The Source Finder (SoFi)

Paul Scherrer Institute

Email to: francesco.canonaco@psi.ch
Source apportionment technique - PMF

bilinear PMF model
- rows of matrix F represent factor profiles
- columns of matrix G represent factor time series
- minimizing Q

advantages
- values in G & F are non-negative
- factors represent sources (POA) / aging (SOA)

disadvantages
- constant factor profiles (mass spectra) over PMF run
- assess number of factors
- assess statistical error (e.g. resampling strategy)
- assess amount of rotational ambiguity

\[ X_{\text{model}} = G \cdot F = G \cdot T \cdot T^{-1} \cdot F = G' \cdot F' \]

SoFi – The interface in Igor Wavemetrics Pro

Canonaco et al., 2013

SoFi 6.1 current version

F. Canonaco, PSI Switzerland

Email to: francesco.canonaco@psi.ch
General aspects

**Standard SoFi 6**

- stable SoFi version, no changes in architecture planned
- compatibility is guaranteed from 6.0 on
- freeware
- current SoFi version is not compatible with Igor 7! It will be made compatible next year
- panel-oriented / user-friendly interface
- fast (exploits multiprocessors for PMF calc. and during the data treatment in Igor)
- storage of PMF runs in HDF files avoids memory problems in Igor and allows to perform and treat many PMF runs (thousands to millions)
- rotational and statistical uncertainty can be quantified (e.g. a value approach / fpeak combined with the resampling strategy «bootstrap»)
number of repeats

still in progress, at least 500

Canonaco et al., in prep., soon in AMTD
General aspects

Add-ons to SoFi 6

- license-based system
- PMF runs (!huge amount!) can be sorted based on user-defined list of criteria (see later)
- rolling mechanism, important for e.g. ambient long-term SA (see later)

Commercial aspects

- spin-off will be opened up in coming months (currently F. Canonaco and C. Bozzetti)
- maintenance of the SoFi package and license system for the add-ons
- immediate troubleshooting service guaranteed by spin-off
- SA service offered
- offering online (i.e. while ACSM is measuring) SA service on Q-ACSM and ToF-ACSM in near future
- Eventually: offering off-line AMS analyses of filters including SA.

Email to: francesco.canonaco@psi.ch
Add-on to SoFi 6

Criteria-based approach

- various criteria, including their thresholds and weights can be easily defined in SoFi
- criterion, e.g. correlation coefficient, explained variation, peak in the diurnal cycle, distance between two factor profile variables, regression model., etc.
- thresholds backed up with e.g. literature data, based on significance test, etc.
- weights currently only for graphical purposes (different criteria distr. overlap)

Example: Zurich ACSM data 2011/2012

<table>
<thead>
<tr>
<th>factor</th>
<th>criterion</th>
<th>thresholds</th>
<th>weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>HOA</td>
<td>Diurnal cycle correlation between HOA and BC_{traffic}</td>
<td>lower limit: 0.62</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>higher limit: -</td>
<td></td>
</tr>
<tr>
<td>COA</td>
<td>lunch peak (11 – 13 am) compared to before and after (9+10+14+15 am)</td>
<td>lower limit: 1.2</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>higher limit: -</td>
<td></td>
</tr>
<tr>
<td>BBOA</td>
<td>explained variation of m/z 60</td>
<td>lower limit: 0.35</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>higher limit: -</td>
<td></td>
</tr>
<tr>
<td>LV-OOA</td>
<td>m/z 44 in the profile</td>
<td>lower limit: 0.10</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>higher limit: -</td>
<td></td>
</tr>
<tr>
<td>SV-OOA / LV-OOA</td>
<td>distance in f44 / f43 space between SV-OOA and LV-OOA</td>
<td>lower limit: 0.05</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>higher limit: -</td>
<td></td>
</tr>
</tbody>
</table>
Example: three criteria from the Zurich ACSM data 2011/2012

- **threshold** determines which PMF runs are considered
- **weight** guarantees that the criteria distr. fall ~in the same x-range

The overlap between criteria distr. forms the total distr. (these PMF runs are further analyzed in SoFi).

SoFi result analysis package
Criteria-based approach

- criteria can be either active (score relevant for the ranking of the PMF runs criteria discussed before) or passive (score not relevant for the ranking of the PMF runs)

Example: Zurich ACSM data 2011/2012

Selection of PMF results on the active criterion (HOA) shows the $R_{\text{Pearson}}$ values on two passive criteria.
Add-on to SoFi 6

Rolling mechanism

- PMF window can be automatically shifted over the PMF input matrix
- after every shift the PMF runs are reinitialized (seed, a value, fpeak, bootstrap, etc.)
- length of window and shift are user-defined parameters

Example: Zurich ACSM data 2011/2012
Update to standard SoFi 6

SoFi 6.2 (November)

- many small improvements in the graphs / bugs fixed
- problems during the iteration, occurring in some HR-AMS/elemental datasets solved
- calc. of average PMF result is optional (saves time/memory in Igor)
- customized colors for factors
- customized selection of OM:OC calc. (Canagaratna or Aiken) relevant for HR-AMS
- x-axis labelling for all variable plots can be enabled, especially important for elemental/compounds data/HR-AMS
- rapid inspection of raw / daily averaged factor time series
- PMF runs can be merged automatically in SoFi (manual «rolling PMF»)

-separate PMF runs

-important: rapid inspection of averaged factor profiles / t-dependent factor profiles
Update to standard SoFi 6

SoFi 6.2 (November)

- feature of selecting certain time ranges, hours and/or days available in all time plots, including correlation table/graph, pie charts (here only shown for a scatter plot)

Selecting only:
- weekdays
- rush hours
- period May/June 2011
Acknowledgments

ACSM project supported by the Swiss Federal Office for the Environment (FOEN)

Thank you all for the attention...