Response to reviewers for manuscript

Dear Senior Editor,

Thank you for the opportunity to revise our manuscript, “An uncertainty methodology for solar occultation flux measurements: ammonia emissions from agriculture.” We greatly appreciate the insightful comments from both reviewers, which will enhance the quality and clarity of our paper. In response to Reviewer #1's comments on the structure and novelty of our work, we have further clarified our contributions. Besides the innovative plume height methodology, we also emphasize our advancements in the use of SOF for NH_3 quantification, the application of the GUM methodology for uncertainty estimation particularly and blind validation tests. These aspects collectively underscore the novelty and significance of our research in this field. Reviewer 2# had fewer comments, with the emphasis on the deposition of NH_3, which we explained that it diverges from the goal of the paper, to dive deeper in the deposition impact on the quantification. Regarding Reviewer #2's focus on NH_3 deposition, we have added a brief discussion to clarify that while deposition impacts are outside the primary scope of our paper, our findings indirectly inform this area. This addition aims to acknowledge the relevance of deposition in the broader context of NH_3 quantification without diverting from the main objectives of our study. As feedback from the reviewer #2, the title of the manuscript was changed to “An uncertainty methodology for solar occultation flux measurements: ammonia emissions from livestock production.”

We look forward to your decision.

Best Regards,

Johan Mellqvist, Nathalia Thygesen Vechi, and co-authors

Reviewer Comment	Author' Response	Revised Text – Line numbers refer to clean (without track changes) version of the revised paper.
Reviewer #2		
The authors address a comprehensive approach to quality uncertainties of measured ammonia emissions from livestock sources by using SOF & MeFTIR instruments, which is well fit for the scope of AMT. Meanwhile, scientifically speaking, understanding the level of uncertainty is crucial for the reliability and validity of the results in the Nitrogen domain. The authors validated such error estimation method using tracer release experiment (single point source) and later also applied it into multiple livestock farms case studies (area sources), which can further help mitigate the possible causes of errors in real-world. In all, it is worthy of scientific publication. But before that, I do have some specific comments related to the current error propagation methodology and the structural of the paper as well, which I'd hope the authors can remedy certain issues and deepen the discussion.	**Thank you for the comments and feedback**	
Some specific comments: 1. It is well known that ammonia flux has its bidirectional character. Although notably the authors stated it is out of scope for the present work, I'd like to know how problematic would it be to fully	**Clarification provided** The paper was focused on presenting the method. As well as the error was focused on the methodology error, so it would not be correct to add in the error budget. We could have tried to estimate, but it would vary a lot from case by case, so it would	**Added line 506:** Regarding NH_3 deposition, it will very largely according to the conditions of each site. In California for example, where both case studies 2 and 3 were performed, previous studies measured a deposition of 15% in the first 3 km, while others estimated to be from 8 – 15 % (Miller et al., 2014). For Denmark and France, these number might be higher because there was likely less convection during the measurement days,

	ignore it from the error propagation? How large will it affect the error ranges? For example, the authors directly compared the tracer emitted NH_3 value with the SOF measured value ~200 meter away from the source at the downwind direction. This possibly causes the current error propagation scheme (methodology) to tentatively overestimate the actual 'error', because the ammonia 'deposition loss' is not corrected before the error propagation starts. In the other word, it might not be fair to name ammonia deposition loss as part of the total measurement error, isn't it?	need a whole methodology to estimate that in addition to what we already had. We added some extra information on this in the last section.	but this paper focused more in describing SOF, rather than investigating the emissions sources.
2.	Some comprehensive discussions are missing from the current paper structure. I'd think it can add some extra for this paper if the authors can properly address the above mentioned issue, by adding some comparison cases studies, or discuss both limitations and advantages of the current method and results in the structural way.	**Not implemented** We are uncertain about the specific point the reviewer is addressing. Our understanding is that it may be an extension of the question regarding deposition. Consequently, we refer to our response to that question and the additional text provided in the previous comment for clarification.	**No modification**
3.	Both in the abstract and conclusion parts, the authors emphasized the SOF method estimated NH_3 emission can be as low as 1 kg/h ± 21%. Does it somehow indicate the SOF method no longer trustworthy if the	**Clarification provided** It is most likely the quantification limit from the method, considering the precision of the instrument. It might work for lower emissions at certain meteorological conditions. We added the 0.5 information in the abstract,	**Added line 23:** (~0.5 - 1 kg h^{-1}) and

ammonia source emission below 1 kg/h? As it can influence the applications for potential users, can the authors justify your statement? Is such statement applicable for all agricultural sources or need further testing?	because we believe that even at the 0.5 kg/h the measurements had good quality.	
Some specific technical comments are listed below: 1. The title of this paper is 'An uncertainty methodology for solar occultation flux measurements: ammonia emissions from agriculture'. It seems too broad to cover all agricultural domain. The agricultural activities include fertilizer application, livestock operations, and other agricultural processes as well. However, none of the study cases listed in the present work demonstrated its application for fertilizer or manure field emission measurement yet. All cases are focused on livestock farms. Would it be more appropriate to narrow down the title to 'livestock' ?	**Suggestion implemented** We agree we the reviewer suggestion, the title should be changed.	**Paper's title was modified**
1. In Line 19, how does the plume height estimation reducing the measurement uncertainty? I could not find detailed discussion in the main text. please reconsider its	**Clarification provided** We have not calculated the effect explicitly but used the height for our wind calculations.	**No modifications**

value when it is mentioned in the abstract?		
2. The last sentence in the abstract should be removed. The main paper does not provide concrete evidence or results demonstrating the applicability of the methodology to other gaseous species or purposes, it would not be appropriate to make such claims in the abstract. The abstract should accurately summarize the scope and results of the study based on what is presented in the main paper. It's better to phrase it as a potential avenue for future research or exploration rather than making definitive statements. Mentioning the potential for broader applications in the discussion section of the paper, along with the need for further research in those areas, would be more appropriate.	**Suggestion implemented** The sentence was changed	**Rephrased Line 25:** This paper's findings offer potential for broader applications, such as measuring NH_3 fluxes from fertilized fields, as well as in the oil and gas sector. However, these applications would require further research to adapt and refine the methodologies for these specific contexts.
3. The caption of Figure 2 mentioned case study C3 out of blue without any other context in the previous text, please add some extra information or consider reorder the main text structure.	**Clarification provided** We have removed the information that it was from case 3	**Rephrased figure 2 legend**: a) Example of wind profiles, the grey lines show …
4. In line 170, both 2D sonic anemometer and vane wind meter are used in multiple campaigns. How do different	**Clarification provided** The measurements error of the wind sensor were not taken into account since they	**No modification**

types of wind meters contribute to the final measurement error? Do you take the instrument system error into account when estimating the wind profile error?	were considered small compared to the absolute wind uncertainty.	
5. In line 238, absorption strength uncertainty of 2% was assigned from a previous study, is it always a fixed value for all SOF instrument in various application? If not so, can it be properly estimated or not?	**Clarification provided** This is based on the cited reference according to the text and is applied in all the instruments using HITRAN NH3 absorption cross sections.	**No modification**
6. In line 255 equation 7, the 'abs' value is used. Can you please define what is it? As in equation 6 there is another abs(960-968). Are they the same? Similarly, in equation 8, it is Aabs. They become bit confusing.	**Clarification provided** There was ambiguity in the accompanying text and in equations 6 and 7; this has been fixed.	**Modifications:** Definition of abs_{avg} has been corrected in the text (line 269) Terminology has been changed in Eq.6 to abs_{avg} Terminology has been changed in Eq.7 to abs_{avg}
7. Some confusing statement occurred in both line 304 and 308. From the previous context, in line 304 C2 should be C1; and in line 308 C1 in fact refers to C2. Please double check.	**Suggestion implemented** The reviewer is correct, the suggestion was implemented.	**Rephrased line 324:** In the validation test and C1, two wind masts … **Line 328:** For C2, only one 10 m mast was used …
8. In line 317 Table 1. Case study C2 integrated wind profile is estimated using C3 data. Can you explain why C2 can use C3 wind profile data? I assume they were not measured in the same day nor the	**Clarification provided** The campaigns were performed in nearby area and same month but different year. The flux calculation was not performed with the C3 data, instead, information from C3 campaign was used for the uncertainty	**Rephrased line 328:** For C2, only one 10 m mast was used to measure the wind, so we estimated the error of choosing different vertical profiles by using information from another study at the same geographic location and at similar time of the year, because of the lack of data to estimate the real wind profile …

same location, so why do they share the same IWPavg value?	calculation in the estimation of the wind speed average profile.	
9. Line 360, MeFTIR and SOF both used to estimate plume height. However, the current method did not mention MeFTIR measurement uncertainty at the ground level. How large uncertainty can be generated from MeFTIR system? and how large can it contribute to the total error prorogation?	**Clarification provided** The uncertainty of the MeFTIR concentration measurements is detailed in Vecchi 2023, typically ranging between 5-10%. We have not estimated the uncertainty of the plume height estimation in this paper, and therefore, the MeFTIR concentration uncertainty is not relevant here. Note that we only used the standard deviation of the individual plume height measurements as a measure of their variability, as illustrated in Figure 6.	**No modification**
10. Line 367, the authors stated PTVS method result is slightly lower than VCGS method, but they are similar". I would not agree so. In Figure 6b, it clearly shows PTVS results can be 2 times bigger than VCGC method in many cases.	**Clarification provided** We phrased the sentence so instead we used the calculated difference between the two methods. Highlighting that small difference might not large interfere in the calculation, especially if they are at higher altitudes. The differences might be also due to the some of the plumes not have an homogenous mixing.	**Rephrased line 387:** … was utilized and compared to the VCGC method, showing that the first method produced in average emissions 35% higher than the second, where only one of the farms (farm 8) had a large difference.
11. Figure 6b, what do you mean "farms" in the title of x-axis? It never mentioned in the main text. Or do you mean "transect" number from certain case study ? please clarify.	**Clarification provided** We added this information in the figure legend. It means that we used the data from SOF measurements at nine different farms to illustrate the plume height methodology. Because the measurements were made at similar conditions (wind speed and distance) the expected plume height at all the respective farms measurements would expect to be similar.	**Rephrased in figure 6 legend**: b) Examples of average plume height calculation from measurements at nine farms using …
12. Line 383, Figure 7b and line 403 all related to the Sep 22		**No modification**

measurement result, which was measured in a cloudy day. Does such data point still validate? Firstly, it is conflict with the previous statement that SOF is better used in sunny and less cloudy day. Secondly, even without 'deposition loss' correction, SOF measured downwind value is higher than the actual tracer release value. This strongly suggests such data may be too faulty to be trusted.	**Clarification provided** High-level clouds and haze can lead to greater measurement variability, but all measurements taken on September 22 were valid according to the quality criteria of the measurements	
13. In Table 4 , for C2 and C3a studies, 3 and 7 transects were measured, respectively. Elsewhere, the author stated that 12 to 16 transects should be applied. If so, will the small sampling size significantly enlarge the final error? Can you demonstrate it further?	**Clarification provided** According to the text, the 12 to 16 transect are for measurements in refineries according to the referred normative. For farms, 5 to 6 transects are ideal because there is less variation.	**No modification**
14. In Figure 8b caption, the flux on the figure corresponded to 0.55 kg/h. However, in the main text line 425 said 'the SOF could measure concentration as low as 1 kg/h with an uncertainty of 21%'. But 0.55 kg/h is below such threshold. which statement is true? What is the lowest detection limitation to use SOF measuring Nh3 sources? Shall it be further explored or is it proved by current study?	**Clarification provided** Yes this is inconsistent. We added 0-5-1 kg/h in the text.	**Added line 23:** (~0.5 - 1 kg h^{-1}) and