Rolling vs. *Seasonal* PMF: Real-world multi-site and synthetic dataset comparison

Marta Via[1,2], Gang Chen[3], Francesco Canonaco[3,4], Kaspar R. Daellenbach[3], Benjamin Chazeau[5], Hasna Chebaicheb[6,7], Jianhui Jiang[8], Hannes Keernik[9,10], Chunshui Lin[11], Nicolas Marchand[5], Cristina Marin[12,13], Colin O'Dowd[11], Jurgita Ovadnevaite[11], Jean-Eudes Petit[14], Michael Pikridas[15], Véronique Riffault[6], Jean Sciare[15], Jay G. Slowik[3], Leïla Simon[7,13], Jeni Vasilescu[12], Yunjiang. Zhang[7,13], Olivier Favez[7], André S. H. Prévôt[3], Andrés Alastuey[1], María Cruz Minguillón[1]

[1]Institute of Environmental Assessment and Water Research, Barcelona, 08034, Spain
[2]Department of Applied Physics, University of Barcelona, Barcelona, 08028, Spain
[3]Laboratory of Atmospheric Chemistry, Paul Scherrer Institute, CH-5232 Villigen PSI, Switzerland
[4]Datalystica Ltd., Park innovAARE, 5234 Villigen, Switzerland
[5]Aix Marseille Univ., CNRS, LCE, Marseille, France
[6]IMT Nord Europe, Institut Mines-Télécom, Univ. Lille, Centre for Energy and Environment, 59000 Lille, France
[7]Institut National de l'Environnement Industriel et des Risques, Parc Technologique ALATA, 60550, Verneuil-en-Halatte, France
[8]Shanghai Key Lab for Urban Ecological Processes and Eco-Restoration, School of Ecological and Environmental Sciences, East China Normal University, 200241 Shanghai, China
[9]Air Quality and Climate Department, Estonian Environmental Research Centre, Marja 4d, 10617 Tallinn, Estonia
[10]Department of Software Science, Tallinn University of Technology, 19086 Tallinn, Estonia
[11] School of Physics and Centre for Climate and Air Pollution Studies, Ryan Institute, National University of Ireland Galway, University Road, H91CF50 Galway, Ireland
[12]National Institute of Research and Development for Optoelectronics INOE2000, Atomistilor 409, RO77125 Magurele, Romania
[13]Department of Physics, Politehnica University of Bucharest, 313 Spl. Independentei Str., Bucuresti, Romania
[14] Laboratoire des Sciences du Climat et de l'Environnement, Orme des Merisiers, 91190 Gif-sur-Yvette, France
[15] Climate and Atmosphere Research Center, The Cyprus Institute, Nicosia, 2121, Cyprus

Correspondence to: Marta Via (marta.via@idaea.csic.es)

Abstract.

Particulate Matter (PM) has become a major concern in terms of human health and climate impact. In particular, the source apportionment (SA) of organic aerosols (OA) present in submicron particles (PM_1) has gained relevance as an atmospheric research field due to the diversity and complexity of its primary sources and secondary formation processes. Moreover, relatively simple but robust instruments such as the Aerosol Chemical Speciation Monitor (ACSM) are now widely available for the near real-time online determination of the composition of the non-refractory PM_1. One of the most used tools for SA purposes is the source-receptor Positive Matrix Factorization (PMF) model. Even though the recently developed *rolling PMF* technique has already been used for OA SA on ACSM datasets, no study has assessed its added value compared to the more common *seasonal* PMF method from a practical approach yet. In this paper, both techniques were applied to a synthetic dataset and to nine European ACSM datasets in order to spot the main output discrepancies between methods. The main advantage of the synthetic dataset approach was that the methods' outputs could be compared to the expected 'true' values, i.e. the original synthetic dataset values. This approach revealed similar apportionment results amongst methods, although the *rolling* PMF profile adaptability feature has been proven advantageous as it generated output profiles moving nearer to the truth points. Nevertheless, these results highlighted the impact of the profile anchor on the solution, as the use of a different anchor with respect to the truth led to significantly different results in both methods. In

44 the multi-site study, while differences were generally not significant when considering year-long periods, their

45 importance grew towards shorter time spans, as in intra-month or intra-day cycles. *Rolling* PMF performed better

46 than *seasonal* PMF globally for the ambient datasets investigated here as far correlation with external

47 measurements is concerned, especially in periods between seasons. The results of this multi-site comparison

48 coincide with the synthetic dataset in terms of *rolling-seasonal* similarity and *rolling* PMF reporting moderate

49 improvements. Altogether, the results of this study provide solid evidence of the robustness of both methods and

50 on the overall efficiency of the recently-proposed *rolling* PMF approach.

51 **1 Introduction**

52 Air pollution is one of the biggest current and future environmental threats to human health and climate change.

53 Results from Chen and Hoek, (2020) notably relate an increased risk for all-cause mortality due to fine aerosol

54 ($PM_{2.5}$, particulate matter of aerodynamic particle diameter below 2.5 micrometres) exposure. Also, even for

55 concentrations below the WHO guidelines threshold (annual means of 10 $\mu g \cdot m^{-3}$ for $PM_{2.5}$ at the time that article

56 was published), the life expectancy of the population of Europe has been reduced by about 8.6 months on average.

57 In turn, fine atmospheric aerosols also play a role in climate change (IPCC, 2021) due to both their direct (through

58 radiation) and indirect (through cloud interaction) effects.

59 Exposure to submicron particulate matter (PM_1, particulate matter with an aerodynamic particle diameter less than

60 1 μm) is known to have severe impacts on the respiratory system (Yang et al., 2018), and even pass the blood-

61 brain barrier to act directly on the central nervous system (Shih et al., 2018; Yin et al., 2020). Impact mitigation

62 strategies must be designed to both reduce emissions (primary aerosols) and prevent the formation of non-directly-

63 emitted (or secondary) aerosols, but also target the most harmful components, especially since recent studies have

64 demonstrated the mitigation strategies might be more effective in tackling specific PM sources rather than the bulk

65 PM (Daellenbach et al., 2020). With the purpose of identifying the most appropriate reduction strategies, source

66 apportionment (SA) methodologies, designed for identifying pollutant sources, must be constantly improved. One

67 of the most widely used receptor models for SA is the Positive Matrix Factorization (PMF) model (Paatero and

68 Tapper, 1994) along with the ME-2 engine (Paatero, 1999). This model can handle various types of data, such as

69 online and offline PM datasets (Amato et al., 2016; Crippa et al., 2014; Rai et al., 2020, respectively), VOCs (Yuan

70 et al., 2012), multi-wavelength absorption of refractory carbon (Forello et al., 2019) etc.; assemble different types

71 of pollutants (Ogulei et al., 2005); and also be coupled to machine learning techniques (Heikkinen et al., 2020;

72 Rutherford et al., 2021).

73 Since organic species account for 20–90% of the total submicron aerosol mass (Chen et al., 2022; Jimenez et al.,

74 2009) scientific interest has been set on the characterization of these pollutants by offline and online techniques.

75 The use of ACSM (Aerosol Chemical Speciation Monitor, Aerodyne Research Inc., Billerica, MA, USA) for

76 continuous monitoring and quantification of submicron non-refractory compounds has become a key approach for

77 air quality (AQ) assessment. The application of PMF to long-term ACSM submicron organic aerosol (OA) datasets

78 (Sun et al., 2018; Zhang et al., 2019) under the Source Finder (SoFi) Pro software package (Datalystica Ltd.)

79 allows to quantify and identify the contribution of major groups of organic compounds. The formerly

80 recommended methodology for OA SA was *seasonal* PMF, which requires splitting the dataset into seasons to

81 perform PMF independently, providing *seasonal* but not an intra-seasonal variation of factor profiles, as reported

82 in Canonaco et al. (2015). The more recently developed *rolling* PMF (Canonaco et al., 2021; Parworth et al., 2015)

83 applies the model on moving/rolling windows of a selected length and therefore it accounts for the temporal

84 evolution of the OA source fingerprints. The current state-of-the-art supports that *rolling* PMF should be more

85 accurate and/or suitable than *seasonal* PMF due to its profile-adaptation feature, and lower computational and

86 evaluation time, which will be the base hypothesis of this study. Nevertheless, only a few individual studies

87 (Chazeau et al., submitteed; Chen et al., 2020; Tobler et al., 2021) and an intercomparison (Chen et al., 2022) have

88 been published so far with this technique and no thorough *seasonal vs. rolling* comparison has been conducted

89 thus far to the best of our knowledge.

90 This research aims to contribute to a deeper understanding of the advantages and weaknesses of the *rolling* and

91 *seasonal* methods, assessing the differences regarding site or dataset characteristics and evaluating the

92 environmental reasonability of their outcomes. This task is of great importance as the knowledge of the strengths

93 of each method will come in handy to choose the best one for each study necessity, e.g., the better SA method for

94 specific OA sources outbreaks. Furthermore, conclusions from this analysis will also impact the quality of health,

95 climate and modelling studies by means of an improved description of the main OA pollution sources.

96 **2 Methodology**

97 **2.1 Instrumentation and datasets.**

98 This study is one of the outcomes of the Chemical On-Line cOmpoSition and Source Apportionment of fine

99 aerosoL (COLOSSAL) project (https://www.costcolossal.eu/) supported by the COST programme, and based on

100 measurements performed within the ACTRIS network. It is closely related to the overview study of Chen et al.

101 (2022), in which 22 more than one-year-long PMF datasets were joined for a *rolling* PMF intercomparison.

102 Participants of the WG2 of the COST COLOSSAL Action contributed to the preparation of a protocol for SA with

103 the purpose of homogenization of the PMF application (Chen et al., 2022). Nine of the 22 datasets from that study,

104 whose main characteristics can be found in Table 1, were also provided for this *rolling-seasonal* comparison. Some

105 of them contain site-specific sources related to instrument artefacts or proximity to pollution hotspots. The factors

106 identified at all sites are Hydrocarbon-like OA (HOA), Biomass Burning OA (BBOA, except for the Dublin site),

107 Less-Oxidised Oxygenated OA (LO-OOA), More-Oxidised Oxygenated OA (MO-OOA) and Oxygenated OA

108 (OOA), which represents the sum of LO-OOA and MO-OOA. Other factors are only present at one or two sites:

109 Cooking-Like OA (COA; in Barcelona-Palau Reial and Marseille-Longchamp), 58-related OA (58-OA; in

110 Magadino), Shipping and Industry OA (SHINDOA; in Marseille-Longchamp), Wood Combustion, Coal

111 Combustion and Peat Combustion OA (WCOA, CCOA, PCOA respectively; in Dublin). The 58-related OA, as

112 explained in Chen et al. (2021), is a factor dominated by nitrogen fragments (m/z 58, m/z 84, m/z 94) which

113 appeared as an artefact after the filament replacement in that instrument.

114 All data presented in the multi-site intercomparison were obtained from ACSMs, which use a mass spectrometer

115 to measure the composition of non-refractory submicron particulate matter (NR-PM$_1$) in near real-time. It works

at a lower mass-to-charge resolution but it is more robust compared to the Aerosol Mass Spectrometer (AMS, Aerodyne Research Inc, Billerica, MA, USA) allowing for long-term deployment. Both quadrupole (Q-ACSM) and Time-of-Flight (ToF-ACSM) ACSMs were used, further described respectively in Fröhlich et al. (2013), Ng et al. (2011). The resolution of ToF-ACSM datasets (10 minutes) was averaged to 30 minutes (resolution of the Q-ACSM) to have harmonised timestamps. The analysis software (version 1.6.1.1 for Q-ACSM and version 2.3.9 for ToF-ACSM), implemented in Igor Pro (WaveMetrics, Inc.), was provided by Aerodyne Research Inc. The treatment of the multi-site ACSM data to generate PMF input matrices is summarised in SI Table S1 and more details can be found in the publications cited therein.

Ancillary measurements consisted of i. SO_4^{2-}, NO_3^-, NH_4^+ and Cl^- measurements from ACSM; ii. Black Carbon (BC) from filter-based absorption photometer AE33 from Magee Scientific (Drinovec et al., 2015), except for Cyprus Atmospheric Observatory – Agya Marina and Magadino, in which the AE31 was used. BC concentrations were differentiated according to their main sources, into fossil fuel (BC_{ff}) and wood burning (BC_{wb}) BC by applying the Sandradewi model (Sandradewi et al., 2008); iii. NO_X concentrations; iv. Ultra-fine particles (range 20-1000 nm) at the Marseille - Longchamp site. Details on the complementary instrumentation at each site can be found in SI Table S2.

2.2 Synthetic dataset

Although the principal aim of this article is to inspect the differences in the methods amongst these European sites, firstly a synthetic dataset comparison was tackled. The main advantage of this procedure is that it allows mimicking real-world environmental measurements already classified in OA sources so that PMF results can be compared with the incoming synthetic data. We created a synthetic dataset that mimics OA mass spectral analyses of a ToF-ACSM in Zurich. For that purpose, we used source-specific OA mass spectra retrieved from the AMS Spectral database (Crippa et al., 2013; Ng et al., 2011b; Ulbrich et al., 2009) and OA source concentration time series generated by the air quality model CAMx (Comprehensive Air Quality Model with Extensions) previously published by Jiang et al. (2019). Details are described in the SI section A. The dataset used to calculate the error matrix is that from the Zurich site which ranges from February 2011 until December 2011. Hence, the same CAMx outcoming time series period was used to generate the concentration matrix. The represented OA sources are HOA, BBOA, SOA from biogenic emissions (SOA_{bio}), SOA from biomass burning (SOA_{bb}) and SOA from traffic and other anthropogenic sources (SOA_{tr}).

The first step for the synthetic dataset creation was to select p (number of factors), POA and SOA spectral profiles from the High-Resolution AMS Spectral database (Crippa et al., 2013; Ng et al., 2010; Ulbrich et al., 2009) and multiply them by the time series of the same sources from the model output. The error matrix was generated following the same steps as for real-world data and real-world parameters were used as detailed in SI. For this purpose, the dataset used is that from the Zurich site which ranges from February 2011 until December 2011. Hence, the same CAMx outcoming time series period was used to generate the concentration matrix. Gaussian noise was subsequently added to the outcoming matrix. The resulting matrices were used as *rolling* and *seasonal* PMF input. Before the comparison to original factors, several tests as in the multi-site comparison have been performed to check the quality of the output, such as the mass closure test or the scaled residuals profile revision.

2.3 Positive Matrix Factorization

The Positive Matrix Factorization model (Paatero and Tapper, 1994) describes the measured matrix $\mathbf{X}$ of n timestamps and m variables as a product of two matrices, $\mathbf{G}$ and $\mathbf{F}$, plus a residual matrix $\mathbf{E}$ for a given number of factors p:

$$x_{ij} = \sum_{i,j}^{n,m} g_{ik} \cdot f_{kj} + e_{ij} \tag{1}$$

The matrices $\mathbf{G}$ and $\mathbf{F}$ can be randomly initialised with *a priori* information. The model then iterates until the quantity

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

where σ_{ij} represents the uncertainties of the input matrix $\mathbf{X}$, is minimised with respect to all model variables.

The use of *a priori* information reduces the rotational ambiguity of the model, consisting of a degeneration of solutions associated with a given Q value (Canonaco et al., 2013), and it is usually done from the *a*-value approach. This consists of initialising $\mathbf{F}$ (or $\mathbf{G}$) with reference profiles (or time series) and multiplying them by the percentage of variation a, $a \in [0,1]$, where 0 and 1 would represent total constraint and freedom, respectively. The Source Finder (SoFi Pro, versions 6.8 and 8.04, Datalystica Ltd., Villigen, Switzerland) applies this algorithm through the Multi-linear Engine 2 (ME-2) (Paatero, 1999) within the Igor Pro software environment (Wavemetrics, Inc., Portland, OR, USA). SoFi is also a powerful software package for preparing the rolling conditions for the input matrices prior to the PMF algorithm and post-processing the outcomes afterwards.

2.3.1 *Seasonal* PMF

In order to apply *seasonal* PMF, the input matrix is divided into season-long submatrices and PMF is applied independently, adjusting the number of necessary factors to the requirements of each subperiod. In order to reach an environmentally-reasonable local Q minimum, the implementation of constraints on Primary Organic Aerosol factors (POA), has been performed according to the COLOSSAL guidelines for source apportionment (COLOSSAL, COST Action CA16109, 2019) and the protocol from Chen et al. (2022). After unconstrained results exploration, which allowed for some marker identification, constraints based on the a-value approach were applied to primary OA factors. The systematic exploration of the a-value space has been performed for each season with the aim of determining the combination of a-values which maximises the correlations between factors and external correlations and represents an environmentally-reasonable OA explanation, hereafter referred to as the base case solution. The random a-value ranges and the reference profiles employed can be found in SI Table S1 and Table S3 a.

With respect to the synthetic dataset, the 11 months from 2011 data were split into three periods (and not four seasons to avoid running PMF over too short periods): February – May, June – August and September – December. The real-world Marseille dataset also used the co-located SO_2 time series to force an Industry + Shipping factor to

185   emerge as reported in previous studies (Bozzetti et al., 2017; El Haddad et al., 2013). The seasonal averaging of

186   the remaining runs were complemented by bootstrapping to estimate the statistical error of the solution.

187   Bootstrapping (Efron, 1979) the *seasonal* PMF input together with random *a*-value resampling allows for

188   statistical and rotational uncertainty assessment. The application of criteria-based selection, which will be deeply

189   explained in 2.3.2, was also used to discard those runs which do not comply with the user-defined standards. The

190   outcome of this technique consists of *p* (number of factors) mass spectra and time series including their uncertainty

191   assessment combined season-wise together to obtain period-long OA sources. This might lead to possible factor

192   discontinuity. Moreover, from this approach, source fingerprints are static throughout a whole season and cannot

193   adapt to OA processes of lifetimes below a meteorological season (i.e., 90 days), but can nevertheless evolve from

194   one season to another.

195   **2.3.2 *Rolling* PMF**

196   *Rolling* PMF runs the model on subsets of the input matrix with a user-defined (window) length in days. Then, the

197   window is shifted by a number of days (also chosen by the user) and PMF is applied again (Parworth et al., 2015).

198   Consequently, many PMF runs are performed in each window length period, so in the post-analysis, one can

199   automatically discard the runs that do not meet certain user-defined criteria (Canonaco et al., 2015). To select the

200   most environmentally-reasonable runs, the remaining solutions are averaged to generate the final solution, which

201   will be provided with statistical and rotational uncertainties based on random input resampling (bootstrap) and

202   random a–value resampling, respectively.

203   A length of 14 days and a shift of 1 day were used in the current study for the synthetic dataset and for 7 out of the

204   9 datasets, which is a good compromise between Q/Q_{exp} values and the percentage of modelled points as suggested

205   in Canonaco et al. (2021). Window lengths of 28 days were also assessed, but the correlations to ancillary

206   measurements deteriorated in most of the cases. Exceptions to this rule were the SIRTA and Tartu sites, for which

207   the 28-day window offered better correlations. These window lengths are consistent with the life cycle lengths of

208   atmospheric aerosols (Textor et al., 2006) and their outcomes do not differ significantly. The application of

209   constraints in PMF, as advised in the protocols, consists of setting random *a*-values within a reasonable range and

210   accepting only the runs that comply with the criteria. This procedure will lead to the selection of *a*-values which

211   induce more environmentally-reasonable solutions, whose average will provide the final number. In some cases,

212   the reference profiles used in *rolling* PMF are those from the *seasonal* solution, as the protocol is flexible regarding

213   this choice. However, this constraint can have an impact on the solution, and in order to identify its implications,

214   the profiles used in each case are detailed in Table S1, Table S3a.

215   A criteria-based selection was developed to automatically inspect the large number of PMF runs provided by the

216   *rolling* method (Canonaco et al., 2021). This consists of the application of certain criteria to be fulfilled by the

217   PMF outcoming factors. The acceptance/rejection of a run can be dictated by the thresholds retrieved from

218   bootstrapped *seasonal* solutions, or more advisably, from a double-tailed Welch's t-test hypothesis evaluation with

219   *p*-values (Chen et al., 2021) chosen by the user (not exceeding 0.05). This procedure allows for factor discontinuity,

220   as one can run PMF for two consecutive numbers of factors, and choose a certain criterion upon which to select

221   one more (or less) factor depending on the outbreak/vanishing of a factor marker. The list of criteria is specified

222   in SI Table S3 b for the synthetic dataset and in the respective publications for each real-world site.

2.4 SA procedure and dataset homogenization

224   A method to compare source apportionment performance, analogous to Belis et al. (2015) while adjusted for our

225   specificities, was developed in this study. The first step consisted of preliminary checks, in which the minimum

226   requirements for solution acceptance must be satisfied, such as the mass closure and reasonability of profiles.

227   Secondly, the characterization of discrepancies between methods was addressed in order to confirm the presence

228   or absence of significant differences between *rolling* and *seasonal* PMF. The decision on which method is more

229   suitable for certain dataset particularities was *a posteriori* based on the quantification of the performance goodness

230   of both methods by means of correlation to external measurements and residual analysis. This flow process was

231   applied to both the multi-site analysis and the synthetic dataset.

232   All the participants of the multi-site comparison applied the SA protocol to their own datasets, benefiting from the

233   expertise in the previous OA SA studies at their sites. The analysis of differences between source contribution

234   estimates by both methods was performed for each site individually and overall. The similarity of time series from

235   one method to the other was assessed not only for each whole dataset but also in a 'rolling' fashion, that is to say,

236   by calculating some metrics on windows of a given number of days with 1-day shifts between windows. This

237   approach allowed for the identification of significant discrepancies between both approaches for the set PMF

238   window lengths (14 days for *rolling*, 90 days for *seasonal*), a feature which was not evident in the whole long-

239   term time series. It also enabled watching intra-daily differences, by setting period lengths of 1 day.

240   A detailed study of model residuals was also beneficial to quantify the accuracy of each technique performance.

241   Scaled residuals represent the model error (e_{ij}) normalised by the uncertainty matrix (σ_{ij}):

242    $$Scaled\ residuals_{ij} \ = \ \frac{e_{ij}}{\sigma_{ij}} \tag{3}$$

243   and their i,j sum has been reported in Paatero and Hopke (2003) to describe a unimodal histogram within a ±3

244   range under good model performance conditions. The output Q quantity has both been compared in a raw and in

245   a normalised way. This normalisation aims to deprive the impact of the degrees of freedom that normally depend

246   on the input size and on the number of factors, computing hence the quantity Q/Q_{exp}, where

247    $$Q_{exp} = m \cdot n - p \cdot (m + n) \tag{4}$$

248   Then, various PMF runs can be compared in a more fundamental way. The expressions used for the normalisation

249   arrangement have to be adapted to the particular degrees of freedom of each method:

250    $$\left(\frac{Q}{Q_{exp}}\right)_{Rolling} = \frac{Q}{m \cdot n - p \cdot (n + m \cdot n_{14})} \tag{5}$$

251    $$\left(\frac{Q}{Q_{exp}}\right)_{Seasonal} = \frac{Q}{m \cdot n - p \cdot (n + m \cdot n_{90})} \tag{6}$$

252 The parameters $n_{14\ days}$ and $n_{90\ days}$ refer to the number of periods throughout the dataset of 14 and 90 days,

253 respectively.

254   For the synthetic dataset, the comparison between methods had to consider the error of each. For this purpose, the

255 metric presented in Belis et al. (2015), the uncertainty-normalised Root Mean Squared Error (RMSE$_u$) was used:

256
$$RMSE_u = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(m_i - r_i)^2}}{2u} \qquad (7)$$

257 In this expression, *m* represents the modelled values, *r* the reference values and *u* the mean uncertainty of the

258 model.

259 **3 Results and discussion**

260 **3.1 Synthetic dataset**

261 This section aims to assess the quality of the outcome of the *rolling* and *seasonal* PMF methods. Relying on

262 synthetic ToF-ACSM data offers the opportunity to compare the PMF outputs to the truth which is not available

263 for real-world measurements. We focus here on the OA sources' (factors') mean concentration and their temporal

264 variability as well as the mean chemical composition and their temporal variability.

265 Regarding the OA apportionment vs. the input OA scatterplot, Table S4 presents the fitting coefficients for several

266 resolutions with no substantial difference between methods. Figure 1 shows the relative factor contributions to the

267 apportioned OA for both methods. The POA factors do not differ substantially between the SA methods, but they

268 are underestimated with respect to the truth, (25% of OA in *rolling* and *seasonal*, 35% in truth). Also, whilst the

269 LO-OOA-to-MO-OOA ratio is nearly one in the *rolling* case, it presents a much fresher secondary aerosol for the

270 *seasonal* (1.5). Compared to the *truth*, PMF using a priori information on POA's chemical composition (HOA,

271 BBOA) underestimates POA and overestimates SOA.

272 Figure 2 presents the time series, diel cycles of the truth, and *rolling* and *seasonal* methods, as well as the, scatter

273 plots between the corresponding PMF time series. In time series and diel plots, it is noticeable that SOA is

274 overestimated by PMF at the expense of POA (Fig. 2(c)). Squared Pearson correlation coefficients and slopes were

275 similar for both rolling and seasonal, respectively, for HOA (0.89,0.88) and OOA (0.95, 0.97), but not for BBOA

276 which *seasonal* resolves better (0.55,0.72). Welch's t-tests between *rolling* and *seasonal* time series rejected the

277 similarity of all factors' concentrations. This test applied to both methods against the truth also rejected the

278 hypothesis of significantly similar means, discarding a good method representation of truth results. This could be

279 explained by the fact that truth profiles are static and the methods were trying to adjust to moving fingerprints and

280 the anchor profiles might have influenced the results. The uncertainty-biased RMSE (RMSEu, equation (7)) values

281 are 1.10, 0.90 for HOA, 0.95, 1.98 for BBOA and 0.05 and 0.33 for SOA, respectively for *rolling* and *seasonal*

282 PMF results. Values under one represent values within the range of PMF uncertainty, and therefore acceptable

283 values, all in this case except for *rolling* HOA and *seasonal* BBOA. These exclusions could be explained by two

non-exclusive hypotheses: i. The dissimilarity between methods and truth is large; ii. The uncertainties of the methods might be underestimated. In all cases except for HOA, *seasonal* presents higher RMSEu and therefore a worse fit to the truth. Besides, the statistical Welch's t-test was performed on the synthetic dataset PMF results, testing the null hypothesis of statistically similar means with different variances.

The difference in Pearson squared correlation coefficients between factors and their potential markers is shown as a histogram in Figure S1 for each of the methods. The truth results show the worst correlation with ancillary measurements compared to modelled PMF. *Rolling* and *seasonal* results are very similar, although these correlations seem greater for *rolling* POA factors and for *seasonal* SOA factors. Slightly higher correlation coefficients were found for *rolling* in transition periods (i.e., ±7 days before/after the change of seasons) for the *rolling*: 0.88 and 0.77 for HOA vs. BC, 0.74 and 0.65 for HOA vs. NO_X, 0.52 and 0.52 for OOA vs. NH_4 and 0.07 and 0.07 for MO-OOA vs. SO_4.

Profiles (Figure 3(a)) did not show remarkable discrepancies between PMF methods, but these could be nonetheless noticed when compared to the truth profiles. The cosine similarity method, though, revealed high similarity from both methods to synthetic profiles (1.00 and 1.00 for HOA, 0.91 and 0.91 for BBOA and 1.00 and 1.00 for SOA for *rolling* and *seasonal* PMF, respectively). However, it is noteworthy that both HOA's and BBOA's chemical compositions were constrained. Model HOA profiles were very similar to the truth, except for the lower m/z 44 and higher m/z 57 of truth, and other HOA markers regarding models. Modelled BBOA presented significant differences from the truth to modelled profiles, the truth profiles contained a lower m/z 44-to-m/z 43 ratio, the lower influence of HOA markers and a much higher m/z 60 and m/z 73, BBOA tracers. Hence, modelled BBOA contained a higher proportion of other OA factor markers and lower of their own, meaning modelled profiles might have resulted in less clean than the true ones. SOA PMF modelled profiles contained lower m/z 43 and m/z 44 than the truth profiles, although the rest remains very alike. In short, PMF results present a BBOA factor with more SOA and HOA influence as the main profile. The underestimation of POA is therefore understood due to a poorer modelization of the key source identifiers, leading to a less pure profile and hence, a lower mass apportionment compared to truth.

The influence of reference profile constraints might have enhanced the misattribution of the profiles, for example, imposing m/z44-to-m/z43 ratios has led to a significant difference in the degree of oxidation solution with respect to truth. Nevertheless, constraining profiles has provided more accurate solutions than unconstrained set-ups, as shown in Figure S2. These plots show how *seasonal* constrained PMF launches always present higher similarity to *truth* in terms of key ions ratios. Moreover, OA sources of unanchored runs were less robust due to lower reproducibility along accumulation of runs. By extension, rolling results are expected to reproduce the same results as it has been proven that both techniques' outcomes converge sufficiently.

The adaptability of the models can be assessed from Figure 3 (b), where the 60/55 vs. 44/43 (which are proxies for the BBOA-HOA differentiation and the SOA oxidation, respectively) is plotted for the truth and both methods. Here we use m/z 55 since it is known to be a key marker for HOA. *Rolling* is shown to be as a continuous time series, as the profiles for this method are time-dependent, whilst these ratios for *seasonal* only vary from season to season. In HOA, the modelled points circle the actual truth and anchor profile points (which are similar or

321 equal), but this is not the case for BBOA, in which the *rolling* and *seasonal* points are near the anchor profile but

322 distant from the truth. This implies that the anchor profile, which was selected ignoring the truth profile

323 characteristics, plays an important role in terms of adaptability to the actual solution. Overall, even for OA sources

324 with nominally constant chemical composition (here HOA, BBOA), the factors resolved by PMF exhibit a varying

325 chemical composition. Therefore, caution is required in interpreting the variability in sources of chemical

326 variability resolved by rolling PMF. Oppositely, the SOA profiles, as they were unconstrained, can be compared

327 more fairly. Both the *rolling* and *seasonal* dots are within the truth markers except for some points of high 60/55

328 for which the highest disparity to truth is found for *seasonal*. This is suggesting a poorer PMF OOAs chemical

329 composition profile apportionment, which in turn, might be influenced by the POA anchoring deficiencies.

330 The benefits of the continuity of the *rolling* profiles are reflected in time series as can be seen in Figure 3 (c), in

331 which the behaviour of *seasonal* points is unrealistically drastic depending on the season. The profile adaptability

332 of the *rolling* method represents a more resolute approach to positively representing the truth. Contrarily, the

333 *seasonal* approach, although can be plotted for each timestamp as SOA is the sum of two OOAs, can only vary in

334 lines of equal 44/43 ratio as the profiles are constant all over a season. In short, it can be stated that as opposed to

335 *seasonal* PMF or in general batch-wise PMF analysis, *rolling* PMF offers the potential to interpret changes (e.g.

336 seasonal) in OA sources' chemical composition, but the anchor profile selection has been shown to generate

337 significant discrepancies to the truth for both methods requiring caution in interpreting such variability.

338 The SA method used has a severe impact on model scaled residuals. Figure S3(a) shows the histogram of the

339 scaled residuals for all the resolutions. In all cases, the *rolling* PMF histogram is significantly sharper, more centred

340 to zero. Also, the same effect is visible in the transition periods (Figure S3(b)). Regarding Q values, the *rolling*

341 value (3838356) is lower than the *seasonal* value (24665377), as expected due to a higher extent of degrees of

342 freedom of the former method. Q/Q_{exp} values, computed from equations (5) and (6), are 7.08 and 37.58,

343 respectively for *rolling* and *seasonal* PMF. The fact that normalising by the model-specific degrees of freedom the

344 Q/Qexp is lower for *rolling* than for seasonal leads to the conclusion that the minimization of uncertainty-weighted

345 errors is better achieved by the *rolling* method.

3.2 Multi-site comparison

3.2.1 Preliminary tests

348 Preliminary tests were performed to check the consistency of the reported results as well as the actual difference

349 between the methods reported. An important performance metric is the closure of the OA mass, that is to say, the

350 difference between the sum of all OA factors concentrations vs. the input mass. Table S4 provides the fit statistics

351 of the input OA vs. the outcome OA for all the sites and four different time spans for the whole period, a season,

352 a fortnight and a day. All squared correlation coefficients are higher than 0.88, and slopes are within the 0.92-1.09

353 range. This ensures the quality of the PMF performance at all time resolutions and for both methods. A closer

354 inspection of the table shows slightly higher correlation coefficients and slopes closer to one for *rolling*.

In order to confirm or reject the existence of systematic disparity between both methods, a two-tailed Welch's t-test was performed under the null hypothesis of the time series having statistically similar expected values. In Table S5, all cells marked represent the runs that reject the null hypothesis i.e., for which the factors retrieved from *rolling* and *seasonal* are not statistically equal (p-values over 0.05). The row 'all' refers to the concatenation of all the dataset time series. Apportioned OA presents the highest acceptance of the hypothesis rate, implying that the global apportionment means are not significantly different. The factors with higher rejection rates are LO-OOA, MO-OOA and HOA and in this order. OOA factors, as they are unconstrained, might be rather sensitive to source outbreaks or variations, which can have been caught or not by one model, although their sum (OOA) remains coherent. Period-long figures get the highest rejection rates, which decay rapidly from lower to higher resolutions, meaning the seasonal and fortnightly averages are still high despite their rapid resolution while for the daily resolution this rate is very low. This fact highlights that the methods yet present significant differences in means in intermediate resolutions.

Figure S4 compares the relative difference of the *rolling* minus *seasonal* concentrations for each factor in the all-sites ensemble. The factors with higher errors are MO-OOA, LO-OOA, tilted to positive values, this is, resulting in higher concentrations for *rolling*. This is probably related to the lack of anchors which promotes higher freedom and hence, higher difference between methods. Also, BBOA presents significant positive whiskers, but as mean concentrations in Figure 1 are equal, we suspect these are linked to sporadic high concentration outbreaks, which might only have been caught by the *rolling* method. Besides, the other factors are not significantly different from zero.

The pie charts in Figure 4 show the amount of mass apportioned by the main OA sources in all datasets. These pies do not account for site-specific sources, they present the relative contribution of the all-sites ones scaled to account for the 360 degrees. OA is mainly driven by secondary organic aerosols in both cases, although the ratio of fresh-to-aged aerosol contribution, that is, LO-OOA over MO-OOA, is much higher for *rolling* (0.62) rather than for *seasonal* (0.54). The ratio POA over SOA is higher for *seasonal* than for *rolling* (0.58 and 0.37, respectively), and the ratio BBOA over HOA is considerably different (1.17 and 1.45, respectively). The fact that wood burning exceeds fossil fuel is consistent with the average ratio for BC_{wb} and BC_{ff} of 3.1, implying that PMF reproduces this relation. Hence, *rolling* describes a more oxidised SOA, which is less prevailing than that from *seasonal*. In both, POA is governed by the biomass burning OA. Figure S5 shows the individual apportionment pies, in which the same trends can be generally recognized. In general terms, these results do not coincide with the synthetic dataset ones, in which the POA/SOA and LO-OOA / MO-OOA ratios were, respectively, equal and higher for *rolling*.

Figure 5 shows both the monthly and diel cycles of *rolling* minus *seasonal* concentrations of the ensemble of sites for the main factors. In general terms, the intra-year variation is not remarkable as mainly all the boxes are crossing the zero line. Besides, the differences between mean and median could indicate the adaptability to spiky events. The fact that HOA and BBOA are remarkably different throughout the whole period coincides with Welch's t-tests aforementioned. Moreover, the mean for HOA in January and December and for BBOA in July and August are positively set beyond the boxes, which could imply that the most extreme events are better captured by the *rolling*. This fact reinforces the hypothesis of a more precise capture of intra-month events. SOA factors present fewer

393 clear trends, although an alternate sign between warm and cold months can be recognised. Figure S6 depicts the

394 behaviour of the remaining factors, which are nearly zero except for 58-OA, which is significantly negative in

395 summer and SHINDOA, which alternates from positive to negative from summer to winter. In the case of 58-OA,

396 this indicates a summertime under/overestimation of one of the methods, and for the SHINDOA, a differing

397 capturability of events along the year.

398 Regarding diel cycles (Figure 5), the differences are evident in HOA and BBOA at night, implying there is where

399 the mixing between POA sources is aggravated. The SOA factors reveal one of the methods overestimates the

400 other throughout the daily cycle: *rolling* is greater for MO-OOA and OOA and lower for LO-OOA. Whilst these

401 differences do not have an impact on the Welch's t-test for LO-OOA, they do for the rest, even for HOA for which

402 they are not very uneven, probably due to compensation of differences while averaging. Figure S6 shows similar

403 behaviour for both methods in all the factors except for the WCOA, PCOA, CCOA, which present higher

404 differences at night. *Seasonal* concentrations, though, are remarkably higher for 58-OA throughout the period/daily

405 cycle, a bit superior in the COA 8h peak and inferior in the SHINDOA afternoon. While these results do not have

406 an impact on the p-value for the Barcelona - Palau Reial and Magadino sites, it does for the Marseille-Longchamp

407 site.

408 **3.2.2 PMF goodness evaluation**

409 3.2.2.1 Correlation with ancillary measurements

410 In order to assess the quality of each PMF method outcome, the correlation of factors with their potential markers

411 was monitored from a single and global perspective. The pairs of variables compared were: HOA and BC_{ff}, HOA

412 and NOx, BBOA and BC_{wb}, MO-OOA and SO_4^{2-}, OOA and NH_4^+. SHINDOA was compared to ultrafine particles

413 of diameters 10 - 20 nm coming from shipping/industry, differentiated according to Chazeau et al. (2021),

414 Rodríguez and Cuevas, (2007). The correlation of LO-OOA vs. NO_3^- has been excluded in this study due to the

415 plentiful sources of NO_3-; besides organonitrates would hamper the traceability of LO-OOA from this compound.

416 This analysis has not been extended for the rest of the OA sources due to the lack of appropriate tracers available.

417 Figure 6 presents the Pearson-squared correlation coefficient for all the pairs of markers and factors retrieved from

418 *rolling* and *seasonal* PMF. Even though these marker time series are not deprived of errors, the hypothesis is that

419 better agreement leads to better adaptation of the model to the OA source emitting these tracers. Overall, the *rolling*

420 boxes are centred to higher correlation values than the *seasonal* ones but their whiskers always reach the maximum

421 value of one in both cases. The difference between methods is small since medians do not differ more than 0.05,

422 however, the *seasonal* performance is slightly underscoring these correlations. This finding would support the

423 hypothesis of the superior performance of the *rolling*, although HOA is evenly characterised in both methods,

424 which is consistent with the great similarity in the apportionment of OA shown in Figure 4. The histogram for the

425 difference of Pearson-squared correlation coefficients is plotted as a histogram for all sites in Figure S7. Positivity

426 in this graph reflects better *rolling* results matched with co-located measurements, and the histogram spreads the

427 range of correlations. The amount of shoulders in the right half of these histograms is higher than those in the left,

428   which implies systematic improvement of the *rolling* method with respect to *seasonal* in terms of correlation with
429   ancillary measurements.

430   Periods of transition from one season to another are strategically relevant for this comparison since the *seasonal*
431   method, due to its profile staticity definition, could yield to discontinuities in the time series of the different
432   components. The change of OA factors spectra for *rolling* is smooth, therefore no abrupt changes should be
433   expected in the season edges. Moreover, the *rolling* technique is capable of introducing factors depending on
434   criteria compliance, therefore, their concentration edges are not as sharp as they would be for the *seasonal* PMF.
435   This is the case for BBOA appearance in the cold months in Barcelona - Palau Reial and Marseille - Longchamp,
436   and the 58-OA outbreak after Q-ACSM filament replacement in Magadino. From these premises, one could expect
437   to find better correlation coefficients relating to factors and their markers for the *rolling* method, which could
438   better represent these periods. Table S6 shows the correlation of the OA factors and their markers only for these
439   periods both for the *rolling* and *seasonal* PMF. In all cases, the differences between methods are not extensive.
440   However, it can be seen how the 'whole' dataset figures are always greater for *rolling* rather than for *seasonal*.
441   This finding supports the conjecture that the *seasonal* method presents greater difficulty representing the edges of
442   the seasons. The relevance of this conclusion is to be considered especially in the datasets in which the number of
443   days near season changes is important due to data gaps.

444   3.2.2.2 Model residuals

445   Figure 7 shows the normalised scaled residuals distribution for both methods in a concatenated dataset including
446   all the sites. Given that the uncertainty matrix was the same for both techniques, scaled residuals reflect the
447   capacity of each technique to apportion the most similar quantity of OA as was entered as input. Boxplots show a
448   tendency towards negative values for both methods, implying a systematic bias towards the overestimation of the
449   input matrices. *Seasonal* errors present a higher spread and lower mean and medians; hence, *seasonal* results are
450   less accurate and precise than those from *rolling*, overall. However, the span of both distributions does not exceed
451   the $\pm 3\sigma$ threshold in any of the cases, meaning the results are acceptable for both techniques. Figure S8 shows the
452   same plot for each of the participant sites. In general, the *rolling* histograms are more centred to zero and their
453   sharpness is higher with respect to *seasonal* distributions. An exception to this behaviour is Marseille –
454   Longchamp, which presents negatively-shifted distributions probably related to the model's difficulty to
455   differentiate between BBOA and LO-OOA.

456   Scaled residuals for the season transition periods are presented in Figure S9. Both histograms extend largely
457   beyond the (-3, 3) domain, implying both methods struggle in this kind of period, but the *seasonal* distribution of
458   scaled residuals is much wider than that of *rolling*. Also, in the zoomed (-3, 3) range, *seasonal* results seem to
459   present a wider distribution. Distribution shoulders are present in both, negative in *rolling* and positive in *seasonal*,
460   respectively, indicating *rolling* overestimation and *seasonal* underestimation of input concentrations. These
461   findings would imply that even if the methods provide a substantial error in the transition periods, the *rolling* better
462   captures the season change due to its profile adaptability.

463   Regarding Q values, the differences between techniques are presented in Table 2. Unweighted Q values show a
464   clear pattern on lower values for *rolling* PMF except for one site. The SIRTA datasets were treated by two different

465 users, which might have led to different PMF steps and unreliable result. The generally greater minimisation of Q
466 performed by the *rolling* PMF method can be explained by the major quantity of runs performed compared to
467 *seasonal* PMF because of the proper definition of the method. Depriving the Q of the degrees of freedom effect as
468 shown in equations (5), (6), the minimisation of both methods is signified. The trend generally points to lower
469 figures for the *rolling* method, but whilst the Q minimization of the unweighted Q was an expected fact, the implicit
470 error reduction cannot be ensured within a theoretical frame. However, the majority of sites (excluding the
471 aforementioned SIRTA) show lower Q/Q_{exp} values for the *rolling* method.

472 3.2.2.3 Adaptability tests

473 Adaptation tests were designed to inspect how much the methods comply with the input data. One of the main
474 concerns to assess is the adaptability of the output profiles to short-lifetime events (order of magnitude of days),
475 as it is the hypothesis onto which the *rolling* PMF is based. For this purpose, the check was based on the difference
476 between main ion ratios, calculated from input values and the apportioned amounts of these ions by OOA factor
477 profiles for both methods, e.g. $(m/z44 / m/z43)_{input}$ – OOA $(m/z44 / m/z43)$Rolling or Seasonal. This can be seen
478 in Figure S10 in a time-series form for each site. Because m/z44 and m/z43 are also part of POA profiles, one
479 should not expect to find there a perfect match between the raw and the OOA profiles ratios, but a qualitative idea
480 of how well the profiles adapt to the degree of aerosol ageing. In these plots, generally the *rolling* profile variation
481 seems to adapt better rather than *seasonal*, which is a straight line along a season. Under the same logic, Figure 8
482 shows the site histograms of the $(m/z44 / m/z43)_{input}$ – OOA $(m/z44 / m/z43)$Rolling or Seasonal values only for
483 periods around the change of season, which have been proven to be tricky for the PMF model.

484 Figure S9 shows in general terms how the *rolling* adapts to the main 44-to-43 trends, whilst *seasonal* can only
485 present a single value for a whole season. Even though the *rolling*/*seasonal* SOA and the input time series are not
486 expected to match perfectly, the main features of the variability are usually caught by the *rolling*. By taking a look
487 only at the transition periods in Figure 8, the tendency is that the difference between the input ratio and the *rolling*
488 ratio is closer to zero or sharper around it than with *seasonal*. These qualitative appreciations bolster the
489 aforementioned conclusion that *rolling* is adapting the SOA profiles to specific singularities of the input time
490 series, generating thus a more accurate solution.

491 **4 Conclusions**

492 The present study aimed at performing a comprehensive comparison between the two methodologies of fine
493 organic aerosol (OA) source apportionment through the Positive Matrix Factorization (PMF) model: *rolling* and
494 *seasonal* PMF.

495 The synthetic dataset *rolling* and *seasonal* outputs assessment has been rather fruitful for this comparison. The
496 main highlight of this approach is that the modelled sources could be compared to the 'truth' ones, this is the
497 artificially chosen OA sources during the dataset tailoring. Contrasting PMF results against the *truth* highlighted
498 the model overestimation of SOA and underestimation of POA (here in case of using a priori information on
499 POA's chemical composition) for both *rolling* and *seasonal* and different degrees of SOA oxidation between

methods. Nevertheless, the correlation of *rolling* and *seasonal* with the *truth* time series and profiles show very similar results in terms of concentrations. The temporal variability of OA sources' chemical composition has been shown to oscillate even for POA components with temporally invariant chemical composition and to be severely impacted by the selection of the profile anchors, as it differed significantly from the truth results when the anchor was significantly different to the truth profile. However, the use of profile constraints still provided solution closer to the truth than unconstrained PMF. Besides, the *rolling* method has been proven to give a more sensitive representation of the continuous OA fingerprints variation. Scaled residuals minimisation also supported that the *rolling* solution was mathematically superior to *seasonal*.

The following multi-site comparison pretends to contrast both PMF methods in real-world datasets treated homogeneously under Chen et al. (2022) protocol to observe general performance trends. The *rolling* method generally presents a comparatively similar proportion of Primary OA (POA), and a secondary OA (SOA) of lower oxygenation degree, i.e. the ageing state. The double-tailed Welch's t-test showed that the narrower the window of inspection, the higher the differences between factors retrieved from one method to the other. Moreover, towards weekly or daily periods, SOA factors differ more than POA factors. This fact is likely due to the absence of constraints for the SOA factors during PMF. Contrastingly, POA factors are more dissimilar period-wise. The ratio of BBOA-to-HOA differs considerably from *rolling* to *seasonal* (1.45, 1.17, respectively) for the ensemble of sites, but in any case, it is over 1, as the ratio of BC_{wb}-to-BC_{ff} suggests.

In general terms, *rolling* results correlate better with ancillary measurements than those from *seasonal* for almost each of the considered external datasets at all sites. This is particularly true in the days surrounding the change of season, in which the *seasonal* profiles change drastically from one time point to the following. Model residuals also point to a better minimisation for the *rolling* PMF, although regarding scaled residuals both methods comply with the (-3, 3) range advised by the protocol. The time series of key ions also quantitatively pointed to a better adequation of the *rolling* SOA profiles to the oxygenated OA key ions. Finally, the errors also proved to be more stable for the r*olling* method, while it should be noted that the individual sites' discrepancies from the overall trends have not been discussed in this study.

Overall, these results confirm the hypothesis that the *rolling* PMF can be considered more accurate and precise, globally, than the *seasonal* one, although both meet the standards of quality required by the source apportionment protocol. Moreover, the *rolling* method was already recognised to involve less user subjectivity and computational time as well as to be more suitable for long-term and evolving SA analysis, such as semi-automated online SA. This study, therefore, promotes the acceptance of this novel *rolling* method as an improved approach suitable for source apportionment studies. An additional conclusion stemming from this comparison is that the selection of anchor profiles is highly influencing the OA factors, so local reference profiles are encouraged to minimise this impact.

5 References

Amato, F., Alastuey, A., Karanasiou, A., Lucarelli, F., Nava, S., Calzolai, G., Severi, M., Becagli, S., Gianelle, V. L., Colombi, C., Alves, C., Custódio, D., Nunes, T., Cerqueira, M., Pio, C., Eleftheriadis, K., Diapouli, E., Reche,

C., Minguillón, M. C., Manousakas, M. I., Maggos, T., Vratolis, S., Harrison, R. M. and Querol, X.: AIRUSE-LIFE+: A harmonized PM speciation and source apportionment in five southern European cities, Atmos. Chem. Phys., 16(5), 3289–3309, doi:10.5194/acp-16-3289-2016, 2016.

Belis, C. A., Pernigotti, D., Karagulian, F., Pirovano, G., Larsen, B. R., Gerboles, M. and Hopke, P. K.: A new methodology to assess the performance and uncertainty of source apportionment models in intercomparison exercises, Atmos. Environ., 119, 35–44, doi:10.1016/j.atmosenv.2015.08.002, 2015.

Bozzetti, C., El Haddad, I., Salameh, D., Daellenbach, K. R., Fermo, P., Gonzalez, R., Minguillón, M. C., Iinuma, Y., Poulain, L., Elser, M., Müller, E., Slowik, J. G., Jaffrezo, J. L., Baltensperger, U., Marchand, N. and Prévôt, A. S. H.: Organic aerosol source apportionment by offline-AMS over a full year in Marseille, Atmos. Chem. Phys., 17(13), 8247–8268, doi:10.5194/acp-17-8247-2017, 2017.

Canonaco, F., Crippa, M., Slowik, J. G., Baltensperger, U. and 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., 6(12), 3649–3661, doi:10.5194/amt-6-3649-2013, 2013.

Canonaco, F., Slowik, J. G., Baltensperger, U. and Prévôt, A. S. H.: Seasonal differences in oxygenated organic aerosol composition: Implications for emissions sources and factor analysis, Atmos. Chem. Phys., 15(12), 6993–7002, doi:10.5194/acp-15-6993-2015, 2015.

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. and 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., 14(2), 923–943, doi:10.5194/amt-14-923-2021, 2021.

Chazeau, B., Temime-Roussel, B., Gille, G., Mesbah, B., D'Anna, B., Wortham, H. and Marchand, N.: Measurement report: Fourteen months of real-time characterisation of the submicronic aerosol and its atmospheric dynamics at the Marseille-Longchamp supersite, Atmos. Chem. Phys., 21(9), 7293–7319, doi:10.5194/acp-21-7293-2021, 2021.

Chen, G., Sosedova, Y., Canonaco, F., Fröhlich, R., Tobler, A., Vlachou, A., Daellenbach, K., Bozzetti, C., Hueglin, C., Graf, P., Baltensperger, U., Slowik, J., El Haddad, I. and Prévôt, A.: Time dependent source apportionment of submicron organic aerosol for a rural site in an alpine valley using a rolling PMF window, Atmos. Chem. Phys. Discuss., (December), 1–52, doi:10.5194/acp-2020-1263, 2020.

Chen, G., Sosedova, Y., Canonaco, F., Fröhlich, R., Tobler, A., Vlachou, A., Daellenbach, K., Bozzetti, C., Hueglin, C., Graf, P., Baltensperger, U., Slowik, J., El Haddad, I. and Prévôt, A.: Time dependent source apportionment of submicron organic aerosol for a rural site in an alpine valley using a rolling PMF window, Atmos. Chem. Phys. Discuss., 43, 1–52, doi:10.5194/acp-2020-1263, 2021.

Chen, G., Canonaco, F., Tobler, A., Aas, W., Alastuey, A., Atabakhsh, S., Aurela, M., Baltensperger, U., Bougiatioti, A. and Joel, F.: European Aerosol Phenomenology - 8 : Harmonised Source Apportionment of

570   Organic Aerosol using 22 Year-long ACSM / AMS Datasets, 2022.

571   Chen, J. and Hoek, G.: Long-term exposure to PM and all-cause and cause-specific mortality: A systematic review
572   and meta-analysis, Environ. Int., 143(June), 105974, doi:10.1016/j.envint.2020.105974, 2020.

573   Crippa, M., Decarlo, P. F., Slowik, J. G., Mohr, C., Heringa, M. F., Chirico, R., Poulain, L., Freutel, F., Sciare, J.,
574   Cozic, J., Di Marco, C. F., Elsasser, M., Nicolas, J. B., Marchand, N., Abidi, E., Wiedensohler, A., Drewnick, F.,
575   Schneider, J., Borrmann, S., Nemitz, E., Zimmermann, R., Jaffrezo, J. L., Prévôt, A. S. H. and Baltensperger, U.:
576   Wintertime aerosol chemical composition and source apportionment of the organic fraction in the metropolitan
577   area of Paris, Atmos. Chem. Phys., 13(2), 961–981, doi:10.5194/acp-13-961-2013, 2013.

578   Crippa, M., Canonaco, F., Lanz, V. A., Äijälä, M., Allan, J. D., Carbone, S., Capes, G., Ceburnis, D., Dall'Osto,
579   M., Day, D. A., DeCarlo, P. F., Ehn, M., Eriksson, A., Freney, E., Ruiz, L. H., Hillamo, R., Jimenez, J. L.,
580   Junninen, H., Kiendler-Scharr, A., Kortelainen, A. M., Kulmala, M., Laaksonen, A., Mensah, A. A., Mohr, C.,
581   Nemitz, E., O'Dowd, C., Ovadnevaite, J., Pandis, S. N., Petäjä, T., Poulain, L., Saarikoski, S., Sellegri, K.,
582   Swietlicki, E., Tiitta, P., Worsnop, D. R., Baltensperger, U. and Prévôt, A. S. H.: Organic aerosol components
583   derived from 25 AMS data sets across Europe using a consistent ME-2 based source apportionment approach,
584   Atmos. Chem. Phys., 14(12), 6159–6176, doi:10.5194/acp-14-6159-2014, 2014.

585   Daellenbach, K. R., Uzu, G., Jiang, J., Cassagnes, L. E., Leni, Z., Vlachou, A., Stefenelli, G., Canonaco, F., Weber,
586   S., Segers, A., Kuenen, J. J. P., Schaap, M., Favez, O., Albinet, A., Aksoyoglu, S., Dommen, J., Baltensperger, U.,
587   Geiser, M., El Haddad, I., Jaffrezo, J. L. and Prévôt, A. S. H.: Sources of particulate-matter air pollution and its
588   oxidative potential in Europe, Nature, 587(7834), 414–419, doi:10.1038/s41586-020-2902-8, 2020.

589   Efron, B.: Bootstrap Methods: Another Look at the Jackknife, Ann. Stat., 7(1), doi:10.1214/aos/1176344552,
590   1979.

591   Forello, A. C., Bernardoni, V., Calzolai, G., Lucarelli, F., Massabò, D., Nava, S., Pileci, R. E., Prati, P., Valentini,
592   S., Valli, G. and Vecchi, R.: Exploiting multi-wavelength aerosol absorption coefficients in a multi-time source
593   apportionment study to retrieve source-dependent absorption parameters, Atmos. Chem. Phys. Discuss., 1–26,
594   doi:10.5194/acp-2019-123, 2019.

595   Fröhlich, R., Cubison, M. J., Slowik, J. G., Bukowiecki, N., Prévôt, A. S. H., Baltensperger, U., Schneider, J.,
596   Kimmel, J. R., Gonin, M., Rohner, U., Worsnop, D. R. and Jayne, J. T.: The ToF-ACSM: A portable aerosol
597   chemical speciation monitor with TOFMS detection, Atmos. Meas. Tech., 6(11), 3225–3241, doi:10.5194/amt-6-
598   3225-2013, 2013.

599   El Haddad, I., D'Anna, B., Temime-Roussel, B., Nicolas, M., Boreave, A., Favez, O., Voisin, D., Sciare, J.,
600   George, C., Jaffrezo, J. L., Wortham, H. and Marchand, N.: Towards a better understanding of the origins, chemical
601   composition and aging of oxygenated organic aerosols: Case study of a Mediterranean industrialized environment,
602   Marseille, Atmos. Chem. Phys., 13(15), 7875–7894, doi:10.5194/acp-13-7875-2013, 2013.

603   Heikkinen, L., Äijälä, M., Daellenbach, K., Chen, G., Garmash, O., Aliaga, D., Graeffe, F., Räty, M., Luoma, K.,

604   Aalto, P., Kulmala, M., Petäjä, T., Worsnop, D. and Ehn, M.: Eight years of sub-micrometre organic aerosol
605   composition data from the boreal forest characterized using a machine-learning approach, Atmos. Chem. Phys.
606   Discuss., 1–47, doi:10.5194/acp-2020-868, 2020.

607   IPCC: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth
608   Assessment Report of the Intergovernmental Panel on Climate Change [Masson-Delmotte, V., P. Zhai, A. Pirani,
609   S. L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, Cambridge Univ. Press, (In Press), 3949 [online] Available
610   from: https://www.ipcc.ch/report/ar6/wg1/downloads/report/IPCC_AR6_WGI_Full_Report.pdf, 2021.

611   Jiang, J., Aksoyoglu, S., El-Haddad, I., Ciarelli, G., Denier Van Der Gon, H. A. C., Canonaco, F., Gilardoni, S.,
612   Paglione, M., Minguillón, M. C., Favez, O., Zhang, Y., Marchand, N., Hao, L., Virtanen, A., Florou, K., O'Dowd,
613   C., Ovadnevaite, J., Baltensperger, U. and Prévôt, A. S. H.: Sources of organic aerosols in Europe: A modeling
614   study using CAMx with modified volatility basis set scheme, Atmos. Chem. Phys., 19(24), 15247–15270,
615   doi:10.5194/acp-19-15247-2019, 2019.

616   Jimenez, J. L., Canagaratna, M. R., Donahue, N. M., Prevot, A. S. H., Zhang, Q., Kroll, J. H., DeCarlo, P. F.,
617   Allan, J. D., Coe, H., Ng, N. L., Aiken, A. C., Docherty, K. S., Ulbrich, I. M., Grieshop, A. P., Robinson, A. L.,
618   Duplissy, J., Smith, J. D., Wilson, K. R., Lanz, V. A., Hueglin, C., Sun, Y. L., Tian, J., Laaksonen, A., Raatikainen,
619   T., Rautiainen, J., Vaattovaara, P., Ehn, M., Kulmala, M., Tomlinson, J. M., Collins, D. R., Cubison, M. J., Dunlea,
620   E. J., Huffman, J. A., Onasch, T. B., Alfarra, M. R., Williams, P. I., Bower, K., Kondo, Y., Schneider, J., Drewnick,
621   F., Borrmann, S., Weimer, S., Demerjian, K., Salcedo, D., Cottrell, L., Griffin, R., Takami, A., Miyoshi, T.,
622   Hatakeyama, S., Shimono, A., Sun, J. Y., Zhang, Y. M., Dzepina, K., Kimmel, J. R., Sueper, D., Jayne, J. T.,
623   Herndon, S. C., Trimborn, A. M., Williams, L. R., Wood, E. C., Middlebrook, A. M., Kolb, C. E., Baltensperger,
624   U. and Worsnop, D. R.: Evolution of organic aerosols in the atmosphere, Science (80-.)., 326(5959), 1525–1529,
625   doi:10.1126/science.1180353, 2009.

626   Ng, N. L., Herndon, S. C., Trimborn, A., Canagaratna, M. R., Croteau, P. L., Onasch, T. B., Sueper, D., Worsnop,
627   D. R., Zhang, Q., Sun, Y. L. and Jayne, J. T.: An Aerosol Chemical Speciation Monitor (ACSM) for routine
628   monitoring of the composition and mass concentrations of ambient aerosol, Aerosol Sci. Technol., 45(7), 770–
629   784, doi:10.1080/02786826.2011.560211, 2011a.

630   Ng, N. L., Canagaratna, M. R., Jimenez, J. L., Chhabra, P. S., Seinfeld, J. H. and Worsnop, D. R.: Changes in
631   organic aerosol composition with aging inferred from aerosol mass spectra, Atmos. Chem. Phys., 11(13), 6465–
632   6474, doi:10.5194/acp-11-6465-2011, 2011b.

633   Ogulei, D., Hopke, P. K., Zhou, L., Paatero, P., Park, S. S. and Ondov, J. M.: Receptor modeling for multiple time
634   resolved   species: The   Baltimore   supersite,   Atmos.   Environ.,   39(20),   3751–3762,
635   doi:10.1016/j.atmosenv.2005.03.012, 2005.

636   Paatero, P.: The Multilinear Engine—A Table-Driven, Least Squares Program for Solving Multilinear Problems,
637   Including   the   n-Way   Parallel   Factor   Analysis   Model,   J.   Comput.   Graph.   Stat.,   8(4),   854–888,
638   doi:10.1080/10618600.1999.10474853, 1999.

639   Paatero, P. and Hopke, P. K.: Discarding or downweighting high-noise variables in factor analytic models, Anal.
640   Chim. Acta, 490(1–2), 277–289, doi:10.1016/S0003-2670(02)01643-4, 2003.

641   Paatero, P. and Tapper, U.: Positive matrix factorization: A non-negative factor model with optimal utilization of
642   error estimates of data values, Environmetrics, 5(2), 111–126, doi:10.1002/env.3170050203, 1994.

643   Parworth, C., Fast, J., Mei, F., Shippert, T., Sivaraman, C., Tilp, A., Watson, T. and Zhang, Q.: Long-term
644   measurements of submicrometer aerosol chemistry at the Southern Great Plains (SGP) using an Aerosol Chemical
645   Speciation Monitor (ACSM), Atmos. Environ., 106, 43–55, doi:10.1016/j.atmosenv.2015.01.060, 2015.

646   Rai, P., Furger, M., El Haddad, I., Kumar, V., Wang, L., Singh, A., Dixit, K., Bhattu, D., Petit, J. E., Ganguly, D.,
647   Rastogi, N., Baltensperger, U., Tripathi, S. N., Slowik, J. G. and Prévôt, A. S. H.: Real-time measurement and
648   source apportionment of elements in Delhi's atmosphere, Sci. Total Environ., 742, 140332,
649   doi:10.1016/j.scitotenv.2020.140332, 2020.

650   Rodríguez, S. and Cuevas, E.: The contributions of "minimum primary emissions" and "new particle formation
651   enhancements" to the particle number concentration in urban air, J. Aerosol Sci., 38(12), 1207–1219,
652   doi:10.1016/j.jaerosci.2007.09.001, 2007.

653   Rutherford, J. W., Larson, T. V., Gould, T., Seto, E., Novosselov, I. V. and Posner, J. D.: Source apportionment
654   of environmental combustion sources using excitation emission matrix fluorescence spectroscopy and machine
655   learning, Atmos. Environ., 259(May), 118501, doi:10.1016/j.atmosenv.2021.118501, 2021.

656   Shih, C.-H., Chen, J.-K., Kuo, L.-W., Cho, K.-H., Hsiao, T.-C., Lin, Z.-W., Lin, Y.-S., Kang, J.-H., Lo, Y.-C.,
657   Chuang, K.-J., Cheng, T.-J. and Chuang, H.-C.: Chronic pulmonary exposure to traffic-related fine particulate
658   matter causes brain impairment in adult rats, Part. Fibre Toxicol., 15(1), 1–17, doi:10.1186/s12989-018-0281-1,
659   2018.

660   Sun, Y., Xu, W., Zhang, Q., Jiang, Q., Canonaco, F., Prévôt, A. S. H., Fu, P., Li, J., Jayne, J., Worsnop, D. R. and
661   Wang, Z.: Source apportionment of organic aerosol from 2-year highly time-resolved measurements by an aerosol
662   chemical speciation monitor in Beijing, China, Atmos. Chem. Phys., 18(12), 8469–8489, doi:10.5194/acp-18-
663   8469-2018, 2018.

664   Textor, C., Schulz, M., Guibert, S., Kinne, S., Balkanski, Y., Bauer, S., Berntsen, T., Berglen, T., Boucher, O.,
665   Chin, M., Dentener, F., Diehl, T., Easter, R., Feichter, H., Fillmore, D., Ghan, S., Ginoux, P., Gong, S., Grini, A.,
666   Hendricks, J., Horowitz, L., Huang, P., Isaksen, I., Iversen, T., Kloster, S., Koch, D., Kirkevåg, A., Kristjansson,
667   J. E., Krol, M., Lauer, A., Lamarque, J. F., Liu, X., Montanaro, V., Myhre, G., Penner, J., Pitari, G., Reddy, S.,
668   Seland, Stier, P., Takemura, T. and Tie, X.: Analysis and quantification of the diversities of aerosol life cycles
669   within AeroCom, Atmos. Chem. Phys., 6(7), 1777–1813, doi:10.5194/acp-6-1777-2006, 2006.

670   Tobler, A. K., Canonaco, F., Skiba, A., Styszko, K., Nęcki, J., Slowik, J. G. and Baltensperger, U.: Characterization
671   and source apportionment of PM 1 organic aerosol in Krakow , Poland, , (April), 8299, 2021.

672 Ulbrich, I. M., Canagaratna, M. R., Zhang, Q., Worsnop, D. R. and Jimenez, J. L.: Interpretation of organic

673 components from Positive Matrix Factorization of aerosol mass spectrometric data, Atmos. Chem. Phys., 9(9),

674 2891–2918, doi:10.5194/acp-9-2891-2009, 2009.

675 Yang, M., Chu, C., Bloom, M. S., Li, S., Chen, G., Heinrich, J., Markevych, I., Knibbs, L. D., Bowatte, G.,

676 Dharmage, S. C., Komppula, M., Leskinen, A., Hirvonen, M. R., Roponen, M., Jalava, P., Wang, S. Q., Lin, S.,

677 Zeng, X. W., Hu, L. W., Liu, K. K., Yang, B. Y., Chen, W., Guo, Y. and Dong, G. H.: Is smaller worse? New

678 insights about associations of PM 1 and respiratory health in children and adolescents, Environ. Int., 120(August),

679 516–524, doi:10.1016/j.envint.2018.08.027, 2018.

680 Yin, P., Guo, J., Wang, L., Fan, W., Lu, F., Guo, M., Moreno, S. B. R., Wang, Y., Wang, H., Zhou, M. and Dong,

681 Z.: Higher Risk of Cardiovascular Disease Associated with Smaller Size-Fractioned Particulate Matter, Environ.

682 Sci. Technol. Lett., 7(2), 95–101, doi:10.1021/acs.estlett.9b00735, 2020.

683 Yuan, B., Shao, M., De Gouw, J., Parrish, D. D., Lu, S., Wang, M., Zeng, L., Zhang, Q., Song, Y., Zhang, J. and

684 Hu, M.: Volatile organic compounds (VOCs) in urban air: How chemistry affects the interpretation of positive

685 matrix factorization (PMF) analysis, J. Geophys. Res. Atmos., 117(24), 1–17, doi:10.1029/2012JD018236, 2012.

686 Zhang, Y., Favez, O., Petit, J. E., Canonaco, F., Truong, F., Bonnaire, N., Crenn, V., Amodeo, T., Prévôt, A. S.

687 H., Sciare, J., Gros, V. and Albinet, A.: Six-year source apportionment of submicron organic aerosols from near-

688 continuous highly time-resolved measurements at SIRTA (Paris area, France), Atmos. Chem. Phys., 19(23),

689 14755–14776, doi:10.5194/acp-19-14755-2019, 2019.

690

691 **Acknowledgements.**

692 IDAEA-CSIC is a Centre of Excellence Severo Ochoa (Spanish Ministry of Science and Innovation, Project

693 CEX2018-000794-S). The authors gratefully acknowledge the Romanian National Air Quality Monitoring

694 Network (NAQMN, www.calitateaer.ro) for providing NOx data.

695 **Financial support.**

696 This work was supported by COST Action CA16109 COLOSSAL, the Generalitat de Catalunya (grant no.

697 AGAUR 2017 SGR41), the Spanish Ministry of Science and Innovation 70 through the CAIAC project (grant no.

698 PID2019- 108990RB-I00), and FEDER funds through EQC2018-004598-P. This work has received funding from

699 the RI-Urbans project from European Union's Horizon 2020 research and innovation programme under grant

700 agreement No 101036245. This study was partially supported by the European Union's projects ACTRIS (EU

701 FP7-262254) and ACTRIS-2 (EU Horizon 2020-654109). This work was supported by a grant of the Romanian

702 Ministry of Research, Innovation and Digitalization, CNCS - UEFISCDI, project number PN-III-P1-1.1-TE-2019-

703 0340 and through Program 1- Development of the national research-development system, Subprogram 1.2 -

704 Institutional performance - Projects to finance the excellent RDI, Contract no. 18PFE/30.12.2021, by Romanian

705 National Core Program contract 18N/2019. IMT Nord Europe acknowledges financial support from the Labex
706 CaPPA project, which is funded by the French National Research Agency (ANR) through the PIA (Programme
707 d'Investissement d'Avenir) under contract ANR-11-LABX-0005-01, and the CLIMIBIO project, both financed
708 by the Regional Council "Hauts-de-France" and the European Regional Development Fund (ERDF). The ATOLL
709 and SIRTA sites are part of ACTRIS-France facilities and also contribute to the CARA program of the LCSQA
710 funded by the French Ministry of Environment. JS and MP acknowledge funding from the European Union's
711 Horizon 2020 Research and Innovation Programme (under grant agreement no. 856612) and the Cyprus
712 Government. National University of Ireland Galway work was supported by the EPA AEROSOURCE Project
713 (2016-CCRP-MS-31) and AC3 network, the Department of Environment, Climate and Communications. GC and
714 AP has been supported by the European Research Council, H2020 European Research Council (ERA-PLANET
715 (grant no. 689443)) and a COST-related project of the Swiss National Science Foundation, the Schweizerischer
716 Nationalfonds zur Förderung der Wissenschaftlichen Forschung (SAMSAM (grant no. IZCOZ0_177063)). K.R.D
717 acknowledges support by the Swiss National Sciencc Foundation Ambizione grant PZPGP2_201992.

718 **Figures and tables**

719 **Table 1. Participant sites.**

Site	Type	Latitude	Longitude	Height (m a.s.l)	ACSM type	Network	Period
Barcelona – Palau Reial (BCN-PR)	Urban background	41° 23' 11.48'' N	02° 07' 05.00'' E	80	Q	ACTRIS, GAW	09/2017 – 10/2018
Cyprus Atm. Obs. - Agia Marina Xyliatou (CAO-AMX)	Remote	35° 2' 19.35''N	33° 3' 27.95'' E	352	Q	ACTRIS, GAW	03/2015 – 01/2017

Dublin (DUB)	Urban background	53° 18' 19.08'' N	06° 13' 4.52'' W	35	Q	ACTRIS, AQ network in Ireland: http://www.macehead.org/	09/2015 – 08/2017
ATmospheric Observations in liLle (ATOLL)	Suburban	50° 36' 40.0 N	03° 08' 25.4'' E	70	Q	ACTRIS, CARA programme (French AQ network)	10/2016 – 09/2017
Magadino (MGD)	Rural	46°9′37″ N	8° 56′ 2″ E	204	Q	GAW	08/2013 – 10/2014
Magurele – INOE (INO)	Peri-urban	44° 20' 52.98'' N	26º 1' 43.93'' E	93	Q	ACTRIS, GAW	09/2016 - 09/2017
Marseille-Longchamp (MRS-LCP)	Urban background	43° 18' 18.84'' N	5° 23' 40.89'' E	71	ToF	CARA programme (French AQ network)	01/2017 – 04/2018
SIRTA (SIR)	Suburban	48° 42'' 36'' N	2° 9'' 0'' E	163	Q	ACTRIS	01/2016 - 05/2017

Tartu (TAR)	Urban background	58° 22' 14.16'' N	26º 44' 5.64'' E	39	Q		National Air monitoring station	09/2016 – 07/2017

720

721

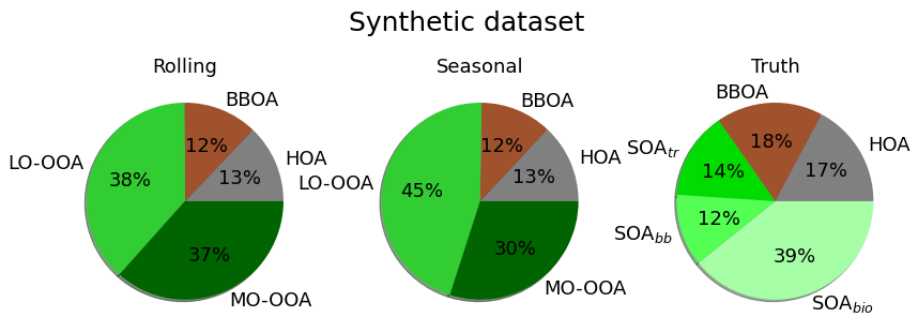



Figure 1. OA apportionment results for *rolling* and *seasonal* methods and truth output.

723

724

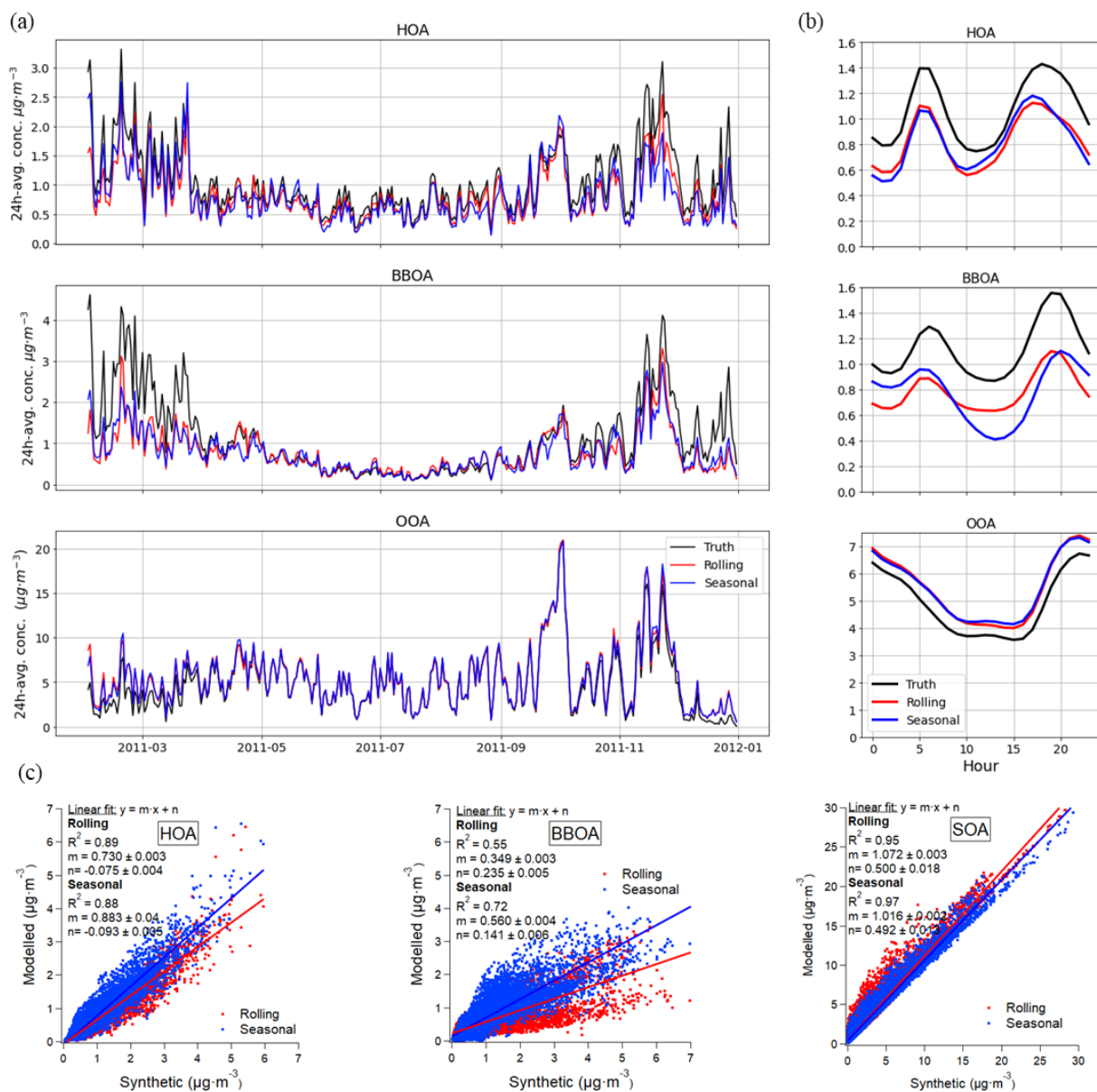



Figure 2. Rolling, seasonal and truth (synthetic dataset original values) (a) time series (in hourly averages for the sake of clarity), (b) diel profiles and (c) scatter plots.

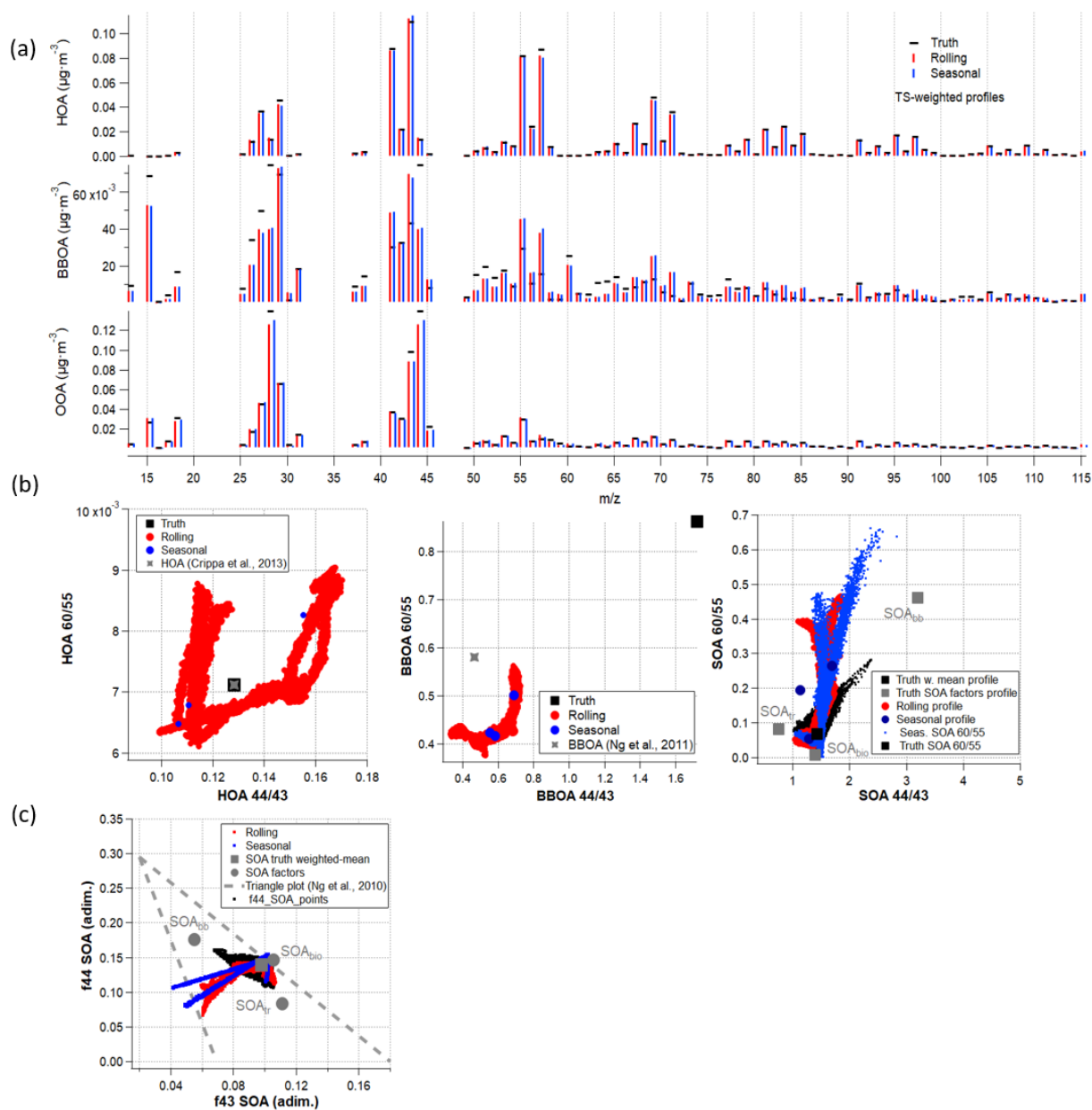
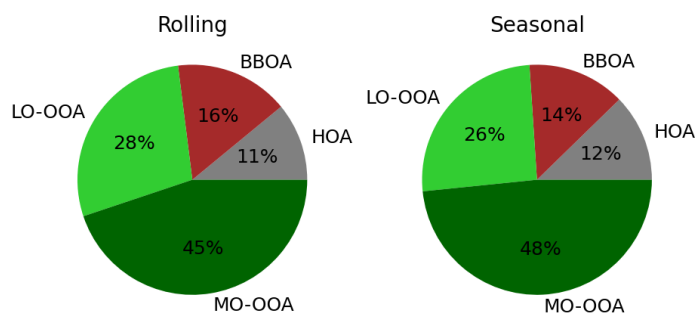
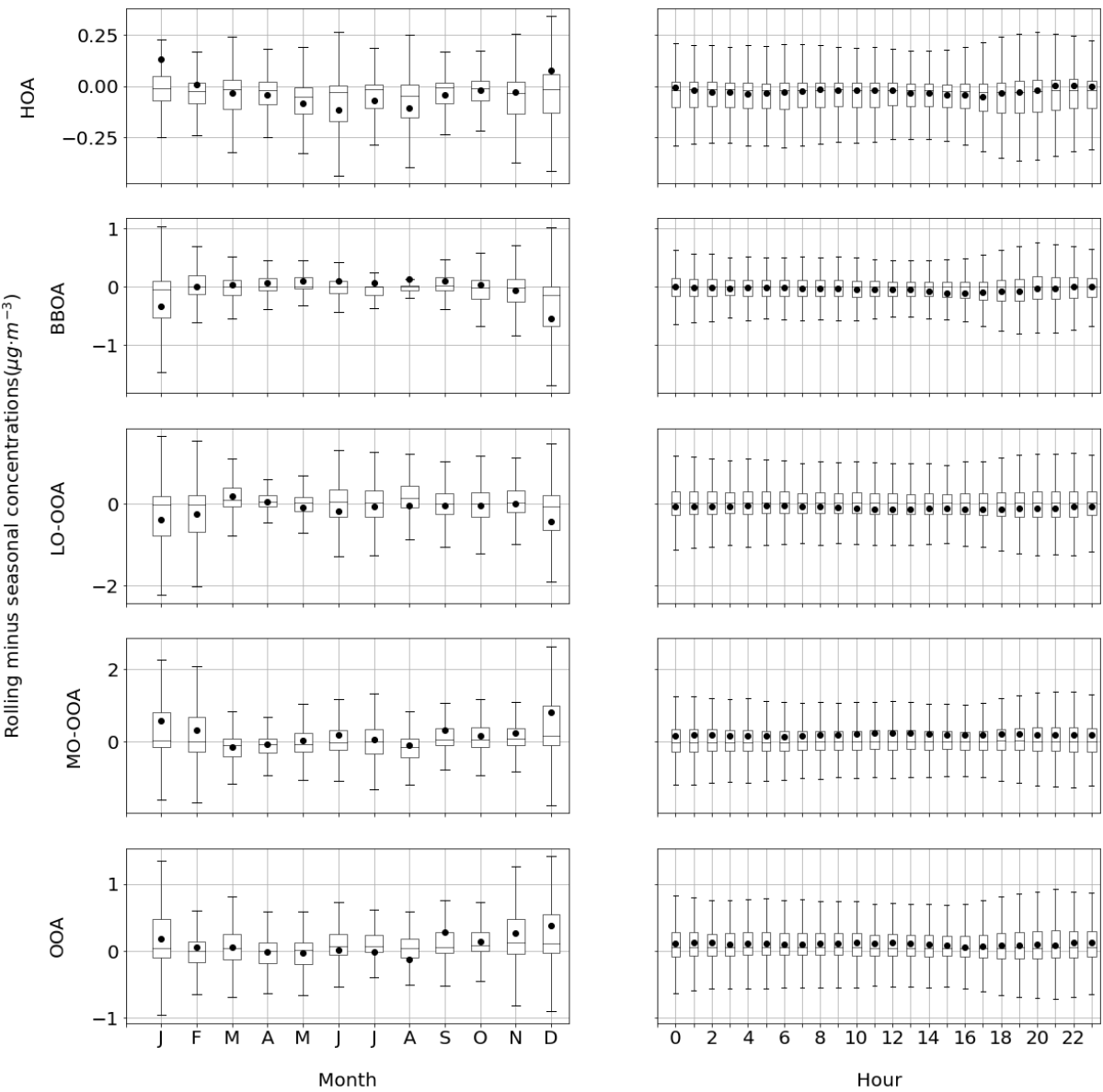



Figure 3. Synthetic dataset solution (a) Profiles; (b) Time-dependent profile variability of ratios 60/55 vs. 44/43; (c) Triangle plot of f44 vs. f43; for *rolling* PMF (red), *seasonal PMF* (blue) and *truth* (black).

731 **Figure 4. Pie charts of the mean concentrations of the main factors for the ensemble of all sites.**

732

733

734 **Figure 5. Boxplots of *rolling* minus s*easonal* factor absolute concentrations (in μg·m⁻³) per month and hour.**

735 **Boxes show the Q1-to-Q3 range with the median (horizontal line) and the average (full circles), and whiskers**

736 **extend up to the range of the data.**

737

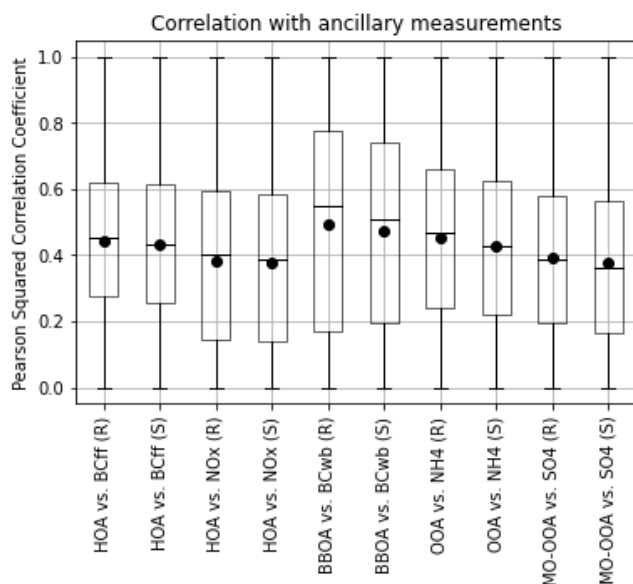



Figure 6. *Rolling* **(R) and** *seasonal* **(S) boxplots of the Pearson-squared correlation coefficient of each OA source with**

its respective markers for all sites.

740

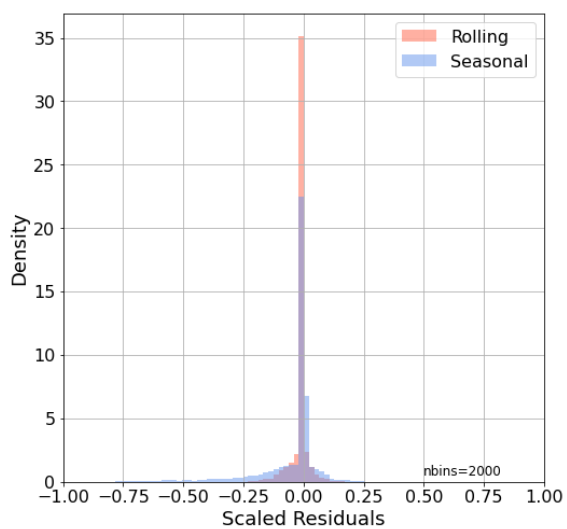



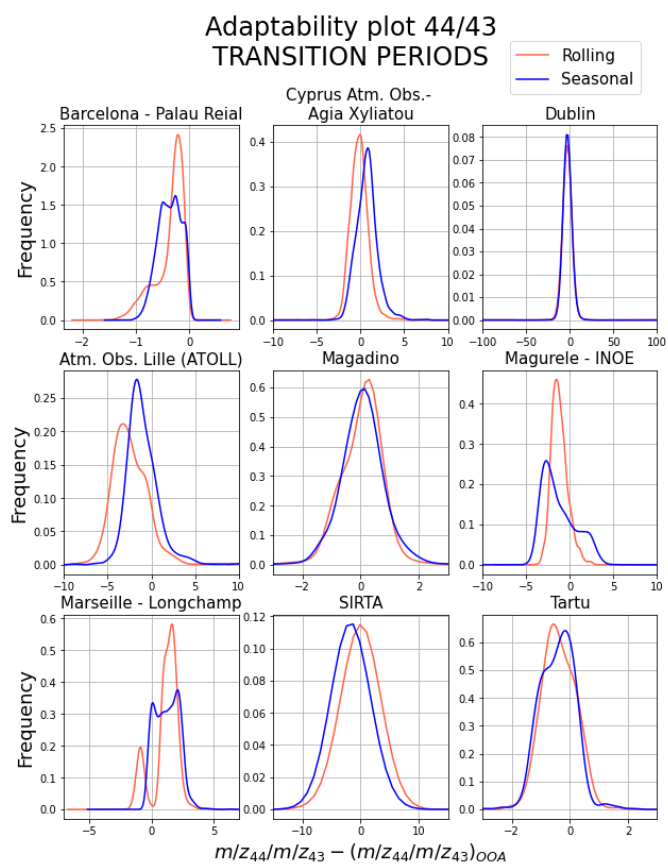
Figure 7. Normalised scaled residuals histogram for both PMF techniques.

Table 2. Q/Qexp values for *rolling* **and** *seasonal* **solutions.**

Q	Raw		Normalised	
	Rolling	Seasonal	Rolling	Seasonal
BCN - PR	**481008**	1766588	**0.35**	1.14
CAO - AMX	**57337**	5101949	**0.04**	2.87

DUB	**1031616**	1261451	**1.14**	1.19
ATOLL	**465480**	477145	0.84	**0.69**
MGD	**8463251**	3117660	**0.75**	2.46
INO	**6138684**	25404272	**4.58**	17.24
MRS – LCP	**57337**	5101949	17.24	**2.87**
SIR	558044	**44965**	0.47	**0.10**
TAR	**82742**	152343	0.59	**0.34**

743
744
745
746

747

748   **Figure 8. Kernel density estimation of the histograms of the subtraction of the m/z 44-to-43 ratio from the**

749   **raw (from input matrirces) time series data minus the apportioned quantity in profiles. These plots only**

750   **contain those time lapses among the change of season (transition periods)**

751

752