Evidence for a Significant Proportion of Secondary Organic Aerosol from Isoprene Above a Maritime Tropical Forest: Supplementary Material

Robinson, N. H., et al. Atmos. Chem. Phys. Discuss.

Scope

This purpose of this document is to describe the methods used to generate the PMF factor outputs used in the main article. This includes how the data was handled, the outputs generated and the methods used to choose the most appropriate solution set. No attempt is made to interpret chemically the outputs of the analysis beyond assigning names for the factors based on marker peaks or comparisons with previously published material. The analysis was handled in a similar manner to Allan et al. (2010), in that solutions (according to number of factors, fpeak value and starting seed) were assessed according to uniqueness, numerical stability and whether the derived factors had realistic mass spectral profiles.

Data pretreatment and PMF operation

A matrix of organic mass spectral data was generated using the 'fragmentation table' technique (Allan et al., 2004) and errors generated using the error model described by Allan et al. (2003), with an additional arbitrary 'electronic noise' parameter included (equivalent to a standard error of 32 digitiser bits per data capture), which served to prevent unrealistically small error estimates being generated for low-signal channels. In order to prevent excessive weighting of the m/z=44 variable, the signals of m/z<20 were removed (these were added back in before further analysis). In addition, the signals of m/z>120 were found to hamper convergence. Given their low signal-to-noise ratios (<3, reducing to <1 after around m/z=250) and the fact they contributed little to the derived spectra, they were deemed 'weak' and removed (Paatero and Hopke, 2003).

The PMF Evaluation Toolkit (PET), as introduced by Ulbrich et al. (2009), was used to process the data, which acted as a front-end for the PMF2 version 4.2 executable (Paatero, 1997). The default conversion criteria were used. An initial run was performed and 'bad' runs were removed from the time series, based on the overall Q/Q_{exp} for the row being greater than 20. The rows removed were mainly in the form of spikes in the time series, either from local contamination or logging software malfunction. The analysis was then rerun with the filtered data.

Choice of number of factors

The analysis was performed with various numbers of factors and it was found that 5 or more yielded unsatisfactory results. This is illustrated by the 5-factor (fpeak=0) case below:
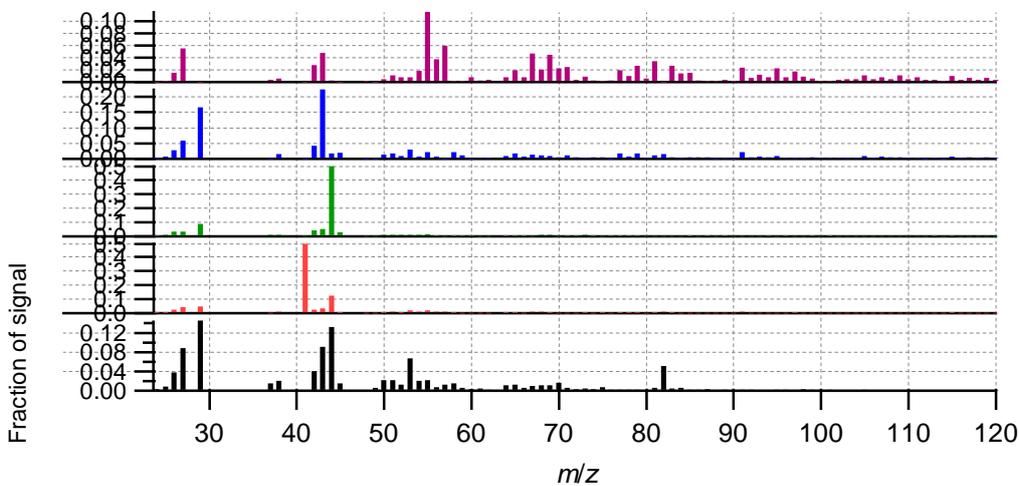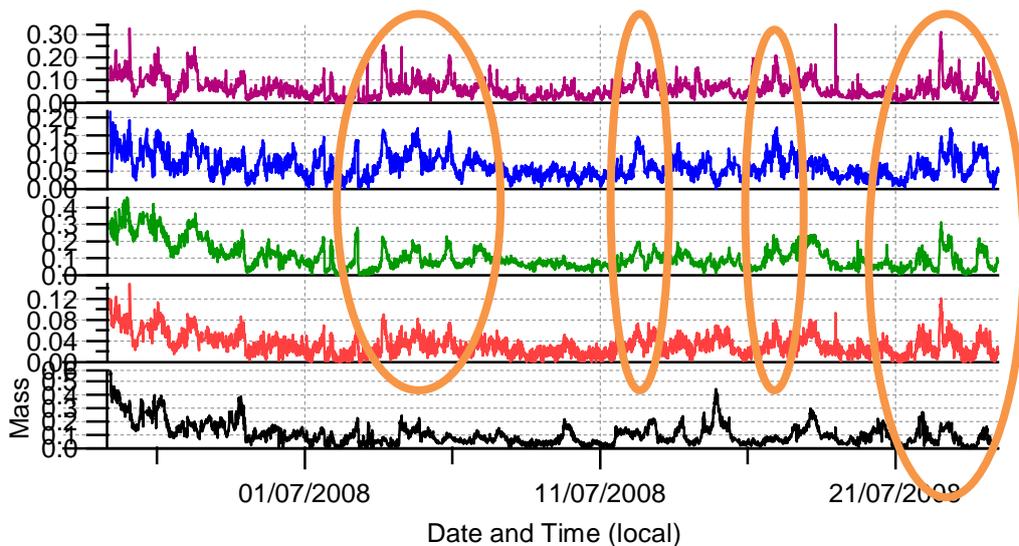
This solution was deemed unacceptable for two reasons; firstly, there were many common features, in particular in the time series of the top four time series (as plotted). Certain features are circled for clarity, but there were few features in the respective time series that were unique to a single factors. This was taken as indicative of 'mixing' taking place (Ulbrich et al., 2009). Secondly, a number of the mass spectral profiles (specifically the middle three) consisted of one or two peaks with little other signal at other *m/z* channels. Given that AMS mass spectra (in common with other EI mass spectrometers) typically consist of numerous fragment ions, it was thought that these were unlikely to be physically meaningful. In principle, it may have been possible to eliminate these behaviours by adjusting the pfeak parameter, however it was found that different values caused the algorithm to fail to converge.

Choice of pfeak

Within the 4 factor solution set, the fpeak=0 solution was found to have similar mixing and physicality issues as the 5 factor solution:
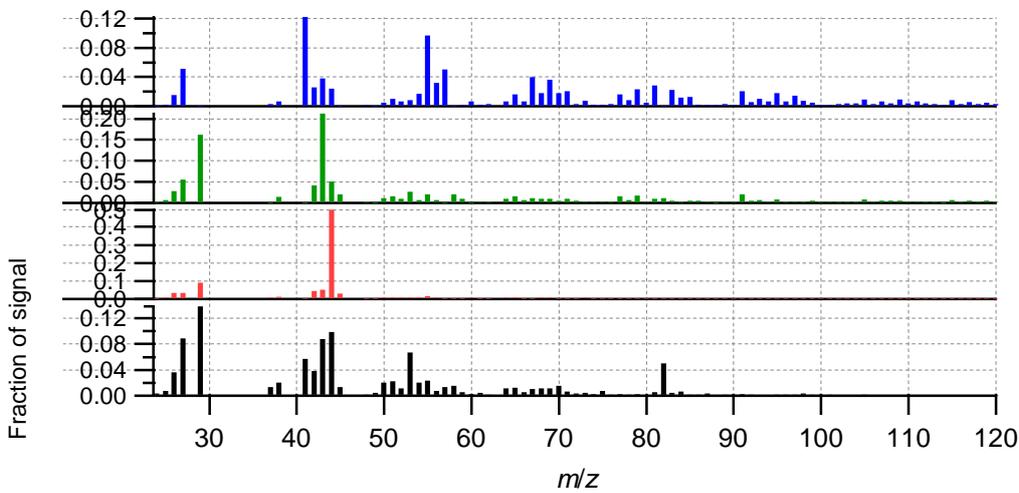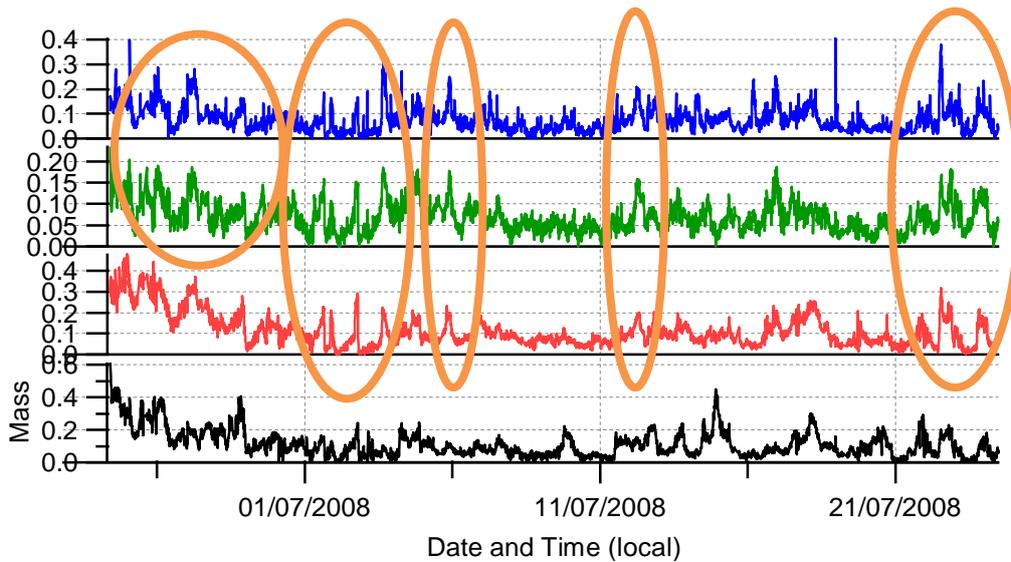
Note that there are many common features in the top three factors' time series and that the middle two profiles feature similar one- or two-peak spectra as the 5 factor solution. However, with the 4 factor case, other values of fpeak down to -1 were found to converge and the fpeak=-1 was deemed to be much more acceptable for a modest overall increase in Q/Q_{exp} (3.30 instead of 3.29):
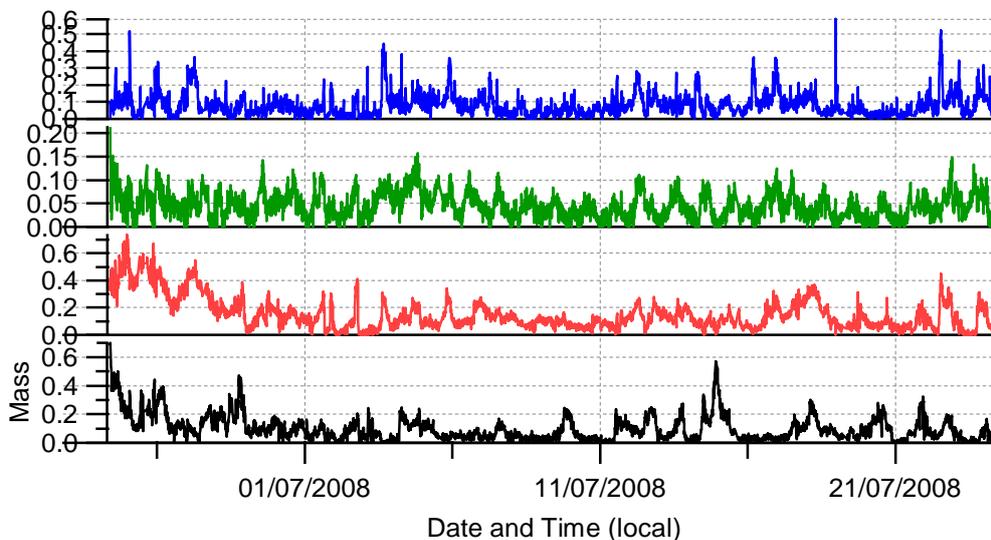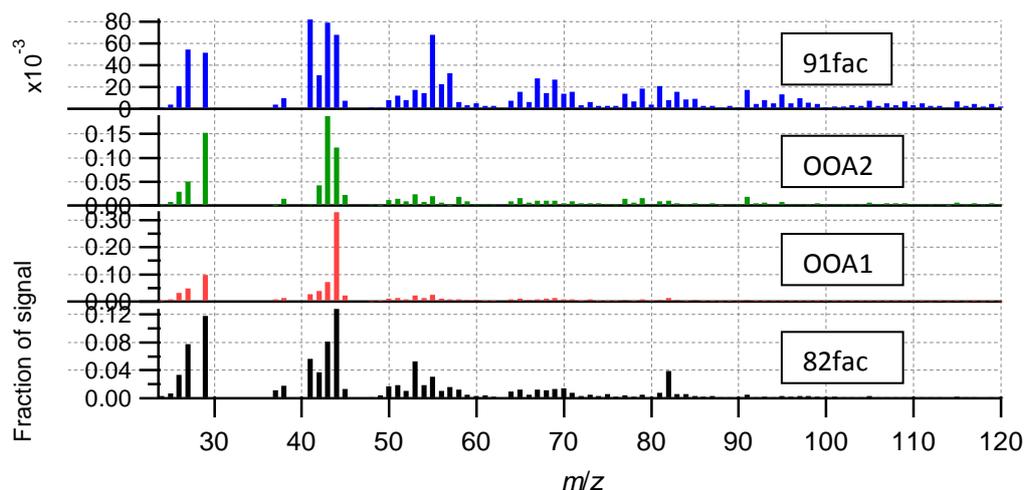
Note that the time series show many more unique features. Additionally, the mass spectra bear more of a resemblance to those previously published. The middle two spectra have been labelled 'OOA1' and 'OOA2' based on their similarity to previously published spectra (Lanz et al., 2007; Ulbrich et al., 2009). These have been linked to low volatility (LV) and semivolatile (SV) oxygenated organic aerosol (Jimenez et al., 2009), but without a direct measure of volatility, it is difficult to draw this conclusion with this data. The top factor bears a resemblance in some of its profile to biomass burning aerosols (Alfarra et al., 2007; Allan et al., 2010), although it does not possess the strong signals at m/z 60 and 73 often associated with levoglucosan, which could be expected (Jordan et al., 2006). It is possible that this particular chemical marker is not present in sufficient quantities, due to burn conditions or plume age, but it is difficult to tell with this data taken in isolation. Therefore, this factor is simply labelled as '91fac' for now. The remaining factor, is labelled '82fac', based on its most prominent marker peak. This is not a feature that has been widely reported in previous AMS datasets.

Stability of the '82fac' factor

The solution above (4 factors, peak=-1) was deemed to be the most satisfactory, however it was generally found that the 82fac factor, which is the subject of the main article, was remarkably stable across different solutions. For instance, while it represented 26% of the mass in the chosen solution (before correction for the missing mass below m/z 20), this figure only changed to 30% for the corresponding fpeak=0 solution and 28% for the 5 factor solution. When the numerical stability was checked by varying the starting seed, it was found that there was negligible differences between the solutions that converged. In all of the cases so far described, the mass spectrum was largely unaffected with key peaks at m/z 43, 44, 53 and 82. Also note that the peak at m/z=82 was mostly absent from the other factors. This shows that not only is the 82fac factor very stable and unambiguous, it has a very strong association with the 82 peak in isolation. While it may be possible to generate different solutions through further modification of the data treatment or algorithm function, we believe that any changes to this factor would be relatively small. Note that this does not apply to any other factors derived.

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, Environ. Sci. Technol., 41, 5770-5777, Doi 10.1021/Es062289b, 2007.

Allan, J. D., Jimenez, J. L., Williams, P. I., Alfarra, M. R., Bower, K. N., Jayne, J. T., Coe, H., and Worsnop, D. R.: Quantitative sampling using an aerodyne aerosol mass spectrometer - 1. Techniques of data interpretation and error analysis, J. Geophys. Res.-Atmos., 108, 4090, doi:10.1029/2002JD002358, 2003.

Allan, J. D., Coe, H., Bower, K. N., Alfarra, M. R., Delia, A. E., Jimenez, J. L., Middlebrook, A. M., Drewnick, F., Onasch, T. B., Canagaratna, M. R., Jayne, J. T., and Worsnop, D. R.: A generalised method for the extraction of chemically resolved mass spectra from aerodyne aerosol mass spectrometer data, J. Aerosol. Sci., 35, 909-922, 2004.

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, Atmos. Chem. Phys., 10, 647-668, 2010.

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, 10.1126/science.1180353, 2009.

Jordan, T. B., Seen, A. J., and Jacobsen, G. E.: Levoglucosan as an atmospheric tracer for woodsmoke, Atmos. Environ., 40, 5316-5321, DOI 10.1016/j.atmosenv.2006.03.023, 2006.

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, Atmos. Chem. Phys., 7, 1503-1522, 2007.

Paatero, P.: Least squares formulation of robust non-negative factor analysis, Chemometr Intell Lab, 37, 23-35, 1997.

Paatero, P., and Hopke, P. K.: Discarding or downweighting high-noise variables in factor analytic models, Anal Chim Acta, 490, 277-289, Doi 10.1016/S0003-2670(02)01643-4, 2003.

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, Atmos. Chem. Phys., 9, 2891-2918, 2009.