Organic aerosol factor analysis of long term ACSM data

Qi Zhang¹, Caroline Parworth¹, Jerome Fast², Tim Shippert², Chitra Sivaraman², Fan Mei³, and Alison Tilp³

¹University of California - Davis
²Pacific Northwest National Laboratory
³Brookhaven National Laboratory

Contact: dkwzhang@ucdavis.edu

Funding: DE-FG02-11ER65293

Aerosol Chemical Speciation Monitor (ACSM)

- Long term, continuous measurements
- Time resolution: 15 – 30 min
- D. L. (µg m⁻³) for 30 min averaging: org = 0.15, sulfate = 0.024, nitrate = 0.012, ammonium = 0.28, chloride = 0.011
- Organic MS: unit mass resolution, 10 – 150 amu
- 3 systems at DOE Atmospheric Radiation Measurement (ARM) sites
  - The Southern Great Plains (SGP), Oklahoma
  - Tropical Western Pacific in Darwin, Australia
  - MAOS mobile facility
Organic Factor Analysis of OA Spectra

$$\overrightarrow{m_{\text{measured}}} = \overrightarrow{c_a} \cdot \overrightarrow{m_a} + \overrightarrow{c_b} \cdot \overrightarrow{m_b} + \overrightarrow{c_c} \cdot \overrightarrow{m_c} + \ldots$$

Multivariate analysis methods:
- Positive Matrix Factorization (PMF)
- Multilinear Engine (ME-2)
- Tracer-based multilinear decomposition
- Spectra-based linear decomposition (CMB-style) ...

Products:
- Factors: HOA, OOA (SV-, LV-), BBOA, ...
- Time-resolved concentration time series of OA factors (OA$_i$):
  $$\sum \text{OA}_i \approx \text{Organics}$$
- Mass spectra of OA factors that bear some information of their chemical properties, e.g., f44 $\rightarrow$ O/C

OA factor data derived from AMS and ACSM field data are useful for model validation.
Ng, N. L. et al. (2011), An Aerosol Chemical Speciation Monitor (ACSM) for routine monitoring of the composition and mass concentrations of ambient aerosol, Aerosol Science and Technology, 45(7), 770-784, 10.1080/02786826.2011.560211

ACSM Data: Intercomparisons (New York City)

ACSM Data: Intercomparisons (Long Island, NY)

S. Zhou, Q. Zhang, M. Fan et al.
HR-TOF-AMS vs. PILS, Long Island, NY

(a) Chloride
(b) SO$_4$
(c) NH$_4$
(d) NO$_3$

AMS Species (µg/m$^3$)
PILS Species (µg/m$^3$)

S. Zhou, Q. Zhang, M. Fan et al.

ARM SGP Site

Oklahoma Gas and Electric Company
Coal-fired power station
ARM ACSM Data Processing Flow Chart

sampling → A1 Product
  raw data

mentor input → B1 Product
  OM, SO₄, NO₂, NH₄, Cl, QA, and other ancillary data needed for multivariate analysis

Data Management Facility
  run script every 2 weeks that performs multivariate analysis

plots

developer input
  user access A1, B1, and C1 products via archive

1 year SGP ACSM Data
SGP ACSM Data Sanity Check

SGP ACSM Data Sanity Check
Basic PMF Reminder

\[ X = GF + E \]
\[ x_{ij} = \sum_p g_{ip} f_{pj} + e_{ij} \]
\[ Q(E) = \sum_{i=1}^{m} \sum_{j=1}^{n} \left( \frac{e_{ij}}{\sigma_{ij}} \right)^2 \]

**Assumption:** the ACMS organic aerosol data matrix represents the linear combination of OA factors with constant profiles that have varying contributions across the dataset.

- One may choose # of factors (P) base on evaluating \( Q \) vs. \( P \)
- # of factors extractable determined by chemical and temporal resolution.

Ulbrich et al., 2009
Proposed Outline for PMF Analysis

• Obtain background information about site and surrounding areas
  – Emission sources
  – Wind and meteorological patterns

• Characterize existing data
  – Determine uncertainty associated with data
  – Average MS
  – Time series of species
  – Contributions of each species to entire mass loading

• Prep data for PMF analysis
  – Determine best pretreatment steps

• Perform PMF on prepared data

Error Pretreatment

• Goal 1: Compare removal of noisy m/z with down-weighing of m/z
  – Determined that down-weighing noisy m/z by factor of 10 is more preferable than deleting these m/z

• Goal 2: What S/N criteria should be used when down-weighing m/z
  – Eg: Down-weigh m/z that have S/N < 0.1 by factor of 10 vs down-weigh m/z with S/N < 0.2 by factor of 10
Error Pretreatment: Objective 1

- Removal of noisy m/z
- Down-weighing of noisy m/z

Criteria:
- \( S/N < 0.2 \) down-weighted by factor of 10 (bad m/z)
- \( 0.2 < S/N < 2 \) down-weighted by factor of 2 (weak m/z)
Evaluation and selection of PMF solutions


Data Pretreatment Steps

1. Apply minimum error of $5 \times 10^{-3}$ to $\text{ORG}_{\text{err}}$
2. Remove spikes in $\text{ORG}$ and $\text{ORG}_{\text{err}}$, spike threshold set by $v_{\text{spike\_thr}}$
3. Down-weigh m/z associated with m/z 44 by a factor of 2, includes m/z 16, 17, 18 and 44
4. Down-weigh m/z in ListofBadAmus (mean S/N < 0.2) by factor of 10
5. Down-weigh m/z in ListofWeakAmus (mean $0.2 < \text{S/N} < 2$) by factor of 2
6. Down-weigh m/z in ListofSmallError ($10^{th}$ percentile S/N <- 2) by factor of 5
7. Remove all NaNs in data and error matrices
Rolling Window Analysis

Calculation of $Q/Q_{exp}$ and Residuals

1. Use weigh "water" $m_1/2$
2. Use weigh "all" $m_1/6$
3. Use weigh $m_1/6$
4. Calculate mean extracted mass from PAF result, where $m_2$ represents the $m_7$ factor time series and $m_3$ represents the $m_8$ factor mass spectrum.
5. Calculate residual of each element of mean extracted matrix $m_4 = m_2 - m_3$.
6. Calculate scaled residual matrix with elements $m_5$.
7. Replace columns corresponding to "bad" $m_4$'s in the scaled residual matrix with blanks.
8. Replace columns corresponding to $m_4$'s associated with $m_4/6$ in the scaled residual matrix with blanks.
9. Total number of columns with blanks: $n_1$
10. Calculate squared scaled residual matrix with elements $m_6$.
11. Treat elements of squared scaled residual matrix at $m_6_{ij}$ by replacing with $0 / (0)^2$.
12. Calculate Q/C
   1. $Q = \sum_{i=1}^{n_1} \frac{m_6_{ij}}{m_5_{ij}}$
   2. $Q_{exp} = degrees\ of\ freedom = n - (n - \pi) - p\times(n - \pi)$, where columns with blanks are accounted for.