Reply to reviewer number 2, Anonymous review

Dear Anonymous Referee,

Thank you for your thorough review of our manuscript. Hereby, I present the replies to all your comments (in green).

General comments

Dear Editor,

In this article, the authors present a declination of an already existing methodology of methane emissions quantification with the use of an alternative instrumentation. They present laboratory and real fields tests of this instrumentation, to better characterize its specificities. They also performed real field quantifications of methane sources, with both the reference instrumentation and the newly tested instrumentation in parallel, which allows a direct comparison of the performances of each method. They also performed a detailed statistical analysis of some of the main sources of uncertainties associated with this emissions quantification, based on one of the tests performed on the field.

They highlight the strengths of this methodology and the strong interest of such low-cost instrumentation for a generalization of this type of measurements of methane emissions.

This paper has a great interest in the current context where many groups are developing airborne methods based on Uncrewed Aircraft Systems (UAS) to monitor methane emissions at the facility scale, filling a significant gap in the greenhouse gases emissions reporting. Furthermore, the emphasis on the uncertainty analysis is very beneficial as this is crucial for the validation of such method.

The topic of this article corresponds to the scope of Atmospheric Measurement Techniques. This paper is generally well written and very clear. However, there are still minor aspects which could be improved before final publication.

General comments

The authors presented intercomparaison of emissions quantifications based on Axetris and active AirCore measurements. For the data treatment of the Axetris measurements, the authors applied an H2O correction based on AirCore H2O measurements.

- What is the reliability of H2O measurements of an active AirCore? In the case where the Axetris would be used as a stand-alone instrument without AirCore, the authors suggest that the H2O values can be derived from model or nearby meteorological station values. But they don't present any estimates of the performance of the quantification method using such values, compared to AirCore H2O measurements. It would be interesting to evaluate the performances of the method without having access to AirCore H2O measurement, as the Axetris should at the end be employed as a standalone instrument. An alternative could also be to embark an extra H2O sensor on-board the UAS on top of the Axetris. It would also be interesting to compare the

AirCore H2O measurements with the humidity values measured by the embarked Trisonica mini, if they have been logged.

- Thank you for your comment, and that is a fair point of critique. However, it should be noted that the water vapour concentration observed by the LI-7810 were used, and not from the AirCore directly. During field deployment, the AirCore samples ambient air through a chemical dryer, removing most of the water before analysis. The AirCore measurements are therefore corrected differenty compared to the Axeteris, which still needs a water vapour dilution correction.
- The water vapour concentration during this field experiment is determined by the LI-7810, which was able to sample ambient air in between flights. The average value obtained from this was used for the general water vapour dilution correction. We added the additional suggestion to obtain the approximate [H2O] from RH measurements reported by local meteorological sites, in case one does not have a LI-7810 at hand. This convenient approximation contributes only the tiniest error to the method (e.g., if using, at 10°C, 65% RH instead of a true 70% RH, the commensurate [H2O] is 0.79% vs 0.85% – resulting in an error in the mass emission rate estimate of 0.06%).
- The addition of an extra H_2O sensor is a really good suggestion and is also something that we are considering for the next version of the Axetris package. We do log the humidity values from the TriSonica mini but believe the meteorological data from the nearby stations is sufficient.

- Despite the good quality of the monitoring methodology and the uncertainty analysis, there are still some gaps in the protocol and the analysis, which are still limiting the scope of this study. One would expect more emphasis on these limits in the discussion.
  - We added some ideas for future studies in the Discussion to address how to ensure some of these limitations can be brought to light: “Future studies should concentrate on a systematic comparison between the Axetris and high-precision, high-frequency in situ instrument. Similar to this study, both sensors should be deployed simultaneously, to allow direct comparison to mass emission rate estimates using comparable data analytics. Furthermore, large-scale controlled release experiments could provide a quantitative assessment of the sensors’ performance.”
  - Additionally, we added some additional points to the discussion of out uncertainty analysis to highlight what was not considered during this analysis: “The statistical analysis provides insight into potential sources of uncertainty in UAS-based sampling. The OU-simulation is appropriate for the specific flight trajectory presented in this study. The uncertainties derived using the simulation remain reasonable when real-world flights are conducted comparable to those simulated. Deviations may arise when observed mass emission rates are highly variable over a short time (i.e. one hour) or when the flight path (specifically: sampling density) is sparser than the modelled configuration. In those cases, the simulation results may no longer be representative. Flight paths of all flights in this study are comparable to flight 1, therefore OU-derived uncertainties can be applied to all flights.
    However, not all sources of uncertainty have been considered. First, this study did not include the effects of spatial sampling density, which can also affect

the accuracy of the mass emission rate measurements. Significant data gaps negatively affect the MBA, leading to systematic underestimation of the mass emission rate. This limitation has been addressed in detail by Mohammadloo et al. (2025). Higher-frequency observations in the horizontal and vertical directions decrease the estimated emission rate error. Based on their findings and our own flight track featuring relatively dense horizontal and vertical coverage, we expect to have only an additional 5-10% error. Furthermore, wind direction uncertainties are also not included in the statistical analysis. Changes in wind direction influence the dispersion and detection of atmospheric plumes and is often referred to as the meandering effect (Wietzel et al., 2025). As mentioned before, wind variability (e.g. changed in wind speed and wind direction) propagate non-linearly through the MB calculations, therefore affecting the mass emission rate estimates in varying ways, depending on the error. Future work should combine optimised flight-path design with real-time plume detection to further minimise spatial sampling bias."

- Concerning the characterization of the instrument: the authors performed an experiment to evaluate the stability of the analyser to variations of the temperature of the air surrounding the instrument, but the temperature of the incoming air is not mentioned (probably laboratory temperature) and also not varying. This would be useful to test the influence of the temperature of the incoming air, as it might probably influence the cell temperature stability: different temperature gaps between the instrument regulated temperature and the incoming air temperature could be tested, to simulate the behaviour of the instrument in different types of environments or seasonality. One could also extent these tests to measure the impact of temporally varying temperatures on one side (simulating the variations which might be encountered if measurements are performed in an industrial environment with warm plumes (under the wind of natural gas flares for example). Regarding the humidity sensitivity experiment, the authors proposed a humidity generation method based on a heated wet paper towel. This allows a strong variation of the humidity of the incoming air, but the variation of the temperature of the incoming air is not discussed. Could it have an impact of the spectroscopic response of the instrument? Ideally, humidity sensitivity tests rather be performed at constant temperature. This also justifies the need of the already mentioned sensitivity test to incoming temperature.
    - Thank you for your comment. This is indeed something we did not consider directly. However, we believe that having the active temperature control to stabalise the cell temperature and therefore also limit the temperature variability. Of course it can be possible that the sensor is sensitive to the temperature of the incomming air, but seeing the great improvement with this temperature control we believe this effect to be negligible. Even when varying air temperatures are sampled, the sensor readings are more stable with the active control compared to without.
    - Numerical example: at a pumping rate of 200 sccm, air of 10°C flowing into the cavity of 25°C would have a cooling capacity of 0.05 W. If the incoming air temperature would vary by 5°C (with time, altitude, flight direction or similar), that would vary the cooling capacity between 0.03 and 0.07 W. That variability certainly is of absolutely no consequence. Any very gradually accumulating effect is effectively compensated by the controlled heater (max. power: ~2 W).

- Besides the smapling pathway during the water vapour experiment was more than 2 meters, which already equilibrates the air temperature before it reaches the cell temperature.

- Concerning the uncertainty analysis, it is based on the analysis of one case based on a single flight. Although it provides an overall good interpretation of the relative contributions of the different sources of uncertainty for this case, the conclusions are therefore difficult to generalize to any monitoring case, where many parameters might be different such as types of sources (punctual or diffuse, low or high elevations, with/without ejection speed, single or multiple sources), different topographies, presence of buildings/obstacles, varying distances between source and monitoring plane, different wind conditions (mean speed and turbulence). This uncertainty analysis could also be further developed to study the influence of other sources of uncertainties in this case, such as the quality of the wind direction measurements, or of the quality of the H2O measurements.
  - Thank you for your comment, this is of course correct. We tried to implement as many possible sources of uncertainty that are known sources of uncertainty with higher precision sensors to determine how this might influence the sensors uncertainty during UAS deployment. You don't optimise flight for the terrain roughness or similar, but for plume shape and variability thereof. If that's wide or complex (e.g. due to topography or ejection velocity), that plume is in principle still well-quantifiable with our method, provided sufficient repeats or flight duration.
  - It is true that this uncertainty analysis is optimised based on one specific case and each case is slightly different. However, we believe that an uncertainty analysis like this can still give a good approximation of the biggest sources of uncertainty and does give a good uncertainty range for the UAS applications with a single source and curtain flight pattern.

- To put the uncertainty analysis of this single case in a broader context, it would be interesting to compare these uncertainties with the variability of the quantifications obtained between repeated flights on the same source. Are the estimated levels of uncertainties coherent with the real observed variability?
  - This is correct, and indeed a good addition to the manuscript. We did not explicitly state this in the first version of the manuscript, but we do believe this to be a nice bridge between the statistical analysis and real-life scenarios. We specified this in the updated version and implemented as:
  - "This simulation was used to estimate the overall combined uncertainty during our field campaign for a single flight. With the above-mentioned setup, the obtained flux was 10.0 kgCH_4/hr ([3.3–19.1 kgCH_4/hr; 95% CI), with a standard deviation of ± 40% (1σ). During field deployment, the variability of the observed plume was 39% (1σ, n=4). The observed simulation uncertainties and the field example agree well."

- The proposed uncertainty analysis is limited to one case and does not allow a quantification of uncertainty for any single flight. I would also like to see in the discussions if the authors think that the proposed uncertainty analysis method could be applied to any single flight to evaluate the uncertainty of the flux quantification for each individual flight.
    - Thank you for this comment. We added an additional paragraph to the discussion to address this topic. We believe that our uncertainty analysis does give a good representation of sensor uncertainty during UAS deployment. However, applying this uncertainty to any flight is not possible. (There are limitations). The following paragraph was added to the discussion to address this point:
    - "The statistical analysis provides insight into potential sources of uncertainty in UAS-based sampling. The OU-simulation is appropriate for the specific flight trajectory presented in this study. The uncertainties derived using the simulation remain reasonable when real-world flights are conducted, comparable to those simulated. Deviations may arise when observed fluxes are highly variable over a short time (i.e. one hour) or when the flight path (specifically: sampling density) is sparser than the modelled configuration. In those cases, the simulation results may no longer be representative. Flight paths of all flights in this study are comparable to flight 1, therefore OU-derived uncertainties can be applied to all flights."

Specific comments

Wording: I think the gender-neutral policies of EGU suggests using the term "Uncrewed Aircraft Systems (UAS)" instead of "UAV". Also replace "drone" with the appropriate term (either UAV or UAS) throughout the text.

- Revised and replaced all instances of "UAV" with "UAS" where applicable. The reference list still mentions the term "UAV" due to references used

Line 131: "the ground time » > the ground team?

- Revised and replaced "the ground time" with "the ground team". This was indeed a spelling mistake.

Line 200: It seems that the LI-7810 Is used to monitor the dry-air CH4 concentrations, but the Licor is technically also measuring wet air. What correction did you use to calculate the dry-air CH4 concentrations from Licor measurements? Is it a manufacturer correction? Did you validate this function?

- This is indeed something to consider. The analyser has a build-in water vapour dilution correctoin. We validated this correctoin works appropriately (i.e. the instrument reports "bone-dry" mole fractions, irrespective of the actual water vapour content).

Lines 212-215: I agree with the following remarks that the errors introduced by water corrections would be negligible compared to other uncertainties. However, wouldn't it be possible to use the same protocol at different CH4 concentrations to check the stability of this correction?

- This would indeed be a good possibility. We performed the same protocol at a different CH_4 concentration, close to ambient, to determine the effect at lower concentrations. This additional experiment is added to the Appendix (Appendix; Figure C3 and C4) and shows that at lower concentrations the water vapour effect cannot be distinguished from the sensors noise.

Line 250 (Figure 2): It is regrettable that the time series of the in-flight tests are not shown here, as for the laboratory tests.

- Thank you for your comment. We added the time series of the in-flight test to Figure 2, in a similar manner as the laboratory tests. We also updated the color scheme so the time series match the Allan deviation graphs.

Line 255-265: It is not clear whether the incoming air sample temperature was stable or varying as well during the tests. On the field, one does not necessarily expect strong air temperature changes during a flight (at least as long as no warm source such as fires or industrial sources are monitored), but there might also be a strong temperature difference between the cell temperature and the outside air temperature. Would it also affect the measurements or is the cell temperature regulation sufficient to compensate important outside-air to cell-temperature gradients?

- Thank you for this comment. During the temperature dependency experiments, we sampled the air from a cylinder. The temperature of the incomming air is not significantly different compared to the cell temperature and we therefore do not believe this influences the results much. The active temperature control is sufficient to compensate for this difference, since we did not experience any visible effects.

Line 324: "was derived as the average the WindMaster Pro observations" > "the average of the"

- Revised and replaced "was derived as the average the..." to "was derived as the average of the…"

Line 358-370: Information is missing here about the distance between source and observational plane.

- Thank you for pointing this out. We added a sentence stating the average distance between the observational plane and the source.
- "The flights were conducted under light to moderate wind conditions (3.7 ± 1.4 m/s, NW to WNW; Appendix E) and consisted of 12 transects between 4 and 60 m above ground level at approximately 120 m downwind of the farm (Appendix H)."

Line 430: This method might work but is subject to interpretation and is time consuming. Alternative baseline fitting methods exist which could be employed to estimate the background concentrations, even with noisy datasets once the appropriate parameters have been found. They could be applied either to individual transects or the complete time series of observations at once.

- Thank you for this suggestion. We concur that the baseline fitting method possibly is more robust than the method used in this manuscript and less prone to errors. An additional benefit is the applicability of this method for the Axetris and the AirCore. However, due to our limited testing of the baseline method, we opt to retain in this manuscript the originally presented method but will certainly strive to incorporate the baseline method into future work.

Line 460: « completed eliminated » > « completely eliminated »

- Revised and replaced "completed eliminated" to "completely eliminated"

Line 467-468: In this complete section, the authors present the example of the second flight only, but here there is a mix of results from all flights and individual flights. It should be more coherent. Furthermore, the mean difference of all flights is discussed, but the average value does not appear either in the text at this point or in the Table. I suggest moving these remarks to the following paragraphs treating all flights (lines 486-499).

- I am not entirely sure what this comment refers to. The lines mentioned contain information about the average wind conditions of the day and naturally contain information about all the flights. The paragraph above (Lines 461- 466 in the original manuscript contains information about Figure 3 and refers to flight 2. However, the last sentence might be the one you refer to, stating: "The emissions estimates derived by the two methods (AirCore, Axetris) differed by less than 10%, both for individual flights as for the mean of flights." We therefore removed this sentence here and merged it to form the following sentence: "The correlation between the individual flight observations is strong and the emissions estimates derived by the two methods (AirCore, Axetris) differed by less than 10%, both for individual flights as for the mean of all flights (Table 1)."

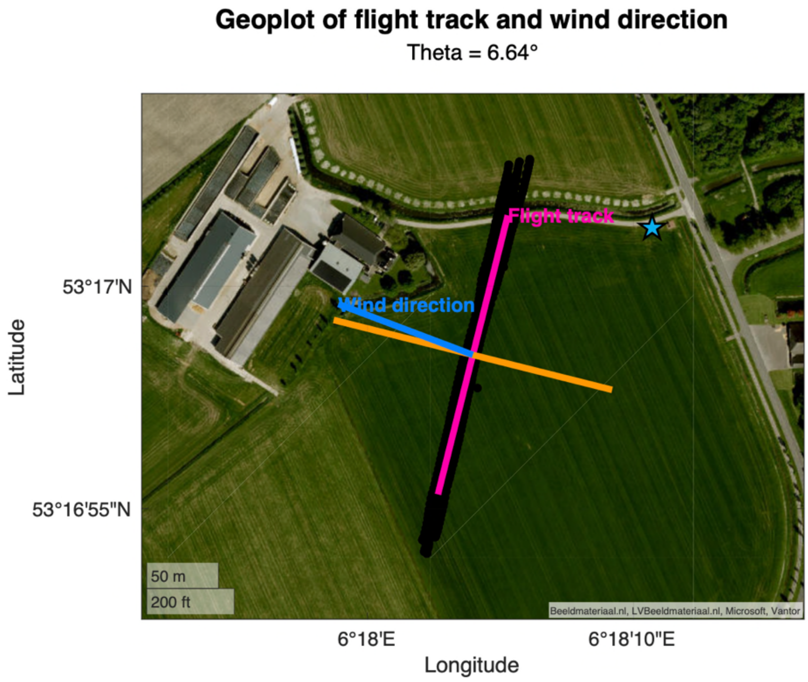
Line 470: The authors are referring to Figure D1 (appendix), where the legend could be more precise in terms of period of observations: what time period is exactly represented (it is the complete day or only the period when the flights occurred)? From observations at which frequency?

- We changed the caption of Figure F1 (Figure D1 in the original version) to include the notion of the Windmaster Pro data. The caption now reads: "*Figure F1: Wind rose of Windmaster pro observations during the campaign day at 29-07-2025*". As stated, the observations are from the complete day, and the Windmaster Pro collects observations at 10 Hz. The additional figures (Figures E2 to E6) are flight-specific.

  They could also refer to Table 1 where the mean wind speed and directions of each flight is presented.
- Table 1 gives flight-specific data, while Appendix E1 shows the observations during the entire day and shows that the wind speed and direction did not vary significantly during the day. Since we are discussing the daily average, we will reference Figure E1 rather than Table 1.

Line 471-472: The information about flight plan should be presented earlier in the "Fight strategy" section, in particular the distance between source and observations which is lacking in this section. You are also referring to Appendix G, where the figure are lacking a colorbar for CH4 concentrations (or you could only present UAS positions instead of CH4 values, which are hardly readable).

- Revised and we moved the information about the flight plan to section 3.5 (flight strategy). This section now mentions the distance between the source and the observations and also the number of transects and height above ground: "The flights were conducted under light to moderate wind conditions (3.7 ± 1.4 m/s, NW to WNW; Appendix E) and consisted of 12 transects between 4 and 60 m above ground level at approximately 120 m downwind of the farm (Appendix H)."
- Thank you for pointing it out. We changed all figures in Appendix H to only show the UAS position (black) instead of the CH_4 concentrations, and we changed the caption to complement this. We agree that the CH_4 concentrations were hard to read and adding them to this top down view does not add any valuable information. All figures now look something like this:



-

Line 615-625: Wind speed uncertainty here only represents the uncertainty of the mean wind speed. However, wind uncertainty can also affect the wind direction, which is not considered here. This would be interesting to add this additional uncertainty into the analysis.

- You are indeed correct and this is certainly something that should be discussed more clearly. We spend some sentence to address this by adding onto the paragraph describing the influence of spatial sampling density to the discussion. We decided to leave it out of the analysis due to difficulty modeling the correct interpretation onto the statistical analysis to represent real life movements of the plume.
- We did add some discussion about this point into the revised manuscript: "Furthermore, wind direction uncertainties are also not included in the statistical analysis. Changes in wind direction influence the dispersion and detection of atmospheric plumes and is often referred to as the meandering effect (Wietzel et al., 2025). As mentioned before, wind variability (e.g. changed in wind speed and wind direction) propagate non-linearly through the MB calculations, therefore affecting the mass emission rate estimates in

varying ways, depending on the error. Future work should combine optimised flight-path design with real-time plume detection to further minimise spatial sampling bias."

Line 631: correct "may stem from to changing"

- Revised and replaced "may stem from to changing" to "may stem from changing"

Line 639: "to accurate resolve" > "to accurately resolve"

- Revised and replaced "to accurate resolve" to "to accurately resolve"

Line 642: "improved spatial resolution the nominal plume" > "improved spatial resolution of the nominal plume"?

- Revised and changed the complete sentence to: "Low-noise sensors enable improved detection of smaller plumes and enhance the spatial resolution of the nominal plume"

Line 676-691: I agree with the interpretation that the uncertainties associated with noise level and background concentrations estimate will probably be the factors that play a role in the dependency of the uncertainty to the emission rate. However, this would have been preferable to perform the same analysis for each individual source of uncertainty, to really justify this assumption.

- Thank you for pointing this out and we understand the critique. However, we tried to address this concern by pointing out that by increasing the flight repeats (e.g. fly more during the same day) the uncertainty at low emission rates benefits more compared to higher emission rates by stating "random errors are effectively averaged out through repeated sampling." Additional meteorological errors play a bigger role when emission rates increase, and noise and background uncertainty have less of an influence (since percentually, it would be difficult to be off more). The uncertainty of each individual source was conducted in section 4.2, but at lower emission rates.
- We did add an additional Figure to the Appendix L showing the effect of each individual source of uncertainty for a source of an emission mass rate of 1 kg/hr and refered to it at the end of the statement:

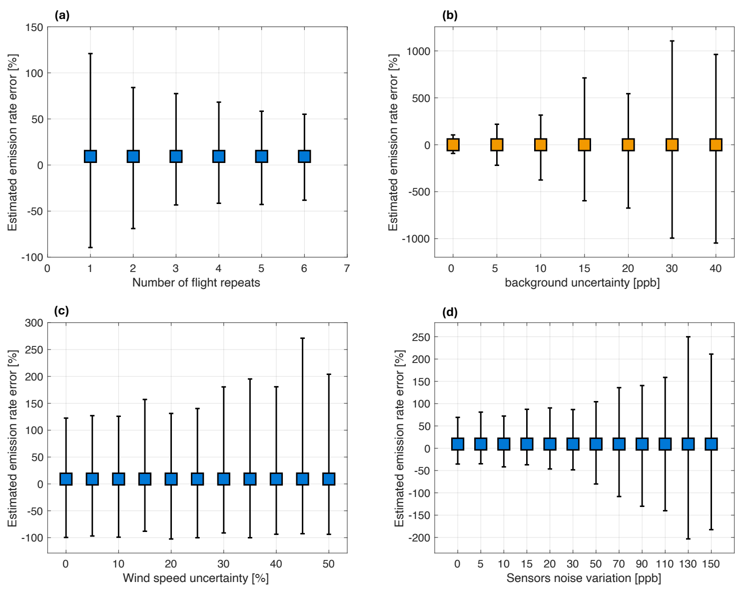


Figure L1: Overview of the individual sources of uncertainty and their respective effect on the estimated error based on an emission mass emission rate of 1kg/hr.

Line 732: define “IGA »

- IGA is already defined earlier in the paper (line 72)

Line 740 : « the use if a method » > “the use of a method”

- Revised and replaced “the use if a method” with “the use of a method”