

Final author comments

We would like to thank the reviewer for taking the time to carefully read and evaluate our manuscript, and for the positive assessment of the importance and usefulness of this intercomparison study. We believe that the comments and suggestions have helped us further clarify and improve the manuscript.

Below, we provide **point-by-point responses to each comment** (in red) and indicate the **corresponding changes made to the manuscript** (in blue).

In addition to the changes made in direct response to the reviewers' comments, we have also revised the "Supplement" section as requested following the initial submission and corrected a few typographical errors identified during the revision process. We have also made a few additional changes to the manuscript that were not specifically requested by the reviewers. The schematic of the Eisele-type inlet in Figure 1 has been partially modified. Figure 3 has been revised by adding labels to the peaks included in the nitrate-mode peak list and moving the legend to the top of the figure. Figure 8 has also been modified by excluding the data collected by EL on August 2 after 05:40, consistent with the approach used in the previous figures and analysis, for the reasons explained in the text. In addition, we have added two funding sources and one co-author that were unfortunately omitted from the first submission. Finally, some table, figure, and reference numbers in the text have changed as a result of the new additions made following the reviewers' suggestions.

Cecilia Righi, on behalf of all co-authors

Reviewer #1 comment

This paper presents an intercomparison of CIMS instruments at SMEAR II. The results show variable intercomparability between CIMS instruments, where the same sulfuric acid calibration is carried over. This is unsurprising. More surprisingly, the situation for bromide is substantially better, adding further evidence that it is a good reagent ion to use. This is an extremely useful paper that addresses a real gap in knowledge for CIMS users, contains a wealth of knowledge and useful data, and I recommend it for publication. A few minor comments are provided below.

Specific comments

Line 40: "studies indicate that neutral new particle formation (NPF) dominates over ion-induced pathways in the continental boundary layer (Kerminen et al., 2010)". I believe this is true, but it may be worth referencing more recent results (e.g. Zhang et al., <https://doi.org/10.1038/s41586-024-07547-1>), which would support the statement across a wider range of mechanisms.

We have included the reference suggested – Zhao et al. (2024), based on the DOI – and revised the passage as follows:

Although the species and mechanisms leading to the nucleation of aerosol particles vary with geographical location and elevation, studies indicate that neutral new particle formation (NPF) involving multiple neutral compounds generally dominates over ion-induced pathways in the continental boundary layer (Kerminen et al., 2010; Zhao et al., 2024).

Line 53: I'd suggest a different word to "numerous", which could be read as meaning high in concentration and so contradict point (1), where the species are described as being at low concentration. Also "low in concentration" is a bit nebulous.

We have revised the passage accordingly:

These neutral compounds are challenging to measure because (1) they are present at extremely low concentrations (often at parts-per-trillion to parts-per-quadrillion volume mixing ratios), (2) they readily condense onto surfaces due to their low volatility, and (3) they require charging before detection, which introduces additional complications, as different compounds may be charged and detected with different efficiencies. Adding to these challenges, (4) they comprise a large number of chemically diverse species: [...]

Line 53: The species span not only a large m/z range, functional groups and degrees of oxidation, but also a large range of carbon number, which is worth adding.

We have further revised the passage to include this point:

[...] even when considering only the monoterpene-derived HOMs commonly targeted within this class, they cover a broad m/z range and include molecules with varying numbers of carbon atoms and different functional groups and degrees of oxidation.

Line 73: "the operator can alter the fragmentation...". As far as I'm aware this applies mostly to the fragmentation of clusters rather than of molecules, so it might be worth specifying this.

We agree and have clarified the sentence by replacing "molecules" with "clusters".

Table 1: CEAM is not defined.

We have now defined the acronym CEAM in Table 1.

Line 122: The SI reference should point to the specific SI figure (here & throughout)

We have revised the references to the Supporting Information throughout the manuscript to point to a specific section, table, or figure.

Line 145: "CI-Fi1" should read "CI-FI1" (capital I).

We have corrected "CI-Fi1" to "CI-FI1".

Line 185: It would be nice to quote the mass resolution of each instrument, perhaps the resolution at 201 m/Q (for nitrate) for nitrophenol.

We have added Table 3, which reports the full width at half maximum (FWHM) and the corresponding mass resolution calculated for each instrument at m/z = 201 (nitrate mode).

We have also added the following sentence to introduce the table:

For the NO₃⁻ mode measurements presented in this study, the mass resolution at m/z = 201, calculated from the full width at half maximum (FWHM) of the corresponding peak, is reported for each CI-API-ToF in Table 3.

Line 200: By "performance" do you mean sensitivity? And is this necessarily true, given the results of Hyttinen? the binding energies do scale linearly, but that doesn't necessarily translate to performance or sensitivity once a threshold is passed where one collision = one cluster

We thank both reviewers for pointing this out. We agree that the term “performance” was not sufficiently specific here and could imply a direct relationship with sensitivity. We have therefore revised the sentence as follows:

These reagent ions are highly selective toward H_2SO_4 and other gas-phase acids stronger than HNO_3 , as well as toward oxidation products of VOCs containing more than five oxygen atoms, with the binding strength of the reagent ion to the target molecule increasing almost linearly with the number of oxygen atoms (Hytinen et al., 2018).

Line 225: It would be nice to have a brief justification for why the less-oxygenated species (e.g. $\text{C}_{10}\text{H}_{16}\text{O}_{4,5,6}$) weren't intercompared for nitrate, since these measurements ARE sometimes used for those of us who can't run multiple chemistries simultaneously, and intercomparability would likely be even worse for these, which would be bad news.

We thank the reviewer for the comment. As mentioned in Section 2.3.1, the peak lists were designed based on the selectivity of the reagent ion employed. Although nitrate ion chemistry can sometimes be used to detect less-oxygenated species when other ionization schemes are not available, nitrate is generally not considered a suitable reagent ion for the efficient detection of these compounds. Therefore, these species were not included in the nitrate-mode intercomparison. Nevertheless, in this study we also considered some less-oxygenated species detected using nitrate when relevant for comparisons between different reagent-ion modes. Some of these comparisons are also shown in Section S15 of the SI.

Line 237: CI-API-ToF is used without introduction. A sentence above might help, e.g. "API-ToF refers to the atmospheric pressure instruments without ionisation; CI-API-ToF to those with ionisation; CIMS is the more common parlance, as it generally also includes instruments that do not operate at atmospheric pressure."

We agree that this terminology should be clarified before the first use of CI-API-ToF. We have therefore added a sentence in Section 2.1, when presenting the mass spectrometers involved in the campaign:

Hereafter, the API-ToF mass spectrometers operated with chemical ionization are referred to as CI-API-ToFs, while CIMS is used as a general term for all instruments included in the intercomparison, since this terminology also includes mass spectrometers that do not operate at atmospheric pressure, such as B-ToF.

Lines 265–268: the assumption that all the small acid peaks sit in sulfuric acid's mass range and therefore get $C_{\text{trans}} = 1$ I think might need some consideration; iodic acid, nitrophenol, MSA and malonic acid go a fair bit higher in m/z than 97, especially $\text{HIO}_3\text{NO}_3^-$, and given how steeply the transmission curves move, blanket-assigning unity to all of them could introduce a systematic bias. This might even contribute to the poor-to-moderate HA agreement in Table 6, so it's worth at least acknowledging (but probably not worth re-doing any analyses)

We thank the reviewer for pointing this out. We would first like to clarify that the statement that the HA peaks fall within the sulfuric acid mass range refers to all the peaks considered to calculate the sulfuric acid concentration. As shown in Eqs. (1) and (3), this includes not only the bisulfate peak at $m/z = 97$, but also the peaks at $m/z = 160$ and $m/z = 195$, for which the transmission is also constrained to 1. Thus, only the nitrate clusters of nitrophenol and iodic acid ($m/z = 201$ and $m/z = 238$, respectively), together with deprotonated methanesulfonic acid ($m/z = 95$), are excluded from the range of sulfuric acid peaks.

That said, we acknowledge that this approximation might affect the intercomparison outcomes for such species, and therefore we performed a correlation analysis for the species measured in nitrate mode, based on concentrations calculated by applying only C_{SA} and by applying all corrections. The results, reported in Tables S7 and S8 of the SI, show that applying C_{trans} does not improve the agreement between instruments for such species, except for nitrophenol. Based on these results, we did not find evidence that the non-application of the C_{trans} correction is responsible for the poor agreement. Nevertheless, we agree that this possibility is worth acknowledging, and we have therefore added a paragraph on HAs to the Conclusions:

(3) For organic and inorganic acids of relatively low molecular mass (below m/z 250) considered in this study (HAs), the agreement among instruments was generally poorer than for sulfuric acid and HOMs. The reason for these larger differences could not be clearly identified. One possible explanation considered was the lack of correction for relative transmission differences for these species; however, additional analysis showed that applying this correction did not improve the agreement among instruments. The results therefore suggest that the measurement of these compounds remains challenging, and that further work is needed to better understand and improve their quantitative determination by CIMS.

Line 310: A HEPA filter was used by Häkkinen, but that was in front of a VIA. It would be good to explain why it works here. I presume it's the huge surface area essentially scavenging all of the H_2SO_4 and HOM? A quick SI figure showing the signals going to zero with the filter in place would be helpful. Does it also work for malonic acid and other compounds? I remember seeing this done before, but not with a HEPA filter, and the signals nosedived, but didn't go to total zero.

We thank both reviewers for their suggestions regarding the use of the HEPA filter for background measurements. In response, we have included Figure S3 in the SI, as requested, and modified Section 2.3.5 to clarify the rationale and limitations of applying this approach to gas-phase measurements:

The use of HEPA filter for this type of measurements has been reported in previous studies (Boyer et al., 2024; Häkkinen et al., 2023) and is straightforward to apply during field measurements. Although HEPA filters are conventionally used to remove particles in particle-phase measurements, they can also be used for background measurements of gas-phase species when the compounds of interest are sufficiently low volatile and are efficiently removed by interaction with the filter material. The efficiency of this removal is, however, compound-dependent and the filter should therefore not be considered a universal gas-phase zeroing method. To demonstrate the effectiveness of the approach for the compounds targeted in this study, measurements with and without the HEPA filter are shown in Figure S3 in the SI. The signals of the investigated compounds were strongly suppressed when the filter was installed, although the signals did not decrease completely to zero and the extent of signal decrease differed among sulfuric acid, HAs, and HOMs.

Line 320: The CIMS LOD does indeed depend on electronic noise, but for many species it also depends on the height of an adjacent peak. For example, if quantifying iodous acid with a large HOM sitting right next to it, the tail of the HOM can swamp the iodous acid, so the LOD is a little more complex than described. I don't think this needs explicit treatment here, but it might be worth raising as a caveat of atmospheric mass spectrometry. (Optional.)

To address the comments from both reviewers on this point, we have modified the paragraph to more clearly describe the method used to characterize the general background signal and to acknowledge the potential influence of nearby peaks, as follows:

To further characterize differences in noise levels between instruments, we calculated a more general LOD, referred to as LOD_{CIMS} , using the same formula but a different definition of μ : for each HOM peak included in the NO_3^- mode peak list, a nearby region where no peak was expected was selected, and the signal in this region was fitted and taken as μ . For example, for a peak at nominal masses 308 and 494 Th, the signal at 308.5 and 494.5 Th were used, and similarly for other peaks up to a nominal mass of 620 Th. Since the HOM peaks span a broad m/z range, LOD_{CIMS} arguably provides a more general characterization of the background signal and noise level, rather than a compound-specific detection limit, such as LOD_{SA} , and can thus help in interpreting differences in measurement outcomes between instruments. It should be noted that the detection limit of an individual compound can also be affected by the presence and intensity of nearby peaks, an effect that is not accounted for in this approach and may become increasingly relevant at higher m/z .

Line 327: "AIM" should be capitalised.

We have capitalized "AIM" here and throughout the manuscript.

Line 345: Building on the earlier comment, these LODs seem very high, so it would help to back this up with the SI figure showing signals going to zero. Since the CIMS essentially operates as an ion counter, the LOD is also very sensitive to the averaging time, so this discussion could use more nuance. We also tend to be able to get sensible-looking time series below $1e5$ (for example, h_2so_4 signals decreasing towards $1e4$ as the sun goes down and rising back above $1e5$ as it comes up again, which gives credence to these LODs being high).

We thank the reviewer for this additional comment. As it was requested in a previous comment, measurements with and without the HEPA filter are shown in Figure S3 of the SI. The dependence of the LOD on the averaging time is already discussed in this paragraph, where we highlight the differences in averaging periods between the studies used for comparison. We have now included in the discussion one more sentence to clarify that the reported LODs should not be considered strict limits below which the measurements are necessarily unreliable:

However, differences in the background measurement procedure and averaging time need to be considered when comparing these values. Boyer et al. (2024) used HEPA-filter background measurements, as in the present study, and 5-minute averaged data, whereas He et al. (2023) introduced pure nitrogen or synthetic air into the CI inlet and calculated the LOD from 1 h of data averaged to 1 min intervals. In addition, possible incomplete removal of sulfuric acid or sulfuric-acid-containing clusters during HEPA-filter background measurements may have contributed to an overestimation of the derived LOD_{SA} values. For these reasons, the reported LOD values should not be considered strict limits below which the measurements are necessarily unreliable.

Line 355: Very useful point. We tend to set and forget our calibrations, and this is good reason not to.

We thank the reviewer for highlighting this point.

Figure 2: The spread of the data is very wide, and some points sit very far from the curve. I understand this is normal, but would it be possible to offer some practical recommendations for obtaining these curves as cleanly as possible?

We thank the reviewer for the suggestion. Having a larger number of data points and repetitions would help to obtain cleaner relative transmission curves. As mentioned in Section 2.3.3, time limitations during the campaign, together with the constraints of operating in field rather than laboratory environment, were the main reasons for using only two compounds (C5 and C7) for this calibration, out of the six suggested by Heinritzi et al. (2016). For the same reasons, it was not possible to obtain more than two repetitions for each compound, which would have provided more data points for fitting the curves.

In response to the related request from RC2, we have now included more details on the calibration experiment and some practical recommendations in Section 2.3.3:

Transmission calibrations were performed on all CI-API-ToFs in NO_3^- mode, following the depletion method described by Heinritzi et al. (2016), and briefly here. A Union Tee piece of 3/4" was attached to the inlet tubes of the CI-API-ToFs, using adaptors to 1" for instruments with larger inlet tube diameter. The third outlet of the Tee was closed with a cap, in which a small quantity of the perfluorinated acids used for the calibration was placed, either in liquid or solid form, one compound at a time. The calibration experiment consisted of heating the cap with a heat gun to induce evaporation or sublimation of the calibrant, resulting in depletion of the reagent ions. After the heating cycle, the signal from the perfluorinated acid decreased as the system cooled and the primary reagent ions recovered. The perfluorinated acids used in the experiments are listed in Table 4. During the ACTRIS CI-FI1 campaign, C5 and C7 were employed, with C7 excluded for EH2, whereas C7, C8, and C9 were used for ML2, which was calibrated post-campaign in the ACTRIS Laboratory of the University of Helsinki. The use of different calibrant sets was mainly related to time limitations during the campaign, as well as to ongoing testing and optimization of the method. Based on the experience gained during the campaign, we recommend performing at least three heating cycles for each calibrant and aiming for different degrees of reagent ion depletion, as also suggested by Heinritzi et al. (2016).

Section 3.3: There is a reference error that should point to Fig. 4 (here and throughout).

We have corrected the reference error throughout the manuscript.

Line 439: I agree it's wise to exclude the AB instrument from this figure. Would it be possible to show the H_2SO_4 signal time series from AB in the supplement, though? It might be useful for those of us interested in that instrument. (also totally optional)

We thank the reviewer for the suggestion. Unfortunately, no clear sulfuric acid signal was observed with the AB instrument, and therefore we are not able to provide the requested time series in the SI.

Figure 4: A great figure. A few comments. On my laptop screen the horizontal lines are hard to distinguish, so I'd recommend removing the major y-axis gridlines, which would remove two lines per panel. The caption is clear, but the figure itself could be more intuitive, especially the middle panels; I'd either label all panels (a)–(f) and clarify in the caption that the middle ones are medians calculated with different LODs, or add y-axis labels to the middle-top and middle-bottom reading "Concentration (LOD_SA)" and "Concentration (LOD_CIMS)". I'd also set the y-axis limits of the right-hand panels equal to those of the left-hand panels, so the reader can move their eye horizontally across the data. Also, the fact that the choice of LOD affects the median concentrations and the shape of the distributions

so strongly is very important. Note too that the sulfuric acid concentrations rise at 08/02 00:00; this appears consistently across all instruments, and so is likely a real atmospheric phenomenon, yet occurs at concentrations well below LOD_SA, which lends further credence to LOD_SA not being the true limit of detection. Finally, I presume that in the right-hand panels the points below LOD are not included in the fits?

Figure 4 has been modified in response to the comments from both reviewers. The following changes were made:

- Major y-axis gridlines were removed, as suggested by RC1.
- Each panel was labelled individually, and the caption was modified accordingly, as suggested by RC1.
- The caption now specifies that points below the corresponding LOD are excluded from the fits shown in the scatter plots, as suggested by RC1.
- Data from instrument EL that were not considered usable were removed from the graphs, as suggested by RC2.
- The median information in the middle panels was shown in bold to improve readability, as suggested by RC2.

We decided not to adopt the suggestion from RC1 to set the y-axis limits of the right-hand figures equal to those of the corresponding left-hand figures, since we believe that keeping the current y-axis limits allows better readability of the point distributions and the associated R^2 values.

Figures 5 and 6, as well as Figures S5 and S6 in the SI, were modified similarly.

Figure 4: I remember an intercomparison done at the TROPOS chamber led by Peter Mettke showing that a lot of the instrument-to-instrument variability was, in fact, due to dodgy calibrations, and when everyone adopted some flat calibration number, the instruments agreed better. Not sure whether that's just papering over the cracks (so to speak) but it might be worth quickly checking for the H₂SO₄ time series here.

We thank the reviewer for this interesting comment.

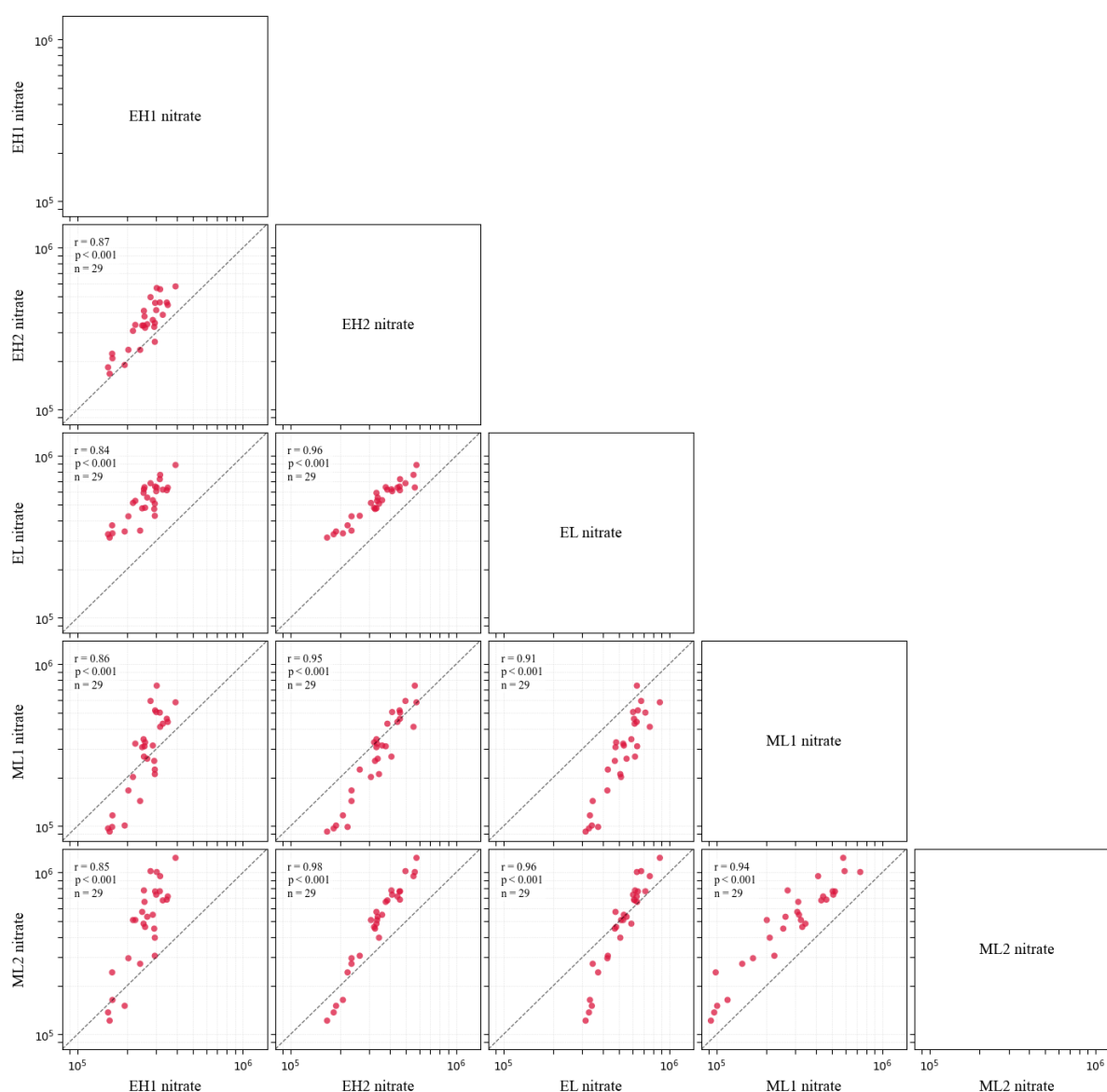
Line 455: "Nice agreement" reads as slightly too casual.

We have moved the definition of good, moderate, and poor agreement before this sentence, as suggested in the comment of Line 504. We have also changed "nice agreement" to "good agreement" in this section.

Optional: it might be nice to include a correlation matrix showing the scatter of every instrument against every other, e.g. for readers interested in how the EH instruments compare specifically against the ML instruments in the SI

We thank the reviewer for the suggestion. We have decided to include here, for the reviewer to see, the requested figure for the case of dimers detected in nitrate mode. As discussed in the paper, EH1 and EH2 show good agreement, and EL also behaves similarly, although with a small offset. For ML1 and ML2, instead, the points lie on the diagonal when looking at their correlation with each other, but not when compared to the CI-API-ToFs with Eisele-type inlets.

dimers



Line 495: The monomer signals behave super well, but this is the sum of many peaks. What is the agreement like peak by peak? It would be interesting if the individual signals disagreed while the sum did not.

We show here the comparison for some individual closed-shell organonitrate monomers. The good agreement observed for the summed signal is also generally maintained for the individual species signals.

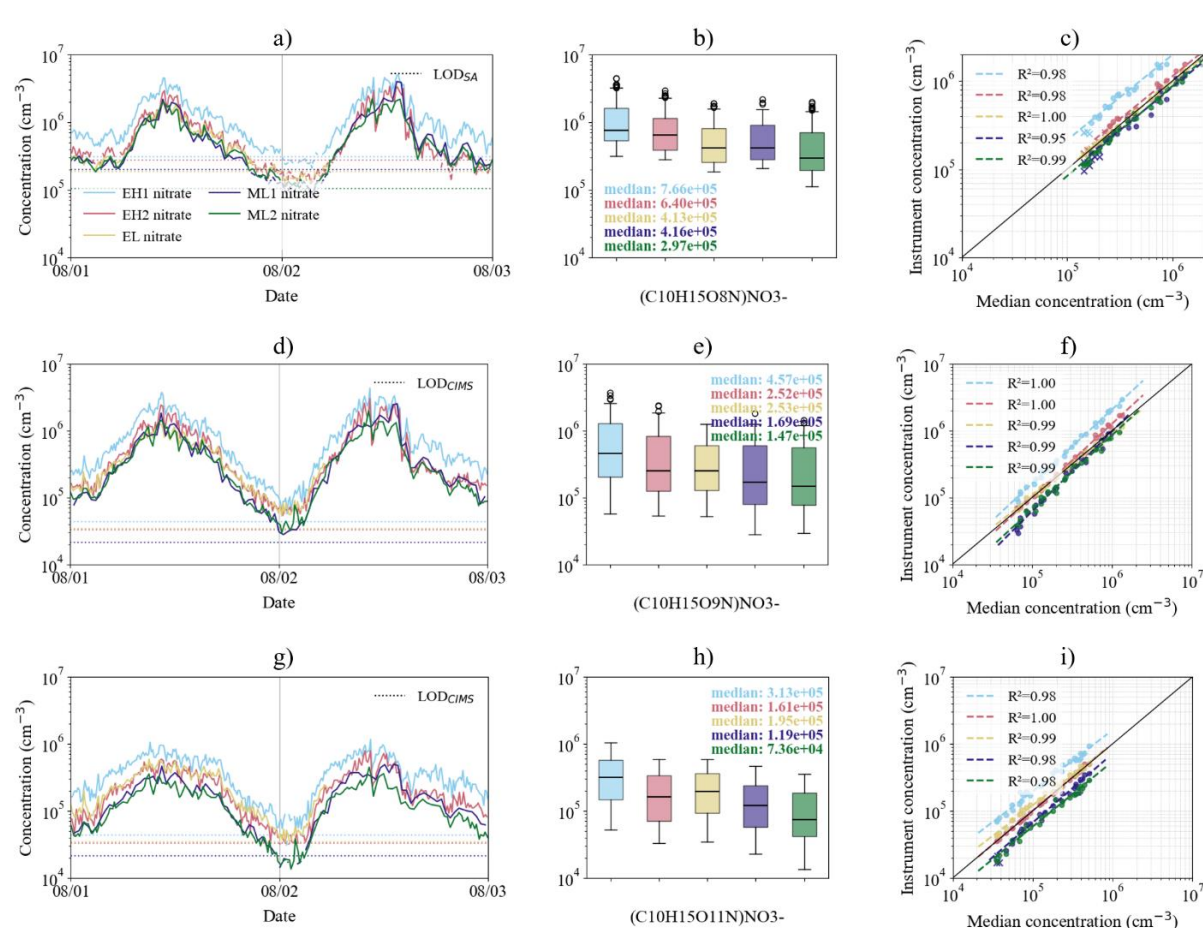


Table 6: Is this the correlation of each instrument individually against the mean of all instruments? And which LOD is used for them? It's worth clarifying this in the table caption. Also, very interesting that the agreement is so poor for organic peroxy radicals and for nitrophenol. Is this due to low signals?

The results reported in the table are obtained from the correlation analysis, where the hourly median concentrations measured by the five instruments are directly compared with the corresponding measurements from the other instruments, rather than with the mean or median value of such hourly median concentrations. Values below LOD_{CIMS} were excluded from the correlation analysis, and this information has now been added to the table caption. The captions of Tables S7-S9 in the SI were modified accordingly.

We do not have a clear explanation for the poorer agreement observed for nitrophenol and organic peroxy radicals compared with the other compound groups. However, a comparison with Table S9 in the SI, where EH1 is completely excluded from the analysis, shows higher R^2 values for these two groups (0.681 and 0.320, respectively, compared with 0.314 and 0.089 in Table 6). This indicates that the different behavior of EH1 is an important contributor to the poorer agreement observed for these species.

Line 504: It would be useful to set up what is meant by "good agreement" and so on before line 455 so you can say good agreement as a defined term.

Addressed as described in our response to the comment on Line 455.

Line 529: Text says $R^2=0.853$, figure shows a higher number. Also, the caption of figure 6 says it uses means while others use medians. is this right, and if so, why?

The R^2 value given in the text is obtained from the correlation analysis, the results of which are reported in Table S11 of the SI, where the hourly median concentrations measured by ML1 and ML2 are directly compared. In Figure 6, instead, the R^2 values are calculated separately for each instrument by comparing its hourly median concentration with the mean of the hourly median concentrations from the two instruments. This difference in the calculation explains the different R^2 values.

Concerning the caption of Figure 6, we acknowledge that the wording was not fully clear. On the x-axis, we show the mean of the hourly median concentrations from the two instruments, which is equivalent to the median when only two values are considered. We have therefore rephrased this part of the caption of Figure 6 (and similarly that of Figures 4 and 5) as follows:

The right figure ((c) and (f)) shows the hourly median concentration measured by each instrument against the mean of the corresponding hourly median concentrations from the two instruments.

Line 547: Two compounds are named (sulfuric acid and $C_{10}H_{16}O_6$), not three.

We have corrected “three” to “two”.

Line 595: I would be a little harsher about the I⁻ results, which are rather poor. (This is surprising!)

We have modified the section according to the reviewer's suggestion:

By comparison, under the present sampling conditions and instrument configuration, I⁻ mode showed substantially poorer agreement with the other instruments and a limited ability to resolve variations in $C_{10}H_{16}O_5$ concentrations.

Conclusions: Great, but, this is somewhat of a repetition of the abstract. It may lie outside the scope of an intercomparison paper, but if you deem it appropriate, it could be worth reframing this section as a set of practical recommendations for instrument operation and intercomparability that go beyond the usual good-practice steps we already follow.

We thank the reviewer for this positive comment and suggestion. We also considered reframing this section as a set of practical recommendations, but decided to retain the current structure, as we believe that the practical recommendations that can be directly drawn from this campaign are already discussed throughout the manuscript and summarized in the Conclusions. These mainly concern the importance of performing regular calibrations and carefully applying the appropriate correction factors, as discussed in the relevant sections of the manuscript.

Final author comments

We would like to thank the reviewer for taking the time to carefully read and evaluate our manuscript, and for highlighting the usefulness of the presented guidelines for the CIMS community and their potential contribution to harmonizing condensable vapor measurements. We believe that the comments and suggestions have helped us further clarify and improve the manuscript.

Below, we provide **point-by-point responses to each comment** (in red) and indicate the **corresponding changes made to the manuscript** (in blue).

In addition to the changes made in direct response to the reviewers' comments, we have also revised the "Supplement" section as requested following the initial submission and corrected a few typographical errors identified during the revision process. We have also made a few additional changes to the manuscript that were not specifically requested by the reviewers. The schematic of the Eisele-type inlet in Figure 1 has been partially modified. Figure 3 has been revised by adding labels to the peaks included in the nitrate-mode peak list and moving the legend to the top of the figure. Figure 8 has also been modified by excluding the data collected by EL on August 2 after 05:40, consistent with the approach used in the previous figures and analysis, for the reasons explained in the text. In addition, we have added two funding sources and one co-author that were unfortunately omitted from the first submission. Finally, some table, figure, and reference numbers in the text have changed as a result of the new additions made following the reviewers' suggestions.

Cecilia Righi, on behalf of all co-authors

Reviewer #2 comment

The paper "*Towards harmonized measurements of condensable vapors: insights from the intercomparison of six chemical ionization mass spectrometers at a boreal forest site*" by Righi, et al. describes not only the intercomparison between six different chemical ionization mass spectrometers (CIMS) but also details the steps taken to characterize the instrument performance and enable a comparison. I believe the guidelines shown will be very helpful for the CIMS community and indeed provide an important stepping stone to harmonize condensable vapor measurement. Thus, I recommend this paper for publication after some issues are addressed in a revision.

- 1) In the introduction in L.77 the issues in comparing "quantitative measurement outcomes" are introduced. I think here it needs to be introduced first that for CIMS the standard procedure is not to calibrate each analyte for each instrument individually, as only with that in mind the differences in set-up start impacting quantitative measurement outcomes.

We thank the reviewer for pointing this out. We have modified the section accordingly, as follows:

The flexibility of CIMS, which makes it a powerful technique in atmospheric studies, also represents a challenge for data interpretation. Since these instruments are generally not calibrated for every targeted analyte, different instrument setups and operating configurations, even when the same reagent ion is employed, can

introduce substantial variability in quantitative measurement outcomes and complicate direct comparisons across studies

- 2) In Section 2.3.4 “Data processing” the calculation of the C_{inlet} factor is introduced. I would recommend to add a table with all calculated C_{inlet} factors to the supplement as this would be helpful to gauge the impact of this factor. For the information given in the supplement, I also recommend to add a bit more explanation to the equations to calculate C_{inlet} : How were the equations derived? For example, supplement L.80: Where does the factor 3.6 come from? And is there literature that could be cited in this context?

We agree that having the calculated C_{inlet} factors included in the Supplementary Information can be useful for the reader, and we have added Table S5 in Section S6 for this purpose.

As for the equations in Section S6, Eq. S1 was taken from Fuller et al. (1969), as indicated in the text, while Eqs. S2–S5 are based on the calculation of first-order diffusion losses in a cylindrical inlet tube. We have now added further explanation of the origin of factor 3.6 in Eq. S2.

The factor 3.6 is the geometric factor accounting for diffusion-limited transport to the wall of a cylindrical tube under laminar-flow conditions (Tang et al., 2014).

- 3) It would be good to provide a bit more additional background/information on the “Transmission calibration” Section 2.3.3:
The reader would profit from a short explanation of the depletion method -why does this work to determine transmission? What is the actual measurement set-up? E.g. how were stable concentrations of the perfluorinated acids generated and supplied to the instruments, how was their concentrations varied? How was the data analyzed? The text only mentions the fit applied to the “depletion data” but what was this depletion data practically? Some details might be better suited for the supplement, but as a reader it would be great to have this information. Additionally, I think some restructuring of the text would be useful: Section 2 describes the data processing, but in case of the transmission calibration the actual fit function and how it relates to the C_{trans} factor is only introduced in section 3.2. Therefore, I would include this information in section 2.3.3.

We thank the reviewer for these useful suggestions. While the sulfuric acid calibration setup used during the ACTRIS CI-FI1 campaign differs in some aspects from the setup described by Kürten et al. (2012) and has therefore been described in more detail in the SI, the setup used for the transmission calibration using the depletion method is the same as that described in Heinritzi et al. (2016, Figure 1c), which is already referenced in the text. We therefore did not include an additional schematic of the setup in the manuscript or the SI. We have now added some explanation of the method, together with the fit function, to Section 2.3.3:

Transmission calibrations were performed on all CI-API-ToFs in NO_3^- mode, following the depletion method described by Heinritzi et al. (2016), and briefly here. A Union Tee piece of 3/4" was attached to the inlet tubes of the CI-API-ToFs, using adaptors to 1" for instruments with larger inlet tube diameter. The third outlet of the Tee was closed with a cap, in which a small quantity of the perfluorinated acids used for the calibration was placed, either in liquid or solid form, one compound at a time. The

calibration experiment consisted of heating the cap with a heat gun to induce evaporation or sublimation of the calibrant, resulting in depletion of the reagent ions. After the heating cycle, the signal from the perfluorinated acid decreased as the system cooled and the primary reagent ions recovered. The perfluorinated acids used in the experiments are listed in Table 3. During the ACTRIS CI-FI1 campaign, C5 and C7 were employed, with C7 excluded for EH2, whereas C7, C8, and C9 were used for ML2, which was calibrated post-campaign in the ACTRIS Laboratory of the University of Helsinki. The use of different calibrant sets was mainly related to time limitations during the campaign, as well as to ongoing testing and optimization of the method. Based on the experience gained during the campaign, we recommend performing at least three heating cycles for each calibrant and aiming for different degrees of reagent ion depletion, as also suggested by Heinritzi et al. (2016).

[Table 4]

Curves describing the relative transmission as a function of m/z were derived by least-squares fitting of a 2-fold Gaussian distribution (Eq. 2) to the depletion experiment data:

$$f(x) = a_1 \cdot \exp\left(-\left(\frac{x-b_1}{c_1}\right)^2\right) + a_2 \cdot \exp\left(-\left(\frac{x-b_2}{c_2}\right)^2\right), \quad (2)$$

where x corresponds to the m/z of the selected species.

We have not included how the fit function related to C_{trans} in this section, as the correction factor is introduced only later in Section 2.3.4. However, we have now added a brief explanation of how C_{trans} is calculated in Section 2.3.4:

Finally, C_{trans} is the mass-dependent transmission correction factor, obtained as the inverse of the relative transmission function fitted for each instrument and resolved for each species as detailed in Section 2.3.3.

- 4) I also recommend to rework some of the “Background measurements and detection limit” Section 2.3.5:
 - a. For the sulfuric acid background measurement technique an explanation for the application of a HEPA filter is missing for me: The instruments are measuring in the gas phase but a particulate filter is used to determine background concentrations. In the referenced study by Häkkinen et al., this technique has been used for the VIA inlet, so for measurements in the particle phase. Later Boyer et al is also cited with applying this technique for gas-phase measurements. I think a clear separation between gas-phase and particle-phase measurement should be made here and the use of a particulate filter for gas-phase species determination has to be explained. If the argument is that the species of interest are also removed by a HEPA filter due to their low volatility, then I think it needs to be clearly stated. Some of this is touched on by the discussion in line 346-351 but I think the discussion of weak points and different possibilities should be done at the introduction of the method. To evaluate the method, it would be best to show measurements that illustrate that the filter indeed removes the analyte. In any case I recommend to add an explanation why this method is possible for gas-phase analytes and under which prerequisites it is possible and useful.

- b. I would also suggest to more strongly highlight the method to determine the general background by the noise of the instrument.
- c. For the introduction of the limit of detection (LOD) I recommend to mention that other definitions of LOD exist that use both a blank measurement and the slope of a calibration curve and that the method used here and as cited by He, et al is applicable to HOMs and similar analyte groups because it does not require an individual calibration curve for each analyte.
- a. We thank both reviewers for their suggestions regarding the use of the HEPA filter for background measurements. In response, we have included Figure S3 in the SI, as requested, and modified Section 2.3.5 to clarify the rationale and limitations of applying this approach to gas-phase measurements:

The use of HEPA filter for this type of measurements has been reported in previous studies (Boyer et al., 2024; Häkkinen et al., 2023) and is straightforward to apply during field measurements. Although HEPA filters are conventionally used to remove particles in particle-phase measurements, they can also be used for background measurements of gas-phase species when the compounds of interest are sufficiently low volatile and are efficiently removed by interaction with the filter material. The efficiency of this removal is, however, compound-dependent and the filter should therefore not be considered a universal gas-phase zeroing method. To demonstrate the effectiveness of the approach for the compounds targeted in this study, measurements with and without the HEPA filter are shown in Figure S3 in the SI. The signals of the investigated compounds were strongly suppressed when the filter was installed, although the signals did not decrease completely to zero and the extent of signal decrease differed among sulfuric acid, HAs, and HOMs.

- b. To address the comments from both reviewers on this point, we have modified the paragraph to more clearly describe the method used to characterize the general background signal and to acknowledge the potential influence of nearby peaks, as follows:

To further characterize differences in noise levels between instruments, we calculated a more general LOD, referred to as LOD_{CIMS} , using the same formula but a different definition of μ : for each HOM peak included in the NO_3^- mode peak list, a nearby region where no peak was expected was selected, and the signal in this region was fitted and taken as μ . For example, for a peak at nominal masses 308 and 494 Th, the signal at 308.5 and 494.5 Th were used, and similarly for other peaks up to a nominal mass of 620 Th. Since the HOM peaks span a broad m/z range, LOD_{CIMS} arguably provides a more general characterization of the background signal and noise level, rather than a compound-specific detection limit, such as LOD_{SA} , and can thus help in interpreting differences in measurement outcomes between instruments. It should be noted that the detection limit of an individual compound can also be affected by the presence and intensity of nearby peaks, an effect that is not accounted for in this approach and may become increasingly relevant at higher m/z .

- c. Following the suggestion, we have modified the section as follows:

The limit of detection (LOD), defined as the minimum analyte signal that can be distinguished from background noise, can be determined using different approaches. Here, following He et al. (2023), the LOD was calculated for H₂SO₄ was calculated as:

$$LOD_{SA} = (\mu + 3\sigma) \cdot C_{SA}, \quad (6)$$

where μ represents the mean value of the normalized sulfuric acid signal during the background measurement, and σ is the corresponding standard deviation. As described by He et al. (2023), for species without a direct calibration, such as HOMs, the LOD calculated using the sulfuric acid calibration factor can be applied, although this generally represents a lower limit of the actual LOD.

- 5) I would restructure the information flow regarding the participation of the AB instrument and say from the beginning that for data comparison only the I- mode was use.
- a. This could also help explain why no sulfuric acid calibration was performed. In L.243 the reason given for not calibrating the instrument for H₂SO₄ was that the ambient H₂SO₄ concentration would have been too low to measure it with instrument AB. But I was wondering if it would have been possible to calibrate instrument AB with H₂SO₄ in higher than ambient concentration to get a calibration factor to apply besides the later introduced collision limit? I am not so familiar with I-CIMS measurement but is the reason that the expected sensitivities to the different analytes vary too much?

We thank the reviewer for this suggestion. We have clarified in Section 2.2.5 that, although instrument AB operated with three reagent ions, only the iodide-mode data are considered in the present work.

As for the second part of the comment, the sulfuric acid calibration performed for the other instruments actually requires the use of concentrations substantially higher than those typically measured in the ambient (especially at a Boreal forest site), for reasons detailed in Kürten et al. (2012), so this would not have been a limitation in that sense.

In the case of instrument AB, it was decided not to target sulfuric acid in nitrate mode, as the concentrations expected at the measurement site were considered too low for the instrument to detect them, as explained in Section 2.3.2.

- 6) My question in point 5a) touches on a subject that I generally miss in the data processing section: The different sensitivities of one reagent ion to different analytes. This is discussed when introducing the reagent ion chemistry and later on in the results & discussion on specific analyte signals, but I think it should also be clarified when introducing the calculation for analyte quantification that the application of one calibration factor for whole groups of analytes creates significant uncertainty. And that only part of the different sensitivities is removed by applying transmission correction. I know that this is state-of-the-art in CIMS measurements right now, but we should always be aware that this is an approximation, and that it works better in some cases than others.

We thank the reviewer for raising this point, and we have decided to include this consideration in the Conclusions. We have therefore modified the last paragraph as follows:

The ACTRIS CI-FI1 campaign proves that (a) comparable measurements of condensable vapors by different CIMS instruments are achievable when all relevant calibration and correction factors are critically applied, and (b) inter-instrument comparability is strongly influenced not only by reagent ion chemistry, but also by inlet design and instrument operating conditions, including operating pressure. At the same time, it is worth acknowledging that, since these instruments are generally not calibrated for every targeted analyte, the use of a common calibration factor for groups of species remains an approximation because the sensitivity of a given reagent ion can vary between them, and the application of additional correction factors can only partially account for these differences, which may contribute to the less consistent agreement observed for some species.

Hence, detailed guidelines on measurement principles and calibration procedures are urgently needed to ensure good comparability and quantification of condensable vapors across different observations, thereby improving our ability to evaluate and interpret new particle formation processes.

Detailed content-related comments:

- 1) L.72-74: It would be nice to add a reference to the effect of changing voltages in case the reader wants to know more, maybe you could reference a publication on voltage scanning?

We agree that adding some reference on voltage scanning can be of use to the reader and we have cited Lopez-Hilfiker et al. (2016) and Wang et al. (2024).

- 2) L.146-147: How are neutral compounds kept in the O region?

In the case of Eisele-type inlet, neutral reagent gas is not kept in the O region as it is in the case of MION(*) and MION2 inlets, but the neutral gas reaches the I region and travels coaxial to the flow ionized sample gas. The sample gas is then directed to the pinhole (albeit with much lower efficiency than in the case of MION2 inlet for instance, 10-20% vs 90%, see Finkenzeller et al. 2024, Table 1), while the reagent gas not ionized exists together with excess flow (see excess/exhaust flow arrow – purple one in the schematic in Fig. 1).

Here are some schematics that might help you: Fig. 7 of Finkenzeller et al. 2024 (doi:10.5194/amt-17-5989-2024) and Fig. 3 of Sipila et al 2015 (doi:10.5194/amt-8-4001-2015).

- 3) L.135 in description of Fig. 1 and L.162: Is the purge flow different between the MION* and the MION2? In Figure 1 it is marked with an asterisk but only in the MION* panel not in the MION2 panel does this denote a difference between the two purge flows?

We thank you for raising this point. As briefly mentioned in the caption of Fig. 1, the MION* inlet differs from the original MION described by Rissanen et al. 2019

because of the introduction of this purge flow, which is an upgrade to make it more similar to MION2 (so the flows are the same in the two MIONs). As explained in the dedicated paper, MION has in place of this purge flow some counterflows applied through the ion source orifices to prevent the mixing of the electrically neutral reagent precursors with the sample flow. Since this purge flow is essentially what distinguishes MION from MION*, we have decided to mark these flows with the asterisk as well. To make it clearer, we have now modified the caption of the figure as follows:

Schematic representation of the chemical ionization inlets employed during the ACTRIS CI-FI1 campaign: Eisele-type inlet, AIM Reactor, MION* inlet, and MION2 inlet. The MION* inlet incorporates an added purge flow compared to the original MION design described by Rissanen et al. (2019), and is therefore denoted with an asterisk.

- 4) L.200: Is performance the same as sensitivity here? So, does increasing performance mean a change in calibration factor or what exactly does increase in performance mean here?

We thank both reviewers for pointing this out. We agree that the term “performance” was not sufficiently specific here and could imply a direct relationship with sensitivity. We have therefore revised the sentence as follows:

These reagent ions are highly selective toward H_2SO_4 and other gas-phase acids stronger than HNO_3 , as well as toward oxidation products of VOCs containing more than five oxygen atoms, with the binding strength of the reagent ion to the target molecule increasing almost linearly with the number of oxygen atoms (Hytinen et al., 2018).

- 5) L.226: The text here mentions only species with seven or more oxygen were used, but then on the next page is introduced that species with as little oxygen as three were used, where the analytes with less oxygen only used for the comparison to bromide? If yes it would be could to denote this in Table S2.

The reviewer is correct that species containing fewer than seven oxygen atoms were included specifically for the comparison between the bromide (and iodide) and nitrate modes. This is explained later in the same section (Section 2.3.1), where we state: “However, due to the higher sensitivity of Br^- toward less oxidized HOMs, a set of additional organic compounds containing three to six oxygen atoms were included in both the nitrate and bromide peak lists to enable a more meaningful comparison between the two reagent ion modes in instruments operating with both.”

This is currently also indicated in the caption of Table S2, which states that “In addition to highly oxygenated organic molecules (HOMs), semi-volatile organic compounds (SVOCs) were included to account for less oxidized organic compounds detectable in NO_3^- mode.” To make this distinction clearer, we have rephrased the relevant part of the caption as follows:

In addition to highly oxygenated organic molecules (HOMs), semi-volatile organic compounds (SVOCs) were included to account for less oxidized organic compounds observed in all the reagent ion modes considered.

- 6) L.235 and the following “Sulfuric acid calibration” section: What concentrations of H₂SO₄ were used for calibration? It would be great to also include this in the supplement.

We have now included the range of sulfuric acid concentrations used for the calibration in Section S7 of the Supporting Information.

- 7) L. 435/436: Did you have any hypothesis why the transmission curve of instrument EH1 showed such low transmission factors for higher m/z? Do you think it is an artefact of the transmission calibration method? After all, if I understood correctly, you were able to see some analyte signal over noise level in this m/z range, right?

We suspect that there may have been an issue during the transmission calibration experiment itself, and that the resulting calibration curve does not represent the actual transmission behavior of instrument EH1. As the reviewer correctly points out, and as discussed in Section 3.2, based on the obtained calibration curve, EH1 would be expected to show little to no signal above m/z = 200. However, the normalized mass spectra shown in Figure 3 indicate that EH1 does show measurable signal in the m/z 240–390 range, where monoterpene-derived HOM monomers detected with nitrate can be found, with analyte peaks clearly distinguishable from the background noise.

- 8) L.456: I am not sure that “higher background signal” is the right terminology here, because this is referring to the real measurement timeseries and not the background determination. Therefore, I would suggest to reformulate it to indicate that at low ambient H₂SO₄ concentrations the Eisele-type inlets have higher measurement signal and then correlate that to the higher LOD determined. A similar use of background signal that I suggest to change can be found in L.491 for iodic acid and then in the discussion of the bromide mode measurements in L.532, where again I think “background level” refers to low ambient concentrations.

We thank the reviewer for pointing this out, since the use of “background” in this context is effectively misleading, and we have modified all the sections indicated accordingly:

i) The instruments show good agreement during periods of peak H₂SO₄ concentrations in the central part of the day, whereas larger differences are observed at lower H₂SO₄ concentrations. Specifically, in the latter case, the time series show higher signal for the instruments with Eisele-type inlets compared to those with MION inlets.

ii) The difference at low analyte concentrations between instruments with Eisele-type inlets and to those with MION inlets is even more pronounced in the iodic acid plots (Figure 5a) than in the sulfuric acid case.

iii) With the settings used in the ACTRIS CI-FI1 campaign, ML1 appears to more effectively resolve both the lowest and the peak concentrations for these species.

- 9) L.505 and the description of Table 6 that follows: I am missing some conclusions/discussion on these results, why do some analyte groups show worse

agreement than others? What are the conclusions for the measurement of these compounds and maybe could something be done to increase the robustness of the measurement if one is especially interested to measure such compounds?

We agree that the results in the table would benefit from some additional discussion, in particular regarding the poor agreement observed for the HA species. We have therefore added further considerations in the discussion of the results presented in this section and in the Conclusions. In the discussion of Table 7* (previously Table 6), we now note that the reason for the poorer agreement for HAs could not be clearly identified. In the Conclusions, we have added the following paragraph, discussing the results for HAs and the possible influence of relative transmission differences (as requested by RC1):

(3) For organic and inorganic acids of relatively low molecular mass (below m/z 250) considered in this study (HAs), the agreement among instruments was generally poorer than for sulfuric acid and HOMs. The reason for these larger differences could not be clearly identified. One possible explanation considered was the lack of correction for relative transmission differences for these species; however, additional analysis showed that applying this correction did not improve the agreement among instruments. The results therefore suggest that the measurement of these compounds remains challenging, and that further work is needed to better understand and improve their quantitative determination by CIMS.

- 10) L.537 and point (3) in the conclusion (L.633): Here I don't quite agree that the good agreement of the two Br-mode measurements show that the measurement is robust at different RH levels. The consistent behavior could also just indicate that the two instruments are impacted in the same way by a change in RH. In my opinion the next section comparing it to the NO₃-mode measurements is a much stronger argument that the Br-mode is robust even with changing RH levels. So, I would suggest to give this conclusion after the comparison of Br- and NO₃-mode or to add more arguments why the two Br-modes could not just be influenced in the same way by RH.

We agree with the reviewer on this point and have modified the relevant sections (Sections 3.4, Section 3.5, and Conclusions) accordingly.

S3.4 | Nevertheless, for species that are typically targeted using Br⁻, the two instruments show consistent behavior, suggesting that bromide-mode measurements can may remain robust under field conditions and may be less susceptible to variations in ambient moisture than previously anticipated. However, it cannot be excluded that ML1 and ML2 were similarly affected by the variations in RH, and this possibility is further considered in the following section by comparing the Br⁻ and NO₃⁻ measurements.

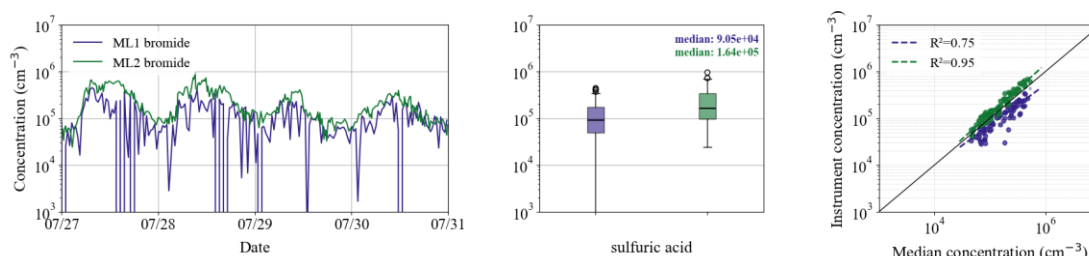
S3.5 | The RH bars in all panels further suggest that RH does not play a major role in driving the performance of Br⁻ mode relative to NO₃⁻ mode, despite the reduced Br⁻ sensitivity expected under the elevated RH experienced during the measurements, further supporting the interpretation proposed in the previous section.

Conclusions | The bromide mode comparison yielded encouraging results for the selected SVOCs, for which ML1 and ML2 showed good agreement ($R^2 = 0.853$; Section S13 in the SI) despite the elevated RH conditions encountered during the

campaign. The comparison of nitrate and bromide mode measurements further indicates that the elevated and variable RH conditions did not strongly affect the relative behavior of Br^- mode for the selected compounds. These observations suggest that bromide-mode measurements can remain robust under ambient conditions for compounds forming sufficiently stable clusters with Br^- ions.

- 11) L.570 and following: Here the differences between the two ML instruments are described and I was wondering: How was their agreement for the H_2SO_4 measurement? Would it be possible to add a figure showing this to the supplement? Maybe a comparison either with timeseries as done in the Fig. 6 left-side plot or in the style of Fig. 7.

We have here included the requested figure. As can be seen, there was an issue with the acquisition of sulfuric acid in ML1 during the campaign. We have investigated this issue, but have not been able to identify a satisfactory explanation for the observed behavior. For this reason, we decided not to discuss the agreement between the two instruments operating in bromide mode in terms of their sulfuric acid measurements.



The poor quality of the ML1 sulfuric acid measurements can also be seen in Figure 7a of the manuscript text, where we have nevertheless discussed the comparison between sulfuric acid measurements obtained in bromide and nitrate modes for each instrument separately.

- 12) And directly following up with a detail in L.571: Should differences in transmission characteristics not be removed by the application of the transmission factor?

We agree with the reviewer, and we thank them for raising this point. We have now removed transmission characteristics as a possible source of the observed differences.

- 13) Figure 8a): Is the dashed colored line in each plot the LOD_{CIMS} value? If not, what is it and could the LOD be added? Why is the line missing for ML2 bromide and AB iodide?

Yes, the dashed lines indicate the LOD_{CIMS} values, and this is now explicitly specified in the caption of Figure 8 (and Figures S7 and S8 in the SI). We thank the reviewer for bringing it to our attention. The LOD_{CIMS} lines are only shown for the nitrate-mode measurements because background measurements were conducted only in nitrate mode (Section 2.3.5), and the LOD_{CIMS} values were calculated based on these measurements. Therefore, corresponding dashed lines are not shown for the bromide and iodide measurements in Figure 8, nor in the time series shown in Figure 6.

- 14) Still regarding Fig. 8: Why was it chosen to not show ML1 bromide when both ML nitrate measurements were shown? If it was only for space reasons I would suggest to reconsider and see if it is possible to show it as well.

ML1 bromide was not shown because, as mentioned in Section 3.1, an abrupt shift in the behavior of the ML1 bromide reagent ions occurred on July 31st. Consequently, the data acquired thereafter in this mode are not considered reliable for the remainder of the ACTRIS CI-FI1 campaign. Since the data shown in Figure 8 correspond to August 1–2, the ML1 bromide measurements were therefore not included.

- 15) L.595 and the following comparison for $C_{10}H_{16}O_5$: Here, I am missing the discussion of the actual analyte concentration as only the variation in concentration is discussed. Why does the iodide AB instrument measurement show the highest concentration? I would assume that this has to do with the difference in calibration factor determination, or whatever the reason case it would be good to mention and explain it.

We agree that it is worth acknowledging the difference in the absolute concentrations reported by the instruments, and we have therefore added a corresponding consideration to the text. At the same time, for the purpose of the intercomparison, we believe that the different response of AB to variations in $C_{10}H_{16}O_5$ concentrations is particularly relevant:

By comparison, under the present sampling conditions and instrument configuration, I^- mode showed substantially poorer agreement with the other instruments and a limited ability appears to resolve variations in $C_{10}H_{16}O_5$ concentrations. This behavior is not unexpected, given the fundamental differences in operating pressure of the inlet between the bipolar instrument and the other mass spectrometers and, consequently, the higher LOD expected for AB. The substantial difference between this instrument and the others may also contribute to the overall higher concentrations reported by AB, although other factors, such as differences in data processing and concentration calculation, may also play a role. In this regard, for the purpose of the intercomparison, the different response of AB to variations in $C_{10}H_{16}O_5$ concentrations is more relevant than the difference in absolute concentration.

- 16) L.615: Ten measurement modes were used during the campaign, but this paper only utilizes data from eight, if I counted correctly, right? If yes, I would mention this as well.

We have revised accordingly, as follows:

In this study, we compared data from six chemical ionization mass spectrometers, some of which operated using multiple reagent ions, resulting in a total of ten measurement modes, of which eight were analyzed in this work.

Formalities:

- 1) When referring to information in the SI it would be beneficial to add the referred section instead of just referring to SI.

We have revised the references to the Supporting Information throughout the manuscript to point to the specific section, table, or figure.

- 2) L.129: “dedicated publications” -Here it would be great to directly refer to the references in Table 2.

We thank the reviewer for the suggestion. Since the references in Table 2 refer only to the CI inlets and not to the mass spectrometers, we have clarified this point by adding the following sentence at the end of the paragraph:

The relevant references are cited in the respective subsections below and, for the CI inlets, are also listed in Table 2.

- 3) L.199: “volatile organic compounds” -as VOCs was already introduced as an abbreviation I would use it here.

We have corrected as suggested.

- 4) L.209: HO₂, the two is not subscript.

We have corrected the formatting.

- 5) L.253: I think the abbreviation MCP was not introduced.

We have specified the full expression before the acronym.

- 6) Supplement Section S5: The text is differently formatted with tighter spacing.

We have adjusted the formatting as suggested.

- 7) Table 4: I would find the unit clearer if the calibration factor included the unit of the signal, e.g. something like cm⁻³·ncps⁻¹

We thank the reviewer for this suggestion. However, we would prefer to keep the unit of the calibration factor as cm⁻³ (or molecules cm⁻³), as this is the unit conventionally used for sulfuric acid calibration factors (see, e.g., Kürten et al., 2012).

- 8) Figure 4: I suggest to remove the data from instrument EL that is not usable from the graphs.

See our reply to request 9.

- 9) In Figure 4 and 5 the median information in the middle panel is quite hard to read, maybe the font could be bold?

Figure 4 has been modified in response to the comments from both reviewers. The following changes were made:

- Major y-axis gridlines were removed, as suggested by RC1.
- Each panel was labelled individually, and the caption was modified accordingly, as suggested by RC1.

- The caption now specifies that points below the corresponding LOD are excluded from the fits shown in the scatter plots, as suggested by RC1.
- Data from instrument EL that were not considered usable were removed from the graphs, as suggested by RC2.
- The median information in the middle panels was shown in bold to improve readability, as suggested by RC2.

We decided not to adopt the suggestion from RC1 to set the y-axis limits of the right-hand figures equal to those of the corresponding left-hand figures, since we believe that keeping the current y-axis limits allows better readability of the point distributions and the associated R^2 values.

Figures 5 and 6, as well as Figures S5 and S6 in the SI, were modified similarly.

10) L.426: “indicating improved despite” -this formulation needs to be reworked.

We have revised the section as follows:

[...] indicating improved agreement despite the lower R^2 .

11) L.472: “only the sulfuric acid concentrations measured by EL remain above the detection limit” -it would be good to add something like “at all times here”.

We have revised the section as follows:

Notably, when considering LOD_{SA} , only the sulfuric acid concentrations measured by EL remain above the detection limit over the period considered, although they approach it during nighttime, whereas measurements from all other instruments fall clearly below it. When applying LOD_{CIMS} , however, measurements from all instruments are above the detection limit.

12) L.481/482: “excluding EH1” -I suggest to mention here again that the difference for EH1 is that no C_{trans} factor applied.

We have revised the section as follows:

[...] whereas HOM concentrations were calculated according to Eq. (5); for EH1, the C_{trans} factor was not applied in this case (Section 3.2).

13) L.523: I suggest to directly write data of which two mass spectrometers are shown.

We have corrected the caption as suggested. We have also made the same correction for Figure S6 in the SI.

14) Some Figure/Table references are broken (L.438 and L.517).

We thank the reviewer for pointing this out.