Atmospheric Measurement Techniques Discussions

Atmos. Meas. Tech. Discuss., referee comment RC1
https://doi.org/10.5194/amt-2022-107-RC1, 2022

Comment on amt-2022-107

Anonymous Referee #1

Referee comment on "Intercomparison of in situ measurements of ambient NH_3: instrument performance and application under field conditions" by Marsailidh M. Twigg et al., Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2022-107-RC1, 2022

Review of "In-situ measurements of NH_3: instrument performance and applicability" by Marsailidh M. Twigg et al.

Summary: Ammonia (NH_3) is not regularly monitored in networks. It is also an unregulated pollutant in many countries. Identifying a standard among the available technology for routine gas-phase ammonia monitoring is vital to our understanding of this pollutant and how future regulation policy can be shaped. This paper describes a comprehensive intercomparison of 13 instruments for sampling gas-phase NH_3. The study is done in-situ at an agricultural field site in Scotland. The instruments represent a variety of currently available technologies for measuring gas-phase NH_3. The authors perform a comprehensive analysis of the available data and show that variability in the ensemble is within 20%. This alone is an interesting finding from an intercomparison of 13 independent NH_3 instruments of varying techniques and time responses. While there are still nuances of instrument setup, maintenance, and operations to be determined, the observations reported in this work are a step forward towards developing standardized practices for NH_3 monitoring.

The paper presents findings from a new field study and the topic is highly relevant towards addressing current air quality and climate concerns. Thus, it fits the scope of AMT. I recommend this paper for publication in AMT following minor revisions.

General Comments:

In general, the flow of the paper could be improved by moving around some of the sections and refocusing the key points in the introduction to match the outcomes. For example, the introduction of this paper implies that the authors will provide recommendations on how to achieve high quality future routine monitoring of NH_3. However, the conclusions do not provide specific recommendations for what the optimal inlet setup and operating/maintenance procedures could be for a monitoring site. In

contrast, a key conclusion of this work that is not identified earlier in the paper is that the variability in the ensemble of 13 instruments is within 20%. Such a tight cluster of NH_3 measurements from 13 independent instruments actually seems pretty good, especially considering that gas-phase NH_3 can be challenging to measure. It would help to contextualized this finding better in the text in terms of what this could mean for monitoring networks comprised of a few types of NH_3 instrument techniques.

There are some very detailed instrument descriptions in the methods section and some that are rather vague. It would be helpful to include a similar level of detail for all instruments, especially in cases where references are not available. Please also clarify the purpose and design of the experiments using identical instruments with different inlets earlier in the paper. This is an important factor in the final recommendation, but there is little information leading up to these results to prepare the reader to understand these findings.

Please include more context in the methods section about how the calibration sources were used in these experiments and what instruments they were used with. Were these calibrations only applied in the field, or were instruments calibrated individually in a laboratory before/after the field experiment? Please clarify in the text. It could also help the flow of the paper to move up "Section 3.9 Ammonia calibration system" to between Sections 3.2 and 3.3.

Zeros were routinely performed for the QCLAS, but what about the other instruments? Were span calibrations also performed? Are zeros and span calibrations stable over time? Were corrections associated with these zeros and spans applied to the data prior to analysis in Section 3? Please clarify in the text.

Instruments were in the field from 22 Aug to 2 Sept. Is there a reason that only 23 Aug and 29 Aug is used for the intercomparison study and not the full period? There are inconsistences in the dates in several places throughout. For example, Figures 8 and 9 and Table 4 suggests a start date of 22 Aug instead of 23 Aug. There are also inconsistencies in the dates in some of the figure captions and table headers in the SI.

A major concern for Section 3 is that the bulk of the analysis relies on comparing each instrument to an ensemble median rather than a true, validated reference (e.g., Figures 8 and 9). While this might be the best approach for this case study, it should be specified in the text that the median is not an independent variable (i.e., it is by definition driven by the range of observed values from the different instruments). If the ensemble median is not independent, then the correlation analysis might be better described using 2-sided linear least squares fits (aka. orthogonal distance regression). See additional specific comments below.

Specific Comments:

Title: Could the title be more specific to the experiments and findings? Maybe something like "Exploring best practices for gas-phase NH_3 monitoring: An in-situ comparison of 13 instruments in Southeast Scotland".

Abstract: The Abstract could be streamlined a bit by relocating some of the context to the Introduction. It could also be added that NH_3 is an unregulated pollutant in many countries. Listing each instrument make and model could be solely addressed in the methods section.

P3, L15. Flip phrase to read as "not all NH_3 is captured"

P3, L35. There are many similarities between this work and a prior report by von Bobrutzki et al., 2010, which is referenced several times throughout this work. From the text it seems that the main advantages of this intercomparison are the addition of newer instruments/technology and the evaluation of traceable gas standards. Are there other advancements in this work in terms of the experiment objectives, experiment design, and application of lessons learned from von Bobrutzki? For example, this work uses pairs of identical instruments outfitted with different inlets to characterize artefacts due to the setup itself. This is a unique feature of this experiment that ought to be clearly outlined as a focus of this paper in the intro. The experiments related to this comparison and any additional setup should be clearly described in the methods section.

Figure 1. This is a nice photo of the field site. It gives a lot of perspective for the experiments. It would help to add a wind rose or an arrow to show the predominant wind direction during the study period.

Table 1 is missing accuracy, precision, and range information for the miniDOAS #1 - add symbols like n/a or (-) unknown. The LGR #1 and LGR #2 response times look to be flipped. The LGR#2 with the higher flow rate should have a faster response time of 1 s. There are 13 instruments compared in this study, yet there are 15 rows in the table. Should the OGS and the ALPHA sensors be separated from this table or distinguished in some way. Maybe this is simply fixed with a footnote to clarify the usage of the OGS and ALPHA samplers. It would also be helpful to further explain the dependency of the Picarro#2 and the OGS in a footnote in the table. Further, the acronym OGS has not been defined yet in the manuscript.

Table 2. Add another row to specify if the inlet components are heated or not. For instruments that do have heated inlets, at what temperature are they maintained? For consistency, change "N" to "No" for the AiRRmonia#1 filter. Is "diameter" meant to be the inner diameter of the tubing or the nominal outer diameter? The i.d. of the tubing will be the most relevant for your residence time calculations, so that could be the better parameter to include in this table. In either case, please clarify.

Is there a reason that the manifold inlet and manifold itself are made from different materials? Is there any research on how NH_3 sticks to uncoated Pyrex surfaces?

I find it interesting that all of the instruments on the common inlet use a filter. Was that planned?

Some additional explanation of the parameters for the unique inlet system associated with the QCLAS are likely needed in Table 2 and Section 2.2.3. This instrument setup is unique in that it uses a heated inertial inlet with a critical orifice to separate particles from the airstream. It also requires a rather large capacity scroll pump to create a sample flow rate of 13 l min^{-1}. It is also important to distinguish the size of the critical orifice (~1 mm) located inside the inertial inlet compared to the size of the tubing (typically 3/8" o.d., ¼" i.d.).

P8, L9: How often are the passive ALPHA samplers collected and analyzed? Please include this information in the description in the methods section.

P8, L34: It isn't clear here why the correction was not included. I think this is explained later in section 2.3.1. It would help to add a reference here to the explaination in the later section.

P8, L37. I suggest moving the last sentence of this paragraph to Section 2.3.1 where the OGS is described. This information gets lost here and is better served in the other section.

P9, L16: Background subtraction from routinely measured zeros seems to be another unique feature of this instrument's operations. Were any other instruments routinely zeroed throughout the measurement campaign? It seems like this would be a fairly important step for all instruments to accurately report NH_3 mixing ratios. It is also interesting that N_2 was used for the zeros instead of zero air. While using N_2 should not impact a zero calibration, a prior study showed that span calibrations on top of N_2 compared to zero air produced a spectroscopic artefact up to 10% (Pollack et al., 2019; https://amt.copernicus.org/articles/12/3717/2019/).

Section 2.3.1: What wavelength does the OGS operate at? What is the uncertainty of the absorption cross section (or line strength) at this wavelength? How does this impact the overall uncertainty of the OGS calibration system?

P11, L36: "associated with long stabilization times"

Section 2.3.2: Please specify the carrier gas used. N2 versus air can have different effects on NH_3 permeation devices and spectroscopic artefacts.

Figure 2. Please add something (e.g., shaded bar, horizontal arrow, red dashed lines) to indicate the period of data used for the instrument intercomparison (which I think is 23 Aug to 29 Aug). Is there a reason that the remaining data in the timeseries up to 02 September was not included in the intercomparison?

Figures 3 and 5 make me wonder how the instruments were time aligned prior to correlation analysis. Were inlet delays account for prior to averaging to 1 hour? Please clarify in the text.

Figure S2. This is an interesting figure. It seems as if there was less atmospheric stability during the intercomparison period, yet much larger stability during the period immediately after the study period (29 Aug thru 2 Sept). How do the instruments compare during the more stable period? Are there any observable differences? This would be an interesting comparison to add to this study that would be within the scope of this study. Also, there is a typo in the caption; the date range should be from 22/08/2016 to 03/09/2016.

Figure 4 could be moved to the SI if you feel this paper is too lengthy.

Section 3.4: Can you comment on the effects of a heated vs. unheated inlet? See Ellis et al., 2010 for reference. Some additional comments about the utility of a heated inertial inlet for filter less separation of particles could be included here.

Section 3.3: Does the ensemble median include LGR#1? If yes, how does the ensemble median and related statics change if this measurement is excluded?

Figure 6. Change CV limit to 20% in caption to be consistent with figure and discussion text.

P18, L13: Is the response time truly different under ambient conditions? Without the same level of fine structure in Figure 7b as in Figure 7a, it is difficult to accept the levels of smoothing applied to the DOAS under ambient conditions to match the profiles of the AiRRmonia and Picarro instruments. Can you include another trace in 7b to show the DOAS signal with the same level of smoothing as in 7a? This would help highlight whether additional smoothing of the DOAS is indeed needed to match the features of the AiRRmonia and Picarro instruments under ambient conditions.

Figure 8. The met filter looks like it could have induced some bias in some of the fits. The met filter was applied to eliminate low wind speed and unstable conditions that could have led to inhomogeneity between the inlets at the field site. But didn't the ALPHA samplers indicate that there was homogeneity during the study period? What do the fits look like without the met filter? At a minimum, you should comment in the text about any differences in fits with and without the met filter applied.

Section 3.6: What would you consider to be a reasonable deviation from 1 for the slopes? How does this compare to the deviation in the ensemble or with the reported measurements uncertainties in Table 1? For example, if the spread around the ensemble median is 20%, would slopes ranging from 0.80 to 1.20 be considered good?

It would also be interesting to see how the fast instruments compare with an ensemble average of only the fast instruments (like on a 1 min average timescale). Was this something you tested? Can you comment on whether the results would be different?

Data in Figures 8 and 9 are split into NH_3 > 10 ppb and NH_3 < 10 ppb. This was done to be able to best compare the NH_3 < 10 ppb data with the findings in von Bobrutzki. Please include additional discussion of how this work compared with the findings in the prior work. Does separating the data points by NH_3 mixing ratio prior to intercomparison analysis generate any bias? It would be helpful to see the results of the intercomparison fits using all data, which could be a nice figure in general to include in the SI.

At the risk of cluttering the figure, it also seems appropriate to also include the actual fits in each plot.

P24, L1: Are the least squares regressions 1-sided or 2-sided? Please specify. See general comment above.

P24, L6: It seems as if the LGR#1 instrument was having some issues during this experiment. Would it be better to exclude the data from LGR#1 from the paper altogether? Do you have specific reasons for keeping it in this intercomparison? Was the LGR#1 was included in the ensemble median? Please clarify your reasoning in the text.

Figure 10. What does the color scale correspond to? Please clarify in the caption. Add labels "a" and "b" to plots.

P26, L15: What the was the temperature of the LGR#2 inlet? Was it high enough to thermally dissociate enough NH_4NO_3 to impact the measurement? Gentle heating (<40 degC) might not have a huge impact on the measurement on the sampling timescale

(e.g., Fig 4a in Huffman et al., 2009; https://doi.org/10.5194/acp-9-7161-2009).

Figure 11. Since the instruments compared in this figure all have relatively fast response times (1 minute or less), would it be more realistic to compare them on a 1-minute timescale?

Figure 12. Please clarify what height means in the caption. Does it have units? It would be helpful as a quick reference to include labels in the figure about what the different grouping are. You could easily add this by changing the outline colors to match a legend or a description in the caption?

P29, L11: Do you mean Figure S4?

Table 4: Fix table header to correspond with proper figures in the SI (S4 and S5). Add a footnote to highlight the relationship between Picarro#2 and OGS.

P34,L17: Based on lessons learned from this study, can you provide a recommendation for how often routine calibrations (zeros and spans) ought to be performed if one or more of these measurement techniques are used at a surface monitoring network site?

Section 4.3: Based on lessons learned in this study, can you provide additional recommendations about inlet setup? It is not surprising that instrument manufacturers do not specify a schedule for calibrating and servicing, as it is largely dependent on how, where, and under what conditions the instrument is utilized. Instead, there could be a list of indicators to watch for that could signal a user to perform routine maintenance. For example, a prior works (Pollack et al., 2019; Ellis et al, 2010) showed that increases in the response time to a step change in NH_3 from a calibration source was a good indicator of when the instrument inlet needed to be cleaned.

Figure 15. There are two panel b's in this figure.

Figure S1. There are inconsistencies in the labels used in the caption (a, i, 2, 3). The time resolutions in brackets are missing for some instruments in the figure. What do you mean by "raw" data in the timeseries? "Raw" typically implies uncalibrated data. Please clarify.

Table S3. Header has a reference error. The dates in the period do not make sense. Are these response times truly meant to be in units of minutes, or should they be seconds? E.g., 100 minutes for a 1/e time for the Picarro seems rather long.

Figure S3. Were the datapoints associated with breakthrough eliminated from any intercomparisons with the MARGA? Please clarify.