Supporting Information for: Real-time Source Apportionment of Organic Aerosols in Three European Cities

Gang Chen\textsuperscript{a#}, Francesco Canonaco\textsuperscript{b}, Jay G. Slowik\textsuperscript{a}, Kaspar R. Daellenbach\textsuperscript{a}, Anna Tobler\textsuperscript{b}, Jean-Eudes Petit\textsuperscript{c}, Olivier Favez\textsuperscript{d}, Iasonas Stavroulas\textsuperscript{e}, Nikolaos Mihalopoulos\textsuperscript{e}, Evangelos Gerasopoulos\textsuperscript{e}, Imad El Haddad\textsuperscript{a}, Urs Baltensperger\textsuperscript{a}, and André S.H. Prévôt\textsuperscript{a*}

ADDRESS

\textsuperscript{a}Laboratory of Atmospheric Chemistry, Paul Scherrer Institute, Villigen, Aargau, 5232, Switzerland

\textsuperscript{b}Datalystica Ltd., Park innovAARE, Villigen, Aargau, 5234, Switzerland

\textsuperscript{c}Laboratoire des Sciences du Climat et de l’Environnement, CEA/Orme des Merisiers, Gif-sur-Yvette, 91191, France

\textsuperscript{d}INERIS, Parc Technologique ALATA, Verneuil-en-Halatte, 60550, France
Institute for Environmental Research and Sustainable Development, National Observatory of Athens, Athens, 15236, Greece

*Email: andre.prevot@psi.ch

Now at: G.C.: MRC Centre for Environment and Health, Environmental Research Group, Imperial College London, London, W12 0BZ, United Kingdom

KEYWORDS: Real-time; Source Apportionment; Organic Aerosol; Urban Pollution
Table of Figures

Figure S1. The step-by-step flow chart of instructions for SoFi RT ...............................................3

Figure S2. Yearly averaged source apportionment comparison for Paris data for factor profiles (left with mirrored y-axis), time series with 24-hour resolution (middle), and diurnal cycles (right). Blue colours indicate real-time results, black colours indicate best estimate results.................................5

Figure S3. Yearly averaged source apportionment comparison for Zurich data for factor profiles (left with mirrored y-axis), time series with 24-hour resolution (middle), and diurnal cycles (right). Blue colours indicate real-time results, black colours indicate best estimate results. .....................6

Figure S4. Scatter plots for real-time vs. best-estimate results in mass concentrations for each factor at all three sites (color-coded by date and time).....................................................................................7

Figure S5. Scatter plots for real-time vs. best-estimate results in relative contributions for each factor at all three sites (color-coded by date and time).....................................................................................8

Figure S6. (Real-time)/(Best Estimate) vs. mass fraction of each OA factor binned into 10 groups based on best-estimate results. The markers and lines represent the median, and the error bars the interquartile ranges. ..................................................................................................................................9
Figure S 7. Time series of uncentered correlations between real-time and best-estimate results for all OA factors.

Figure S 8. Histogram of scaled residual of three datasets retrieved from best-estimate results (left) and real-time results.
42 **S1. Step-by-step instructions for SoFi RT**

**Figure S1** The step-by-step flow chart of instructions for SoFi RT

The detailed instructions of SoFi RT are shown in Figure S1. Specifically, seasonal PMF pre-tests (typically 1-2 month data) need to be conducted if there is no a priori information (number of factors and representative factor profiles) available for the current dataset. Then, the rolling PMF settings initialized by inputting necessary parameters to automate the analysis. The initialization details are provided in Section 1.1 below. With co-located complementary measurements (e.g., AE33, NOx, CO) available, SoFi RT can automatically pre-treat the data and make them available for validation of the source apportionment results using the criteria-based selection function within SoFi RT. The user shall modify the criteria list accordingly based on available complementary data from the suggested criteria list by Chen et al. (2022)^1. With continuous measurements from
the ACSM, SoFi RT will update the PMF input for rolling PMF analysis on a daily basis. After averaging the selected rolling PMF results (meeting the requirements of all criteria), the factor profiles will be ready to conduct CMB-like analysis for all incoming data on the following day with all factor profiles fully constrained at $a$ value = 0. Once the current day is over, the rolling PMF will advance to the next 14-day time window and repeat the previous analysis to provide averaged and updated source profiles for CMB-like analysis on future data.

### S1.1 Initializing SoFi RT

There are parameters users need to prepare and input before starting running SoFi RT:

- **a.** The calibration parameters (CE/CDCE, RIE);
- **b.** The number of factors and factor profiles of primary organic aerosol (POA) factors (from a priori information or seasonal PMF pre-tests);
- **c.** Rolling PMF constrains the POA factors (from a priori information or seasonal PMF pre-tests) using a random $a$-value approach with the upper $a$ value of 0.5 and step of 0.1 (as suggested by Canonaco et al. (2021));
- **d.** Rolling PMF time window =14 days, step = 1 day, repeats =50
S2. Residual calculation for SoFi RT

The residuals are retrieved from Equation S1 as follows:

\[ x_{ij} = \sum_{k=1}^{p} g_{ik} \times f_{kj} + e_{ij} \]  

S1

SoFi RT still uses PMF but with all profiles fully constrained with \( \alpha \) value=0. Therefore, the \( f_{kj} \) will be fixed for the current calendar day, the concentration of each factor will be modelled accordingly, and the residuals will be calculated through this equation. The residual distribution is examined for normality.

S3. Signs for updating \textit{a priori} information

In order to get robust real-time source apportionment results, users need to monitor constantly the SoFi RT result panel (including distribution of scaled residual, pie chart of OA fractions, time series of OA factors, and averaged OA factor profiles) that is updated as new data feed in. If there
are any following signs observed, seasonal pre-tests are highly recommended to update a priori information for future real-time source apportionment.

1. Non-unimodal distribution of scaled residual

2. Contribution of a source approaches to 0%

3. Consistent chemical markers appearing in the LO-OOA (i.e., typically the fresher OOA factor among two unconstrained factors) factor profile

---

**Figure S2.** Yearly averaged source apportionment comparison for Paris data for factor profiles (left with mirrored y-axis), time series with 24-hour resolution (middle), and diurnal cycles (right). Blue colours indicate real-time results, black colours indicate best estimate results.
Figure S3. Yearly averaged source apportionment comparison for Zurich data for factor profiles (left with mirrored y-axis), time series with 24-hour resolution (middle), and diurnal cycles (right). Blue colours indicate real-time results, black colours indicate best estimate results.
Figure S4. Scatter plots for real-time vs. best-estimate results in mass concentrations for each factor at all three sites (color-coded by date and time).
Figure S5. Scatter plots for real-time vs. best-estimate results in relative contributions for each factor at all three sites (color-coded by date and time)
**Figure S6.** (Real-time)/(Best Estimate) vs. mass fraction of each OA factor binned into 10 groups based on best-estimate results. The markers and lines represent the median, and the error bars the interquartile ranges.

**Figure S 7.** Time series of uncentered correlations between real-time and best-estimate results for all OA factors.
Figure S 8. Histogram of scaled residual of three datasets retrieved from best-estimate results (left) and real-time results.