Supporting Information for:

On the chemical nature of the Oxygenated Organic Aerosol: implication in the formation and aging of α-pinene SOA in a Mediterranean Environment, Marseille.

Imad El Haddad [a,1,*], Barbara D'Anna [b], Brice Temime-Roussel [a], Melanie Nicolas [b], Antoinette Boreave [b], Olivier Favez[b,2], Didier Voisin [c], Jean Sciare [d], Christian George [b], Jean-Luc Jaffrezo [c], Henri Wortham [a], and Nicolas Marchand [a,*]

[a] Aix-Marseille Univ, Laboratoire Chimie Environnement, 13331, Marseille cedex 03, France
[b] Université Lyon 1, Lyon, F-69626, France; CNRS, UMR5256, IRCELYON, Institut de Recherches sur la Catalyse et l'Environnement de Lyon, Villeurbanne, F-69626, France
[c] Universités Joseph Fourier-Grenoble 1-CNRS, UMR 5183, Laboratoire de Glaciologie et Géophysique de l'Environnement, Saint Martin d'Hères, 38402, France.
[d] Laboratoire des Sciences du Climat et de l'Environnement, CEA-CNRS-UVSQ-IPSL, Gif sur Yvette, 91191, France.

[1] Now at Laboratory of Atmospheric Chemistry, Paul Scherrer Institut, 5232 Villigen PSI, Switzerland.
[2] Now at INERIS, DRC/CARA/CIME, Parc Technologique Alata, BP2, Verneuil-en-Halatte, 60550, France.

[*] Corresponding authors:
I. El Haddad – email: imad.el-haddad@psi.ch, phone: +41 5 63 10 27 85, fax: +41 5 63 10 45 25
N. Marchand – email: Nicolas.Marchand@univ-amu.fr, phone: +33 4 13 55 10 51, fax: +33 4 13 55 10 60

S.1 Site description and sampling strategy

Description of the sampling site and meteorological conditions encountered during the study are thoroughly detailed in El Haddad et al., 2011a and b, and only a brief outline follows. Field measurements were conducted in summer 2008 (30 June-14 July), at an urban background site located in a downtown park, in Marseille (43°18′20″N, 5°23′40″E, 64 m.a.s.l.). Marseille, the second most populated city in France, comprises the most important port of the Mediterranean Sea. It is also in the vicinity of a large petrochemical and industrial area, located 40 km northwest of the metropolitan area. The main industries include petroleum refining, shipbuilding, steel facilities and coke production. Owing to a particular air mass circulation in this region, industrial emissions can directly impact the metropolitan area of Marseille, particularly in sea breeze conditions. This area is also well known for its photochemical pollution, especially regarding ozone (Flaounas et al., 2009), and evidence of rapid formation of secondary organic aerosol has been pointed out within the frameworks of ESCOMPTE experiment (Cachier et al., 2005) and BOND project (Petäjä et al., 2007). Fig.S1 illustrates the air masses impacting the sampling site during the measurement period, showing that these are mostly associated with long range transport from the Atlantic and the Mediterranean.

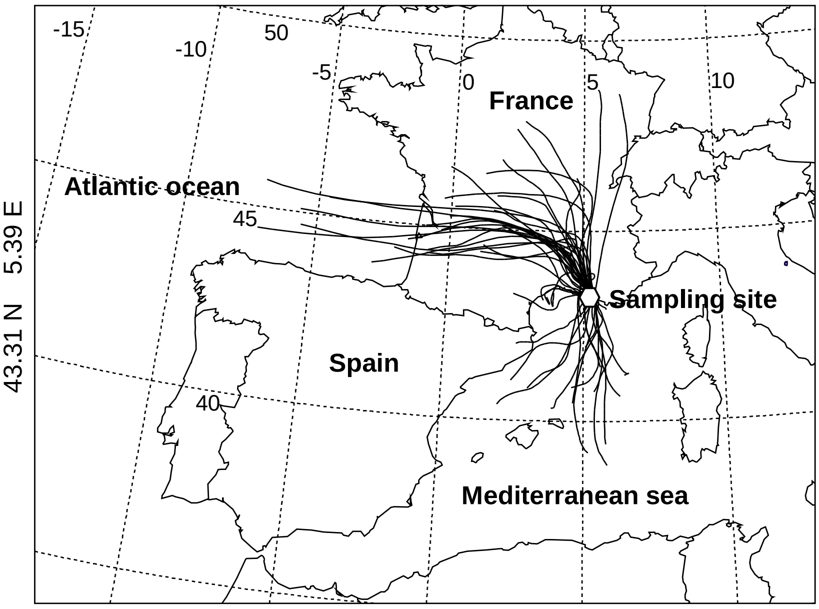

Fig.S1: 24h HYSPLIT air masses backward trajectories at 100m above sea level (Rolph, 2010) illustrating the overall air masses circulation occurring during the entire measurement campaign. Backward trajectories are confirmed by both MM5 modeling and local wind measurements.

$PM_{2.5}$ collection was performed using high volume samplers (HiVol, Digitel DA80) operating at a flow rate of 30 m^3 h^{-1}. The samples were collected continuously on a 12-hour basis (5:30 to 17:30 UT, and 17:30 to 05:30 UT, total number of 30 samples) onto 150mm-diameter quartz fiber filters (Whatman QMA), pre-heated at 500 °C during 3 h. Samples were then stored at -18 °C in aluminium foil, sealed in polyethylene bags until analysis. Six field blank samples were also prepared following the same procedure. $PM_{2.5}$ were also collected on a 24-h timescale onto pre-heated 25 mm-diameter quartz filters using a Dekati 13-stage low pressure cascade impactor (LPI) at a flow rate of 30 l min^{-1}, for size resolved EC/OC analysis.

The chemical composition of fine PM was also investigated every 2 min using a compact time-of-flight (c-TOF, Tofwerk) Aerodyne Aerosol Mass Spectrometer (AMS, Aerodyne). This instrument allows real-time measurements of PM_1 non-refractory components (OA, NH_4, NO_3 and SO_4) combining thermal vaporization and electron ionization (Drewnick et al., 2005). Aerosol size distribution (mobility diameters from 11 to 1083 nm), was investigated using a Scanning Mobility Particle Sizer (SMPS, L-DMA and CPC5403, GRIMM). Semi-continuous hourly concentrations of elemental carbon (EC) and organic carbon (OC) in $PM_{2.5}$ were obtained in the field from an OC/EC Sunset field instrument (Sunset Laboratory, Forest Grove, OR, USA; Bae et al., 2004) running at 8l min^{-1}. AMS results are all corrected for the collection efficiency, using a common factor of 0.65±0.14, estimated based on SMPS and EC data. Fig.S2 displays EC, OA, NH_4, NO_3 and SO_4 time series recorded over the period of study.

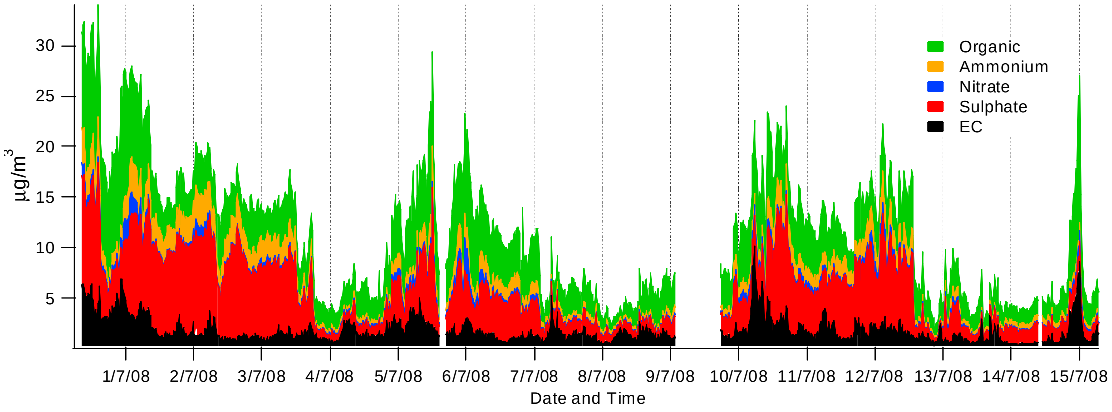

Fig.S2:. Time series of the main PM_1 components (EC, OA, NH_4, NO_3 and SO_4) during the period of study. Due to technical issues, AMS measurements are not available between the 9th and 10th of July.

In addition to aerosol sampling, HS-PTRMS (High Sensitivity Proton Transfer Reaction Mass Spectrometer, Ionicon Analytic, (Lindinger et al., 1998) was deployed in order to quantify volatile organic compounds (VOCs) and oxygenated VOCs. In this paper, a special focus is placed on the isoprene oxidation products (Methacroleine, MACR and Methyl Vinyl Ketone, MVK), tracers of aged biogenic air masses impacting the site. Finally, 15 min-averaged NO_X, O_3, SO_2 and $PM_{2.5}$ concentrations were also measured with the standard equipment of the air quality monitoring network.

S.2 Offline chemical analyses

$PM_{2.5}$ collected onto 150mm-diameter filters was comprehensively characterized. Technical description of the analysis techniques can be found in El Haddad et al., 2011a and b and only a brief outline follows.

EC/OC, ions, WSOC, $HULIS_{WS}$ and elements: The carbonaceous content was analyzed for EC and OC using a Thermo-Optical Transmission method on a Sunset Lab analyzer (Birch and Cary, 1996), following both NIOSH (Schmid et al., 2001) and EUSAAR-2 (Cavalli et al., 2010) protocols. It is well established that different protocols result in very different values for EC (Schmid et al., 2001). We based our analysis (i.e. Chemical Mass Balance analysis and multiple regression analysis, see below sections S.4 and S.8) on concentrations determined following NIOSH protocol, as source profiles were determined based on this protocol. Biases arising from discrepancies between the two protocols are all discussed in S.8.2 of the supporting material.

Sample fractions of 11.34 cm^2 were extracted into 15mL ultrapure Milli-Q water by 30 min short vortex agitation for the analyses of major ions (NH_4^+, SO_4^{2-}, NO_3^-), water-soluble organic carbon (WSOC) and water-soluble humic like substances (HULIS). HULIS analysis was performed following the method described in Baduel et al. (2009, 2010). This method involves extraction of HULIS by adsorption onto DEAE resin (GE Healthcare®, HiTrapTM DEAE FF, 0.7 cm ID×2.5 cm length) and its subsequent quantification with an OI Analytical 700 total organic carbon analyzer.

Finally, fifty elements were measured using ICP-MS (Agilent 7500ce) following complete dissolution of filter aliquots in a mixture of high-purity concentrated HF and HNO_3. Element concentrations were then calculated using the rock reference material BR (Chauvel et al., 2010).

<u>*Radiocarbon measurements:*</u> Radiocarbon (^{14}C) measurements were conducted on high volume quartz filter fractions (~40 cm^2) using ARTEMIS Accelerator Mass Spectrometry. Each sample was first packed into a prefired quartz tube containing CuO and Ag powder to be combusted at 850°C in a muffle furnace for 4 hours. Carbon dioxide was collected and purified before its conversion into graphite by hydrogen reduction at 600°C using Fe catalyst. The modern fraction (f_m) was determined as the ratio of $^{14}C/^{12}C$ in aerosol sample to $^{14}C/^{12}C$ in the NBS Oxalic Acid standard (NIST-SRM-4990B).

In order to account for the thermonuclear weapon tests of the late 1950s and early 1960s (Levin et al., 1985), the modern fraction (f_m) is divided by a ratio of 1.1 to get a corrected non-fossil fraction (f_{nf}) (Levin and Hesshaimer, 2000). This value is subsequently subtracted from one to obtain the fossil fraction (f_f).

<u>*Organic speciation:*</u> A chemical derivatisation/gas chromatography-mass spectrometry (GC-MS) approach was used to quantify primary and secondary organic markers, including α-pinene oxidation products, a major focus of this study. The approach is fully described in El Haddad et al., 2011b and will be only outlined in the following.

Prior to extraction, filters were spiked with known amounts of two isotope-labelled standards: tetracosane-d50 and cholesterol-d6. Organic species were extracted from filters with a dichloromethane/acetone mix (1/1 v:v) using an accelerated pressurized solvent extraction device (ASE, Dionex 300). Extracts were then reduced to a volume of 500µL using a Turbo Vap II concentrator. The remainder was split into two fractions. The first fraction was directly injected, whilst the second fraction was subjected to derivation for 2 h at 70 °C before GC-MS analysis, using N,O-Bis(trimethylsilyl)-trifluoroacetamide containing 10% trimethyl-chlorosilane. The two fractions were analyzed following the same GC-MS conditions: Aliquots of 2 µL were analyzed using a Thermo Trace GC chromatograph interfaced to a Polaris Q ion trap mass spectrometer fitted with an external electron ionization source. The chromatographic separation was accomplished on a TR-5MS capillary column (Thermo Electron, 30 m × 0.25 mm i.d. × 0.25 µm film thickness). Field blank filters were also treated with the same procedure and none of the target compounds were detected in these field blanks.

Primary organic markers including n-alkanes, hopanes, polycyclic aromatic hydrocarbon (PAH) and levoglucosan were quantified and used as inputs in the CMB analysis to apportion primary sources. These can all be found in El Haddad al. (2011a). α-pinene oxidation products quantified by this method are used to estimate α-pinene SOA contributions. As described in El Haddad al. (2011b), we identified and quantified 9 α-pinene SOA markers, whose structures are presented in Fig.1. These include pinic (PA) and pinonic (PNA) acid identified and quantified using authentic standards. Seven other multifunctional compounds (A1-A7), for which native standards are not available, were tentatively identified by examining their retention times and MS characteristics (for more details refer to El Haddad et al., 2011b). They include 3-hydroxyglutaric acid (A1), 3-(2-hydroxyethyl)-2,2-dimethylcyclobutane carboxylic acid (A2), 3-hydroxy-4,4-dimethylglutaric acid (A3), 3-acetylglutaric acid (A4), 3-acetyladipic acid (A5), and 3-isopropylglutaric acid (A6) and 3-methyl-1,2,3-butanetricarboxylic (A7). These compounds were quantified using the response factor of malic acid as surrogate. Relative standard deviation of the concentrations based on duplicate analysis is between 5 and 15%.

S.3 Comparison between offline and online measurements

The aim of this section is to evaluate biases and artefacts associated with the offline and online measurements of OA (e.g. AMS particle collection efficiency, adsorption artefacts onto filters). Fig.S3 conveys the comparison between AMS (PM_1) and filter based ($PM_{2.5}$) measurements for the two major aerosol components: SO_4 and OA. SO_4 is expected to primarily occur in the PM_1 fraction and to be mostly associated with ammonium sulfate and bisulfate (very little influence from sea salt), and thus to be quantitatively analysed by the AMS. A very good agreement can be observed between the AMS-SO_4 and the HiVol-SO_4 (s~1, i~0 and $R^2>0.9$; Fig.S3), substantiating our AMS measurements and the particle collection efficiency factor, CE=0.65, estimated based on the SMPS and EC data. HiVol-OA was derived from OC concentrations measured on filter samples, corrected for differences in diameter cut-offs between the AMS and the HiVol sampler; **it constitutes our best estimate of offline PM_1OA**. The calculation of PM_1OA proceeds as follows: Based on size resolved EC/OC measurements performed on the LPI samples, the fraction of PM_1OC in $PM_{2.5}OC$ was retrieved: $PM_1OC/PM_{2.5}OC$ = 0.82±0.06. PM_1OC is then scaled by an average OM/OC ratio of 1.67±0.05, obtained by comparing the AMS-OA to the LPI PM_1OC. The comparison between PM_1OA and AMS-OA reveals that both fractions exhibit similar variability (R^2>0.7), with a slope close to 1. However, a negative intercept of -1.3±0.7 µg m^{-3} can be observed, implying that filter based measurements (PM_1OA) are systematically associated with a positive bias 1.3 µg m^{-3} engendered by adsorption artefacts onto filter samples. As a result, filter based measurements tend to overestimate the absolute concentrations of OA by up to 28%[†]. In contrast, such artefacts would have a minor influence on our apportionments, providing that they evenly impact the different components of OA.

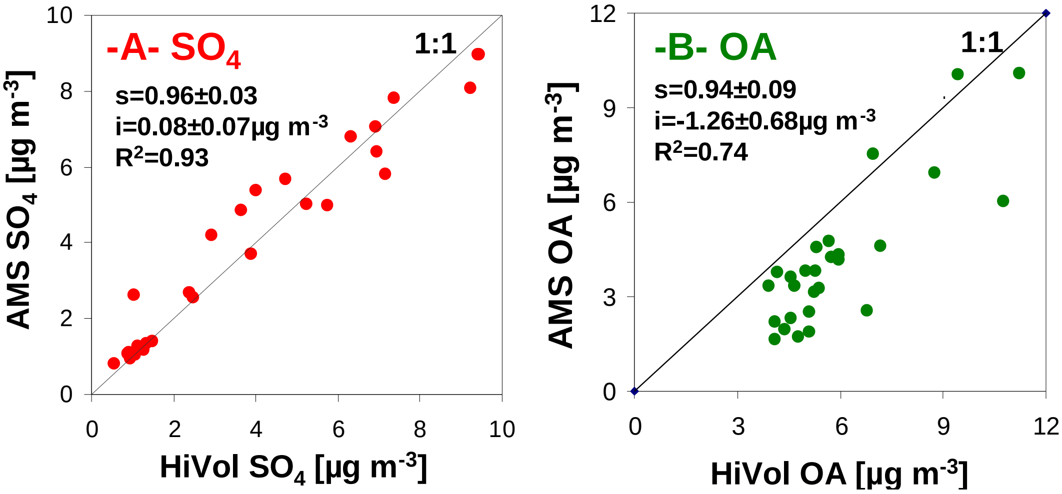

Fig.S3: Comparison between AMS and offline measurements for SO_4 (**A**) and OA (**B**). HiVol-OA was corrected for differences in the diameter cut-offs between the AMS and the HiVol sampler (see text); it refers to the PM_1 fraction. Also, shown are the 1:1 line and the slopes (s), intercepts (i) and coefficients of determination (R^2) obtained by linear fits of the data.

S.4 CMB analysis

Available data used here also include source contributions to OC, apportioned using a Chemical Mass Balance analysis (CMB) in conjunction with organic marker concentrations, as fully described in El Haddad et al. (2011a). CMB model is based on the mass conservation of individual organic markers. In the mass conservation equations, known concentrations (C_{ik}) of specific markers of primary sources at receptor site k are written as the product of known source profiles a_{ij} and unknown primary source contributions s_{jk} (Watson et al., 1998) as expressed in equation 1:

[†] The following estimation of adsorption artefacts onto HiVol filter samples (positive artefacts of 28%) is obtained by assuming no volatilisation artefacts occurring during sampling with the LPI. Negative artefacts are common for sampling under low pressure and under the latter assumption we tend to overestimate the positive artefacts onto the HiVol samples.

$$C_{ik} = \sum_{j=1}^{m} a_{ij} s_{jk} \tag{1}$$

where m denotes the total number of emission sources and a_{ij} is represented by the fractional abundances of chemical species in the source emissions, expressed as marker-to-OC ratios. The set of linear equations generated by equation 1 is solved with an effective variance weighted least square method using Environmental Protection agency EPA-CMB8.2 software.

Primary markers and source profiles selection is detailed in El Haddad et al., 2011a. Primary markers include: levoglucosan as a specific marker for biomass burning (BBOC), EC and three hopanes (i.e., 17(H),21(H)-norhopane, 17(H),21(H)-hopane and 22S,17(H), 21 (H)-homohopane) as key markers for vehicular emissions. In addition, a series of C27-C32 n-alkanes was selected since this range demonstrates high odd-carbon preference, specific to primary biogenic sources. In order to apportion industrial emissions, four PAH (benzo[b,k]fluoranthene, benzo[e]pyrene, indeno[1,2,3-cd]pyrene, and benzo[ghi] perylene), V, Ni and Pb were included as fitting species. Source profiles comprise vehicular emissions derived from a tunnel study held in Marseille (El Haddad et al., 2009), biomass burning emissions (Fine et al., 2002), vegetative detritus (Rogge et al., 1993a) and natural gas combustion (Rogge et al., 1993b). Three industrial-emission-related profiles were chosen, including metallurgical coke production (Weitkamp et al., 2005), HFO combustion/shipping (Agrawal et al., 2008), and steel manufacturing (Tsai et al., 2007).

In this study, emissions from the three industrial processes are lumped together under the term “industrial OA”. Biomass burning, vegetative detritus and natural gas combustion contributed very little OC during the period of measurements (El Haddad et al., 2011a) and were not considered in the comparison between CMB and AMS/PMF results (see section S.7). CMB technique is unable to directly apportion secondary sources; however, the fraction of OC non-attributed to primary sources is considered as an upper bound estimate of secondary OC (SOC).

In order to compare CMB and AMS/PMF results, primary OA associated with vehicular and industrial emissions were calculated applying an OM-to-OC ratio of 1.2 (based on Aiken et al., 2008). SOA is considered as the difference between the total OA, determined by scaling the total OC by an OM-to-OC ratio of 1.67[‡], and the primary OA. An OM-to-OC ratio of 1.82 can be inferred for SOA (i.e. SOA-to-SOC), consistent with an overwhelmingly secondary origin of this fraction (Aiken et al., 2008).

S.5 estimation of α-pinene SOA contributions

Contributions of α-pinene SOA to ambient OA is estimated following the marker-based approach developed by Kleindienst et al. (2007). This approach consists of converting the measured concentrations of marker compounds derived from a given precursor hydrocarbon (hydrocarbon, *HC*) into a SOA concentration in μg m^{-3} ($[SOA]_{HC}$), using the laboratory-generated mass fractions of the same markers ($f_{SOA,HC}$) determined by Kleindienst et al. (2007). $[SOA]_{HC}$ can be subsequently calculated as follows:

$$[SOA]_{HC} = \frac{\sum_{i}^{J} [M]_i}{f_{SOA,HC}} \qquad (2)$$

where $[M]_i$ is the concentration of the marker *i* and *J* is the total number of markers derived from the hydrocarbon *HC*. For the apportionment of α-pinene SOA contributions, pinonic acid, pinic acid and A1-A7 were considered as markers and a $f_{SOA,HC}$ factor of 0.168±0.08 was used (Kleindienst et al., 2007).

S.6 Positive matrix factorization applied on AMS measurements

S.6.1 General principle

The AMS OA dataset was analyzed applying Positive Matrix Factorization (PMF2) (Ulbrich et al. 2009). A complete description of PMF2 and the subsequent evaluation tools can be found in Paatero and Tapper (1994), Lanz et al. (2007) and Ulbrich et al. (2009). Time series of organic mass spectra, arranged as a matrix (*X*), are deconvolved into a linear combination of smaller matrices such that

$$X_{ij} = \sum_{p} g_{ip} f_{pj} + e_{ij} \qquad (3)$$

[‡] OM-to-OC ratio of 1.67 is calculated by comparing AMS OA with LPI OC measurements, see section S.3.

where X_{ij} represents the elements of matrix X, p the number of factors in the solution, g_{ip} and f_{pj} the elements of matrices G and F representing respectively time series and mass spectra of each factor, and e_{ij} the elements of matrix E of residuals not fitted by the model for each data point. This model includes both the data matrix of organic fragments and an instrumental error matrix, both obtained from the AMS-data-analysis Squirrel software. In this study, the data matrix is composed of 4043 data points (time series) of 276 m/z. The error matrix calculated in Squirrel software was modified following the recommendations of Ulbrich et al. (2009) and references therein.

S.6.2 Number of factors: Assigning factors to emissions and processes

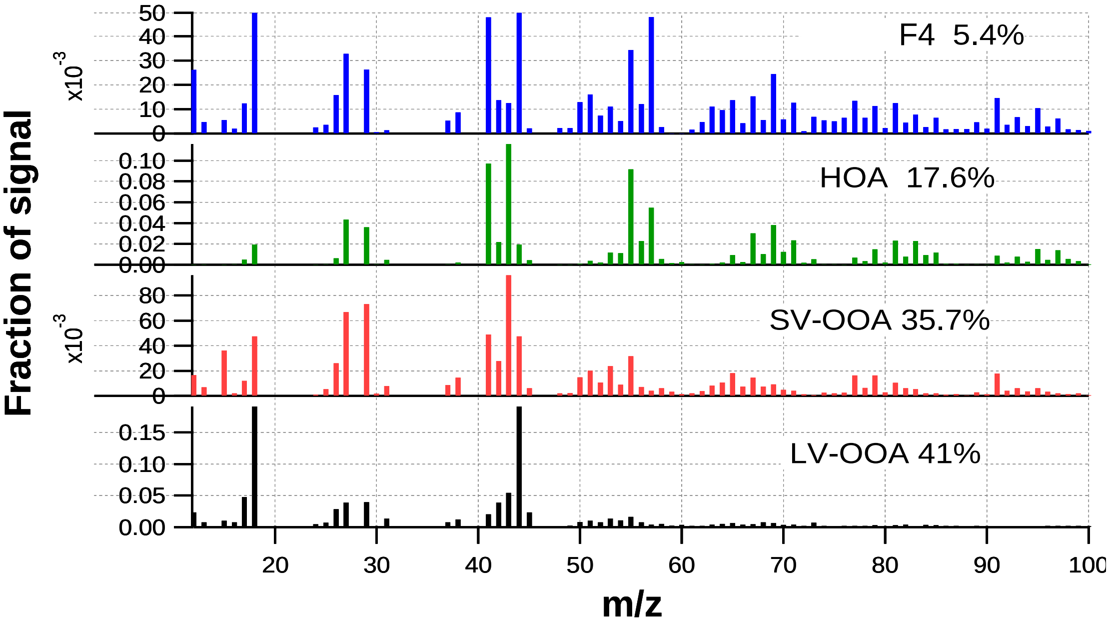

Fig.S4: Factor spectral profiles derived from the 4 factor solution PMF2 analysis and their average mass contribution over the period of study for “FPEAK”=0 and “seed”=0.

The number of factors has been evaluated taking into account the physical meaning of the factors, good correlations with external time series (Fig.2 in the manuscript) and with mass spectra from laboratory and ambient samples (see for example Tab.S1 for the 4 factor solution). Diurnal patterns of the factors were examined as well. Ultimately, we retained a 4 factor solution; the mass spectra, time series and diurnal variations of these factors a represented in Fig.S4, Fig.2 and Fig.S5, respectively.

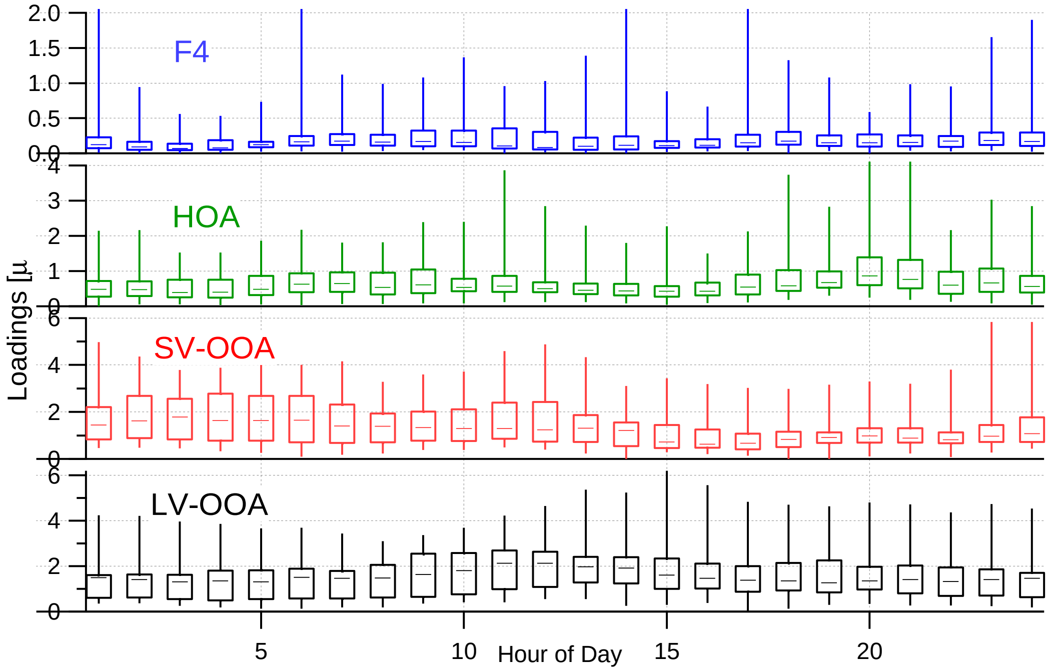

Fig.S5: Diurnal patterns of the OA four factors identified during FORMES campaign in Marseille (solution for FPEAK=0 and SEED=0).

We first examined a 3 factor solution, apportioning the OA to hydrocarbon like OA (HOA), related to traffic emissions, and to two oxygenated fractions related to semi-volatile moderately oxidized (SV-OOA) and low volatility highly oxidized (LV-OOA) OA. Adding one more factor enabled revealing the industrial influence, evidenced by the correlation of the corresponding factor (termed F4) with heavy metals (e.g. Pb, Cs, Mo, Fe, La, V, Zn, Ni and Co measured on filter samples) and polycyclic aromatic compounds (PAH, measured online by the AMS and offline). This factor exhibited remarkable variations similar to that of industrial OA apportioned using CMB, both characterised by episodic ten-fold enhancements in their contributions (Fig.2 in the manuscript). The wide variability in the contribution of industrial emissions is due to the fact that it is a local point source, whose influence is strongly dependent on local meteorology (El Haddad et al., 2011a). The mass spectrum of F4 exhibits a similar pattern as HOA, with high contribution from m/z 55, 57, and 69 and higher molecular weight fragments. F4 mass spectrum is also associated with high contribution from oxygenated fragments, e.g. m/z44 (CO_2^+), suggesting to some extent the aging of the emissions while transported to the site. Unfortunately, the low resolution of the C-TOF precludes an enhanced chemical characterisation of this factor.

Tab.S1: Pearson correlation coefficients (R) between Factor's mass spectra and reference mass spectra drawn from previous ambient, emission and chamber studies.

	HOA	**SVOOA**	**LVOOA**	**Factor 4**
HOA (Zhang et al., 2005)	**0.97**	0.57	0.16	
Diesel bus exhaust (Canagaratna et al., 2004)	**0.96**	0.66	0.25	
α-pinene SOA (Bahreini et al., 2005)	0.81	**0.96**	0.59	
OOA-type II (Lanz et al., 2007)	0.82	**0.97**	0.62	
Aged Rural (Alfarra et al. 2004)	0.55	0.79	**0.94**	
OOA-type I (Lanz et al., 2007)	0.41	0.71	**0.99**	

Overall, for this solution, the factor's mass spectra exhibit typical patterns, very similar to those obtained in previous studies (Tab.S1). SV-OOA correlates fairly well with volatile and semi-volatile secondary components such as nitrate and isoprene gas phase oxidation products (MVK+MACR, see Fig.2). In contrast, LV-OOA correlates with less volatile secondary components, including HULIS and sulfate (Fig.2), which clearly relates this fraction with low volatile highly aged SOA. HOA is characterised by a prominent diurnal pattern, with increasing contributions during rush hours (Fig.S5), correlating with vehicular tracers (EC and NO_X, Fig.2). Fig.S6 presents a scatter plot of HOA vs. EC, from which it is possible to investigate more thoroughly the sources of this fraction. As HOA and EC arise majorly from the same source, i.e. vehicular emissions (El Haddad et al., 2011a), it is expected that the data point cluster around one line with a slope corresponding to the HOA/EC ratio at the point of emission. However, more scattering is observed with three different clear patterns: Most of the data scatter around one line characterised by a ratio of HOA/EC of ~0.45, representative of average vehicular emissions at typical ambient concentrations (see for e.g. Chirico et al., 2011 and references therein). This is a clear indication that HOA is mostly related to traffic. A second part of the data scatter around another line characterised by a lower ratio of HOA/EC (<0.25), concomitant with the dilution of the emissions occurring with the development of the boundary layer after 10:00 and the enhancement of the photochemical activity, favouring the oxidation of HOA. In this regard, the depletion of traffic emission markers with respect to EC due to photochemistry was previously demonstrated in our conditions (El Haddad et al., 2011a). The third part of the data presents more scattering, with higher HOA/EC ratios (around 0.75), occurring mostly during night-time. This enhancement in HOA over EC might be due to the fact that HOA emitted during the night peak is less impacted by photochemistry and partitioning to gas phase with dilution. Another likely explanation that cannot be ruled out is the contamination of HOA by another source exhibiting similar spectral profiles, such as cooking emissions. Based on the comparison between EC and HOA, this contamination is on average less than 20%, in perfect agreement with the very low concentrations of cholesterol (0.13-3.32 ng m^{-3}, El Haddad et al., 2011a).

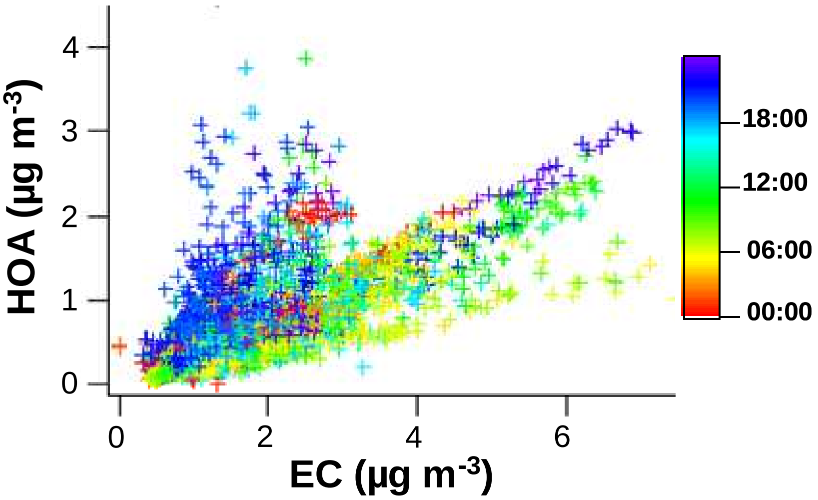

Fig.S6: HOA vs. EC. Color scale: hour of the day

Increasing further the numbers of factors resulted in a splitting/mixing of SVOOA, LVOOA and HOA and some ambiguity in assigning the factor spectral profiles to specific sources. The resulting factors shows less correlations with external tracers, hindering the factors attribution to specific sources. Therefore, the 4 factor solution was considered as the best solution.

S.6.3 Robustness of the selected solution

To assess the robustness of the 4 factor solution, rotational ambiguity has been investigated by varying FPEAK from -2 to 2 with 0.1 steps. Two main groups of solutions are identified, the first one corresponding to “FPEAK” values below 0, for which unrealistic zero time series values are observed for LVOOA, and the other one corresponding to “FPEAK” above 0. Robust solutions are found between 0 to 1 FPEAKs, with very little variability in the factor’s time series and mass spectra (Fig.S7).

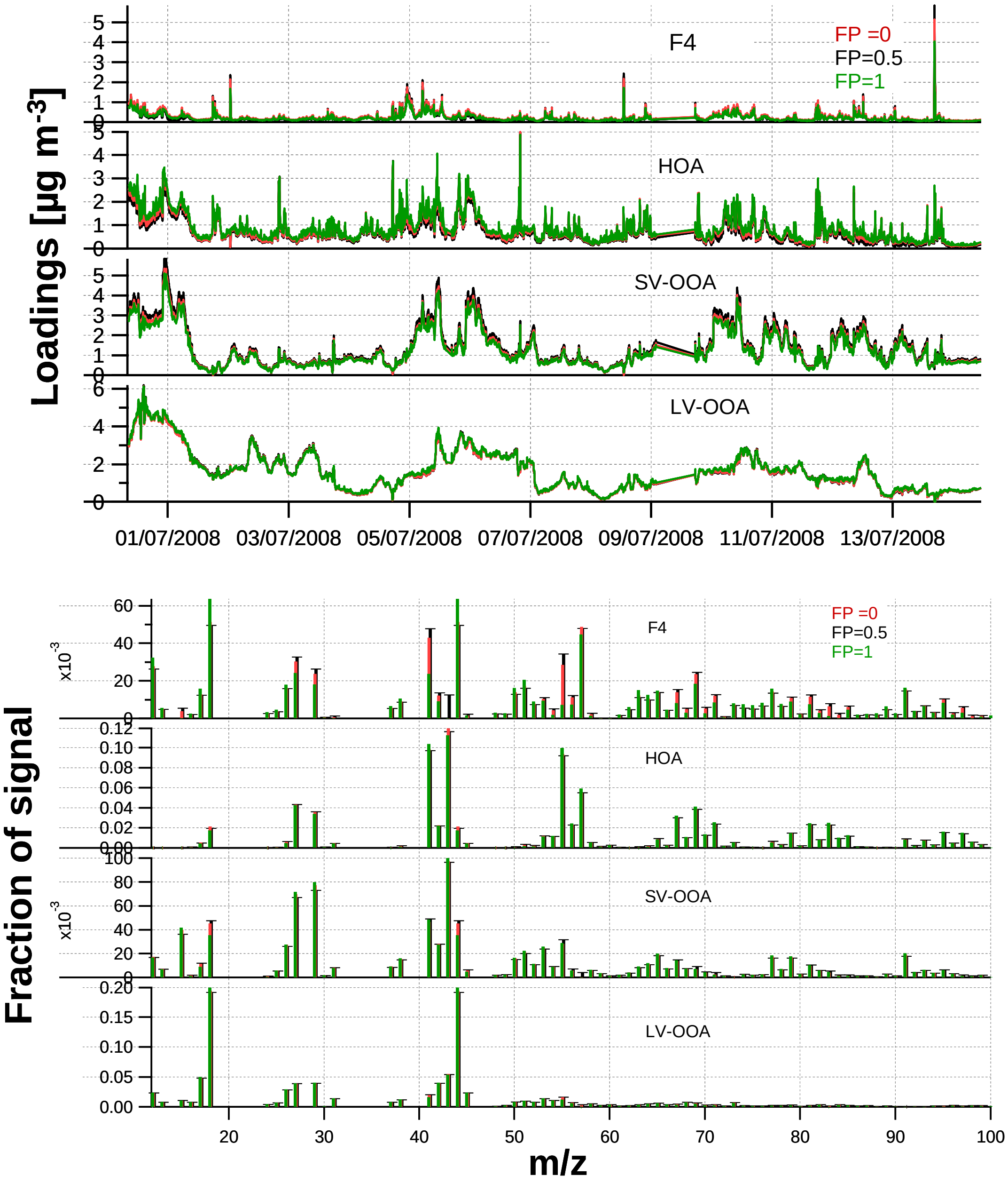

Fig.S7: Spectra calculated for four factorial PMF2factors and their mass contribution for the 29th June to 13th July 2008 period, with "FPEAK" = 0 and "seed" = 0.

The influence of the initial conditions "SEEDS" (corresponding to pseudorandom starting-points of the PMF2 algorithm) ranging from 0 to 60 (with steps of 1) was also verified (Fig.S8). No influence of different SEEDS is observed, an evidence of the robustness of the chosen solution.

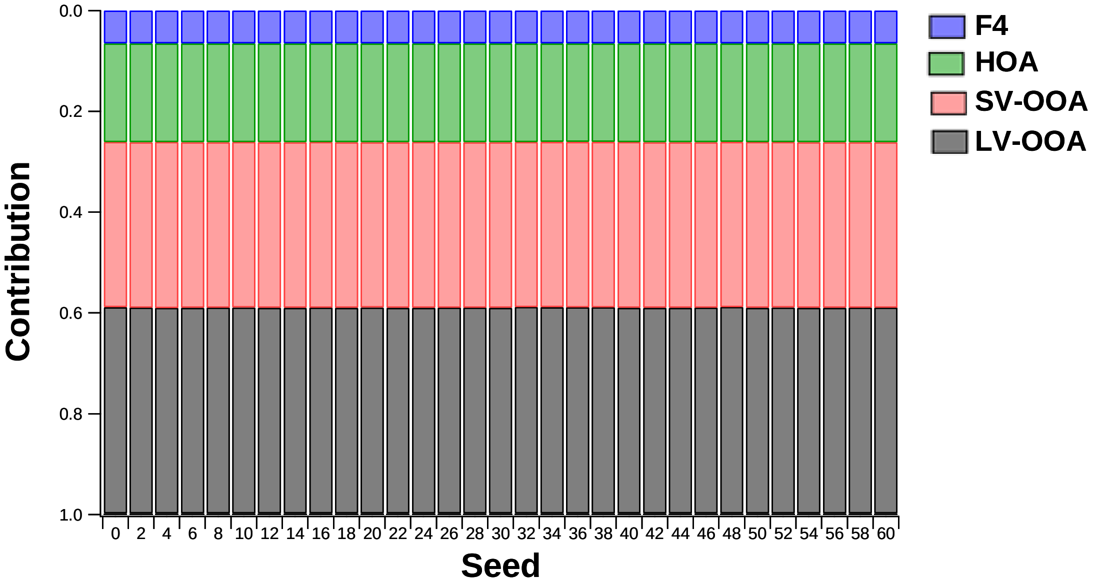

Fig.S8: Influence of the initial conditions "SEEDS" (corresponding to pseudorandom starting-points of the PMF2 algorithm) ranging from 0 to 60 (with steps of 1) for FPEAK=0.

S.7 AMS vs. CMB

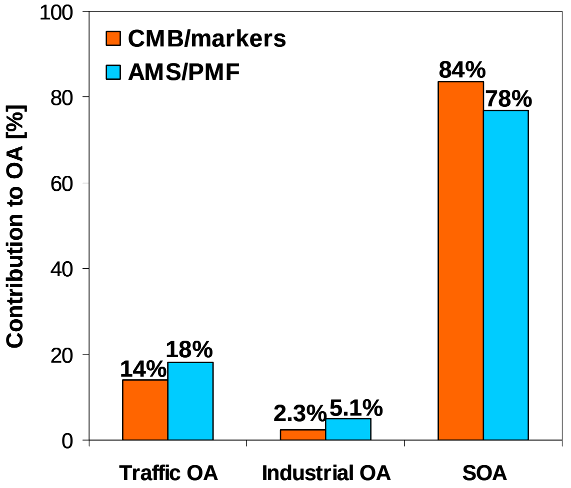

Fig.S9: Inter-comparison of AMS and CMB results. For the AMS/PMF2, Traffic OA, industrial OA and SOA denote the HOA, F4 and the sum of SV-OOA and LV-OOA. For the CMB model, industrial OA represents the aggregate contribution from three processes: coke production, metal smelting and shipping/oil burning and SOA is the fraction un-apportioned to the primary sources.

Regardless a great number of uncertainties, artefacts and assumptions underlying the apportionments delivered by AMS/PMF and CMB[§], the inter-comparison between both

---

[§] Uncertainties/errors include: uncertainties in the measurement of markers and mass spectra, uncertainties in the PMF model (e.g. number of factors considered) and source profile selection in CMB (e.g. non-representative or non-considered profiles), uncertainties in OM/OC ratios, reactivity of organic markers, adsorption artefacts onto filters and differences in size cut-offs between HiVol samplers ($PM_{2.5}$) and AMS (PM_1).

models is excellent (Fig.S9). Further, as shown in Fig.2 in the manuscript for the case of industrial emissions, both models capture the same variability in the factor's contributions. However, AMS/PMF2 tends to provide less contribution from SOA (sum of OOAs), as unlike POA, this component is likely to be enhanced in the $PM_{2.5}$-PM_1 fraction and thus would not be measured by the AMS: ~20% of the carbon mass is estimated to occur in the $PM_{2.5}$-PM_1 fraction, mostly related to secondary matter and primary dust and biological particles. On the other hand, the CMB model under-apportions the industrial source compared to the AMS/PMF2, owing most probably to omitted profiles for fugitive industrial emissions in the CMB and to the occurrence of aged OA in the industrial plumes. These emissions are often concomitant with new particle formation events, when the site is impacted by land and sea breezes allowing the aging of the emissions (El Haddad et al., 2011a).

S.8 Apportionment of fossil and non-fossil OOA and related uncertainties

S.8.1 Multiple regression model

AMS/PMF apportionments and ^{14}C measurements are combined using a multiple regression model to estimate the fossil and the non-fossil contributions to both SVOOA and LVOOA. It is worthwhile to note that such a combination is not straightforward, involving a certain number of assumptions resulting in considerable uncertainties.

First of all, ^{14}C measurements are conducted on ***$PM_{2.5}OC$*** onto filter samples that are subjected to well-known but non systematic adsorption artefacts of gas phase organic compounds. In contrast, AMS provides real-time measurements of ***PM_1OA*** with little interference from gas phase organics. However, particle collection efficiency (CE) of the AMS estimated, here, (CE=0.65) can be highly variable, depending on the aerosol chemical nature and mixing state.

Second of all, ^{14}C measurement conducted in this study relates to the total carbon (TC) mass that can be oxidized at 850 °C under oxygen, i.e. organic carbon (OC) and elemental carbon (EC), whereas AMS quantifies OA that consists of OC and the associated heteroatom (H, N, O, S …). This fundamental difference engenders two major limitations for the assessment of fossil and non-fossil contributions of the OOA fractions. First, as the TC apportioned by ^{14}C measurements also includes EC, assumptions related to the origins of the latter must be made. Furthermore, since the separation between EC and OC measured using a Sunset analyser is method-dependent, biases associated with EC determination can impact the estimations. Second, the AMS measurement of OA includes heteroatom that can be unevenly distributed between the fossil carbon and the non-fossil carbon. Here also, this distribution is not empirically accessible and assumptions have to be considered to achieve our estimations. Further uncertainties can arise from PMF calculations and residuals and from variability in the biomass $^{14}C/^{12}C$ ratio. The assumptions made to achieve the apportionment are explicitly presented in this section and the resulting biases and uncertainties are thoroughly discussed in section S.8.2.2.

The procedure goes as follows: First, in order to estimate the fossil and non-fossil fractions of OC, EC was assumed to be entirely related to fossil carbon (assumption founded on Chemical Mass Balance calculations reported in El Haddad et al., 2011a). This is described in equations (4) and (5):

$$OC_f = TC_f - EC_f = TC_f - EC \left(EC_f = EC \parallel EC_{nf} = 0 \right) \tag{4}$$

$$OC_{nf} = TC_{nf} - EC_{nf} = TC_{nf} \left(EC_f = EC \parallel EC_{nf} = 0 \right) \tag{5}$$

Where TC_f, OC_f and EC_f correspond to the fossil TC, OC and EC, respectively, and TC_{nf}, OC_{nf} and EC_{nf} to the non-fossil TC, OC and EC, respectively.

The second step involves the transition from $PM_{2.5}OC$ measured on filter samples to PM_1OA determined by the AMS. This conversion is achieved for each of the samples by scaling OC_f and OC_{nf} to fossil OA (OA_f) and non-fossil OA (OA_{nf}), respectively, by a factor α:

$$OA_f = \alpha \times OC_f \left(\alpha = PM_1OA / PM_{2.5}OC \right) \tag{6}$$

$$OA_{nf} = \alpha \times OC_{nf} \left(\alpha = PM_1OA / PM_{2.5}OC \right) \tag{7}$$

where α is the ratio between AMS PM_1OA and filter $PM_{2.5}OC$. α is variable depending on the considered sample, but has an average of 0.92±0.21. The assumption underlying this scaling is that $PM_{2.5}OC$ and PM_1OA are associated with the same proportions of fossil and non-fossil mass (i.e. the ratio fossil/total is the same for $PM_{2.5}OC$ and PM_1OA). Biases arising form this assumption are discussed in the section S.8.2.2.

OA_f and OA_{nf} obtained in equations 6 and 7 can be accordingly expressed as a linear combination of the AMS/PMF OA fractions derived from fossil ($fOA_i : POA_f$, $SVOOA_f$ and $LVOOA_f$) and non-fossil ($nfOA_i : POA_{nf}$, $SVOOA_{nf}$ and $LVOOA_{nf}$) sources, respectively:

$$OA_f = \sum_i^l fOA_i = POA_f + SVOOA_f + LVOOA_f \tag{8}$$

$$OA_{nf} = \sum_i^m nfOA_i = POA_{nf} + SVOOA_{nf} + LVOOA_{nf} \tag{9}$$

where l and m are the total numbers of fOA_i and $nfOA_i$ fractions, respectively. As fOA_i and $nfOA_i$ are not directly accessible, equations 7 and 8 can be written in terms of the OA factors (OA_i) determined by AMS/PMF analysis and the respective share of fossil and non-fossil fractions to these factors such that

$$OA_f = \sum_i^l a_i \times OA_i = a_1 \times POA + a_2 \times SVOOA + a_3 \times LVOOA \quad \left(a_i = \frac{fOA_i}{fOA_i + nfOA_i} \right) \tag{10}$$

$$OA_{nf} = \sum_i^m b_i \times OA_i = b_1 \times POA + b_2 \times SVOOA + b_3 \times LVOOA \quad \left(b_i = \frac{nfOA_i}{fOA_i + nfOA_i} \right) \tag{11}$$

where a_i and b_i denote the relative share of fossil and non-fossil fractions to OA_i factors, respectively (a_1 for POA_f, a_2 for $SVOOA_f$, a_3 for $LVOOA_f$, b_1 for POA_{nf}, b_2 for $SVOOA_{nf}$ and b_3 for $LVOOA_{nf}$). This system of linear equations can be visualised as the following matrix equation:

$$\begin{bmatrix} POA & SVOOA & LVOOA & 0 & 0 & 0 \\ 0 & 0 & 0 & POA & SVOOA & LVOOA \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} OA_f \\ OA_{nf} \\ 1 \\ 1 \\ 1 \end{bmatrix} \tag{12}$$

As POA is assumed to be strictly related to fossil sources (i.e. $POA_f = POA = HOA + F4$), the parameter a_1 can be assumed to equal 1, implying that b_1 is equal to 0 (i.e. $POA_{nf} = 0$, the sensitivity of the results to this assumption is assessed in section S.8.2.2). Equation 12 can be then simplified as follows:

$$\begin{bmatrix} SVOOA & LVOOA & 0 & 0 \\ 0 & 0 & SVOOA & LVOOA \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} a_2 \\ a_3 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} SOA_f \\ SOA_{nf} \\ 1 \\ 1 \end{bmatrix} \qquad (13)$$

With SOA_f and SOA_{nf} denoting the fossil and non-fossil fractions of SOA, respectively, calculated as: $SOA_f = OA_f - POA_f$ and $SOA_{nf} = OA_{nf} - POA_{nf}$. The equations of the resulting linear system (equation 13) are not independent and thus the system has no solution. For that reason, a multiple linear regression analysis was applied instead to solve equation (13) finding the average values for a_i and b_i that fit best the equation, in the sense of solving the quadratic minimization problem. In equation (13) AMS/PMF OA_i vectors are included as independent variables and OA_f and OA_{nf} as dependant variables.

It should be noted that the apportionment procedure followed here is not exclusive. As this study focuses mainly on AMS measurements, the method chosen here is AMS data oriented in that the resulting apportionments would exhibit the same variability as the AMS/PMF factors (e.g. $SVOOA_f + SVOOA_{nf} = SVOOA$). This is the result of the equation (13), in which AMS/PMF OA_i vectors were chosen as independent variables. As this equation is a self-consistent system (i.e. $OA_f + OA_{nf} - POA = SVOOA + LVOOA$), another approach is also possible giving the same average result but orienting the variability towards ^{14}C measurements, by including OA_f and OA_{nf} as independent variables.

S.8.2 Output quality control, uncertainty assessments and potential biases

S.8.2.1 Output quality control and residual analyses

One of the major drawbacks of the multiple regression analysis applied here is that it considers a constant contribution of fossil and non-fossil sources to each of the OA_i factors (i.e. constant a_i and b_i ratios), while these contributions may significantly vary over the course of the measurements. Accordingly, a_i and b_i ratios should be regarded as average contributions of fossil and non-fossil sources to OA_i. These ratios are reported in Tab.S2.

Tab.S2: a_i and b_i ratios for the POA, SVOOA and LVOOA fractions.

	POA	SVOOA	LVOOA
a_i	1.0	0.33±0.11	0.082±0.085
b_i	0.0	0.67±0.11	0.92±0.08

Fig.S10 compares the measured and the modeled concentrations for the total fossil and non-fossil fractions. It shows that the model capture quite well the amounts and the variability of the measured concentrations, especially in the case of the fossil fraction (Fig.S10-A). In the case of the non-fossil fraction, the model tends to slightly underestimate (overestimate) the measured levels at low (high) concentrations (Fig.S10-B). It should be noted though that most of the variability observed in panels A and B of Fig.S10, arises from discrepancies between the $PM_1\ OA$ and filter-based $PM_{2.5}\ OC$, as shown in panel C of the same figure. Differences between the 2 measurement techniques were accounted for in equations 6 and 7 prior to the multiple regression analyses by the coefficient α that encompasses various conversion factors. Overall, these comparisons validate the representativeness of a_i and b_i obtained in the multiple regression analyses.

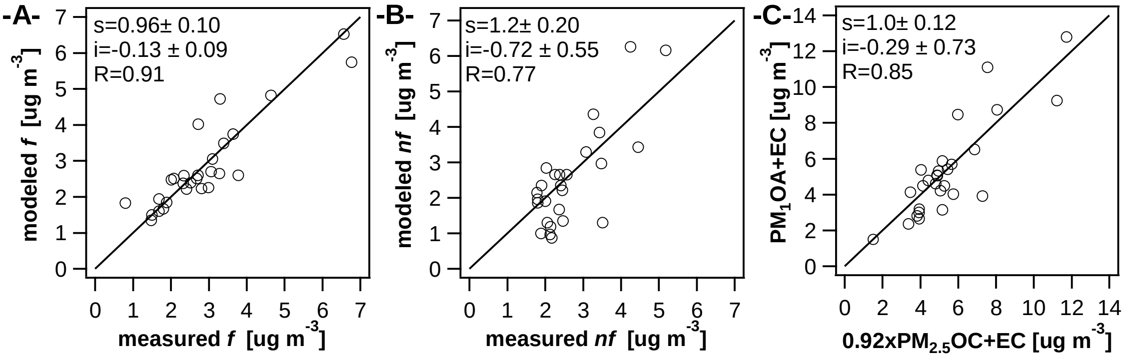

Fig.S10: Scatter plots of modeled vs. measured data for the fossil (**A**) and non-fossil (**B**) fractions. Measured fossil fraction=EC+0.92xOC$_f$, modeled fossil fraction=EC+HOA+F4+ SVOOA$_f$+LVOOA$_f$, Measured non-fossil fraction=0.92xOC$_f$, and modeled fossil fraction= SVOOA$_{nf}$+LVOOA$_{nf}$. The 0.92 value is the average value of the factor α used in equations 6 and 7. The comparison between filter measurements and AMS measurements is shown in panel C. For all panels, the slope of the linear regression (s), its intercept (i) and its coefficient of determination (R) are also indicated (n=28 samples for each plot).

Assumptions underlying the residual’s distributions are examined for the fossil and non-fossil fractions in Fig.S11. Residuals follow normal distributions with mean values statistically equal to zero, implying that errors are homoscedastic (variance = 0) and are not correlated. From Fig.S11, it is possible to estimate the uncertainties related to the total fossil and non-fossil OA fractions. Fossil and non-fossil OA are accordingly estimated to contribute 1.52±0.31 $\mu g\ m^{-3}$ (implying 20% of errors) and 2.52±0.78 $\mu g\ m^{-3}$ (implying 31% of errors), respectively. These uncertainties include: (1) measurement differences between filter-based $PM_{2.5}$ TC and AMS PM_1 OA + EC (Fig.S10 - panel C) and (2) variability in a_i and b_i obtained in the multiple regression analyses (see the related uncertainties in Tab.S2).

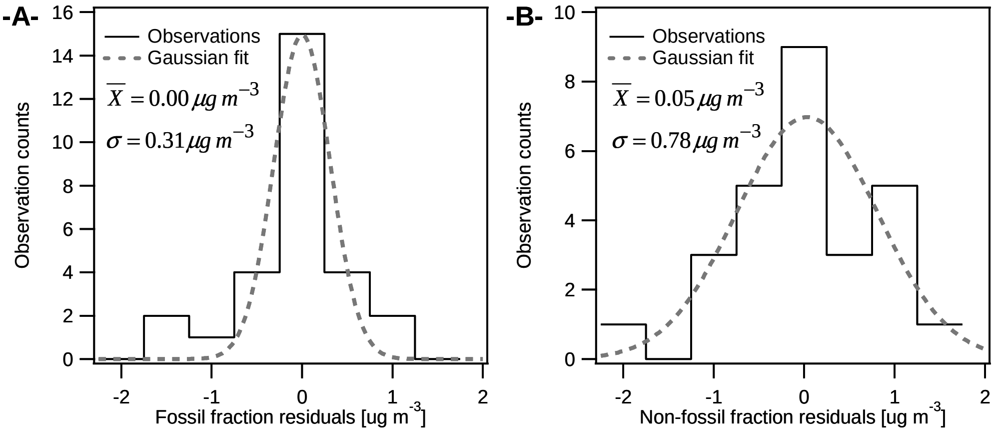

Fig.S11: Residuals normal distributions (modelled – measured) derived from the multiple regression approach applied above for the fossil (**A**) and the non-fossil (**B**) fractions. Residuals are fitted using a Gaussian fit, from which the mean $\left(\overline{X}\right)$ and the standard deviation *(σ)* are calculated for both fractions.

S.8.2.2 General assessment of uncertainties and biases

It is worthwhile to note that a great part of the uncertainties assessed for the absolute concentrations of fossil and non-fossil OA arises from discrepancies between AMS and filter measurements and hence not representative of the statistical significance of each of the fractions. The statistical significance of the relative contributions of HOA, F4, $SVOOA_f$, $LVOOA_f$, $SVOOA_{nf}$ and $LVOOA_{nf}$ are assessed through a sensitivity test, using a random selection technique. Inputs to the calculation are the PMF factor mass concentrations, ^{14}C data and OC/EC measurements. The calculation is performed based on equations (4-13); it proceeds as follows:

- For each of the input parameters a range is assigned, within which these can vary (see BoxTab.S1). The criteria on which we based our assessment of this range are developed below, in the TextBox.S1.
- The parameters are then allowed to randomly vary within the range predetermined in the previous step, assuming a normal distribution. This approach is somewhat similar to Monte Carlo calculations and allows vast numbers of combinations of input parameters to be computed. A Monte Carlo simulation would involve testing all possible combinations of input parameters, which is proven prohibitive in terms of processing time. In contrast, random sampling is much more effective and for our purposes provides essentially the same results as a full Monte Carlo analysis.
- Following the approach described above, fifty sets of parameters are generated randomly and used subsequently in the equations 4-13 to calculate the inputs for the multiple linear regression analysis (i.e. SOA_f , SOA_{nf} , $SVOOA$ and $LVOOA$).

  This provides for each set of parameters average values for a_i and b_i plus the corresponding uncertainties.
- For each set of the coefficients a_i and b_i previously generated, the average values of these coefficients are varied assuming a binomial distribution, derived based on the corresponding uncertainties provided by the multiple linear regression analyses. In this step and for each set of parameters, ten values are generated for a_i and b_i and used to compute the contributions of $SVOOA_f$, $LVOOA_f$, $SVOOA_{nf}$ and $LVOOA_{nf}$. This gave in total 500 different solutions.

A great advantage of this approach is that combinations of parameters which are very unlikely (e.g. that only the minimum-possible values from each parameter were used) will represent only a small percentage of the output. The obtained 500 solutions are presented graphically as a probability density (frequency distribution) of possible solutions to the source apportionment problem we have set up, as shown in Fig.S12.

This analysis provides strong support to our results, allowing the assessment of the uncertainties underlying our measurements and assumptions and offering a measure to our ability in separating the different components (statistical significance of each component and the likelihood between them). It shows that the uncertainties on our estimations depend on the component considered. Depending strictly on the PMF analysis errors, the uncertainties associated with the contributions of POA (HOA and F4) are less 10%. Conversely, for OOA components the uncertainties are less homogenous. For non-fossil OOAs the uncertainties are around 10% as these are well resolved by the regression model, whereas for fossil OOAs uncertainties are higher (~36% and ~58% for SVOOA$_f$ and LVOOA$_f$, respectively), as these are poorly resolved by the regression model and strongly dependant on the EC measurements and the assumptions on POA. All 6 fractions are statistically significant with contributions higher than 0 (Z equal 23, 7.4, 2.8, 1.7, 9.2, 12 for HOA, F4, SVOOA$_f$, LVOOA$_f$, SVOOA$_{nf}$ and LVOOA$_{nf}$ respectively, with Z=average/uncertainty).

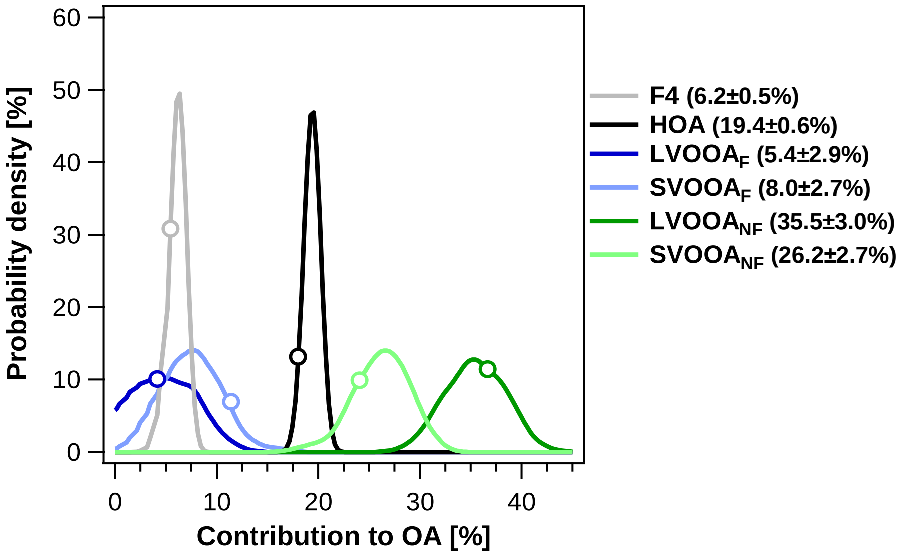

Fig.S12: Frequency distributions for the contribution of the different components assessed through the sensitivity test. The plot gives access to the average contribution and the associated standard deviation for the different components, indicated in the figure legend (Avg±1σ).

The sensitivity test offers as well the assessment of the biases on the apportionments presented in the manuscript. It suggest that we might underestimate the contributions of F4, HOA, SVOOA$_{nf}$ and LVOOA$_f$ by 9%, 6%, 5% and 23%, respectively and overestimate the SVOOA$_f$ and LVOOA$_{nf}$ by 34% and 5%, respectively. The main conclusion to be derived from this analysis is the robustness of the results presented and discussed in the manuscript. For example, OOA$_{nf}$ is clearly the biggest contributor to OA. It is also clear that LVOOA derives predominately from non-fossil precursors (LVOOA$_{nf}$ /LVOOA=89±7%), whereas SVOOA encompasses a bigger fraction of fossil SOA

($SVOOA_{nf}$/SVOOA=75±8%). Given the wide range of uncertainties used in the sensitivity test, these results demonstrate that in general we can clearly identify the contribution from different components.

TextBox.S1: Calculations of the different parameters entered as inputs to the uncertainty calculation

For the parameters in BoxTab.S1, the ranges were established as follows:

- For EC/OC measurements, the range is designed to encompass biases and uncertainties associated with the separation between EC and OC. This range is bounded by measurements determined following NIOSH and EUSAAR2 protocols, respectively. For EC and OC, a constant bias between the 2 protocols is obtained as 40±8% and 6±5%.
- The average uncertainty for the discrimination between fossil and non-fossil TC is 4%, including uncertainties on ^{14}C measurements and errors on the correction for ^{14}C inputs from the bomb testing.
- An assumption made in equations 4 and 5 relates to the origin of EC, estimated to only pertain to the fossil fraction. The transgression of this assumption would bias high the contributions of fossil sources to the secondary OC fractions. As there was too little influence from biomass burning, we assumed an upper limit contribution of non-fossil sources to EC of 15%, (based on Minguillón et al., 2011 and references), and varied this contribution between 0 and 15%.

BoxTab.S1: Ranges [Low, High] of the different parameters entered as inputs to the uncertainty calculation. E* denotes equations 4-13.

E*	Parameters	Variables	Low	High	Remarks
4, 5	OC_f, OC_{nf}	OC/EC	NIOSH	EUSAAR2	
		F_f	$0.96\times F_{nf}$	$1.04\times F_{nf}$	Uncertainties on measurements of ^{14}C in TC
		EC_f	$0.85\times EC$	EC	Origin of EC: $EC= EC_f + EC_{nf}$
6, 7	$\alpha=PM_1OA/PM_{2.5}OC$ see BoxTab.S2	α_1^{OC}	0.76	0.88	Diameter cut-offs: $\alpha_1^{OC}=PM_1OC/PM_{2.5}OC$
		α_2^{OC}	CI(-)	CI(+)	Positive artefacts based on Fig.S3: CI(-) and CI(+) are the upper and lower bounds of the confident interval on the linear regression
10,11	AMS/PMF2 OA	*Factors*	*FPEAK0*	*FPEAK1*	AMS/PMF2 results obtained for FPEAKs between 0 and 1.
12,13	POA_f, POA_{nf}	$a_1\times POA$	$0.75\times HOA+F4$	$0.9\times HOA+F4$	$a_1\times POA$ is the fraction of fossil POA. Its uncertainty is constrained based on Fig.S6

- The conversion from $PM_{2.5}OC$ measured on filter samples to PM_1OA determined by the AMS was performed in equations 6 and 7, using a common factor α (with $\alpha = PM_1OA/PM_{2.5}OC$) for both fossil and non-fossil OC. This factor encompasses three key corrections related to differences between the two measurement techniques, including differences in diameter cut-offs between AMS and filter sampling (referred to as $PM_1OC/PM_{2.5}OC$ ratio), the adsorption artefacts on the filters and the OM/OC ratio. The assumption underlying the $PM_{2.5}OC$ to PM_1OA conversion is that the aggregate of the aforementioned corrections is similar for both fossil and non-fossil OC and well represented by α.

TextBox.S1: Continues.

Indeed, these corrections can vary greatly between the primary and the secondary fractions, which exhibit variable contributions to the fossil and non-fossil OC. To address the biases resulting from the application of a single conversion factor α, the latter is deconvolved in equation (14) into several factors, such that

$$PM_1OA = \alpha \times PM_{2.5}OC = \sum_{i}^{n}\left(\prod_{j}^{p}\alpha_{ij}\right)\times OC_i \qquad (14)$$

In this equation, α_{ij} denotes the factor used for a conversion (*j*) applied to an OC_i fraction. n is the total number of OC fractions; in our case, it is limited to two fractions representing the primary and the secondary OC. p is the total number of α_i conversions applied to OC_i; in our case, p is equal 3, accounting for the $PM_1OC/PM_{2.5}OC$ ratio (*j*=1), for the sampling artefacts (*j*=2) and for the OM/OC ratio (*j*=3). The following is an example performed for the campaign average value, representing the matrix of α_{ij} factors for primary and secondary OC (BoxTab.S2). Similar calculations were performed for all the data set to achieve the sensitivity test. In this calculation, α_{ij} were measured for the total OC$\left(\alpha_j^{OC}\right)$, assumed for the primary OC $\left(\alpha_j^{POC}\right)$ and inferred for secondary OC$\left(\alpha_j^{SOC}\right)$. The different conversions include the following:

- α_1^{OC} denotes the $PM_1OC/PM_{2.5}OC$ ratio, estimated using size resolved OC measurements, i.e. 0.82±0.06%. POC was assumed to pertain entirely to the PM$_1$ fraction$\left(\alpha_1^{POC}=1\right)$, resulting in an α_1^{SOC} of 0.77 (i.e. 77% of the PM$_{2.5}$ SOC are comprised in the PM$_1$ fraction).
- α_2^{OC} is the ratio allowing the correction for the sampling artefacts retrieved from Fig.S3, with an average value of 0.72. Artefacts are assumed to be evenly distributed between the primary and the secondary fractions, i.e. $\alpha_2^{POC}=\alpha_2^{SOC}$ (BoxTab.S2).
- α_3^{OC} is the OM/OC ratio, obtained from the comparison between PM$_1$ AMS and LPI measurements. An average α_3^{OC} value of 1.67 was found, and assuming an α_3^{POC} value of 1.2 for primary OC a value of 1.81 can be inferred for the α_3^{SOC}.

TextBox.S1: Continues.

BoxTab.S2: α_{ij} factors estimated for POC, SOC and total OC fractions.

α_{ij}	α_j^{POC}	α_j^{SOC}	α_j^{OC}
α_{i1}	1.0	0.77	0.82
α_{i2}	0.72	0.72	0.72
α_{i3}	1.20	1.81	1.67
$\prod_j^p \alpha_{ij}$	0.84	0.98	0.95

From α_{ij} matrix, overall conversion factors $\prod_j^p \alpha_{ij}$ can be inferred for POC, SOC and total OC, which are equal to 0.84, 0.98 and 0.95, respectively. Estimated $\prod_j^p \alpha_j^{OC}$ (0.95) is comparable to the average α empirically determined and used in equations (6) and (7) to convert from $PM_{2.5}OC$ measured on filter samples to PM_1OA determined by the AMS $(\alpha = 0.92 \pm 0.21)$.

Using $\prod_j^p \alpha_j^{POC}$ and $\prod_j^p \alpha_j^{SOC}$ obtained above, one can apply different conversion factors to POA and SOA. In the sensitivity test such a calculation has been made by considering a range of α_1^{OC} and α_2^{OC} (BoxTab.S2).

- In the calculation of equations 10 and 11, we considered for the apportionments the AMS/PMF2 results, including *POA*, *SVOOA* and *LVOOA*. One approach to assess the uncertainties on the AMS/PMF2 apportionments consists of varying FPEAK within a reasonable range, in our case between FPEAK=0 and FPEAK=1.
- In the calculation of equations 12 and 13, we considered that POA pertains only to the fossil fraction, while we observed some evidence of some inputs from cooking (a non fossil primary source) to HOA. To take this into account in the uncertainty calculations, we considered that these inputs contribute between 10% and 25% of total HOA, based on Fig.S6.

References

Agrawal, H., Welch, W. A., Miller, J. W., and Cocker, D. R.: Emission measurements from a crude oil tanker at sea, Environ. Sci. Technol., 42, 7098–7103, 2008.

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., and Jimenez, J. L.: O/C and OM/OC Ratios of Primary, Secondary, and Ambient Organic Aerosols with HighResolution Time-of-Flight Aerosol Mass Spectrometry, Environ. Sci. Technol., 42, 4478–4485, doi:10.1021/es703009q, 2008.

Alfarraa, M.R., H.Coe, J. D. Allan, K. N. Bowerb, H. Boudriesc, M.R. Canagaratna, J. L. Jimenez, J.T. Jayne, A. A. Garforth, S-M. Li, D. R. Worsnop, Characterization of urban and ruralorganic particulate in the Lower Fraser Valley using two Aerodyne Aerosol Mass Spectrometers, Atmos. Env. 2004, 38, 5745-5758.

Bae, M.-S., Schauer J.J., DeMinter, J.T., Turner, J.R., Smith, D. and Cary, R.A.: Validation of a semi-continuous instrument for elemental carbon and organic carbon using a thermal-optical method, Atmos. Environ., 38, 2885–2893, 2004.

Bahreini, R., M.D. Keywood, N.L. Ng, Varutbangkul V., S. Gao, R. C. Flagan, J.H. Seinfeld, D.R. Worsnop and J.L. Jimenez, Measurements of Secondary Organic Aerosol from Oxidation of Cycloalkenes, Terpenes, and m-Xylene Using an Aerodyne Aerosol Mass Spectrometer, Environ. Sci. Technol., 39, 5674-5688, 2005.

Cachier, H., Aulagnier, F., Sarda, R., Gautier, F., Masclet, P., Besombes, J. L., Marchand, N., Despiau, S., Croci, D., Mallet, M., Laj, P., Marinoni, A., Deveau, P. A., Roger, J. C., Putaud, J. P., Van Dingenen, R., Dell'Acqua, A., Viidanoja, J., Santos, S. M. D., Liousse, C., Cousin, F., Rosset, R., Gardrat, E., and GalyLacaux, C.: Aerosol studies during the ESCOMPTE experiment: an overview, Atmos. Res., 74, 547–563, 2005.

Canagaratna, M.R., Jayne, J.T., Ghertner, D.A., et al.: Chase studies from in-use New-York city vehicles, Aerosol Sci. Technol., 38, 555-573, 2004.

Cavalli, F., Viana, M., Yttri, K. E., Genberg, J., and Putaud, J.-P.: Toward a standardised thermal-optical protocol for measuring atmospheric organic and elemental carbon: the EUSAAR protocol, Atmos. Meas. Tech., 3, 79–89, doi:10.5194/amt-3-79-2010, 2010.

Chauvel, C., Bureau, S., and Poggi, C.: Comprehensive chemical and isotopic analyses of basalt and sediment reference materials, Geostand. Geoanalyt. Res., 35, 125–143, doi:10.1111/j.1751-908X.2010.00086., 2010.

Chirico, R., A.S.H. Prevot, P.F. DeCarlo, M.F. Heringa, R. Richter, E. Weingartner, and Urs Baltensperger. Aerosol and Trace Gas Vehicle Emission Factors Measured in a Tunnel Using an Aerosol Mass Spectrometer and other On-line Instrumentation. Atmospheric Environment, 45(13), 2182-2192, 2011.

Drewnick, F., Hings, S. S., DeCarlo, P., Jayne, J.T., Gonin, M., Fuhrer, K., Weimer, S., Jimenez, J.L., Demerjian, K.L., Borrmann, S., Worsnop, D.R.: A new time-of-flight aerosol mass spectrometer (TOF-AMS) – Instrument description and first field deployment, Aerosol Sci. Tech., 39, 637–658, 2005.

El Haddad, I., Marchand, N., Dron, J., Temime-Roussel, B., Quivet, E., Wortham, H., Jaffrezo, J. L., Baduel, C., Voisin, D., Besombes, J. L., and Gille, G.: Comprehensive primary particulate organic characterization of vehicular exhaust emissions in France, Atmos. Environ., 43, 6190–6198, 2009.

El Haddad, I., Marchand, N., Wortham, H., Piot, C., Besombes, J. -L., Cozic, J., Chauvel, C., Armengaud, A., Robin, D., Jaffrezo. J. -L., Primary sources of PM2.5 particles in an industrial Mediterranean city, Marseille, Atmospheric Chemistry and Physics, 11, 2039–2058, 2011a.

El Haddad, I., Marchand, N., Temime-Roussel, B., Wortham, H., Piot, C., Besombes, J. -L., Baduel, C., Voisin, D., Armengaud, A., Jaffrezo, J. -L., Insights into the secondary fraction of the organic aerosol in a Mediterranean urban area: Marseille, Atmospheric Chemistry and Physics, 11, 2059–2079, 2011b.

Fine, P. M., Cass, G. R., and Simoneit, B. R. T.: Chemical Characterization of Fine Particle Emissions from the Fireplace Combustion of Woods Grown in the Southern United States, Environ. Sci. Technol., 36, 1442–1451, doi:10.1021/es0108988, 2002.

Flaounas, E., Coll, I., Armengaud, A., Schmechtig, C.: The representation of dust transport and missing urban sources as major issues for the simulation of PM episodes in a Mediterranean area, Atmos. Chem. Phys., 9, 8091–8101, doi:10.5194/acp-9-8091- 2009, 2009.

Kleindienst, T. E., Jaoui, M., Lewandowski, M., Offenberg, J. H., Lewis, C. W., Bhave, P. V., and Edney, E. O.: Estimates of the contributions of biogenic and anthropogenic hydrocarbons to secondary organic aerosol at a southeastern US location, Atmos. Environ., 41, 8288–8300, 2007.

Lanz, V.A., Alfarra, M.R., Baltensperger, U., Buchmann, B., Hueglin, C., and Prévôt, A.S.H.: Source apportionment of submicron organic aerosols at an urban site by factor analytical modelling of aerosol mass spectra, Atmos. Chem. Phys., 7, 1503-1522, 2007.

Levin, I., Kromer, B., Schochfischer, H., Bruns, M., Munnich, M., Berdau, D., Vogel, J. C., and Munnich, K. O.: 25 Years of Tropospheric C-14 Observations in Central-Europe, Radiocarbon, 27, 1-19, 1985.

Levin, I., and Hesshaimer, V.: Radiocarbon - A unique tracer of global carbon cycle dynamics, Radiocarbon, 42, 69-80, 2000.

Lindinger, W., Hansel, A., and Jordan, A.: On-line monitoring of volatile organic compounds at pptv levels by means of protontransfer-reaction mass spectrometry (PTR-MS) – Medical applications, food control and environmental research, Int. J Mass. Spectrom., 173, 191–241, 1998.

Paatero, P. and Tapper, U.: Positive matrix factorization: a nonnegative factor model with optimal utilization of error estimates of data values, Environmetrics, 5, 111-126, 1994.

Petaja, T., Kerminen, V.-M., Dal Maso, M., Junninen, H., Koponen, I. K., Hussein, T., Aalto, P. P., Andronopoulos, S.,Robin, D., Hameri, K., Bartzis, J. G., and Kulmala, M.: Submicron atmospheric aerosols in the surroundings of Marseille and Athens: physical characterization and new particle formation, Atmos. Chem. Phys., 7, 2705–2720, doi:10.5194/acp-7-2705-2007, 2007.

Rolph, G. D.: Real-time Environmental Applications and Displays Ystem (READY) Website (http://ready.arl.noaa.gov). NOAA Air Resources Laboratory, Silver Spring, MD, 2010.

Rogge, W. F., Hildemann, L. M., Mazurek, M. A., Cass, G. R., and Simoneit, B. R. T.: Sources of Fine Organic Aerosol .4. Particulate Abrasion Products from Leaf Surfaces of Urban Plants, Environ. Sci. Technol., 27, 2700–2711, 1993a.

Rogge, W. F., Hildemann, L. M., Mazurek, M. A., Cass, G. R., and Simoneit, B. R. T.: Sources of Fine Organic Aerosol.5. NaturalGas Home Appliances, Environ. Sci. Technol., 27, 2736–2744, 1993b.

Schmid, H., Laskus, L., Jurgen Abraham, H., Baltensperger, U., Lavanchy, V., Bizjak, M., Burba, P., Cachier, H., Crow, D., Chow, J., Gnauk, T., Even, A., ten Brink, H. M., Giesen, K.-P., Hitzenberger, R., Hueglin, C., Maenhaut, W., Pio, C., Carvalho, A., Putaud, J.-P., Toom-Sauntry, D., and Puxbaum, H.: Results of the “carbon conference” international aerosol carbon round robin test stage I, Atmos. Environ., 35, 2111–2121, 200.

Tsai, J.-H., Lin, K.-H., Chen, C.-Y., Ding, J.-Y., Choa, C.-G., and Chiang, H.-L.: Chemical constituents in particulate emissions from an integrated iron and steel facility, J. Hazard. Mater., 147, 111–119, 2007.

Ulbrich, I.M., Canagaratna, M.R., Zhang, Q., Worsnop, D.R., and Jimenez, J.L.: Interpretation of organic components from Positive Matrix Factorization of aerosol mas spectrometric data, Atmos. Chem. Phys., 9, 2891-2918, 2009.

Watson, J. G., Robinson, N. F., Fujita, E. M., Chow, J. C., Pace, T. G., Lewis, C., and Coulter, T.: CMB8 Applications and Validation Protocol for PM2.5 and VOCs, US EPA, USA, 1998.

Weitkamp, E. A., Lipsky, E. M., Pancras, P. J., Ondov, J. M., Polidori, A., Turpin, B. J., and Robinson, A. L.: Fine particle emission profile for a large coke production facility based on highly time-resolved fence line measurements, Atmos. Environ., 39, 6719–6733, 2005.

Zhang, Q., Alfarra, M.R., Worsnop, D.R., Allan, J.D., Coe, H., Canagaratna, M.R., and Jimenez, J.L.: Deconvolution and quantification of hydrocarbon-like and oxygenated organic aerosols based on aerosol mass spectrometry, Environ. Sci. Technol., 39, 4938-4952, 2005.