

Review of “An inter-laboratory comparison of aerosol inorganic ion measurements by Ion Chromatography: implications for aerosol pH estimate” by J. Xu et al.

### **General Comments:**

This manuscript presents an intercomparison of IC measurements made by 10 independent research laboratories spanning three countries. The analysis includes ambient PM filter extracts and certified reference material (CRM). The ambient filter-IC results are also compared to two different ACSMs, revealing some interesting trends. Ultimately, the results are analyzed to understand how differences in filter-IC measurements produce different aerosol pH values modeled with ISORROPIA. In my opinion, this is a novel and quite significant study. The manuscript is well-organized and the writing is clear. There are a few issues that need to be addressed before publication, these are detailed below.

### **Specific Comments:**

The first major issue relates to Section 3.5.1 (“Anion and Cation Equivalence Ratio”) and Section 3.5.2 (“Ion Balance”). The authors are referred to several recent papers that have analyzed the use of these proxies for aerosol acidity (e.g., Guo et al., 2015; Hennigan et al., 2015). The findings are summarized in a recent review on aerosol acidity (Pye et al., 2020). In short, these methods are flawed representations of particle acidity in most environments, and it is advised that they not be used. I suggest removing these two sections, or at the very least, a major revision of these sections is required to accurately reflect the updated assessment of their inability to represent particle acidity.

Another major issue is with Section 3.4 and the method of using the CRM “recoveries”. From what I understand, these are not recoveries in the traditional sense that the word is used with filter measurements (e.g., material spiked onto the filter, then measured in the filter extract), but it is instead more like a QA standard that one would run with a batch of samples. It seems that the “correction” applied in Section 3.4 is actually more like a single-point calibration. The justification seems to be that the CRM solutions may have been prepared more recently than a lab’s calibration standards, allowing less time for artifacts to develop from contamination or volatilization (of water or analytes). However, it is presumed that each of the labs, themselves, used a CRM for their calibrations, though this is never stated. If true, then perhaps the result here speaks more to guidance on the frequency of calibrations. Also, it is good analytical chemistry protocol to run blanks and QA standards regularly interspersed with sample analyses. The QA standards should be made from CRM, as well. There are typically pre-defined limits of acceptable performance, if the QA standards fall within this predefined accuracy, there is no adjustment to the sample results. To summarize: each lab’s calibration procedures should be detailed (likely in the SI); as should their QA/QC procedures for running a batch of filter samples. If it is found that a lab’s calibration is off, then adjustments are warranted (again the QA standards should signal this), but if QA standards fall within a predetermined accuracy, then a “recovery” correction is questionable.

While not the central focus of the manuscript, the differences between the ICs and the ACSMs are significant and deserve more attention. The explanation in lines 312-314 seems highly

unlikely (volatilization from filters) since sulfate had very similar behavior. Two ideas that were not discussed but which could have contributed to the discrepancies are: 1) differences in the performance of the two PM<sub>2.5</sub> cut-point selectors, which could lead to different transmission of particles in the 2-3 μm range, and 2) the collection efficiency applied for the ACSMs. On the second point, the collection efficiency is often assumed to be 0.5, and is applied to an entire data set even though this changes with aerosol composition and meteorological conditions. Therefore, an applied CE of 0.5 when the actual CE was closer to 1 would also produce the observed results. Further, line 339-340 states that “it is essential that the filter-based observations are robustly quality controlled before any ACSM and IC intercomparison,” but what about robust quality control of the ACSM measurements? What was done for this study, and what improvements should be made?

### Technical Corrections:

- Line 48: suggest “factor of three” instead of “3 times”
- Line 69: suggest deleting the second ‘to’ in this sentence
- Line 77-78: sentence awkward as written, suggest rephrasing
- Line 84: suggest replacing “nowadays” with “at present” (or similar)
- Line 183: change to “COD value equals to...”
- The COD value of 0.269 (line 187-189, Figure 4) seems highly arbitrary. There really is no discussion about the 0.20 vs. 0.269 – I suggest picking only one.
- Line 320: was an statistical outlier test actually performed? Be careful excluding data without reason, especially for a smaller *n* like this study.
- Line 328: suggest “factor of three” instead of “3 times”
- Suggest adding 1:1 lines to Figure 2
- Section 3.2.3 is quite short, and could be omitted, or incorporated into another section. Fig. 3 does not add much of substance to the overall discussion.
- Line 523-524: what is the hypothesized reason for the highly different pH value from Lab-9’s data on this date?
- Line 565: I disagree that there are “substantial” uncertainties, especially in the major ions measured by IC. The uncertainties are quantifiable, and they seem to be in line with previously published values. For the “minor” ions, the threshold for major/minor is never stated, but clearly this is true for Ca<sup>2+</sup>, while most of the others seemed to perform quite well.
- Line 582: yes, but I would note here that the median ammonium recovery was close to 100%.

### References:

Guo, H., Xu, L., Bougiatioti, A., Cerully, K. M., Capps, S. L., Hite Jr., J. R., Carlton, A. G., Lee, S.-H., Bergin, M. H., Ng, N. L., Nenes, A., and Weber, R. J.: Fine-particle water and pH in the southeastern United States, *Atmos. Chem. Phys.*, 15, 5211–5228, <https://doi.org/10.5194/acp-15-5211-2015>, 2015.

Hennigan, C. J., Izumi, J., Sullivan, A. P., Weber, R. J., and Nenes, A.: A critical evaluation of proxy methods used to estimate the acidity of atmospheric particles, *Atmos. Chem. Phys.*, 15, 2775–2790, <https://doi.org/10.5194/acp-15-2775-2015>, 2015.

Pye, H. O. T., Nenes, A., Alexander, B., Ault, A. P., Barth, M. C., Clegg, S. L., Collett Jr., J. L., Fahey, K. M., Hennigan, C. J., Herrmann, H., Kanakidou, M., Kelly, J. T., Ku, I.-T., McNeill, V. F., Riemer, N., Schaefer, T., Shi, G., Tilgner, A., Walker, J. T., Wang, T., Weber, R., Xing, J., Zaveri, R. A., and Zuend, A.: The acidity of atmospheric particles and clouds, *Atmos. Chem. Phys.*, 20, 4809–4888, <https://doi.org/10.5194/acp-20-4809-2020>, 2020.