Environmental Science & Technology

pubs.acs.org/est

Article

Real-Time Source Apportionment of Organic Aerosols in Three European Cities

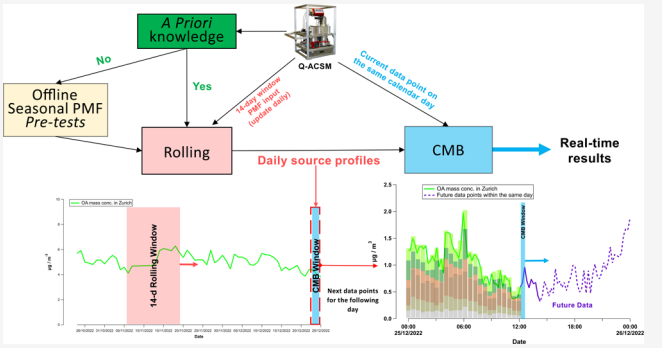
Gang Chen, Francesco Canonaco, Jay G. Slowik, Kaspar R. Daellenbach, Anna Tobler, Jean-Eudes Petit, Olivier Favez, Iasonas Stavroulas, Nikolaos Mihalopoulos, Evangelos Gerasopoulos, Imad El Haddad, Urs Baltensperger, and André S. H. Prévôt*

Cite This: https://doi.org/10.1021/acs.est.2c02509

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: 97% of the urban population in the EU in 2019 were exposed to an annual fine particulate matter level higher than the World Health Organization (WHO) guidelines (5 $\mu g/m^3$). Organic aerosol (OA) is one of the major air pollutants, and the knowledge of its sources is crucial for designing cost-effective mitigation strategies. Positive matrix factorization (PMF) on aerosol mass spectrometer (AMS) or aerosol chemical speciation monitor (ACSM) data is the most common method for source apportionment (SA) analysis on ambient OA. However, conventional PMF requires extensive human labor, preventing the implementation of SA for routine monitoring applications. This study proposes the source finder real-time (SoFi RT, Datalystica Ltd.) approach for efficient retrieval of OA sources. The results generated by SoFi RT agree remarkably well with the conventional rolling PMF results regarding factor profiles, time series, diurnal patterns, and yearly relative contributions of OA factor on three year-long ACSM data sets collected in Athens, Paris, and Zurich. Although the initialization of SoFi RT requires a *priori* knowledge of OA sources (i.e., the approximate number of factors and relevant factor profiles) for the sampling site, this technique minimizes user interactions. Eventually, it could provide up-to-date trustable information on timescales useful to policymakers and air quality modelers.



KEYWORDS: *real-time, source apportionment, organic aerosol, urban pollution, SoFi RT*

1. INTRODUCTION

Atmospheric aerosols are mixtures of liquid or solid particles suspended in the air. They strongly affect the climate, the ecosystem, and public health. The mass concentration of particulate matter with a diameter smaller than 2.5 μm ($PM_{2.5}$) has been associated with oxidative stress in many studies.[1−3] The World Health Organization (WHO) updated the air quality guideline of the annual average of $PM_{2.5}$ to 5 $\mu g/m^3$ in 2021,[4] an exposure limit exceeded by 99% of the global population. However, reducing the overall PM levels is extremely challenging and may not be the most cost-effective way to reduce the corresponding mortalities because different sources/compositions of PM have different health impacts.[5] Organic aerosol (OA) is a major fraction (20−90%) of fine PM.[6−9] In addition, compared to inorganic aerosol, the sources and compositions are much more complex. Therefore, it is more important than ever to quantify OA sources to reduce aerosol-related premature deaths.

The chemical components of OA are typically identified and quantified by applying positive matrix factorization (PMF[10]) on OA mass spectra collected by aerosol mass spectrometers (AMS) or aerosol chemical speciation monitors (ACSM) (both from Aerodyne Research, Inc.).[9,11,12] OA chemical components, commonly refers to as OA sources, are typically identified by their characteristic mass spectra, which one commonly refers to as OA sources. Recently, rolling PMF analysis was shown to be a powerful tool to retrieve more robust and accurate OA source apportionment (SA) results in long-term (i.e., multiseason) data sets[13−16] by considering temporal variations of the factor profiles. However, these analyses remain time-consuming for long-term datasets due to the complex preprocessing, rolling PMF procedures, and post-PMF analysis of the data.[9] Typically, it takes a well-trained analyst several months to a year to retrieve high-quality SA results for a year-long data set from scratch. These requirements are difficult to reconcile with the practical constraints

Received: April 10, 2022
Revised: September 26, 2022
Accepted: September 28, 2022

governing routine monitoring applications, where a single technician may be responsible for a large number of diverse instruments and/or lack specialized expertise in source apportionment. However, the ability of SA to reduce a large and highly complex input data set to a small number of sources linked to real-world processes is an optimal output for such monitoring initiatives, providing valuable information for timely responses to extreme air quality events, evaluation of policy initiatives, and development and validation of air quality forecasts.

In this study, we propose an automated and state-of-the-art SA technique that can provide robust OA source information in real-time once 14 days of measurements (typically in 30 min resolution) have been collected. This technique has been implemented and validated using data sets collected from three European cities, including Athens,[17] Paris,[18] and Zurich.[9,19]

2. METHODS

2.1. Instrumentation and Measurement Sites. ACSMs[20,21] are designed for the quantification of the mass concentration of nonrefractory constituents of ambient aerosol particles (NR-PM) (i.e., sulfate, nitrate, ammonium, chloride, and organics). These instruments are based on AMS technology, but they are optimized for long-term monitoring purposes via reduction in both cost and user operation/maintenance requirements. However, they are less sensitive, are normally capable of only integer mass resolution, and lack a measurement of particle size.

PM is sampled through a critical orifice (100 μm inner diameter) at a flow rate of 1.4 cm^3/s (at 1 atm. and 20 °C). Then, an aerodynamic lens focuses the sampled particles into a narrow beam to impact them on a tungsten or an iridium surface (~600 °C), where the NR-PM is vaporized and ionized by an electron impact source at 70 eV. The quadrupole ACSM (Q-ACSM) detects ions up to a mass-to-charge ratio (m/z) of 148 Th,[20] while the time-of-flight ACSM (ToF-ACSM) provides higher mass resolution and detects ions up to m/z = 300 Th.[21,22] In this study, Q-ACSMs were used with a PM_1 aerodynamic (standard) lens and the standard vaporizer for all data sets, which were collected in three European cities (Athens, Paris, and Zurich), and specifically, the Athens data range from July 2016 to July 2017, the Paris data range from January 2016 to May 2017, and the Zurich data range from August 2016 to July 2017. The sampling sites are described in more details in previous publications[17−19] and are also presented as part of Chen et al.[9]

2.2. Receptor Models. PMF and chemical mass balance (CMB) are two receptor models that have been widely used to identify OA sources. Specifically, both PMF[23] and CMB are bilinear factor models with non-negativity constraints with the following mass balance equation:

$$x_{ij} = \sum_{k=1}^{p} g_{ik} \times f_{kj} + e_{ij} \quad (1)$$

where x_{ij} is the measured OA at the i^{th} time point and the j^{th} variable, while in terms of model outputs, g_{ik} represents the factor time series for the k^{th} factor, f_{kj} represents the factor profiles, and e_{ij} represents the residual. The PMF models iteratively minimize the following quantity Q:

$$Q = \sum_{i=1}^{n} \sum_{j=1}^{m} \left(\frac{e_{ij}}{\sigma_{ij}} \right)^2 \quad (2)$$

Here, σ_{ij} is the measurement uncertainties, corresponding point-by-point to x_{ij}. However, the PMF is subject to rotational ambiguity, meaning that different solutions can have similar values of Q. Although they are of similar quality mathematically, they may not be of equal environmental relevance. To address this, the multilinear engine (ME-2)[24] has been demonstrated to be a powerful tool to direct the model toward environmentally meaningful matrix rotations[13,19] by using the a *priori* information of corresponding mass spectral or time series (the a-value approach). The a value determines the extent to which the final mass spectrum or time series is allowed to deviate from the anchor (mass spectral/time series) during the PMF model iteration. The classical CMB is a simplified version of the PMF model with predetermined factor profiles and number of factors but with a different cost function as shown below:

$$Q = \sum_{i} \sum_{j} \frac{e_{ij}^{\,2}}{s_{ij}^{\,2} + \left(\sum_k g_{ik} \sigma_{f_{jk}} \right)^2} \quad (3)$$

This cost function distinguishes the classical CMB from the CMB-like method used here, which is described in Section 2.4.1.

2.3. Rolling PMF with ME-2. While the conventional PMF only allows having static factor profiles during the PMF window of analysis, in reality some factor profiles are expected to vary with time. To take these temporal variations of factor profiles into account, PMF can be performed over a small time window (e.g., 14 days).[16] Then, this window is moved to the next position with a step size of one day,[16] in which the factor profiles are allowed to adapt. Canonaco et al. (2021)[13] utilized this technique, the so-called rolling PMF,[16,23] to improve the quality of long-term (multiseason) SA results. At the same time, the rotational ambiguity of the PMF model can be minimized using the multilinear engine (ME-2),[19] which users need to provide a *priori* information of factor profiles/time series as constraints for the PMF model. Chen et al.[9] then developed a standardized protocol to retrieve robust and comparable SA results across Europe using the rolling PMF with the ME-2.

2.4. Real-Time SA Technique. *2.4.1. Overview of Process.* To address the challenges mentioned before, a real-time SA technique is proposed in this study. Generally, the seasonal PMF *pre-tests*[9] are required to identify the number of factors and/or representative factor profiles, especially for the new data set to account for some unknown and/or additional (primary or secondary) sources are expected. Second, the rolling PMF (using the a *priori* information from the last step) is conducted for the previous 14-day time window to initialize the source profiles. Finally, assuming that the source profiles on the current day are consistent with the factor profiles retrieved from the previous 14-day rolling PMF window, the real-time SA results are generated through the CMB-like analysis (i.e., the PMF model with all factors fully constrained at a value = 0) for the upcoming measured points on the current day. By advancing the rolling window in 1-day steps, the source profile will be updated to generate a more robust real-time SA results. All these processes have been integrated

https://doi.org/10.1021/acs.est.2c02509
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

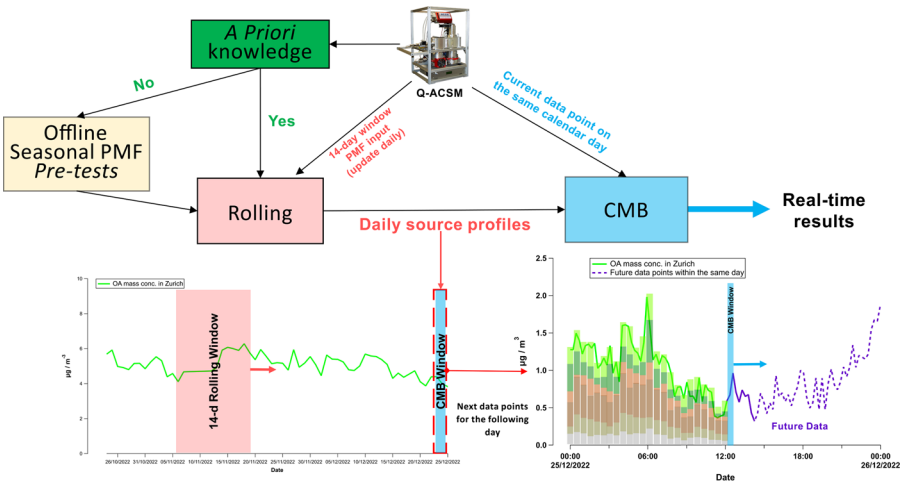


Figure 1. Schematic flow diagram of the real-time source apportionment. CMB is implemented through the PMF model with all factors fully constrained at *a* value = 0.

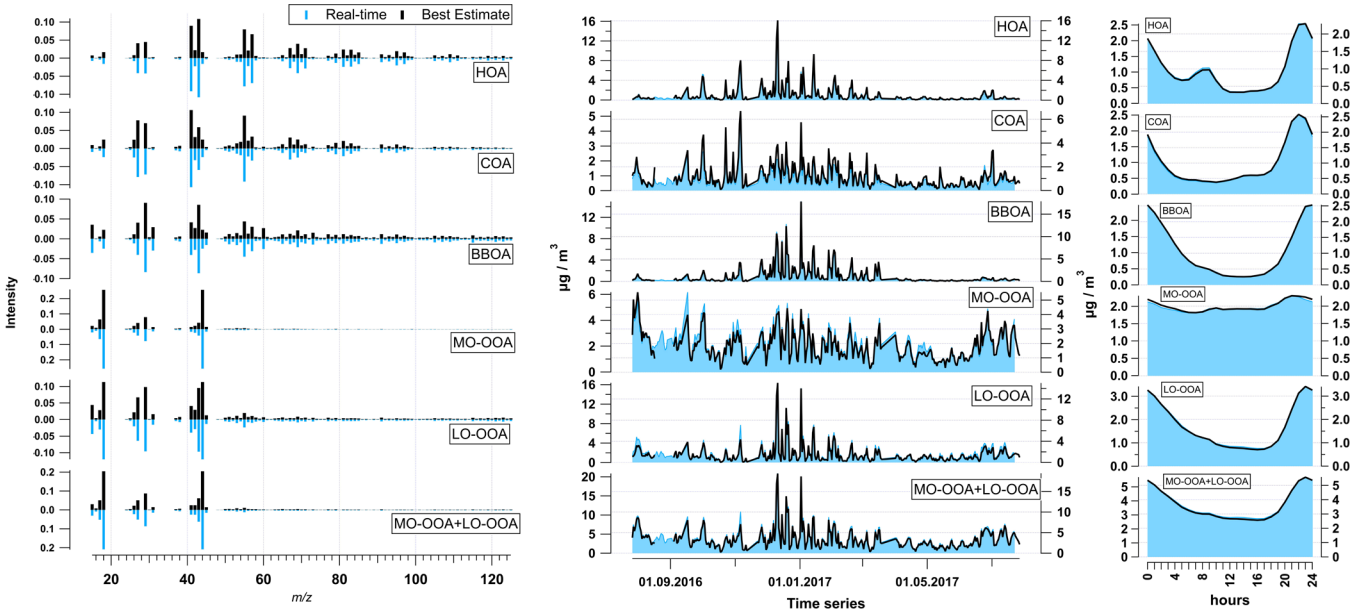


Figure 2. Yearly averaged source apportionment comparison for Athens data for factor profiles (left, with mirrored *y*-axis), a time series with a 24 h resolution (middle), and diurnal cycles (right). The blue color indicates the real-time results and the black color indicates the best-estimate results.

into a software application, the source finder real-time (SoFi RT, Datalystica, Ltd., Switzerland), which is commercially available and distributed by Datalystica.

2.4.2. Data Preparation and PMF Initialization/Setup. Several steps are required to prepare the PMF input for real-time SA. First, the SoFi RT automatically calls the functions within the ACSM Local software (ACSM Local 1.6.1.3, Aerodyne, Inc.) to apply correct fragmentation tables and ion transmission efficiency corrections and to calculate the error matrix to prepare the PMF input. Then the SoFi RT automatically applies the collection efficiency (CE) correction (user-defined, i.e., the constant CE or composition-dependent CE (CDCE)[25]), which can be defined in advance.

Moreover, to retrieve robust SA results, a few additional steps are required in advance. As mentioned above, seasonal PMF *pre-tests*[9] are normally required to gain some a *priori* OA source information of the sampling site. There is a lot of knowledge on OA sources available from many SA studies across the world. For Europe, Chen et al.[9] performed the necessary analysis for 22 sites. Therefore, in many data sets it should be straightforward to define the number of sources and representative source profiles. For example, the OA sources in the three data sets used in this work have been well studied. In addition, since rolling PMF generally yields lot of PMF runs, a set of criteria and corresponding statistical tests for probabilistic thresholds should be defined in advance to select environmentally reasonable solutions by following the standardized protocol as suggested by Chen et al.[9] For criteria based on external data, like black carbon, NO_x, etc., the SoFi RT can also automatically retrieve and preprocess those data and make them ready for criterion evaluations.

2.4.3. Automated RT Analysis. As soon as more than a full time window of data (here 14 days) has been collected with the above initialization and setup, the SoFi RT is able to retrieve the robust real-time OA sources automatically. As shown in Figure 1, a PMF input of a 14-day time window is created and used to conduct rolling PMF using the specified timing/repeat parameters, which here follow the 14-day

https://doi.org/10.1021/acs.est.2c02509
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

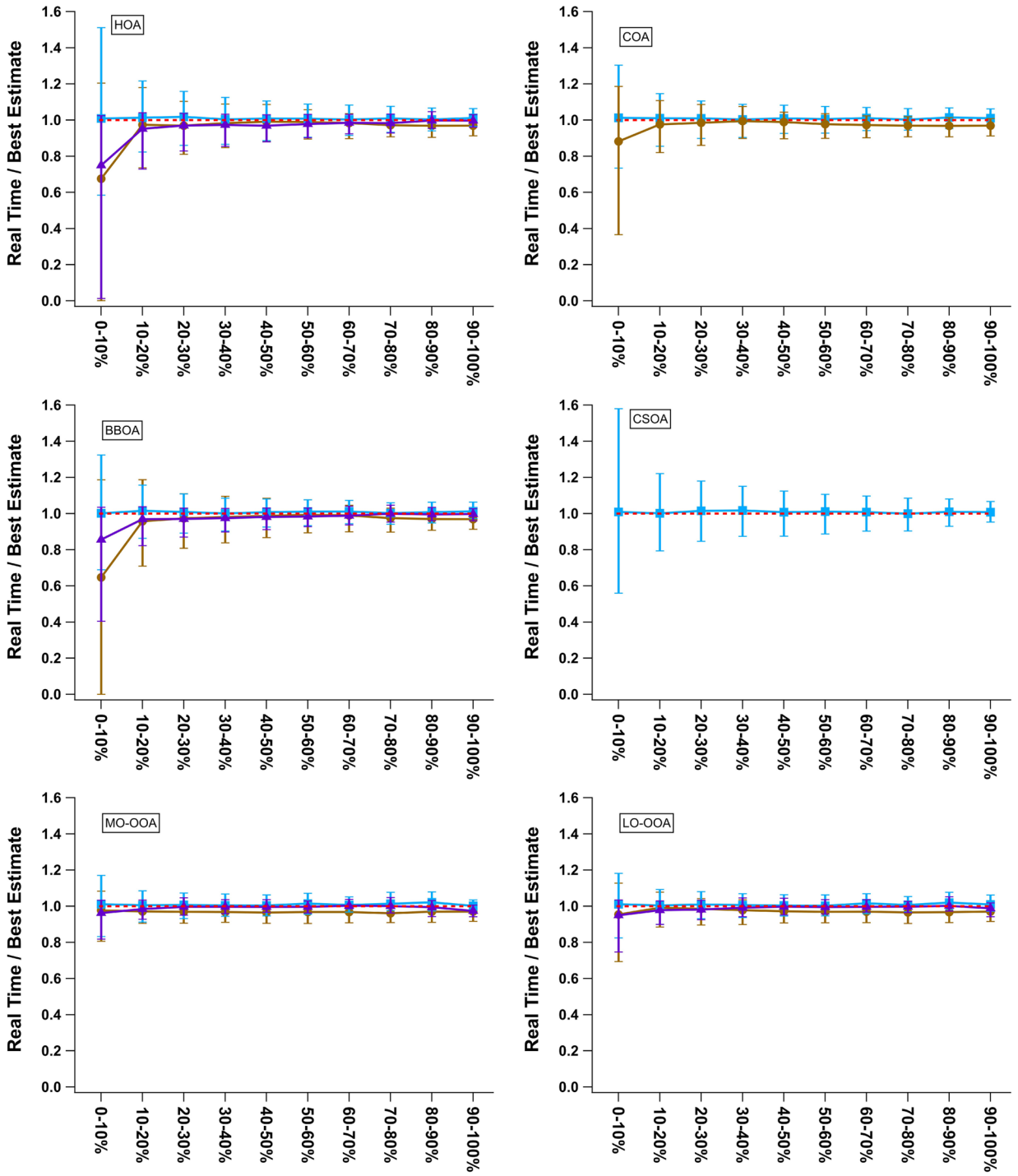


Figure 3. (Real-time)/(best estimate) vs absolute mass concentration of each OA factor binned in 10 groups based on the best-estimate results. The markers and the lines represent the medians, and the error bars represent the interquartile ranges.

window, 1-day step, and 50 repeats proposed by Canonaco et al.[13] In addition, the SoFi RT utilizes ME-2[26,27] to constrain the primary organic aerosol (POA) factor profiles (to reduce the rotational ambiguity) based on the profiles previously identified for the site. The profiles are constrained by a random *a* value in a range of 0−0.4.[13] By averaging the environmentally reasonable solutions from this 14-day rolling window, the source profiles can be generated for the CMB-like analysis on the upcoming data point (from the current calendar day) to retrieve real-time SA results. With the 1-day step of the rolling PMF window, the factor profiles for the CMB-like analysis are updated on a daily basis. Therefore, the factor profiles used in the CMB-like analysis can actually adapt by using recent data, which considerably decreases the uncertainties of this method. More detailed instructions of the SoFi RT are described in SI and Figure S1.

3. SOFI RT IMPLEMENTATION AND VALIDATION

In this study, we used year-long ACSM data sets collected from three European cities (Athens, Paris, and Zurich) to validate the real-time SA technique. The offline rolling results were obtained by following the standardized rolling PMF protocol

https://doi.org/10.1021/acs.est.2c02509
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

developed by Chen et al.,[9] which is considered as the most advanced and robust OA source apportionment (and is, therefore, referred to as the "best estimate" in the text as follows). All three sites resolved a hydrocarbon-like OA (HOA) factor, a biomass burning OA (BBOA) factor, a more-oxidized oxygenated OA (MO-OOA) factor, and a less-oxidized oxygenated OA (LO-OOA) factor. In addition, a cooking-like OA (COA) factor was retrieved in Athens and Zurich. Moreover, a cigarette smoking OA (CSOA) factor was identified in the Zurich data set. To validate the real-time technique, we conducted head-to-head comparisons between the best-estimate results and the real-time results regarding source profiles, time series, diurnal cycles, and relative contributions, as outlined below.

4. RESULTS AND DISCUSSION

Figure 2 compares Athens' yearly averaged factor profiles, diurnal cycles, and time series (24 h average) between the best-estimate results and the real-time results. It clearly shows a high consistency between these two methods. The same conclusions were drawn from the comparisons of the Paris and Zurich data sets (Figures S2 and S3). The linear regression between the time series of each factor retrieved from the two methods was conducted using a 30 min time resolution (Figure S4). We found that the factors of all stations generally show a very good agreement (with $R^2_{Pearson} > 0.94$ except LO-OOA from the Paris data set with an $R^2_{Pearson}$ of 0.89) and biases <10%. The POA factors generally show better agreements with the higher $R^2_{Pearson}$ and slopes closer to 1 than the OOA factors. This is expected since the POA factor profiles are constrained during the offline rolling PMF. Similarly, we also compared the relative contributions of each factor (with 30 min resolution) between the two methods, as shown in Figure S5. It shows good consistencies (with the $R^2_{Pearson} > 0.73$ except the LO-OOA in Paris with $R^2_{Pearson} = 0.59$) and <10.6% differences from the 1:1 line. This agreement is not as good as for the absolute concentrations, but this is expected since the correlation coefficient for the fraction comparisons can be affected considerably by the low mass concentration points (often with large uncertainty), which eventually results in relatively worse agreements between these two methods.

In addition, the agreement between the real-time results and the best-estimated results is better for total OOA (MO-OOA + LO-OOA) than for either MO-OOA or LO-OOA alone in terms of mass concentration and relative contribution (Figures S4 and S5). This means that the separation of OOA into its subfractions is somewhat less robust than the separation of the POAs and total OOA. This is expected since the separation between MO-OOA and LO-OOA is often based on gradients in photochemical age, volatility, etc. instead of discrete sources.[6,11,28] Thus, it has more cross-talk between the vectors in response to perturbations of the OOA fraction.[6,11,28] The disagreement in LO-OOA in Paris during the warm periods (yellow points in Figure S4) is most likely caused by this reason. Nevertheless, this technique still shows satisfactory agreement when separating OOA into LO-OOA and MO-OOA as discussed in the previous paragraph.

To understand the causes of these discrepancies in the abovementioned comparisons, we divided the data into 10 bins based on the percentile absolute mass concentration of each OA factor retrieved from the best-estimate results (x-axis of Figure 3). Then we investigated how the mass concentration of the OA factor affects the differences between the two methods (Figure 3). In general, the real-time to best-estimate ratio approaches 1 (with smaller standard deviations) as the mass concentration of OA factor increases. Also, the bottom 10% of the mass concentrations of each OA factor shows the largest discrepancies (median) and the largest interquartile range. When the mass concentration of the OA factor gets too small, it approaches the detection limit of the ACSM, which leads to a high measurement uncertainty. As shown in Figure S6, when we group the data based on the mass fraction of each OA factor, the bottom 10% bin exhibits the largest discrepancies, except for the LO-OOA in Paris (due to the inherent inseparability between the LO-OOA and the MO-OOA as discussed above). It suggests that when the contribution of key ions for the corresponding OA factors is low, it results in larger uncertainties of both the PMF and the real-time model. Therefore, the discrepancies at a low factor mass concentrations/fractions between the two methods are most likely caused by a combination of the high uncertainties of the measurements and/or PMF/real-time models.

Since the CMB-like profiles are taken from the previous 14-day rolling window while the offline rolling PMF from the best-estimate solution keeps updating its profiles, the discrepancies between the CMB-like factor profiles (used to generate real-time SA results) and the factor profiles for the offline rolling PMF are expected for a given time point. Therefore, we compared the CMB-like factor profiles with the offline rolling PMF factor profiles at each time point using the uncentered correlation (R)[29] (Figure S7) to validate the assumption that the source profiles for the current day are basically identical to the source profile from the real-time approach. Figure S7 shows high values for the averaged uncentered correlation, generally with $R > 0.92$. In addition, the POA (except CSOA in Zurich, with a maximum and a minimum R of 0.92 and 0.42, respectively) and the MO-OOA factor profiles are almost constant over time, with an average R greater than 0.99. The low correlation of CSOA is explained by the fact that cigarettes' smoke is not continuously present, which leads to large uncertainties when the source is absent. LO-OOA shows relatively larger discrepancies but still with an average R greater than 0.95. This is expected by the fact that the LO-OOA factor profile from the rolling analysis often has a larger temporal variation than the one of MO-OOA.[9] Also, LO-OOA could be quite sensitive to pollution episodes since it is a rather "fresh" and unconstrained factor. Nevertheless, the comparison in Figure S7 demonstrates that the assumption of this technique is valid.

Overall, these two techniques provide very similar yearly averaged OA factor contributions for all three data sets, with differences smaller than 2.7 percentage points (Figure 4). The POA contributions appear to be generally more consistent ($0.1 < \Delta < 0.9$) than those of the OOA factors ($0.5 < \Delta < 2.7$). This is expected since the OOA factors were unconstrained for the best-estimate model (offline rolling PMF) but fully constrained in the real-time approach. As shown in Figure S8, the real-time results have similar or slightly larger-scaled residuals than the best-estimate results. In all cases, the distributions of scaled residuals meet the expectations.

In this study, we have developed a state-of-the-art source apportionment technique that can minimize human interactions once the model has been initialized and allowing us to efficiently retrieve robust OA sources in real-time, i.e., minutes after the data is collected. The technique was validated using

https://doi.org/10.1021/acs.est.2c02509
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

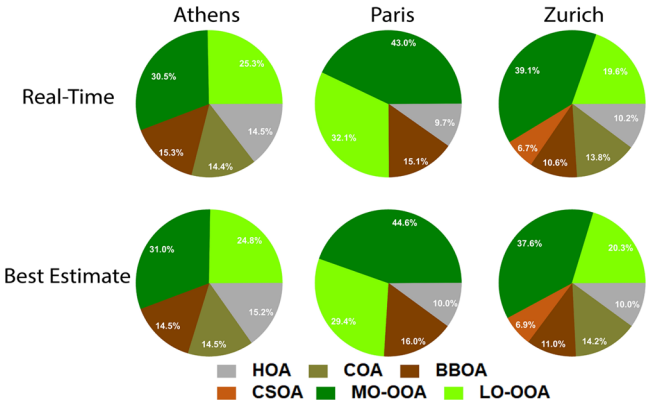


Figure 4. Yearly average mass fractions of OA factors at each site using the real-time and rolling window (denoted "best estimate") methods.

three year-long Q-ACSM data sets collected in three European cities. We found very high consistencies between the real-time results and the best-estimate results for the factor profiles (uncentered correlation >0.92), the time series ($R^2_{Pearson}$ > 0.89 and bias <12.5%), the diurnal patterns, and the yearly relative contributions (Δ < 2.7 percentage points). The discrepancies between the two methods were found at low factor mass concentrations and are mainly due to a combination of the high uncertainties of the measurements and/or the PMF/real-time model. Even though this technique requires some a *priori* knowledge of OA sources (i.e., the number of factors and the relevant factor profiles as constraints, which can be done by conducting seasonal *pre-tests*[9]) for the sampling site, the robustness of this technique is remarkable. For appropriate ACSM/AMS data sets and corresponding OA source apportionment studies, which are already available worldwide, this technique could ideally be implemented and start generating robust SA results instantly in a real-time manner. However, this method may still suffer from a changing number/type of factors during the analysis. Therefore, more work should be invested to allow the software to automatically adjust the constraints of rolling PMF accordingly. For instance, integrating clustering algorithms (e.g., k-means and hierarchical clustering[30,31]) into SoFi RT could potentially contribute in this respect and might be investigated further. Currently, SoFi RT is also applicable of handling online metal (i.e., Xact, Cooper Environmental Inc.) and black carbon (e.g., AE33, Magee Scientific Corp.) data. More efforts should be put in coupling the ACSM, Xact, and AE33 data to improve the SA results in real-time. Finally, these more timely and high-quality OA source information in real-time could help policymakers design and validate mitigation strategies. These results could also improve air quality models by providing extra inputs and eventually pave the way for a next-generation air quality forecast, especially relevant for safeguarding public health in megacities.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.est.2c02509.

S1: Step-by-step instructions for SoFi RT; S2: Residual calculation for SoFi RT; S3: Signs for updating a *priori* information; Figure S1: The step-by-step flow chart of instructions for SoFi RT; Figure S2: yearly averaged source apportionment comparison for Paris data for factor profiles, time series with 24-hour resolution, and diurnal cycles; Figure S3: yearly averaged source apportionment comparison for Zurich data for factor profiles; Figure S4: scatter plots for real-time vs. best-estimate results in mass concentrations for each factor at all three sites; Figure S5: scatter plots for real-time vs. best-estimate results in relative contributions for each factor at all three sites; Figure S6: (Real-time)/(best estimate) vs. mass fraction of each OA factor binned into 10 groups based on best-estimate results; Figure S7: time series of uncentered correlations between real-time and best-estimate results for all OA factors; Figure S8: histogram of scaled residual of three datasets retrieved from best-estimate results and real-time results (PDF)

■ AUTHOR INFORMATION

Corresponding Author

André S. H. Prévôt − *Laboratory of Atmospheric Chemistry, Paul Scherrer Institute, Villigen, Aargau 5232, Switzerland;* orcid.org/0000-0002-9243-8194; Phone: +41 56 310 4202; Email: andre.prevot@psi.ch

Authors

Gang Chen − *Laboratory of Atmospheric Chemistry, Paul Scherrer Institute, Villigen, Aargau 5232, Switzerland;* Present Address: MRC Centre for Environment and Health, Environmental Research Group, Imperial College London, London, W12 0BZ, U.K. (G.C.); orcid.org/0000-0002-1507-4622

Francesco Canonaco − *Datalystica Ltd., Park innovAARE, Villigen, Aargau 5234, Switzerland*

Jay G. Slowik − *Laboratory of Atmospheric Chemistry, Paul Scherrer Institute, Villigen, Aargau 5232, Switzerland*

Kaspar R. Daellenbach − *Laboratory of Atmospheric Chemistry, Paul Scherrer Institute, Villigen, Aargau 5232, Switzerland*

Anna Tobler − *Datalystica Ltd., Park innovAARE, Villigen, Aargau 5234, Switzerland*

Jean-Eudes Petit − *Laboratoire des Sciences du Climat et de l'Environnement, CEA/Orme des Merisiers, Gif-sur-Yvette 91191, France*

Olivier Favez − *INERIS, Parc Technologique ALATA, Verneuil-en-Halatte 60550, France*

Iasonas Stavroulas − *Institute for Environmental Research and Sustainable Development, National Observatory of Athens, Athens 15236, Greece*

Nikolaos Mihalopoulos − *Institute for Environmental Research and Sustainable Development, National Observatory of Athens, Athens 15236, Greece*

Evangelos Gerasopoulos − *Institute for Environmental Research and Sustainable Development, National Observatory of Athens, Athens 15236, Greece*

Imad El Haddad − *Laboratory of Atmospheric Chemistry, Paul Scherrer Institute, Villigen, Aargau 5232, Switzerland;* orcid.org/0000-0002-2461-7238

Urs Baltensperger − *Laboratory of Atmospheric Chemistry, Paul Scherrer Institute, Villigen, Aargau 5232, Switzerland*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.est.2c02509

Notes

The authors declare the following competing financial interest(s): F.C. and A.T. are employed by Datalystica Ltd., which is the official distributor of the SoFi Pro and SoFi RT licenses.

■ ACKNOWLEDGMENTS

This research was supported by the COST action CA16109 Chemical On-Line cOmpoSition and Source Apportionment of fine aerosoLs (COLOSSAL), the SNF COST project SAMSAM IZCOZO_177063. This project has received funding from the European Union's Horizon 2020 Framework Programme via the ACTRIS-2 project under grant NO. 654109, the RI-URBANS project under grant NO. 101036245, the ERA-PLANET, and transnational projects SMURBS and iCUPE (grant agreement NO. 689443).

■ REFERENCES

(1) Beelen, R.; Hoek, G.; van den Brandt, P. A.; Goldbohm, R. A.; Fischer, P.; Schouten, L. J.; Jerrett, M.; Hughes, E.; Armstrong, B.; Brunekreef, B. Long-Term Effects of Traffic-Related Air Pollution on Mortality in a Dutch Cohort (NLCS-AIR Study). *Environ. Health Perspect.* **2008**, *116*, 196−202.

(2) Krewski, D.; Jerrett, M.; Burnett, R. T.; Ma, R.; Hughes, E.; Shi, Y.; Turner, M. C.; Pope, C. A.; Thurston, G.; Calle, E. E.; Thun, M. J.; Beckerman, B.; DeLuca, P.; Finkelstein, N.; Ito, K.; Moore, D. K.; Newbold, K. B.; Ramsay, T.; Ross, Z.; Shin, H.; Tempalski, B.Extended Follow-up and Spatial Analysis of the American Cancer Society Study Linking Particulate Air Pollution and Mortality. *Res. Rep. 140, Health. Eff. Inst.*, 2009.

(3) Pope, C. A., III Lung Cancer, Cardiopulmonary Mortality, and Long-Term Exposure to Fine Particulate Air Pollution. *JAMA* **2002**, *287*, 1132.

(4) World Health Organization. *WHO Global Air Quality Guidelines. Particulate Matter ($PM_{2.5}$ and PM_{10}), Ozone, Nitrogen Dioxide, Sulfur Dioxide and Carbon Monoxide*: Geneva, 2021; vol *368*.

(5) Daellenbach, K. R.; Uzu, G.; Jiang, J.; Cassagnes, L.-E.; Leni, Z.; Vlachou, A.; Stefenelli, G.; Canonaco, F.; Weber, S.; Segers, A.; Kuenen, J. J. P.; Schaap, M.; Favez, O.; Albinet, A.; Aksoyoglu, S.; Dommen, J.; Baltensperger, U.; Geiser, M.; El Haddad, I.; Jaffrezo, J.-L.; Prévôt, A. S. H. Sources of Particulate-Matter Air Pollution and Its Oxidative Potential in Europe. *Nature* **2020**, *587*, 414−419.

(6) Jimenez, J. L.; Canagaratna, M. R.; Donahue, N. M.; Prevot, A. S. H.; Zhang, Q.; Kroll, J. H.; DeCarlo, P. F.; Allan, J. D.; Coe, H.; Ng, N. L.; Aiken, A. C.; Docherty, K. S.; Ulbrich, I. M.; Grieshop, A. P.; Robinson, A. L.; Duplissy, J.; Smith, J. D.; Wilson, K. R.; Lanz, V. A.; Hueglin, C.; Sun, Y. L.; Tian, J.; Laaksonen, A.; Raatikainen, T.; Rautiainen, J.; Vaattovaara, P.; Ehn, M.; Kulmala, M.; Tomlinson, J. M.; Collins, D. R.; Cubison, M. J.; Dunlea, J.; Huffman, J. A.; Onasch, T. B.; Alfarra, M. R.; Williams, P. I.; Bower, K.; Kondo, Y.; Schneider, J.; Drewnick, F.; Borrmann, S.; Weimer, S.; Demerjian, K.; Salcedo, D.; Cottrell, L.; Griffin, R.; Takami, A.; Miyoshi, T.; Hatakeyama, S.; Shimono, A.; Sun, J. Y.; Zhang, Y. M.; Dzepina, K.; Kimmel, J. R.; Sueper, D.; Jayne, J. T.; Herndon, S. C.; Trimborn, A. M.; Williams, L. R.; Wood, E. C.; Middlebrook, A. M.; Kolb, C. E.; Baltensperger, U.; Worsnop, D. R. Evolution of Organic Aerosols in the Atmosphere. *Science* **2009**, *326*, 1525−1529.

(7) Murphy, D. M.; Cziczo, D. J.; Froyd, K. D.; Hudson, P. K.; Matthew, B. M.; Middlebrook, A. M.; Peltier, R. E.; Sullivan, A.; Thomson, D. S.; Weber, R. J. Single-Particle Mass Spectrometry of Tropospheric Aerosol Particles. *J. Geophys. Res.: Atmos.* **2006**, *111*, D23.

(8) Zhang, Q.; Jimenez, J. L.; Canagaratna, M. R.; Allan, J. D.; Coe, H.; Ulbrich, I.; Alfarra, M. R.; Takami, A.; Middlebrook, A. M.; Sun, Y. L.; Dzepina, K.; Dunlea, E.; Docherty, K.; DeCarlo, P. F.; Salcedo, D.; Onasch, T.; Jayne, J. T.; Miyoshi, T.; Shimono, A.; Hatakeyama, S.; Takegawa, N.; Kondo, Y.; Schneider, J.; Drewnick, F.; Borrmann, S.; Weimer, S.; Demerjian, K.; Williams, P.; Bower, K.; Bahreini, R.; Cottrell, L.; Griffin, R. J.; Rautiainen, J.; Sun, J. Y.; Zhang, Y. M.; Worsnop, D. R. Ubiquity and Dominance of Oxygenated Species in Organic Aerosols in Anthropogenically-Influenced Northern Hemisphere Midlatitudes. *Geophys. Res. Lett.* **2007**, *34*, L13801.

(9) Chen, G.; Canonaco, F.; Tobler, A.; Aas, W.; Alastuey, A.; Allan, J.; Atabakhsh, S.; Aurela, M.; Baltensperger, U.; Bougiatioti, A.; De Brito, J. F.; Ceburnis, D.; Chazeau, B.; Chebaicheb, H.; Daellenbach, K. R.; Ehn, M.; El Haddad, I.; Eleftheriadis, K.; Favez, O.; Flentje, H.; Font, A.; Fossum, K.; Freney, E.; Gini, M.; Green, D. C.; Heikkinen, L.; Herrmann, H.; Kalogridis, A.-C.; Keernik, H.; Lhotka, R.; Lin, C.; Lunder, C.; Maasikmets, M.; Manousakas, M. I.; Marchand, N.; Marin, C.; Marmureanu, L.; Mihalopoulos, N.; Močnik, G.; Nęcki, J.; O'Dowd, C.; Ovadnevaite, J.; Peter, T.; Petit, J.-E.; Pikridas, M.; Matthew Platt, S.; Pokorná, P.; Poulain, L.; Priestman, M.; Riffault, V.; Rinaldi, M.; Różański, K.; Schwarz, J.; Sciare, J.; Simon, L.; Skiba, A.; Slowik, J. G.; Sosedova, Y.; Stavroulas, I.; Styszko, K.; Teinemaa, E.; Timonen, H.; Tremper, A.; Vasilescu, J.; Via, M.; Vodička, P.; Wiedensohler, A.; Zografou, O.; Cruz Minguillón, M.; Prévôt, A. S. H. European Aerosol Phenomenology − 8: Harmonised Source Apportionment of Organic Aerosol Using 22 Year-Long ACSM/AMS Datasets. *Environ. Int.* **2022**, *166*, No. 107325.

(10) Paatero, P. Least Squares Formulation of Robust Non-Negative Factor Analysis. *Chemom. Intell. Lab. Syst.* **1997**, *37*, 23−35.

(11) Zhang, Q.; Jimenez, J. L.; Canagaratna, M. R.; Ulbrich, I. M.; Ng, N. L.; Worsnop, D. R.; Sun, Y. Understanding Atmospheric Organic Aerosols via Factor Analysis of Aerosol Mass Spectrometry: A Review. *Anal. Bioanal. Chem.* **2011**, *401*, 3045−3067.

(12) Crippa, M.; Canonaco, F.; Lanz, V. A.; Äijälä, M.; Allan, J. D.; Carbone, S.; Capes, G.; Ceburnis, D.; Dall'Osto, M.; Day, D. A.; DeCarlo, P. F.; Ehn, M.; Eriksson, A.; Freney, E.; Hildebrandt Ruiz, L.; Hillamo, R.; Jimenez, J. L.; Junninen, H.; Kiendler-Scharr, A.; Kortelainen, A.-M.; Kulmala, M.; Laaksonen, A.; Mensah, A. A.; Mohr, C.; Nemitz, E.; O'Dowd, C.; Ovadnevaite, J.; Pandis, S. N.; Petäjä, T.; Poulain, L.; Saarikoski, S.; Sellegri, K.; Swietlicki, E.; Tiitta, P.; Worsnop, D. R.; Baltensperger, U.; Prévôt, A. S. H. Organic Aerosol Components Derived from 25 AMS Data Sets across Europe Using a Consistent ME-2 Based Source Apportionment Approach. *Atmos. Chem. Phys.* **2014**, *14*, 6159−6176.

(13) Canonaco, F.; Tobler, A.; Chen, G.; Sosedova, Y.; Slowik, J. G.; Bozzetti, C.; Daellenbach, K. R.; El Haddad, I.; Crippa, M.; Huang, R.-J.; Furger, M.; Baltensperger, U.; Prévôt, A. S. H. A New Method for Long-Term Source Apportionment with Time-Dependent Factor Profiles and Uncertainty Assessment Using SoFi Pro: Application to 1 Year of Organic Aerosol Data. *Atmos. Meas. Tech.* **2021**, *14*, 923−943.

(14) Tobler, A. K.; Skiba, A.; Canonaco, F.; Močnik, G.; Rai, P.; Chen, G.; Bartyzel, J.; Zimnoch, M.; Styszko, K.; Nęcki, J.; Furger, M.; Różański, K.; Baltensperger, U.; Slowik, J. G.; Prevot, A. S. H. Characterization of Non-Refractory (NR) PM_1 and Source Apportionment of Organic Aerosol in Kraków. *Atmos. Chem. Phys.* **2021**, *21*, 14893−14906.

(15) Chen, G.; Sosedova, Y.; Canonaco, F.; Fröhlich, R.; Tobler, A.; Vlachou, A.; Daellenbach, K. R.; Bozzetti, C.; Hueglin, C.; Graf, P.; Baltensperger, U.; Slowik, J. G.; El Haddad, I.; Prévôt, A. S. H. Time-Dependent Source Apportionment of Submicron Organic Aerosol for a Rural Site in an Alpine Valley Using a Rolling Positive Matrix Factorisation (PMF) Window. *Atmos. Chem. Phys.* **2021**, *21*, 15081−15101.

(16) Parworth, C.; Fast, J.; Mei, F.; Shippert, T.; Sivaraman, C.; Tilp, A.; Watson, T.; Zhang, Q. Long-Term Measurements of Submicrometer Aerosol Chemistry at the Southern Great Plains (SGP) Using an Aerosol Chemical Speciation Monitor (ACSM). *Atmos. Environ.* **2015**, *106*, 43−55.

(17) Stavroulas, I.; Bougiatioti, A.; Grivas, G.; Paraskevopoulou, D.; Tsagkaraki, M.; Zarmpas, P.; Liakakou, E.; Gerasopoulos, E.; Mihalopoulos, N. Sources and Processes that Control the Submicron Organic Aerosol Composition in an Urban Mediterranean Environment (Athens): A High Temporal-Resolution Chemical Composition Measurement Study. *Atmos. Chem. Phys.* **2019**, *19*, 901−919.

https://doi.org/10.1021/acs.est.2c02509
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

(18) Zhang, Y.; Favez, O.; Petit, J.-E.; Canonaco, F.; Truong, F.; Bonnaire, N.; Crenn, V.; Amodeo, T.; Prévôt, A. S. H.; Sciare, J.; Gros, V.; Albinet, A. Six-Year Source Apportionment of Submicron Organic Aerosols from near-Continuous Highly Time-Resolved Measurements at SIRTA (Paris Area, France). *Atmos. Chem. Phys.* **2019**, *19*, 14755−14776.
(19) Canonaco, F.; Crippa, M.; Slowik, J. G.; Baltensperger, U.; Prévôt, A. S. H. SoFi, an IGOR-Based Interface for the Efficient Use of the Generalized Multilinear Engine (ME-2) for the Source Apportionment: ME-2 Application to Aerosol Mass Spectrometer Data. *Atmos. Meas. Tech.* **2013**, *6*, 3649−3661.
(20) Ng, N. L.; Herndon, S. C.; Trimborn, A.; Canagaratna, M. R.; Croteau, P. L.; Onasch, T. B.; Sueper, D.; Worsnop, D. R.; Zhang, Q.; Sun, Y. L.; Jayne, J. T. An Aerosol Chemical Speciation Monitor (ACSM) for Routine Monitoring of the Composition and Mass Concentrations of Ambient Aerosol. *Aerosol Sci. Technol.* **2011**, *45*, 780−794.
(21) Fröhlich, R.; Cubison, M. J.; Slowik, J. G.; Bukowiecki, N.; Prévôt, A. S. H.; Baltensperger, U.; Schneider, J.; Kimmel, J. R.; Gonin, M.; Rohner, U.; Worsnop, D. R.; Jayne, J. T. The ToF-ACSM: A Portable Aerosol Chemical Speciation Monitor with TOFMS Detection. *Atmos. Meas. Tech.* **2013**, *6*, 3225−3241.
(22) Fröhlich, R.; Cubison, M. J.; Slowik, J. G.; Bukowiecki, N.; Canonaco, F.; Croteau, P. L.; Gysel, M.; Henne, S.; Herrmann, E.; Jayne, J. T.; Steinbacher, M.; Worsnop, D. R.; Baltensperger, U.; Prévôt, A. S. H. Fourteen Months of On-Line Measurements of the Non-Refractory Submicron Aerosol at the Jungfraujoch (3580 m a.s.l.) − Chemical Composition, Origins and Organic Aerosol Sources. *Atmos. Chem. Phys.* **2015**, *15*, 11373−11398.
(23) Paatero, P.; Tapper, U. Positive Matrix Factorization: A Non-Negative Factor Model with Optimal Utilization of Error Estimates of Data Values. *Environmetrics* **1994**, *5*, 111−126.
(24) Paatero, P.; Hopke, P. K.; Song, X.-H.; Ramadan, Z. Understanding and Controlling Rotations in Factor Analytic Models. *Chemom. Intell. Lab. Syst.* **2002**, *60*, 253−264.
(25) Middlebrook, A. M.; Bahreini, R.; Jimenez, J. L.; Canagaratna, M. R. Evaluation of Composition-Dependent Collection Efficiencies for the Aerodyne Aerosol Mass Spectrometer Using Field Data. *Aerosol Sci. Technol.* **2012**, *46*, 258−271.
(26) Paatero, P. The Multilinear Engine—A Table-Driven, Least Squares Program for Solving Multilinear Problems, Including the n -Way Parallel Factor Analysis Model. *J. Comput. Graph. Stat.* **1999**, *8*, 854−888.
(27) Paatero, P.; Hopke, P. K. Rotational Tools for Factor Analytic Models. *J. Chemom.* **2009**, *23*, 91−100.
(28) Ng, N. L.; Canagaratna, M. R.; Jimenez, J. L.; Chhabra, P. S.; Seinfeld, J. H.; Worsnop, D. R. Changes in Organic Aerosol Composition with Aging Inferred from Aerosol Mass Spectra. *Atmos. Chem. Phys.* **2011**, *11*, 6465−6474.
(29) Ulbrich, I. M.; Canagaratna, M. R.; Zhang, Q.; Worsnop, D. R.; Jimenez, J. L. Interpretation of Organic Components from Positive Matrix Factorization of Aerosol Mass Spectrometric Data. *Atmos. Chem. Phys.* **2009**, *9*, 2891−2918.
(30) Äijälä, M.; Heikkinen, L.; Fröhlich, R.; Canonaco, F.; Prévôt, A. S. H.; Junninen, H.; Petäjä, T.; Kulmala, M.; Worsnop, D.; Ehn, M. Resolving Anthropogenic Aerosol Pollution Types − Deconvolution and Exploratory Classification of Pollution Events. *Atmos. Chem. Phys.* **2017**, *17*, 3165−3197.
(31) Heikkinen, L.; Äijälä, M.; Daellenbach, K. R.; Chen, G.; Garmash, O.; Aliaga, D.; Graeffe, F.; Räty, M.; Luoma, K.; Aalto, P.; Kulmala, M.; Petäjä, T.; Worsnop, D.; Ehn, M. Eight Years of Sub-Micrometre Organic Aerosol Composition Data from the Boreal Forest Characterized Using a Machine-Learning Approach. *Atmos. Chem. Phys.* **2021**, *21*, 10081−10109.

https://doi.org/10.1021/acs.est.2c02509
Environ. Sci. Technol. XXXX, XXX, XXX−XXX