

We thank the Reviewer for the detailed and insightful comments and suggestions. The feedback provided, particularly regarding the LC-MS method optimization and semi-quantification approach, helped us enhance the scientific quality of the revised article. Our point-by-point responses to the specific comments, together with the list of changes made in the revised article, are provided below.

Reviewer 2 comments:

Overall, the manuscript is timely and provides an important comparison of multiple NTA workflows. This work will provide a benchmark for others to work towards when undertaking NTA and highlights the type of optimising and testing which should take place. The link with toxicity is a great step forward, especially highlighting the lack of correlation with general aerosol composition (O:C, H:C etc).

However, while I agree with the approach of NTA optimisation and semi-quantification, I have some major comments regarding the LC-MS and semi-quantification methods used. Overall, the manuscript should be published in ACP once comments are addressed.

Specific comments:

Reviewer 2: L125- What mobile phase (A or B) does the 5% apply to?

Author's response: The 5% and other percentage values apply to the mobile phase B component, organic solvent.

Changes in the revised manuscript: This paragraph was revised to clarify that the gradient elution program refers to the percentage of mobile phase component B in the mobile phase “The gradient elution program involved adjusting the amount of eluent B as follows:”

Reviewer 2: L135 – How were the methods optimised? Was it based on a mix of all 59 standards and/or spiked into the sample matrix?

Author's response: The methods were optimized using filter extracts to obtain quality MS/MS spectra by step-wise adjustment of CE at low, medium, and high levels in both ESI polarities. The threshold values in DDA methods were chosen to obtain approximately 4000 MS/MS spectra, based on previously published data reporting the number of unique molecules detected in similar samples. We agree that the optimization of MS and MS/MS conditions wasn't reported in sufficient detail.

Changes in the revised manuscript: An explanation about adjusting MS/MS conditions was added in Section 2.2: “After adjusting the DDA conditions using BrC filter extracts, a quality MS and MS/MS

spectra were obtained for approx. 4000 features, which is comparable with similar, combustion-related samples (Brege et al., 2021; Divisekara et al., 2023; Young et al., 2021; Graham et al., 2002)."

Reviewer 2: L137 – How were the 59 organic molecules determined? Based on previous literature and/or commercial availability? How representative of the overall composition are the standards? Are you biasing towards this functionality?

Author's response: The model compounds were selected to best mimic the molecular composition of biomass-related, water-soluble BrC based on the previously published data and availability of standards. Most of the model compounds are aromatics, functionalized (aromatic)acidic, phenols, or N-containing molecules. We note that our approach may be biased due to the analytical technique employed. Electrospray ionization (ESI) is more selective towards polar compounds, forming $[M+H]^+$ or $[M-H]^-$ ions. Providing further insights into the non-polar fraction of BrC (PAHs, alkanes, alkylated benzenes, etc.) would likely require the use of GC-EI/MS, which was beyond the scope of our work.

Changes in the revised manuscript: Table S1 was added in the SI, with references justifying the use of selected standards as (model)tracers of organic molecules detected in combustion-related emissions. This table is now referenced at the beginning of Section 2.3: "First, the workflows based on MS-FINDER, CFM-ID, Metaboanalyst, GNPS, and MZmine were tested using 59 model compounds. These molecules were selected to mimic the molecular composition of water-soluble BrC emitted by BB based on previously published results and standards availability (Tables S1 and S2), and included derivatives of cinnamic acid, nitrophenols, and polycarboxylic, furoic, and fatty acids."

Reviewer 2: L155/6 – Not entirely sure what this means. How was the structure annotated just based on MS1?

Author's response: The annotation workflow is based on the MS-DIAL scoring approach, which combines MS1 similarity, MS2 similarity, retention time (RT) similarity, and isotope ratio similarity (Tsugawa et al., 2015)

Score_{MS – DIAL}

$$= \frac{MS2\ similarity + MS1\ similarity + RT\ similarity + 0.5\ isotope\ ratio\ similarity}{3.5} \times 100$$

However, in our data, retention time and isotope ratio information were unavailable. Therefore, the scoring function was simplified to include only MS1 and MS2 similarities

$$Score_{this\ work} = \frac{MS2\ similarity + MS1\ similarity}{2} \times 100$$

For cases with MS1-only data (no MS2 spectra acquired or no match found in MS2 library), structure suggestions were based solely on high-resolution mass matching (Tsugawa et al., 2015).

$$Score_{MS1\ only} = MS1\ similarity = \exp \left\{ -0.5 \times \left(\frac{Mass_{sample} - Mass_{library}}{\delta_{mass}} \right)^2 \right\}$$

where δ_{mass} is mass tolerance parameter, set to 0.01 Da in this study.

Thus, the software proposes candidate structures whose theoretical exact masses and isotope distribution profiles in databases match the observed MS1 spectra within the specified tolerance. However, without MS/MS fragmentation data or retention time information, such matches are inherently ambiguous, reflecting lower confidence. “Suggested structures” describe a “grey zone”, where evidence exists for possible structures, but insufficient information is available to achieve a higher confidence level (Schymanski et al., 2014).

Changes in the revised manuscript: To improve clarity, we have included these mathematical formulations and a discussion of the MS1-only annotation method in the SI (a new Section S3.1.5 was added), and it is now referenced in the main text: “The scoring approach implemented in MS-DIAL is described in more detail in Section S3.1.5.”

Reviewer 2: Section 2.5, You have run 59 surrogate standards for library building/identification. When authentic standards were identified (level 1) in your BB samples, did you use the authentic standard for quantification? If so, why not? Could you have assessed the error of the semi-quantification technique based upon comparisons between authentic and surrogate standard quantification?

Author’s response: The quantitative results were further analyzed to assess the difference between quantification results using authentic and surrogate standards. Furthermore, additional calibration was performed for model compounds detected in water-extractable BrC. The quantification biases ranged from -50 to 44% with an average -14 %, consistent with previously published data and with 50% uncertainty assumed for all concentrations derived using surrogate standards, as described in Section 2.5 of the main text.

No.	Ion mode	Compound detected in BrC	Concentration (mg/L)		Bias = $\left(\frac{C1 - C2}{C1}, \% \right)$
			Authentic (C1)	Surrogate (C2)	
1	(-)	Glutaric acid	5.0	7.4	44
2		Coumaric acid	15.1	18.9	25
3		Benzoic acid	2.2	1.1	-50
4		Azelaic acid	12.7	10.5	-17
5		Suberic acid	4.1	2.1	-49
6		Homovanillic acid	3.4	3.1	-8.8

No.	Ion mode	Compound detected in BrC	Concentration (mg/L)		Bias $= (\frac{C1 - C2}{C1}, \%)$
			Authentic (C1)	Surrogate (C2)	
7		Orcinol	3.3	4.0	21
8	(+)	Syringic acid	3.5	1.8	-49
9		Cinnamic acid	3.4	2.1	-38
10		Syringaldehyde	18.6	15.4	-17

At the same time, we believe that providing further insights into the results of semi-quantification (for instance for all surrogate standards) would require a separate, dedicated investigation, as this is a large, complex and dynamically evolving area (Evans et al., 2024; Malm et al., 2021; Kruve, 2019).

Changes in the revised manuscript: A new section was added in the SI, “S2. Semi-quantification using surrogate standards”. The above Table (now Table S4) was added in the SI and is now referenced in Section S2 “This uncertainty is consistent with the quantitative data obtained using authentic and surrogate standards for the compounds detected in water-soluble BrC (Table S4) and higher than the effects of the sample matrix (Table S5).”

Reviewer 2: How were the 5 surrogate standards chosen?

Author’s response: The five surrogate standards (for each ionization mode) were chosen to cover the range of retention times of water-soluble tracers detected in BrC. We prioritized the molecules with good chromatographic properties under the LC/MS analysis conditions used. Initially, a larger number of candidates were tested before narrowing the list to the 5 standards.

Changes in the revised manuscript: A justification for selecting the specific surrogate standards was added in the SI, before Table S3 “The five surrogate standards (for each ionization mode) were chosen to cover the range of retention times of water-soluble organics in BrC. Initial tests were performed on a larger number of molecules, and the list was narrowed down, prioritizing the standards with good chromatographic properties under the LC/MS analysis conditions used – Table S3”.

Reviewer 2: Why does tetradodecanoic acid cover the tr range of above 35 minutes with a retention time of 26 minutes?

Author’s response: This is a typing error.

Changes in the revised manuscript: The t_r for tetradecanedioic acid in Table S3 (formerly Table S2) was corrected to 39.33 min.

Reviewer 2: Why weren't all 59 standards used for semi-quantification? Your range used by each standard is large; including more standards would have narrowed the windows.

Author's response: We agree that using more standards could yield more accurate results. At the same time, large uncertainties are an inherent feature of semi-quantification with surrogate standards. As such, we believe that in the context of our work, using more surrogate standards wasn't critical, as we wanted to obtain an estimate of the amount of water-soluble BrC using the LC/MS data. Even if the quantification results from the three methods agree very well (as presented in Fig. 7 in the main text), the LC/MS data is still implied to a large uncertainty, as underlined in Section 3.3.

Changes in the revised manuscript: Description in Section 2.5 was revised as "Five surrogate standards were used in each ionization mode, with retention times from 2 to 40 min (Table S3), covering the elution window for WSOCs detected in BrC_{aq}." to underline that the retention times of surrogate standards used converted the retention times of the investigated analytes.

Reviewer 2: Did you spike the standards into the sample matrix to determine matrix effects on concentrations, given the sample complexity?

Author's response: Additional analyses were performed to analyze the matrix effects, particularly the impact on ionization efficiency and detection sensitivity (Williams et al., 2023). Please note also that the filter extracts were diluted 10 times before LC/MS analyses to reduce matrix effects (Pihlström et al., 2021) and avoid instrument contamination.

The matrix effect for each analyte was evaluated by comparing the analyte response (peak area) in the diluted sample extract with that in a pure solvent standard, and calculated using the following equation

$$\% \text{ Matrix Effect} = \frac{(\text{Peak area}_{\text{sample}} - \text{Peak area}_{\text{solvent}})}{\text{Peak area}_{\text{solvent}}} * 100$$

Concentration (mg/L)	Negative ion mode			Positive ion mode			
	Adipic acid	Caffeic acid	Sebacic acid	Tryptophan	Isatin	Fenuron	Isoproturon
1	19	10	9.2	-20	-40	-20	-8.6
2	8.4	2.4	3.9	-13	-32	-19	-7.5
5	1.4	6.9	4.6	-8.2	-27	-20	-9.2
10	6.9	3.4	3.5	-8.8	-22	-14	-8.4
20	2.0	9.3	8.4	-11	-13	-11	-7.2

The above Table summarizes the matrix effect values for surrogate compounds measured in both negative and positive ionization modes across different concentration levels (1-20 mg/L). The results

indicate that the matrix effect for all compounds remained below 20%, except isatin, which exhibited higher suppression ranging from -13% to -40%. According to the criteria defined in SANTE/11312/2021v2, these results indicate that matrix effects were generally within acceptable limits for most analytes in this workflow (Pihlström et al., 2021). Furthermore, matrix effects were lower than the uncertainty resulting from the use of surrogate standards.

Changes in the revised manuscript: The above discussion, Table S5 were added in a new Section S2. “Semi-quantification using surrogate standards” in the SI.

Line was added in Section 2.1: “The extracts were diluted tenfold before LC/MS analysis to prevent ion source contamination and minimize matrix effects on the semi-quantitative results (see Section 2.5).”

Lines were added in Section 2.5, regarding the use of surrogate standards and the results of semi-quantitative analyses: “For compounds detected in BrC_{aq}, this value is consistent with the results obtained using authentic and surrogate standards (Table S4). Furthermore, the imposed 50% uncertainty exceeds the matrix effects observed for most surrogate standards (Table S5). To prevent introducing a positive bias into the data, molecules detected in both ionization modes were identified by comparing annotation results, and their average concentrations were used to obtain the amount of BrC_{aq}.”

Reviewer 2: Section 3- Many of the analytes will ionise in both the negative and positive mode. How were these species dealt with in terms of final concentrations and composition?

Author’s response: We agree that detecting the same molecules in both modes would lead to a positive bias in the semi-quantification results. The molecules detected in both modes were identified by comparing the annotation results, and the average concentrations were used for such compounds.

Changes in the revised manuscript: A clarification was added in Section 2.5. “To prevent introducing a positive bias into the quantitative data, molecules detected in both ionization modes were identified by comparing annotation results, and their average concentrations were used to obtain the amount of BrC_{aq}.”

Reviewer 2: L235 – What sample did you use to determine this? I.e a mixture of standards/spiked sample

Author’s response: The workflows were assessed using a standard mixture of 59 model compounds.

Changes in the revised manuscript: The caption of Figure 5 was revised as “Workflow performance evaluation with the standard mixture of 59 model compounds (Table S1). Number of correctly identified analytes (A) and number of analytes identified only by a given software (B)”

Reviewer 2: L239 The use of “identifies” here makes it seem like this is a known compound, maybe change to annotated?

Author’s response: We agree that the term “annotated” is more appropriate in the context of NTA.

Changes in the revised manuscript: This sentence was revised as “The combination of MS-DIAL and MS-FINDER also annotated 5 unique features not recognized by any other workflow, whereas MZmine annotated only one unique analyte – Fig. 5.”

Reviewer 2: Fig 6 – How representative are your standards to the overall detected composition?

Author’s response: Naturally, it is challenging to cover the molecular diversity of BrC using commercially available standards. At the same time, the values of Kendrick Mass and Kendrick Mass Defect, double-bond equivalent, O/C and H/C ratios, C+N, and C numbers for the selected standards are all within the range of the values presented in Fig. 6 for the main components (represented by the largest cycles) of water-extractable BrC.

Changes in the revised manuscript: Table S2 was added in the SI, reporting the values of the parameters presented in Fig. 6 and elemental composition of all 59 model compounds, including an average and two standard deviation values for each parameter.

A paragraph was added in the SI, before Table S2: “The general characteristics of surrogate standards used to evaluate the workflows were derived as described in Section S4 and are listed in Table S2. Covering the high molecular diversity of the water-extractable BrC with commercially available standards is challenging. At the same time, values of Kendrick Mass and Kendrick Mass Defect, double-bond equivalent, O/C and H/C ratios, C+N, and C numbers for the selected standards are all within the range of the values presented in Fig. 6 in the main text, for the main components of water-extractable BrC.”

Line was added in the caption of Fig. 6 in the main text: “These general characteristics (Section S4) of water-extractable BrC were comparable to those of the surrogate standards used to evaluate the NTA workflows (Table S2).”

Table S2 is now referenced in Section 2.3.

Cited references

Brege, M. A., China, S., Schum, S., Zelenyuk, A., and Mazzoleni, L. R.: Extreme Molecular Complexity Resulting in a Continuum of Carbonaceous Species in Biomass Burning Tar Balls from Wildfire Smoke, *ACS Earth and Space Chemistry*, 5, 2729-2739, 10.1021/acsearthspacechem.1c00141, 2021.

Divisekara, T., Schum, S., and Mazzoleni, L.: Ultrahigh performance LC/FT-MS non-targeted screening for biomass burning organic aerosol with MZmine2 and MFAssignR, *Chemosphere*, 338, 139403, <https://doi.org/10.1016/j.chemosphere.2023.139403>, 2023.

Evans, R. L., Bryant, D. J., Voliotis, A., Hu, D., Wu, H., Syafira, S. A., Oghama, O. E., McFiggans, G., Hamilton, J. F., and Rickard, A. R.: A Semi-Quantitative Approach to

Nontarget Compositional Analysis of Complex Samples, *Anal. Chem.*, 96, 18349-18358, 10.1021/acs.analchem.4c00819, 2024.

Graham, B., Mayol-Bracero, O. L., Guyon, P., Roberts, G. C., Decesari, S., Facchini, M. C., Artaxo, P., Maenhaut, W., Köll, P., and Andreae, M. O.: Water-soluble organic compounds in biomass burning aerosols over Amazonia 1. Characterization by NMR and GC-MS, *Journal of Geophysical Research: Atmospheres*, 107, LBA 14-11-LBA 14-16, <https://doi.org/10.1029/2001JD000336>, 2002.

Kruve, A.: Semi-quantitative non-target analysis of water with liquid chromatography/high-resolution mass spectrometry: How far are we?, *Rapid Commun Mass Spectrom*, 33, 54-63, 2019.

Malm, L., Palm, E., Souihi, A., Plassmann, M., Liigand, J., and Kruve, A.: Guide to Semi-Quantitative Non-Targeted Screening Using LC/ESI/HRMS, *Molecules*, 26, 10.3390/molecules26123524, 2021.

Pihlström, T., Fernández-Alba, A. R., Amate, C. F., Poulsen, M. E., Hardebusch, B., Anastassiades, M., Lippold, R., Cabrera, L. C., de Kok, A., and O'Regan, F.: Analytical quality control and method validation procedures for pesticide residues analysis in food and feed SANTE 11312/2021, *Sante*, 11312, 2021.

Schymanski, E. L., Jeon, J., Gulde, R., Fenner, K., Ruff, M., Singer, H. P., and Hollender, J.: Identifying Small Molecules via High Resolution Mass Spectrometry: Communicating Confidence, *Environmental Science & Technology*, 48, 2097-2098, 10.1021/es5002105, 2014.

Tsugawa, H., Cajka, T., Kind, T., Ma, Y., Higgins, B., Ikeda, K., Kanazawa, M., VanderGheynst, J., Fiehn, O., and Arita, M.: MS-DIAL: data-independent MS/MS deconvolution for comprehensive metabolome analysis, *Nat Methods*, 12, 523-526, 10.1038/nmeth.3393, 2015.

Williams, M. L., Olomukoro, A. A., Emmons, R. V., Godage, N. H., and Gionfriddo, E.: Matrix effects demystified: Strategies for resolving challenges in analytical separations of complex samples, *Journal of Separation Science*, 46, 2300571, 2023.

Young, T. M., Black, G. P., Wong, L., Bloszies, C. S., Fiehn, O., He, G., Denison, M. S., Vogel, C. F. A., and Durbin-Johnson, B.: Identifying Toxicologically Significant Compounds in Urban Wildfire Ash Using In Vitro Bioassays and High-Resolution Mass Spectrometry, *Environmental Science & Technology*, 55, 3657-3667, 10.1021/acs.est.0c06712, 2021.