Reply to Reviewer number 3, Anynymous 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

The paper discusses initial validation tests with the cost-effective medium-precision Axetris sensor compared to a high-precision established sensor. Lab tests by Van Ettinger et al highlighted the importance of temperature control to reduce sensor noise. The Axetris sensor was then deployed on a drone to measure the methane emission rate from a dairy farm; the testing was conducted in parallel with the established AirCore method and the results show potential of the Axetris sensor for use in the mass balance approach. The authors conduct a detailed sensitivity analysis, which highlights the main factors of error in the method.

I think the work is relevant for many researchers and environmental scientists in the field of methane emission measurement. The novelty of the paper lies in the insight that a cost-effective medium-precision sensor is capable of providing accurate results on methane emission quantification. The uncertainty analysis is more extensive than what is usually found in the literature and that's very helpful. The paper is very well written in general, although I think that some more deeper discussion can be provided on a few specific topics (listed below). I also find that a few graphs are unclear at the moment; a better description may be needed there. Overall, I would like to recommend the following minor revisions before publication in AMT.

Specific comments

The title is very generic at the moment, and does not capture some of the main points of the paper. From a practical perspective, I think the key point is that the new Axetris sensor is cost-effective but still can provide promising results; I would therefore add something along these lines of "a cost-effective in situ methane sensor". Furthermore, on the title, I think "point source emissions" is too limiting: first of all, the Mass Balance Approach also applies to diffuse sources within a certain domain. And second, one could argue that a dairy farm used in the experiment produces more diffuse emissions (less of a point source, indeed) than, say, a vent stack in an oil & gas facility. I would therefore recommend to replace "quantify point source emissions" to something like "quantify emissions at facility level" or just "quantify emissions".

- We changed the title to be more specific. It currently reads: Evaluating the performance of a cost-effective in situ methane sensor for UAS-based systems and its ability to quantify facility-scale emissions.

Drone-based methane sensing is a rapidly developing field, and I see that many references to existing measurement campaigns are from several years ago. For some latest references, I would recommend to read and refer to, e.g., the following recent papers and references therein:

- AMT – Controlled release testing of commercially available methane emission measurement technologies at the TADI facility (Audrey McManemin, Catherine Juéry, Vincent Blandin, James L. France, Philippine Burdeau, and Adam R. Brandt, 2026)

- Validation and demonstration of a drone-based method for quantifying fugitive methane emissions - ScienceDirect
(C. Scheutz, J.E. Knudsen, N.T. Vechi, J. Knudsen, 2025)

I will come back to these references below, because some of the conclusions in these papers are in line with what Van Ettinger et al reported so it will be good to make that connection.

- We added some newer references to the paper. Based on the references suggested above. We decided to only include references talking about specific sensor development and no papers talking about a wider application, but we did read through the suggestions above. The paper from Scheutz et al. 2025 also uses the Axetris sensor, but leaves out the sensor characterization and focuses more on a "new" flux quantification technique.

If the key novelty in van Ettinger et al's paper, over other established methods using expensive high-precision sensors used in the references above, is in the Axetris sensor that is low-cost but requires temperature control, then a few considerations need to be discussed in further depth. First, it is unclear now in how far Axetris themselves have already done lab testing before bringing the sensor on the market. What has already been tested, and how does the lab testing by the Ettinger et al build on that? Second, for anyone wanting to deploy the sensor on a drone themselves, it will be important to know how the temperature control is implemented. What is required? Is this available on the market or do people have to develop it themselves? What is the complexity of this solution, i.e. how much would it cost approximately?

- Thank you for your comment. This is indeed something to consider. We have had some contact with Axetris during the writing process, and they helped us better understand the sensor. However, as fas as we know, Axetris themselves have not published a sensor characterisation details pertaining to temperature stability. Most of the information about the sensor is available in their brochure, where they state the detection range etc. They are aware of the need for temperature stability and mention (in the brochure) that with the help of an included temperature sensor, the cell temperature is measured and the concentration observations are consequently compensated. We establish that this "compensation" is not a functional part of the final design, which is why we advice the additional active temperature control.

- To address the second part of this comment, we are planning on making the design of the hardware involved in the Axetris package open source, allowing other laboratories and companies to implement it into their system design. Currently, we are making some improvements to the design to clean it up and add some additional sensors to broaden its use. As soon as this process is complete we will make the final design available. As stated in the description of the Axetris package, the required hardware is absolutely minimal. Component costs are less than €100,= (MOSFET + heating wire + tape). The only requirement is that a microprocessor is available to execute the control loop. In practice, this criterion is commonly met.

Another issue is the wind measurement implementation. I understand from Van Ettinger et al that they did measure the wind at the location of the UAV, but in the end the data was not used for the mass balance calculations. Some discussion is warranted about the implementation that Van Ettinger et al tested. Why did it not give useful results? Is this something that can be improved in future? Let me say that it is not easy to measure wind at the location of the UAV accurately, but e.g. Scheutz et al (2025) do manage it and show the value of the data in the methane emission quantification analyses. Some discussion is needed about how important this may be and what can be done in the future to improve wind measurements at the drone.

- Thank you for your comment. We decided to use the 'trisonica' wind measurements at the UAS location only for illustrative purposes, since we are not confident that our UAV-movement correction method yields sufficiently accurate data. Using the coordinate-specific wind data can be a great improvement if implemented correctly (as can be read in Scheutz et al., 2025). And we are also planning on implementing this for the next publication. However, we still used the uncorrected TriSonica wind measurements to verify our logarithmic profile (as can be seen in Appendix E).

In the field trial, the Axetris sensor is compared to the AirCore method. I understand that this may have been a practical choice due to availability of the equipment, but it does raise the problem that there are two distinct differences have to be evaluated: the difference in performance between the sensors, and the potential difference induced by the use of the data analytics required on the AirCore data (sampling, mapping of concentration data onto the original UAV trajectory, etcetera). In an ideal scenario, I think it would have been preferable to compare the Axetris sensor to another (expensive but established) in situ sensor that gives high-precision and high-frequency data, mounted on the same UAV: then the readings for the concentration measurements can be directly compared to each other, and this will probably lead to further insights on theperformance of the Axetris sensor. Of course I am not suggesting that such an experiment should be included in the current paper, but I think it does deserve to be mentioned as a suggestion for future work.

- Yes you are correct. Unfortunatley we were limited to the equipment that we had access to. In other work done by Morales et al. the Active Aircore was compared to an open path in situ QCLS. We added the following sentence to the discussion to addres this point: “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 flux estimates using comparable data analytics. Furthermore, large-scale controlled release experiments could provide a quantitative assessment of the sensors’ performance.”

Based on the observations above, I feel that some conclusions are a bit overstated, in particular “The close agreement between the techniques validates the use of the Axetris sensor for robust flux quantification” (line 769). I think it is premature to use words like “validates” and “robust” for a method that has been tested on only five flights downwind (of which four flights were successful) of a dairy farm, from which the mass emission rate was not known at source level. I agree with the authors that the Axetris implementation has potential and that the results are worth reporting on in a scientific paper, but further testing will be required in different wind conditions and at different facilities before a robust validation can be claimed. In this light, I think further test campaigns, especially with controlled releases of methane, can be suggested as next steps in the Discussion or the Conclusions, to further evaluate the Axetris implementation – in addition to the side-by-side testing with a high-precision sensor as suggested above.

- We changed the tone of this statement so we do not overstate the conclusion. Right now it reads: “The close agreement between the techniques, despite using different background determination methods tailored to each technique, validates the use of the Axetris sensor for reliable flux quantification.”

- The following sentences were added to the discussion and mentions future study objectives: “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 flux estimates using comparable data analytics. Furthermore, large-scale controlled release experiments could provide a quantitative assessment of the sensors’ performance.”

Technical question/ issues/ suggestions

- Line 66: examples are from 2016 and 2021. Please consider adding more recent papers on UAV deployments. E.g. the two papers mentioned above and references therein.
  - Revised and added some additional references.
- Line 74: I know that some papers in the literature do use the word “flux” as a synonym for “mass emission rate” but from a fluid dynamics perspective the concepts are distinct. A flux is a mass emission rate per unit of area (e.g. in g/(m2s), for a diffuse emission from a large area like a landfill or biological processes). Here and henceforth, I would therefore recommend to use “mass emission rate” for the quantity that has actually been measured from the farm – not “flux”.
  - Revised and changed all the instances of “flux” to “mass emission rate”

- Line 131: what does “ground time” mean?
  - This was a spelling mistake, and we revised it to “ground team”

- Line 175 (equation): the equation does not go through (0,0). Is this a problem, for instance, when extrapolating these results to higher or lower concentrations than observed in the field trial?
  - We understand your concern, but this should not be a problem. The test was conducted over a large range of concentrations (2 – 20 ppm). 2 ppm is a background level one would usually measure during ambient air experiments, since this is close to the mean global concentration. You would therefore not really encounter any concentrations that are significanlty lower as the those considered during this test. Of course, at very high concentrations, the system may not behave linearly, but since we measure emission spikes, these events are short-lived and should not be strongly affected by nonlinearity.

- A related question on the Figure B1 (page 30): why do the outlier points in Figure B1 seem to lie on a hyperbolic tangent-like curve rather than on random distribution around the straight fitting line? Perhaps it’s my limited understanding, but I don’t think there’s an explanation in lines 167-181 for this hyperbolic tangent-like data.
  - This is an artifact of the difference in the response time. The data included the entire time series of the experiment, which starts with sampling ambient air and then rises to the high emission inside the flask. The Axetris simply reaches that plateau sooner. To remove the hyperbolic-tangent-like data (which are not meaningful for the experiment), we decided to remove the observations associated with the rise. This resulted in an updated Figure B1, which can be found in the updated manuscript. Removing the tangent-like behaviour also changed the linear equation, which is updated as a result of the change.

- Line 187: as I mentioned under the “Specific comments” , I am curious to know if Axetris hasn’t performed any such measurement themselves on their sensor before bringing it to market. If they did do these tests, what were the results?
    - To our knowledge, they did not do such tests themselves. We have been in frequent contact with them and brought the extent of the temperature dependency to their attention. Axetris AG mentions being aware of some effect, but were unaware of its severity. They state that the oscillations we describe are negligible for the (i.e., their) intended use case of the sensor (at much higher concentrations than we see)

- Lines 301: “for direct comparison of our results with theirs”. It should be added that there is no evidence that the methane emission rate from the site was the same though. In fact, this point is made later in the paper, but it would be appropriate to mention it here already as well. This is not a controlled release test where the emissions are controlled and known to be the same as what Vinkovic measured.
    - Revised to: “Emissions from the farm were previously quantified by Vinković et al. (2022), allowing us to identify similarities or differences in the emission pattern of the farm.” As you mentioned, it is not possible to do a direct comparison, but comparing the data can still give us insight into the differences between the observation days.

- Line 326 and below: “This logarithmic wind profile is used for calculation of fluxes” . As mentioned under “Specific comments” , some discussion may be needed about the causes of the noise in the TriSonica wind measurements that made that these in situ measurements could not be used in the mass balance approach (compare for instance with the discussion in Scheutz et al (2025)).
    - Thank you for pointing this out. As stated before, we flew with the TriSonica attached, but we did not use the direct observations from the TriSonica because they still required additional corrections that were not ready before we sent this manuscript for publication. We will add an additional sentence, also referring to Scheutz et al. (2025), stating that including point-specific wind data can help improve emission quantification.

- Line 385: Equation uses x and y for coordinates in vertical plane; z is used for the observed value in equation 3. On the next page (Equation 4), the parameter z is used for the vertical coordinate. This needs to be made consistent.
    - Revised and ensured to keep referring to y as a coordinate in the vertical plane.

- Line 405, equation 4: The original mass balance method has cos(theta) inside the summation, because the method allows for changing wind directions to be included (see e.g. Mohammadloo (2025) for a derivation, which shows the dot product with the wind velocity and the vertical curtain’s orientation). Taking cos(theta) outside of the summation is an additional assumption that needs to be noted down.
    - Thank you for this comment. You are indeed correct. In the original method (as explained in Mohammadloo et al., 2025), variations in wind direction along the flight track are allowed, which is why the term appears inside the summation. In our case, we assume a constant wind direction and use a single value for the entire flight. This results in the cosine term being taken outside of the summation. We clarified this assumption in Section 3: “In this work, a

constant wind direction was assumed during flight, allowing the corresponding term to be placed outside of the summation."

- Line 430 "on visual inspection": this approach looks subjective, and a possible cause for further uncertainty in the results. I didn't see a clear discussion in Chapter 4 on this: can the visual inspection approach lead to errors or uncertainties not already accounted for in the sensitivity analysis?
  - Of course, this method can introduce additional error, especially when plume boundaries are not reached. However, each background determination method has its flaws, since the exact background is never known and can also change on both sides of the plume. Even using the 10th-percentile method can introduce an offset relative to the "true" background. In a way, this uncertainty is included in section 4.2.2, where we try to quantify the added error for a false representation of the background. As mentioned there, we assume we can determine the background with a 10 ppb uncertainty.

- Line 475: "Performing multiple passes across different altitudes, these plume-displacements can be captured and accounted for in the final flux estimate." This observation is in line with Scheutz et al (2025), see e.g. their Fig 7; a reference can be made to that paper.
  - Thank you for pointing this out, we added Scheutz et al . (2025) to this statement to improve the strenght of this argument.

- Figure 3 top plot: How was the LOWESS smoothing done? In Figure 3's top plot, it seems that the original Axetris data has much higher peaks than the smoothed data, but the troughs are similar. Was the LOWESS smoothing really mass-conserving? Perhaps an equation can be provided to show this.
  - The LOWESS smoothing was done within MATLAB. You are correct, and the function is not inherently mass-conserving. The difference in the determined integral of the smoothed Axetris data and the raw Axetris data is 103 ppb, which constitutes a percentual difference of 0.001% over the complete time series.

- Figure 3 caption: "The middle of the barn is at 0m horizontal distance". Many people will intuitively understand what is meant here, but the wording a bit sloppy and can be improved. The barn is not actually located in the plane flown by the drone, of which we are seeing the integrated contour plots. It may also be nice to mention the drone's downwind distance to the farm here.
  - Thank you for pointing this out, we updated the caption to make it a bit more clear and also mention the downwind distance. It now reads: "*The horizontal distance is referenced to the projected location of the barn centre (defined as 0 m). The sampled plane is located 120 m downwind of the farm.*"

- Line 577: "Increasing the flight duration lowers the CI bounds and the standard deviation". This finding seems to be again in line with Scheutz et al (2025).
  - Thank you for pointing this out, we added the reference to this statement.

- Lines 627-632: The current explanation for the difference with the observations by Mohammadloo et al is not clear. In particular, does "changing atmospheric conditions" relate to a spatial difference between the ground-based anemometer and the wind speed at the location of the drone? Otherwise, could it perhaps be related to how Ettinger et al have used a mean wind direction in the MBA rather than the actual time-varying wind direction (see my earlier comment on line 405)?

  - This contradiction pertains solely to wind speed uncertainty (and not wind direction). Mohammadloo et al separated the effects of wind speed and wind direction (before combining them). Our results show a clear increase in wind speed variability, whereas Mohammadloo et al. stated that this should not be a problem since concentrations scale directly with increasing wind speed. This statement is correct, but it fails to account for the use of an incorrect wind speed in ones flux calculations.
  - Say you think the wind speed is 3 m/sbased, since on your on-site wind measurements, while in reality it is 4 m/s. This would mean you would multiply the measured mole fractions by a wind speed that is too low, resulting in an underestimation of the true flux. We believe this was not properly considered in Mohammadloo et al., resulting in the difference.
  - The final statement was poorly worded and updated to: “However, this overlooks the impact of an erroneous wind speed on emission flux determinations. If the wind speed is underestimated (e.g. 3m/s instead of the true wind speed 4 m/s), the resulting flux will also be underestimated, since the flux scales proportionally with the wind speed.”

- Line 747-751: I completely agree with this paragraph, but the authors could perhaps add that the effectiveness of their current solution can also be improved in future by simply removing the heavy Aircore equipment from the drone. This will enhance the flight endurance of the system and would allow for more repeated measurements.
  - This is a fair point and is certainly something to consider. We therefore added the following sentence: “Besides, the sensor's effectiveness can be further improved when used as a stand-alone, since flight duration and flight repeats increase due to a lower payload and less post-flight processing.”

- Line 756: “We improved the sensor's performance by properly insulating the sensor and applying active thermal regulation to maintain a constant cell temperature” . As noted under the Specific comments, some discussion is needed on how difficult/costly it is to implement such an enhancement for temperature control.
  - Thank you for this comment. As mentioned, we plan to make the design open source to address this. However, the implementation cost is very low (< 100 €).

- Line 1003: this is presumably flight 2, not 1.
  - You are correct. Revised and changed it from flight 1 to flight 2.