Geophysical Research Abstracts
Vol. 20, EGU2018-16758-1, 2018
EGU General Assembly 2018

Applying combined exploratory data reduction and classification methods to analysis of aerosol mass spectral data

Mikko Äijälä (1), Liine Heikkinen (1), Markku Kulmala (1), Douglas Worsnop (2), and Mikael Ehn (1)
(1) Department of Physics, University of Helsinki, Helsinki, Finland, (2) Aerodyne Research Inc., Billerca, U.S.A

Atmospheric aerosol is a complex mixture of carrier gas and particles, and comprises both complex organic compounds as well as inorganics. As aerosol chemical composition affects the interactions between aerosols and climate (by altering cloud condensation nucleus activation, hygroscopic properties, aerosol optics, volatility and condensation), as well as aerosol effects on human health (influencing toxicity, carcinogenicity, particle morphology), measurement and analysis of atmospheric fine particle composition are an essential part in understanding these important interactions.

Real-time aerosol composition measurements (using e.g. an Aerosol Mass Spectrometer AMS; Canagaratna et al., 2007) are today commonplace. Also the analysis tools required for processing large amounts of aerosol mass spectral data have developed rapidly in recent years. Data reduction methods such as exploratory and constrained factorisation (for example Positive Matrix Factorization; PMF; Paatero, 1999) are already widely used, and exploratory classification tools such as clustering seem to be a viable option for moving towards holistically machine learning oriented approaches for analysis of aerosol chemical data. (Äijälä et al., 2017).

In this study we analysed data from 8 AMS measurement campaigns of about one month each, measured at the forested background station in Hyytiälä, Juupajoki, Southern Finland. We combined both the inorganic and organic parts of the aerosol spectra, and applied a harmonized data processing and analysis methodology to all the data sets, to be able to summarise and combine the results. Our method was to perform rough, initial runs of exploratory PMF analysis with a fixed number of factors (10) to each of the sets, then automatically classify all the 80 factors obtained, and use the information to identify the most obvious commonalities between the data sets. In the second round we restrict the PMF algorithm, using constraints and variabilities based on the initial results, and again classify the resulting factors we obtain, hoping to distinguish the subtler underlying structures present in data. Finally, we interpret the results in light of aerosol chemical knowledge available.

In the initial analysis we identified some inorganic factors that were to be expected to be present in ambient aerosols, such as ammonium sulphate and ammonium nitrate, a range of organic aerosol factors, such as hydrocarbon-like organic aerosol, as well as some more surprising findings indicative of alkali metals Rubidium an Cesium from a local source, possibly from coal burning. The final round of analysis and interpretation is currently ongoing.

Canagaratna, M. et al. (2007). Mass Spectrom Rev., 26:185-222.
Paatero, P. (1999). J Comput Graph Stat, 8: 854-888.
Äijälä, M. et al. (2017). Atmos. Chem. Phys., 17, 3165–3197.