Dear reviewer 2,

Thank you for the detailed review. We have carefully considered each statement and evaluated whether the manuscript would benefit from the reviewers suggestion. We think that the overall manuscript improved. Please find the point-by-point response below. We have written the reviewers´ comments in **black**, our response in **blue** and corrections in the manuscript in **red**.

Reviewer 2

Review of Karbach et al. (2026): Non-Target Analysis of Atmospheric Organic Aerosols as a Tool to Discriminate Anthropogenic Contribution in Mixed Air Masses during the ACROSS campaign

Overall Assessment

This manuscript presents a wind-informed non-target LC-HRMS analysis of organic aerosol collected during the ACROSS 2022 campaign and applies a K-means clustering framework to organize molecular signatures associated with anthropogenic and biogenic source influences. However, in its current form, the manuscript requires substantial revision before it is suitable for publication in ACP. While the results are promising, the manuscript currently sits somewhat between a methodological contribution and a measurement/report study, without fully developing either aspect.

If the primary contribution is methodological, additional analyses are needed to demonstrate the specific advantages of the proposed wind-informed clustering framework relative to simpler approaches, such as direct enrichment-based classification or clustering performed without wind-based source partitioning. In particular, it remains unclear what information is uniquely gained from the K-means step, given that source-associated behavior is already quantified through the wind-based feature construction. The manuscript would benefit from a clearer discussion of the role of clustering in the overall workflow and evidence demonstrating its added value relative to alternative approaches. Furthermore, while the reduced-sample clustering exercise suggests internal consistency of the derived cluster assignments, it does not constitute an independent validation of the framework. Consequently, claims regarding the applicability of the method to determine cluster contributions in “truly unknown organic aerosol samples” appear overstated and would require validation using independent datasets or a more rigorous predictive evaluation strategy.

Alternatively, if the primary contribution is the chemical characterization of anthropogenic and biogenic molecular signatures, the discussion would benefit from a more comprehensive comparison with published ACROSS studies. In particular, the manuscript does not discuss results from Pereira et al., 2025 (https://doi.org/10.5194/acp-25-4885-2025), which reports LC-HRMS measurements conducted during the same campaign and sampling periods, including observations at both the suburban-forest site and the Paris site that is interpreted here as

the dominant anthropogenic source region. Such comparisons could provide important validation of the cluster assignments, place the findings within the broader ACROSS context, and substantially strengthen the atmospheric interpretation of the identified molecular signatures. Given the regional nature of the dataset, these comparisons are also important for demonstrating how the reported findings advance or extend existing knowledge, which is a key consideration for ACP studies with a predominantly local focus.

As reviewer 2 suggested, we have tried to clarify the primary contribution of this manuscript. Our initial goal was the chemical characterization of anthropogenic and biogenic (more accurately: regional background; see later comments) signatures. Due to the limited size of the dataset with only 37 samples, we could not employ a typical clustering method as e.g. described in Thoma et. al. 2025. We therefore resorted to a more simplistic clustering algorithm using K-means clustering. However, the main goal of this work remained the chemical characterization of the two different signatures.

We understand the comment of Reviewer 2 and have tried to clarify this in the manuscript. We have added a comparison with Pereira et al., 2025. As suggested, this addition should place our findings in a broader context and validate our results. The comparison is placed in a separate chapter directly behind the clustering results.

Pereira et al. (2025) provide a complementary molecular characterization of aerosol samples collected during the same ACROSS campaign. During the background period on 3 and 4 July, they observed fewer CHON and aromatic compounds at Rambouillet and identified CHO formulas which are compatible with alpha-pinene oxidation products. This agrees with the CHO-rich regional-background clusters identified in this work, which show an accumulation at H/C = 1.2-1.7 and O/C = 0.15-0.5, consistent with oxidized terpene products. In addition, Pereira et al. found that polluted periods were associated with northeastern air masses crossing Paris and with increasingly similar aerosol compositions at the urban and forest sites, indicating transport of anthropogenic influence to Rambouillet. They additionally detected the organosulfur compounds attributed to alpha-pinene oxidation in the presence of SO_2. These observations are consistent with the enrichment of CHOS compounds and nitro, sulfonic, and aromatic functionalities in the anthropogenic cluster identified in the present study. The agreement between the independently processed datasets supports the atmospheric interpretation of the clusters and their attribution to regional-background and Paris-influenced conditions.

Specific Comments

Scientific Language and Organization

The manuscript would benefit from careful editing to improve clarity, consistency, and precision throughout. The examples provided below are illustrative rather than exhaustive, and a thorough review of the manuscript is recommended with these issues in mind. Sentences such as “Mixed anthropogenic biogenic environments add a layer:” (line 72); “…which prioritizes abundant precursors however.”(line 92); “After exporting the data, the blank was subtracted

from the data three times…" (line 140); and "The concentration of those compounds is strongly increased…" (line 327); could be rewritten for clarity and scientific formality. Many claims would benefit from literature citations. For instance, lines 70-79 discuss organic markers in atmospheric aerosols but without any references. Several acronyms and abbreviations are introduced without prior definition (e.g., lines 46, 58, 66, 67, 85, etc.), and terminology is not always used consistently. For example, non-anthropogenic sources are variously described as biogenic, natural, and mixed, while sampling periods are alternately referred to as day/night and daytime/nighttime. Figure references are also inconsistent, with panels labeled alphabetically in the figures but described as left and right in the text.

We have revised language to use a more formal tone. All adaptions can be found in the track-changes file.

The sample nomenclature is also difficult to follow. There is currently no clear reference linking the filter identifiers presented in Table S1 and the backward trajectory figures in the Supplement to the sample identifiers (S1, S2, etc.) used throughout the manuscript. It is therefore unclear whether F1 corresponds directly to S1 or whether the samples were renumbered after the exclusion of six filters. This ambiguity becomes particularly apparent when discussing time series, where the x-axis extends to 43 despite only 37 samples being included in the analysis. I recommend using a single, consistent naming convention (e.g., "samples" or "filters") throughout the manuscript and clearly explaining in the main text how the excluded filters were handled.

We agree with the reviewer. Two different sample identifiers (S for sample and F for filter) could be difficult to follow. We have adjusted the sample identifiers in the SI. All samples are now labeled S1, S2, S3, … throughout the manuscript, field blanks retain the prefix B. In the first version of the manuscript, S1 did directly correspond to F1, so nothing in the data or the data interpretation has changed.

Adapted sample nomenclature in the SI.

In addition, Section 2.1 is titled Sample Preparation but also includes sample collection procedures. Consider revising the section title or separating collection and preparation into distinct subsections.

We agree with the reviewer and think that the heading should be adapted. We therefore changed the title to "Sample collection and preparation".

Changed title to: Sample collection and preparation

More importantly, most statements describing differences between groups remain qualitative. Where possible, differences should be quantified and accompanied by appropriate statistical metrics. Similarly, statements regarding significance, enrichment, or separation should be supported by suitable statistical analyses rather than qualitative descriptions alone.

We agree that terminology implying statistical significance should not be used without formal testing. Rather than adding a broad inferential analysis that is not supported by the present study design, we revised the manuscript to describe these differences as qualitative or semi-quantitative observations. We also clarify that the bootstrap analysis assesses internal cluster-assignment stability rather than independent validation.

The introduction in its current form is overly descriptive and method-oriented and does not yet clearly establish a focused scientific hypothesis, knowledge gap, or study objective. As a result, the reader is left uncertain about the central scientific question being addressed. Several sections read more as methodological background than scientific motivation, which weakens the conceptual framing of the study. In addition, the connection between the ACROSS campaign and the specific objectives of the manuscript is not sufficiently explicit, making the scientific purpose of the work somewhat diffuse. The introduction is also organized as a series of largely independent thematic sections rather than a coherent narrative that progressively develops toward a clearly articulated research question. Consequently, the broader atmospheric significance, novelty, and expected scientific advances of the study are not sufficiently evident from the introduction.

Here we do not agree with the reviewer. We do not think that an introduction should not be descriptive and method oriented. We think that the main purpose of an introduction should be to introduce the main topic to readers so that some fundamental parts are explained and allow new researchers and researchers from other specialities to be able to understand the manuscript and underlying methods used. We understand however, that the objectives of the study can be more clearly established in the introduction. As suggested by the reviewer, we have therefore added a chapter where we addressed the “broader atmospheric significance, novelty and expected scientific advances” of the study.

Scientific objectives and atmospheric relevance

Mixed urban-forest environments represent a particular challenge for molecular source attribution because anthropogenic pollutants and biogenic precursors interact during transport and produce secondary compounds that cannot be assigned to isolated source categories solely from their elemental composition. The ACROSS campaign was designed to investigate such interactions by examining the transport of emissions from Greater Paris into the surrounding vegetated region, with the Rambouillet forest providing a receptor site that is alternately exposed to Paris-influenced and regional-background air masses. Within this framework, the central question of the present study is whether these contrasting air-mass conditions produce distinguishable and chemically interpretable molecular signatures in organic aerosol collected at a single receptor site. We hypothesize that compounds showing consistent enrichment

during air-mass transport from Greater Paris differ systematically in elemental composition, oxidation state, and fragmentation characteristics from compounds associated with other transport directions. To test this hypothesis, LC-Orbitrap MS non-target analysis is combined with HYSPLIT back trajectories, wind-correlation filtering, clustering, and fragment-based characterization of selected marker compounds. The novelty of the study lies in resolving source-associated molecular groups from a comparatively small set of 37 field samples while retaining compound-level chemical information rather than relying only on bulk aerosol properties or predefined tracers. The expected scientific advance is the characterization of regional-background and Paris-influenced signatures, including molecular evidence for the anthropogenic modification of biogenic secondary organic aerosol through nitrogen- and sulfur-containing products. More broadly, this work contributes to understanding how megacity emissions alter organic aerosol composition beyond the urban boundary and provides a basis for interpreting source contributions in other mixed anthropogenic-biogenic environments.

The methodological description is incomplete and not presented in a sufficiently sequential manner. For example, the exact coordinates of the sampling site are not reported; the rationale for the OA molecular characterization approach and clustering framework, including the basis for selecting the number of clusters, is not clearly defined; the directional-dominance metric is introduced later without a clear description of its calculation; and the bootstrap validation approach is insufficiently described, with key details such as the resampling strategy and number of iterations omitted. As a result, the workflow cannot be readily reproduced by an independent researcher based on the current description.

We agree with the reviewer and, also in accordance with reviewer 1, have added chapters to further explain the methods used in this work. This includes chapter “2.5 Determining the wind directionality of individual compounds”, additions to chapter “2.6 Apportionment of the sum formulas origin” where it is explained, why three clusters were chosen, and a complete chapter “2.6.1 Clustering naming conventions” where the naming conventions used in this work are explained. A “Bootstrapping” chapter was also added where the verification workflow is explained in more detail.

“Determining the wind directionality of individual compounds

In order to perform filtering on the compounds that undergo clustering, a wind directionality is calculated. The wind directionality is the fraction of the total concentration of a given compound found within the dominant 45° direction bin and its two neighboring bins. A value near 1 means most of the compound originates from only one specific wind direction, while a value < 0.4 means the concentration is widely spread across different directions. Wind directionality measures directional consistency. During the ACROSS campaign, the wind direction measured at the sampling site, at the time of sampling usually agrees with the back trajectory as determined by the HYSPLIT model for the sampled air mass. In order to reduce complexity of the analysis, the wind direction as was measured at the sampling site, during the sampling period was used to determine the wind directionality.

Apportionment of the sum formulas origin

The wind data, determined as described in chapter 2.4, is used to assign sources to individual compounds, where compounds whose intensity peaks during N-NE (red shaded areas) influenced wind conditions likely originate from the greater Paris area and are therefore likely to be more anthropogenically influenced, whereas samples taken when the airmass was coming from the other directions are, in this study considered to be less anthropogenically influenced (green shaded areas). As Paris is the largest anthropogenic accumulation in France, the air masses that were passing over Paris are considered to be more anthropogenically influenced than samples not passing over Paris, creating two types of samples: one type of sample that experienced a higher anthropogenic activity along its trajectory and one that experienced lower anthropogenic activity. For the sake of this paper, the two types are called anthropogenic influenced and regional background air mass.A Python script was written to automatically perform the clustering based on wind direction (more specifically naturally influenced air mass or anthropogenic influenced air mass) and compound intensity.

Peak areas per assigned molecular formula were normalized to sampled air volume (to account for different sampling durations) and to total signal per sample (to account for different total particle concentrations, e.g. due to rain). The normalization helps to remove meteorological influences in the dataset and keep directional effects preserved, generally improving directional clustering results. Each detected compound´s normalized intensity profile was partitioned into user labeled natural and anthropogenic periods, and, the mean normalized contribution during natural and anthropogenic periods and their difference (natural minus anthropogenic) was calculated. Compounds with insufficient summed signal, or when more than 15 samples were below detection limit, were excluded. The compounds were then clustered with K-means (k=3, random_state=0). After initial testing with only two clusters, it was found that in order to improve robustness, allowing a third cluster for compounds without a strong wind directionality improved results. Cluster membership was interpreted as source related groups with cluster 1 and 2 being the regional background, and cluster 3 being the anthropogenic cluster. Compounds within the regional background clusters with weak wind directionality were assigned to cluster 2 and compounds with strong wind directionality were assigned to cluster 1. This can be seen in the results in Figure 7.

Clustering naming conventions

For the sake of this work, time periods where the air masses have passed over the greater Paris area are called anthropogenically influenced, and consequently, compounds that show the highest concentration during those time periods are called anthropogenic compounds. The other time periods, where the air did not pass over the greater Paris area are called regional background time periods. The term regional background was chosen, as for those time periods, the air did not experience the strong anthropogenic influence of Paris, but some anthropogenic influence cannot be excluded, as other cities and anthropogenic agglomerations can be found along the trajectories of those air masses. Those time periods represent a mixture of

biogenic and anthropogenic influences. Compounds that do show a higher concentration during those time periods are called regional background compounds.

Bootstrapping

Cluster-assignment stability was evaluated using a compound-level bootstrap applied to the three standardized clustering features: the mean contribution during regional-background periods, the mean contribution during Paris-influenced periods, and their difference; non-finite values were replaced by feature-wise medians. The original K-means solution obtained with "random_state = 0" served as the reference assignment. A total of 1000 bootstrap replicates were generated using a random seed of 42, with 80 % of the compounds sampled with replacement in each replicate. A new standardization and K-means model were fitted to every resampled dataset and subsequently used to predict the cluster membership of all compounds. The resulting labels were aligned to the reference solution using the Hungarian algorithm to account for the arbitrary numbering of K-means clusters. Only out-of-bag predictions were used to calculate compound-specific membership probabilities, with stability defined as the probability of reassignment to the reference cluster and uncertainty quantified by normalized Shannon entropy. The procedure was performed for both the three-cluster solution and the two-cluster control analysis and assesses internal assignment reproducibility rather than validation against an independent dataset. "

Finally, the selection of compounds highlighted for detailed discussion requires clarification. Given the large feature space (1894 biogenic and 1023 anthropogenic compounds), it is unclear on what basis four anthropogenic and one biogenic compound were chosen for further interpretation. Without a clearly defined and objective selection criterion (e.g., cluster membership strength, enrichment metrics, abundance, occurrence frequency, or statistical ranking), it is difficult to assess how representative these examples are of their respective clusters. As currently presented, the selection appears subjective and may bias the discussion toward compounds that are more readily interpretable from a chemical perspective. Providing transparent selection criteria or a more systematic summary of cluster composition would improve reproducibility and strengthen confidence in the atmospheric interpretation of the identified molecular signatures.

We understand the concerns of reviewer 2. The example compounds chosen in chapter 3.3.2 and 3.3.3 were indeed "subjectively" chosen to visualize key concepts and applications of the resulting dataset. The examples should not act as validation but rather show an example of how the researcher can use the generated data to derive further conclusions and come to new findings. The three compounds shown as anthropogenic examples all have a similar structure, as could be shown by the fragmentation experiments, additionally, they all differ by only a R-CH2 group suggesting a homologous series and they all experience a similar time series pattern. These results clearly show the relation of the three individual compounds and highlight the anthropogenic significance. In this example, the significance is also supported by literature as could be shown by the various references provided in the manuscript.

Presentation and Figures

Many figures and figure legends would benefit from revision. Font sizes are inconsistent across figures and within individual panels, which reduces readability. A more uniform graphical style throughout the manuscript would improve clarity and visual consistency.

We chose to use different font sizes in the figures to order information based on importance - with the largest font size being the most important. Using a consistent font size throughout the figures, would have resulted in either having too much text in the figures, or leaving out information that might be relevant for the reader. As ACP is an online journal, the reader is easily able to zoom into graphics to more easily read smaller fontsizes within the figures. We do understand however, that for printed journals, a consistent font size is more important for easier readability. Where different font sizes were retained to avoid crowding, the corresponding information is now explained in the caption rather than relying on visual prominence alone.

A map showing the study region and sampling location would be helpful to provide geographical context for readers unfamiliar with the ACROSS campaign domain.

We agree that additional geographical context is helpful. We added the exact sampling-site coordinates, clarified that the star marks the sampling site, and described the position of Greater Paris relative to the site in the Figure 1 caption. Because Figure 1 and the HYSPLIT panels already show the regional setting, we retained the existing figure rather than adding a separate map.

Figure 1. It would be helpful to add cardinal directions (e.g., N, NW, S, SW) alongside the degree scale on the y-axis, as wind direction is referred to in this format throughout the text rather than in degrees. In addition, the HYSPLIT trajectory plots are difficult to interpret due to small font sizes and limited geographical context. The sampling location is not immediately identifiable, which makes it challenging to assess air mass origins. It may be helpful to include a marker for the sampling site, an outline of the Greater Paris region, and a marker indicating the location of Paris.

We have added the suggested cardinal directions (right y-axis), as well as a marker for the greater Paris area in the figure (red cross).

Figure 2. It may improve comparability to use a common y-axis scale for panels (a) and (b). In addition, a more chemically resolved representation (e.g., stacked contributions of CHO, CHON, CHOS, CHONS, and related classes) could provide additional insight into how compound classes are distributed across m/z and retention time.

This is beyond the scope of this manuscript. The two plots in Figure 2 should give a first impression about the underlying dataset. Further and more in-detail analysis is done in the results chapter of the manuscript. The panels retain independently optimized y-axis limits because they display different distributions; the chemically resolved composition is presented in the subsequent results figures.

Figure 3. The Van Krevelen diagram is difficult to interpret due to the large legend occupying substantial plot space. The marker size is not defined, and the marginal distributions along the axes are not clearly explained. It may be useful to separate molecular formula classes into individual Van Krevelen plots to improve interpretability. The equation used for OSc should also be provided in the main text, figure caption, or Supplement.

During initial writing we considered separating the molecular-formula classes, but retained the combined plot because Figure 3 is intended to summarize the full chemical space before class-resolved interpretation in the subsequent figures. To address the reviewer's concern, we expanded the caption and removed the description of the shaded regions from the legend. The description is written in detail in the caption.

We think that calculating the carbon oxidation state should be sufficiently known to chemists. A description of the method can however still be found in the reference provided in the figure caption “Kroll et al., 2011”.

Figures 3–6, 11. The acronym MCR is not defined in the manuscript, although the corresponding regions shown in these figures appear to be defined. These regions are not discussed in the main text. The authors may wish to clarify their meaning and relevance, or consider whether they are necessary if they are not directly used in the interpretation of the results.

Thank you for identifying this inconsistency. We defined maximum carbonyl ratio (MCR) in the relevant captions and clarified that the shaded regions provide a common visual reference across different Van Krevelen plots; they are not used as an independent classification criterion.

Figure 6. This figure is difficult to interpret due to relatively small panels, overlapping legends, and elements that are not fully described (e.g., clustering structure and axis subscripts).

The discussion in Section 3.3.1 also appears somewhat limited relative to the complexity of the figure and could be expanded to better support interpretation.

We agree that Figure 6 contains substantial information. We expanded Section 3.3.1 and the caption to define the bootstrap stability metric, cluster labels, and the relation between the three- and two-cluster solutions. We also clarified that the analysis assesses internal assignment stability rather than predictive validation.

"Bootstrapping was performed to assess internal cluster-assignment stability. The analysis indicated that the assignments were internally stable under the applied resampling procedure, with only a small number of compounds exhibiting an alternative cluster assignment in some bootstrap iterations (Fehler: Verweis nicht gefunden). "

Figure 7. It may improve interpretability to harmonize panel formatting using a common y-axis scale.

Due to different cluster sizes it is not applicable to use a common y-axis scale without.

A chemically resolved presentation of the directional-dominance analysis (e.g., by molecular formula class) could help identify which compound classes contribute most strongly to the observed patterns.

This is beyond the scope of the manuscript.

Figure 9. This figure is not explicitly referenced in the main text. If retained, it may be helpful to clearly define terms such as "XIC" and distinguish between identified peaks and observed time series. Improvements in font size and color contrast could also enhance readability.

We added an explicit reference to Figure 9 in Section 3.3.2 and expanded XIC as extracted-ion chromatogram in the caption.

Figure 10. It may be helpful to align the structure of this figure with Figure 9 to facilitate comparison.

There is no benefit in comparing Figure 9 and Figure 10. They are different compounds from different groups. The resemblance of the two figures is due to the fact, that they were created using the same Python workflow.

Figure 11. It may improve clarity to separate the HYSPLIT trajectory plots from the Van Krevelen plots, as combining them contributes to visual congestion.

Removing the HYSPLIT trajectory from the plot would be possible, as this plot is also contained in the SI. However, here it was chosen to include a thumbnail of the trajectory of the individual sample to quickly visualize the differences in trajectories and facilitate an easier interpretation of the data. We do not think that this congests the plot, but rather makes interpretation easier for the reader.

Line Comments

Line 11: The phrase "particle phase of Earth's atmosphere" is misleading, as particles are suspended within the atmospheric gas phase rather than constituting a phase of the atmosphere itself. Consider revising to clarify that organic aerosol is a major component of atmospheric particulate matter.

We agree that the sentence structure can be misleading for some readers, hence, we adapted the sentence:

"The organic fraction is a major component of atmospheric fine particulate matter and .... "

Line 34: Consider broadening the end of this sentence. Restricting the toxic components of OA to "combustion-related constituents" overlooks a wide range of other health-adverse organic species that comprise the OA mass.

We have made it more clear that redox active substances and combustion related constituents are just to be understood as examples.

"OA is also linked to adverse health outcomes because fine particles penetrate deep into the respiratory tract and can carry a variety of health adverse organic species like for example *redox active organic functionality and toxic combustion related constituents (Nel, 2005; Pope et al., 2009)."*

Line 36: The sentence says OA is a source of uncertainty because it is complex and dynamically evolving. But the complexity itself isn't what causes the uncertainty; it's our incomplete understanding and poor modeling of that complexity. OA would still be complex even if we predicted it perfectly. Consider rephrasing to focus on these modeling and observational constraints, and add a few representative references to back it up.

We agree that the uncertainty arises from limitations in observation and modeling rather than from complexity alone. We revised the sentence accordingly while keeping the Introduction concise.

"Despite decades of research, OA remains a major source of uncertainty in air-quality prediction and aerosol radiative forcing because current observational and modeling approaches do not fully resolve its chemically complex, multiphase, and dynamically evolving composition."

Line 39: Instead of listing specific sources for POA, consider grouping them more broadly into anthropogenic and biogenic categories. Because new OA sources are continuously being identified, summarizing them this way will improve the long-term relevance of the text.

We have added the two classes anthropogenic and biogenic to the sentence.

"*Ambient OA comprises primary organic aerosol (POA) emitted* directly by either anthropogenic sources like *traffic, biomass burning, cooking, and industrial sources*, and by biogenic processes like fungal spore emissions. *Secondary organic aerosol (SOA) is formed in the atmosphere from oxidation of volatile, intermediate volatility, and semi volatile organic compounds (VOCs, IVOCs, SVOCs) followed by gas particle partitioning and particle phase chemistry (Gouw and Jimenez, 2009)."*

Lines 43-48: The paragraph sets up a strict dichotomy where high temporal resolution is limited to bulk composition (e.g., AMS) and molecular speciation is restricted to offline filter analysis. This characterization overlooks significant advances in state-of-the-art in situ instrumentation. Advanced online platforms, such as the Thermal Desorption Aerosol Gas Chromatograph (TAG), bridge this gap by providing molecularly speciated OA measurements at hourly resolution or better. Additionally, derivatization methods and online two-dimensional GC-MS aerosol system (i.e., 2D-Q-TAG, VAPS) allow for the measurement of polar species present in OA. Consider revising this section to acknowledge these advanced online speciation techniques, as the current text presents an overly simplified view of modern aerosol analysis capabilities.

As was already written above, that the introduction is already too descriptive, we have chosen not to go into more detail here. There certainly are advances made in the field of online speciation techniques, however it is not within the scope of this manuscript to detail those new techniques.

Line 54: Consider omitting the description of the NTA workflow or move it to methods as it does not add significance in the introduction.

We agree with the reviewer and have removed the detailed description of the workflow at this part of the introduction. The used workflow is still contained in the Experimental section of the work.

Line 55: The statement "Because OA datasets are high dimensional…" is overly broad. OA datasets are not necessarily high dimensional (e.g., bulk OA mass measurements), although chemically resolved datasets are often highly dimensional. Please revise to reflect this distinction.

This is a good catch. We have added:

"*Because* chemically resolved *OA datasets are high dimensional and sample series reflect...*"

Line 62: Please acknowledge the campaign coordinators in the Acknowledgements section.

Thank you. We have removed the coordinators at this point. They were already in the Acknowledgements section.

Line 113: Please change the date format to the Copernicus standard: " dd month yyyy".

We changed the dates to the suggested Copernicus format.

"Atmospheric organic aerosol were sampled during the ACROSS campaign from 28 June 2022 *to the* 19 July 2022 *at the ACROSS sampling site..."*

Line 200: It is unclear how nitrogen-containing compounds are included in the CHOS subset. Since nitrogen is not part of CHOS, please clarify whether CHON or CHONS compounds are intended here.

We understand the concerns and agree with reviewer 2. The sentence could lead to a misunderstanding. What we meant is that within the anthropogenically influenced air masses, nitrogen containing compounds (CHON subset) and sulfur containing compounds (CHOS subset) are both more pronounced than compared to the regional background. In our dataset the CHOS subset is even stronger pronounced than the CHON subset. We have removed the reference to the CHOS subset as this might lead to confusion and the meaning of the sentence can be understood without it.

"*Later in this study it is shown, that the sulfur* containing compounds (CHOS subset) *and* the *nitrogen containing compounds* (CHON subset) *are* more pronounced *in anthropogenically influenced air masses* as compared to regional background. As the air was sometimes strongly anthropogenically influenced this *explains the observed high abundances in the CHOS and CHON subset (e.g. (Rincón et al., 2012; Wang et al., 2017)).*"

Line 207: The inference of a temporal evolution ("increasing time in the atmosphere") from the Kroll plot is not justified, as the figure does not track air mass aging over time but instead shows a cross-sectional distribution of compounds with different oxidation states.

We understand that this might be misleading for some readers. While with increasing time in the atmosphere compounds do generally move upwards and towards the right in the Kroll plot, due to functionalization (upwards movement) and fragmentation (towards the right), we still understand that at this point, this information can be omitted.

"*The Kroll plot shown in Figure 3 (right (Kroll et al., 2011)) highlights the evolution of organics in the atmosphere, with fragmentation, functionalization and oligomerization, leading to the typically shaped curve. Oxidation of organics leads to functional groups being introduced into the molecule, ultimately increasing COS (vertical movement). When oxidation progresses, molecules ultimately lose parts of its structure (fragmentation), leading to horizontal movement in the right direction.* These processes explain the curved shape which is typical for mixed air masses and can also be seen in our dataset. "

Line 226: The interpretation that differences between samples S6 and S22 can "solely be attributed to different meteorological conditions" is too strong. While wind direction suggests different air mass origins (W–SW vs NE), other factors such as chemical aging, boundary layer dynamics, and regional mixing may also contribute to the observed differences. The attribution should therefore be softened to reflect these additional uncertainties.

We agree with reviewer 2. We have softened the statement.

"*Both samples are nighttime samples, so differences* should mostly *be attributed to different meteorological conditions during sampling* as opposed to different diurnal effects."

Line 264: The statement that "similar clustering results" were obtained compared to Thoma et al. (2025) is unclear, as it is not specified whether similarity refers to cluster structure, compound class distributions, or source attribution patterns. A more precise description of the basis for this comparison would be helpful, and, where possible, should be supported with quantitative or clearly defined criteria. In addition, a comparison with Pereira et al. (2025), which

reports LC-HRMS measurements from the same ACROSS campaign and sampling periods, would be highly relevant.

We agree with reviewer 2. The implemented changes now explain that the similarity of the two clustering results is solely based on a visual comparison of the two Van Krevelen plots. Despite being subjective to the viewer, we still think that the accumulations of compounds is visible in both compared works.

"This observed similarity is solely on the basis of a visual comparison of the Van Krevelen plots of the two studies, where in both cases, the distinct accumulations described in chapter 3.3 can be seen."

We have also added a chapter in the SI where this similarity is shown in more detail.

"Comparison of the clustering results with Thoma et al. 2025

Comparison of the clustering results presented in Thoma et al. 2025 with the results presented in this work reveals that a significant portion of the identified compounds within the "biogenic" (see Thoma et al. 2025; Fig. 3, top, green bordered) or regional background (our work; Fig. 5, top part) fall within $1.2 \leq H/C \leq 1.7$ and $0.15 \leq O/C \leq 0.5$. Most of the compounds there are in the CHO subset, As explained by Thoma et al. 2025 and in our work, this area corresponds to oxidation products of terpenes/isoprenoids (H/C = 1.6), further verifying the biogenic origin of the compounds.

In the anthropogenic cluster on the other hand, two accumulations of compounds within other subsets can be found. The first, most prominent accumulation of compounds within the anthropogenic subset is located at $1.5 \leq H/C \leq 2.2$; $0.2 \leq O/C \leq 1.0$ and mostly consists of compounds within the CHOS subset. This accumulation can again be seen in both works (see Thoma et al. 2025; Fig. 3, bottom, magenta bordered; and our work; Fig. 5, bottom part). Due to the high H/C ratio, those compounds are likely aliphatic alkanes with a sulfur containing functional group, as might originate from fossil fuel combustion. The second accumulation of compounds within the anthropogenic cluster is located at $0.4 \leq H/C \leq 0.8$; $0.2 \leq O/C \leq 0.8$. This accumulation is not as pronounced as the CHOS accumulation, however it is still present in both works. Compounds within this accumulation are likely to be nitrogenated aromatic compounds, explaining the low H/C ratio."

Line 270: The rationale for selecting samples S6, S21, S22, and S38 for exclusion in the validation clustering experiment is not explained, and it is unclear whether this selection was based on predefined criteria, randomness or on inspection of the clustering results.

We agree with reviewer 2. We have added a part that explains the rationale for choosing the 4 samples. We have added this description in chapter 3.3.4. The samples were not chosen randomly. In the "testing" dataset, we wanted to have samples that passed over the greater Paris area (N-NE), as well as samples with air coming from SW and SE to cover a range of

different wind directions and therefore also a range of different anthropogenic contributions to the sample.

We have added this in the manuscript, along with a short description.

"The four samples were chosen primarily based on wind direction of the samples. Sample S21 and S22 showed a clear trajectory passing over the greater Paris area, whereas sample S6 approached the sampling site from the SW direction and sample S38 approached from SE. Furthermore, the sample S21 and S22 are consecutive day / night samples, choosing those consecutive samples therefore ensured, that there was no diurnal effect was accidentally contained in the clustering results."

Line 325: Please correct the typo from "and likely related" to "and are likely related." Additionally, consider replacing the vague term "related" with a more precise descriptor, such as "chemically related" or "members of a homologous series." As written, "related" is ambiguous and could refer to source origins, chemical structures, or transformation pathways.

Thank you. We have corrected the sentence.

Line 327: The term "concentration" should be re-assessed, as the reported values represent signal intensity rather than calibrated concentrations, and this should be corrected throughout the manuscript. The claim that these signals are "strongly increased" is not quantitatively supported, and the statement that this "validates the automatic clustering results" is too strong given the lack of statistical evaluation. The wording should be revised to more appropriately reflect a qualitative or semi-quantitative comparison.

We agree. Peak areas are interpreted only as relative within-compound signal intensities; no calibrated concentrations are reported. We replaced "concentration" with "signal intensity" where appropriate.

Line 328: The interpretation of the fragmentation data for $C10H18NO8S^-$ would benefit from clarification. While the observed neutral loss of HNO3 is consistent with the presence of an organonitrate functionality, the terminology "nitro group" is chemically imprecise in this context (i.e., nitro vs. nitrate ester/organonitrate). In addition, the extrapolation that the "whole compound class" carries this functionality is not fully supported by a single MS/MS observation and should be stated more cautiously or supported with additional evidence.

We have changed the wording from "nitro group" by "nitrate ester group". We agree that the neutral loss observed for a single compound does not establish the functionality of the full

series. However, due to the similar behavior of the three individual compounds, we assumed chemical relationship between the compounds and therefore also likely a nitrate ester group for the two other compounds of the compound class.

Line353: The interpretation of the C10 fragment evidence requires clarification. The parent ion is identified as $C11H17O6^-$, from which C10 fragments can arise through common fragmentation pathways, such as loss of small neutral or alkyl groups. Therefore, the presence of C10 fragments is not diagnostic of a monoterpene (C10) precursor, but rather a general and expected outcome of fragmentation of a C11-containing molecule. Although these compounds are assigned to the biogenic cluster, the conclusion that the observed fragments directly support a biogenic source attribution conflates structural fragmentation behavior with source identification and should be stated more cautiously or supported by additional diagnostic evidence.

We agree that the C10 fragments are not diagnostic on their own. The text now states that the biogenic interpretation is supported jointly by the regional-background intensity profile and published formation pathways, while the fragmentation pattern is only consistent with, rather than proof of, a monoterpene-related precursor.

Line 374: The proposed calculation of "anthropogenic influence" as the relative abundance of cluster-assigned compounds requires clearer methodological definition, as cluster membership in a non-target feature space does not directly translate into source apportionment. The reduced-sample clustering exercise provides a useful internal consistency check but does not represent an independent validation, as it is performed within the same dataset and analytical framework. Therefore, the interpretation that it validates "truly unknown organic aerosol samples" is overstated. In addition, it is unclear how the clustering step improves source attribution compared to the wind-based classification, which already appears to capture the primary source information.

We acknowledge that the term "calculation" may imply a level of quantitative rigor that is not intended here. We have therefore replaced "calculation" with "evaluation." This revised wording clarifies that the proposed approach is not an exact quantitative analytical method, but rather a means of identifying changes within a set of related samples. One potential application would be the identification of periods associated with elevated production activity in industrial areas located upwind of the sampling site. Another example would be to use this metric in a continuous monitoring environment where new pollution sources should be quickly identified. However, we deliberately refrained from including a specific example in the manuscript, as this could unnecessarily restrict the perceived scope and applicability of the metric.

Consequently, the reduced clustering approach should not represent an independent validation, but rather show that the clustering works also on previously unseen samples (coming from the same dataset though). We understand that the formulation "truly unknown organic aerosol samples" is somewhat overstated and have softened it in the text.

"*...showing the applicability* of this method *to determine cluster contributions of* previously unseen *organic aerosol samples.*"

The clustering based approach provides an easy way of rapidly clustering thousands of compounds in a non-target dataset. The method that we presented here, is indeed a wind-based classification approach. Wind data is paired with the individual samples, after which the clustering is carried out, which then clusters the compounds into one of the three subsets. Due to the location of the ACROSS sampling site with influences from anthropogenic (greater Paris area) and biogenic / regional background regions, the dataset allowed to pair wind directions with the time series of the individual compounds and derive information about the likely source of the compound. This yielded the two compound classes with compounds in the anthropogenic cluster having a significantly higher concentration during NE wind conditions – when the air mass passed over the greater Paris area.