

Submicron organic aerosol in Tijuana, Mexico, from local and Southern California sources during the CalMex campaign

S. Takahama ^a, A. Johnson ^a, J. Guzman Morales ^a, L.M. Russell ^{a,*}, R. Duran ^b, G. Rodriguez ^b, J. Zheng ^c, R. Zhang ^c, D. Toom-Sauntry ^d, W.R. Leaitch ^d

^a Scripps Institution of Oceanography, University of California San Diego, La Jolla, CA 92093-0221, USA

^b Universidad de Autónoma de Baja California, Tijuana, Mexico

^c Texas A&M University, College Station, TX, USA

^d Science and Technology Branch, Environment Canada, Toronto, ON, Canada

HIGHLIGHTS

- Submicron OM composition was measured by FTIR, ACSM, and STXM during CalMex.
- The average OM concentration was $3.3 \mu\text{g m}^{-3}$ in Tijuana.
- Vehicular sources contributed approximately 40% to the submicron OM.
- OM measured in Tijuana was highly oxygenated (average O/C = 0.6).
- The Southern California Air Basin was a large source of oxygenated OM in Tijuana.

ARTICLE INFO

Article history:

Received 30 October 2011

Received in revised form

18 July 2012

Accepted 19 July 2012

Keywords:

Organic aerosol

Air pollution

FTIR

ACSM

STXM

NEXAFS

Border region

ABSTRACT

The CalMex campaign was conducted from May 15 to June 30 of 2010 to study the properties and sources of air pollution in Tijuana, Mexico. In this study, submicron organic aerosol mass (OM) composition measured by Fourier Transform Infrared Spectroscopy (FTIR), Aerosol Chemical Speciation Monitor (ACSM), and X-ray spectromicroscopy are combined with statistical analysis and measurements of other atmospheric constituents. The average (\pm one standard deviation) OM concentration was $3.3 \pm 1.7 \mu\text{g m}^{-3}$. A large source of submicron aerosol mass at this location was determined to be vehicular sources, which contributed approximately 40% to the submicron OM; largely during weekday mornings. The O/C ratio estimated from ACSM measurements was 0.64 ± 0.19 ; diurnal variations in this value and the more oxygenated fraction of OM as determined from Positive Matrix Factorization and classification analyses suggest the high degree of oxygenation originates from aged OM, rather than locally-produced secondary organic aerosol. A large contribution of this oxygenated aerosol to Tijuana from various source classes was observed; some fraction of this aerosol mass may be associated with non-refractory components, such as dust or BC. Backtrajectory simulations using the HYSPLIT model suggest that the mean wind vector consistently originated from the northwest region, over the Pacific Ocean and near the Southern California coast, which suggests that the origin of much of the oxygenated organic aerosols observed in Tijuana (as much as 60% of OM) may have been the Southern California Air Basin. The marine aerosol contribution to OM during the period was on average $23 \pm 24\%$, though its contribution varied over synoptic rather than diurnal timescales. BB aerosol contributed $20 \pm 20\%$ of the OM during the campaign period, with notable BB events occurring during several weekend evenings.

© 2012 Elsevier Ltd. All rights reserved.

1. Introduction

California and Baja California share a common air basin along the ~ 200 km border, with sister city pairs of Tijuana–San Diego

and Mexicali–Calexico accounting for the highest observed ozone and aerosol concentrations in the urban areas of the border region (Border, 2012). These cities are linked through international commerce that is expected to contribute to increased atmospheric burdens of particulate matter (PM) and gases; rising trends in urbanization suggest that a larger share of the population will be adversely affected by these air pollutants. To study the implications

* Corresponding author.

E-mail addresses: lmrussell@ucsd.edu, lmrussell@ucsd.edu (L.M. Russell).

of current air quality relevant for climate and policy analysis, the CalMex campaign was conducted in Tijuana, Mexico, from May 15 to June 30, 2010, in conjunction with the CalNex campaign that focused on similar issues north of this border.

Atmospheric PM is a mixture of many components, often containing some combination of nitrate, sulfate, ammonium, crustal elements, sea salt, black carbon, and organic molecules (Seinfeld and Pandis, 2006). Of this airborne PM, the number of molecules comprising the organic aerosol fraction are numerous (Hamilton et al., 2004) and undercharacterized (Kanakidou et al., 2005). Our lack of knowledge regarding the constituents and characteristics of this fraction imparts large uncertainties in prediction of aerosol lifetimes (and thus atmospheric concentrations) and climate impacts incurred through aerosol interactions with solar radiation and water vapor. Higher ambient PM mass concentrations have been associated with increased mortality and morbidity (Dockery et al., 1993), posing a risk from a public health perspective.

In this study, we report characterization of the major source classes of organic aerosol to Tijuana and determine the significance of local and transported regional PM to the air quality at this site. We use several analytical methods to characterize the organic aerosol, including FTIR (organic functional groups) and ACSM (non-refractory molecular mass fragments). These measurements are combined with mathematical techniques, additional measurements, and backtrajectory analyses, which have effectively been used for source identification and apportionment for OM in previous campaigns (e.g., Russell et al., 2009; Liu et al., 2009; Hawkins and Russell, 2010; Schwartz et al., 2010; Frossard et al., 2011; Takahama et al., 2011).

2. Methods

2.1. Sampling site

The measurements (Table S1) were collected in a trailer installation at Parque Morelos, located in the Southeast region of Tijuana (32.52° , -117.037° , 6 m.a.s.l.). The park is an ecological reserve and recreational park surrounded by roadways, shopping centers, and residential neighborhoods, and situated less than 15 km from the San Ysidro border station.

2.2. Organic functional group composition

Atmospheric particles were sampled on Teflon filters [with filter changes at approximately 06h00, 14h00, and 18h00 Pacific Standard Time (PST) daily] and analyzed by Fourier Transform Infrared Spectroscopy (FTIR) for organic functional group (OFG) composition. The filters were analyzed immediately after removal from storage for FTIR analysis. Infrared sample spectra were obtained with a Tensor 27 spectrometer (Bruker, Billerica, MA), baselined, and fitted with peaks to identify OFG using the method described by Russell et al. (2009). Using this method, FTIR spectroscopy provides OFG concentrations, including alkane, carboxylic acid, organic hydroxyl, primary amine, and carbonyl groups through chemical bond-based measurements in atmospheric particles collected on a substrate (Russell et al., 2009). Alkene, aromatic, and organosulfate groups were below detection limit for all samples. Our discussion of OM will therefore neglect these compounds; their contribution to the actual OM is estimated to be between 2 and 5%. Ketone group contributions are estimated from a comparison of moles of carboxylic COH and total carbonyl quantified; non-acid carbonyl contributions (including aldehydes and ketones) are determined by the moles of carbonyl present in excess of quantified moles of carboxylic COH. The non-acid carbonyl is determined to be ketonic rather than aldehyde carbonyl, as absorption bands

between 2700 cm^{-1} and 2860 cm^{-1} indicative of aldehydic hydrogen were not observed in the sample spectra. The quantified non-acid carbonyl will therefore be referred to as ketones in this manuscript. The uncertainty and detection limit of ketones are therefore estimated through a contribution of the estimated carboxylic COH and total carbonyl (Russell et al., 2009). Estimation of mass is based on the analysis by Russell (2003), where moles of measured bonds are converted to the moles of comprising atoms, and values of OM are calculated from the sum of moles of atoms multiplied by their respective molecular weights. Using this approach, the uncertainty in OM has been calculated to be on the order of 21% (Russell, 2003).

2.3. Non-refractory organic and inorganic mass concentrations

The Aerodyne Aerosol Chemical Speciation Monitor (ACSM; Ng et al., 2011b) with a Pfeiffer Prisma quadrupole was used to scan over 150 mass/charge (m/z) units at a rate of 0.5 m/z per second. The ACSM is similar to the Aerodyne Aerosol Mass Spectrometer (AMS; Jayne et al., 2000) with the primary differences being the absence of a particle time-of-flight measurement and reduced sensitivity in the ion signals. Differences in measured signal between 6 to 12 aerosol-containing and aerosol-free air (using filter mode) are averaged over 15–30-min intervals for each measurement. The absolute and relative ionization efficiencies of nitrate and ammonium, respectively, are calibrated with nearly monodisperse particles (10:1 sheath to sample flowrate on a TSI, Inc. 3071 differential mobility analyzer and counted with a 3010 Condensation Particle Counter), and the relative ionization efficiencies for sulfate and organics are based on the AMS (Ng et al., 2011b). A diffuse naphthalene source is nominally used as an internal standard to normalize the measurements with respect to temperature and pressure; this source was depleted toward the end of the campaign so the closed signal of the m/z 32 associated with the airbeam is used for the corrections instead. Software provided by Aerodyne, Inc. is used to process the measurements to obtain concentrations of sulfate, nitrate, ammonium, chloride, and total organic aerosol, and organic mass fragment spectra. Further details are discussed by Ng et al. (2011b). A correction factor of 0.3 is used to scale the final ACSM concentrations such that the reported OM agrees with the FTIR; corrections of this type are anticipated to account for non-unit transmission and particle collection efficiencies associated with the ACSM (Ng et al., 2011b). The scaling is linear (no intercept) and therefore does not affect the conclusions regarding OM composition presented in this manuscript, which are based primarily on relative fractions and temporal variations of ACSM-measured components of OM. This correction factor has also been applied to non-refractory PM_{1} (NR- PM_{1}) and its components hereafter reported in this manuscript. Following conventions adopted by the AMS community (e.g., Ng et al., 2010, 2011b), the mass fragment concentrations are reported either in C_x [the OM-equivalent mass concentration of mass fragment (m/z) x], f_x [the OM-normalized mass concentration of mass fragment (m/z) x], or as integrated values of SO_4 , NO_3 , NH_4 , Cl , and OM (also commonly referred to as "OA" or "Org").

2.4. Dimension reduction methods

Positive Matrix Factorization (PMF; Paatero and Tapper, 1994) is commonly used for the analysis of environmental measurements, and more specifically, for the inversion of observed infrared and mass fragment spectra into a set of constituent components. Given a column matrix of row vectors (i.e., the sample spectra), \mathbf{X} , the forward model is expressed as $\mathbf{X} = \mathbf{GF} + \mathbf{E}$. \mathbf{G} and \mathbf{F} are matrices comprising component strengths and profiles, respectively; \mathbf{E} is the

residual matrix. The Q (or χ^2) value to be minimized is defined by the canonical objective function $Q = \sum \sum e_{ij}^2/s_{ij}^2$, where e_{ij} are the residuals (elements of E) and s_{ij} define the weighting for the fit. Details are described in Section S1.1.

Conceptually, PMF analysis permits determination of the relative proportion of each component in a sample mixture, provided that the inverse problem is well-constrained. When a possible set of components are hypothesized a priori, classification of the mixture data allows us to determine the components most similar to each reported observation according to a prescribed algorithm, and provides an additional tool by which we can evaluate among the possible solutions obtained through PMF decomposition. For this purpose, two supervised learning methods – k -nearest neighbor (k -nn) and linear discriminant analysis (LDA) – are also used for the classification of ACSM spectra against a set of training spectra. The training spectra consist of two subcategories of oxygenated organic aerosol (OOA): OOA-1, OOA-2; hydrocarbon-like organic aerosol (HOA); and biomass burning organic aerosol (BBOA), which are PMF components archived from various AMS studies (Ulbrich et al., 2009, 2012). Further details are included in Sections 3 and S1.3; results are summarized in Table S2.

2.5. Elemental analysis

Samples used for FTIR analysis were also used for measurement of elemental composition by X-ray fluorescence (XRF). Ninety of the same filter samples used for FTIR were sent to Chester LabNet (Tigard, Oregon) for mass quantification of Na and heavier elements (EPA Protocol 6).

2.6. Black carbon number concentrations

An 8-channel Single Particle Soot Photometer (SP2; Droplet Measurement Technologies, Inc.) measures individual particle responses upon interaction with a Nd:YAG laser (1064 nm). Particles containing black carbon (BC) absorb the energy and thermally irradiate as they vaporize at approximately 4000 K. The emitted thermal radiation (at visible wavelengths) is detected with red and blue sensitive photomultipliers. A calibration curve was generated by SP2 response to particles atomized from Aquadag in aqueous solution and size selecting them by using a TSI Electrostatic Classifier. Thresholds for event detection for high-gain and low-gain channels were set to 2500 and 100 counts, respectively, based on analysis of signal count histograms. The lower detection limit has been reported to be on the order of 0.7 fg, or 90 nm in mass-equivalent diameter (Schwarz et al., 2010).

2.7. X-ray spectromicroscopy analysis

Scanning Transmission X-Ray Microscopy with Near-Edge X-Ray Absorption Fine Structure (STXM-NEXAFS; Stöhr, 1992) with processing algorithms described by Takahama et al. (2010) was also used to examine single particle morphology and composition for a limited number of particles. With this method, X-rays generated at the Advanced Light Source at Lawrence Berkeley National Laboratories (Beamline 5.3.2) are used to probe the electronic structure of individual particles at a spatial resolution of approximately 30 nm. Collectively, 39 particles were analyzed from the campaign period.

2.8. Proton-Transfer-reaction mass spectrometry

A Proton-Transfer-Reaction Mass Spectrometer (PTR-MS) was used to measure the concentrations of volatile organic compounds (VOCs) by facilitating proton-transfer reactions in a drift-tube

reaction chamber between gas-phase species and H_3O^+ ions generated by an ion source (Lindner et al., 1998; Warneke et al., 2003). Gas-phase concentrations of benzene, toluene, and C_8 and C_9 aromatics measured by this instrument (Zhao and Zhang, 2004; Zhao et al., 2005; Fortner et al., 2009) are used for this study.

2.9. Backtrajectory analysis

Airmass back and forward trajectories were obtained from the HYbrid Single Particle Lagrangian Integrated Trajectory Model (NOAA HYSPLIT; Draxler and Rolph, 2010). Three-day back-trajectories were calculated every hour for the duration of the campaign at heights of 10, 100, and 500 m above ground level to consider the uncertainty of trajectories due to initial conditions. Mixing heights calculated by the HYSPLIT model were also extracted from the trajectory simulations.

3. Results and discussion

The non-refractory submicron aerosol composition measured by ACSM during the campaign was $5.8 \pm 5.1 \mu\text{g m}^{-3}$ on average (\pm one standard deviation shown), with contributions from sulfate ($14 \pm 7\%$), nitrate ($17 \pm 10\%$), ammonium ($12 \pm 7\%$), chloride ($3 \pm 3\%$), and organic aerosol ($53 \pm 14\%$) (Fig. 1). Trends in diurnal variations indicated organic aerosol, nitrate, and NR-PM₁ in general peaked during morning hours (06h00–10h00 PST; Fig. 2), with similar traffic activity observed during both weekdays and weekends. Mixing layer heights estimated by HYSPLIT indicate that the boundary layer begins to rise around 06h00 PST, so this increase in concentration is observed in spite of a rising boundary layer (which has a diluting effect). The following sections will describe the organic aerosol fraction in more detail and discuss inferred source classes and regions.

3.1. Submicron organic aerosol composition

3.1.1. FTIR analysis

The organic aerosol fraction as measured by FTIR during the campaign was $3.3 \pm 1.7 \mu\text{g m}^{-3}$ on average, with contributions from alkane ($44 \pm 9\%$), organic acid ($26 \pm 7\%$), organic hydroxyl ($22 \pm 13\%$), primary amine ($5 \pm 2\%$), ketone ($3 \pm 5\%$), and organonitrate ($<1\%$) groups (Fig. 3). The average organic aerosol composition was similar to that measured during previous campaigns in urban areas: Houston (Texas GoMACCS), Mexico City (MILAGRO), and Atlantis (CalNex) off the coast Southern California, while different from those measured in more remote areas (Fig. 4). These campaigns were conducted during different times and months (with the exception of CalNex Atlantis, which was measured concurrently), but the similarity of measured composition to Houston and Mexico City may reflect the large contribution from anthropogenic combustion sources which are the major source classes to these locations. As discussed in Section 3.3, the similarity of composition to Atlantis with the exception of ketone groups from burning episodes may also be attributed to common sources or airmasses.

3.1.2. ACSM analysis

The organic mass fragments from ACSM suggest a highly oxygenated aerosol, with an average O/C value of 0.64 ± 0.19 estimated for the campaign using the f_{44} parameterization by Aiken et al. (2008). This value is relatively high with respect to O/C ratios compiled by Aiken et al. (2008) for urban locations, and would be more closely associated with values observed in urban outflows. There is a possibility that a bias is introduced in estimating O/C from a parameterization developed for the AMS in

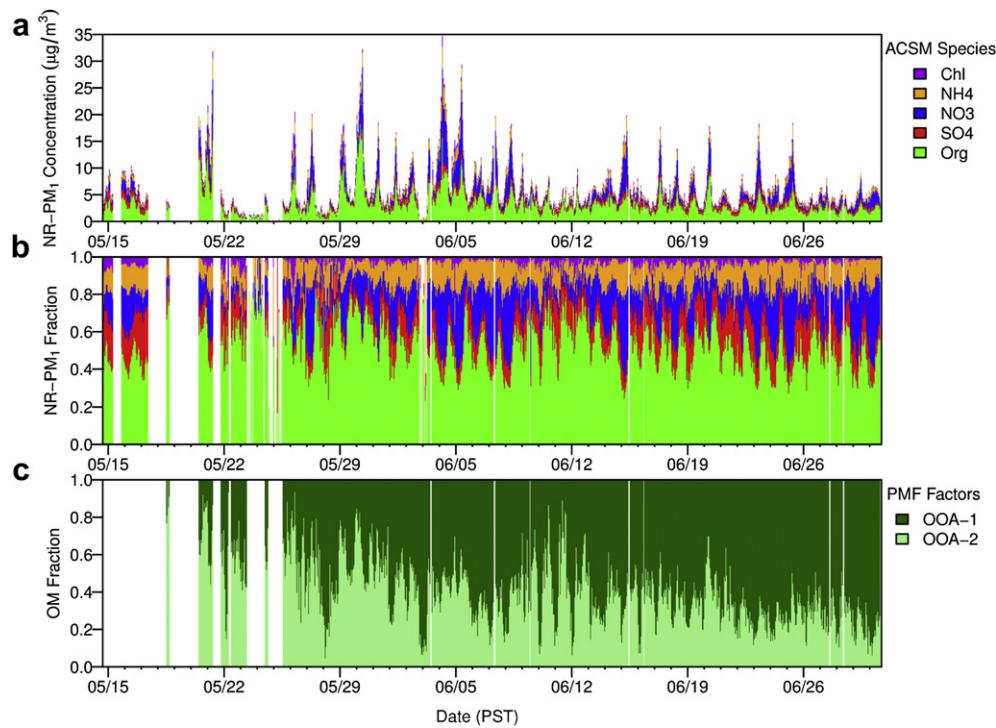


Fig. 1. Time series of (a) NR-PM₁ concentrations, (b) NR-PM₁ component fractions (normalized by sum of NR-PM₁) (c) OM fractions of PMF components.

Mexico City to the ACSM in Tijuana. Results reported by Ng et al. (2011b) suggest that AMS and ACSM spectra interpretation are comparable at Queens, NY, but the existence of some bias due to measurement of aerosol populations at a new geographical region with different size and composition cannot be ruled out. The average O/C measured by FTIR is 0.52 ± 0.16 (Fig. 5a), though this

average and much of its variation are due to large contribution from the marine fraction of OM. A direct O/C comparison with ACSM indicates a mild correlation with non-hydroxyl O/C ($r = 0.46$; Fig. 5b); some correlation between these two values is expected as the mass fragment on which O/C is parameterized for the ACSM is the m/z 44 (COO^+) fragment. This fragment is assumed to be

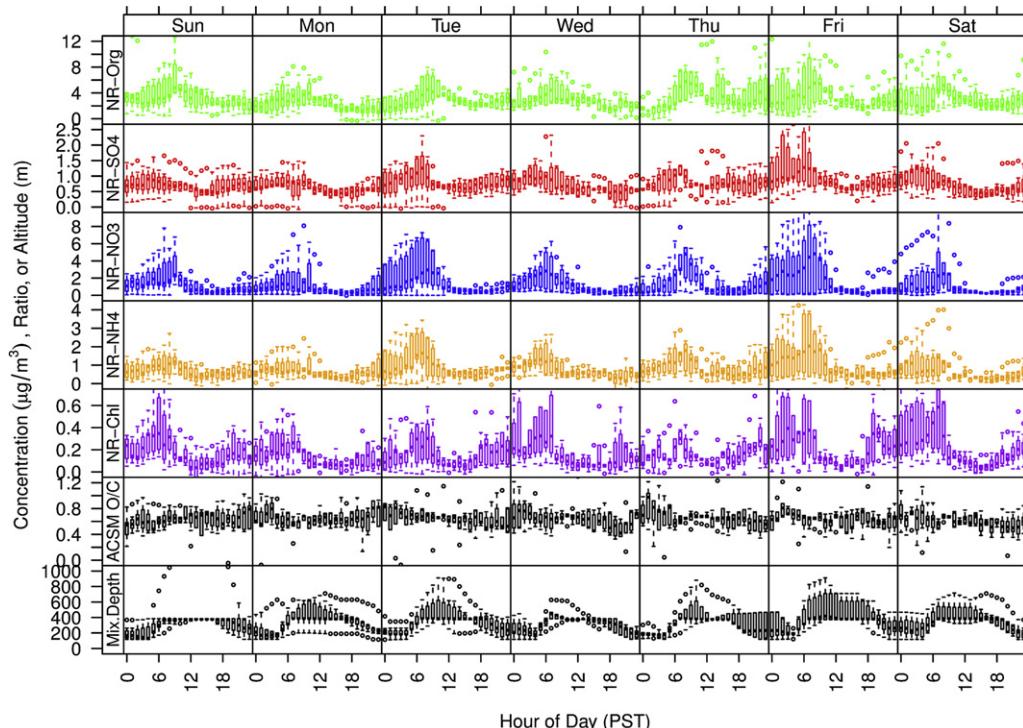


Fig. 2. Diurnal variations by day of week for non-refractory components, O/C ratio estimated with ACSM, and mixing heights (Mix.Depth) retrieved from NOAA's HYSPLIT model.

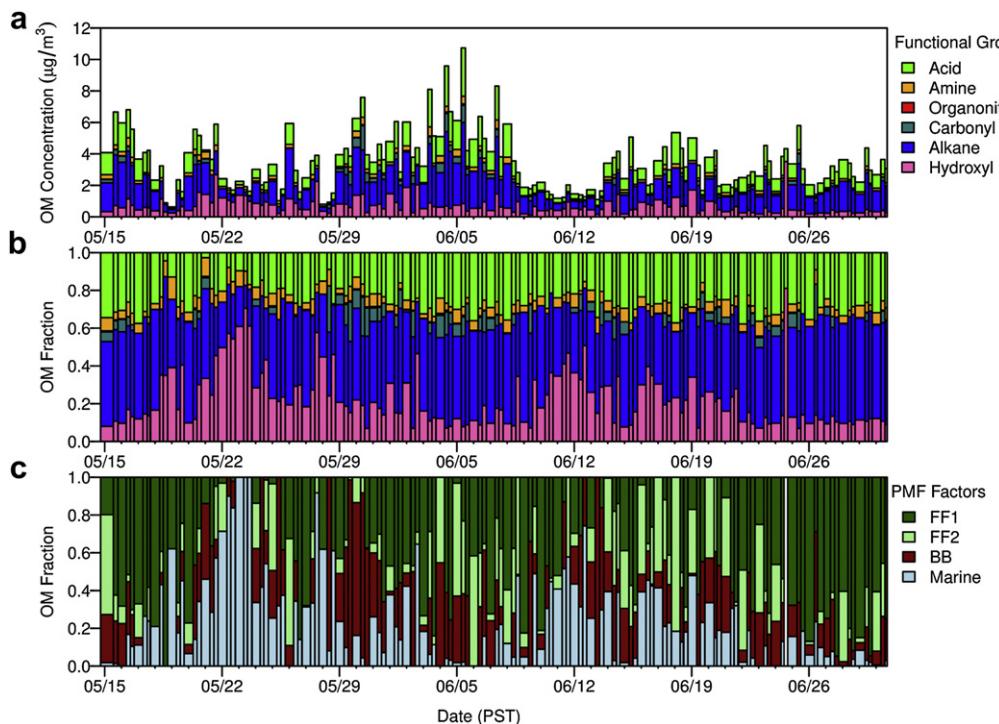


Fig. 3. Time series of (a) OFG concentrations, (b) OFG concentrations normalized by total OM, and (c) OM fraction of PMF components.

produced largely through the fragmentation of carboxylic acid groups (Russell et al., 2010; Ng et al., 2011b) rather than hydroxyl groups in alcohol or phenol compounds. Without O/C contributions from hydroxyl groups, the estimated FTIR O/C is approximately 0.3 ± 0.1 , which is low but closer to what has been reported for urban areas by AMS observations (Aiken et al., 2008). As discussed in Section 3.2, the marine fraction of OM as determined from PMF analysis has an O/C ratio of ~ 1 , but the rest of the contributing components range between 0.4 and 0.6, with the most

predominant component (fossil-fuel combustion) having an O/C of about 0.4 (Fig. 6). The magnitude of O/C overall is considerably lower for FTIR than estimated for the ACSM (especially when hydroxyl functional groups are excluded); this bias has also been reported in a previous FTIR-AMS comparison for the Texas GoMACCS campaign (Russell et al., 2009), and may reflect the differences in the analytical techniques and estimation methods used by spectroscopic and mass fragment analysis approaches to organic aerosol measurement. The ACSM O/C ratio most notably does not appear to show a strong diurnal variation (Fig. 1) as observed in Mexico City, where larger O/C values were observed in the afternoon periods corresponding to formation of SOA (Aiken et al., 2008; de Gouw et al., 2009).

If we examine the chemical composition in $f_{43}-f_{44}-f_{57}-f_{60}$ space (Fig. 7; based on precedent by Ng et al., 2010; Cubison et al., 2011; among others), we expect that the measurements in Tijuana most frequently fall within bounds delineated by oxygenated organic aerosol (OOA), more oxygenated OOA (OOA-1), and less oxygenated OOA (OOA-2). Though OOA is conceptually a superset of OOA-1 and OOA-2, the labels reflect categorizations provided by individual researchers in separate campaigns. This observation is more formally supported by classification analyses using LDA and k -nn in this domain (Fig. S2). These supervised learning methods suggest that most observations ($\sim 90\%$) would be classified as some type of OOA, with approximate ensemble compositions around 69–77% OOA-1 and 20–26% OOA-2 (Fig. 8a). The ensemble analysis of classified ACSM mass spectra indicate that the OOA-2 is more frequently observed during morning (06h00–09h00 PST) and evening (17h00–20h00 PST) hours, consistent with the picture that a higher fraction of hydrocarbons are present in organic aerosols associated with vehicle-related emissions. The fraction of the more oxygenated OOA-1 is higher during late evening and early morning hours, which may reflect the varying contribution of OOA-2 to the total OM in addition to the effect of continued aging of organic aerosol after hydrocarbon emissions from anthropogenic activity have been reduced.

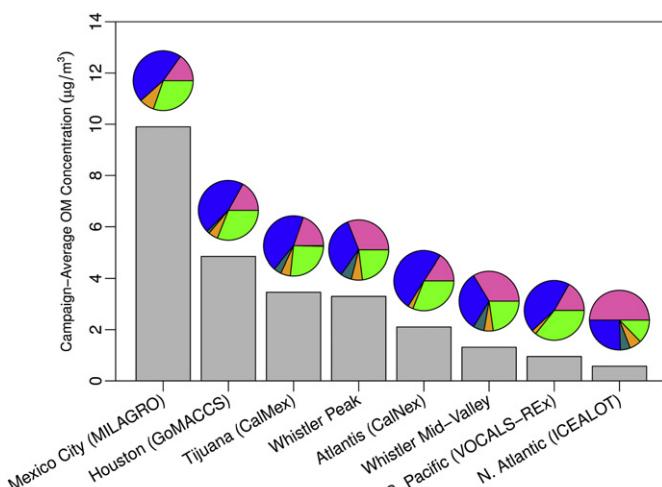


Fig. 4. Average OM concentrations (bar plots) and fractional contributions of OFG (pies) for various campaigns for which comparable FTIR measurements are available. Colors indicate contributions of individual functional groups to OM, and are defined in Fig. 3. Analysis for MILAGRO by Liu et al. (2009), GoMACCS by Russell et al. (2009), Whistler peak by Takahama and Russell (2011), Whistler Mid-Valley by Schwartz et al. (2010), VOCALS-REx by Hawkins and Russell (2010), and ICEALOT by Frossard et al. (2011). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

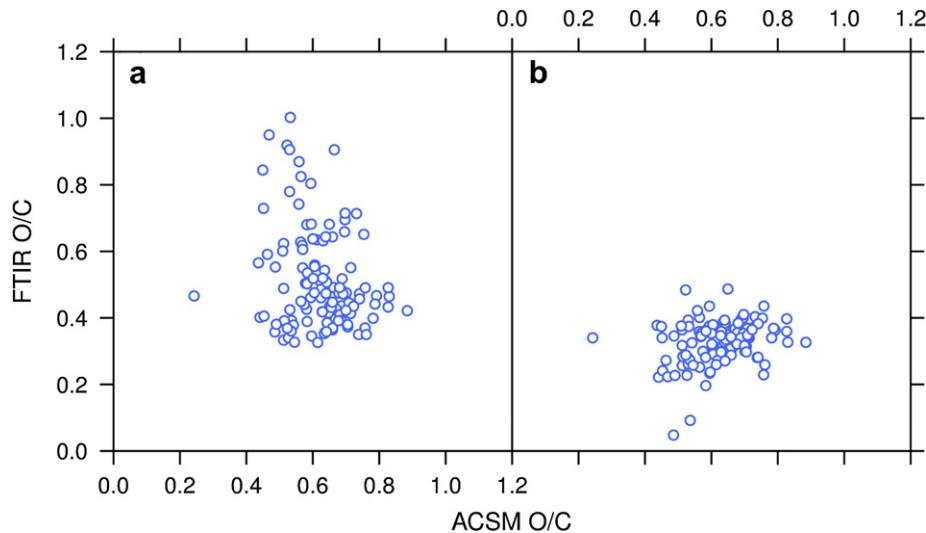


Fig. 5. O/C estimates for ACSM and (a) FTIR, and (b) FTIR excluding hydroxyl functional-group contributions to O/C values.

3.2. Source classes of PM_1 organic aerosol at Parque Morelos

3.2.1. PMF analysis of FTIR spectra

A four-factor solution was selected from PMF decomposition of FTIR spectra matrix (Fig. 9). The first factor was correlated with elemental tracers commonly associated with anthropogenic combustion (S and V; Fig. 10) and its spectra profile is similar to that of fossil fuel combustion from previous campaigns. A second factor was also correlated with S (Fig. 10) but was more oxygenated with an O/C of ~0.6 (as opposed to 0.4 for the first factor; Fig. 6). These factors will be referred to as FF1 and FF2 following the convention used by Russell et al. (2011) for fossil-fuel combustion-derived components. Alkane and carboxylic acid functional groups were the

major components of these two types of fossil fuel combustion classes; this has been attributed to emission of hydrocarbons and oxidative processes in the atmosphere producing organic acids. Another factor is correlated with elemental tracers for marine aerosol (Na and Cl, presumably from sea salt; Fig. 10) and its spectral profile is similar to that of marine aerosol from previous campaigns and laboratory standards of saccharides Russell et al. (2010, 2011). This marine factor accounted for $45 \pm 32\%$ of the hydroxyl functional group mass concentrations measured during the campaign. The last factor was correlated with K (Fig. 10), an elemental tracer for burning, and will therefore be referred to as biomass burning (BB). This factor was also correlated with elements not commonly associated with burning of biomass such as S and V (Fig. 10), so this component may also incorporate burning events from incineration, fuel-burning, or other anthropogenic activity. This factor accounted for all of the detectable ketonic carbonyl measured during the campaign. Elevated concentrations of ketones have been observed in burning samples in San Diego (Hawkins and Russell, 2010) and Whistler Peak (Takahama et al., 2011). The relative contribution of primary amine groups to the OFG concentration was not substantially higher during the BB periods as observed aboard the Ron Brown during the ICARTT campaign (Bahadur et al., 2010) or at Whistler Peak (Takahama et al., 2011). Elevated amine concentrations have been reported in BB aerosol elsewhere (e.g., Laskin et al., 2009), but variations in their contributions to the BB aerosol fraction may indicate differences in emissions or gas-particle conversions (Ge et al., 2011). Regions of the spectra ($\sim 860 \text{ cm}^{-1}$) associated with organonitrate absorption (Garnes and Allen, 2002; Day et al., 2010) were not included in the PMF decomposition, but a posteriori correlation of organonitrate mass concentrations with BB OM ($r = 0.8$) indicates the presence of organic nitrate compounds in the aerosol phase during BB events. While an elevation in organonitrate aerosol concentrations was detected, the overall concentration of organonitrate compounds in the condensed phase remained less than one percent of the OM mass during these periods.

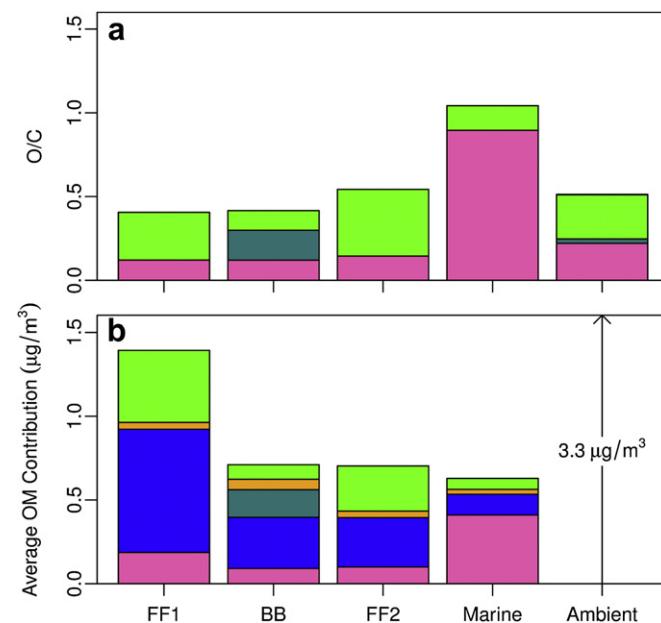


Fig. 6. Estimates of (a) O/C ratios and (b) average contributions to OM during the campaign for FTIR PMF components. Colors indicate contributions of individual functional groups to O/C or OM, and are defined in Fig. 3. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3.2.2. PMF analysis of ACSM spectra

While HOA and various types of OOA are commonly found in other locations (e.g., Zhang et al., 2005, 2007; Jimenez et al., 2009), the degree of oxygenation as indicated by the ACSM mass fragment measurements and classification analyses discussed in Section 3.1

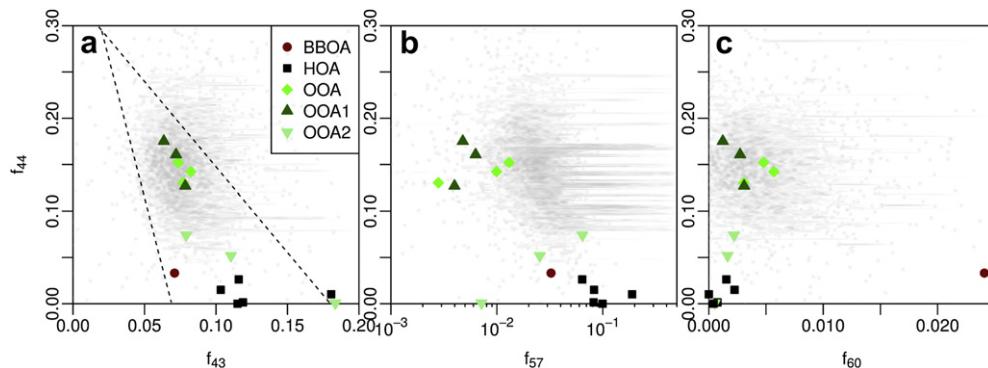


Fig. 7. Normalized mass fragment concentrations (Section 2) projected into f_{43} – f_{44} – f_{57} – f_{60} space. Gray points indicate Tijuana ambient measurements. Colored symbols represent PMF components extracted from previous campaigns in other locations (provided through AMS Spectral Database; Ulbrich et al., 2009, 2012). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

suggest that an all-OOA PMF solution is a plausible interpretation of the ACSM spectra for Tijuana during this campaign period (Fig. 11). A possible deconvolution (in the spectral sense) into HOA and OOA can be forced through regression analysis with mass fragment apportionment parameters proposed from previous campaigns (Ng et al., 2011a) (Section S2.1). However, no HOA and OOA solution was found to be feasible within the specification of our PMF analysis (Sections 2 and S1.1). Similarity of selected PMF factor profiles (with O/C ratios of 0.7 and 0.4, respectively) within the composition space of OOA-1 and OOA-2 reported previously (Ng et al., 2010;; with slight differences) argues in favor of adopting these naming conventions for our components. It is perhaps more appropriate to propose that a two-component solution yields a pair of oxygenated

and less-oxygenated components along of continuum of oxygenation, as the temporal variations of each component (according to PMF and regression analysis) are approximately the same (Section S2.1). Within this robust finding, we can find a systematic difference in the OM mass apportioned between the two components (Section S1.5), but primarily discuss our interpretation from the perspective of the PMF analysis.

Our selected PMF solution for ACSM spectra suggests that the aerosol mixture composition is approximately $40 \pm 19\%$ OOA-2 and $60 \pm 19\%$ OOA-1 during the campaign, though the OOA-2 fraction is higher during the earlier period of the campaign when the OM concentrations are higher ($r = 0.46$; Fig. 1c). Elevated m/z 60 concentrations are observed during BB events, but the frequency of

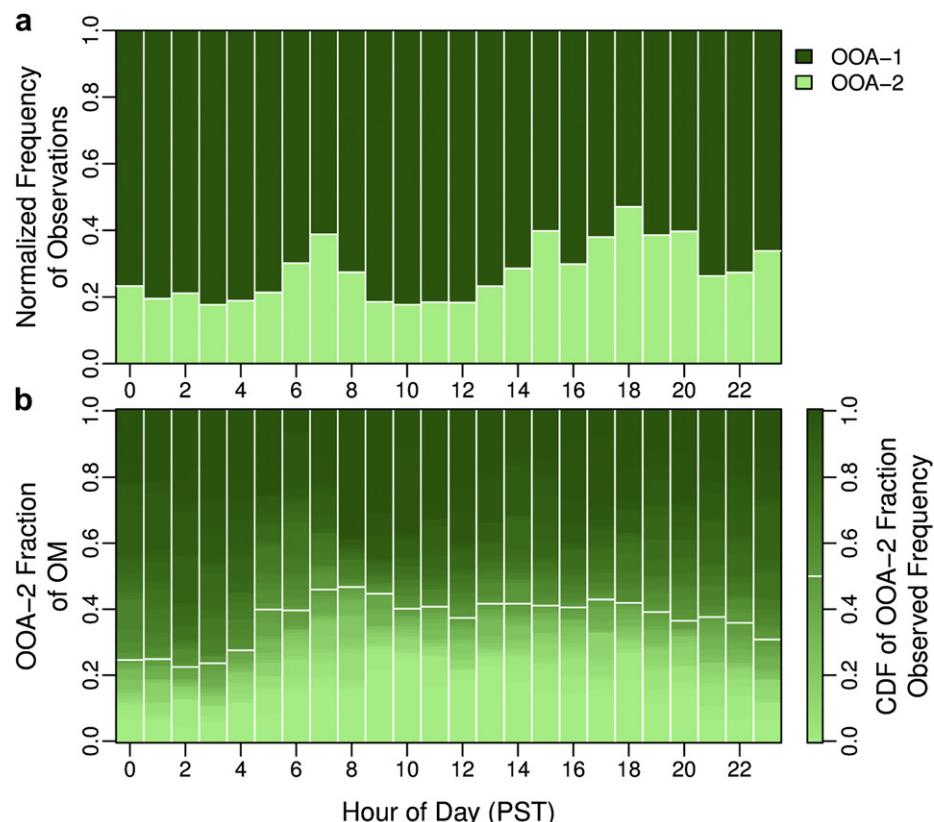


Fig. 8. Ensemble statistics of OOA-1 and OOA-2 contributions to hour of day during the campaign as estimated by (a) LDA and (b) PMF. Color gradient in panel (b) corresponds to cumulative distribution function (CDF) of OOA-2 fraction of OM, with a white line demarcating the location of the median. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

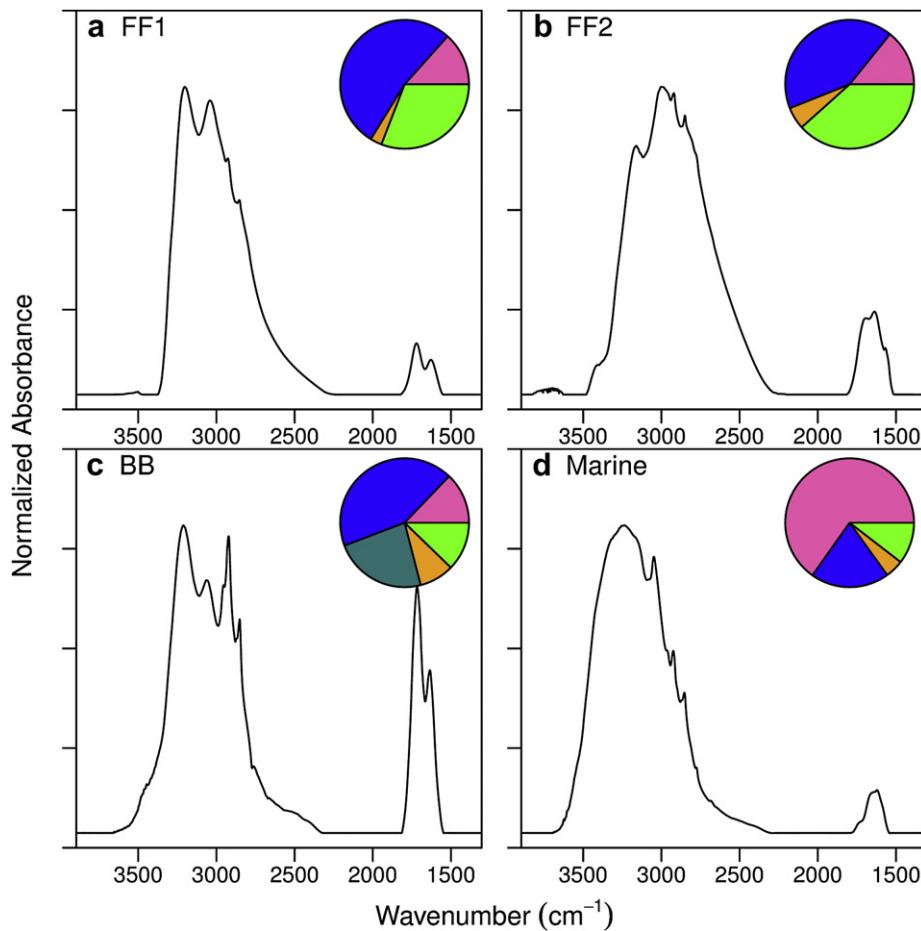


Fig. 9. FTIR PMF factor profile spectra and relative functional group composition. Colors indicate contributions of individual functional groups to OM, and are defined in Fig. 3.

observation and relative magnitudes of contribution are too low for a BB PMF component to be separated from the ACSM mass spectra. Statistics on the ensemble indicate that the OOA-2 fraction is frequently higher during elevated traffic activity periods at 06h00–09h00 (PST) and late afternoon (Fig. 8b) and lower during evening

and early morning hours. Several aspects of this diurnal profile are qualitatively consistent with trends indicated by the classification analysis (Fig. 8b). Robust features include the rise of OOA-2 around periods of peak traffic activity and higher fraction of oxygenated aerosol during late evening and early morning hours. An additional

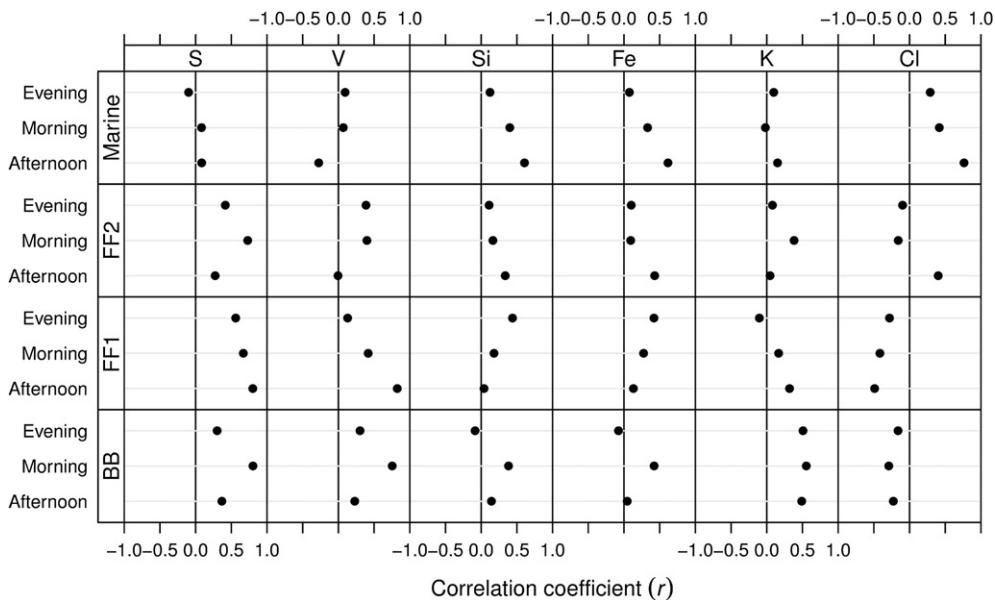


Fig. 10. FTIR PMF factor OM correlations with elements from XRF analysis.

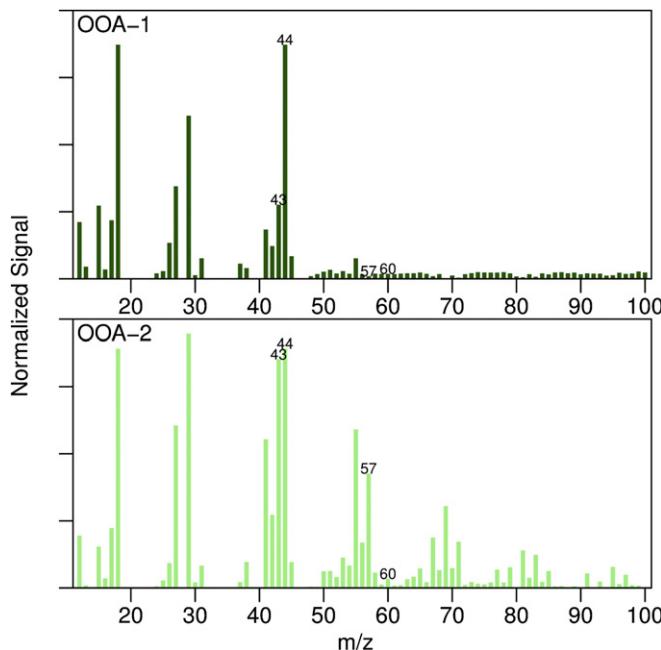


Fig. 11. ACSM PMF factor mass spectra.

regular feature includes small increases of oxygenated aerosol fraction (OOA-1) around 12h00 PST, which is the period of maximum solar intensity and therefore UV-initiated photochemistry. An alternative method of separation in which apportionment is forced into HOA and OOA categories using OM-equivalent concentrations of m/z 44 and m/z 57 mass fragments is discussed in Section S1.5, and supports the diurnal variations captured by this PMF solution (while in this case the less-oxygenated HOA fraction is estimated to be significantly lower). The combined interpretation from these analyses is that the ACSM mass spectra are separable into two components, one of which is less oxygenated than the other. The correlation of factor strengths ($r = 0.77$ for OOA-2 and HOA, and 0.86 for OOA-1 and OOA, respectively) suggests that the less oxygenated component is likely to be associated with primary emissions that are emitted locally (based on the assumption that HOA can be interpreted as POA; e.g., Liggio et al., 2010). The estimate of the primary fraction differs by 28% (of $3.3 \mu\text{g m}^{-3}$) depending on which interpretation is selected, but is robust with respect to the temporal trends that are represented by this component (Section S1.5).

3.2.3. Source identification from time-resolved measurements

Examination of time-resolved OM from ACSM with black carbon number concentrations from the SP2 for morning periods indicate that co-emission of OM from traffic combustion sources may be a large contributor to overall OM, as indicated by correlations with m/z 57, and OOA-2 (Fig. 13). The correlations of these components with BC for weekday mornings are milder than for the weekend mornings because there is larger variability in OM concentrations relative to BC, possibly suggesting additional sources during weekday morning periods. In contrast, mild correlations of OM and m/z 57, OOA-2, and additionally m/z 60 with SP2 BC number concentrations (Fig. 13) during several weekend evenings (notably 5/30, 6/13, and 6/20 corresponding to periods between Saturday night and Sunday mornings) suggest that the BB sources are large contributors to OM during these periods. Correlations of BC concentrations with FTIR factors support these statements, with $r = 0.72$ with the BB factor on weekend evenings, and $r = 0.58$ and

0.46 with the FF1 factor on weekday mornings and weekend afternoons, respectively. Examination of diurnal variation in VOC concentrations further support attribution of OOA-2 to vehicular combustion. Concentrations of VOCs traditionally associated with anthropogenic emissions – benzene and heavier (C_8, C_9) aromatic compounds, often associated with diesel exhaust (Jobson et al., 2005) – have correlation of $r = 0.51$ – 0.53 with OOA-2 within the same hour (Fig. S5).

3.2.4. The nature of BB aerosol at the Parque Morelos site

Hydroxyl functional groups present in levoglucosan and anhydrous sugars reported in fresh BB sources (Fine et al., 2002; Sullivan et al., 2008) are found in small quantities in the FTIR BB factor, while m/z 60 concentrations from ACSM are still correlated ($r = 0.70$) with this component. This observation is consistent with findings by Hawkins and Russell (2010) and Takahama et al. (2011). The BB samples reported in these past studies had been aged over long transport times; given the large number of fires reported in Tijuana during the campaign period and instances of visible smoke at the sampling site, the BB aerosol sampled here are thought to be attributed to more local sources during most periods. The relative fraction of hydroxyl groups is slightly higher in the BB factor observed during the CalMex campaign, and this may reflect differences in age (Hennigan et al., 2010). While the m/z 60 signal did increase during BB events, the OM-normalized f_{60} fraction was not correlated ($|r| < 0.2$) with BB OM or carbonyl functional groups measured by FTIR. Additionally, periods during which f_{60} increased above the background levels of $0.3 \pm 0.06\%$ (defined by Cubison et al., 2011) did not correspond to periods of BB in our samples. Our mean and standard deviation for f_{60} values for the campaign were $0.3 \pm 0.14\%$ (Fig. 7), and no significant differentiation in this metric between BB and non-BB periods as determined by FTIR was observed. This observation may possibly reflect the presence of additional organic mass in the BB aerosol in Tijuana. Co-occurrence of anthropogenic elements reported by XRF analysis (Fig. 10) suggests that the composition of this BB aerosol may be different from BB aerosol previously reported in the literature. Combined with the observation of increase in total organic aerosol concentration during these BB periods, it is possible that the composition domain is not sufficiently distinguished by f_{60} for the burning events observed in Tijuana during this campaign.

3.2.5. Synthesis and OM apportionment

From the FTIR PMF analysis, we find that the largest contributor to submicron OM to Tijuana during this campaign is FF combustion, comprising $40 \pm 28\%$ less oxygenated and $17 \pm 19\%$ more oxygenated fractions, with marine and BB components also contributing approximately 20% each (Fig. 6). ACSM mass fragment analysis indicates the presence of very oxygenated aerosol overall (mean O/C of 0.64 ± 0.19), and the more oxygenated of the two ACSM PMF components (OOA-1) is, on average, the larger contributor to the organic aerosol fraction of submicron OM. Its contributions are most frequently highest during evening hours, suggesting nighttime production and/or accumulation under a shallow planetary boundary layer. The less oxygenated OOA (OOA-2) is found to peak most frequently during morning hours, corresponding to traffic activity in Tijuana and consistent with expectations that aerosol containing more hydrocarbon-like molecules would be present during periods influenced by this source category. Because FTIR sampler filter changes occurred each morning at 07h00, the peak loadings from this source (traffic) were sampled on both the evening and morning samples. The relative contributions of OOA-1 and OOA-2 to Tijuana OM largely characterize the degree of aerosol oxygenation from a mixture of source classes, as indicated with correlations with FTIR PMF components (BB, $r = 0.70, 0.63$; FF1,

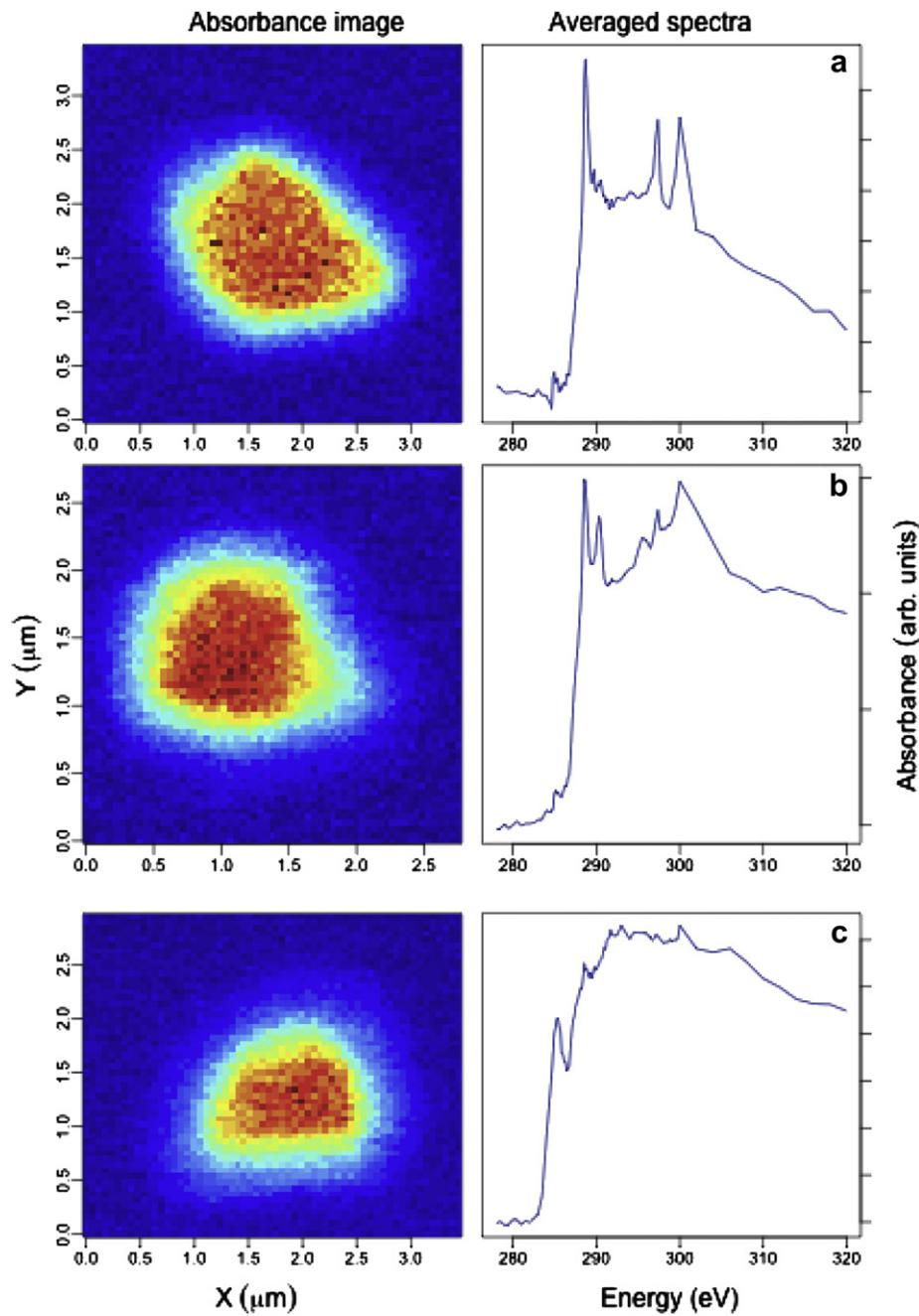


Fig. 12. Example STXM-NEXAFS images and spectra for (a) organic acid-dominated, (b) organic-dust, and (c) organic-black carbon particles.

$r = 0.36, 0.03$; FF2, $r = 0.35, 0.16$ for OOA-1 and OOA-2, respectively). These correlations also suggest that the split by degree of oxygenation for the FTIR FF factors and ACSM OOA factors is relative to the measurement by each instrument and not necessarily universal or absolute, though the coarse time resolution of the FTIR may also have an effect in the resolvability of component contributions (Henry, 2003). As discussed in Section S2.1, approximately 50–60% of the OM associated with FF factors is associated with OOA-1, suggesting an estimate for the anthropogenic contribution from remote sources (Sections 3.3 and S2.1).

There is some evidence that the weak (or lack of) correlation between FTIR FF2 and ACSM PMF factors may be due to the difference in the fraction of OM that is quantified by the two instruments. The stronger correlation of FF2 with dust elements in

the afternoon (Fig. 10) may suggest the formation of SOA on existing dust particles, which would not be included in the ACSM OM as this component would be associated with non-refractory material or larger particles not efficiently transmitted through the ACSM inlet. This hypothesis is consistent with the relatively small increase in the OOA-1 fraction observed by ACSM in the afternoon ($7 \pm 19\%$ on average), though it is also possible that the majority of SOA production from local emissions happens downwind of the measurement site. The STXM-NEXAFS single-particle analysis (example images and spectra shown in Fig. 12) supports the hypothesis that some of the discrepancy in the FTIR and ACSM measurements of organic aerosol may be explained by the fraction of OM associated with non-refractory material; 33 out of 39 particles analyzed indicated organic aerosol mixed with dust

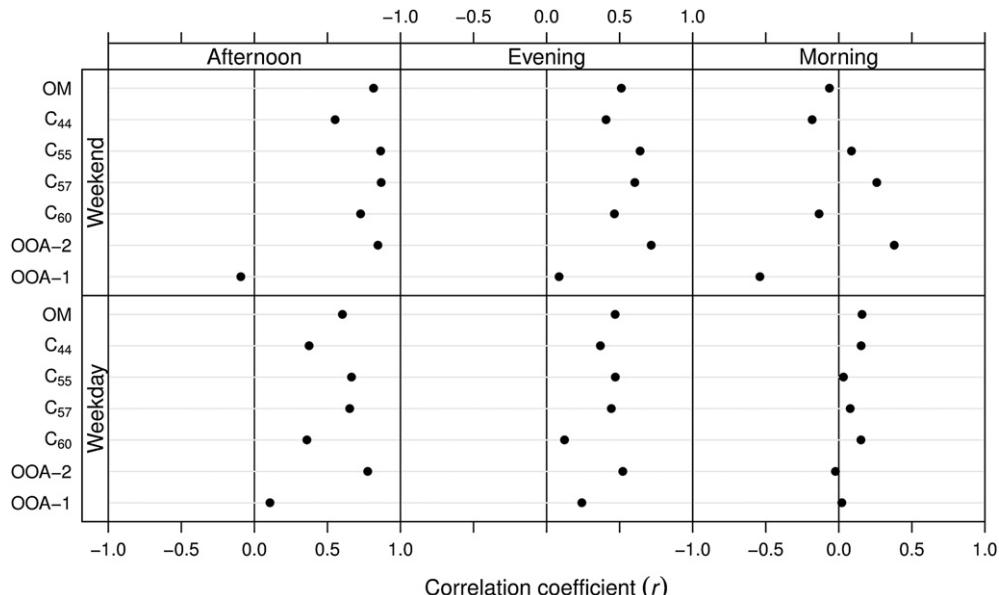


Fig. 13. Correlations of BC with ACSM OM, mass fragments, and PMF components. C_x is the OM-equivalent mass concentration of mass fragment (m/z) x .

($n = 15$; 1–3.5 μm) or BC ($n = 30$; 0.2–3.1 μm), and six carboxylic-acid dominated particles (Takahama et al., 2007) (1.4–4.6 μm). Additionally, the marine aerosol is not correlated with ACSM OM ($r = 0.10$), presumably because the OM is associated with non-refractory, larger sea salt particles.

3.3. Geographical origins of source contributions

HYSPLIT backtrajectories from Tijuana indicate that the airmass consistently originated from the west and northwest quadrant (Fig. 14) without much variation by time of day, week, or speed. These quadrants include San Diego, Tijuana, and the intermediate border region, which are likely origins of the less oxygenated OM that correlates in with BC and VOC emissions that are likely to be

from automobiles. In previous studies, the more oxygenated fraction of OM has been apportioned to locally-produced SOA and background OM based on diurnal occurrence and correlation with VOC tracers (de Gouw et al., 2009; Liggio et al., 2010; Liu et al., 2011). A delayed correlation of OOA-1 with anthropogenic VOCs (Fig. S5; maximum $r = 0.26$ delayed by one or two hours) may suggest some production of SOA at the site; however, the increase in OOA-1 fraction of OM between morning and afternoon periods at this measurement site is on average $7 \pm 19\%$ (the actual OOA-1 and OM concentrations decrease, on average, but largely due to differences in boundary-layer heights). The relatively uniform O/C ratio and OOA-1 fractions suggest that a large fraction of OM (estimated as 60% for an upper bound from the OOA-1 fraction) is transported to Parque Morelos from more regional sources (and contains a mixture of sources, as described in Section S2.1). There are several possible explanations, which include advection of urban aerosol to the Pacific Ocean during the daytime and returning during the night due to land-sea breeze (LSB) circulation. This effect has been reported by Cass and coworkers (Shair et al., 1982; Cass and Shair, 1984) for the Los Angeles (LA) area, and this sub-grid-scale phenomenon is not captured by the HYSPLIT backtrajectories as the model calculates mean wind vectors derived from archived GDAS meteorology assimilated with a spatial resolution of 40 km. Despite the regularity in the temporal cycle of anthropogenic emissions, examination of the marine aerosol fraction of OM from PMF on FTIR measurements does not indicate a strong diurnal pattern, which would be expected to be low during the day and higher during the evening for a strong, local LSB influence. The marine aerosol fraction is highest during successive periods extending over several days, indicating the larger influence of synoptic-scale meteorology in regulating the marine fraction of aerosol. While Ault et al. (2009) and Hawkins and Russell (2010) report large contributions of LA aerosol to the mass budget of OM in San Diego, backtrajectories from Tijuana and forward trajectories from LA during this period indicate that the major path of advection from LA was not directed through San Diego and Tijuana during the CalMex campaign period. However, it is possible that the source of the oxygenated aerosol may be Southern California pollution advected to sea through LSB circulation and transported to this measurement site in Tijuana. LSB circulation may also explain the

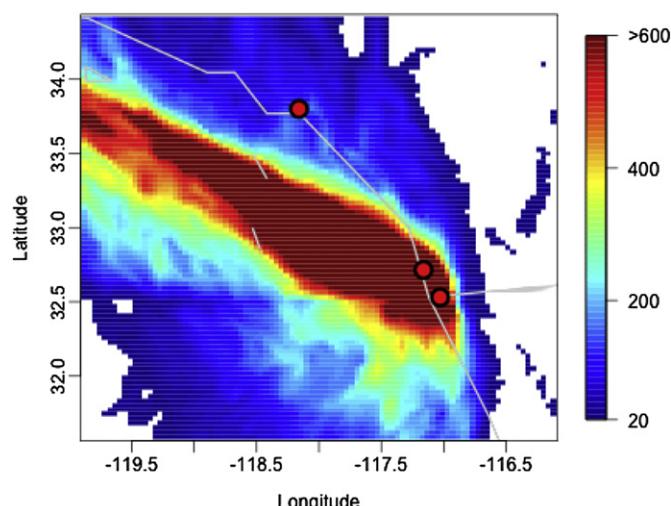


Fig. 14. Density image of backtrajectories from Tijuana for starting altitudes of 10, 100, and 500 m a.g.l. The color scale bounds values between 10th and 90th percentile of the hourly-trajectory points falling in a 40 km \times 40 km grid in the displayed domain. The circles indicate the location of major cities, from north to south: LA, San Diego, and Tijuana. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

observed increases in concentrations of inorganic aerosol nitrate during the early morning periods (Fig. 2), which may be transported to the site following production of nitric acid through the N_2O_5 pathway in a humidified environment over the ocean. Entrainment of free-troposphere aerosol from above the shallow marine boundary layer (200–300 m on average during daytime) cannot be ruled out, but the source of pollution in FT is unlikely to be large and consistent enough to account for the OM measured. Transport from west of Los Angeles would require 10–24 h (for wind speeds of 2–5 m s^{-1} estimated from HYSPLIT calculations; Section S3); this interpretation is consistent with the observed O/C values indicative of urban outflow (Aiken et al., 2008), and lack of strong diurnal variations observed in the OOA-1 fraction during this campaign. This meteorological pattern is not likely to be an isolated phenomena observed only during this campaign; simulation and reanalysis over longer timescales (six years) suggest that north easterly winds due to the Pacific Subtropical High (PSH) system is a common meteorological pattern in the region (Zhao et al., 2011).

4. Conclusions

PM_1 organic aerosol composition was measured by FTIR, ACSM, and STXM-NEXAFS. Combining these measurements with statistical analysis methods (factor and classification analyses), additional measurements (e.g., BC, VOCs) and meteorological models, the OM mass was apportioned to various source classes and a dominant source region. Synthesizing organic aerosol measurements from various instruments, we can make several statements about the concentration, sources, and origins of OM in Tijuana. The campaign-average OM concentration was $3.3 \pm 1.7 \mu\text{g m}^{-3}$ (approximately $53 \pm 14\%$ of NR- PM_1), with anthropogenic combustion, biomass burning, and marine sources contributing to this aerosol burden. The aerosol was very oxygenated according to ACSM analysis, with average O/C ratios of 0.64 ± 0.19 . The PMF and classification analyses suggest that the more oxygenated fraction of the OM is likely to be aged OM transported to Tijuana rather than locally-produced SOA. A large contribution of oxygenated aerosol from various source classes was inferred; some fraction of oxygenated (secondary) organic aerosol may be associated with non-refractory components, such as dust or BC. Backtrajectory simulations using the HYSPLIT model suggest that the mean wind vector consistently originated from the northwest region, over the Pacific Ocean. Much of the aged aerosol (as much as 60% of the OM and containing a mixture of sources, based on quantification of the highly oxygenated fraction) is likely to have been transported to Tijuana from pollution advected to sea by a sub-grid scale land-sea breeze circulation off the coast of Southern California. Anthropogenic combustion is estimated to contribute approximately 40–60% to the OM mass at Parque Morelos, and 40–50% of this is likely to be due to local sources, such as automobile traffic. The marine aerosol contribution to OM during the period was on average $23 \pm 24\%$, though its contribution varied over synoptic rather than diurnal timescales. OM from BB aerosol contributed $20 \pm 20\%$ and the OM during the campaign period, with notable BB events occurring during several weekend evenings. OM observed during this period were less oxygenated (apportioned to the OOA-2 factor) and different in composition from previously reported BBOA (according to magnitudes of the estimated f_{60} metric), consistent with their co-occurrence with elements normally associated with industrial or anthropogenic emissions (S, V). The more oxygenated component of anthropogenically-derived OM reported by FTIR PMF analysis appears to be SOA formed on mineral dust surfaces, which was not captured by the ACSM (similarly with marine aerosol) as it was associated with non-refractory particles.

Acknowledgments

The authors would like to thank: the Molina Center and the participants of the CalMex campaign who assisted with logistics and sample collection (particularly Luisa T. Molina, Hugo Barrera, Miguel Zavala, Claudia Rivera, Aldo Cortez, and Bianca Puckita); Zhan Zhao for her expertise in meteorological analysis; Ashley Corrigan for helpful comments on the PMF analysis; John T. Jayne for ACSM operation in the field; and Phil Croteau and Sally Ng for assistance and software for preparation of measurement error matrix for PMF analysis of ACSM mass spectra. Funding for this work was provided by NSF AGS-1009408, with logistical costs supported by CARB 09-356. The statements and conclusions in this paper are those of the researchers (contractor) and not necessarily those of CARB. The mention of commercial products, their source, or their use in connection with material reported herein is not to be construed as actual or implied endorsement of such products.

Appendix A. Supplementary data

Supplementary data related to this article can be found at <http://dx.doi.org/10.1016/j.atmosenv.2012.07.057>.

References

- Aiken, A.C., Decarlo, P.F., Kroll, J.H., Worsnop, D.R., Huffman, J.A., Docherty, K.S., Ulbrich, I.M., Mohr, C., Kimmel, J.R., Sueper, D., Sun, Y., Zhang, Q., Trimborn, A., Northway, M., Ziemann, P.J., Canagaratna, M.R., Onasch, T.B., Alfarra, M.R., Prevot, A.S.H., Dommen, J., Duplissy, J., Metzger, A., Baltensperger, U., Jimenez, J.L., 2008. O/C and OM/OC ratios of primary, secondary, and ambient organic aerosols with high-resolution time-of-flight aerosol mass spectrometry. *Environmental Science & Technology* 42 (12), 4478–4485.
- Ault, A.P., Moore, M.J., Furutani, H., Prather, K.A., May 2009. Impact of emissions from the Los Angeles port region on San Diego air quality during regional transport events. *Environmental Science & Technology* 43 (10), 3500–3506.
- Bahadur, R., Upfinger, T., Russell, L.M., Sive, B.C., Cliff, S.S., Millet, D.B., Goldstein, A., Bates, T.S., 2010. Phenol groups in northeastern US submicrometer aerosol particles produced from seawater sources. *Environmental Science & Technology* 44 (7), 2542–2548.
- Border, 2012. US Mexico Environmental Program, Indicators Report, 2005. Tech. rep. <http://www.epa.gov/usmexicoborder/docs/BorderIndicatorsReportApril2006.pdf>.
- Cass, G., Shair, F., 1984. Sulfate accumulation in a sea breeze/land breeze circulation system. *Journal of Geophysical Research* 89 (D1), 1429–1438. URL: <http://www.agu.org/journals/ABS/1984/JD089iD01p01429.shtml>.
- Cubison, M.J., Ortega, A.M., Hayes, P.L., Farmer, D.K., Day, D., Lechner, M.J., Brune, W.H., Apel, E., Diskin, G.S., Fisher, J.A., Fuelberg, H.E., Hecobian, A., Knapp, D.J., Mikoviny, T., Riemer, D., Sachse, G.W., Sessions, W., Weber, R.J., Weinheimer, A.J., Wisthaler, A., Jimenez, J.L., 2011. Effects of aging on organic aerosol from open biomass burning smoke in aircraft and laboratory studies. *Atmospheric Chemistry and Physics* 11 (23), 12049–12064. URL: <http://www.atmos-chem-phys.net/11/12049/2011/>.
- Day, D.A., Liu, S., Russell, L.M., Ziemann, P.J., 2010. Organonitrate group concentrations in submicron particles with high nitrate and organic fractions in coastal southern California. *Atmospheric Environment* 44 (16), 1970–1979.
- de Gouw, J.A., Welsh-Bon, D., Warneke, C., Kuster, W.C., Alexander, L., Baker, A.K., Beyersdorf, A.J., Blake, D.R., Canagaratna, M., Celada, A.T., Huey, L.G., Junkermann, W., Onasch, T.B., Salcido, A., SJostedt, S.J., Sullivan, A.P., Tanner, D.J., Vargas, O., Weber, R.J., Worsnop, D.R., Yu, X.Y., Zaveri, R., 2009. Emission and chemistry of organic carbon in the gas and aerosol phase at a sub-urban site near Mexico city in march 2006 during the Milagro study. *Atmospheric Chemistry and Physics* 9 (10), 3425–3442.
- Dockery, D.W., Pope, C.A., Xu, X.P., Spengler, J.D., Ware, J.H., Fay, M.E., Ferris, B.G., Speizer, F.E., Dec. 1993. An association between air-pollution and mortality in 6 United States cities. *New England Journal of Medicine* 329 (24), 1753–1759.
- Draxler, R., Rolph, G., 2010. Hysplit (Hybrid Single-particle Lagrangian Integrated Trajectory) Model Access via NOAA Arl Ready Website. <http://ready.arl.noaa.gov/hysplit.php>.
- Fine, P.M., Cass, G.R., Simoneit, B.R.T., Apr. 2002. Chemical characterization of fine particle emissions from the fireplace combustion of woods grown in the southern United States. *Environmental Science & Technology* 36 (7), 1442–1451.
- Fortner, E.C., Zheng, J., Zhang, R., Knighton, W.B., Volkamer, R.M., Sheehy, P., Molina, L., Andre, M., 2009. Measurements of volatile organic compounds using proton transfer reaction – mass spectrometry during the Milagro 2006 campaign. *Atmospheric Chemistry and Physics* 9 (2), 467–481.

- Frossard, A.A., Shaw, P.M., Russell, L.M., Kroll, J.H., Canagaratna, M.R., Worsnop, D.R., Quinn, P.K., Bates, T.S., 2011. Springtime arctic haze contributions of submicron organic particles from European and Asian combustion sources. *Journal of Geophysical Research-atmospheres* 116.
- Garnes, L.A., Allen, D.T., Oct. 2002. Size distributions of organonitrates in ambient aerosol collected in Houston, Texas. *Aerosol Science and Technology* 36 (10), 983–992.
- Ge, X., Wexler, A.S., Clegg, S.L., Jan. 2011. Atmospheric amines – part i. A review. *Atmospheric Environment* 45 (3), 524–546.
- Hamilton, J.F., Webb, P.J., Lewis, A.C., Hopkins, J.R., Smith, S., Davy, P., Aug. 2004. Partially oxidised organic components in urban aerosol using GCXGC-TOF/MS. *Atmospheric Chemistry and Physics* 4, 1279–1290.
- Hawkins, L.N., Russell, L.M., 2010. Oxidation of ketone groups in transported biomass burning aerosol from the 2008 northern California lightning series fires. *Atmospheric Environment* 44 (34), 4142–4154.
- Hennigan, C.J., Sullivan, A.P., Collett, Jeffrey, L.J., Robinson, A.L., May 2010. Levoglucosan stability in biomass burning particles exposed to hydroxyl radicals. *Geophysical Research Letters* 37, L09806.
- Henry, R.C., 2003. Multivariate receptor modeling by n-dimensional edge detection. *Chemometrics and Intelligent Laboratory Systems* 65 (2), 179–189.
- Jayne, J.T., Leard, D.C., Zhang, X.F., Davidovits, P., Smith, K.A., Kolb, C.E., Worsnop, D.R., 2000. Development of an Aerosol Mass Spectrometer for size and composition analysis of submicron particles. *Aerosol Science and Technology* 33 (1–2), 49–70.
- Jimenez, J.L., Canagaratna, M.R., Donahue, N.M., Prevot, A.S.H., Zhang, Q., Kroll, J.H., DeCarlo, P.F., Allan, J.D., Coe, H., Ng, N.L., Aiken, A.C., Docherty, K.S., Ulbrich, I.M., Grieshop, A.P., Robinson, A.L., Duplissy, J., Smith, J.D., Wilson, K.R., Lanz, V.A., Hueglin, C., Sun, Y.L., Tian, J., Laaksonen, A., Raatikainen, T., Rautiainen, J., Vaattovaara, P., Ehn, M., Kulmala, M., Tomlinson, J.M., Collins, D.R., Cubison, M.J., Dunlea, E.J., Huffman, J.A., Onasch, T.B., Alfarra, M.R., Williams, P.L., Bower, K., Kondo, Y., Schneider, J., Drewnick, F., Borrmann, S., Weimer, S., Demerjian, K., Salcedo, D., Cottrell, L., Griffin, R., Takami, A., Miyoshi, T., Hatakeyama, S., Shimono, A., Sun, J.Y., Zhang, Y.M., Dzepina, K., Kimmel, J.R., Sueper, D., Jayne, J.T., Herndon, S.C., Trimborn, A.M., Williams, L.R., Wood, E.C., Middlebrook, A.M., Kolb, C.E., Baltensperger, U., Worsnop, D.R., Dec. 2009. Evolution of organic aerosols in the atmosphere. *Science* 326 (5959), 1525–1529.
- Jobson, B.T., Alexander, M.L., Maupin, G.D., Muntean, G.G., Aug. 2005. On-line analysis of organic compounds in diesel exhaust using a proton transfer reaction mass spectrometer (PTR-MS). *International Journal of Mass Spectrometry* 245 (1–3), 78–89.
- Kanakidou, M., Seinfeld, J.H., Pandis, S.N., Barnes, I., Dentener, F.J., Facchini, M.C., Van Dingenen, R., Ervens, B., Nenes, A., Nielsen, C.J., Swietlicki, E., Putaud, J.P., Balkanski, Y., Fuzzi, S., Horth, J., Moortgat, G.K., Winterhalter, R., Myhre, C.E.L., Tsigaridis, K., Vignati, E., Stephanou, E.G., Wilson, J., 2005. Organic aerosol and global climate modelling: a review. *Atmospheric Chemistry and Physics* 5, 1053–1123.
- Laskin, A., Smith, J.S., Laskin, J., May 2009. Molecular characterization of nitrogen-containing organic compounds in biomass burning aerosols using high-resolution mass spectrometry. *Environmental Science & Technology* 43 (10), 3764–3771.
- Liggio, J., Li, S.-M., Vlasenko, A., Sjostedt, S., Chang, R., Shantz, N., Abbatt, J., Slowik, J.G., Bottenheim, J.W., Brickell, P.C., Stroud, C., Leaitch, W.R., Nov. 2010. Primary and secondary organic aerosols in urban air masses intercepted at a rural site. *Journal of Geophysical Research-atmospheres* 115, D21305.
- Lindinger, W., Hansel, A., Jordan, A., Sep. 1998. Proton-transfer-reaction mass spectrometry (PTR-MS): on-line monitoring of volatile organic compounds at PPTV levels. *Chemical Society Reviews* 27 (5), 347–354.
- Liu, S., Day, D.A., Shields, J.E., Russell, L.M., 2011. Ozone-driven daytime formation of secondary organic aerosols containing carboxylic acid groups and alkane groups. *Atmospheric Chemistry and Physics* 11 (16), 8321–8341. URL: <http://www.atmos-chem-phys.net/11/8321/2011/>.
- Liu, S., Takahama, S., Russell, L.M., Gilardoni, S., Baumgardner, D., 2009. Oxygenated organic functional groups and their sources in single and submicron organic particles in Milagro 2006 campaign. *Atmospheric Chemistry and Physics* 9 (18), 6849–6863.
- Ng, N.L., Canagaratna, M.R., Zhang, Q., Jimenez, J.L., Tian, J., Ulbrich, I.M., Kroll, J.H., Docherty, K.S., Chhabra, P.S., Bahreini, R., Murphy, S.M., Seinfeld, J.H., Hildebrandt, L., Donahue, N.M., DeCarlo, P.F., Lanz, V.A., Prevot, A.S.H., Dinar, E., Rudich, Y., Worsnop, D.R., 2010. Organic aerosol components observed in northern hemispheric datasets from aerosol mass spectrometry. *Atmospheric Chemistry and Physics* 10 (10), 4625–4641.
- Ng, N.L., Canagaratna, M.R., Jimenez, J.L., Zhang, Q., Ulbrich, I.M., Worsnop, D.R., Feb. 2011a. Real-time methods for estimating organic component mass concentrations from Aerosol Mass Spectrometer data. *Environmental Science & Technology* 45 (3), 910–916.
- Ng, N.L., Herndon, S.C., Trimborn, A., Canagaratna, M.R., Croteau, P.L., Onasch, T.B., Sueper, D., Worsnop, D.R., Zhang, Q., Sun, Y.L., Jayne, J.T., 2011b. An aerosol chemical speciation monitor (ACSM) for routine monitoring of the composition and mass concentrations of ambient aerosol. *Aerosol Science and Technology* 45 (7). pii: 934555189.
- Paatero, P., Tapper, U., 1994. Positive matrix factorization – a nonnegative factor model with optimal utilization of error-estimates of data values. *Environmetrics* 5 (2), 111–126.
- Russell, L.M., 2003. Aerosol organic-mass-to-organic-carbon ratio measurements. *Environmental Science & Technology* 37 (13), 2982–2987.
- Russell, L.M., Bahadur, R., Ziemann, P.J., 2011. Identifying organic aerosol sources by comparing functional group composition in chamber and atmospheric particles. *Proceedings of the National Academy of Sciences of the United States of America* 108 (9), 3516–3521.
- Russell, L.M., Hawkins, L.N., Frossard, A.A., Quinn, P.K., Bates, T.S., 2010. Carbohydrate-like composition of submicron atmospheric particles and their production from ocean bubble bursting. *Proceedings of the National Academy of Sciences of the United States of America* 107 (15), 6652–6657.
- Russell, L.M., Takahama, S., Liu, S., Hawkins, L.N., Covert, D.S., Quinn, P.K., Bates, T.S., 2009. Oxygenated fraction and mass of organic aerosol from direct emission and atmospheric processing measured on the R/V Ronald Brown during TEXAQS/GoMACCS 2006. *Journal of Geophysical Research-atmospheres* 114.
- Schwartz, R.E., Russell, L.M., Sjostedt, S.J., Vlasenko, A., Slowik, J.G., Abbatt, J.P.D., Macdonald, A.M., Li, S.M., Liggio, J., Toom-Sauntry, D., Leaitch, W.R., 2010. Biogenic oxidized organic functional groups in aerosol particles from a mountain forest site and their similarities to laboratory chamber products. *Atmospheric Chemistry and Physics* 10 (11), 5075–5088.
- Schwarz, J.P., Spackman, J.R., Gao, R.S., Perring, A.E., Cross, E., Onasch, T.B., Ahern, A., Wrobleski, W., Davidovits, P., Olfert, J., Dubey, M.K., Mazzoleni, C., Fahey, D.W., 2010. The detection efficiency of the single particle soot photometer. *Aerosol Science and Technology* 44 (8). pii: 924376370.
- Seinfeld, J.H., Pandis, S.N., 2006. *Atmospheric Chemistry and Physics*, second ed. John Wiley & Sons, New York.
- Shair, F., Sasaki, E., Carlan, D., Cass, G., Goodin, W., Edinger, J., Schacher, G., 1982. Transport and dispersion of airborne pollutants associated with the land breeze–sea breeze system. *Atmospheric Environment* (1967) 16 (9), 2043–2053. URL: <http://linkinghub.elsevier.com/retrieve/pii/000469818290275X>.
- Stöhr, J., 1992. *NEXAFS Spectroscopy*. Springer-Verlag, Berlin.
- Sullivan, A.P., Holden, A.S., Patterson, L.A., McMeeking, G.R., Kreidenweis, S.M., Malm, W.C., Hao, W.M., Wold, C.E., Collett, J.L., 2008. A method for smoke marker measurements and its potential application for determining the contribution of biomass burning from wildfires and prescribed fires to ambient PM_{2.5} organic carbon. *Journal of Geophysical Research-atmospheres* 113.
- Takahama, S., Liu, S., Russell, L.M., 2010. Coatings and clusters of carboxylic acids in carbon-containing atmospheric particles from spectromicroscopy and their implications for cloud-nucleating and optical properties. *Journal of Geophysical Research-atmospheres* 115.
- Takahama, S., Pathak, R.K., Pandis, S.N., Apr. 2007. Efflorescence transitions of ammonium sulfate particles coated with secondary organic aerosol. *Environmental Science & Technology* 41 (7), 2289–2295.
- Takahama, S., Russell, L.M., Jan. 2011. A molecular dynamics study of water mass accommodation on condensed phase water coated by fatty acid monolayers. *Journal of Geophysical Research-atmospheres* 116, D02203.
- Takahama, S., Schwartz, R.E., Russell, L.M., Macdonald, A.M., Sharma, S., Leaitch, W.R., 2011. Organic functional groups in aerosol particles from burning and non-burning forest emissions at a high-elevation mountain site. *Atmospheric Chemistry and Physics* 11 (13), 6367–6386.
- Ulbrich, I., Lechner, M., Jimenez, J., 2012. AMS Spectral Database. URL: <http://cires.colorado.edu/jimenez-group/AMSSd>.
- Ulbrich, I.M., Canagaratna, M.R., Zhang, Q., Worsnop, D.R., Jimenez, J.L., 2009. Interpretation of organic components from positive matrix factorization of aerosol mass spectrometric data. *Atmospheric Chemistry and Physics* 9 (9), 2891–2918.
- Warneke, C., De Gouw, J.A., Kuster, W.C., Goldan, P.D., Fall, R., Jun. 2003. Validation of atmospheric VOC measurements by proton-transfer-reaction mass spectrometry using a gas-chromatographic preseparation method. *Environmental Science & Technology* 37 (11), 2494–2501.
- Zhang, Q., Alfarra, M.R., Worsnop, D.R., Allan, J.D., Coe, H., Canagaratna, M.R., Jimenez, J.L., Jul. 2005. Deconvolution and quantification of hydrocarbon-like and oxygenated organic aerosols based on aerosol mass spectrometry. *Environmental Science & Technology* 39 (13), 4938–4952.
- Zhang, Q., Jimenez, J.L., Canagaratna, M.R., Allan, J.D., Coe, H., Ulbrich, I., Alfarra, M.R., Takami, A., Middlebrook, A.M., Sun, Y.L., Dzepina, K., Dunlea, E., Docherty, K., DeCarlo, P.F., Salcedo, D., Onasch, T., Jayne, J.T., Miyoshi, T., Shimono, A., Hatakeyama, S., Takegawa, N., Kondo, Y., Schneider, J., Drewnick, F., Borrmann, S., Weimer, S., Demerjian, K., Williams, P., Bower, K., Bahreini, R., Cottrell, L., Griffin, R.J., Rautiainen, J., Sun, J.Y., Zhang, Y.M., Worsnop, D.R., Jul. 2007. Ubiquity and dominance of oxygenated species in organic aerosols in anthropogenically-influenced northern hemisphere midlatitudes. *Geophysical Research Letters* 34 (13), L13801.
- Zhao, J., Zhang, R.Y., May 2004. Proton transfer reaction rate constants between hydronium ion (H_3O^{+}) and volatile organic compounds. *Atmospheric Environment* 38 (14), 2177–2185.
- Zhao, J., Zhang, R.Y., Misawa, K., Shibuya, K., Dec. 2005. Experimental product study of the OH-initiated oxidation of m-xylene. *Journal of Photochemistry and Photobiology A-chemistry* 176 (1–3), 199–207.
- Zhao, Z., Chen, S.-H., Kleeman, M.J., Tyree, M., Cayen, D., Jul. 2011. The impact of climate change on air quality-related meteorological conditions in California. Part i: present time simulation analysis. *Journal of Climate* 24 (13), 3344–3361.

Supplemental Information for
“Submicron organic aerosol in Tijuana, Mexico, from local and
Southern California sources during the CalMex campaign”

Contents

S1 Dimension reduction methods	1
S1.1 PMF description	1
S1.2 Augmentation of the ACSM PMF standard deviation matrix	2
S1.3 Comparison of ACSM spectra classification analysis by LDA and k -nn algorithms	2
S1.4 Using classification analysis to approximate ensemble mixture proportions	3
S1.5 Additional methods for ACSM spectra analysis: SVD and regression analysis (CMB)	4
S2 Apportionment of OM	4
S2.1 Synthesis of FTIR and ACSM PMF solutions	4
S2.2 Mass fragment and VOC ratios	7
S2.2.1 Lag-time correlations	7
S2.2.2 Toluene-to-benzene ratios	7
S3 Meteorological analysis	7

¹ S1 Dimension reduction methods

² S1.1 PMF description

³ Given a column matrix of row vectors (i.e., the sample spectra), \mathbf{X} , the forward model is expressed as
⁴ $\mathbf{X} = \mathbf{G}\mathbf{F} + \mathbf{E}$. \mathbf{G} and \mathbf{F} are matrices comprising component strengths and profiles, respectively; \mathbf{E} is
⁵ the residual matrix. The Q (or χ^2) value to be minimized is defined by the canonical objective function
⁶ $Q = \sum_i \sum_j e_{ij}^2 / s_{ij}^2$, where e_{ij} are the residuals (elements of \mathbf{E}) and s_{ij} define the weighting for the fit.
⁷ These weights are derived from mechanistic estimates of the measurement error (Polissar et al., 1998) for
⁸ both FTIR (Russell et al., 2009b) and ACSM (Ng et al., 2011). The goodness-of-fit metric (Q) is evaluated
⁹ against a theoretical or expected value, which is approximated by the degrees of freedom in the system:
¹⁰ $Q_{expected} \approx \nu = m \cdot n - p(m + n)$ (Paatero et al., 2002). For data sets with large number of n variables (e.g.,
¹¹ FTIR spectra) or large number of m samples (e.g., ACSM), is essentially the number of elements in the data
¹² matrix, \mathbf{X} (Ulbrich et al., 2009). The assumption in this model is that \mathbf{X} is composed of a signal and noise,
¹³ and we wish to represent the signal with factor components and allow the residual term to carry the noise.
¹⁴ A commonly prescribed criterion for selecting a solution is that $Q/Q_{expected} \approx 1$, according to the premise
¹⁵ that \mathbf{S} represents the magnitude of noise component, and factor components are fitting only the true signal
¹⁶ such that $\mathbf{E} \approx \mathbf{S}$.

17 The four-factor FTIR solution discussed has a $Q/Q_{expected}$ value of 0.8 and is chosen as it satisfies the
18 criteria outlined by Russell et al. (2009b): the factor strengths are not strongly correlated ($|r| < 0.5$) and
19 the reconstructed X matrix adequately reproduces the original spectra (Explained Variation > 95%). Factor
20 components are presumably a mixture of compounds; source classes are inferred for each of these mixtures
21 by examining correlations with elemental tracers and confirmed by comparison to factor spectra obtained
22 from previous campaigns (Russell et al., 2011). For ACSM, we choose a two-factor solution under a modified
23 \mathbf{S} matrix (discussed in Section S1.2). The PMF analysis for both FTIR and ACSM is performed using the
24 PMF2 algorithm by Paatero and Tapper (1994). Exploration of parameters and evaluation of solutions are
25 performed using a set of scripts written in *R* (R Development Core Team, 2012).

26 **S1.2 Augmentation of the ACSM PMF standard deviation matrix**

27 If x = the measured signal (OM), u the true signal, and ϵ represents the instrument noise, a scalar repre-
28 sentation of their relationship can be written as

$$x = u + \epsilon . \quad (\text{S1})$$

29 Var(ϵ) is estimated from measurement errors compounded for ion counts converted to analog signals (Ng
30 et al., 2011). The PMF statement is equivalent to Equation (S1) except that ϵ represents the residuals
31 (fitting error), which is assumed to be approximately equal to the measurement error in magnitude for
32 the purpose of finding a suitable solution (Paatero and Tapper, 1994). Upon an initial iteration of PMF
33 decomposition, we find that the $Q/Q_{expected}$ does not converge to unity even for a large number of factors
34 (out to 12), suggesting that \mathbf{S} may be underestimated for this data set ($Q/Q_{expected}$ is approximately 3
35 for a four-factor solution). This underestimation leads to solutions in which some of the factors appear
36 to representing additional noise rather than an underlying component of the true signal. To remedy this
37 situation, we select a solution in which the model is purposely over-fitting the data such that the factors
38 with the lowest loadings are fitting the noise. These components are taken as estimates of the unaccounted
39 measurement error and append this to the original \mathbf{S} matrix. In formula, this is expressed as $x = u + \delta + \epsilon$.
40 In this solution, a factor component, δ , contributes little to explaining the systematic variation in x (median
41 correlation with any m/z signal ~ 0.1). Therefore, the measurement error matrix is augmented by assuming
42 that $\text{Var}(\delta) = \delta$ (limit of Poisson statistics) and deriving a new estimate of the standard deviation matrix.
43 In terms of the matrix formulation of PMF, the first decomposition yields $\mathbf{X} = (\mathbf{G}\mathbf{F})_u + (\mathbf{G}\mathbf{F})_\delta + \mathbf{E}$; the
44 standard deviation matrix \mathbf{S} is augmented as $\mathbf{S}_{(new)} = \sqrt{(\mathbf{G}\mathbf{F})_\delta^2 + \mathbf{S}^2}$. This reduces the Q -value from 3.9
45 to 1.7 for a two-factor solution, but more importantly, eliminates the generation of solutions in which a δ
46 term or terms are included (i.e., components appear to represent true signals). Residuals (e_{ij}) normalized
47 by the corresponding standard deviation are shown in Figure (S1).

48 **S1.3 Comparison of ACSM spectra classification analysis by LDA and k -nn 49 algorithms**

50 Linear discriminant analysis (LDA) is a method for determining linear decision boundaries (hyperplanes) for
51 multidimensional, continuous variables, to delineate regions of observations belonging to a particular class
52 or category. The k -nearest neighbor (k -nn) method assigns categories to multivariate observations based on
53 proximity (the Euclidean distance metric is most commonly used) to elements in the training set. Details of
54 these statistical learning methods are described by (Hastie et al., 2009).

55 As described in Section 2, ACSM mass fragment spectra are compared to a training set of unit mass
56 resolution mass spectra from AMS (Ulbrich et al., 2009, 2012). As a projection of the measurements in the
57 space of normalized mass fragment concentration pairs (including $f_{43}, f_{44}, f_{57}, f_{60}$; each plotted against f_{44} is
58 shown in Figure 7) indicates that a segregation of the training set naturally lends itself to a partitioning of the
59 composition space such a method as linear discriminant analysis (LDA), we can obtain a first-order estimate
60 of the types of samples present. Given the number of assumptions required for LDA (equal covariances,
61 multi-normality) that are difficult to assess for the small number of samples in the training set, k -nearest

62 neighbor (k -nn) classification is also used to provide another method of classification subject to different a set
 63 of constraint criteria. We use a reduced set for the classification based on our understanding that m/z 43, 44,
 64 57, and 60 are the important mass fragments that differentiate among OOA, HOA, and BBOA aerosol (Ng
 65 et al., 2011) and avoid using the full spectrum where collinearity may inflate the errors in the classification.
 66 We apply a square-root transformation of the feature vector so that normal distributions are approximated
 67 by the data, and scale each vector by its 2-norm distance. Both LDA and k -nn classifications are performed
 68 in R (R Development Core Team, 2012) using the MASS and class libraries, respectively. Figure S2 shows
 69 the similarities in classified fractions according to LDA and k -nn analysis.

70 S1.4 Using classification analysis to approximate ensemble mixture proportions

71 Let us assume the existence of two end-member states, M_1 and M_2 . x represents the proportion of component
 72 M_1 in a sample, obtained from feature vector \mathbf{v} of an arbitrary number of dimensions using a mapping
 73 represented by g_{M_1} . If we allow \mathcal{M}_1 and \mathcal{M}_2 to denote categories with approximately similar composition
 74 to their respective end members, we can also define an indicator function $I_{\mathcal{M}_1}(\mathbf{v})$:

$$I_{\mathcal{M}_1}(\mathbf{v}) = \begin{cases} 1 & \text{if } \mathbf{v} \in \mathcal{M}_1 \\ 0 & \text{if } \mathbf{v} \notin \mathcal{M}_1 \end{cases}.$$

75 According to this formulation, the result of classification analysis is to alternatively represent the feature
 76 vector for each sample with a discrete (binary) value, x' , rather than a continuous variable, x :

$$\begin{aligned} x &= g_{M_1}(\mathbf{v}) \\ x' &= I_{\mathcal{M}_1}(\mathbf{v}) . \end{aligned}$$

77 The expected value of X' (a Bernoulli random variable) can be shown to approximate the ensemble average
 78 of X (a continuous random variable) according to the indicator function. That is to say that we approximate
 79 $E(X)$ with $E(X')$ by replacing x with x' for each sample i :

$$\frac{1}{n} \sum_i^n g_{M_1}(\mathbf{v}_i) \approx \frac{1}{n} \sum_i^n I_{\mathcal{M}_1}(\mathbf{v}_i) = \frac{n_{\mathcal{M}_1}}{n}, \quad (\text{S2})$$

80 where n is the total number of samples and $n_{\mathcal{M}_1}$ is the number of samples classified into category \mathcal{M}_1 . The
 81 right-hand side of Equation (S2) follows from the fact that the estimator for $E(X')$ is $\hat{p} = n_{\mathcal{M}_1}/n$. This
 82 concept can be extended to extended to N exhaustive end-member states by specification of complementary
 83 indicator functions such that

$$\sum_{k=1}^N I_{\mathcal{M}_k}(\mathbf{v}_i) = 1 \quad \forall i = \{1, \dots, n\} .$$

84 To fix ideas in a simple, one-dimensional case with two end members, let $x = \mathbf{v}$ and g_{M_1} be the identity
 85 function. We define the indicator function in terms of a threshold value, $\phi_{\mathcal{M}_1}$, such that $I_{\mathcal{M}_1}(x) = \Phi(-x +$
 86 $\phi_{\mathcal{M}_1})$. Φ is the Heaviside unit step function; $\phi_{\mathcal{M}_1}$ is the threshold below which x is categorized as \mathcal{M}_1 (and
 87 above as \mathcal{M}_2). The probabiliy distribution of the Bernoulli random variable, X' , is related to the continuous
 88 distribution of the original variable such that

$$E(X') = \Pr(X' = 1) = \Pr(X \in \mathcal{M}_1) = \Pr(X \leq \phi_{\mathcal{M}_1})$$

89 since the empirical cumulative distribution function (c.d.f.) of X also has the form $\Pr(X \leq \phi) = \Phi(-x + \phi)$
 90 where ϕ is any value in the domain of X , which is [0,1]. Markov's inequality (Wasserman, 2010) can
 91 approximately constrain the complementary c.d.f. for a continuous distribution from its expected value.
 92 Since in our example the c.d.f. is equal to $E(X')$, this inequality can give us some indication of how $E(X)$
 93 and $E(X')$ are related in this specific case. If we let $\Delta = |E(X') - E(X)|$, then $\Delta \leq E(X') [1 + \phi_{\mathcal{M}_1}] - \phi_{\mathcal{M}_1}$,

94 and we can bound the magnitude of the approximation (between $E(X)$ and $E(X')$) for choice of $\phi_{\mathcal{M}_1}$ and
 95 estimated values of $E(X')$ (and also minimize Δ with respect to $\phi_{\mathcal{M}_1}$). While many classification algorithms
 96 preclude definition of a simple threshold and each estimation method may require different preparation of
 97 feature vectors to satisfy required assumptions, this illustration shows a very simple case in which discrete
 98 categorizations can be used to approximate mixture proportions. In the more general case, if we define
 99 $d = I_{\mathcal{M}_1}(\mathbf{v}) - g_{\mathcal{M}_1}(\mathbf{v})$, then $\Delta = |E(D)|$ (from Equation S2). Therefore, we expect that the bias in estimation
 100 between the two methods would depend on the distribution of D .

101 It should be noted that the correctness of the approximation strongly depends on the classification
 102 scheme which defines the indicator functions. Therefore, the solution of 88/12% OOA/HOA (Section S1.5)
 103 rather than 60/40% OOA-1/OOA-2 (Section 3.2) apportionment cannot be ruled out, especially if the
 104 feature vectors or classification algorithm cannot fully discriminate between HOA and OOA-2 if they are
 105 approximately collinear along a continuum of oxidation states. But as stated in Section 3.2, this partitioning
 106 affects the estimation of the more and less oxygenated fractions, but their respective trends are robust with
 107 respect to either interpretation. Therefore, the source association of the components based on the diurnal
 108 variations showed in Figure 8 remains unchanged.

109 **S1.5 Additional methods for ACSM spectra analysis: SVD and regression anal- 110 ysis (CMB)**

111 Singular Value Decomposition (SVD) on the sample matrix \mathbf{X} (which does not account for the measure-
 112 ment error matrix, \mathbf{S}) indicates that two components explain approximately 80% of the variation in the
 113 measurements, consistent with an Explained Variation of 80% from PMF analysis for a two-component
 114 solution.

115 Regression analysis, often referred to as a Chemical Mass Balance (CMB) approach (Chow and Watson,
 116 2002; Ng et al., 2011), to component apportionment is also used but yielded results not consistent with the
 117 domain delineated by Figure 7 because of the large collinearity among regressands. Approximate apportion-
 118 ment to HOA and OOA (HOA_{est} and OOA_{est} , respectively) from unit-mass resolution AMS measurements
 119 using OM-equivalent concentrations of mass fragments 44 and 57 (C_{44} and C_{57} , respectively) is suggested
 120 by Ng et al. (2011):

$$\begin{pmatrix} HOA_{est} \\ OOA_{est} \end{pmatrix} = \begin{pmatrix} 0 & b & a \cdot b \\ d & 0 & c \end{pmatrix} \begin{pmatrix} 1 \\ C_{57} \\ C_{44} \end{pmatrix}. \quad (\text{S3})$$

121 Using Equation (S3) with coefficients $\{a = 0.095, b = 15.2, c = 6.92, d = 0.07\}$ estimated from median
 122 values of various campaigns (Ng et al., 2011) provides an alternate method of partitioning the OM into
 123 more and less oxygenated fractions. The HOA and OOA apportioned using this method correlate well
 124 with OOA-2 ($r=0.86$) and OOA-1 ($r=0.77$), respectively. Therefore, conclusions dependent on the relative
 125 variation of the less oxygenated and more oxygenated components in time are robust with respect to either
 126 interpretation. However, the regression coefficients $\beta=\{\text{intercept},\text{slope}\}$ are $\{-0.25,0.46\}$ for HOA_{est} on OOA-
 127 2, and $\{-0.5,1.8\}$ for OOA_{est} on OOA-1. This alternative interpretation would suggest that the transported
 128 fraction to OM may be as high as 88% – rather than 60% as suggested using the OOA-1 and OOA-2 PMF
 129 solution – based on the justified assumption that the majority of the more oxygenated fraction of OM is
 130 due to long-range transport (Section 3.3). However, more weight is placed on the OOA-1 and OOA-2 PMF
 131 solution as it is derived for measurements at the Tijuana site specifically, and the solution is supported by
 132 the chemical composition space delineated by Figure 7 and classification analysis discussed in section S1.3.

133 **S2 Apportionment of OM**

134 **S2.1 Synthesis of FTIR and ACSM PMF solutions**

135 As the FTIR and ACSM provides complementary information regarding the organic fraction of ambient
 136 aerosols, we wish to harmonize our interpretation of our measurements by combining the information re-

137 trieval from both instruments. One method is to combine measurements from both instruments into a
 138 single data matrix for multivariate PMF analysis, which poses many challenges. For instance, the num-
 139 ber of variables provided by each instrument are different; appropriate downweighting must be applied to
 140 reduce redundancies for each instrument, and for one set of measurement with respect to the other such
 141 that one is not favored more heavily in the solution. These issues are effectively handled through the ad-
 142 justment of the standard deviation matrix; a bottom-up approach to specification of this matrix, including
 143 the measurement-uncertainty component, requires absolute calibration such that they are comparable across
 144 instruments. Slowik et al. (2010) combined measurements from an AMS and PTR-MS to find covarying
 145 mass-fragments from PMF analysis, and iteratively adjusted the elements of the standard deviation matrix
 146 (collectively by instrument) with a single scaling factor to obtain empirical weightings such that mass frag-
 147 ments from both instruments were well represented and reproduced by the solutions. In this way, Slowik
 148 et al. (2010) retained relative weights (due to measurement uncertainties) for each instrument, but adjusted
 149 overall according to their information content with respect to the other instrument. Combining absorbance
 150 and mass fragment spectra for PMF analysis could possibly be approached in this way to ensure that solu-
 151 tions represent apportionment based on variation, rather than (possibly miscalibrated) variance. However,
 152 in this work, we adopt a naive approach which provides an interpretation of the two measurements after
 153 PMF decomposition has been applied independently.

154 Taking a rather extreme perspective, we assume FTIR PMF components retain information regarding
 155 source class (Russell et al., 2011) and ACSM PMF components indicate information regarding aerosol age
 156 (Jimenez et al., 2009). From this analysis we attempt to approximate age of OM attributed to each source
 157 class (from which notions of local and remote origins emerge). Let OM_A and OM_F represent OM concentra-
 158 tions measured by ACSM and FTIR, respectively, and OM_C represent the OM concentration measured by
 159 both techniques (for a single sample). Disregarding non-systematic disturbances to each set of measurements
 160 (PMF presumably retains the signal to factor components and apportions these disturbances to the residual
 161 term), it is assumed that the primary differences among these metrics of OM are the unmeasured compounds
 162 or organic aerosol fractions due to reasons described in Section 3.2. α is a vector which indicates the frac-
 163 tion of OM attributed to components $\mathcal{A} = \{\text{OOA-1}, \text{OOA-2}\}$ from ACSM PMF analysis, or components
 164 $\mathcal{F} = \{\text{FF1}, \text{FF2}, \text{BB}\}$ from FTIR PMF analysis (as denoted by subscripts). The Marine factor is excluded
 165 from this portion of the analysis as we expect marine organic OM associated with non-refractory sea salt
 166 not to be measured by the ACSM (Section 3.2). We additionally define a vector of all ones, $\mathbf{1}_m$, where m
 167 indicates the number of elements (often defined here by the cardinality of a set, $|\cdot|$, or number of samples,
 168 n). Using these definitions, OM_C can be defined in terms of the fractional sum of components for each
 169 instrument: $OM_C = OM_A \alpha_{\mathcal{A}}^T \mathbf{1}_{|\mathcal{A}|} = OM_F \alpha_{\mathcal{F}}^T \mathbf{1}_{|\mathcal{F}|}$. In this context, $\mathbf{1}_m$ is introduced as a postfix operator
 170 which sums elements of a vector (or rows or columns of matrices, depending on orientation). In our present
 171 specification, $\alpha_{\mathcal{A}}^T \mathbf{1}_{|\mathcal{A}|}$ and $\alpha_{\mathcal{F}}^T \mathbf{1}_{|\mathcal{F}|}$ are not required to equal unity, as the marine component for FTIR has
 172 been excluded from $\alpha_{\mathcal{F}}$, for instance. We can eliminate the explicit references to OM by redefining $\alpha_{\mathcal{A}}$ and
 173 $\alpha_{\mathcal{F}}$ with scaling factors OM_A/OM_C and OM_F/OM_C , respectively, such that

$$1 = \alpha_{\mathcal{A}}^T \mathbf{1}_{|\mathcal{A}|} = \alpha_{\mathcal{F}}^T \mathbf{1}_{|\mathcal{F}|} \quad (\text{S4})$$

174 The proportion of OM_C attributed to any component from one measurement can be described with respect
 175 to contributions from the other, which follows from Equation (S4):

$$\begin{aligned} \alpha_{\mathcal{A}} &= \alpha_{\mathcal{A}} \alpha_{\mathcal{F}}^T \mathbf{1}_{|\mathcal{F}|} \\ \alpha_{\mathcal{F}} &= \alpha_{\mathcal{F}} \alpha_{\mathcal{A}}^T \mathbf{1}_{|\mathcal{A}|}. \end{aligned} \quad (\text{S5})$$

176 The two outer products, $\alpha_{\mathcal{A}} \alpha_{\mathcal{F}}^T$ and $\alpha_{\mathcal{F}} \alpha_{\mathcal{A}}^T$, define the fraction of FTIR components associated with ACSM
 177 components, and vice versa. For instance, the former case (Equation S5) is also written,

$$\begin{pmatrix} \alpha_{\text{OOA-1}} \\ \alpha_{\text{OOA-2}} \end{pmatrix} = \begin{pmatrix} \alpha_{\text{OOA-1}} \alpha_{\text{FF1}} & \alpha_{\text{OOA-1}} \alpha_{\text{FF2}} & \alpha_{\text{OOA-1}} \alpha_{\text{BB}} \\ \alpha_{\text{OOA-2}} \alpha_{\text{FF1}} & \alpha_{\text{OOA-2}} \alpha_{\text{FF2}} & \alpha_{\text{OOA-2}} \alpha_{\text{BB}} \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad (\text{S6})$$

178 where the first matrix on the right-hand side is $\boldsymbol{\alpha}_{\mathcal{A}}\boldsymbol{\alpha}_{\mathcal{F}}^T$. Campaign-averaged contributions for $OM_C\boldsymbol{\alpha}_{\mathcal{A}}\boldsymbol{\alpha}_{\mathcal{F}}^T$ are
 179 shown in Figure S3. Averaged quantities of $\boldsymbol{\alpha}_{\mathcal{A}}\boldsymbol{\alpha}_{\mathcal{F}}^T$ and $\boldsymbol{\alpha}_{\mathcal{F}}\boldsymbol{\alpha}_{\mathcal{A}}^T$ are essentially a measure of the uncentered
 180 covariance scaled by individual measurements of OM. To illustrate this interpretation, let us define two
 181 matrices, $\mathbf{A}_{\mathcal{A}}$ and $\mathbf{A}_{\mathcal{F}}$ for ACSM and FTIR measurements, respectively, which contain all rescaled values of
 182 $\boldsymbol{\alpha}$ for the campaign:

$$183 \quad \mathbf{A}_{\mathcal{A}} = (\boldsymbol{\alpha}_{\mathcal{A},1} \ \boldsymbol{\alpha}_{\mathcal{A},2} \ \dots \ \boldsymbol{\alpha}_{\mathcal{A},n}) = \begin{pmatrix} \boldsymbol{\alpha}_{\text{OOA-1}}^T \\ \boldsymbol{\alpha}_{\text{OOA-2}}^T \end{pmatrix} \quad \text{and} \quad \mathbf{A}_{\mathcal{F}} = (\boldsymbol{\alpha}_{\mathcal{F},1} \ \boldsymbol{\alpha}_{\mathcal{F},2} \ \dots \ \boldsymbol{\alpha}_{\mathcal{F},n}) = \begin{pmatrix} \boldsymbol{\alpha}_{\text{FF1}}^T \\ \boldsymbol{\alpha}_{\text{FF2}}^T \\ \boldsymbol{\alpha}_{\text{BB}}^T \end{pmatrix}.$$

184 Each matrix of \mathbf{A} is defined in two equivalent representations, either as 1) vectors containing component
 185 fractions, $\boldsymbol{\alpha}_{\mathcal{A},i}$ or $\boldsymbol{\alpha}_{\mathcal{F},i}$ – with second indices indicating the sample number – or 2) vectors containing OM
 186 mass fractions for each PMF component in \mathcal{A} or \mathcal{F} , with number of dimensions equal to the number of
 187 measurements, n , in the campaign. We adopt the second notation for illustrating similarities to the array
 188 notation of $\boldsymbol{\alpha}_{\mathcal{A}}\boldsymbol{\alpha}_{\mathcal{F}}^T$ displayed in Equation (S6). We can restate the equality in OM reconstruction for each
 189 sample (Equation S4):

$$\mathbf{1}_n = \mathbf{A}_{\mathcal{A}}^T \mathbf{1}_{|\mathcal{A}|} = \mathbf{A}_{\mathcal{F}}^T \mathbf{1}_{|\mathcal{F}|},$$

190 and define their averages:

$$191 \quad \frac{1}{n} \mathbf{A}_{\mathcal{A}} \mathbf{1}_n = \frac{1}{n} \mathbf{A}_{\mathcal{A}} \mathbf{A}_{\mathcal{F}}^T \mathbf{1}_{|\mathcal{F}|} \quad (\text{S7})$$

$$192 \quad \frac{1}{n} \mathbf{A}_{\mathcal{F}} \mathbf{1}_n = \frac{1}{n} \mathbf{A}_{\mathcal{F}} \mathbf{A}_{\mathcal{A}}^T \mathbf{1}_{|\mathcal{A}|}. \quad (\text{S8})$$

193 A product of $\mathbf{A}_{\mathcal{A}}$ and $\mathbf{A}_{\mathcal{F}}$ will yield uncentered covariances across components (scaled by the number of
 194 observations, n):

$$195 \quad \frac{1}{n} \mathbf{A}_{\mathcal{A}} \mathbf{A}_{\mathcal{F}}^T = \frac{1}{n} \begin{pmatrix} \boldsymbol{\alpha}_{\text{OOA-1}}^T \boldsymbol{\alpha}_{\text{FF1}} & \boldsymbol{\alpha}_{\text{OOA-1}}^T \boldsymbol{\alpha}_{\text{FF2}} & \boldsymbol{\alpha}_{\text{OOA-1}}^T \boldsymbol{\alpha}_{\text{BB}} \\ \boldsymbol{\alpha}_{\text{OOA-2}}^T \boldsymbol{\alpha}_{\text{FF1}} & \boldsymbol{\alpha}_{\text{OOA-2}}^T \boldsymbol{\alpha}_{\text{FF2}} & \boldsymbol{\alpha}_{\text{OOA-2}}^T \boldsymbol{\alpha}_{\text{BB}} \end{pmatrix}.$$

196 Russell et al. (2009a) explored linear relationships between fragments and bonds measured by AMS
 197 and FTIR, respectively, through ordinary regression. We consider this approach for explaining variations
 198 in $\boldsymbol{\alpha}$ s in \mathcal{A} through linear combinations of $\boldsymbol{\alpha}$ s in \mathcal{F} . We extend individual statements of regression to
 199 a multivariate, multiple regression expression, $\mathbf{A}_{\mathcal{A}} = \boldsymbol{\beta}^T \mathbf{A}_{\mathcal{F}} + \mathbf{E}$, where \mathbf{E} is the residual matrix. The
 200 expected value of $\mathbf{A}_{\mathcal{A}}$ can be written as $\hat{\mathbf{A}}_{\mathcal{A}} = \mathbf{A}_{\mathcal{A}} \mathbf{H}^T$. A possible solution for the regression coefficients
 201 are $\hat{\boldsymbol{\beta}} = (\mathbf{A}_{\mathcal{F}} \mathbf{A}_{\mathcal{F}}^T)^{-1} \mathbf{A}_{\mathcal{F}} \mathbf{A}_{\mathcal{A}}^T$, and the hat (projection) matrix can be defined as $\mathbf{H} = \mathbf{A}_{\mathcal{F}}^T \hat{\boldsymbol{\beta}} (\mathbf{A}_{\mathcal{A}}^T)^{-1} =$
 202 $\mathbf{A}_{\mathcal{F}}^T (\mathbf{A}_{\mathcal{F}} \mathbf{A}_{\mathcal{F}}^T)^{-1} \mathbf{A}_{\mathcal{F}}$. Of course, a naive solution is proposed for illustration in this case, but non-negativity
 203 should be considered in the actual specification of $\hat{\boldsymbol{\beta}}$ and \mathbf{H} . To summarize the two approaches based on
 204 Equation S8 and Russell et al. (2009a), we can postfix $\mathbf{A}_{\mathcal{A}} = (\boldsymbol{\alpha}_{\text{OOA-1}}^T)$ with either $\frac{1}{n} \mathbf{A}_{\mathcal{F}} \mathbf{1}_{|\mathcal{F}|}$ to get $(\hat{\boldsymbol{\alpha}}_{\text{OOA-1}}^T)$,
 205 or with \mathbf{H}^T to get $(\hat{\boldsymbol{\alpha}}_{\text{OOA-2}}^T)$. The former statement makes a remark regarding the marginal expectation
 206 $E(X)$ from the joint expectation $E(X, Y)$, and the latter, the conditional expectation $E(X|Y = y)$, where X
 207 and Y are used to denote any pair of covariates in standard scalar notation. The relationship between $\mathbf{A}_{\mathcal{A}}$
 208 and $\mathbf{A}_{\mathcal{F}}$ are embodied in $\mathbf{A}_{\mathcal{A}} \mathbf{A}_{\mathcal{F}}^T$ in the former case, and $\hat{\boldsymbol{\beta}}$ (matrix of regression coefficients) for the latter
 209 case.

210 Since the separation between FTIR FF1 and FF2 are thought to be in degree of oxygenation (Russell
 211 et al., 2011), the expectation is that FF1 and FF2 are less and more aged, respectively, which is not supported
 212 by this analysis if OOA-1 and OOA-2 are strictly interpreted as indicators of age in Tijuana. It is therefore
 213 possible that there are other factors which lead to the separation of ACSM OM into OOA-1 and OOA-
 214 2, and FTIR OM into FF1 and FF2 at this location. For instance, there may be aspects of molecular
 215 composition that are similar (from the perspective of the ACSM mass spectra) across different source types,
 216 and therefore an overlap in functional groups apportioned to OOA-1 and OOA-2. Using the same method

(Equation ??) but letting α_F represent FTIR OFG fractions rather than PMF factor fractions, we can estimate OFG contributions (sans marine OM) to the different ACSM PMF factors and provide support for this interpretation (Figure S4). As discussed in Section 3.2, another interpretation (not mutually exclusive with respect to the previous statement) is that some fraction of FF2 is associated with non-refractory mineral dust (Section 3.2). OM associated with this material is not believed to be sampled by ACSM. Regarding the separation of other components (e.g., BB) between OOA-1 and OOA-2 fractions, the coarse time resolution of the FTIR measurements may also affect the ability to resolve component contributions well (Henry, 2003).

Liggio et al. (2010) apportioned HOA to primary (local) OM, and OOA to a combination of aged background OM and locally-produced SOA. In our case, we conclude that local production of SOA was small at our measurement site, based on estimates of diurnal increases in OM during photochemically active periods (approximately noon each day) and consistently high degree of oxygenation (Sections 3.2 and 3.3). With similar reasoning, we consider our less oxygenated aerosol (OOA-2) to local aerosol and more oxygenated (OOA-1) as aged. As a zeroth-order estimate, we disregard the separation in fossil fuel combustion OM by FTIR PMF into FF1 and FF2 and consider this as a single source component, and evaluate the local and regional contributions based on the degree of oxygenation of mass fragment spectra (Jimenez et al., 2009). In this case, 60% of the FF OM is associated with OOA-1 (50% if we disregard FF2 contribution to OM measured by FTIR in accordance with our previously stated hypothesis); this value may be considered a lower bound on the average contribution of anthropogenic combustion-related material transported to Tijuana (and oxygenated in the process) from regional sources (Section 3.3).

S2.2 Mass fragment and VOC ratios

S2.2.1 Lag-time correlations

The hourly lag time correlations between ACSM mass fragments and selected VOCs from PTR-MS measurements, as discussed in Section 3.3, are shown in Figure S5.

S2.2.2 Toluene-to-benzene ratios

The toluene-to-benzene ratio, $R_{[T/B]}$, is used as a metric of age and a parameter by which oxygenated aerosol is fractioned into primary and secondary, and background OM (e.g., Liggio et al., 2010). While requirements for chemical specificity precludes apportionment using ACSM mass spectra, correlations of $R_{[T/B]}$ with O/C ratio, and OOA-1 and OOA-2 normalized by OM (Figure S6) suggests that constant and point-source emission assumption (Liggio et al., 2010) required for $R_{[T/B]}$ analysis are not likely to be valid. Despite the constant wind direction/ direction of origin of airmasses to Tijuana, it may be the case that that mixing of airmasses or the existence of multiple sources along the trajectory path leads to inconsistent proportions of toluene to benzene with which to measure airmass age.

S3 Meteorological analysis

Wind speed are estimated to vary between 2-5 m/s as calculated by the HYSPLIT model (Figure S7) streamlines with end-heights of 10, 50, and 100 m as described in Section 2. Using these values, The aged aerosol formed ≥ 10 hours prior is determined to lie beyond the Tijuana and San Diego regions (Figure S8).

References

- Chow, J. C., Watson, J. G., Mar. 2002. Review of pm2.5 and pm10 apportionment for fossil fuel combustion and other sources by the chemical mass balance receptor model. Energy & Fuels 16 (2), 222–260.
- Hastie, T., Tibshirani, R., Friedman, J., 2009. The elements of statistical learning: data mining, inference, and prediction. Springer Verlag.

- 255 Henry, R. C., 2003. Multivariate receptor modeling by n-dimensional edge detection. *Chemometrics and*
256 *Intelligent Laboratory Systems* 65 (2), 179–189.
- 257 Jimenez, J. L., Canagaratna, M. R., Donahue, N. M., Prevot, A. S. H., Zhang, Q., Kroll, J. H., DeCarlo,
258 P. F., Allan, J. D., Coe, H., Ng, N. L., Aiken, A. C., Docherty, K. S., Ulbrich, I. M., Grieshop, A. P.,
259 Robinson, A. L., Duplissy, J., Smith, J. D., Wilson, K. R., Lanz, V. A., Hueglin, C., Sun, Y. L., Tian,
260 J., Laaksonen, A., Raatikainen, T., Rautiainen, J., Vaattovaara, P., Ehn, M., Kulmala, M., Tomlinson,
261 J. M., Collins, D. R., Cubison, M. J., Dunlea, E. J., Huffman, J. A., Onasch, T. B., Alfarra, M. R.,
262 Williams, P. I., Bower, K., Kondo, Y., Schneider, J., Drewnick, F., Borrmann, S., Weimer, S., Demerjian,
263 K., Salcedo, D., Cottrell, L., Griffin, R., Takami, A., Miyoshi, T., Hatakeyama, S., Shimono, A., Sun,
264 J. Y., Zhang, Y. M., Dzepina, K., Kimmel, J. R., Sueper, D., Jayne, J. T., Herndon, S. C., Trimborn,
265 A. M., Williams, L. R., Wood, E. C., Middlebrook, A. M., Kolb, C. E., Baltensperger, U., Worsnop, D. R.,
266 Dec. 2009. Evolution of organic aerosols in the atmosphere. *Science* 326 (5959), 1525–1529.
- 267 Liggio, J., Li, S.-M., Vlasenko, A., Sjostedt, S., Chang, R., Shantz, N., Abbatt, J., Slowik, J. G., Bottenheim,
268 J. W., Brickell, P. C., Stroud, C., Leaitch, W. R., Nov. 2010. Primary and secondary organic aerosols in
269 urban air masses intercepted at a rural site. *Journal of Geophysical Research-atmospheres* 115, D21305.
- 270 Ng, N. L., Canagaratna, M. R., Jimenez, J. L., Zhang, Q., Ulbrich, I. M., Worsnop, D. R., Feb. 2011.
271 Real-time methods for estimating organic component mass concentrations from aerosol mass spectrometer
272 data. *Environmental Science & Technology* 45 (3), 910–916.
- 273 Paatero, P., Hopke, P. K., Song, X. H., Ramadan, Z., 2002. Understanding and controlling rotations in factor
274 analytic models. *Chemometrics and Intelligent Laboratory Systems* 60 (1-2), 253–264.
- 275 Paatero, P., Tapper, U., 1994. Positive matrix factorization - a nonnegative factor model with optimal
276 utilization of error-estimates of data values. *Environmetrics* 5 (2), 111–126.
- 277 Polissar, A. V., Hopke, P. K., Paatero, P., 1998. Atmospheric aerosol over alaska - 2. elemental composition
278 and sources. *Journal of Geophysical Research-atmospheres* 103 (D15), 19045–19057.
- 279 R Development Core Team, 2012. R: A Language and Environment for Statistical Computing. R Foundation
280 for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0.
281 URL <http://www.R-project.org/>
- 282 Russell, L. M., Bahadur, R., Hawkins, L. N., Allan, J., Baumgardner, D., Quinn, P. K., Bates, T. S.,
283 2009a. Organic aerosol characterization by complementary measurements of chemical bonds and molecular
284 fragments. *Atmospheric Environment* 43 (38), 6100–6105.
- 285 Russell, L. M., Bahadur, R., Ziemann, P. J., 2011. Identifying organic aerosol sources by comparing functional
286 group composition in chamber and atmospheric particles. *Proceedings of the National Academy of Sciences*
287 of the United States of America
- 108 (9), 3516–3521.
- 288 Russell, L. M., Takahama, S., Liu, S., Hawkins, L. N., Covert, D. S., Quinn, P. K., Bates, T. S., 2009b.
289 Oxygenated fraction and mass of organic aerosol from direct emission and atmospheric processing measured
290 on the r/v ronald brown during texaqs/gomaccs 2006. *Journal of Geophysical Research-atmospheres* 114.
- 291 Slowik, J. G., Stroud, C., Bottenheim, J. W., Brickell, P. C., Chang, R. Y. . W., Liggio, J., Makar, P. A.,
292 Martin, R. V., Moran, M. D., Shantz, N. C., Sjostedt, S. J., van Donkelaar, A., Vlasenko, A., Wiebe,
293 H. A., Xia, A. G., Zhang, J., Leaitch, W. R., Abbatt, J. P. D., 2010. Characterization of a large biogenic
294 secondary organic aerosol event from eastern canadian forests. *Atmospheric Chemistry and Physics* 10 (6),
295 2825–2845.
- 296 Ulbrich, I., Lechner, M., Jimenez, J., 2012. AMS Spectral Database.
297 URL <http://cires.colorado.edu/jimenez-group/AMSSd/>

- 298 Ulbrich, I. M., Canagaratna, M. R., Zhang, Q., Worsnop, D. R., Jimenez, J. L., 2009. Interpretation of
299 organic components from positive matrix factorization of aerosol mass spectrometric data. *Atmospheric*
300 *Chemistry and Physics* 9 (9), 2891–2918.
- 301 Wasserman, L., 2010. All of Statistics: A Concise Course in Statistical Inference (Springer Texts in Statistics).
302 Springer.

303 **Supplemental Tables**

Table S1: Summary of instruments used in this work. Further details are described in Section 2.

Analytical method	Measurement	Time resolution
FTIR	PM ₁ total OM and organic functional group abundance	6-12 hours
ACSM	Nonrefractory (NR) PM ₁ mass (SO ₄ , NO ₃ , NH ₄ , Cl, Organics) and organic molecular mass fragments	15-30 minutes
XRF	PM ₁ elemental composition	6-12 hours
PTR-MS	VOC compound concentrations from mass fragment analysis	10 minutes
SP2	Size-resolved black carbon number and mass concentrations (only total number concentrations used in this work)	Single-particle
STXM-NEXAFS	Individual Morphology and composition of carbonaceous particles	Single-particle

Table S2: Organic aerosol spectra components or classes derived for FTIR absorbance spectra and ACSM mass fragment spectra. Further details are described in Section 3 and text of this document.

Instrument	Method	Components or classes (campaign average \pm standard deviation, or range in estimated value from multiple methods)
FTIR	PMF	FF1 (40 \pm 28%), FF2 (17 \pm 19%), BB (20 \pm 20%), Marine (23 \pm 24%)
FTIR	Nonlinear regression	hydroxyl (22 \pm 13%), alkane (44 \pm 9%), carboxylic acid (26 \pm 7%), ketonic carbonyl (3 \pm 5%), primary amine (5 \pm 2%), and organic nitrate (0.3 \pm 0.3%) functional groups
ACSM	PMF	OOA-1 (60 \pm 19%), OOA-2 (40 \pm 19%)
ACSM	Linear regression	OOA (88 \pm 19%), HOA(12 \pm 19%)
ACSM	Classification	OOA-1 (69-77%), OOA-2 (20-26%), BBOA (2-4%), HOA (1-2%)

304 Supplemental Figures

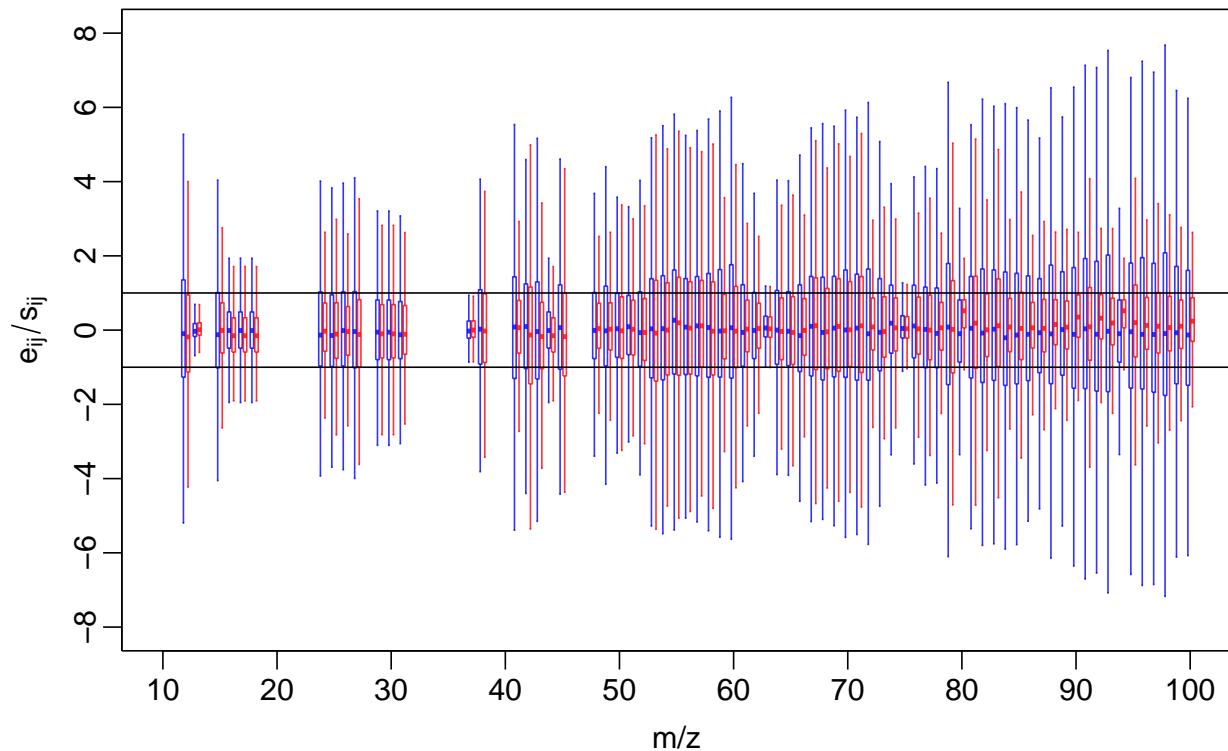


Figure S1: Residuals normalized by corresponding standard deviation values (original in blue; revised in red). Thick, horizontal lines indicate median values, boxes span interquartile range, and whiskers span 1.5 times the interquartile range. Black, horizontal lines are drawn at values of ± 1 .

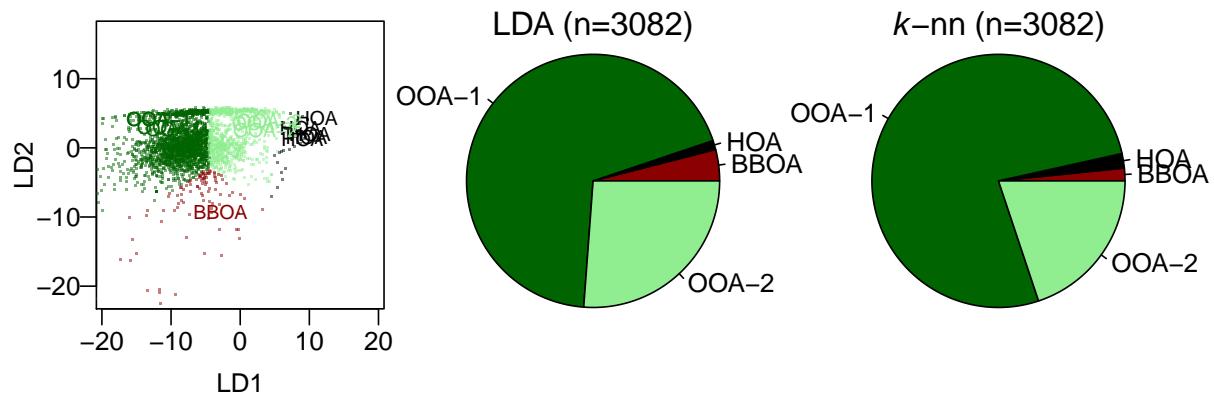


Figure S2: Feature vectors projected onto the space of the first two linear discriminants, and classified fractions of sample spectra according to LDA and k -nn methods (from left to right).

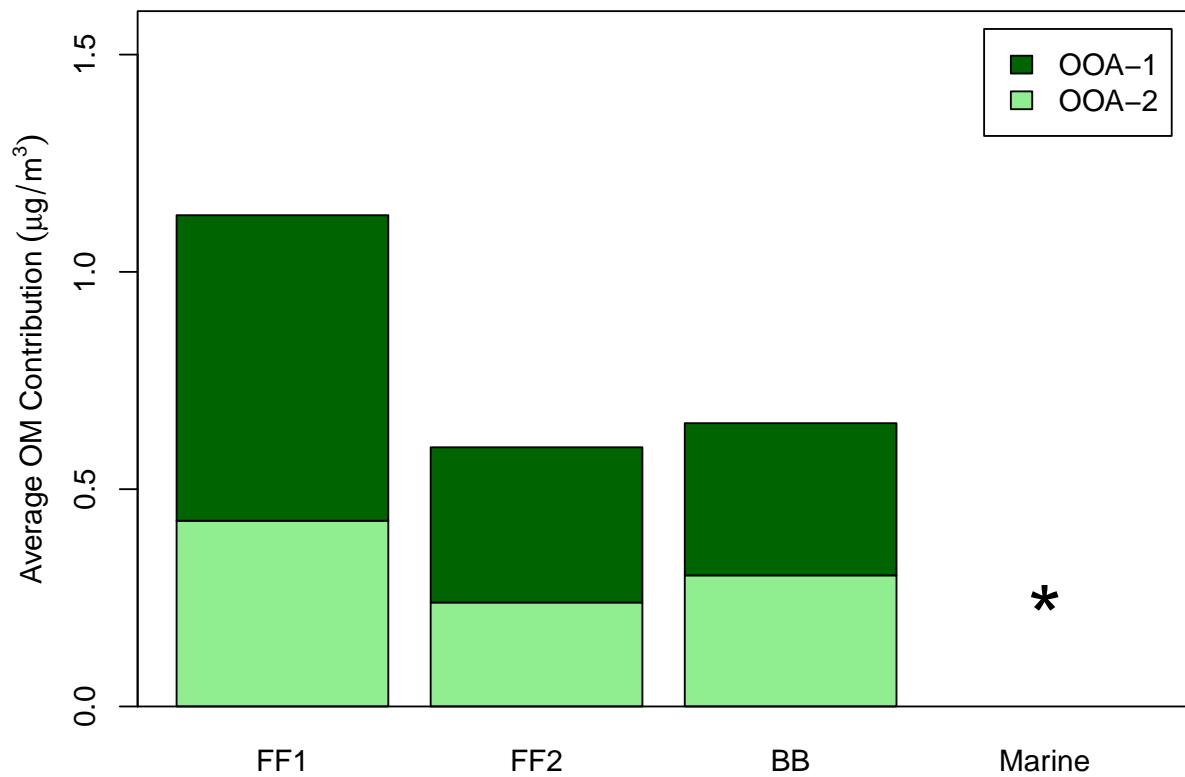


Figure S3: Proportion of FTIR PMF factors associated with ACSM PMF factors, estimated from Equation (S6). Asterisk indicates component not included in OM sum.

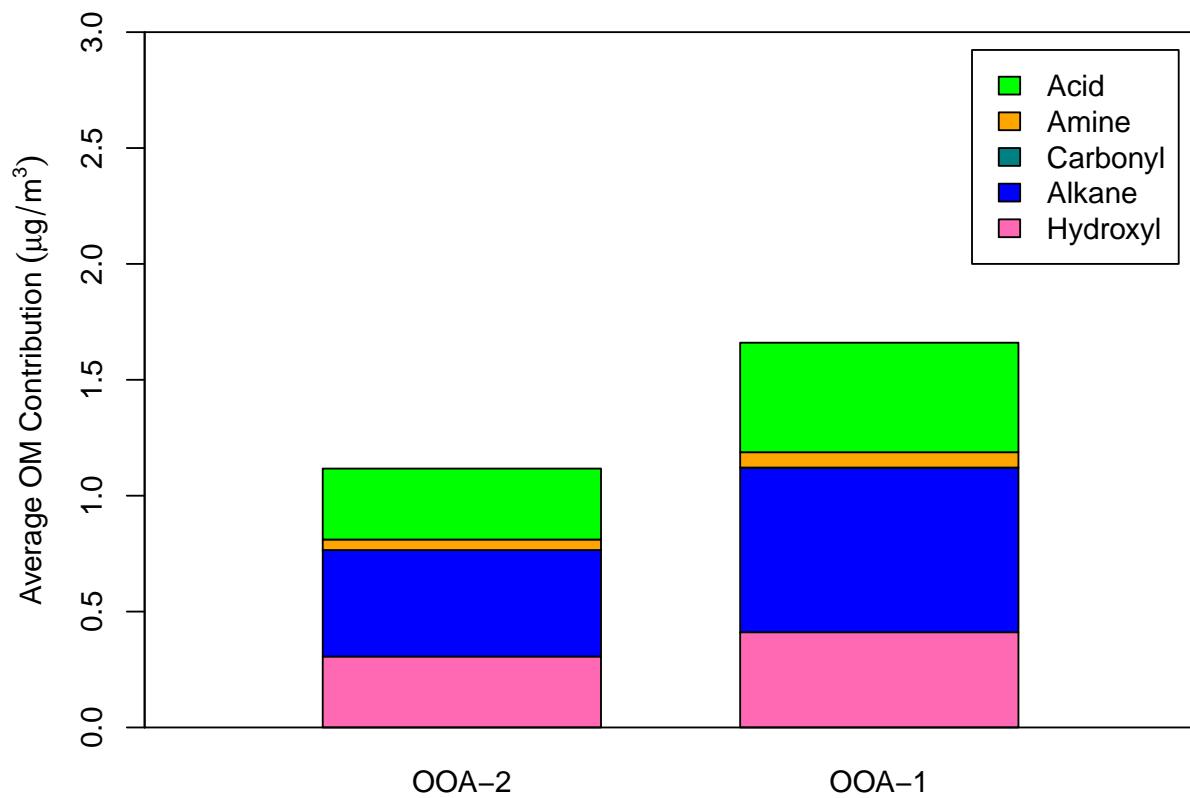


Figure S4: Proportion of FTIR (non-marine) OFG associated with ACSM PMF factors, estimated from Equation (S6) with modifications described in Section S2.1.

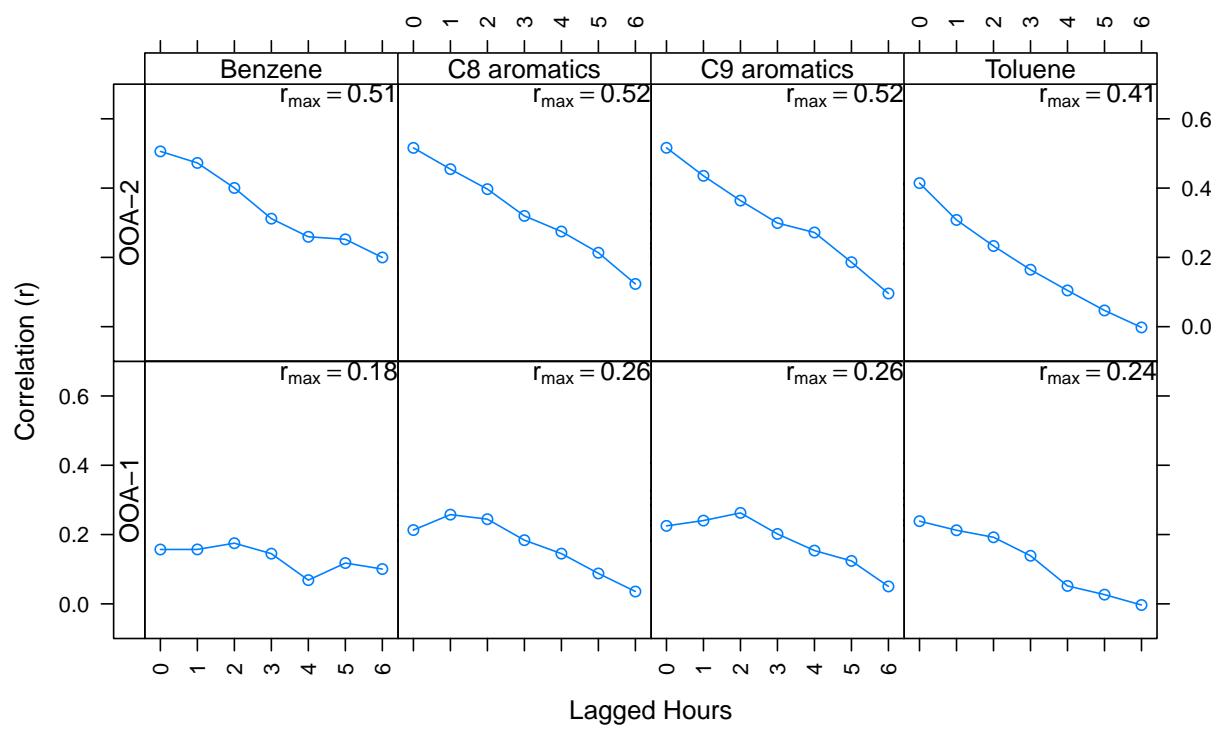


Figure S5: Lagged correlations of ACSM PMF factors with VOCs measured by PTR-MS.

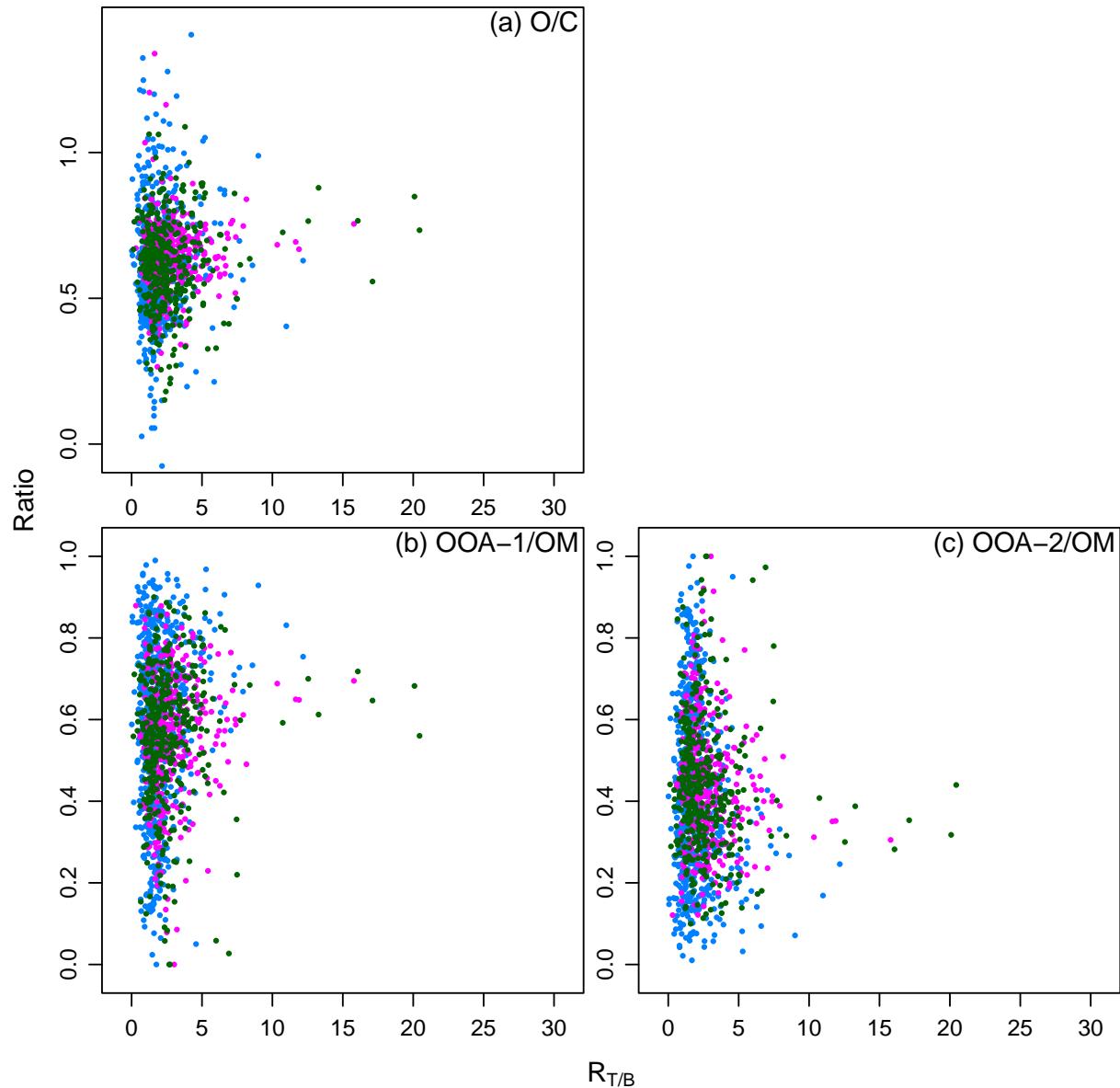


Figure S6: O/C ratio and PMF-factor fractions of OM as a function of Toluene-to-benzene ratio ($R_{[T/B]}$). Color indicates time of day: evening (blue), morning (pink), afternoon (green).

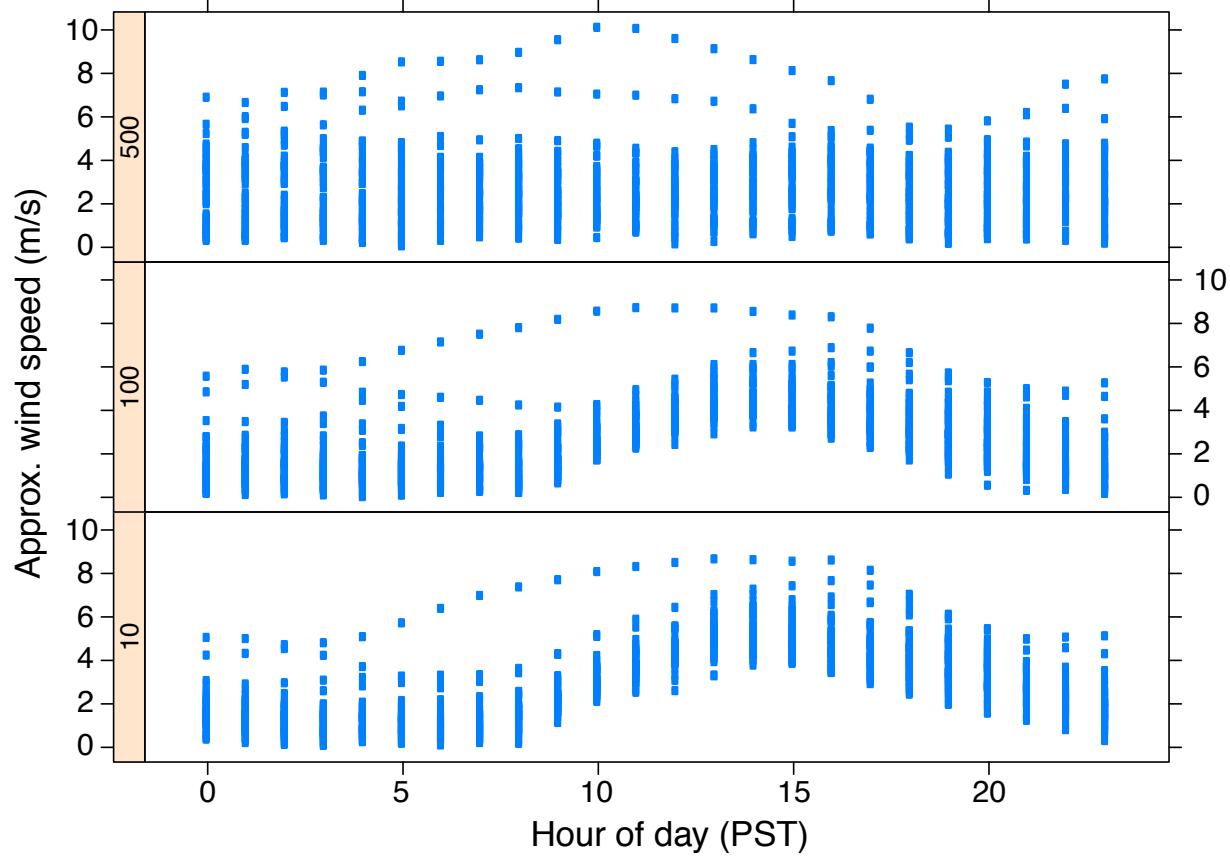


Figure S7: Wind speeds estimated by HYSPLIT model at altitudes of 10, 100, and 500 m shown in vertical panels from bottom to top, respectively.

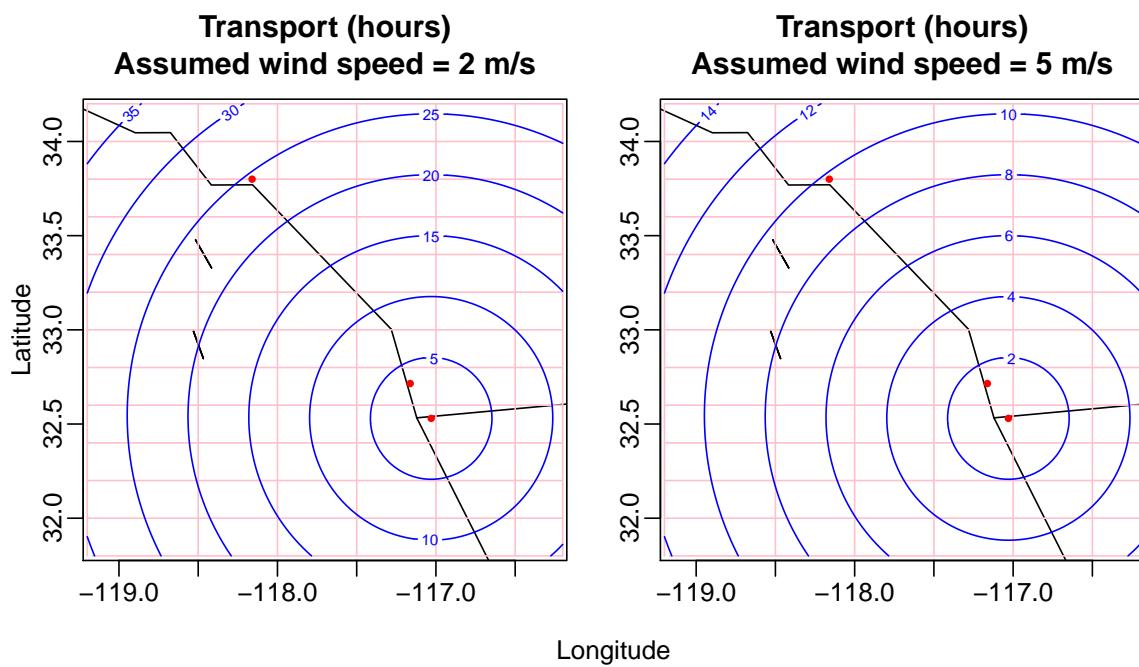


Figure S8: Transport times (blue) required for airmasses to reach Tijuana, estimated for fixed wind speeds of 2 and 5 m/s. Red dots from north to south represent Long Beach, CA, San Diego, CA, and Tijuana (Parque Morelos), Mexico.