EGUsphere, referee comment RC2
https://doi.org/10.5194/egusphere-2022-749-RC2, 2022

Comment on egusphere-2022-749

Anonymous Referee #2

Referee comment on "Theoretical assessment of the ability of the MicroCarb satellite city-scan observing mode to estimate urban CO_2 emissions" by Kai Wu et al., EGUsphere, https://doi.org/10.5194/egusphere-2022-749-RC2, 2022

GENERAL COMMENTS

This study evaluates the ability of the upcoming MicroCarb satellite to retrieve anthropogenic CO2 emissions based on synthetic measurements. The study follows a well-established approach for such observation system simulation experiments: generating perfect observations based on flux inventories and a model of atmospheric transport, perturbing them based on a model of instrument error, using the perturbed observations to retrieve fluxes in an inverse model of atmospheric transport, and analyzing these fluxes.

While the modeling and analysis methods are sound, I have some concerns on the setup of the experiments and, consequently, applicability of the results to real measurements. In addition, I believe that the interpretation of a key result is inaccurate.

The manuscript is mostly easy to read. A few paragraphs, especially on methods and experimental setup, need clarifications and additional information to better understand the results.

In summary, the manuscript is valuable for the CO2 flux estimation community to understand the capabilities of the upcoming MicroCarb satellite. However, especially the major issues summarized above need to be addressed before publication, which might require additional simulations or better communication of the limitations in abstract and conclusions.

SPECIFIC COMMENTS

In my opinion, the **major issues** to address are:
- The authors omit large point sources in their analyses. Therefore, I'm not sure how the results relate to real measurements, which of course see the integrated signal of all sources (see comments on Sect. 2.3).
- I believe that the authors' interpretation of their results on biogenic fluxes is not accurate (see comments on Sect. 3.3).
- Information on the optimization method is missing, namely how the state vector is set up and, related, how the posterior anthropogenic flux component is obtained (see comments on Sect. 2.7)
- The manuscript needs more information on the observation error model (see comments on Sect. 2.7)

Below I elaborate on these major and some other points in detailed comments arranged by section.

Abstract or Introduction:
State somewhere early on that the study analyses only one out of many sources of uncertainty in emission estimation with MicroCarb, i.e. random measurement errors (as discussed in Sect. 4).

2.2. Cloud screening
This is probably a minor issue, but please comment:
The authors decide cloudiness with a random number generator while in reality there should be some spatial correlation when clouds are bigger than satellite pixels. So e.g. for cloud cover 50%, the analysis is valid for clouds that are smaller or similar in size as the satellite pixel. I don't know whether this is a realistic assumption or what impact larger clouds would have.

2.3 Anthropogenic emissions
The authors exclude point sources from the analysis, citing smaller uncertainty than the areal anthropogenic sources they do include. I don't understand that rationale because anthropogenic fluxes are the main focus of the study and the real measurements will of course include all sources. I think that this omission is problematic, because it alters some results of the study:
- As mentioned by the authors, the prior uncertainty of the areally diffuse sources is higher than that of the point sources. Therefore, the random prior uncertainty used in this study is higher than that of the total anthropogenic emissions. Therefore, the reduction of random uncertainty (the "skill" of the observation system) of city-wide integrated anthropogenic emissions is overestimated.
- The atmospheric signal is reduced due to the omission of point sources, which could imply that (relative) biases in posterior emissions are overestimated in this study.
- Omitting point sources means that this study is about an artificial problem that includes only areally diffuse anthropogenic emissions, whereas the abstract includes statements on "total emissions".

Every OSSE underestimates the skill of an observation system because they cannot

account for/quantify all sources of uncertainty, and one may argue that it is fine to only focus on the harder flux component. However, as explained above, some uncertainties are overestimated while some are underestimated w.r.t. total anthropogenic emissions, so how the results relate to the "real" problem of integrated signals is complex. One way out could be to argue that point source plumes may be removed from the real MicroCarb signal (i.e. reducing the real problem to the one studied here), but this would introduce new errors, and their analysis would fill its own dedicated paper.

I think that the best way forward would be to repeat the analyses but with the strong point sources included. Alternatively, the limitations arising from the omission need to be clearly stated and discussed in the manuscript (but in my opinion this would diminish the value of the manuscript because only a partial problem is studied).

2.7 Experiments

The section on measurement errors needs to be expanded: please provide a reference or add a dedicated, if short, section on synthetic measurement errors, including how they were estimated and possible limitations (e.g. if there are error sources that were not considered).

Please clarify:
Are the mean prior fluxes the true fluxes + bias or the true fluxes + bias + random flux error realization? The text sounds like the latter, but in figures 4 and 5, the variations in the prior look very close to those of the true fluxes. Given that the random error in the fossil fluxes is the same as the bias, and the correlation length is small compared to the satellite swath/sweep, I would expect to see differences in the structures more clearly, i.e. on the order of the bias between prior and truth.

More information on the state vector is needed:
- How are anthropogenic emissions inferred from the posterior fluxes? Are individual flux components (e.g. anthropogenic, GPP, respiration) optimized as separate state vector elements? Or are total emissions optimized and then partitioned (how?) into the individual flux components?
- On what resolution are fluxes estimated, i.e. what is the size of the state vector?
- Is there a temporal resolution?
A new section dedicated to the state vector could be helpful for structuring this additional information.

3.3 Biogenic fluxes

Upon first reading the manuscript, only here did I realize that results so far only considered the anthropogenic flux component. In hindsight it's obvious, because figures 4

and 5 don't show negative fluxes/enhancements. Nonetheless, this approach is unexpected and needs to be made clearer throughout the manuscript, e.g. in the introductory paragraph of Sect. 3 (for example, there are 18 scenarios, not 9), in subsection headings, and in figure captions (and perhaps in Sect. 2.7).

I believe that the interpretations of the results in this paragraph (lines 253ff) are not accurate. The Bayesian posterior random flux error is independent of enhancements (see Eq. (4)), so smaller enhancements do not, as the authors state, limit the ability to reduce the random flux errors (note that the posterior flux is of course not independent of the random error or enhancements). What actually limits the ability to retrieve the anthropogenic component should be uncertainty in the biogenic fluxes, not, as the authors state, the fact that they are negative in summer. Since the authors set the uncertainty of the biogenic fluxes at 25% of the flux, it's higher in summer (according to Fig. A2), so the inversion dumps more corrections into this component then (assuming the different flux components are optimized separately, which, as mentioned above, is not clear). This should explain why the posterior anthropogenic fluxes get closer to the truth (Fig. 8) and why the random uncertainty reduction is better without bio fluxes (Fig. 9), especially in summer.

TECHNICAL CORRECTIONS

Recurring
- Please include the DOI in all references that have one

Line 10: I think it would be clearer if the structure of the sentence were switched around, i.e.:
The three-sweep observing strategy, which generally outperforms the two-sweep mode by virtue of its wider scan area that typically yields more cloud-free scenes, can retrieve the total emissions of the truth within 7% over Paris and 21% over London.

Line 14: See comments on Sect. 3.3. I think it's the random error, not the fact that the biospheric signal is negative.

Line 19: Is there a better peer-reviewed reference for this statement? Bulkeley 2013 is not peer-reviewed.

Line 29: I'd remove "integrated".

Line 35: "subject to" doesn't really fit here - perhaps "but they are subject to"

Line 45: "region thereby" -> "region, thereby"

Line 52: "continuity of collecting ..." -> "continuity with other satellites collecting ..."

Line 61: I'd explicitly write out "TCCON"

Line 68: "analyses" -> "analysis"

Line 92: "still exist" -> maybe you mean "may not be predicted"?

Line 98: "is defined as"

Lines 90-106: Add somewhere in the beginning something along the lines of "We follow the approach by Palmer et al. (2011) and briefly outline the method here" - the section is copied (partly verbatim) from there.

Line 143: In the comparison of Eddy data and SMUrF results, please include the cumulative NEE, over one year or perhaps during the months analyzed in the study, alongside r and slope.

Lines 150ff: State how long back in time the backtrajectories are.

Line 189: "Realistic 20% random error": I don't find this number in a quick scan of either of the cited references. Please elaborate a bit on how you arrive at 20% to represent transport error.

Line 197: Please also provide the bias in terms of percentages of the fluxes.

Line 200: "closer to one" -> "close to one"

Line 202: Similar to line 189: Please explain briefly the reasons behind the choice of 10km as correlation length. A quick look at the cited references might not make it obvious to the reader.

Line 215: "Figure 4" -> "Figures 4a, b, e and f"

Lines 281f: Would the results be better if the position of the city scan were adjusted to the wind direction? I.e. scan further South when the wind is from the North? You could add this to the conclusions.

Line 282: See comments on Sect. 3.3. I think it's the random error, not the fact that the biospheric signal is negative.

Lines 285ff: This should be a major conclusion, alongside the call for measuring additional species.

Lines 296f: "Quantifying sub-city scale emissions requires improving the accuracy and precision of satellite-based CO2 measurements." - Does this refer to the MicroCarb skill in this study, or is it a general statement?

Fig. 2:
- Indicate the overpass direction
- Caption: "halve the cloud cover from one to 0.5" is unclear to me. Sounds like cloud cover up to 50% gets a pass, which shouldn't be the case
  - Update to this comment after reading a bit further: It seems that this phrase refers to the artificial halving described in line 109. If so, the phrase is confusing in the caption of Fig. 2 and I would either leave it out or refer the reader to Sect. 2.2.

Fig. 5:
- Just a comment - the figure is really helpful!

Fig. A1:
- Caption: "red star marks" -> "red stars mark"

Fig. A6:
- Do these figures only include anthropogenic or also biogenic fluxes?