Comment on acp-2021-1061
Anonymous Referee #1

Referee comment on "Local to regional methane emissions from the Upper Silesia Coal Basin (USCB) quantified using UAV-based atmospheric measurements" by Truls Andersen et al., Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2021-1061-RC1, 2022

General Comments:

In this paper, Andersen et al. use an AirCore UAV system to quantify CH4 and CO2 emissions from coal vent shafts in Poland using two different methods. These quantified results are compared with directly measured (in stack) emissions, both hourly and aggregated by vent/day. Andersen et al. then use various techniques to upscale the quantified emissions in order to obtain regional emission estimates.

Overall, this work has important implications in the field of methane quantification from coal vents. However, there are non-negligible gaps in the science presented here. Most notably, the authors do not adequately explain the effect of and error introduced by flights when the maximum detected concentrations are on the edge of the “curtain”, indicating that the peak plume concentrations may not have been sampled. Furthermore, the manuscript lacks an important figure which directly compares each flight-based quantification with direct (in stack) measurements. The manuscript can also be improved with a clearer, more coherent argument for the potential impact the manuscript will have on the state of this science. With these changes, this manuscript will be a valuable addition to the literature surrounding methane quantification, AirCore viability, and upscaling procedures. I was glad to get to read this manuscript and provide (hopefully) helpful feedback.

Specific Comments:

I believe there is a lack of explanation in the methodology section regarding the quantification procedures. For the inverse gaussian approach, what point(s) are plugged into the equation? Are multiple points used and compared/averaged? Is the maximum concentration used and assumed to be the center of the plume? How are the dispersion parameters determined (what method), and how do they affect results? Are concentration
peaks dampened by the AirCore method due to mixing in the sampling tube before analysis, and how does this effect quantification?

A critical issue is how you address those flights where the maximum concentration is at the edge of the curtain. How are these flights interpreted? It is hinted at in section 3.2, but I'm confused as to how you are calculating either the IG or MB if the majority of the plume is outside the curtain. This may be clarified by some of the questions in the above paragraph. It would also be nice to see some type of error analysis for each quantification method; that is, how do things like wind variability, peak dampening, dispersion parameters, etc. introduce error and how is this error quantified.

In the same vain, I think there is some issue with how error is represented in the aggregate data. For example, in the aggregation of quantified flux from Pniowek IV (Figure 5c) you claim an error of ±0.2kt/y due to the standard deviation of averaged points. However, in the individual day data for this vent (Figure 6c), the inherent error in each measurement is on the order of 3kt/y. A more robust error propagation analysis would make the aggregate numbers more defensible.

A plot I'd really like to see is the hourly emissions compared with the flight quantified emissions (basically combining Figures 6 and 7). There is a bit of a roundabout comparison in the "hourly inventory" vs UAV quantified analysis (Figure 8b), but the critical representation is missing. The direct comparison is a key figure as it validates your UAV quantification approaches with real, empirical vent emissions data. As you state, emissions vary both intra and interday, so comparing UAV measurements at specific times with the directly quantified vent emissions instead of relying on aggregate data (like that presented in Figure 9) is an important distinction.

In section 2.1, you describe the del13CH4 data collection, but I am confused as how this is conducted. Are you capturing the outlet air of the Picarro upon measuring CH4/CO2 from the AirCore in a bag then analyzing? Some clarification would be helpful.

In my opinion, the argument of "weekends/holidays" does not add any value. If anything, it is confusing, as you postulate a reduction of emissions on these days then go on to show otherwise.

Finally, I think there could be some added discussion about lessons learned and recommendations for future use of AirCore technology to quantify vent emissions. Specifically, expanding beyond why the hourly emissions data and individual flight quantifications may not align well and describing how the methods may be improved would be helpful. Similarly, some discussion of the recommendations for best practices to achieve a certain level of accuracy for quantifying vent/regional emissions using AirCore flights would be helpful; such as, how many flights are needed over how many days...etc.
Technical Corrections:

16: Insert (CH4) after methane.

23: Delete “have”

28: Delete “though”

28-29: Rephrase “As an alternative...” sentence. Make sure verb tenses match and phrasing is clear.

34: Is methane the second “most abundant” or just second most important in terms of climate forcing?

49: Citation for coal being 12% of methane emissions?

52: Change “part of” to “some”

54: Change “releases” to “is released”

56: Insert comma between “mines” and “the”

58: Citation for data loggers lacking accuracy and temporal resolution? It seems that your data shows otherwise... high resolution and temporally resolved fluxes from vents.

64: Sources for other studies using UAVs for methane monitoring?

71: Perhaps add a line describing the Merlin mission and how CoMet ties in?

78: Change “strong ties to hard coal mining” to “containing extensive hard coal mining”
or similar.

83: Period after PRTR

83: Remove “the” after “quantify” and before “emitted”

Paragraphs 70-100: Ensure consistent verb tense. Example: 71 – “goal of CoMet is to provide”, 76 – “CoMet campaign was to quantify” etc.

86/89: It goes from 59 flights to 34 quantifications – consider adding a line about filtering and what section you discuss this, otherwise it is confusing why these numbers don’t match.

89: The quantified emissions are of the shafts using the aircore, not quantified emissions of aircore flights.

95-100: Consider removing the “Section 2 presents …, section 3 contains…” and instead replacing with a strong statement about what your results convey and why they are important.

123-124: The names of the vent shafts have not yet been introduced and I did not know what these names meant. Consider revising to introduce the region and vent shafts before this section (maybe move section 2.3 to beginning of methodology).

126: “First few”: specify how many.

126: What meteorological parameters were collected?

129: Add “Meteorology for flights #5 through …”

Section 2.2: Add details about the height of the meteorological sensors.
135-137: “The CSAT3 has an operating temperature ... small changes in wind direction” is unnecessary.

144: Give some highlights about what the sampling criteria were to consider a “good flight”

144: The intro said 34 flights were used for quantification, this line says 36 fulfilled the criteria – why the discrepancy?

146: Add “technique” between “this” and “effectively”

153-154: Add the altitude range for the flight to go with duration and downwind distances.

Figure 1: Is there any reason for the different colors for each vent shaft? If so label.

179: How do you account for plume rise? In the gaussian equation, I believe $h$ is typically the “effective stack height” which accounts for advective or buoyancy rise effects of the plume.

Section 2.4: How is the local/regional background accounted for?

196: Add “estimate” after “annual emission”. Also, a source citing the E-PRTR inventory would be helpful.

202: Add comma after “active shafts”

210-211: How do you account for the fact that the operating range of the sensors is <100% RH, but the conditions are often over 100%?

215: Should “concentrations” be changed to “fluxes”?

243-244: The sentence “All the isotopic....” Does not make sense.
Figure 6: I’m confused by the color differences – did different flights use different approaches (MB or IG)? I thought each flight was analyzed in both ways? If not be more clear in section 2 about this. Label what the error bars represent. Consider making the x axis on (b), (c), and (d) so that there isn’t so much white space (restrict to sampling time period). Put in caption what the “N:7-5” means. Overall, I think there may be a better way to represent this data, consider reframing.

298-299: “The Borynia VI inventory ‘may therefore not represent...’” I’d think it clearly does not, given the intra and inter day variability in your other data.

Figure 8: Many of the labels are obscured, overlapping, or otherwise can’t be read.

356: “Best statistics” – do you just mean the most flights? If so say that, if not clarify what “best statistics” means.

363: Wording is confusing

366: Again, is “lowest statistic” just fewest flights?

370: “All over”? Confusing

421, 424, others: Replace “linear curve” with “line”

421: Comma between “rate” and “calculated”

450-456: Instead of “comparing” to estimates then talking about how the estimates don’t include coal, perhaps introduce this idea earlier. In reading, it is confusing why the numbers are so different until I realized that the EPTR estimate really doesn't represent coal emissions at all.

462: Add “method” after “upscaleing”
466/471: These lines contradict one another. “does not accurately represent emissions of the whole region” vs “a useful tool for regional emission estimates”. Best to clarify.

500: Delete “have” between “we” and “used”

S.I.: The color scale makes it so that the peak (and most critical part) of each plume is invisible (white).