RC1

General comments

The publication "A helicopter-based mass balance approach for quantifying methane emissions from industrial activities, applied for coal mine ventilation shafts in Poland" by Förster et al. addresses the important aspect of comparing reported coal mine emission rates with those estimated using a top-down approach involving airborne based in situ measurements. To do this, the authors use the HELiPOD platform, which enables (vertically) dense flight patterns also at quite low flight altitudes and close to the source - a feat that would probably be impossible for larger aircraft platforms equipped with in situ instrumentation. The CH_4 measurements, combined with locally acquired meteorological data such as wind information, are subsequently converted to emission rates using the widely adopted mass balance approach. Any usual difficulties encountered during this type of measurement, such as missing ground observations or atmospheric turbulence, are resolved by using mobile ground-based measurements, if available, and/or acquiring multiple densely sampled downwind curtains. Additionally, the method was tested in a controlled release experiment. The measurements were performed in the Upper Silesian Coal Basin (USCB) in Poland - the CH4 hot spot in Europe – which also has been the focus of many other groups in the past, as detailed in the manuscript. Förster et al. found very good agreement between the estimated top-down emission rates and the reported ones, which are based on safety sensors, for the investigated shafts. They were also able to distinguish emissions from ventilation shafts and a nearby drainage station, and found that drainage stations are significant sources of CH_4 emissions in the context of coal mining. The manuscript fits well in the scope of AMT, and I recommend publication following some minor modifications along the line of the comments below.

Overall, the manuscript is well written, the methods are sound, and the results are clearly presented. Especially, the Supplement contains a wealth of detailed information for interested readers. My main comment relates to the introduction, where the authors conflate aircraft and satellite, and in situ and remote sensing measurements. I address this issue again in the conclusion – see the specific comments below.

I would also encourage the authors to add a short section discussing their findings, i.e. the good agreement of top-down and reported shaft emissions, in context with previous studies in the USCB (e.g., has this been the first time that the reported emissions were confirmed by top-down estimates or was it similar in the past?). There is quite a list of studies listed on P2-3, L81-91.

Answer: First of all, we thank Reviewer 1 for the detailed comments which helped to further improve the manuscript. We tried to implement all comments and recommendation. We made the differentiation between aircraft and satellite, and in situ and remote sensing measurements clearer and added a section where we discuss our findings in context with previous studies in the USCB. Further answers to the specific comments are listed below.

Specific comments

P3, L87: Krautwurst et al., 2021, analysed airborne remote sensing measurements, rather than spaceborne measurements.

Answer: Changed.

P3, L92-105: This paragraph is slightly confusing in terms of the references used:

- Swolkien et al., 2022: In their study, they use in-shaft measurements from safety sensors. Consequently, it does not align with the other references provided here, nor with the statement "… performed at distances of several km downwind of the source." (P3L97)

Answer: Corrected and rephrased.

- Chulakadabba et al., 2023: In their manuscript, the main focus is the airborne remote sensing instrument 'MethanAir', which is not a satellite instrument.

Answer: Corrected.

I would suggest distinguishing between satellite and airborne instruments, and extending the discussion in this paragraph to include specific information on airborne remote sensing instruments, given that some of these are already mentioned in the manuscript (Chulakadabba et al., 2023 and Krautwurst et al., 2021), along with some discussion. Depending on the emission strength and the prevailing wind speed, these instruments are also sensitive to the near and far field (or, as described in the manuscript, the "gap between a distance of ~100 m to 3000 m") of an emission plume. Other further prominent examples include the passive imaging remote sensing instrument AVIRISng (e.g., Duren 2019, which suggest detection limits lower than 10 kg/h), airborne thermal imagers (e.g., Tratt et al. 2014; Hulley et al., 2016), the passive imaging remote sensing instrument MAMAP2D-Light, or the active remote sensing instrument CHARM-F (Krautwurst et al., 2024). These instruments are, in principle, also capable of measuring coal mining emissions in the near and far field. Another prominent satellite instrument is EMIT aboard the ISS (Thorpe et al., 2023; Li et al., 2025); which observes CH4 plumes from the source up to several kilometres upwind (e.g., Thorpe et al.,2023, their Fig 2B), given high emission rates and favorable conditions.

Answer: Thank you for the comment. We added a paragraph where we discuss different measurements techniques for CH_4, including remote sensing, and made the differences with advantages and disadvantages between the techniques clearer.

P3, L111f: Does this mean that helicopter operations do not require a flight permit for the measurement area at all?

Answer: No, it is just much easier with local companies. Sentence was rephrased to make it clearer.

P3, L113f: This statement applies well to in situ measurements by small aircraft, but how does it compare to the mentioned spatially high-resolution satellite or airborne imaging instruments?

Answer:

Satellite: Yes, with modern satellites this might be possible. At least to monitor CH_4 concentrations precisely. But they still need accurate wind data. Momentarily wind is inferred from ERA5 or cloud movements which might be a source of uncertainty.

However, the HELiPOD is a perfect tool to validate the new satellite products like GHGSat, MethaneSat or CarbonMapper. First comparisons to GHGSat are promising.

Ground imaging instruments may integrate different sources from a side view, e.g. Knapp et al., (2023)

Airborne imaging instruments: They are generally also able to quantify single sources. But they might have higher uncertainties due to the assumptions of the retrieval method and wind, e.g. Borchardt et al. (2025) as well as limitations in areas with complex albedo or water surfaces (Cusworth et al., 2019; Ayasse et al., 2022).

P6, Table 1: Are there also estimates of the accuracy of the parameters (partly, they are listed in the Supplement)? I assume that accuracy is less important for the methane concentration measurement, given that both instruments were regularly calibrated and for the mass balance approach, it is the methane enhancement above background that is important, rather than the absolute value. But what about the wind measurements?

Answer: We added in the manuscript: "The accuracy of the wind vector of HELiPOD is around 0.1 m/s for horizontal wind speed and 3° for the wind direction."

P10, L318f: Would it be useful to check the consistency by seeing if MBE4, which is currently not used for emission rate estimates due to the mixing of CH4 signals, gives the sum of the emission rates of the individual sources in that area?

Answer: Yes, the total emission rate of MBE 4 is within the uncertainty range of the sum of the individual sources. Sentence added.

P10, L326f: I would think that showing the upstream curtains would greatly enhance confidence in the measurements, as they demonstrate the transportation of clean air masses into the measurement area, as illustrated in Figs. 3 and 4.

Answer: Upwind curtain was added to Fig. 5.

P11, Eq. 4: How does the wind error of the HELiPOD compare to that of other airborne instruments, and are there advantages in the lower flight speeds of helicopters compared to 'normal' research aircraft (e.g., the DLR Cessna Grand Caravan 208B; Fiehn et al., 2020) in terms of the wind speed error? Supplement (P12, L227 and 228): How are the errors for DD and FF computed or where is this described?

Answer:

See Pätzold et al. (2023), P. 7: For calculating the three-dimensional wind vector at 100 Hz, different sensors are involved: A multi-hole probe (Rosemount, USA) was used with pressure sensors to determine the flow field and combined with high resolution attitude data. The differential pressure between total and static pressure (the dynamic pressure) is required for calculating the true airspeed, and the differential pressure between the up- and downward-oriented hole as well as between the left- and right-hand-orientated hole is used for determining the angle of attack, respectively, the angle of sideslip. The technology of determining the wind vector and methods of calibration are equal to the state-of-the-art research aircraft; thus, the uncertainty of the wind measurement is comparable (see, e.g., Lampert et al., 2020a, https://doi.org/10.5194/amt-13-1937-2020). Compared to the wind measurement of the DLR Cessna Grand Caravan 208B (Fiehn et al., 2020), thoroughly described in Mallaun et al. (https://doi.org/10.5194/amt-8-3177-2015), for the HELiPOD the approach of calibration maneuvers is applied, utilizing wind calibration pattern (wind square) flown roughly at each second flight of the campaign.

P16, L514f: What do the authors consider to be low wind speeds here, i.e. smaller than 2 m/s? According to, e.g., Sharan et al., 1996, the transport through a cross-section must be dominated by advection and not by diffusion. However, for wind speeds lower than ~2 m/s, diffusion becomes more prominent.

Answer: Thank you for this comment! Yes, we consider low wind speeds to be below 2 m/s. We added the value.

P19, L599: Given a true release rate of 21 kg/h, doesn't the range of release rate estimates from 14 to 22 kg/h point to a potential underestimation, at least for low emission rates? A follow-up question: Looking at Fig. 7 or Tab. 3 (last column), it seems that there is a tendency for underestimation also at higher emission rates, although well within the error bars. Could the authors please comment on this?

Answer: In general, underestimation could be due to "incomplete" sampling of the plume at the edges with its low methane enhancements with disperse into the background and are not detected by the instrument. However, with a precision of 1 ppb of the Picarro, the undersampling should be only some kg/hr. For lower emission rates, this effect might therefore be more present. For the Licor data, with its precision of 20 ppb, this effect is visible with overall slightly lower mass flux estimates compared to Picarro (Sect. S7).

P19, L607f: Does this statement also apply to research aircraft equipped with remote sensing sensors? Please clarify or be more specific and state, that this statement applies to an aircraft with an in situ payload.

Answer: We adapted the statement to be more general and refer to remote sensing and other techniques at the end of the conclusion.

P20, L624f: What is the basis for the 60-minute estimate, or is it an educated guess? I would imagine that this depends heavily on the day of measurement and the prevailing meteorological conditions, which fluctuate from day to day.

Answer: Yes, it is an educated guess, mainly based on the experience gained from this campaign. During a follow up campaign on the Arabian Peninsula, we saw also faster changing conditions (especially in the morning) when the convection starts and the PBL/inversion layers start to move to higher altitudes. But still the statement of tight temporally probing is valid. The 60 minutes apply to this campaign. I rephrased and specified.

P20, L649: Can it really be generalised that bottom-up estimates for the listed sectors are poor? Please provide references to support this claim.

Answer: Maybe not poor, but often they underestimate emissions and top-down approaches are necessary to validate them. Sentence rephrased and references added.

P20, L658-660: I cannot follow this conclusion here. I think the author should clarify whether they are referring to remote sensing and/or in situ measurements. Satellite measurements (with high spatial resolution) and airborne imaging remote sensing measurements can characterise and separate emission sources in the near field effectively. The separation of emission sources providing emission rates above the instrument's detection limit is a strength of imaging instruments. This comment relates to my comment above (P3, L92-103).

Answer: We adapted this paragraph. We tried to contextualize our method compared to remote sensing instruments and satellites.

Supplement

Fig. S2: Could the authors please refer to Fig. S3 for clarification on what is meant by “MBE2 (H6 to C6)”?

Answer: Done.

Fig. S4: Regarding the plume on the right (black dots) from the drainage station, a distinct signal appears until transect TS6. Above that, the signal disperses and may not be related to it. Till what altitude do the authors integrate for the plume from the drainage station?

Answer: If a plume does not reach to the highest flown transect, we estimate the plume top to be halfway up to the mean altitude of the next higher transect. We added this sentence also to Step 4, since it was not explained yet.

In case of the drainage station between TS 6 and 7 at around 190 m.

P5, L104ff: Could the high variability in wind direction on the ground also be related to the way in which wind direction was measured during the transect? Another possibility is that obstacles such as buildings or trees along the transect could distract or deflect the wind.

Answer: Yes, could be also the way of measurement. Since the ultra-sonic anemometer was mounted on top of the car and the wind was blowing perpendicular to the driving direction, small eddies could have influenced the measurement.

During the release experiment there were no obstacles (since it was at the airfield), but when driving around coal mine ventilation shafts, trees or buildings could also distract the wind.

Technical corrections

P3, L91: Before re-submitting, please check all publications currently labelled as 'in preparation'.

Answer: Done

P5, L162: Please add the date on which the webpage was last accessed.

Answer: Added.

Fig. 2: Please add labels (a), (b), and (c) for the three different panels

Answer: Done.

Ayasse, A. K., Thorpe, A. K., Cusworth, D. H., Kort, E. A., Negron, A. G., Heckler, J., Asner, G., and Duren, R. M.: Methane remote sensing and emission quantification of offshore shallow water oil and gas platforms in the Gulf of Mexico, Environ. Res. Lett., 17, 084039, https://doi.org/10.1088/1748-9326/ac8566, 2022.

Borchardt, J., Harris, S. J., Hacker, J. M., Lunt, M., Krautwurst, S., Bai, M., Bösch, H., Bovensmann, H., Burrows, J. P., Chakravarty, S., Field, R. A., Gerilowski, K., Huhs, O., Junkermann, W., Kelly, B. F. J., Kumm, M., Lieff, W., McGrath, A., Murphy, A., Schindewolf, J., and Thoböll, J.: Insights into Elevated Methane Emissions from an Australian Open-Cut Coal Mine Using Two Independent Airborne Techniques, Environ. Sci. Technol. Lett., 12, 397–404, https://doi.org/10.1021/acs.estlett.4c01063, 2025.

Cusworth, D. H., Jacob, D. J., Varon, D. J., Chan Miller, C., Liu, X., Chance, K., Thorpe, A. K., Duren, R. M., Miller, C. E., Thompson, D. R., Frankenberg, C., Guanter, L., and Randles, C. A.: Potential of next-generation imaging spectrometers to detect and quantify methane point sources from space, Atmospheric Measurement Techniques, 12, 5655–5668, https://doi.org/10.5194/amt-12-5655-2019, 2019.

Fiehn, A., Kostinek, J., Eckl, M., Klausner, T., Gałkowski, M., Chen, J., Gerbig, C., Röckmann, T., Maazallahi, H., Schmidt, M., Korbeń, P., Neçki, J., Jagoda, P., Wildmann, N., Mallaun, C., Bun, R., Nickl, A.-L., Jöckel, P., Fix, A., and Roiger, A.: Estimating CH_4, CO_2 and CO emissions from coal mining and industrial activities in the Upper Silesian Coal Basin using an aircraft-based mass balance approach, Atmospheric Chemistry and Physics, 20, 12675–12695, https://doi.org/10.5194/acp-20-12675-2020, 2020.

Knapp, M., Scheidweiler, L., Külheim, F., Kleinschek, R., Necki, J., Jagoda, P., and Butz, A.: Spectrometric imaging of sub-hourly methane emission dynamics from coal mine ventilation, Environ. Res. Lett., 18, 044030, https://doi.org/10.1088/1748-9326/acc346, 2023.

Pätzold, F., Bretschneider, L., Nowak, S., Brandt, B., Schlerf, A., Asmussen, M. O., Bollmann, S., Bärfuss, K., Harm-Altstädter, B., Hecker, P., Wehner, B., van der Wall, B. G., Sachs, T., Huntrieser, H., Roiger, A., and Lampert, A.: HELiPOD—Revolution and evolution of a helicopter-borne measurement system for multidisciplinary research in demanding environments, Elementa: Science of the Anthropocene, 11, 00031, https://doi.org/10.1525/elementa.2023.00031, 2023.

RC2

This is a very well written, comprehensive research paper. It uses a novel approach to compare insitu airborne measurements to reported emissions. It covers the actual flight measurements at coal mines and a controlled release experiment. The paper is a good fit for the journal.

I recommend publication with some minor modifications.

Answer: We thank Reviewer 2 for the comments! We tried to implement all of the recommended modifications listed below.

L64 coal sector approx ⅓ of global anthropogenic CH_4 emissions - can a reference be applied

Answer: Reference added and sentence corrected. The 1/3 refer to the total CH_4 emissions of the four countries in respect to the global anthropogenic CH_4 emissions. The statement that the coal sector has emissions of ~ 1/3 was not intended here.

L93-103 Can this be expanded - some aircraft can measure in the window stated - bit more clarity would be useful

Answer: Sentence added with reference to the disadvantages of aircraft listed below (L111-121).

L111 - Do helicopters not need flight permissions? Is this local to the region or global?

Answer: Yes, they do, but it is much easier to get them with a local company. Sentence was rephrased to make it clearer.

L115 - Does window of measurements always ensure top of plume is reached, especially when PBL high?

Answer: Yes, we fly vertical profiles before and after every mass balance experiment and found the PBL not to be higher than 3000 m.

Table 1 - could something similar be added for ground measurements?

Answer: We added the information in Sect. 2.6 where the supportive ground-based measurements are described.

L318 - could this be expanded on - why then not the case in other measurements

Answer: Done.

L324 - The upwind curtain is mentioned - what was the outcome - where there additional sources?

Answer: Sentence was added.

L360 - how was PBL defined?

Answer: This is addressed in the second to last paragraph in Sect. 2.3.

L447 - can you clarify what the multiple sources consist of

Answer: Mainly (other) ventilations shafts and drainage stations due to their strong emission rates and mostly low wind speeds (below 2 m s-1 and strongly varying wind direction). I rephrased the sentence and added shafts and drainage stations as main sources.

L481 - why would emissions vary? Is this what was expected?

Answer: Yes, this is what is expected. Emissions can vary with the excavation on the coal seam. If there is no or less excavation, e.g. on holidays, there might be lower emissions. However, during both campaigns the measured rate was relatively constant, also in the measurements of the in-mine sensors, showing constant excavation processes. Sentence was updated accordingly.

L515 - can you clarify what you define “low wind speed”

Answer: We consider low wind speeds to be below 2 m/s. We added the value.

L608 - can see the positives for the HELIPOD and these are clear but are there any limitations - such as payload and where they can operate?

Answer: Yes, the payload might be a limitation, if more (additional) trace species should be measured. For this campaign the payload was about 135 kg. A limitation here is also the flight time of around 2.5 hrs for the Eurocopter AS350.

The HELiPOD can operate, where and when Helicopters are allowed to fly. Obviously during the day and in conditions without much clouds (VFR rules). The HELiPOD can even be operated from a ship as during the MOSAIC campaign in the artic or from offshore oil platforms.

Furthermore, the operation of sling loads is limited over densely populated areas. During take-off and landing the measurements are disturbed by the helicopter downwash. From the operational perspective helicopters are more flexible in adapting the trajectory at the scale of 10^0 km then fixed wing research aircraft, but with less line kilometers covered in the same flight time, making the helicopter-based approach more suitable for point source investigations.

We added these points at the end of the introduction.

RC3

General Comments

This paper was a pleasure to read - it is virtually free of errors and typos and clear about the motivation, setting, sampling, analysis methods, results, discussion and conclusions. It speaks for itself, but to summarize:

The authors present the results of a field deployment of a helicopter payload (HELiPOD) designed to measure atmospheric methane concentrations, meteorological variables and flight telemetry for the purpose of calculating methane emissions from surface sources in a coal mining setting. The well known mass balance approach is used to calculate top-down methane fluxes (kg/hr) over ventilation shafts and drainage stations. The sampling strategy, mass flux uncertainty, and sensitivity (to wind, sampling choices, surface measurements) are clearly described. The mass fluxes compare well with bottom-up calculations based on simultaneously available in-mine methane emission data, given the uncertainties in both approaches. The technique also appears sensitive to emission rates as low as 20 kg/hr, encountered in a controlled release experiment, albeit with a large range of uncertainty (13-70%, L599). Overall, the conclusions are well supported by the results and discussion presented. The one reservation I have is about the authors making the *unqualified* claim in the abstract about unambiguous detection of these small emission rates (20 kg/hr) without any mention (in the abstract) of the large uncertainties, which are 'on the order of the released amount' (L600).

Overall, the manuscript is entirely within the scope of AMT, with a wise division of materials between the main paper body and supplementary materials. The HELiPOD tool represents a substantial contribution to experimental field quantification of methane GHG emissions (big and small) from a variety of emitting sectors. The scientific quality of the approach, method characterization and results is high. The presentation quality is outstanding.

Specific Comments

Yes. Does the paper address relevant scientific questions within the scope of AMT?

Yes. Does the paper present novel concepts, ideas, tools, or data?

Yes. Are substantial conclusions reached?

Yes. Are the scientific methods and assumptions valid and clearly outlined?

Yes. Are the results sufficient to support the interpretations and conclusions?

Yes. Is the description of experiments and calculations sufficiently complete and precise to allow their reproduction by fellow scientists (traceability of results)?

Yes. Do the authors give proper credit to related work and clearly indicate their own new/original contribution?

Yes. Does the title clearly reflect the contents of the paper?

Yes. Does the abstract provide a concise and complete summary?

Yes. Is the overall presentation well structured and clear?

Yes. Is the language fluent and precise?

Yes. Are mathematical formulae, symbols, abbreviations, and units correctly defined and used?

No. Should any parts of the paper (text, formulae, figures, tables) be clarified, reduced, combined, or eliminated?

Yes. Are the number and quality of references appropriate?

Yes. Is the amount and quality of supplementary material appropriate?

Answer: We thank Reviewer 3 for the assessment and the positive comments.

Regarding the “one reservation”: We did not use such a strong formulation as “unambiguous” but stated that the HELiPOD can DETECT (not quantify) emission rates as low as 20 kg/hr. We think this statement is valid also without giving the uncertainty in the abstract. Regarding accurate quantification, there is a higher uncertainty for one of the three methods described in the supplement, yes.

Since this was the first ever conducted release experiment with the HELiPOD, where the aim was to try to detect small emission rates (as written in the text), we would keep this statement in the abstract as it is.

Technical Corrections

L607: "offers an overall" --> "offers overall"

Answer: Changed

There were perhaps 2-3 more grammatical errors of this nature in the paper, but I lost them.

Answer: Thank you for pointing that out. We will try to find them in the proofreading.