emissions. The average $SO_2$ concentration is 0.61 ppb, approximately half of the observed values in São Paulo downtown during springtime in 2013 (Monteiro dos Santos et al., 2016). Few episodes (15% of the dataset) presented $SO_2$ levels above 1 ppb. In the MASP, $SO_2$ emissions are related to industrial sources and the sulfur content in diesel and gasohol.

## 3.2 Aerosol chemical composition

Near real-time submicron mass concentration ($PM_1$) can be obtained by the non-refractory $PM_1$ (NR-$PM_1$) and eBC measurements (Table 2). Average $PM_1 \pm$ standard deviation during the campaign was $11.4 \pm 7.8\ \mu g\ m^{-3}$. These results are similar to observations in São Paulo downtown ($10.8\ \mu g\ m^{-3}$, Brito et al. (2018) during the spring), New York ($11.7\ \mu g\ m^{-3}$ Sun et al. (2011) during the summer), Barcelona ($18.5\ \mu g\ m^{-3}$, Mohr et al. (2012) during the winter) and Santiago ($18.1\ \mu g\ m^{-3}$, Carbone et al. (2013) during the spring). On average, organic aerosols dominated the composition, contributing to 55%, followed by sulfate (15%) and eBC (14%). Ammonium (9%), nitrate (6%) and chloride (1%) presented smaller contributions to the mass loading. The time series of $PM_1$ chemical species in the submicron size range is shown in Figure 2.

Components of OA were identified using PMF analysis following the procedure described by Ulbrich et al. (2009). The choice of the number of factors was based on the quality of the fit parameter, the correlation analysis with external tracers, spectral analysis and comparisons with AMS database mass spectra, and their diurnal variability. Moreover, the examination of rotational ambiguity was done by varying the FPEAK parameter. Solutions with more than four factors were examined and showed only the splitting behavior of existing factors, instead of providing new consistent factors. Details of the PMF analysis procedure are given in the Supplement Material. The identified four OA factors include: a highly oxidized component, linked to secondary processes, named oxygenated organic aerosol (OOA); a component containing typical biomass burning spectral signature, named biomass burning organic aerosol (BBOA); and two hydrocarbon-like OA components ($HOA_I$ and $HOA_{II}$), that represent primary emissions from anthropogenic sources, but with distinct mass spectra and diurnal variability patterns.

The OOA component is dominated by m/z 44, which is mainly the fragment $CO_2^+$ (Zhang et al., 2005), typical from thermal decarboxylation of organic acid groups, previously described as aged aerosol and related to the formation of SOA (Jimenez et al., 2009). The OOA factor is the dominant component, comprising, on average, 40% of OA total mass concentration. The mass spectrum of OOA (Figure 3) correlates strongly with the standard AMS LV-OOA mass spectra database (R=0.99, Table 3), indicating that the OOA factor is dominated by low volatile organic compounds, rather than semi-volatile organic compounds (R=0.72, Table 3). The OOA time series present moderate correlations with oxidant concentration ($O_X$=$NO_2$+$O_3$, Table 4), and secondary inorganic species, such as sulfate, nitrate and ammonium. Moreover, the OOA mass concentration significantly increases in the afternoon, similarly to the ozone diels (Figure 4), indicating that its formation is partially driven by photochemistry. Considering the sum of secondary inorganic aerosols (sulfate, nitrate and ammonium) and SOA (OOA) as a lower limit for the contribution of secondary aerosols to the total of $PM_1$, it is possible to estimate that at least 56% of submicron particles mass loading results from secondary production.

The BBOA component has a mass spectrum dominated by the m/z's 29, 60 and 73 (Figure 3). The signal at m/z 60 is associated with the $C_2H_4O^+$ ion (Alfarra et al., 2007) and correlates with levoglucosan and similar anhydrosugar species (mannosan, galactosan) that result from the pyrolysis of cellulose. The BBOA mass spectrum presents a strong correlation with standard AMS database BBOA (R=0.90, Table 3). The diurnal variability of BBOA (Figure 4), with an average concentration almost three times higher during nighttime, seems modulated by atmospheric dynamics, such as boundary layer height evolution. The boundary layer height decreases during nighttime trapping freshly emitted smoke particles. The time series of BBOA correlates

moderately with eBC (R=0.47), nitrate (R=0.48) and chloride (R=0.56) (Table 4). The average contribution of BBOA is 13% of total OA and almost 7% of $PM_1$, significantly lower than reported in Pereira et al. (2017), that found considerable biomass
180 burning contributions (approximately 18% of $PM_{2.5}$) associated to long-range transport from regional sugarcane burning in São Paulo during wintertime, in addition to local emission sources.

Both HOA components present mass spectra characterized by hydrocarbon-like structures typical of alkanes, alkenes and cycloalkanes (m/z's 27, 29, 41, 43, 55, 57, 67, 69, 71, 81, 83, 85) related to anthropogenic primary emissions (Canagaratna et al., 2004; Zhang et al., 2005). Both mass spectra correlate with the AMS standard HOA mass spectrum (R=0.90 and R=0.89
185 for $HOA_I$ and $HOA_{II}$, respectively). Although both factors are HOA related, it is not reasonable to interpret them as a split of the same source. The $HOA_I$ factor presents an elevated signal at m/z 55 that has been related to cooking OA (COA), an important source of (primary organic aerosol) POA in urban environments (Mohr et al., 2012). The $HOA_{II}$ factor presents a higher signal at m/z 57 than at m/z 55, and higher correlation with eBC ($R_{HOA_{II}}$=0.69, $R_{HOA_I}$=0.45), which is related to vehicular emissions in the MASP, mostly heavy-duty vehicles (de Miranda et al., 2018). The results indicate that $HOA_{II}$ is
190 more consistent with traffic, while the $HOA_I$ seems like a mixture of anthropogenic sources. Together, the HOA factors present an average contribution of 37% to OA (21% from $HOA_I$ and 16% from $HOA_{II}$). For both HOA factors, the diurnal profiles of mass concentrations (Figure 4) increase during the traffic rush-hour time 6h-8h (local time). However, $HOA_I$ also shows a peak between 12h and 14h (local time), probably associated with local cooking activities.

Although the sampling site is located in an industrialized region, a distinct industrial-related OA factor could not be identified
195 in this study. As a matter of comparison, Bozzetti et al. (2017) identified OA related to industrial emissions in an urban background site in Marseille under influence of industries and petrochemical companies. Some similarities can be observed between the mass spectrum of $HOA_I$ in this study and the mass spectrum of industrial-related OA (INDOA) that they found, particularly substantial signals at m/z 27, m/z 29, m/z 43 and m/z 55. This fact can indicate the influence of industrial emissions in $HOA_I$, however, further investigation is necessary. Furthermore, a significant part of ammonium sulfate could be produced
200 from the oxidation of anthropogenic $SO_2$ from industrial sources.

### 3.3 Aerosol size modes

An algorithm in Matlab was developed to fit one to three lognormal modes to the measured particle number size distributions, based on Hussein et al. (2005), using the standard method of least-squares. The center of each mode is allowed to vary from one size distribution to another, constrained to the following diameter ranges: nucleation mode mean geometric diameter
205 (Dpg) must be smaller than 30 nm; Aitken mode Dpg must be ≥ 30 nm and < 90nm; accumulation mode Dpg must be ≥90 nm. There is also a restriction in the geometric standard deviation, which must be within 1.2 and 2.1 for all modes. The algorithm automatically decides between 2 or 3 mode fitting based on the percentual concentration of 10-40 nm particles, and based on goodness of fit proxies like the root mean square error and the ratio between fit and data total particle number concentration. The algorithm is able to automatically reduce the number of fitted modes (from 3 to 2 modes and from 2 to 1
210 mode) if a complete superposition is detected, based on the ratio between the modes Dpg. The algorithm uses the previous fitting parameters as a first guess to the current fitting, so that continuity is favored.

Considering the average PNSD, nucleation, Aitken and accumulation modes were centered at mobility diameters of 22 nm, 50 nm and 122 nm, respectively (Figure 5). For nucleation and Aitken modes these results are very similar to previous observations in São Paulo region during spring and early summer (Backman et al., 2012), however, the diameter for accumulation mode is significantly lower than the reported in such study (>200 nm). The sum of the number concentration of nucleation, Aitken and accumulation modes explains the measured total particle number concentration (slope=0.98 and $R^2$=0.99). The Aitken mode dominated the PNSD with average concentration $\pm$ standard deviation of $6900 \pm 4600 \, \text{cm}^{-3}$ (56% of total number concentration) followed by the nucleation mode, with average particle number concentration of $2800 \pm 2100 \, \text{cm}^{-3}$. The contribution of the accumulation mode is the lowest in terms of particle number (19% of total number concentration), but the highest in terms of particle volume concentration. The nucleation mode presented a peak concentration in the morning rush-hour, similar to eBC, $HOA_{II}$ and $NO_2$ (Fig. 4). Backman et al. (2012) identified similar diurnal cycles for nucleation-mode particles and attributed it to vehicular emissions and new particle formation (NPF) events. The nucleation mode peak in the morning is consistent with a simultaneous drop in the particle mean geometric diameter. Conversely, in the afternoon a strong increase of accumulation mode particles was observed, simultaneously with the larger presence of secondary species such as OOA, sulfate and $O_3$. Similarly, Aitken mode particles increase during the afternoon. According to Backman et al. (2012) the shift from nucleation to Aitken mode can be attributed to the growth of pre-existing nucleation-mode particles into the Aitken regime.

Figure 6 shows median PNSD for nucleation, Aitken and accumulation modes under low (below $25^{th}$ percentile) and high (above $75^{th}$ percentile) $PM_1$ concentrations. From this comparison, total number concentrations increase from $7000 \, \text{cm}^{-3}$ (for $PM_1 < 5.6 \mu\text{g m}^{-3}$) to $18400 \, \text{cm}^{-3}$ (for $PM_1 > 13.7 \mu\text{g m}^{-3}$). Interestingly, the enhancement differs notably between aerosol size modes. The accumulation mode shows the largest increase from low $PM_1$ condition ($1200 \, \text{cm}^{-3}$) to high $PM_1$ condition ($5200 \, \text{cm}^{-3}$). On the other hand, the nucleation mode shows a smaller increase from low $PM_1$ to high $PM_1$ scenario, and therefore the contribution of the nucleation mode particles to the total number concentration is higher under low $PM_1$ (30%) than under high $PM_1$ (16%). Local traffic emissions likely remain at similar levels during both low $PM_1$ and high $PM_1$ episodes. Considering that nucleation mode particles are mainly traffic related (Fig. 4), its relative contribution to the total number concentration tends to increase during low $PM_1$ episodes. The results are similar to Martins et al. (2010), where aerosol size distributions were measured during a transition period between a highly polluted episode and a clean one in São Paulo downtown. The authors found particles distributed in the nucleation and Aitken mode during the clean period, and larger geometric mean diameters during polluted periods, with particles partitioned between Aitken and accumulation modes. At the MASP, the meteorological conditions that favor the occurrence of high PM concentrations are typically a low boundary layer with a low inversion layer, weak ventilation, absence of precipitation and clear sky (Sánchez-Ccoyllo and Andrade, 2002; Santos et al., 2018), favoring aerosol secondary production and particle size increase by condensation. Similarly, Rodríguez et al. (2007) attributed high ultrafine particle events to low $PM_{2.5}$ conditions in Milan, Barcelona and London. The authors also found that high $PM_{2.5}$ pollution events tend to occur when condensation processes produce significant number concentrations of accumulation mode particles.

The strong occurrence of accumulation mode particles under polluted conditions can be explained by the fact that larger surface area of pre-existing particles favors coagulation processes. Consequently, nucleation is suppressed by coagulation loss and particles become larger. Moreover, the submicron aerosol size distribution is strongly influenced by the competition between nucleation of new particles and condensation of gas-phase components onto pre-existing particles Rodríguez et al. (2005). Under polluted conditions, the aerosol surface is enough to favor the condensation of vapors onto pre-existing particles, inhibiting nucleation, and resulting in particle growth. During low $PM_1$ conditions, the available aerosol surface is low, decreasing both condensation and coagulation rates, which favors homogeneous nucleation.

## 3.4 Relationships between particle size and chemical composition of submicron particles

The contribution of aerosol size modes to the ambient concentrations of the $PM_1$ chemical species was assessed by performing a multilinear regression model (MLR). In the MLR model, the time series of volume concentration at Aitken and accumulation modes were used as predictors. $PM_1$ components were used as species of interest. Results of MLR are summarized in Table 5. The model explained more than 90% of the average measured concentrations for the $PM_1$ species. For predictors used in MLR, the calculated variance inflation factor (VIF) was in the range of 1.11 to 2.16. In general, VIFs below 10 indicate no collinearity (Hair et al., 2006), ensuring the reliability of the regression results.

Contribution of aerosol size modes to mass concentrations of $PM_1$ chemical species (Figure 7) was obtained by multiplying regression coefficients (Table 5) and average volume concentrations. Their confidence intervals were calculated according to the confidence intervals of regression coefficients. $PM_1$ mass loadings were reconstructed by the sum of the partial contributions determined for each size mode (Fig. 8), i.e. Aitken (1.05 $\mu g\ m^{-3}$) and accumulation modes (7.23 $\mu g\ m^{-3}$). It explain 80.4% of the mean of measured $PM_1$.

Secondary inorganic species (ammonium, nitrate and sulfate) are partitioned between Aitken and accumulation modes. Similarly, Rodríguez et al. (2007) observed strong correlation between ammonium nitrate and ammonium sulfate with the accumulation mode, attributing it to condensation mechanisms and particle growth processes. A large fraction of inorganic species are in the Aitken mode, and it looks like the most acidic size mode. Carbone et al. (2013) suggested that the presence of nitrate and ammonium in the Aitken mode is likely to result from the reaction between nitric acid ($HNO_3$) and ammonia ($NH_3$) from traffic emissions. Moreover, Backman et al. (2012) associated the growth of pre-existing nucleation mode aerosols to particle coating by sulfates and inorganic nitrates.

The BBOA presents only a small fraction in Aitken mode, and the highest loading attributed to the accumulation mode. This result indicates that this urban site can be influenced by regional biomass burning emissions. The enhancement of accumulation mode particles under the influence of regional biomass burning emissions has been observed elsewhere, such as in Amazon. Kumar et al. (2016) discussed the relevance of unregulated PM sources in MASP, including wood burning in pizzerias that could emit 321 kg day$^{-1}$ of $PM_{2.5}$, according to their estimates.

Submicron mass loading in accumulation mode has a great contribution of OOA (32%), probably resulting from SOA formation and aging. Accumulation mode also presents significant contributions of HOA factors (24%) and BC (21%). Similarly, Costabile et al. (2009) related the accumulation mode to particles containing aged secondary aerosol and also direct anthro-

280 pogenic emissions. Interestingly, most of eBC concentration is in the accumulation mode and only a small fraction of mass is in Aitken mode. The same behaviour is observed for traffic related HOA factors. It can indicate that a substantial fraction of the BC is heavily coated by organics from regional transport, in addition to local vehicular fresh emissions, however it needs further investigations. A suggestion for future studies is the analysis of BC mixing state and BC coating thickness using a single-particle soot photometer (SP2).

## 285  4  Summary and conclusions

Physicochemical properties of aerosols are key parameters in terms of their impacts on human health and climate effects. In this study, a detailed characterization of submicron particles was performed at an urban site in the MASP. The results show $PM_1$ mass concentrations in close agreement with other megacities, with an average mass concentration of 11.4 $\mu$g m$^{-3}$. As expected, chemical composition was dominated by organic aerosols (56%), with significant contributions of sulfate

290  (15%) and black carbon (13%). Using PMF analysis it was possible to identify four OA classes including oxygenated organic aerosol (OOA), biomass burning organic aerosol (BBOA), and two hydrocarbon-like OA components (a typical HOA related to vehicular emissions, and a HOA associated to a mix of anthropogenic sources). Considering the sum of secondary inorganic aerosols and SOA as a lower limit, more than 50% of $PM_1$ mass loading was estimated as resulting from secondary production.

Nucleation, Aitken and accumulation lognormal size modes were fitted to the measured PNSD. Aitken mode dominated

295  the total number concentration with an average concentration of 6,900 cm$^{-3}$ and submicron aerosol size distribution was strongly influenced by the $PM_1$ levels. The accumulation mode shows a large increase from low PM1 conditions to high $PM_1$ conditions. The shift in the particle size distribution to larger sizes provides more aerosol surface, which can favor the condensation of vapors onto pre-existing particles and likely inhibit nucleation. It corroborates with lower contribution of particles at nucleation mode observed under high $PM_1$ conditions. Conversely, the contribution of particles from the nucleation

300  mode to the total number concentration is higher during low $PM_1$ conditions, when the available aerosol surface is low, decreasing both condensation and coagulation rates, and favoring homogeneous nucleation. Because of the high contribution of nucleation particles under low $PM_1$ loadings, $PM_{2.5}$ and $PM_{10}$ (parameters frequently used in the air quality index) may be insufficient to assess human PM exposure in urban areas.

The relationships between size modes and chemical constituents of $PM_1$ were assessed by performing an MLR model. Sec-

305  ondary inorganic species (ammonium, sulfate and nitrate) were partitioned between Aitken and accumulation modes, related to condensation particle growth processes. Submicron mass loading in the accumulation mode included aged secondary organic aerosol and vehicular emissions.

The results presented here emphasize the well-established impact of traffic-related sources in the MASP and make clear the need to reduce emissions rates in the region by applying new technologies such as the EURO VI emission standard. It is

310  also essential to expand mass transportation systems, since the metro system in São Paulo is heavily underdeveloped, resulting in a better transportation system for 20 million people. Additionally, encouraging alternative transportation, implementing strong incentives for electrical vehicles and the restriction of passenger car circulation can improve significantly air quality

in urban areas. Although the implementation of regulatory programs to control stationary and mobile sources in the MASP over the last decades has been successful to reduce primary emissions, secondary processes have been recognized as critical to air quality in the region. The findings presented provide innovative insights on the association between sources and processes governing physicochemical properties of atmospheric aerosol and highlight the key role of SOA formation on the $PM_1$ ambient concentrations in a megacity largely impacted by traffic emissions and extensive biofuel usage.

**11**

# References

Alfarra, M. R., Prevot, A. S. H., Szidat, S., Sandradewi, J., Weimer, S., Lanz, V. A., Schreiber, D., Mohr, M., and Baltensperger, U.: Identification of the Mass Spectral Signature of Organic Aerosols from Wood Burning Emissions, Environmental Science & Technology, 41, 5770–5777, https://doi.org/10.1021/es062289b, pMID: 17874785, 2007.

Andrade, M. F., Miranda, R. M., Fornaro, A., Kerr, A., Oyama, B., de Andre, P. A., and Saldiva, P.: Vehicle emissions and PM 2.5 mass concentrations in six Brazilian cities, Air Quality, Atmosphere & Health, 5, 79–88, https://doi.org/https://doi.org/10.1007/s11869-010-0104-5, 2012.

Andrade, M. F., Kumar, P., de Freitas, E. D., Ynoue, R. Y., Martins, J., Martins, L. D., Nogueira, T., Perez-Martinez, P., de Miranda, R. M., Albuquerque, T., et al.: Air quality in the megacity of São Paulo: Evolution over the last 30 years and future perspectives, Atmospheric Environment, 159, 66–82, https://doi.org/10.1016/j.atmosenv.2017.03.051, 2017.

Backman, J., Rizzo, L. V., Hakala, J., Nieminen, T., Manninen, H. E., Morais, F., Aalto, P. P., Siivola, E., Carbone, S., Hillamo, R., Artaxo, P., Virkkula, A., Petäjä, T., and Kulmala, M.: On the diurnal cycle of urban aerosols, black carbon and the occurrence of new particle formation events in springtime SÃ£o Paulo, Brazil, Atmospheric Chemistry and Physics, 12, 11 733–11 751, https://doi.org/10.5194/acp-12-11733-2012, https://www.atmos-chem-phys.net/12/11733/2012/, 2012.

Beirle, S., Boersma, K. F., Platt, U., Lawrence, M. G., and Wagner, T.: Megacity emissions and lifetimes of nitrogen oxides probed from space, Science, 333, 1737–1739, https://doi.org/10.1126/science.1207824, 2011.

Boucher, O., Randall, D., Artaxo, P., Bretherton, C., Feingold, G., Forster, P., Kerminen, V.-M., Kondo, Y., Liao, H., Lohmann, U., et al.: Clouds and aerosols, in: Climate change 2013: the physical science basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, pp. 571–657, Cambridge University Press, https://doi.org/doi:10.1017/CBO9781107415324.016, 2013.

Bozzetti, C., El Haddad, I., Salameh, D., Daellenbach, K. R., Fermo, P., Gonzalez, R., Minguillón, M. C., Iinuma, Y., Poulain, L., Elser, M., Müller, E., Slowik, J. G., Jaffrezo, J.-L., Baltensperger, U., Marchand, N., and Prévôt, A. S. H.: Organic aerosol source apportionment by offline-AMS over a full year in Marseille, Atmospheric Chemistry and Physics, 17, 8247–8268, https://doi.org/10.5194/acp-17-8247-2017, https://www.atmos-chem-phys.net/17/8247/2017/, 2017.

Brito, J., Wurm, F., Yaannez-Serrano, A. M., de Assuncao, J. V., Godoy, J. M., and Artaxo, P.: Vehicular emission ratios of VOCs in a megacity impacted by extensive ethanol use: results of ambient measurements in Sao Paulo, Brazil, Environmental science & technology, 49, 11 381–11 387, https://doi.org/doi: 10.1021/acs.est.5b03281, 2015.

Brito, J., Carbone, S., dos Santos, D. A. M., Dominutti, P., de Oliveira Alves, N., Rizzo, L. V., and Artaxo, P.: Disentangling vehicular emission impact on urban air pollution using ethanol as a tracer, Scientific reports, 8, 10 679, https://doi.org/10.1016/S0140-6736(17)30505-6, 2018.

Canagaratna, M. R., Jayne, J. T., Ghertner, D. A., Herndon, S., Shi, Q., Jimenez, J. L., Silva, P. J., Williams, P., Lanni, T., Drewnick, F., Demerjian, K. L., Kolb, C. E., and Worsnop, D. R.: Chase Studies of Particulate Emissions from in-use New York City Vehicles, Aerosol Science and Technology, 38, 555–573, https://doi.org/10.1080/02786820490465504, 2004.

Canonaco, F., Crippa, M., Slowik, J. G., Baltensperger, U., and Prévôt, A. S. H.: SoFi, an IGOR-based interface for the efficient use of the generalized multilinear engine (ME-2) for the source apportionment: ME-2 application to aerosol mass spectrometer data, Atmospheric Measurement Techniques, 6, 3649–3661, https://doi.org/10.5194/amt-6-3649-2013, https://www.atmos-meas-tech.net/6/3649/2013/, 2013.

Carbone, S., Saarikoski, S., Frey, A., Reyes, F., Reyes, P., Castillo, M., Gramsch, E., Oyola, P., Jayne, J., Worsnop, D. R., et al.: Chemical characterization of submicron aerosol particles in Santiago de Chile, Aerosol Air Qual. Res, 13, 462–473, https://doi.org/10.4209/aaqr.2012.10.0261, 2013.

Castanho, A. and Artaxo, P.: Wintertime and summertime São Paulo aerosol source apportionment study, Atmospheric Environment, 35, 4889–4902, https://doi.org/10.1016/S1352-2310(01)00357-0, 2001.

CETESB: Relatório de Qualidade do Ar no Estado de São Paulo 2018, Série Relatórios, São Paulo, http://https://cetesb.sp.gov.br/ar/publicacoes-relatorios/, 2019.

Che, H., Zhang, X., Wang, Y., Zhang, L., Shen, X., Zhang, Y., Ma, Q., Sun, J., Zhang, Y., and Wang, T.: Characterization and parameterization of aerosol cloud condensation nuclei activation under different pollution conditions, Scientific Reports, 6, 24 497, https://doi.org/10.1038/srep24497, 2016.

Ching, J. and Kajino, M.: Aerosol mixing state matters for particles deposition in human respiratory system, vol. 8, https://doi.org/10.1038/s41598-018-27156-z, 2018.

Cohen, A. J., Brauer, M., Burnett, R., Anderson, H. R., Frostad, J., Estep, K., Balakrishnan, K., Brunekreef, B., Dandona, L., Dandona, R., et al.: Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the Global Burden of Diseases Study 2015, The Lancet, 389, 1907–1918, https://doi.org/https://doi.org/10.1016/S0140-6736(17)30505-6, 2017.

Costabile, F., Birmili, W., S, K., Tuch, T., Wehner, B., Wiedensohler, A., Franck, U., Koenig, K., and Sonntag, A.: Spatio-temporal variability and principal components of the particle number size distribution in an urban atmosphere, Atmos. Chem. Phys., https://doi.org/10.5194/acp-9-3163-2009, 2009.

de Miranda, R., Andrade, M., F.N.D, R., Francisco, K., and Perez-Martinez, P.: Source apportionment of fine particulate matter by positive matrix factorization in the Metropolitan Area of São Paulo, Brazil, Journal of Cleaner Production, 202, https://doi.org/10.1016/j.jclepro.2018.08.100, 2018.

Ealo, M., Alastuey, A., Pérez, N., Ripoll, A., Querol, X., and Pandolfi, M.: Impact of aerosol particle sources on optical properties in urban, regional and remote areas in the north-western Mediterranean, Atmospheric Chemistry and Physics, 18, 1149–1169, https://doi.org/10.5194/acp-18-1149-2018, https://www.atmos-chem-phys.net/18/1149/2018/, 2018.

Fuzzi, S., Baltensperger, U., Carslaw, K., Decesari, S., Denier van der Gon, H., Facchini, M. C., Fowler, D., Koren, I., Langford, B., Lohmann, U., Nemitz, E., Pandis, S., Riipinen, I., Rudich, Y., Schaap, M., Slowik, J. G., Spracklen, D. V., Vignati, E., Wild, M., Williams, M., and Gilardoni, S.: Particulate matter, air quality and climate: lessons learned and future needs, Atmospheric Chemistry and Physics, 15, 8217–8299, https://doi.org/10.5194/acp-15-8217-2015, https://www.atmos-chem-phys.net/15/8217/2015/, 2015.

Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., Tatham, R. L., et al.: Multivariate data analysis (Vol. 6), Upper Saddle River, NJ: Pearson Prentice Hall, 2006.

Hussein, T., Dal Maso, M., Petaja, T., Koponen, I. K., Paatero, P., Aalto, P. P., Hameri, K., and Kulmala, M.: Evaluation of an automatic algorithm for fitting the particle number size distributions, Boreal Environ. Res., 10, 337–355, 2005.

IBGE: Instituto Brasileiro de Geografia e Estatística (Brazilian Institute of Geography and Statistics), http://cod.ibge.gov.br/QHF, accessed: 2020.04.10, 2020.

Jimenez, J. L., Canagaratna, M., Donahue, N., Prevot, A., Zhang, Q., Kroll, J. H., DeCarlo, P. F., Allan, J. D., Coe, H., Ng, N., et al.: Evolution of organic aerosols in the atmosphere, Science, 326, 1525–1529, https://doi.org/10.1126/science.1180353, 2009.

Kumar, P., de Fatima Andrade, M., Ynoue, R. Y., Fornaro, A., Dias de Freitas, E., Martins, J., Martins, L. D., Albuquerque, T., Zhang, Y., and Morawska, L.: New directions: from biofuels to wood stoves: the modern and ancient air quality challenges in the megacity of São Paulo, Atmospheric Environment, 140, 364–369, https://doi.org/https://doi.org/10.1016/j.atmosenv.2016.05.059, 2016.

Martins, L. D., Martins, J. A., Freitas, E. D., Mazzoli, C. R., Gonçalves, F. L. T., Ynoue, R. Y., Hallak, R., Albuquerque, T. T. A., and de Fatima Andrade, M.: Potential health impact of ultrafine particles under clean and polluted urban atmospheric conditions: a model-based study, Air Quality, Atmosphere & Health, 3, 29–39, https://doi.org/https://doi.org/10.1007/s11869-009-0048-9, 2010.

Middlebrook, A., Bahreini, R., Jimenez, J., and Cana-Garatna, M.: Evaluation of Composition-Dependent Collection Efficiencies for the Aerodyne Aerosol Mass Spectrometer using Field Data, Aerosol Sci. Technol., 46, 258–271, https://doi.org/https://doi.org/10.1080/02786826.2011.620041, 2011.

Mohr, C., DeCarlo, P. F., Heringa, M. F., Chirico, R., Slowik, J. G., Richter, R., Reche, C., Alastuey, A., Querol, X., Seco, R., Peñuelas, J., Jiménez, J. L., Crippa, M., Zimmermann, R., Baltensperger, U., and Prévôt, A. S. H.: Identification and quantification of organic aerosol from cooking and other sources in Barcelona using aerosol mass spectrometer data, Atmospheric Chemistry and Physics, 12, 1649–1665, https://doi.org/10.5194/acp-12-1649-2012, https://www.atmos-chem-phys.net/12/1649/2012/, 2012.

Monteiro dos Santos, D. A., Brito, J., Godoy, J., and Artaxo, P.: Ambient concentrations and insights on organic and elemental carbon dynamics in São Paulo, Brazil, Atmospheric Environment, 144, 226 – 233, https://doi.org/https://doi.org/10.1016/j.atmosenv.2016.08.081, 2016.

Müller, T., Henzing, J., Leeuw, G. d., Wiedensohler, A., Alastuey, A., Angelov, H., Bizjak, M., Collaud Coen, M., Engström, J., Gruening, C., et al.: Characterization and intercomparison of aerosol absorption photometers: result of two intercomparison workshops, Atmospheric Measurement Techniques, 4, 245–268, https://doi.org/https://doi.org/10.5194/amt-4-245-2011, 2011.

Ng, N., Herndon, S., Trimborn, A., Croteau, P., Onasch, T., Sueper, D., Worsnop, D., Zhang, Q., Sun, Y., and Jayne, J.: An Aerosol Chemical Speciation Monitor (ACSM) for Routine Monitoring of the Composition and Mass Concentrations of Ambient Aerosol, Aerosol Science and Technology - AEROSOL SCI TECH, 45, 780–794, https://doi.org/10.1080/02786826.2011.560211, 2011a.

Ng, N. L., Canagaratna, M. R., Jimenez, J. L., Chhabra, P. S., Seinfeld, J. H., and Worsnop, D. R.: Changes in organic aerosol composition with aging inferred from aerosol mass spectra, Atmospheric Chemistry and Physics, 11, 6465–6474, https://doi.org/10.5194/acp-11-6465-2011, https://www.atmos-chem-phys.net/11/6465/2011/, 2011b.

Paatero, P.: Least squares formulation of robust non-negative factor analysis, Chemometrics and intelligent laboratory systems, 37, 23–35, 1997.

Paatero, P.: The Multilinear Engine—A Table-Driven, Least Squares Program for Solving Multilinear Problems, Including the n-Way Parallel Factor Analysis Model, Journal of Computational and Graphical Statistics, 8, 854–888, https://doi.org/10.1080/10618600.1999.10474853, 1999.

Paatero, P. and Tapper, U.: Positive matrix factorization: A non-negative factor model with optimal utilization of error estimates of data values, Environmetrics, 5, 111–126, https://doi.org/https://doi.org/10.1002/env.3170050203, 1994.

Pereira, G. M., Teinilä, K., Custódio, D., Gomes Santos, A., Xian, H., Hillamo, R., Alves, C. A., Bittencourt de Andrade, J., Olímpio da Rocha, G., Kumar, P., Balasubramanian, R., Andrade, M. D. F., and de Castro Vasconcellos, P.: Particulate pollutants in the Brazilian city of São Paulo: 1-year investigation for the chemical composition and source apportionment, Atmospheric Chemistry and Physics, 17, 11 943–11 969, https://doi.org/10.5194/acp-17-11943-2017, https://www.atmos-chem-phys.net/17/11943/2017/, 2017.

Rodríguez, S., Van Dingenen, R., Putaud, J.-P., Martins-Dos Santos, S., and Roselli, D.: Nucleation and growth of new particles in the rural atmosphere of Northern Italy—relationship to air quality monitoring, Atmospheric environment, 39, 6734–6746, https://doi.org/10.1016/j.atmosenv.2005.07.036, 2005.

Rodríguez, S., Van Dingenen, R., Putaud, J.-P., Dell'Acqua, A., Pey, J., Querol, X., Alastuey, A., Chenery, S., Ho, K.-F., Harrison, R., Tardivo, R., Scarnato, B., and Gemelli, V.: A study on the relationship between mass concentrations, chemistry and number size distribution of urban fine aerosols in Milan, Barcelona and London, Atmospheric Chemistry and Physics, 7, 2217–2232, https://doi.org/10.5194/acp-7-2217-2007, 2007.

Romano, S., Perrone, M., Pavese, G., Esposito, F., and Calvello, M.: Optical properties of PM2.5 particles: Results from a monitoring campaign in southeastern Italy, Atmospheric Environment, 203, https://doi.org/10.1016/j.atmosenv.2019.01.037, 2019.

Santos, T. C. d., Reboita, M. S., and Carvalho, V. S. B.: Investigation of the Relationship Between Atmospheric Variables and the Concentration of MP10 and O3 in the State of São Paulo, Revista Brasileira de Meteorologia, 33, 631–645, https://doi.org/10.1590/0102-7786334006, 2018.

Sun, Y.-L., Zhang, Q., Schwab, J. J., Demerjian, K. L., Chen, W.-N., Bae, M.-S., Hung, H.-M., Hogrefe, O., Frank, B., Rattigan, O. V., and Lin, Y.-C.: Characterization of the sources and processes of organic and inorganic aerosols in New York city with a high-resolution time-of-flight aerosol mass apectrometer, Atmospheric Chemistry and Physics, 11, 1581–1602, https://doi.org/10.5194/acp-11-1581-2011, https://www.atmos-chem-phys.net/11/1581/2011/, 2011.

Sánchez-Ccoyllo, O. and Andrade, M. F.: The influence of meteorological conditions on the behavior of pollutants concentrations in São Paulo, Brazil, Environmental pollution (Barking, Essex : 1987), 116, 257–63, https://doi.org/10.1016/S0269-7491(01)00129-4, 2002.

Ulbrich, I., Canagaratna, M., Zhang, Q., Worsnop, D., and Jimenez, J.: Interpretation of organic components from Positive Matrix Factorization of aerosol mass spectrometric data, Atmospheric Chemistry and Physics, 9, 2891–2918, https://doi.org/https://doi.org/10.5194/acp-9-2891-2009, 2009.

Zhang, Q., Alfarra, M. R., Worsnop, D. R., Allan, J. D., Coe, H., Canagaratna, M. R., and Jimenez, J. L.: Deconvolution and Quantification of Hydrocarbon-like and Oxygenated Organic Aerosols Based on Aerosol Mass Spectrometry, Environmental Science & Technology, 39, 4938–4952, https://doi.org/10.1021/es048568l, pMID: 16053095, 2005.

Zhang, Q., Jimenez, J. L., Canagaratna, M. R., Ulbrich, I. M., Ng, N. L., Worsnop, D. R., and Sun, Y.: Understanding atmospheric organic aerosols via factor analysis of aerosol mass spectrometry: a review, Analytical and bioanalytical chemistry, 401, 3045–3067, https://doi.org/10.1007/s00216-011-5355-y, 2011.

Zhu, T., Melamed, M., Parrish, D., Gauss, M., Klenner, L. G., Lawrence, M., Konare, A., Liousse, C., et al.: WMO/IGAC impacts of megacities on air pollution and climate, World Meteorological Organization, 205, 149–151, 2012.
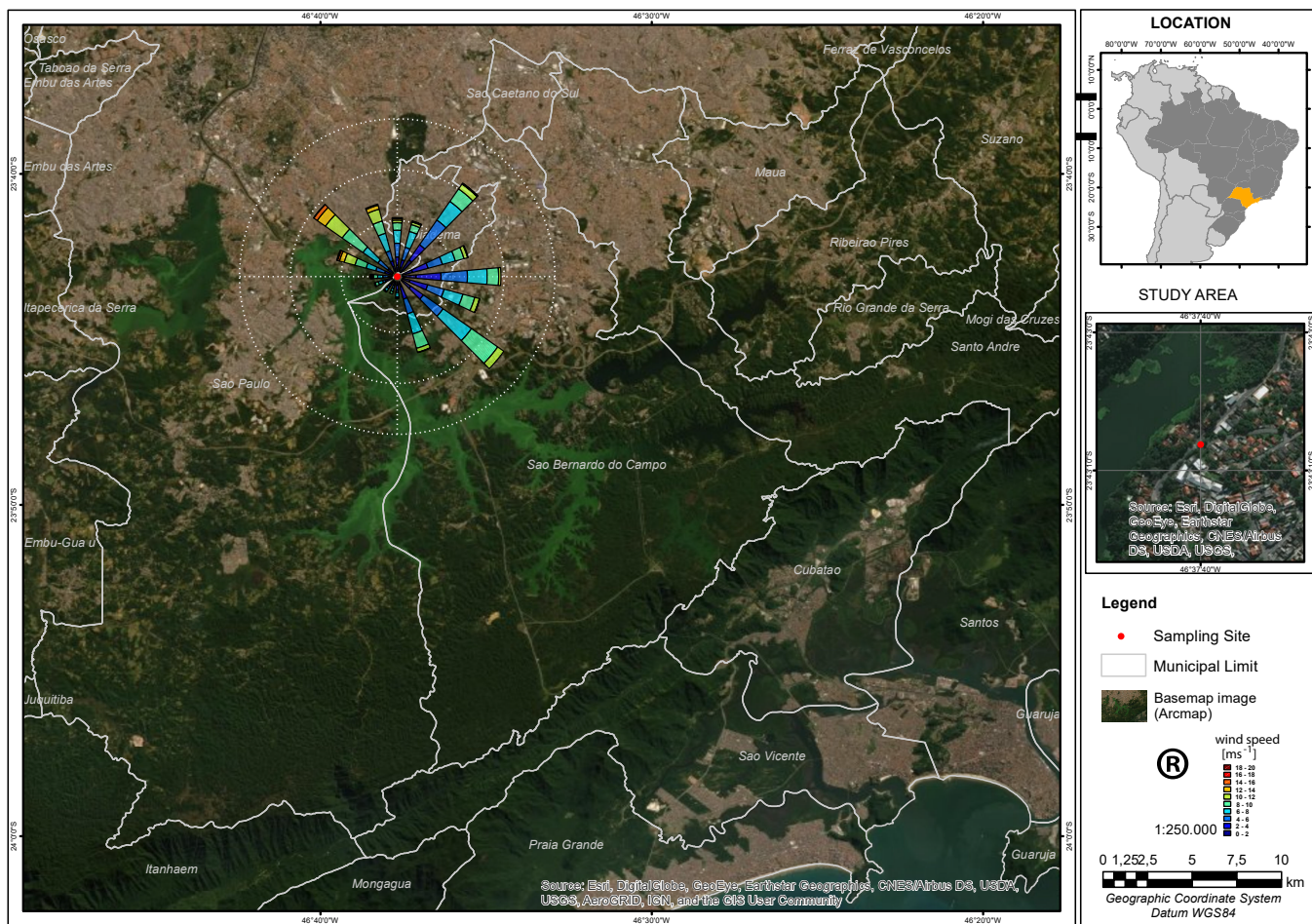
**Figure 1.** Map of Brazil indicating São Paulo state and detailed view of the sampling site (red point), located in the southeastern part of the MASP in Diadema city. Source: Esri, DigitalGlobe, GeoEye, Earthstar Geographics, CNES/Airbus DS, USDA, USGS, AeroGRID, IGN, and the GIS User Community. Wind rose during the campaign was calculated from data from Institute of Astronomy, Geophysics and Atmospheric Sciences of University of São Paulo (IAG) meteorological station.
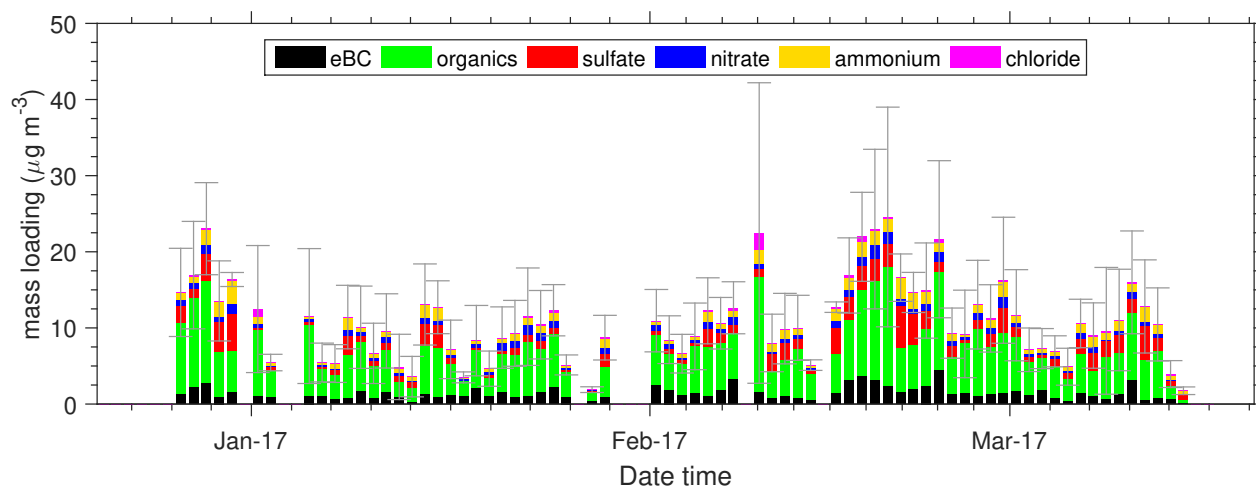
**Figure 2.** Time series of mass concentration of $PM_1$ chemical species. Each bin represents 1 day-average average of measured data. The error bars represent the standard deviation of the sum of all components.
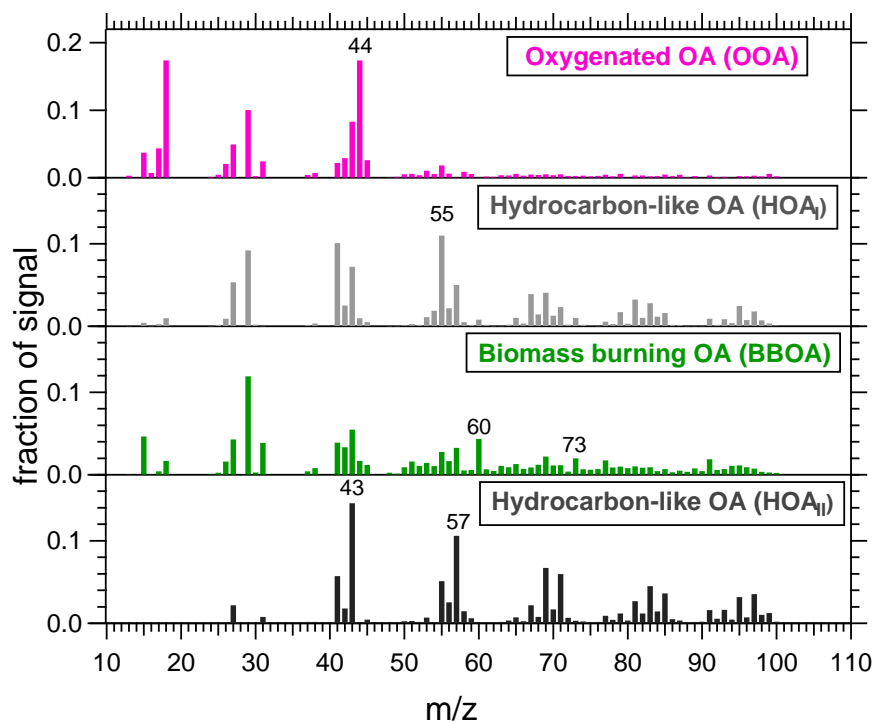
**Figure 3.** PMF mass spectra of the four PMF solutions containing oxygenated organic aerosol (OOA), biomass burning organic aerosols (BBOA) and two hydrocarbon-like organic aerosols (HOA$_I$ and HOA$_{II}$).
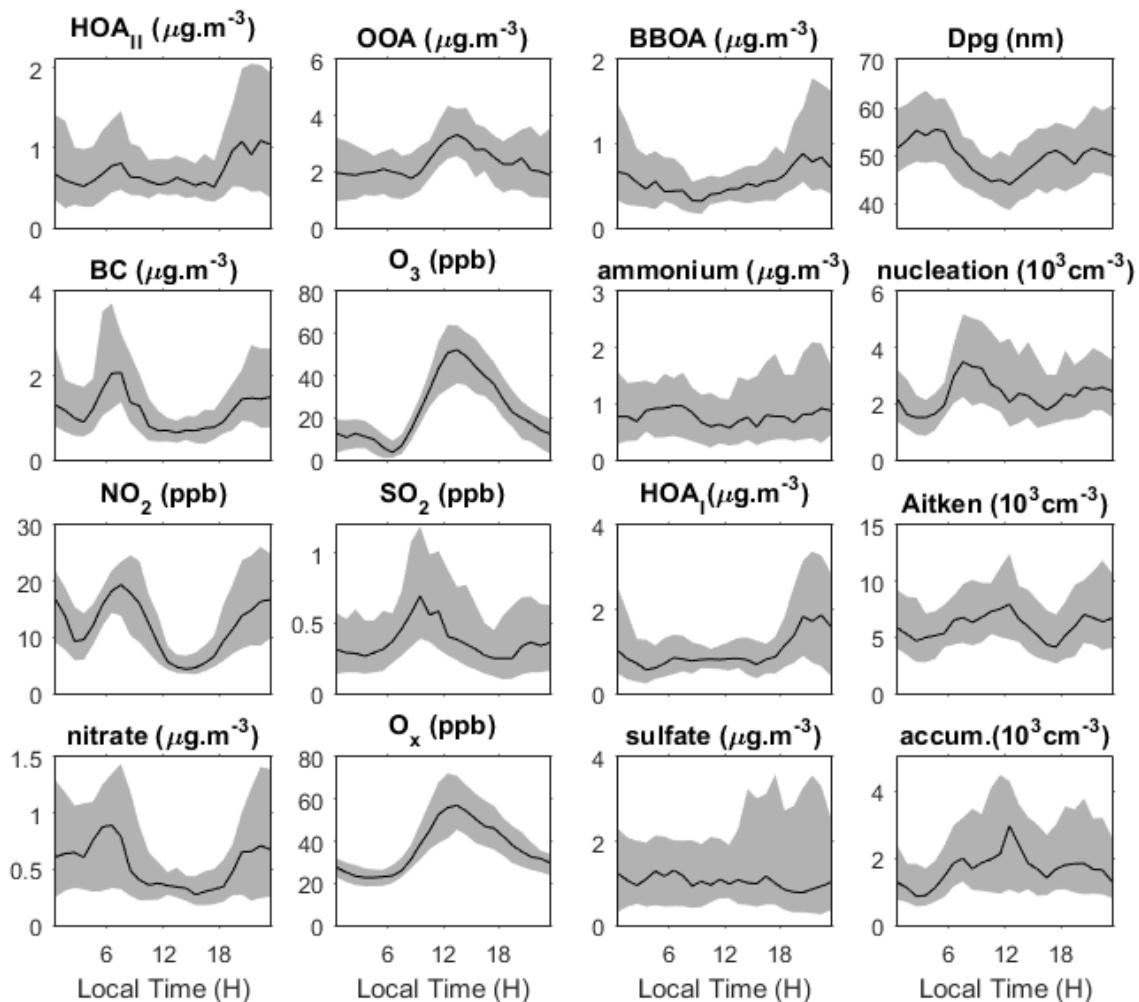
**Figure 4.** Diurnal trends for OA PMF factors (OOA, HOA$_I$, BBOA and HOA$_{II}$), black carbon, sulfate, nitrate, ammonium trace gases (NO$_2$, SO$_2$ and O$_3$), geometric mean diameter and submicrometer aerosol modes (nucleation, Aitken and accumulation). Solid line represents the median and shaded areas represent the interquartile interval.
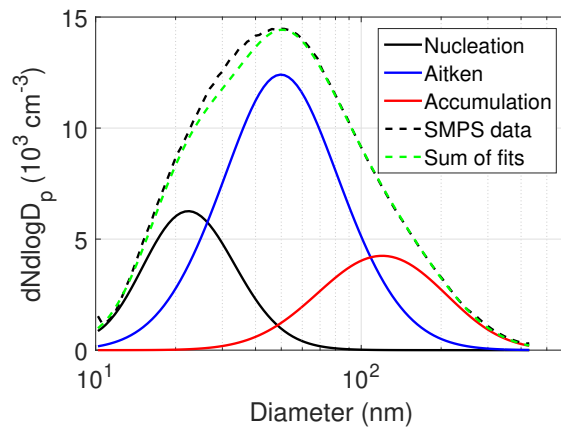
**Figure 5.** Average particle number size distribution for measured and lognormal fits for nucleation, Aitken and accumulation modes. Mean geometric diameter and geometric standard deviation of the modes are 12.6 (1.2) nm, 30.7 (1.3) nm and 83.8 (1.3) nm. The local maximum of the number distribution (mode diameter) is at 22.5 nm for nucleation, at 49.6 nm for Aitken, and at 121.9 nm for accumulation mode. The average $\pm$ standard deviation for particle number concentration is $12500 \pm 7200$ cm$^{-3}$ .
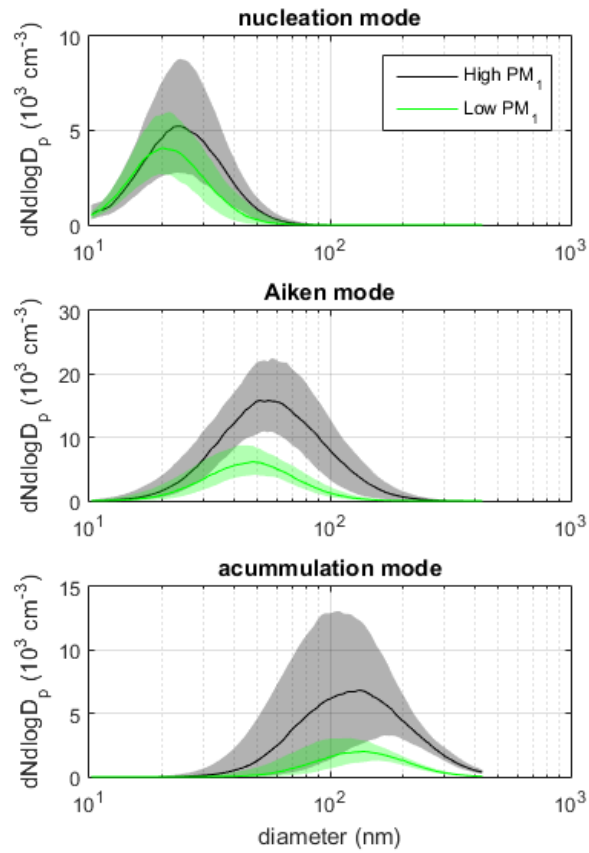
**Figure 6.** PNSD for nucleation, Aitken and accumulation modes under high (black) and low (green) PM$_1$ levels. Solid line represents median PNSD. Shaded areas represent the interquartile interval. PM$_1$ values higher than 75th percentile of time series are considered high PM$_1$ concentrations and PM$_1$ values lower than 25th percentile of time series are considered low PM$_1$ concentrations.
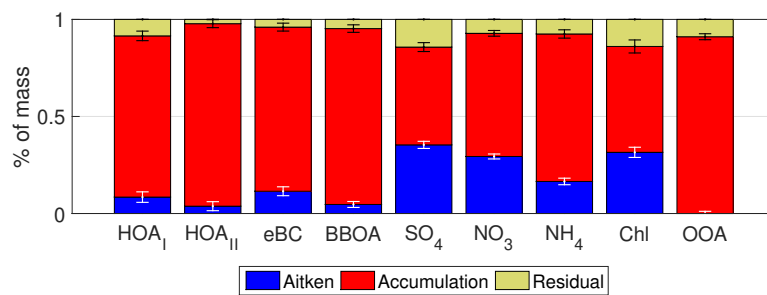
**Figure 7.** Contributions of aerosols size modes into mass concentrations for PM1 chemical species and confidence intervals. Results from the MLR model.
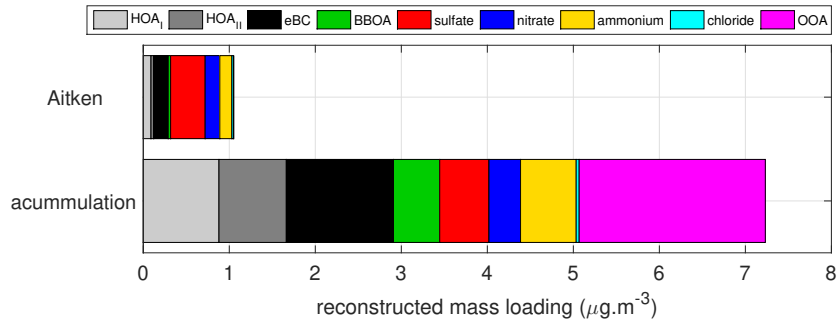
**Figure 8.** Reconstructed mass loading into Aitken and accumulation modes, as attributed by the MLR model.

**Table 1.** Summarized statistics of meteorological data and mixing ratios of trace gases from $20^{th}$ December 2016 to $15^{th}$ March 2017. Statistics includes mean concentration, standard deviation ($\sigma$), median and interquartile interval (IQ).

| variables (and units) | mean (and $\sigma$) | median (and IQ) |
|---|---|---|
| RH (%) | 79 (16) | 85 (68-92) |
| Air temperature (°C) | 23.0 (3.6) | 22.1 (20.5-25.4) |
| mid-day $O_3$ (ppb) | 43.5 (19.8) | 40.6 (29.0-55.0) |
| $NO_2$ (ppb) | 13.2 (8.9) | 11.2 (5.9-18.8) |
| $SO_2$ (ppb) | 0.61 (0.85) | 0.36 (0.19-0.67) |

**Table 2.** Summarized statistics of near real-time aerosol measurements ($PM_1$ and its individual constituents) from $20^{th}$ December 2016 to $15^{th}$ March 2017. Statistics includes mean concentration, standard deviation ($\sigma$), median and interquartile interval (IQ).

| variables (and units) | mean (and $\sigma$) | median (and IQ) |
|---|---|---|
| $PM_1$ ($\mu$g m$^{-3}$) | 11.4 (7.8) | 9.7 (5.9 - 14) |
| eBC ($\mu$g m$^{-3}$) | 1.5 (1.4) | 1.1 (0.6 - 1.9) |
| organics ($\mu$g m$^{-3}$) | 6.3 (5.0) | 5.1 (3.2 - 7.7) |
| sulfate ($\mu$g m$^{-3}$) | 1.6 (1.8) | 1.02 (0.42 - 2.30) |
| nitrate ($\mu$g m$^{-3}$) | 0.70 (0.67) | 0.46 (0.24 - 0.98) |
| ammonium ($\mu$g m$^{-3}$) | 1.03 (0.92) | 0.77 (0.08 - 1.50) |
| chloride ($\mu$g m$^{-3}$) | 0.19 (0.34) | 0.05 (0.01 - 0.13) |

**Table 3.** Pearson correlation coefficients (R) between mass spectra of OA PMF factors and AMS database (http://cires1.colorado.edu/jimenez-group/AMSsd/).

| Pearson correlation (R) with AMS database | OOA | HOA I | BBOA | HOA II |
|---|---|---|---|---|
| HOA (Ng et al., 2011b) | 0.26 | 0.90 | 0.53 | 0.89 |
| SV-OOA (Ng et al., 2011b) | 0.72 | 0.68 | 0.70 | 0.58 |
| LV-OOA (Ng et al., 2011b) | 0.99 | 0.30 | 0.41 | 0.10 |
| OOA (Ng et al., 2011b) | 0.98 | 0.30 | 0.41 | 0.10 |
| BBOA (Ng et al., 2011b) | 0.64 | 0.79 | 0.90 | 0.52 |

**Table 4.** Pearson correlation coefficients (R) between time series of OA factors and time series of aerosol components and trace gases. Aerosol components include sulfate, nitrate, ammonium, chloride and equivalent black carbon time series. Trace gases include sulphur dioxide, nitrogen dioxide, ozone and oxidant concentration ($Ox=O_3+NO_2$) time series.

| Pearson correlation (R) with external tracers | OOA | HOA I | HOA II | BBOA |
|---|---|---|---|---|
| $O_X$ | 0.48 | -0.06 | -0.07 | -0.03 |
| sulfate | 0.46 | 0.04 | 0.04 | 0.15 |
| $SO_2$ | 0.35 | 0.08 | 0.07 | 0.08 |
| ammonium | 0.53 | 0.16 | 0.19 | 0.30 |
| nitrate | 0.63 | 0.43 | 0.58 | 0.48 |
| eBC | 0.40 | 0.45 | 0.69 | 0.47 |
| $NO_2$ | 0.23 | 0.55 | 0.60 | 0.31 |
| $O_3$ | 0.36 | -0.23 | -0.28 | -0.14 |
| chloride | 0.19 | 0.23 | 0.32 | 0.58 |

**Table 5.** Fit parameters of MLR model results between $PM_1$ components (species of interest) and volume of aerosol size modes (predictors) and adjusted $R^2$.

| $PM_1$ species | Aitken mode | | accumulation mode | | Adj. R $^2$ |
|---|---|---|---|---|---|
| | $\beta(\sigma)$ | pValue | $\beta(\sigma)$ | pValue | |
| $HOA_I$ | 0.054 (0.017) | 0.002 | 0.116 (0.003) | <0.001 | 0.46 |
| $HOA_{II}$ | 0.019 (0.012) | 0.104 | 0.103 (0.002) | <0.001 | 0.58 |
| eBC | 0.103 (0.021) | <0.001 | 0.164 (0.004) | <0.001 | 0.57 |
| BBOA | 0.017 (0.006) | 0.002 | 0.071 (0.002) | <0.001 | 0.57 |
| sulfate | 0.244 (0.013) | <0.001 | 0.075 (0.003) | <0.001 | 0.54 |
| nitrate | 0.103 (0.004) | <0.001 | 0.048 (0.001) | <0.001 | 0.74 |
| ammonium | 0.085 (0.009) | <0.001 | 0.085 (0.002) | <0.001 | 0.52 |
| chloride | 0.012 (0.001) | <0.001 | 0.005 (0.001) | <0.001 | 0.35 |
| OOA | -0.057 (0.016) | 0.001 | 0.284 (0.005) | <0.001 | 0.62 |