

Atmos. Chem. Phys. Discuss., referee comment RC2
<https://doi.org/10.5194/acp-2021-1061-RC2>, 2022
© Author(s) 2022. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on acp-2021-1061

Anonymous Referee #2

Referee comment on "Local-to-regional methane emissions from the Upper Silesian 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-RC2>, 2022

Review of "Local to regional methane emissions from the Upper Silesia Coal Basin (USCB) quantified using UAV-based atmospheric Measurements"

General comments:

The authors present a manuscript following a measurement campaign in the Upper Silesian Coal Basin in Poland, where they used an AirCore attached to a UAV in order to measure CH₄ and CO₂ concentrations from several individual coal mine shafts. They make the case that a regional estimate derived from shaft-specific measurements will be superior to those that assign a single number to each mine, which broadly takes an average number across all shafts at any mine. They find good agreement between their methane measurements and high-resolution hourly inventory data in some of the shafts, whereas their flights were not able to reproduce well the coarser inventory numbers based on yearly estimates. They also claim that their CO₂ measurements have found that coal mines may be an overlooked source in the region that is not insignificant.

Overall, the paper is well-written and their ideas are clearly presented. The CH₄ analysis, in particular, is laid out in a straightforward manner that is easy to understand. That said, I feel that there are some important elements that are missing and some important changes that need to be made to this manuscript before it will be ready for publication. In particular, I have strong concerns about the CO₂ analysis, as I do not feel that there is enough data presented to support the conclusion that they have found a missing source of ~1% from the regional inventory, rather than it possibly being an artefact of upscaling. This, combined with my concern about how there is no independent value to compare the CO₂ measurements against, and some other smaller concerns I detail later, means my inclination right now would be to recommend leaving the CO₂ analysis out of this

manuscript entirely. I also may have found an issue with the stated method of how the hourly methane emission rate is calculated, for those numbers that they later compare their UAV measurements against and find relationships with, which I would like to see clarified. And I would also like to see further explanation/justification for the first of the three presented upscaling methods, which upscales the yearly inventory numbers using a relationship identified only with the hourly inventory. Further, I would like to see an expanded discussion of the possible areas of uncertainty including: i) the dangers of upscaling with such a small population of data (which they already do acknowledge briefly), ii) the uncertainties and potential misquantifications inherent in their plume calculation methods (the inverse Gaussian and especially the kriging), iii) the possibility of difficulties introduced by sampling at different times of day and under different atmospheric mixing conditions, and iv) at least some discussion of how the background was defined when calculating the leak rates, along with other ideas the authors may think of themselves. I am curious, as well, as to whether the experimental set-up may mean that the AirCore samples are taken downstream of the rotors of the UAV, and whether that may introduce some dilution into their measurements (which may also help to explain why the measured values tend to be lower than the hourly inventory numbers). There are additionally smaller things that should be quicker to fix, but would also be essential, including double-checking the unit scale-factors on each figure that shows CH₄ mixing ratios (which often seem too high by a factor of 10) and the units on the CH₄/CO₂ ratios. I believe that starting with these changes will make a substantial impact on the quality of the manuscript, and that by the time it is ready for publication, it will be a valuable manuscript to the broader community.

Specific comments:

Lines 47-49: Is there a citation for the numbers in either of these 2 sentences? It's a key statement towards the motivation of the study—even referenced in the abstract—so I think it's important to show where those numbers come from.

Lines 109-110: I am wondering how the AirCore was exactly "attached" to the UAV. I see that the AirCore is coiled up, and there is a reference a couple of lines up to "carbon fibre box housing". Is the AirCore contained within that housing, or is that just the housing of the electronics for the UAV? It would help if there was a picture showing the set-up. Particularly I am wondering how it was ensured that the AirCore was measuring from air that was undisturbed by the rotors of the UAV. It looks like this UAV has 4 vertical rotors, and if the AirCore is taking air from underneath (or otherwise "behind" the rotors), then there may be a risk that the rotors are mixing the air (potentially pulling in more dilute air from the background) just before measurement, and therefore affecting the measured mixing ratios. If so, I would be interested in knowing how much effect this may have on the ultimate measurements. And along those same lines, I would wonder about what the effect on sampling rate is when the UAV is moving, considering the primary driver of intake is the ambient pressure. (Is the AirCore exposed in a way that it would sample more when the intake is pointed towards the direction of movement, because of the higher pressure, and vice versa? If so, how might that affect the results?)

Figure 2: The units on these colorscales seem at least a factor of 10 too high. Were the authors really detecting plumes of 150 ppm of methane?

Line 210: This methane sensor gives output as a percentage concentration? Am I understanding that correctly?

Line 213: I might be misunderstanding this sentence. It says "about 5% of the vented air to the atmosphere is from air inflow via the ventilation shaft closure". I understand that to mean that there is some quantity of vented air in this region, and that 5% of that total ventilated air comes from the shaft closure here. That does not sound like the same thing as saying 5% of the total gas flowing through this shaft gets vented. In order for it to contribute 5% of the total vented gas, we would need to know what the total vented amount is, then we take 5% of that number and use that to see how much of the gas flowing through the shaft would have to be venting. So, have I misunderstood the statement here? If not, then the "95% of the flow-rate" scaling factor would not work.

Lines 230-235: This is a lot of words to describe the math, and I think I got a bit lost. Would it be possible to include the simple formulas for these 3 upscaling techniques?

Line 282: I'm not sure that I'm convinced that there is a potential difference between weekend/holiday and weekdays, given the mass balance numbers. The inverse Gaussian numbers seem more like they could suggest that, but is there a reason to trust these more than the mass balance numbers? Feels like one shouldn't hint at a conclusion either way. (I assume that the phrasing "this may indicate..." is maybe an attempt to stay neutral, but it still reads to me like it's leaning towards the conclusion that there is a relationship.)

Figure 6: Maybe this shouldn't be explained in the caption, exactly, but I'm not finding where in the text it explains why certain flights were deemed worthy of a mass-balance estimate but not of a Gaussian estimate?

Line 297: Instead of assuming, is there anyone who could be contacted/referenced that would have more insight into why this period is missing from the inventory data?

Line 305: So the inventory seems to contradict the hypothesis that there's a difference between weekend/holiday and weekday emissions. To me, though, this seems consistent with the lack of conclusions we could have drawn from the data, anyway?

Figure 7: Looking at Pniowek V, for example because it has the longest timeseries, the inventory would lead me to expect higher measured values on the 19th, 21st, and 28th

compared to the 31st and June 1st, but that's not exactly what was seen in Figure 6, which shows low values recorded on all of the flights of the 28th and potentially high values on June 1st. Do we have an explanation for this discrepancy? (I actually think it might have been nice to combine Figures 6 and 7, so that we see the overlay of the measurements against the reported inventory directly.)

Line 313: Wouldn't we expect that the correlation between individual flights and yearly reported emissions would be very low, though? Because day-to-day variability would be so high, in comparison?

Table 2: Could we convert this to a bar chart, maybe? (One could mark the max/min values separately from the error bars, and include the N numbers at the tops or bottoms of the value bars.)

Figure 8: It's difficult to intuit where the 1:1 line would be with rectangular figures like this. Would it be possible to make these figures square with identical limits on the axes, to really visualize the comparison? Maybe with a dashed 1:1 line, for reference? (I understand that this might necessitate dividing this up into 2 figures, in order to fit on the page.)

Figure 8: It also may be helpful to change the legends of each subplot to indicate that it's the best-fit line from the inverse Gaussian approach, specifically, as is noted in the caption

Line 331: What is the justification for forcing the linear fit through 0?

Line 339: The hourly inventory is going to be used to scale up the UAV-measured concentrations?

Line 341: Of the linear fit from the multiple-days-averaged shaft-specific, inverse Gaussian case?

Figure 9: Is this all the same info from Table 2, it seems? If so, maybe we can just get rid of Table 2 and refer to this instead?

Lines 358-9: Does this also imply that the sample size might not be enough to accurately quantify the other sites?

Line 362: It doesn't look to me like there is overlap at Pniowek IV in the mass balance approach...?

Lines 374-5: I think here is where to mention the possible explanations for lower quantification in the air than what the hourly measurements within the shaft show, rather than lines 362-364, which was specifically talking about Pniowek IV

Line 382: One thing I don't think I understand is, if CO₂ has been measured as well as CH₄ from the AirCore, then why not just calculate the emission rate of CO₂ in the same way as was done with CH₄? Why introduce some linear dependence with methane and throw away the data that does not sufficiently have that linear dependence? Is the thinking that, if there are enhancements seen with CH₄, then it's presumed to originate from the shaft, but if there are enhancements in CO₂, they could also be from elsewhere nearby (are there other CO₂ sources nearby, like running engines)? So this is done in order to ensure that one only looks at CO₂ that is believed to be from the shaft?

Line 383: The authors probably should specify which is the numerator and denominator in "slope", even if it seems obvious.

Lines 385-6: Would it be possible to include these scatter plots in the supplemental info, as well? I'm curious to see what they look like.

Line 387: I'm assuming the units are supposed to be ppb/ppm and not ppm/ppm? Additionally, this caused me to look at the figures in the supplemental info, where the flight tracks are provided, and it looks like the scaling factor on many of the colorbars is listed as 10⁴, but it should be 10³, since background methane should only be around 2ppm, not 20ppm.

Line 392: Can the authors explain the NaNs again here? If there's not enough data to include an upper and lower bound, maybe it's better just to state that than to present it as a NaN value.

Lines 407-412: I don't think I'm following the logic here. Figure 8 showed that there was no clear linear relationship between the measurements and the E-PRTR inventory, but that a relationship may instead be found when comparing against the hourly inventory. Then, here, the linear relationship that was found between the hourly inventory and the measurements is used to scale the E-PRTR inventory? What's the rationale for that?

Lines 414-419: Might want to include an acknowledgement that the number of sampled shafts is small compared to the total number of shafts in the region (and among those

sampled, those that have a large number of samples is even lower), so they may not be representative of the region as a whole.

Lines 421-2: My comment from the last paragraph should apply here, too. Though I think this is a much more sound approach than the first approach (which I would be tempted to toss out altogether without a clearer justification for why the hourly linear relationship would be directly applicable to the E-PRTR estimates).

Line 441: When saying that they aren't statistically different when factoring in the uncertainties, should probably also acknowledge that the uncertainty bars are around 30%, which can be quite large.

Lines 448-460: This illustrates the danger of upscaling to a region from just a few measurements. The authors note that coal mining activities are not a major source of CO₂ in the region, and that their measurements are also very low. The flight paths for the CO₂ enhancements are not included, so it's not apparent how clear or strong the CO₂ plumes really are compared to the background. Although Figure 10 shows that, though many of the quantifications do not have error bars, the ones that do are often quite large (e.g. Pniowek IV and Zofiowka IV). And since the E-PRTR inventory does not include coal mines in their inventory, there appears to be no way to independently check whether these values correspond to what would be expected or not.

Lines 475-476: I do not think that one can conclude that the CO₂ inventory is missing a source of about 1%. Without having more information presented about the nature of the CO₂ plumes that were quantified, it seems within the realm of possibility that contemporaneous CO₂ data recorded with the CH₄ data displayed some stochastic variations (especially if the atmosphere is not well-mixed) that could mistakenly be quantified as small plumes with the inverse Gaussian or kriging techniques, especially if the corresponding background values are not well defined. Then, by scaling up those small numbers to the size of the region, they become an apparently large number (~1%). But this feels to me more like a potential artefact of the upscaling than a real missing piece of the inventory. Would we otherwise have any reason to expect large amounts of CO₂ to come out of coal mines? (If so, this is something that I guess should also be addressed in the introduction?) Overall, it is starting to feel like it may be best to leave out the CO₂ analysis altogether.

Line 496: I thought it was only this large for 25 of the 36 flights? And again I think these units are incorrect.

Lines 509-511: I really disagree with this conclusion without some compelling evidence that it's not just an artefact of the upscaling.

Line 516: Maybe the authors should point out that their data indicated that at least 5—and probably more—good flights were needed for a decent quantification of a single shaft.

Lines 513-520: All of this (good) assessment of uncertainties should have, I think, belonged in the discussion section. It's fine to repeat it here, but it felt like it was lacking above. Additionally, included in the discussion of uncertainties should be a discussion of the inherent uncertainties involved in the techniques applied (especially with kriging, which can be a very uncertain way to quantify a plume!).

Technical Corrections:

Figure 3: One of the labels is cut off—the one attached to the red marker.

Line 241: The isotope numbers should be in units of permil, not percent. It's correct in the figure, but not in the text. (May need to be corrected throughout the manuscript.)

Line 300: "emitted emission" seems redundant

Line 328: "on an hourly basis"

Line 332-3: "not significant"

Line 338: "Our evaluations indicate"

Line 370: "all overlap with"

Line 374: replace "more than one flights" with "multiple flights"

Line 381: remove "emission" from "emitted CO₂ emission"

Figure 10: The caption describes these plots as "histograms". I do not believe that's the case.

Line 403: "As many as"

Line 448: "linear curves" should be "linear fits"

Line 490: "show very low..."? agreement? correlation?

Lines 526-28: This last sentence feels like a long fragment instead of a complete sentence, and should probably be reworked