Milestone M6 (M1.7)

Accuracy check of Online Source Apportionment Service Tool vs. Offline Source Apportionment analysis

RI-URBANS

Research Infrastructures Services Reinforcing Air Quality Monitoring Capacities in European Urban & Industrial AreaS (GA n. 101036245)

By

INERIS, CSIC, CNRS & FMI

31st May 2024

RI-URBANS (www.RIURBANS.eu) is supported by the European Commission under the Horizon 2020 – Research and Innovation Framework Programme, H2020-GD-2020, Grant Agreement number: 101036245

Milestone M6 (M1.7): Accuracy check of Online Source Apportionment Service Tool vs. Offline Source Apportionment analysis.

Authors: Mohamed Gherras & Olivier Favez (INERIS), Marjan Savadkoohi & Marco Pandolfi (CSIC), Jean-Eudes Petit (CNRS) and Hilkka Timonen (FMI)

Work package (WP)	WP1, Novel AQ metrics and advanced source apportionment STs for PM, and nanoparticles
Milestone	M6 (M1.7)
Lead beneficiary	INERIS
Means of verification	Accuracy check of online source apportionment service tool checked by comparing with offline source apportionment analysis: Accuracy check made available
Estimated delivery deadline	M32 (31/05/2024)
Actual delivery deadline	31/05/2024
Version	Final
Reviewed by	WP1 leaders and PCT
Accepted by	RI-URBANS Project Coordination Team
Comments	This document presents the methodologies set-up to evaluate the accuracy of NRT-SA outputs obtained at WP4.1 pilot sites

RI-URBANS (www.RIURBANS.eu) is supported by the European Commission under the Horizon 2020 – Research and Innovation Framework Programme, H2020-GD-2020, Grant Agreement number: 101036245

Table of Contents

1. About this document / scope

In RI-URBANS, WP1.2 defined Service Tools (STs) - to be tested within WP4.1 - for the source apportionment (SA) of fine urban carbonaceous aerosols in Near-Real Time (NRT). These STs are described in Deliverable D4 (D1.4) (Dec. 2022). They employ statistical tools for the analysis of Organic Aerosols (OA) and equivalent Black Carbon (eBC) measurements obtained using Aerosol Chemical Speciation Monitors (ACSM, Aerodyne Res. Inc.) and multiwavelength aethalometers (AE33, Magee Scientific), respectively. Expected outputs include the contributions of primary sources such as traffic, wood burning, and cooking and/or secondary fractions depending on the parameter (i.e., OA and eBC) and on the measurement site.

The present document describes how WP1 (T1.2) provided further support to WP4 (T4.1) in order to evaluate the accuracy of NRT-SA outputs obtained using these STs, which were tested throughout 2023 in various European cities (including Athens, Barcelona, Bucharest, Helsinki, Milano, Paris, Marseille and Zurich; see Figure 1). These pilots encompass diverse European urban environments with the aim to demonstrate - at a real scale - the ability to integrate complementary air quality (AQ) measurement systems in existing AQ monitoring networks (AQMNs).

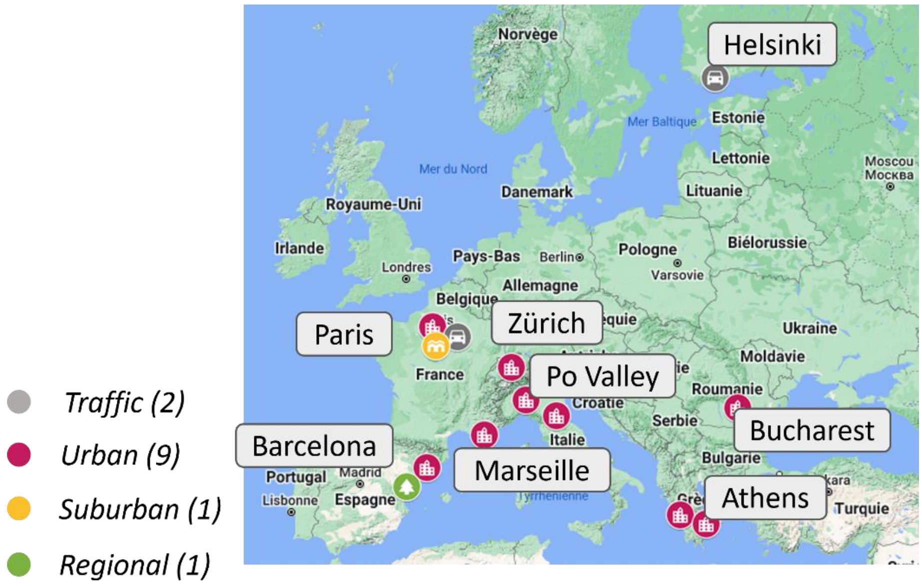

Figure 1. *Pilot sites in RI-URBANS WP4.1*

The evaluation of NRT-SA STs' outputs has been achieved through their comparisons with results obtained from advanced statistical analyses performed a posteriori.

This is a public document, available at the RI-URBANS website, https://riurbans.eu/work-package-1/#milestones-wp1, and distributed to all RI-URBANS partners for their use as well as submitted to the European Commission as an RI-URBANS Milestone M6 (M1.7).

2. Fine organic aerosol Source Apportionment

Source apportionment (SA) of organic aerosols (OA) is highly challenging due to the multiplicity/diversity of their emission sources and (trans-)formation processes, resulting in a tremendous molecular complexity of this major fraction of airborne particles. Along with the worldwide spreading of aerosol-dedicated online mass spectrometry techniques in the two last decades, advanced statistical methods have been developed to deconvolute OA into a handful of subfractions (or factors) according to the shape of their mass spectra.

In particular, Datalystica (based in Switzerland, and born as a spin-off of the Paul Scherrer Institute, Villigen) is proposing a commercial software - namely SoFi Pro - currently and widely considered as the most efficient and easy tool to deconvolute OA mass spectra in a harmonized manner (Chen et al., 2022a). Moreover, recent developments in this software include NRT SA, using a combination of rolling Positive Matrix Factorization (PMF) and Chemical Mass Balance (CMB). The accuracy of this tool has been demonstrated “offline” at three European cities (Zürich, Athens & Paris), as described by Chen et al. (2022b). RI-URBANS piloting activities throughout 2023 offered the opportunity to evaluate the accuracy of this software, embedded in featured NRT data flow and automatic data qualification processes developed to fit RI-URBANS purposes.

As for Chen et al. (2022b), during this pilot, SA outputs obtained from a posteriori advanced analyses have been considered as the reliable “best estimate” of OA fractions, and thus as a reference, to which NRT results should be benchmarked. These comparisons of NRT SA outputs with manual SA results has been set to be conducted monthly (by the data providers), concomitantly with a measurement status report. Furthermore, it has been agreed that reference datasets shall be obtained using the most recent version of the SoFi Pro software, with a harmonized protocol and as commonly performed and described in the literature. This protocol is mostly based on the European phenomenology on source apportionment of organic aerosols described in Chen et al. (2022a). This protocol has been discussed between WP1.2 and WP4.1 coordinators and then with all the participants of the pilot phase. It has been designed to prevent from potential site-specific discrepancies between NRT & a posteriori analyses.

This protocol has already been described in a previous Deliverable D23 (D4.2), available online since September 2023). Results have also been continuously shared and gathered within WP4.1 partners through internal reports issues on a monthly basis by the pilot cities (following recommendations given by Deliverable D22 (D4.1)).

Example of comparison results is presented in Figure 2.

Overall, results from this pilot activity show that consistency of NRT SA outputs with a “reference” results (obtained a posteriori from more robust data treatment analyses) can be reached, as long as the whole chain of the data flow and data handling procedures is operational. Site-to-site discrepancies have been observed, highlighting the key parameters that still need to be framed and harmonized. Indeed, in the first half of the project, most of the efforts have been focused on making the whole chain work, before trying to optimize these parameters. To this end, constant interactions between T4.1 and T1.2 have been on going in order to stabilize data flows, troubleshoot and debug codes. Current efforts within WP4.1 are now dedicated to re-process the 2023 pilot datasets as best as they could have been obtained in NRT if all the successive updates in the ST parametrization would have been in place already before the beginning of the pilot phase. This will be presented in Deliverable D24 (D4.3) (M40).

Marseille, June 2023

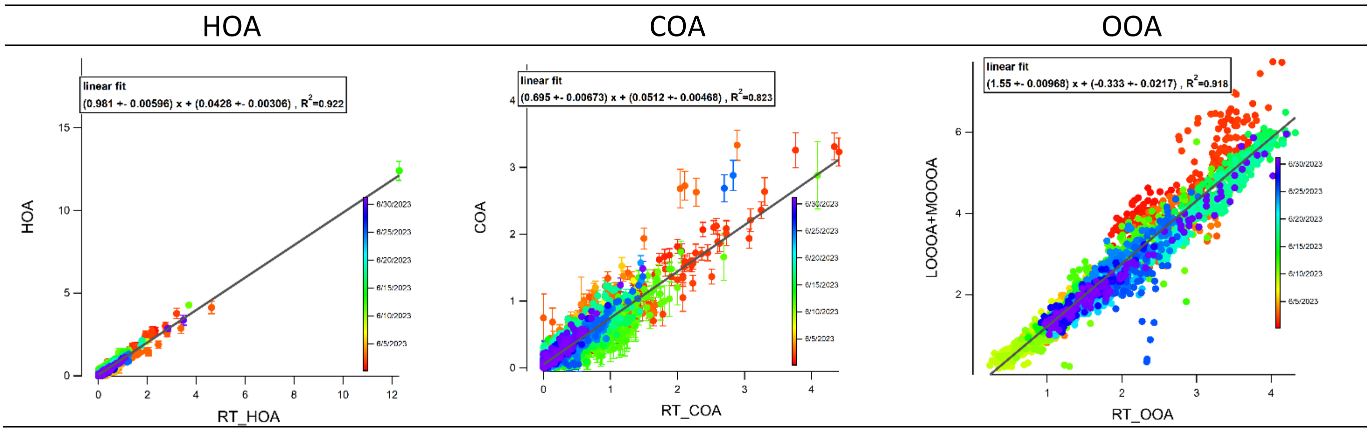

Figure 2. *Example of comparison between fine OA NRT-SA results and a posteriri/manual analysis at Marseille urban background pilot site*

3. Equivalent Black Carbon (eBC) Source Apportionment

Within WP4.1, eBC measurements have been conducted using multi-wavelengths filter-based absorption photometers, and more precisely AE33 devices, at each pilot site. Data were then processed according to the ACTRIS procedures and following the recommendations provided by Savadkoohi et al. (2024) in the frame of the RI-URBANS project. This allowed to harmonize measurements of the wavelength-dependent absorption coefficients ($b_{abs,\lambda}$), recalculated from ATN measurements provided by the AE33 using harmonization factors (H*) to convert the ATN into b_{abs} according to ACTRIS recommendation guidelines. Applying H* factors is necessary because the multiple-scattering enhancement correction (C_0) values differ from those set as a default value within the instrument firmware. For AE33 with filter tapes M8020/M8050, H* is 2.21 and C_0 is 1.57; for the new filter tape, H* is 1.76 and C_0 is 1.39. The wavelengths of 470 nm and 950 nm are chosen for source apportionment.

Once corrected in this way, AE33 measurements could then be used to quantify the two main eBC sub-fractions, using the so-called Aethalometer Model first introduced by Sandradewi et al. (2008). Initially considered to be representative of traffic exhaust and biomass burning emissions, respectively, these 2 eBC subfractions are nowadays rather attributed to liquid fuel vs. solid fuel combustion processes (leading to eBC_{LF} and eBC_{SF} aerosols). Furthermore, the output of this model is highly sensitive to the input of source-specific Angström Absorption Exponents (AAEs). Indeed, the method relies on accurate calibration factors and representative AAEs for different sources, which may introduce uncertainties due to changes in the characteristics of the optical properties of a specific source over time. This is because the AE bilinear model assumes that the absorption properties of eBC remain constant, regardless of aging or mixing with other aerosol components and regardless possible changes in the chemical fingerprints for a given emission source.

Site-specific information can be gathered only after comprehensive measurements have been conducted at this site. In this respect, during the 2023 pilot phase, eBC NRT-SA outputs have been obtained using the same set of default AEE_{LF} and AAE_{SF} values (i.e., 1 and 2, respectively) at all sites. Site-specific values could then be determined - in order to refine the SA outputs and evaluate the accuracy of those obtained in NRT - thanks to thorough a posteriori investigation of the measurements. This notably included comparison with independent data obtained from co-located instruments. In particular, ACSM data, including the signal at mass-to-charge m/z 60, could be of high relevance for such exercises. Indeed, m/z 60 is often associated with certain organic compounds produced by solid fuel combustion, such as polycyclic aromatic hydrocarbons (PAHs) and oxygenated organic aerosols. For instance, m/z 60 can correlate with biomass-burning markers like levoglucosan in many regions. Here, we used m/z60 data from ACSM where available to study its correlation with solid fuel sources. When available, nitrogen oxides (NO_x) datasets were also used for information on the dynamic of traffic emissions.

As a first step of a posteriori data analyses, ambient AAE values (resulting from the mixing of both eBC subfractions) can be computed for each site using frequency distribution to determine the optimal AAE value that best represents the overall distribution. The AAE distribution was filtered with a stringent requirement of $R2>0.99$ for the fit of the natural logarithm of b_{abs} versus wavelength to identify the source and site-specific AAE values (Figure 3). Using data with sufficient period is also important to capture the variability of AAE. Different lower and upper percentiles were calculated using mean and median to determine the central tendency and distribution of AAE values across different sites and seasons. Applying AAE-filtered data's tails provides a threshold for the 1st percentile (AAE_{LF}) and the 99th percentile (AAE_{SF}) of the frequency distribution. This approach was tested for historical and NRT datasets.

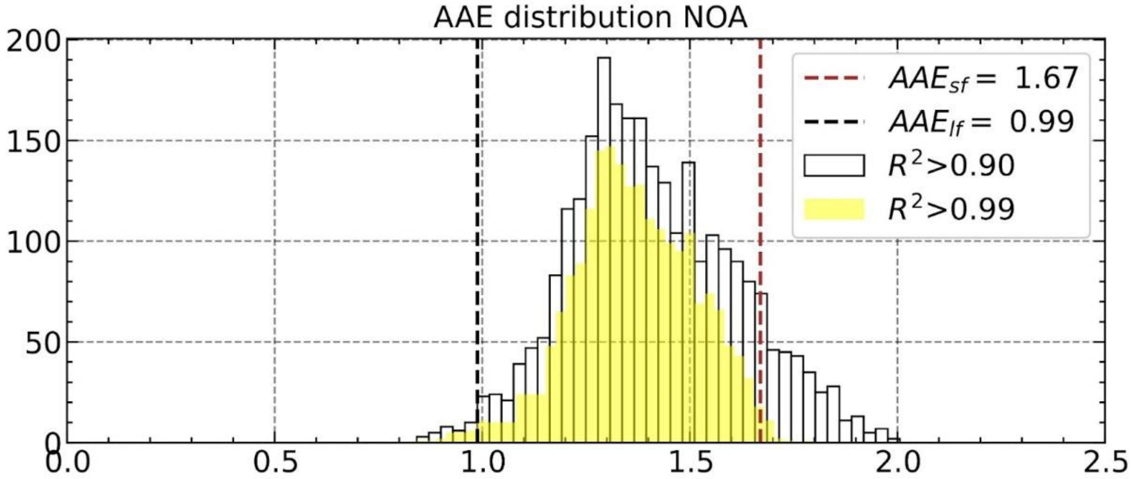

Figure 3. *Example of ambient AAE frequency distributions obtained from measurements performed duing the pilot phase (NOA =one of the 2 Athens pilot sites, urban background station)*

Overall, the 1st percentile of ambient AAE values typically represents the freshest, least-aged emissions from liquid fuels which is reasonably robust and consistent with published values in the literature. In addition, using the 1st percentile filters out noise and outliers in the data that might arise from measurement errors or atmospheric conditions. Moreover, the correlation between solid fuel sources and m/z60 can determine the best fit in terms of AAE_{SF} percentiles (Figure 4). The obtained correlation typically depends on the chemical composition and combustion properties of the solid fuels being burnt (e.g., coal, biomass, wood), combustion conditions, and atmospheric processing (e.g., aging, chemical reactions).

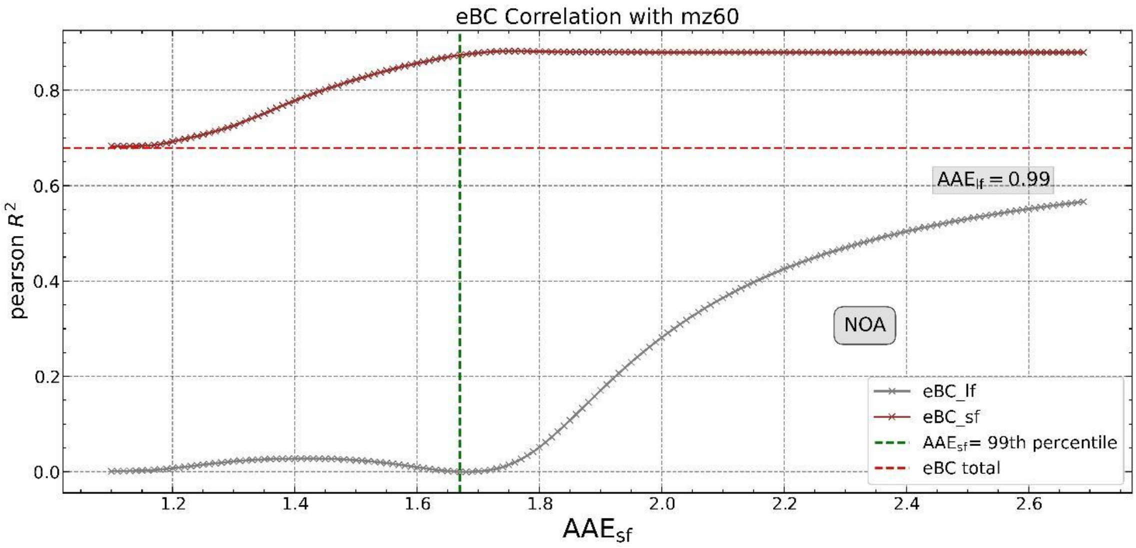

Figure 4. *Example of sensitivity analysis of eBC SA outputs as a function of the chosen AAE_{SF} value*

Based on this methodology, calculating eBC$_{LF}$ and eBC$_{SF}$ was performed across a range of AAE$_{SF}$ values (1.1-2.6), maintaining a fixed AAE$_{LF}$ (1st percentile from the AAE frequency distribution). The objective was to identify the AAE$_{SF}$ value that maximizes the correlation between eBC$_{SF}$ and m/z 60, while intentionally minimizing the correlation for eBC$_{LF}$. This approach leads to a higher correlation between eBC$_{SF}$ and m/z 60, while deliberately seeking a lower correlation for eBC$_{LF}$. Notably, lower AAE$_{SF}$ values correspond to higher proportions of eBC$_{SF}$, highlighting the importance of optimizing this parameter for refining the AE model's outcomes. The Pearson correlation coefficient was used to find the optimum AAE$_{SF}$. This should preferentially be done on wintertime dataset, in order to get sufficient amount of both eBC subfractions in ambient air. Table 1 summarized the set of values determined in this way for the different pilot sites.

Table 1. *Examples of AAEs' sensitivity analysis (using co-located measurements) results and comparison with NRT-SA results using the same set of AAEs values at each site.*

Station	From	To	α_{sf}	α_{lf}	Slope eBC$_{lf}$	Slope eBC$_{sf}$	R^2 BC$_{lf}$	R^2 BC$_{sf}$
Athens	21/12/2022	28/02/2023	1.72	0.97	0.86	1	0.98	0.99
Bucharest	29/11/2022	28/02/2023	1.82	1.15	0.58	1.04	0.93	0.96
Bologna	29/11/2022	28/02/2023	1.75	1.22	0.62	1.29	0.98	0.83
Helsinki	13/12/2022	28/02/2023	1.5	1.03	1.07	1.67	0.96	0.98
Milano	20/12/2022	28/02/2023	1.77	1.23	0.57	1.24	0.95	0.91
Paris Bp Est	04/01/2023	28/02/2023	1.52	0.96	1.05	1.43	0.96	0.99
Paris Chatelet	10/01/2023	28/02/2023	1.57	1.06	1.05	1.57	0.90	0.99
SIRTA	21/11/2022	28/02/2023	1.92	1.11	0.59	0.82	0.92	0.97
Zurich	16/01/2023	28/02/2023	1.74	1.01	0.80	1	0.99	0.99

RI-URBANS (www.RIURBANS.eu) is supported by the European Commission under the Horizon 2020 – Research and Innovation Framework Programme, H2020-GD-2020, Grant Agreement number: 101036245

At some sites, the correlation method did not work properly for different reasons. To effectively separate the two fractions of eBC using m/z 60, we should observe a better correlation between m/z 60 and lower AAE compared to higher AAE values. However, this is not always the case due to issues such as ACSM data quality, non-collocated measurements, or problems with AE33 data. In particular, eBC NRT-SA appeared to be highly challenging, if not impossible, at kerbside due to the overwhelming influence of various types of eBC_{LF} emissions. Therefore, the ST developed and tested with RI-URBANS is deemed to be used preferentially at urban background (AQMN) stations. All the results will be detailed in an upcoming scientific paper currently under preparation, which will combine eBC source apportionment results obtained in the frame of the RI-URBANS project from both the pan-European review of historical data and the piloting activities. For illustration in the present document, Figure 5 displays the eBC SA outputs obtained within WP4.1 for the Bucharest urban background pilot site in 2023.

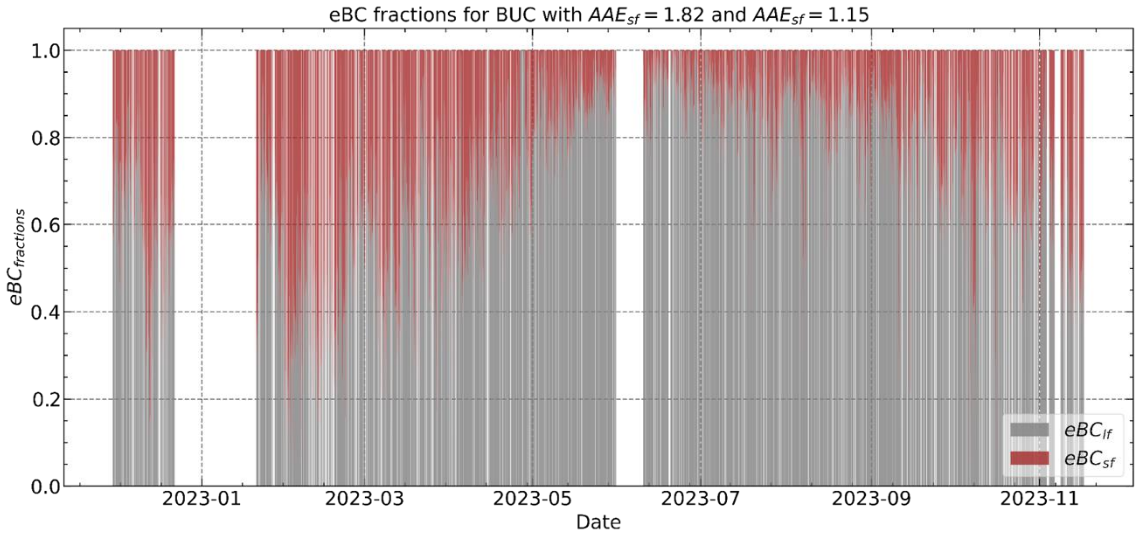

Figure 5. *Relative contributions of* eBC_{LF} *and* eBC_{SF} *to total eBC obtained for the Bucharest pilot site in WP4.1*

4. Conclusion and Outlook

In RI-URBANS, WP1.2 not only provided WP4.1 with STs to be tested for NRT-SA of fine OA and eBC during the pilot phase, but also defined (in collaboration with WP4.1 coordinators) the procedures to evaluate the accuracy of STs' outputs obtained in NRT at pilot sites. The present document summarized these procedures, which could be used in a timely manner.

As mentioned above, eBC SA results obtained following the procedures described herein will be merged with complementary analyses of historical datasets, analysed the same way, to describe the phenomenology of eBC subfractions at the European level. Regarding the ST for OA SA, currently on-going work is dedicated to re-process the 2023 pilot datasets. We will implement all the successive updates in the ST parameterizations and make a harmonized data set describing the current state-of-the-art NRT source apportionment for the organic aerosol fraction. This will be presented in Deliverable D24 (D4.3) (M40).

5. References

Chen et al. (2022a): *European aerosol phenomenology − 8: Harmonised source apportionment of organic aerosol using 22 Year-long ACSM/AMS datasets.* Environ. Int., 166. https://doi.org/10.1016/j.envint.2022.107325

Chen et al. (2022b): *Real-time source apportionment of organic aerosols in three European cities*. Environ. Sci. Technol., 56, 15290-97. https://doi.org/10.1021/acs.est.2c02509

Sandradewi et al. (2008): *Using aerosol light abosrption measurements for the quantitative determination of wood burning and traffic emission contribution to particulate matter*. Environ. Sci. Technol. 42, 3316–3323. https://doi.org/10.1021/es702253m

Savadkoohi et al. (2024): *Recommendations for reporting equivalent black carbon (eBC) mass concentrations based on long-term pan-European in-situ observations*. Environ. Int. 185. https://doi.org/10.1016/j.envint.2024.108553