# Supplement to Aerosol composition and sources in the Central Arctic Ocean during ASCOS

R. Chang et al.

April 26, 2011

## S1 Solutions from positive matrix factorisation

For this study, a four factor solution was determined to best represent the measured aerosol. A two factor solution separates the Ship Emissions factor from the ambient background aerosol, while a three factor solution includes the Organic factor. It is only with four factors that the Marine Biogenic factor is separated from the Continental factor. Additional factors either identify instrumental noise, or split the existing factors. Figure S1 shows the decrease in $Q/Q_{exp}$ as additional factors are included, where $Q_{exp}$ is the expected $Q$. We see that including a $5^{\text{th}}$ factor decreases $Q/Q_{exp}$ by identifying instrumental noise, while additional factors only serve to capture episodic events, often coinciding with ship emissions. Even though $Q/Q_{exp}$ decreased slightly from 3.36 for four factors down to 3.06 for 10 factors, including more factors did not contribute additional information about the measured aerosol. As such, the four factor solution was deemed to give the most information about the measured ambient aerosol.

The robustness of the solution can be explored by either varying the initial seed, which changes the set of pseudorandom values used for the initial point (Paatero, 1997), or by using bootstrapping analysis, in which the rows of $\mathbf{X}$ are randomly sampled and PMF is executed on the new dataset (as described by Reff et al., 2007). Both of these methods were used and the four factor solution at *fPeak* = -0.75 was found to be robust: 100 values for the initial seed parameter in the PMF2 program resulted in 90 of the cases giving the solution presented here, while 100 iterations of the bootstrapping analysis

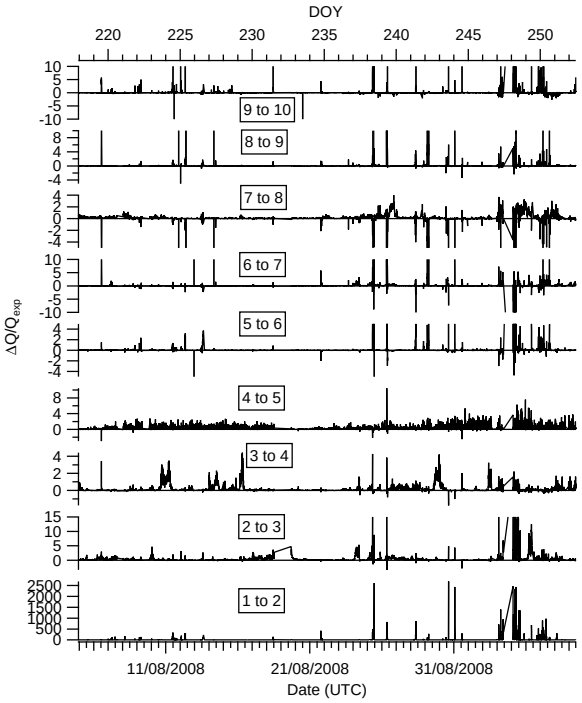
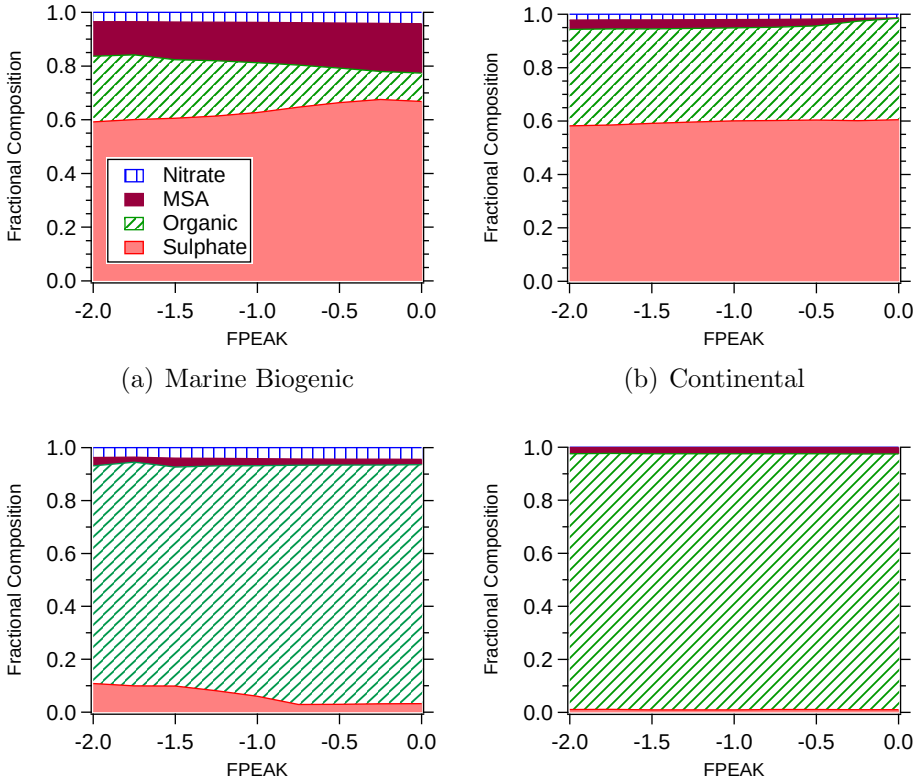



Figure S1: The decrease in $Q/Q_{exp}$ as additional factors are included in the PMF solution.

resulted in deviations of $< 0.01$ (fraction of signal) in the mass spectra and $< 0.015$ $\mu$g m$^{-3}$ in the time series.

Although the solution at *fPeak* = -0.75 is robust, a range of *fPeak* = -1.5 to 0 provide physically reasonable solutions. The degree to which the composition and F44 are dependent on the solution can be seen in Figs. S2 and S3.

Analyses were also performed on a data matrix calculated by adding together the mass spectra of the species of interest (i.e. nitrate, sulphate, organic and MSA) in nitrate equivalent mass with a corresponding error matrix calculated from the individual errors added in quadrature. However, results from the initial runs were similar enough to those calculated from the method described in the main text that only the latter method was pursued.




(a) Marine Biogenic (b) Continental

(c) Organic (d) Ship Emissions

Figure S2: Changes in fractional composition of the factors with varying *fPeak*.

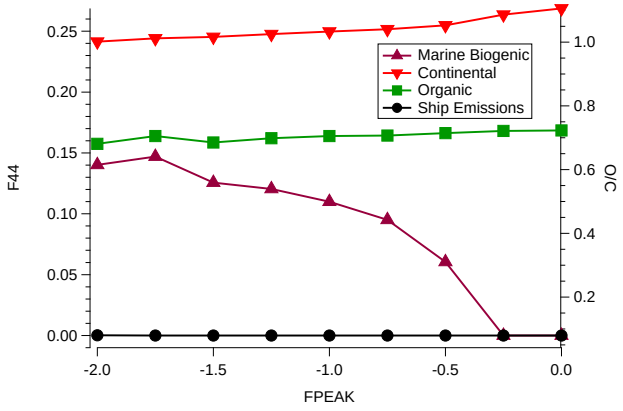



Figure S3: Changes in the F44 with varying *fPeak*.

# S2 Potentail source contribution function

To identify potential source areas of the observed aerosol chemical mass at the location of the ice breaker a receptor model called potential source contribution function (PSCF) was used (Ashbaugh et al., 1985). The PSCF model combines meteorological information with the AMS data to produce probability fields for potential source regions for the observed data.

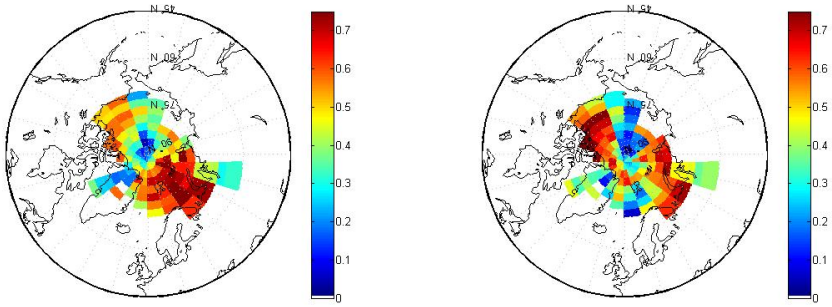
To reconstruct the air parcel movement, three dimensional back trajectory data were calculated from the re-analysis data library using the HYbrid Single-Particle Lagrangian Integrated Trajectory (HYSPLIT4) model (Draxler and Rolph, 2010; Rolph, 2010). The data originated from the National Weather Service's National Centers for Environmental Prediction's (NCEP) Global Data Assimilation System (GDAS).

Five-day back trajectories from 5 August to 8 September 2008 were recalculated hourly during the expedition arriving in the boundary layer, 100 m above *Oden*'s position. Air parcel movement was described by the coordinates of the endpoint of each hourly-segment of the trajectory. Then, to produce the PSCF analyses, each trajectory and its associated segment endpoints were associated with the measured 1h-median AMS aerosol or factor data.

In this study, the northern hemisphere was divided into $18^\circ \times 2.5^\circ$ grid cells. Trajectories with segment endpoints in cell $ij$ were counted as $n_{ij}$ and assumed to collect the chemical mass emitted from that cell. The probability that air from a particular grid had been transported along the trajectory to *Oden*'s position is then given by $P_{ij}(A) = \frac{n_{ij}}{N}$, where $N$ is the total number of trajectory segment endpoints. If the aerosol sample connected to the trajectory has a concentration higher than a selected criterion value (here the $50^{\text{th}}$ percentile of the measured aerosol mass concentration is used) it is set as high, and all segments of this trajectory are considered to be high. The probability that air from a particular grid, $ij$, has a high aerosol concentration, $m_{ij}$, is then given by $P_{ij}(B) = \frac{m_{ij}}{N}$. The conditional probability that the air passing through the $ij^{\text{th}}$ cell had a high aerosol concentration when arriving to *Oden* is then given by the ratio of these two probabilities $PSCF_{ij} = \frac{P_{ij}(B)}{P_{ij}(A)} = \frac{m_{ij}}{n_{ij}}$.

To avoid errors when the total number of segment endpoints in a cell is low, only cells with at least 10 segment endpoints or more are used, and the PSCF values are multiplied with a weighting function

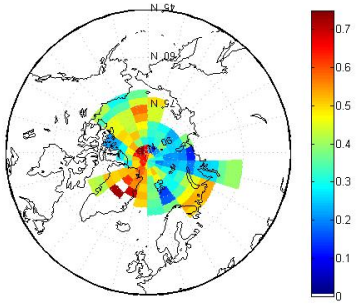
$$W(n_{ij}) = \begin{cases} 1.0 & \text{when } 50 \leq n_{ij} \\ 0.8 & \text{when } 10 \leq n_{ij} < 50 \\ 0 & \text{when } n_{ij} < 10. \end{cases}$$




(a) Marine Biogenic




(b) Continental




(c) Organic

Figure S4: PSCF analysis of the three ambient factors.

Similar approaches have been used in other PSCF studies (Hopke et al., 1995; Yli-Tuomi et al., 2003; Zhang et al., 2010). PSCF analysis on the three factors can be seen in Fig. S4 and are discussed in the main body of the text.

# References


Ashbaugh, L. L., Malm, W. C., and Sadeh, W. Z.: A residence time probability analysis of sulfur concentrations at grand Canyon National Park, Atmospheric Environment (1967), 19, 1263–1270, doi: DOI: 10.1016/0004-6981(85)90256-2, 1985.

Draxler, R. R. and Rolph, G. D.: HYSPLIT (HYbrid Single-Particle Lagrangian Integrated Trajectory) Model access via NOAA ARL READY Website (http://ready.arl.noaa.gov/HYSPLIT.php), access: 21 September 2009, 2010.

Hopke, P. K., Chong, C. L., W., C., and Landsberger, S.: The use of bootstrapping to estimate conditional probability fields for source locations of airborne pollutants, Chemometr. Intell. Lab., 30, 69–79, 1995.

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

Reff, A., Eberly, S. I., and Bhave, P. V.: Receptor Modeling of Ambient Particulate Matter Data Using Positive Matrix Factorization: Review of Existing Methods, J. Air Waste Manage., 57, 146–154, 2007.

Rolph, G. D.: Real-time Environmental Applications and Display sYstem (READY) Website (http://ready.arl.noaa.gov), access: 21 September 2010, 2010.

Yli-Tuomi, T., Hopke, P. K., Paatero, P., Basunia, M. S., Landsberger, S., Viisanen, Y., and Paatero, J.: Atmospheric aerosol over Finnish Arctic: source analysis by the multilinear engine and the potential source contribution function, Atmos. Environ., 37, 4381–4392, 2003.

Zhang, G., Li, J., Li, X., Xu, Y., Guo, L., Tang, J., Lee, C., Liu, X., and Chen, Y.: Impact of anthropogenic emissions and open biomass burning on regional carbonaceous aerosols in South China, Environ. Poll., 158, 3392–3400, 2010.