Airborne measurements of the spatial distribution of aerosol chemical composition across Europe and evolution of the organic fraction: Supplementary material

W. T. Morgan[1], J. D. Allan[1,2], K. N. Bower[1], E.J. Highwood[3], D. Liu[1], G. R. McMeeking[1], M. J. Northway[3], P. I. Williams[1,2], R. Krejci[4] and H. Coe[1]

1. *Centre for Atmospheric Science, University of Manchester, Manchester, UK*
2. *National Centre for Atmospheric Science, University of Manchester, UK*
3. *Department of Meteorology, University of Reading, UK*
4. *Department of Applied Environmental Science, Atmospheric Science Unit, Stockholm University, Sweden*

1 Scope

The supplementary material outlined in this document is provided in order to present the meteorological context of the flight operations and support the analysis techniques and data quantification steps outlined in the main paper. The meteorological fields corresponding to each flying period are presented and further information regarding the photochemical context of the operations is presented. Further details regarding the volume closure between the Aerosol Mass Spectrometer (AMS) and the Passive Cavity Aerosol Spectrometer Probe (PCASP) are discussed. Comparison of the estimated HOA with primary combustion tracers is included. The relationship between the fractional contribution of Low-Volatility Oxygenated Organic Aerosol (LV-OOA) to the organic mass and the normalised organic signal at m/z 44 is also shown. Further information is provided regarding the Positive Matrix Factorisation (PMF) analysis examples from the main text, as well as a summary of some PMF diagnostics for the whole dataset. The PMF analysis was performed using the tools presented by Ulbrich et al. (2009).

2 Meteorological summary

Figs. S1 and S2 display the typical meteorological conditions prevalent during each period considered by the analysis. The periods are relatively consistent in terms of their transport patterns, with the air masses transporting pollution from continental Europe downwind to either the UK region or into the Eastern Atlantic Ocean. Thus the flights are predominantly focused upon either sampling such pollution over continental Europe itself or at a range of scales downwind.

The evolution of the aerosol chemical composition during the LONGREX-2 period was examined based upon the relatively consistent transport patterns prevalent during the period. Fig. S3a displays the back trajectories for each flight during this period based upon Straight and Level Runs (SLRs) during each flight. The trajectories display highly consistent behaviour during the period, which is unsurprising given the relative stability of the high pressure system located over Northern Europe during this period. Fig. S3b highlights the back trajectory from the 14 May 2008, which was initialised from a SLR during B374 in the Eastern Atlantic Ocean. B374 represented the end-point in our operations during this period both in terms of the geographical location of the missions and also the distance from continental European sources i.e. the most aged polluted air mass. The back trajectory indicates that the spatial coverage of the flight operations closely matches the air mass transport during the period leading up to the 14 May 2008. Specifically, flights B370-B374 took place across Northern Europe during this period covering close to 5 days of air mass transport.

3 AMS versus PCASP comparison

Validation of the collection efficiency treatment applied to the dataset following the principles developed by Matthew et al. (2008) is accomplished by comparing the AMS data with the volume estimated concentrations from the PCASP instrument. The AMS total mass concentrations were converted to total volume concentrations using the densities reported by Cross et al. (2007), which correspond to 1.27 $\mathrm{g\,cm^{-3}}$ for organics and 1.77 $\mathrm{g\,cm^{-3}}$ for inorganics. A comparison of the estimated volume from the AMS and PCASP is shown for SLRs below 3000 m in Fig. S4. Over all of the considered flights, the estimated AMS volume concentrations were 26% higher than the estimated PCASP volumes. This average agreement is predominantly determined by the LONGREX flights, which were quite consistent in terms of the agreement from flight-to-flight. The ADIENT flying periods sit on either side of the overall regression slope, with the ADIENT-2 flying indicating that the PCASP volume was 48% of the AMS volume. These discrepancies are considered tolerable given the large uncertainties previously reported in the literature for PCASP volume estimates (e.g. Moore et al., 2004; Hallar et al., 2006) and the uncertainties in the AMS volume estimates.

For B357, the PCASP volume estimate was more than two times greater than the AMS volume estimate, which is outside of the bounds of uncertainty for the two instruments. The reason for this discrepancy is unknown but could reflect an artifact in either instrument or the presence of material that is not detected by the AMS. The discrepancy between the two instruments is also reflected in the calculated volume-scattering

relationship when comparing the measurements with a nephelometer system. The main difference between B357 and the other flights in the dataset is the sampling altitude of the aircraft, where in B357, the aircraft operated at a constant altitude of 200 m for the majority of the flight. The other flights in the dataset operated at altitudes higher in the boundary layer. Potentially, the nephelometer and PCASP measurements could be perturbed by the constant low-level flying in a humid environment as the aerosol sampled may not be sufficiently dried in the inlet lines and by the heater respectively. The AMS volume estimates do not include water, so this could potentially cause the discrepancy. Additionally, the PCASP and nephelometer may be measuring refractory material or particles above the cut off of the AMS aerodynamic lens. This would also lead to the AMS underestimating the volume relative to the PCASP.

4 Photochemical context

The relationship between O_3 and CO with the O_3:NO_x ratio discussed in the main paper is presented in Fig. S5. The results indicate that O_3 increases and CO decreases steadily in the 1-100 O_3:NO_x range, which is a reflection of photochemistry and dilution respectively. Beyond an O_3:NO_x ratio of 100, the concentrations decrease with CO returning to background levels and O_3 remaining relatively constant in the 40-60 ppb range.

5 Positive Matrix Factorisation

Potentially, the most challenging and subjective aspect of PMF analysis is the selection of the appropriate number of factors. For AMS datasets, this is usually accomplished using internal PMF diagnostics, similarity to reference mass spectra and external measurement parameters. An example of an internal diagnostic is the parameter $Q/Q_{expected}$, which is defined as the total sum of the scaled residuals, divided by its expected value. This expected value is derived based upon the error estimates for the data matrix (Ulbrich et al., 2009). A value of unity for the $Q/Q_{expected}$ parameter indicates that the expected variance associated with random errors can be explained by the solution set. Values greater than unity indicate that there is additional variance not accounted for by the solution set. The suite of aerosol and gas phase instrumentation available on the aircraft provides several necessary external parameters to facilitate validation of the solution. Reference spectra utilised in this study are taken from the AMS spectral database (`http://cires.colorado.edu/jimenez-group/AMSsd/`) described in Ulbrich et al. (2009). The factor solutions are interpreted based upon comparisons with external parameters and reference mass spectra. Comparisons between mass spectra are accomplished using the Uncentered Correlation (UC, Ulbrich et al., 2009) coefficient, while the comparisons with external parameters use the Pearson's R coefficient.

Here we include additional information regarding the PMF solutions used in the main manuscript. This includes example PMF solutions from a range of flights in different conditions, a summary of the results for the whole dataset and further details

regarding the estimation of HOA from our data.

5.1 Example PMF solutions

Several PMF solutions will be illustrated in the following section based upon some example flights from the dataset. The flights chosen are B357 (16 April 2008), B362 (6 May 2008), B369 (10 May 2008) and B406 (25 September 2008) as these broadly represent the range of operations conducted in terms of their proximity to pollution sources. B357 was conducted primarily downwind of the major conurbations of Manchester and Liverpool, in the North-West of England. B362 was conducted on a South-North transect originating from Oberpfaffenhofen in southern Germany, culminating in a sequence of Straight-and-Level Runs (SLRs) in the North Sea in the outflow from continental Europe. B406 was conducted in the outflow from continental Europe along the southern coast of the UK. B369 was conducted in the Baltic Sea region and reflects background conditions for comparison with the more polluted examples.

The results for B357 are shown in Fig. S6, for B362 in Fig. S7, for B369 in Fig. S8 and B406 in Fig, S9. All the examples show excellent agreement between the measured and reconstructed organic mass concentrations.

5.1.1 Close to source case study – B357

For B357, a 2-factor solution was deemed most appropriate as increasing the number of factors led to a phenomenon known as "splitting". This leads to multiple factors being assigned with numerically similar factor profiles and time series. In this case, the additional factors merely represented variability within the factors identified in the 2-factor solution. Consequently, the 2-factor solution was retained for analysis. The $Q/Q_{expected}$ was equal to 1.38. Reconstructed OM concentrations made up 98% of the measured OM concentrations. Factor 1 (OOA-1) exhibits a correlation coefficient of 0.89 and 0.98 when compared to reference spectra for an ambient rural case (from Canada) and laboratory generated fulvic acid respectively, which has chemical functionalities that are representative of aged OM (e.g. McFiggans et al., 2005) and hence is indicative of an LV-OOA type factor. Factor 2 (OOA-2) is interpreted as HOA as it exhibits a high correlation (0.91) with the derived HOA mass spectrum from Pittsburgh (Zhang et al., 2005a). Additionally, it has a correlation coefficient of 0.87 when compared to diesel exhaust from a chase study in New York (Canagaratna et al., 2004). The factor 2 time series correlates with NO_x (r=0.65), CO (r=0.73) and BC (r=0.79), with the strong gradients in the time series coincident with large increases in NO_x, CO and BC. Cororally, factor 1 exhibits low correlation (r<0.35) with these combustion tracers. Comparison of the emission ratio of POA to NO_x from B357, which is equal to 29.9 $\mu g\,sm^{-3}$ ppm^{-1}, with the previous studies summarised in Table S1 yields good agreement. Furthermore, the emission ratio relative to CO of 8.0 $\mu g\,sm^{-3}$ ppm^{-1} falls within the literature values summarised in Table S1. Therefore, it appears that factor 2 in this case is likely attributable to primary sources of OM in the form of HOA.

5.1.2 Continental European scale case studies – B362 and B406

While B357 presented a relatively straight forward 2-factor solution, B362 and B406 present highly complex examples of the factor analysis. When more than 3 factors were examined, the "splitting" phenomenon was observed. The 3-factor solutions contained a factor profile strongly resembling LV-OOA, with *m/z* 44 dominating the mass spectrum and a high correlation with the reference spectra for fulvic acid and the ambient rural case. The second factor is characterised by a mass spectrum with similar intensities at m/z43 and 44 and an enhanced base peak at *m/z* 55. Such a spectrum is fairly typical of a SV-OOA type component (Ulbrich et al., 2009; Jimenez et al., 2009). The remaining factor appeared to represent a second SV-OOA mass spectrum but with enhanced mass peaks associated with commonly resolved hydrocarbon peaks (*m/z*'s 27, 29, 41, 43, 55, 57, 69, 71, ...), indicating a contribution of HOA. This third factor was correlated with NO_x, CO and BC but for B362 the slope was found to be approximately a factor of 2 greater than the range of literature values cited in Table S1. A regression for the same factor from B406 with NO_x and BC also yielded a similar discrepancy although a CO measurement was not available during the flight so a comparison with CO was not possible. Enhanced signal is not identified at *m/z* 60 or 73 during either flight. These are typical mass spectral markers for wood burning emissions (Alfarra et al., 2007), therefore solid fuel burning is not considered to be a potential source of the enhanced mass. Consequently, it appears that the 3-factor solution is "blending" more than one distinct organic component into a single factor. This is likely a consequence of there being more chemical variability in the SV-OOA component (arising from evaporation/condensation, aging etc.) than in the HOA component. Thus the PMF solution identifies a factor which represents a 'mathematical mixing' of the more recently formed OOA with a HOA component. This arises due to the commonality between some of the major peaks in their respective mass spectra e.g. *m/z* 29 and 43.

The number of factors was increased under the supposition that an increase beyond 3-factors would reveal a more realistic contribution of HOA to the OM by separating it from the more volatile, fresher OOA fraction. While OOA factor profiles generally display significant variability as a result of differing processes such as aging and partitioning, HOA factor profiles are far more chemically distinct. Additionally, the largely linear association between HOA concentrations and urban primary emission markers makes source identification more straightforward than for OOA. Based upon these criteria, the number of factors was increased and inspected at each step until a single factor was present that most resembled HOA. For B362, this occurred for a 6-factor solution while for B406, a 7-factor solution was chosen. These factors were chosen due to their strong resemblance to HOA based upon their mass spectra and comparisons with combustion tracers. The comparison with the combustion tracers for B362 is shown in Fig. S14b, which indicates that increasing the number of factors does appear to bring the HOA-type component closer to the literature emission ratio. The main deviation from the relationship is from a low-level SLR in a highly moist layer (RH>95%) where total mass concentrations exceeded 50 $\mu g\,sm^{-3}$ and ammonium nitrate was the dominant chemical component. The layer is likely characteristic of freshly formed secondary material and it appears that some of the freshest-OOA mass has been apportioned to the HOA profile. The number of factors was increased

to 15 but this did not significantly alter the mass apportionment during this event.

5.2 Bootstrapping analysis

The numerical stability of the factor solutions was quantitatively evaluated using a bootstrapping analysis (Ulbrich et al., 2009, and references therein), where random resampling of the data matrix is performed in the time dimension. This analysis was performed using 20 iterations, with the results being grouped according to the UC coefficient between mass spectral profiles. The results of this analysis are summarised in Fig. S14a for B362 by comparing the contribution to total mass versus *m/z* 44 for each factor as the solutions are stepped through an increasing number of factors. The derived mean and standard deviations from the bootstrapping analysis are compared with the base solutions. The analysis indicates that by increasing the number of factors, the solutions become increasingly unstable in a numerical sense. This is a consequence of the aforementioned chemical variability inherent in the air masses sampled during the flight operations, which results in large scope for different factor solutions for larger numbers of factors. The bootstrapping analysis suggests that the 2-factor solution is the most appropriate solution, especially as the OOA-1 component is highly robust with a close match between the base solution and the bootstrapping solution. The enhanced standard deviation in the *m/z* 44 for the OOA-2 component is likely a consequence of the variability in the chemical nature of the OM. Very similar results were derived for B406, with the 2-factor solution being more numerically robust than subsequent solutions.

Further results from the bootstrapping analysis are shown for B357 in Fig. S13, for B362 in Fig. S14, for B369 in Fig. S15 and for B406 in Fig. S16. The bootstrapping results for all the flights are summarised in Table S2. The variance of the solutions in both the time series and factor profile dimensions is evaluated using suitable metrics. The time series diagnostic is the mean of the standard deviation for each factor, reported as a percentage of the overall mean mass concentration. For the mass spectra, the greatest standard deviation for each factor profile from the bootstrapping analysis is reported. A mean is not deemed appropriate to evaluate the stability of the mass spectra as the chemical assignment of factors is performed based on a limited number of peaks (i.e. less than 10).

The OOA-1 (LV-OOA) factor profiles are highly robust with little deviation between the average mass spectrum from the bootstrapping analysis and the base solution. Furthermore, the standard deviations are typically low. This is a consistent result throughout the dataset, which is shown by the low scores for the diagnostics in Table S2. The OOA-2 factor profiles are more variable for B362, B369 and B406, with less stability in the signal intensity at m/z 44. This reflects the continuum nature of the OM discussed in the main text, whereby there is significant variability on a given flight in the level of oxidation. Thus by randomly resampling the dataset in the time dimension using the bootstrapping procedure, conditions with either enhanced or diminished m/z 44 in the OOA-2 may be encountered and this is reflected by enhanced standard deviations in both the time series and mass spectra. The normalised standard deviation for the time series of OOA-2 for B369 is much greater than the other flights in the dataset. This is predominantly a result of the low concentrations encountered during the flight,

coupled with large standard deviations which are associated with large changes in the composition of the OOA-2 component in this instance. For much of the flight, signal at m/z 57 is close to zero indicating that HOA makes a minimal contribution to the OOA-2 component. However, during the peak OM event at 15:47 UTC, signal at m/z 57 is enhanced and thus the contribution of HOA is likely elevated. This event coincides with the maximum in the NO_x concentration and is associated with potential sampling of low-level urban outflow from Stockholm into the Baltic Sea. During this period, the standard deviations for the OOA-2 factor increase. Thus such changes in the OOA-2 composition from SV-OOA dominated to HOA dominated are reflected by large increases in the standard deviation values from the bootstrapping analysis. The much lower normalised standard deviation values associated with the time series of the OOA-2 components for the rest of the dataset suggest this feature to be atypical. This is a reflection of the regional nature of the measurements, with few instances of prolonged exposure to intense urban signatures.

The results presented here demonstrate the robustness of the chosen 2-factor solutions in terms of both the mass spectra and time series of the components. The OOA-1 (LV-OOA) components are highly numerically stable, while the decreased numerical stability of the OOA-2 (SV-OOA and HOA) components is entirely consistent with the continuum of oxidation/aging discussed in the main text.

5.3 Application to the entire dataset

The remaining flights in the dataset were analysed in an identical manner to the framework established in the previous section. This resulted in broadly similar behaviour in terms of the inability to accurately and quantitatively resolve HOA. A consistent theme was that increasing the number of factors in order to attempt to separate the HOA contribution led to a numerically unstable solution. Thus we chose to use the 2-factor solutions as these consistently represented a more quantitative solution set.

A summary of the $Q/Q_{expected}$ parameter is shown in Fig. S12, which indicates that the parameter is generally less than 2. Four flights had $Q/Q_{expected}$ values greater than 2, with the largest value being 5.01 (B374). Such values are greater than is generally considered optimal if attempting to produce a perfect characterisation of the dataset but given the difficulty in deriving robust results when more than 2 factors are chosen, it is not possible to reduce Q further. Consequently, the additional Q contribution prevalent in the dataset is considered as 'model error'.

The available solutions include some rotational ambiguity, which is explored by varying a parameter known as "fPeak" (Ulbrich et al., 2009, and references therein). An fPeak range from from -2.5 to 2.5 is investigated in order to explore the numerical variability in factor profiles and time series for small changes in $Q/Q_{expected}$. Investigation of the rotational freedom in the solutions using fPeak was accomplished by inspecting the mass spectra and time series in relation to external tracers for a subset of fPeak values from -2.5 to 2.5. The most appropriate value was then chosen, which for this dataset was determined to be zero in all cases. A test of the numerical uniqueness of the solution sets is the dependence upon the initiation seed, which is described by Ulbrich et al. (2009). Each of the 2-factor solutions was examined using this technique and little variation was exhibited for a range of different seeds.

5.3.1 Estimation of HOA and comparison with OOA components

Included in Fig. S11a are correlations of the estimated HOA concentration with Black Carbon (BC), NO_x and CO. These indicate that for 8 of the flights, the correlations of the estimated HOA with these primary emission tracers are greater than 0.5. Correlations lower than 0.5 are generally encountered on flights where these tracers and the estimated HOA are very low, thus the correlations break down at values when the relationships exhibit enhanced noise due to low signal. This is demonstrated in Fig. S11b and c, where at low concentrations the relationships are relatively flat but at enhanced concentrations, the correlation is greater. Given the simple nature of the HOA estimate, these correlations and relationships do provide some confidence that the estimated HOA provides a qualitative indicator of the contribution of HOA to the OM burden. The HOA estimate using this approach is likely an upper limit as the contribution of any oxidised fragments at *m/z* 57 has not been removed.

Also shown in Fig. S11a are the correlations between the Low-Volatility Oxidised Organic Aerosol (LV-OOA) organic mass fraction and the m/z 44:OM ratio described in the main paper.

References

Alfarra, M. R., Prevot, A. S. H., Szidat, S., Sandradewi, J., Weimer, S., Lanz, V. A., Schreiber, D., Mohr, M., and Baltensperger, U.: Identification of the mass spectral signature of organic aerosols from wood burning emissions, Environmental Science & Technology, 41, 5770–5777, 2007.

Allan, J. D., Williams, P. I., Morgan, W. T., Martin, C. L., Flynn, M. J., Lee, J., Nemitz, E., Phillips, G. J., Gallagher, M. W., and Coe, H.: Contributions from transport, solid fuel burning and cooking to primary organic aerosols in two UK cities, Atmospheric Chemistry and Physics, 10, 647–668, 2010.

Canagaratna, M. R., Jayne, J. T., Ghertner, D. A., Herndon, S., Shi, Q., Jimenez, J. L., Silva, P. J., Williams, P., Lanni, T., Drewnick, F., Demerjian, K. L., Kolb, C. E., and Worsnop, D. R.: Chase studies of particulate emissions from in-use New York City vehicles, Aerosol Science and Technology, 38, 555–573, 2004.

Cross, E. S., Slowik, J. G., Davidovits, P., Allan, J. D., Worsnop, D. R., Jayne, J. T., Lewis, D. K., Canagaratna, M., and Onasch, T. B.: Laboratory and Ambient Particle Density Determinations using Light Scattering in Conjunction with Aerosol Mass Spectrometry, Aerosol Science and Technology, 41, 343–359, URL `http://www.informaworld.com/10.1080/02786820701199736`, 2007.

de Gouw, J., Middlebrook, A., Warneke, C., Goldan, P., Kuster, W., Roberts, J., Fehsenfeld, F., Worsnop, D., Canagaratna, M., Pszenny, A., Keene, W., Marchewka, M., Bertman, S., and Bates, T.: Budget of organic carbon in a polluted atmosphere: Results from the New England Air Quality Study in 2002, Journal of Geophysical Research-Atmospheres, 110, doi:10.1029/2004JD005623, 2005.

Hallar, A. G., Strawa, A. W., Schmid, B., Andrews, E., Ogren, J., Sheridan, P., Ferrare, R., Covert, D., Elleman, R., Jonsson, H., Bokarius, K., and Luu, A.: Atmospheric Radiation Measurements Aerosol Intensive Operating Period: Comparison of aerosol scattering during coordinated flights, J. Geophys. Res., 111, D05S09, doi:10.1029/2005JD006250, 2006.

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., E., Dunlea, J., Huffman, J. A., Onasch, T. B., Alfarra, M. R., Williams, P. I., 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., and Worsnop, D. R.: Evolution of Organic Aerosols in the Atmosphere, Science, 326, 1525–1529, doi:10.1126/science.1180353, 2009.

Kirchstetter, T. W., Harley, R. A., Kreisberg, N. M., Stolzenburg, M. R., and Hering, S. V.: On-road measurement of fine particle and nitrogen oxide emissions from light- and heavy-duty motor vehicles, Atmospheric Environment, 33, 2955–2968, 1999.

Lanz, V. A., Alfarra, M. R., Baltensperger, U., Buchmann, B., Hueglin, C., and Prevot, A. S. H.: Source apportionment of submicron organic aerosols at an urban site by factor analytical modelling of aerosol mass spectra, Atmospheric Chemistry and Physics, 7, 1503–1522, 2007.

Matthew, B. M., Middlebrook, A. M., and Onasch, T. B.: Collection efficiencies in an Aerodyne Aerosol Mass Spectrometer as a function of particle phase for laboratory generated aerosols, Aerosol Science and Technology, 42, 884–898, 2008.

McFiggans, G., Alfarra, M. R., Allan, J., Bower, K., Coe, H., Cubison, M., Topping, D., Williams, P., Decesari, S., Facchini, C., and Fuzzi, S.: Simplification of the representation of the organic component of atmospheric particulates, Faraday Discussions, 130, 341–362, 2005.

Moore, K. G., I., Clarke, A. D., Kapustin, V. N., McNaughton, C., Anderson, B. E., Winstead, E. L., Weber, R., Ma, Y., Lee, Y. N., Talbot, R., Dibb, J., Anderson, T., Doherty, S., Covert, D., and Rogers, D.: A comparison of similar aerosol measurements made on the NASA P3-B, DC-8, and NSF C-130 aircraft during TRACE-P and ACE-Asia, J. Geophys. Res., 109, D15S15, doi:10.1029/2003JD003543, 2004.

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 mass spectrometric data, Atmospheric Chemistry and Physics, 9, 2891–2918, 2009.

Zhang, Q., Alfarra, M. R., Worsnop, D. R., Allan, J. D., Coe, H., Canagaratna, M. R., and Jimenez, J. L.: Deconvolution and quantification of hydrocarbon-like and oxygenated organic aerosols based on aerosol mass spectrometry, Environmental Science & Technology, 39, 4938–4952, doi:10.1029/2007GL029979,, 2005a.

Zhang, Q., Worsnop, D. R., Canagaratna, M. R., and Jimenez, J. L.: Hydrocarbon-like and oxygenated organic aerosols in Pittsburgh: insights into sources and processes of organic aerosols, Atmospheric Chemistry and Physics, 5, 3289–3311, 2005b.

Table S1: Summary of POA:NO_x and POA:CO emission ratios used in this study. Emission ratios are given in μg sm^{-3} ppm^{-1}.

Study	Location	POA:NO_x	POA:CO
Allan et al. (2010)	London, UK	31.6	N/A
Allan et al. (2010)	Manchester, UK	N/A	20.5
de Gouw et al. (2005)	Northeastern USA	N/A	9.4
Kirchstetter et al. (1999)	Tunnel study, USA	11.0	N/A
Lanz et al. (2007)	Zurich, Switzerland	15.9	20.4
Zhang et al. (2005b)	Pittsburgh, USA	N/A	4.3

Table S2: Summary of the diagnostics relating to the bootstrapping analysis from the dataset for each flight.

Flight	SD_{ts}/TS (%)		Max (SD_{ms}) (%)	
	OOA-1	OOA-2	OOA-1	OOA-2
B357	2.0	2.5	0.77	0.85
B362	5.6	9.7	1.00	2.44
B365	2.8	3.0	0.62	1.31
B366	13.4	11.8	1.66	1.46
B369	19.1	45.9	2.23	1.66
B370	2.9	3.6	0.52	0.46
B371	4.8	8.0	0.41	1.82
B373	2.3	3.9	1.06	0.86
B374	1.1	1.7	0.44	0.68
B379	6.6	9.3	0.61	1.63
B380	1.4	2.4	0.36	0.37
B401	11.1	13.3	0.24	0.32
B402	5.5	14.9	0.76	1.23
B406	1.4	1.8	1.06	0.79

Table S3: Summary statistics regarding the AMS chemical composition for each zone referred to the main paper. Concentrations are reported in μg sm^{-3}t at the 25^{th}, 50^{th} and 75^{th} percentiles.

Species	Statistic	Zones						
		1	2	3	4	5	6	7
Organics	25^{th}	3.69	0.57	3.21	3.49	2.65	2.84	0.66
	50^{th}	4.13	0.91	3.69	4.13	3.63	3.68	1.40
	75^{th}	4.55	1.29	4.40	4.83	4.96	4.39	3.40
Nitrate	25^{th}	3.04	0.58	1.15	2.10	0.24	0.33	0.05
	50^{th}	3.51	2.11	2.92	4.84	1.24	1.00	0.14
	75^{th}	4.40	3.02	4.98	6.85	2.72	3.16	1.45
Sulphate	25^{th}	2.78	0.83	2.88	1.99	0.98	1.13	0.49
	50^{th}	3.25	1.20	3.82	3.02	1.93	1.44	0.92
	75^{th}	3.65	1.78	4.89	3.65	4.12	1.93	2.12
Ammonium	25^{th}	2.60	0.55	1.48	1.67	0.51	0.61	0.17
	50^{th}	2.87	1.14	2.06	2.25	1.11	1.23	0.45
	75^{th}	3.06	1.47	3.06	3.36	2.03	1.77	1.46

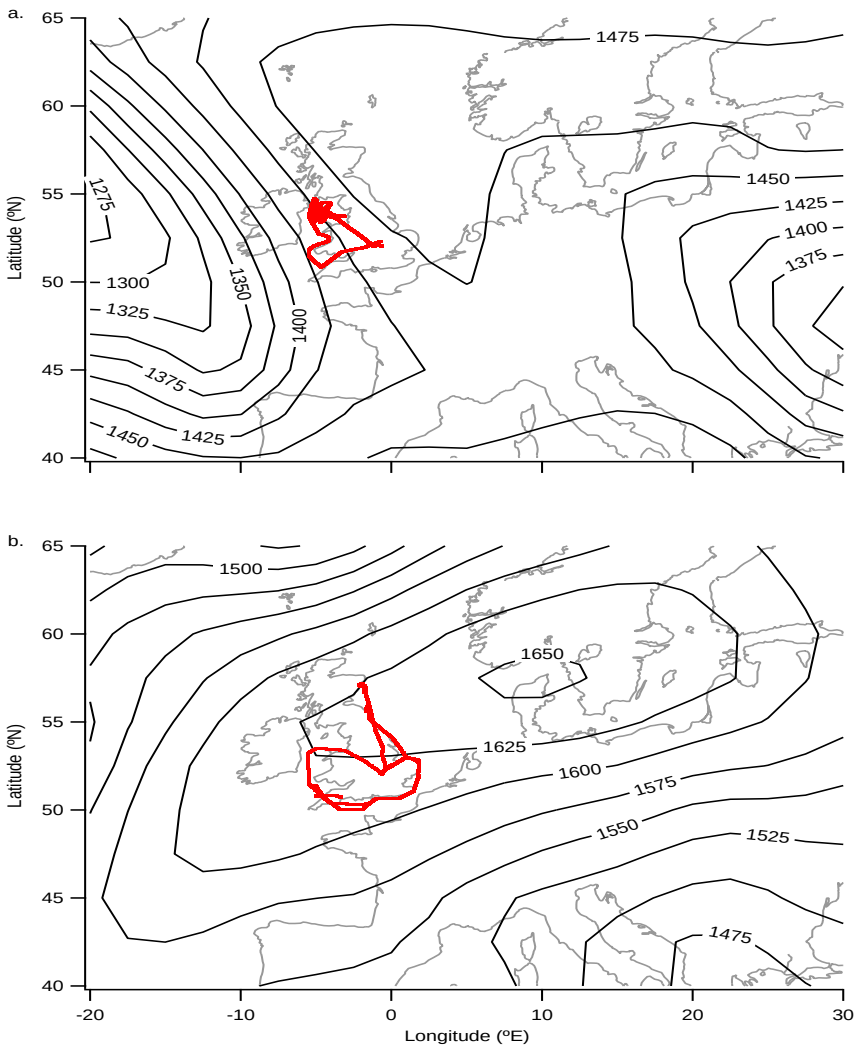

Figure S1: Flight tracks of the BAe-146 considered by this analysis for the ADIENT periods. Also shown are ECMWF 850 hPa geopotential height fields for each period considered by the analysis, where the field is either pertinent to a particular flight or is representative of the overall meteorological regime of the period. All geopotential height fields are for 12UTC. (a) summarises the flights for the UK-based ADIENT flying in April 2008 and the geopotential height field is from 16 April 2008 (B357). (b) summarises the flights for the UK-based ADIENT flying in September 2008 and the geopotential height field is from 25 September 2008 (B406).

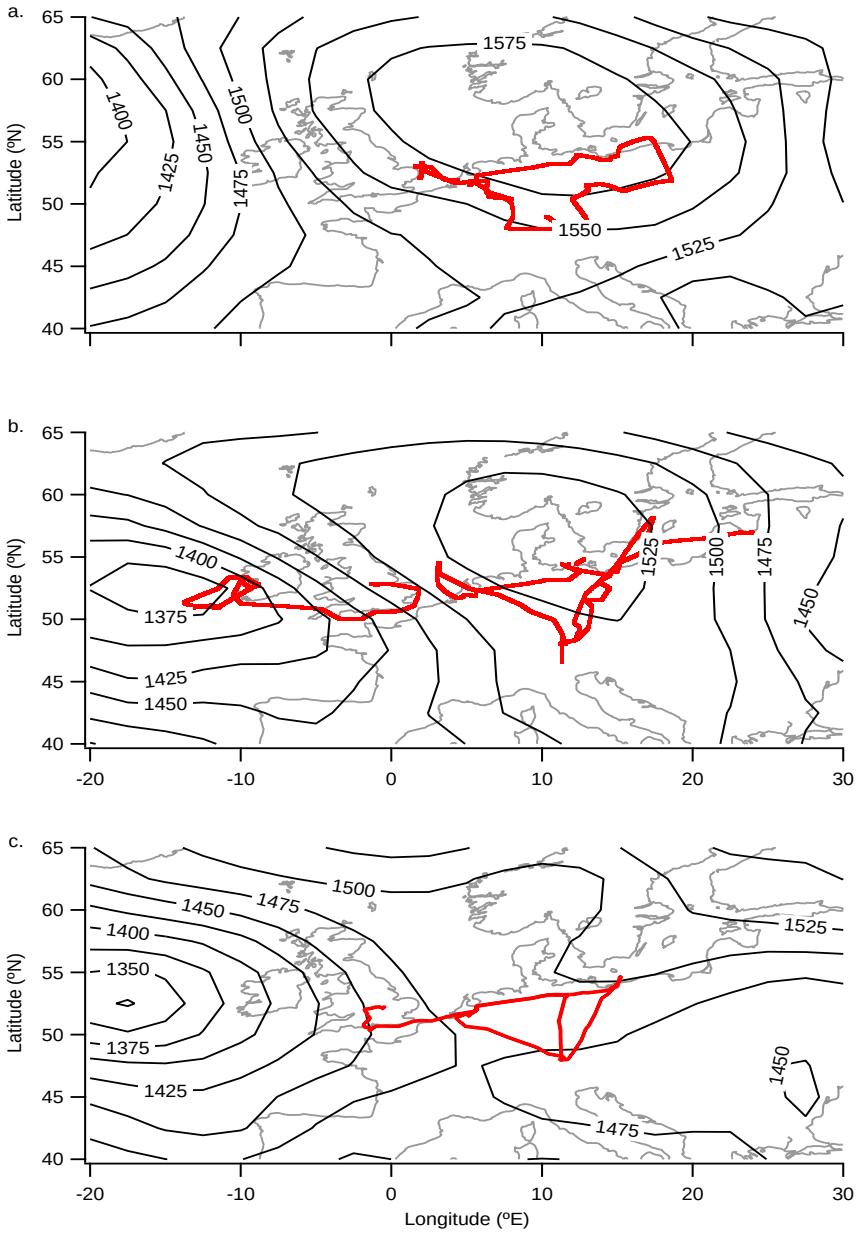

Figure S2: Same plots as Fig. S1 but now for the LONGREX flying period. (a) summarises the flights for the LONGREX-1 period with a geopotential height field for the 06 May 2008. (b) summarises the flights for the LONGREX-2 period with a geopotential height field for the 14 May 2008. (c) summarises the flights for the LONGREX-3 period with a geopotential height field for the 22 May 2008.

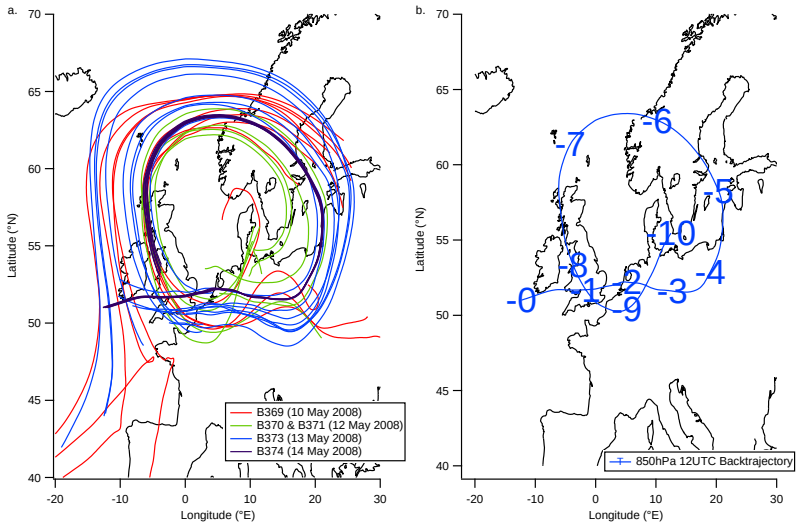

Figure S3: (a) Back trajectories at 850 hPa initialised from 1200 UTC on each flight day from the LONGREX-2 period. The initialisation points correspond to several SLRs during each flight. (b) Air mass back trajectory initialised from 1200 UTC on 14 May 2008 at 850 hPa. The numbered points relate to the number of days passed since the air mass was in that location.

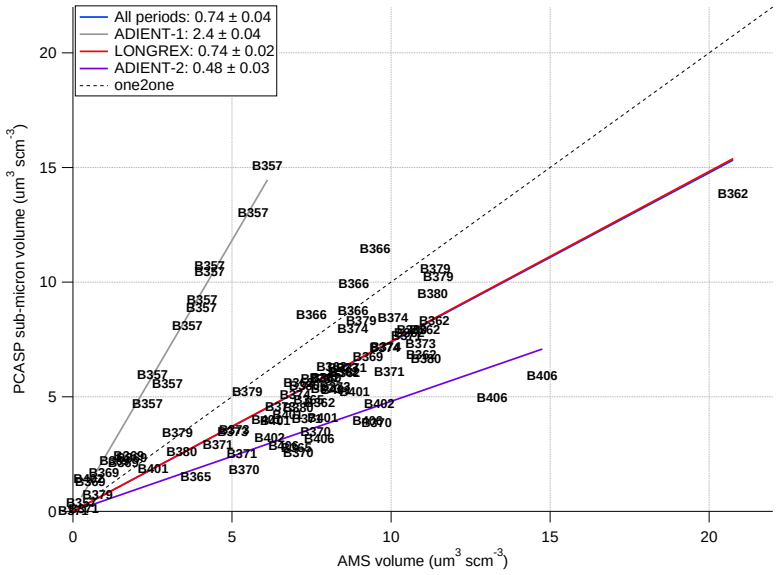

Figure S4: Comparison of inferred volume from the AMS with estimated sub-micron volume derived from the PCASP. The markers refer to SLRs below 3000 m i.e. within the boundary layer. The text markers refer to the flight which the point is from. Linear regression lines are shown for both individual flying periods and the study as a whole.

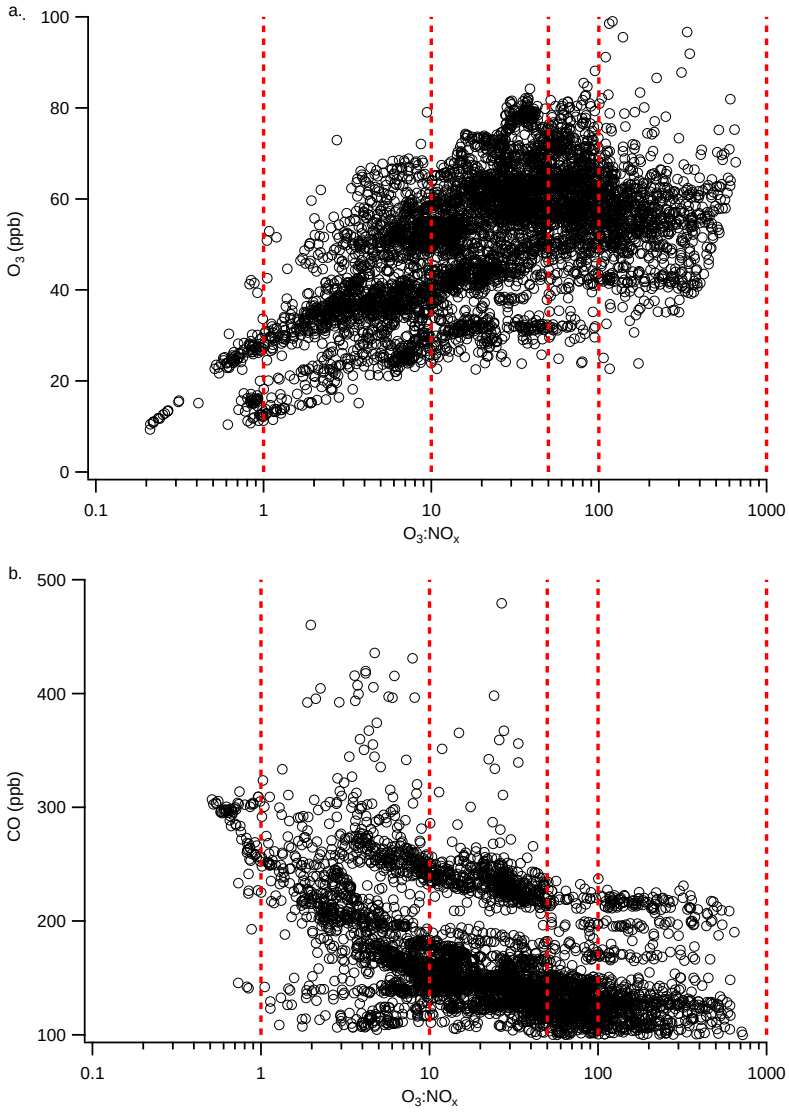

Figure S5: (a) Relationship between O_3 and O_3:NO_x for all flights. (b) Relationship between CO and O_3:NO_x for all flights expect for ADIENT-2. The red dashed lines refer to the distance from source boundaries discussed in the main paper.

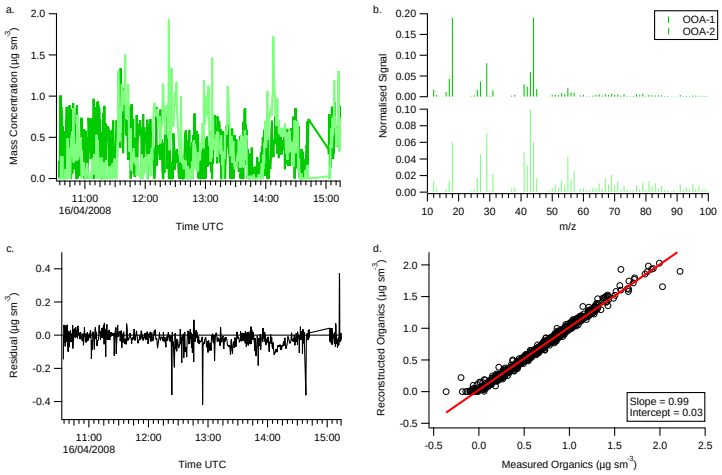

Figure S6: PMF solution for the two factor case from flight B357 including factor component time series (a) and mass spectra (b). The absolute residual is also shown in (c), whilst a comparison between the factor analysis reconstructed mass and measured organic signal is displayed in (d).

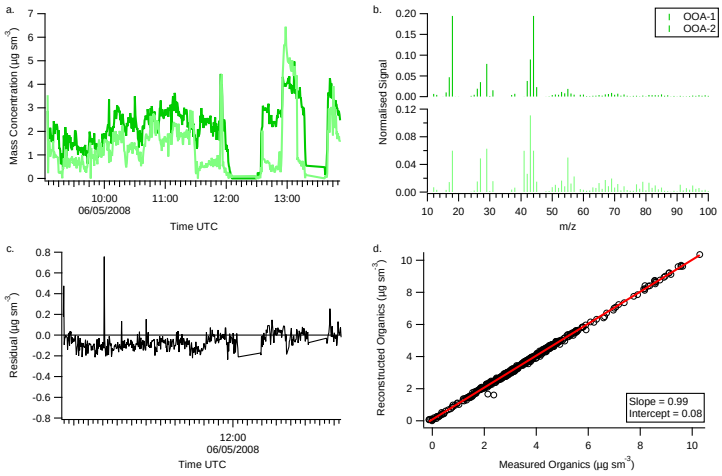

Figure S7: PMF solution for the two factor case from flight B362 including factor component time series (a) and mass spectra (b). The absolute residual is also shown in (c), whilst a comparison between the factor analysis reconstructed mass and measured organic signal is displayed in (d).

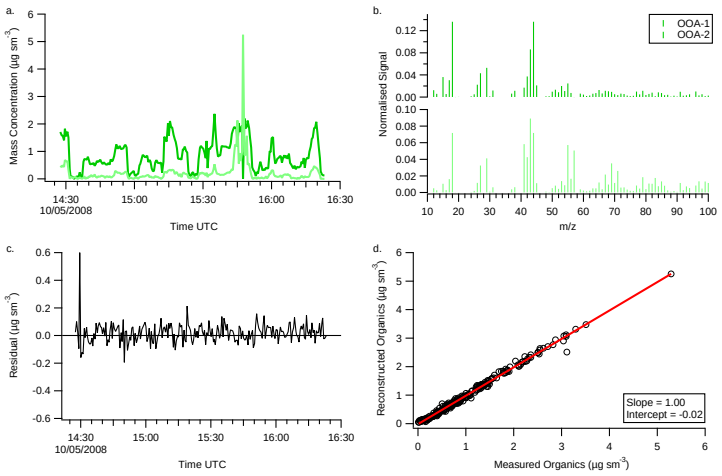

Figure S8: PMF solution for the two factor case from flight B369 including factor component time series (a) and mass spectra (b). The absolute residual is also shown in (c), whilst a comparison between the factor analysis reconstructed mass and measured organic signal is displayed in (d).

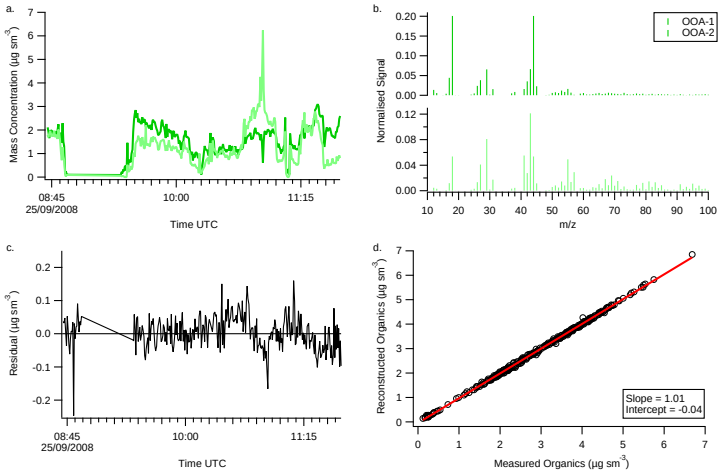

Figure S9: PMF solution for the two factor case from flight B406 including factor component time series (a) and mass spectra (b). The absolute residual is also shown in (c), whilst a comparison between the factor analysis reconstructed mass and measured organic signal is displayed in (d).

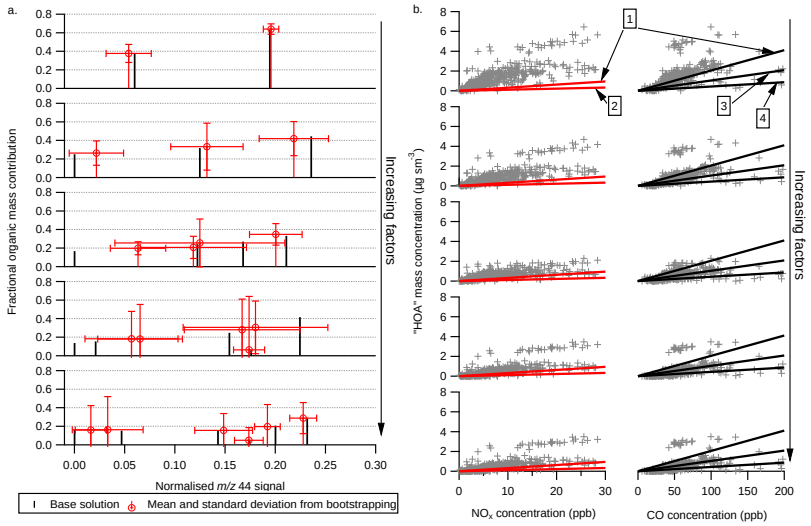

Figure S10: **(a)** Example from flight B362 of the relationship between the fractional mass contribution of a given factor to its normalised signal at *m/z* 44 for PMF solutions from two through seven factors. The black vertical bars refer to the base solution, while the red vertical and horizontal bars are the results from a resampling analysis known as bootstrapping. Increased standard deviations and mismatching between the base and bootstrapping solutions suggest a numerical unstable solution. **(b)** Relationship between the most HOA like factor profile with NO_x (red line) and CO (black line) for the factor solutions in (a). Solid red and black lines refer to literature emission ratios where (1) is from Allan et al. (2010), (2) is from Kirchstetter et al. (1999), (3) is from Lanz et al. (2007) and (4) is from Zhang et al. (2005b).

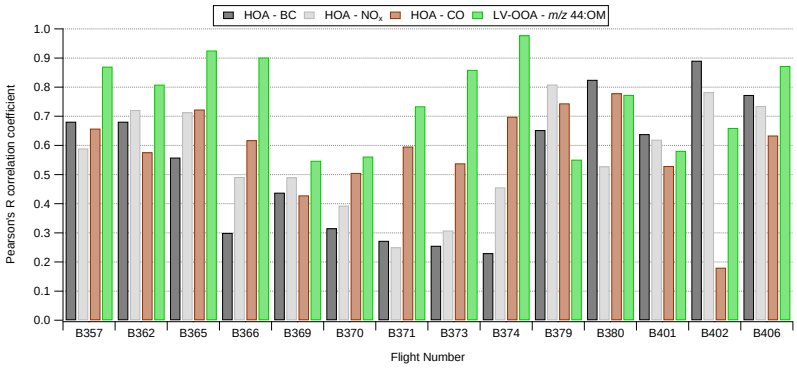

Figure S11: Summary of correlations for estimated HOA with Black Carbon (BC), NO_x and CO. Also shown are the correlations for LV-OOA organic mass fractions with the normalised organic signal at m/z 44.

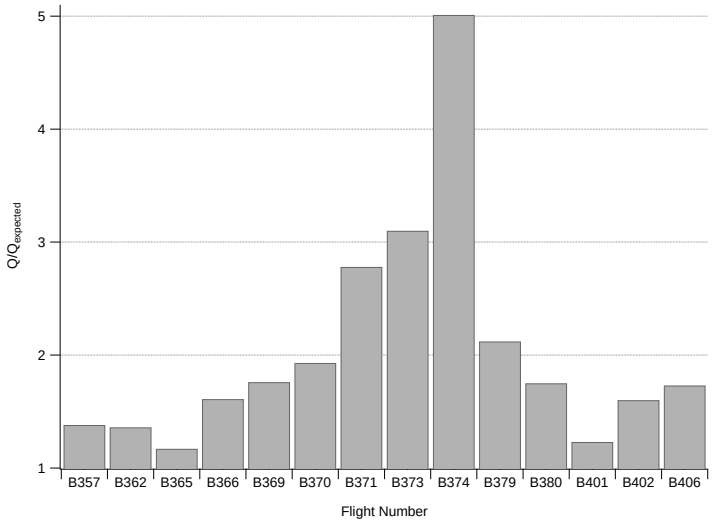

Figure S12: Summary of the $Q/Q_{expected}$ parameter for each flight in the dataset. All values are for the two-factor solutions with an fPeak of zero.

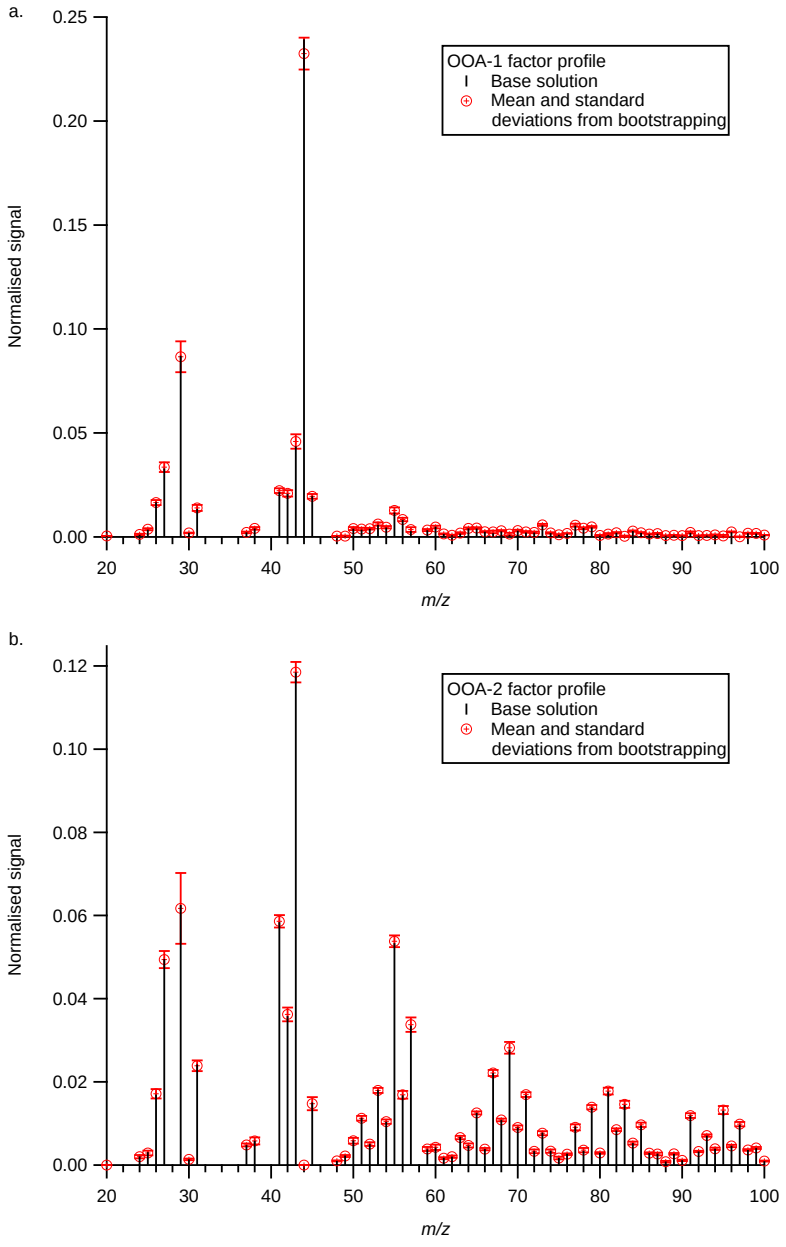

Figure S13: Results of the bootstrapping analysis for the two factor solution mass spectra for flight B357. (a) displays the mass spectrum for OOA-1, while (b) displays the mass spectrum for OOA-2.

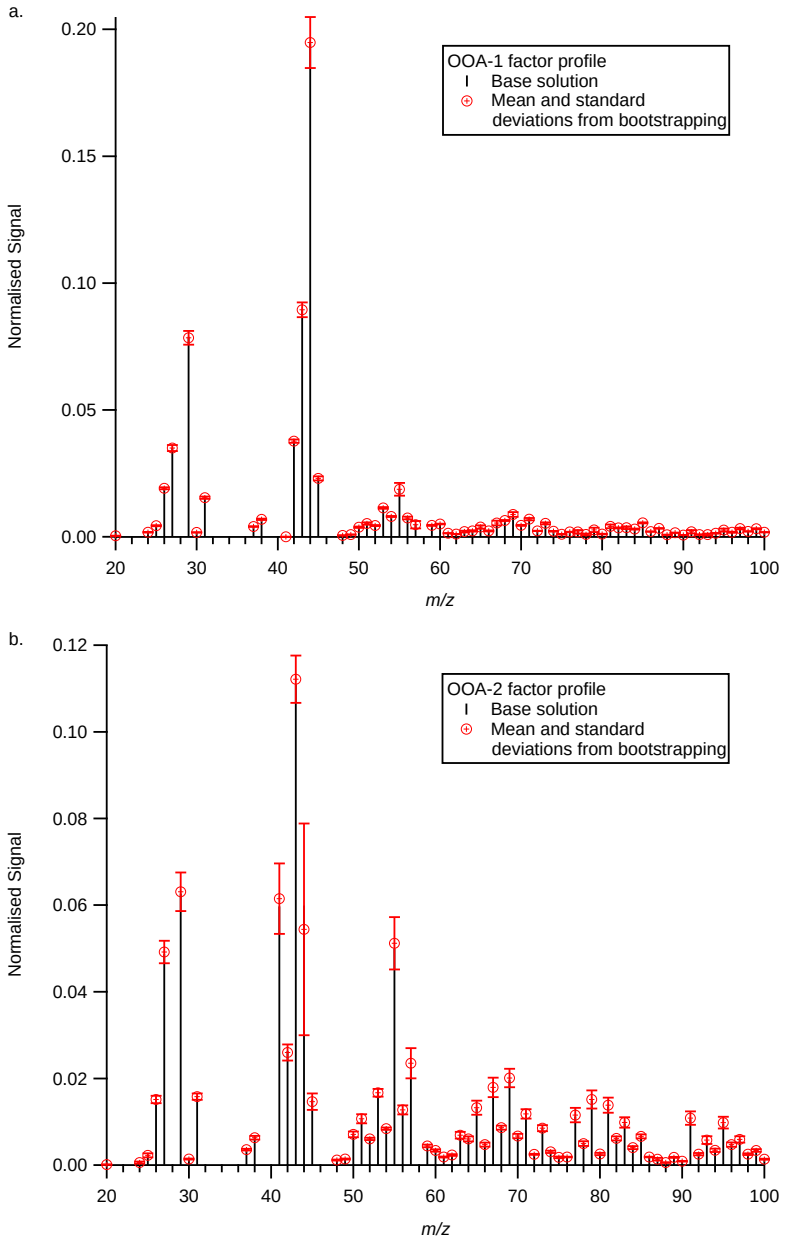

Figure S14: Results of the bootstrapping analysis for the two factor solution mass spectra for flight B362. (a) displays the mass spectrum for OOA-1, while (b) displays the mass spectrum for OOA-2.

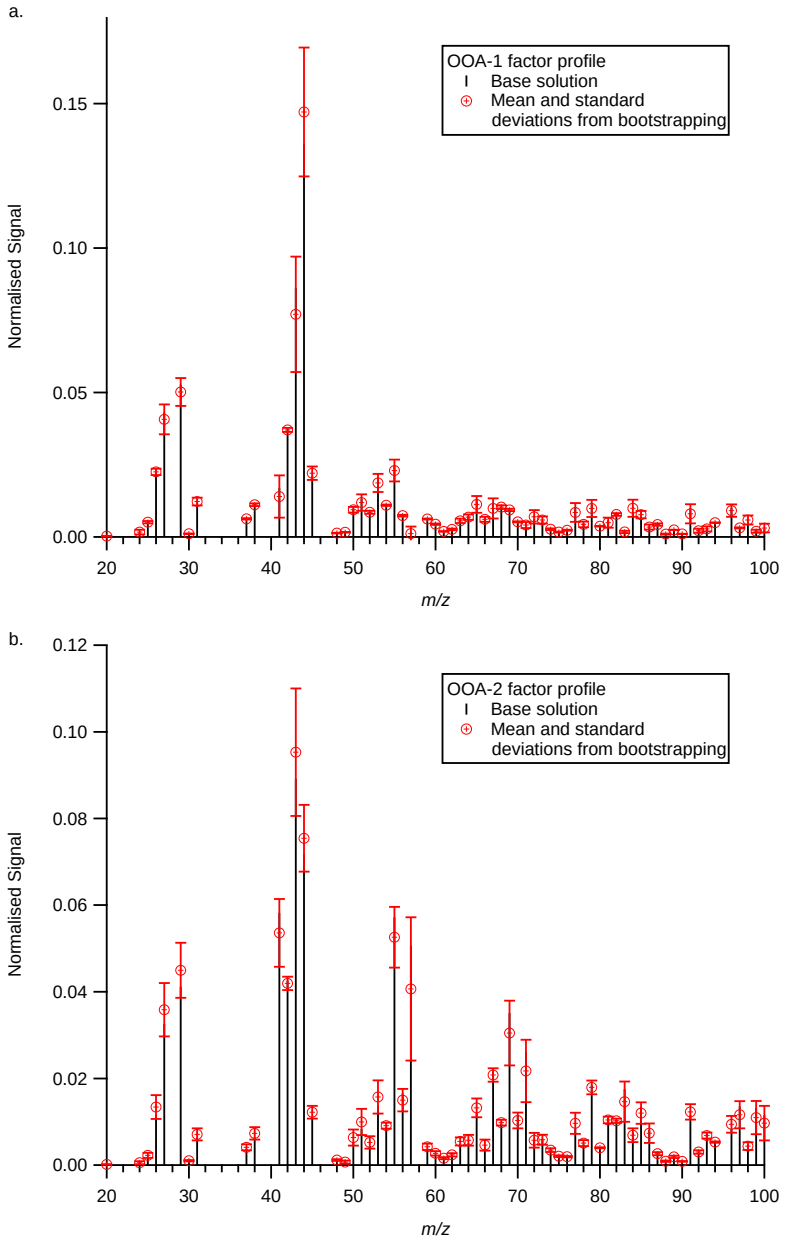

Figure S15: Results of the bootstrapping analysis for the two factor solution mass spectra for flight B369. (a) displays the mass spectrum for OOA-1, while (b) displays the mass spectrum for OOA-2.

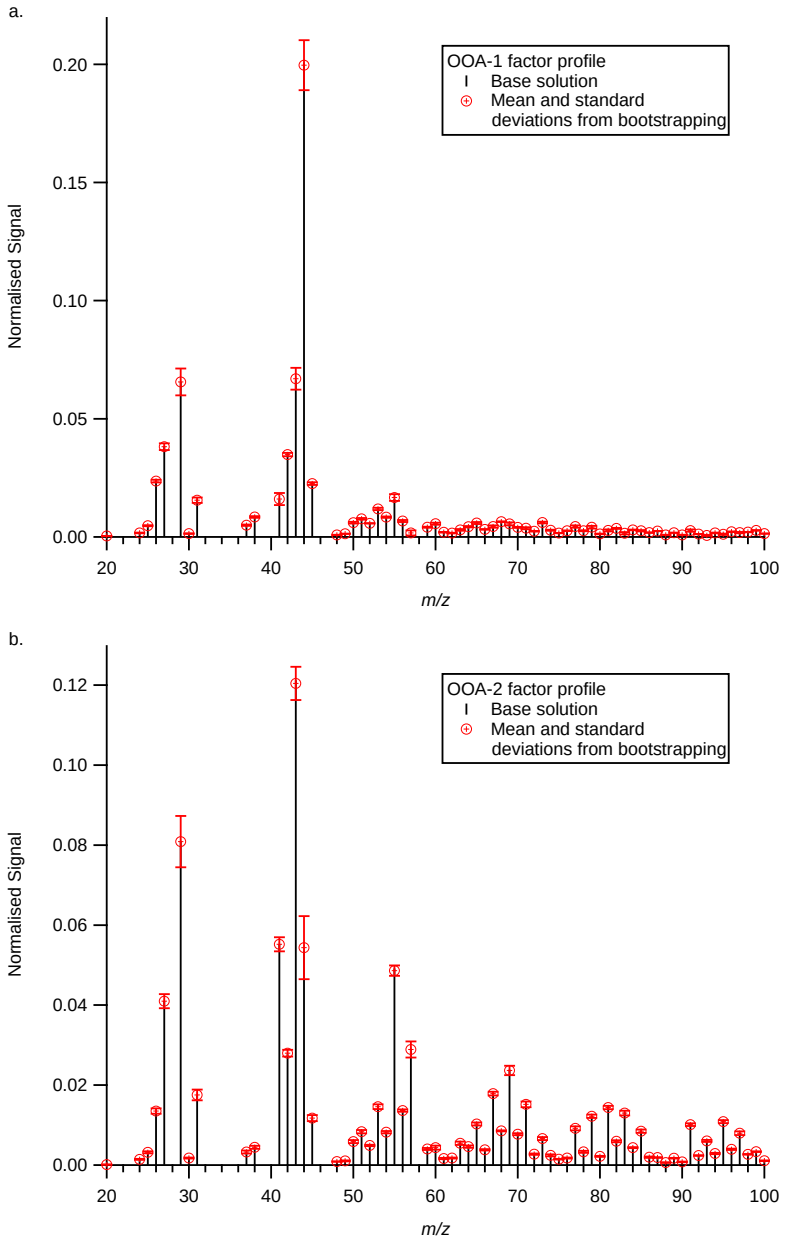

Figure S16: Results of the bootstrapping analysis for the two factor solution mass spectra for flight B406. (a) displays the mass spectrum for OOA-1, while (b) displays the mass spectrum for OOA-2.