

This paper presents GEMINI-UK sensor network and its network design, which consists of 10 EM27/SUNs. The authors have run numerical simulations to assess the potential error reduction given by incorporating the sensor network. However, the study is very theoretical and can be achieved by only tuning the parameters of the inversion. Therefore, a major revision is recommended. The paper can be published after considering the following comments.

**Major comments:**

1. The Authors assess the theoretical performance of GEMINI-UK only via the relative error reduction via prior and posterior error. You don't discuss any other uncertainties which will influence the performance/results.
2. The authors just mention briefly the uncertainty assumptions (transport error, measurement error in line 273 and prior error in line 230). All the error reduction values in chapter 3.3 and figure 6-8 (and in the abstract) depend strongly on those uncertainty assumptions (mainly the measurement/transport error, R).
3. Chapter 3.1 the statistical analysis is not quite clear, hard to understand the numbers, also a repetition of the numbers in line 367. Write it clearer or include a table. Explain what the ERROR and BIAS is (in figure 2/3), either in the figure description or in the text. In the text you write the biases in ppm/ppb and in the figure in %. Line 369 also a repetition of 361. In figure 2/3, it would be easier to understand if you write in the title what is the difference between the three figures (the boundary conditions). The chapter could be improved to make the statistics clearer and easier to understand, both in the text and in the figures.
4. Throughout the manuscript it is written as "CO2 and methane" rather than CO2 and CH4 or carbondioxide and methane. If there is no specific reason for this, I recommend to change all appearances to either one of the two mentioned options.
5. Line 273: transport errors of 3ppm and 15ppb. How did you select these values? What is the reason for choosing exactly these numbers?
6. Line274: scene dependent observational error: Can you please specify the numbers or give examples of these scene dependent standard deviations and explain the difference to the theoretical limit of the EM27/SUN or other studies? For example, Jones et al. 2021 (<https://acp.copernicus.org/articles/21/13131/2021/>) assumed the uncertainty to be 0.2 ppb based on the Allan analysis to the measured column difference when the instruments were co-located (Chen et al. 2016, <https://acp.copernicus.org/articles/16/8479/2016/>).
7. Line 415: whole section 3.4: It is not quite clear why you do this comparison and why it is relevant here? There are existing EM27/SUN vs. TCCON comparisons. What is the benefit of comparing them in your setup, in particular because you mention that the comparison has its weaknesses of different spectral resolutions, as you mention

in line 431. To my knowledge, it is possible to imitate EM27/SUN recordings with an HR125 by reducing the spectral resolution with the HR125. This was not considered here, why?

8. Line 432: Does this mean, that you used different profiles for TCCON and EM27/SUN retrievals? If so, why?
9. Line 442: Can you please specify "higher and further"?
10. Figure 6: It is not quite clear what the units of error reduction are. Please specify. The "chessboard" pattern further implies a high spatial resolution. This would further imply that single grid cells are affected by error reduction through additional observations differently in small spatial scales. However, it appears to be noise. The amount of additional EM27/SUN observations does most likely not influence single grid cells given the spatial resolution you show here. Why are you sure that this is not noise but a true error reduction? It would be helpful to see the simulated EM27/SUN observations in a plot to make further assumptions.
11. Figure 9: The colorbar is capped for a lot of data points. It is difficult to estimate the relevance of certain areas/pixels given that the surrounding has the same value. Further there are pixels capping the colorbar as maximum positive flux in winter but also cap as maximum negative flux in summer (e.g. around Dublin). If there is no other physically relevant explanation, this might point towards an overfitting of these regions. Can you exclude overfitting for your setup and why?
12. A more careful color classification would increase the information content. It is not possible to distinguish flux values within certain areas (e.g. almost the whole of Ireland for CH<sub>4</sub> in summer).
13. Similar to figure 6 it should be made clear, to what degree the visualisation might be noise or related to noise and what content of the figure is most likely based on true signal. Similar to figure 6, the pixels seem to mimic a much higher spatial resolution than the setup is being capable of resolving. A better visualisation of single regions including some visual interpretation of the noise or the errors would give a better understanding of the scenes.
14. Line 172-174: The data from the ICOS network already exists and is public. Furthermore, what would you use this data for
15. Line 182-185: Why are the excluded sectors are not relevant/not included.
16. Line 188-189: The oceans, especially coastal areas are also a large methane source. Since at least half of the modelling domain is ocean including a lot of coasts, this should be considered. <https://www.geomar.de/en/news/article/how-much-methane-is-released-by-the-ocean>

17. Line 344: Optimal regarding that the network-sensitivity should be the same everywhere? How do you know whether it is optimal - did you brute-force it or use some optimization algorithm?
18. Line 396: Have these apriori estimates been introduced in the paper? The inventories for the GEOS-CHEM simulation have been introduced, but if they are the synthetic “true” emissions, what is the prior?
19. Line 407-408: It would be good to have this additional error reduction as an average over the UK land area not just their maximum values.
20. Line 413-414: How does this statement connect to the figure. It just shows one posterior, not the improvement of the posterior.
21. Line 441-442: The optimality criterion for the sensor placement you mentioned was the sensitivity coverage. Is that equivalent?
22. Line 444-445: These are the values if there were not other measurements. You should use the values from figure 8. Please also adjust the numbers in the abstract.
23. Figure 2: The definition of R, bias and error should be mentioned in the text. Are they calculated only using the daytime hours?
24. Figure 2: It would be good to see these fits at the different sites to get a better idea of where each boundary conditions performs better.

**Minor comments:**

1. 144: instruments worldwide
2. 190: described **by** Deng
3. 195: from
4. 197: 2° x 2.5° make the space consistent (e.g. 201)
5. 201: ~~and~~
6. 213: methane. These
7. 279: ~~of~~
8. 286: July instead of June

9. Line 81: "EM27/SUN"
10. Line 87: You should cite TCCON here since it is the first mentioning of it
11. Line 91: "in situ" or "in-situ"? The text uses both variants
12. Line 111: You could also cite <https://doi.org/10.5194/amt-9-2303-2016>
13. Line 118: You could consider to also cite the PROFFAST Software release itself. This is how the PROFFASTpylot cited it: [https://gitlab.eudat.eu/coccon-kit/proffastpylot/-/blob/joss/paper.bib?ref\\_type=heads#L234-L246](https://gitlab.eudat.eu/coccon-kit/proffastpylot/-/blob/joss/paper.bib?ref_type=heads#L234-L246). Maybe they have a more updated way of citing it by now.
14. Line 122: <https://hitran.org/citepolicy/>
15. Line 144: "instruments worldwide"
16. Line 168: The model has not been introduced at this point
17. Line 177-178: What is the nesting? Is there an inner domain? Mention Figure 1b here again.
18. Line 182: Include the "CAM5-REG-v4" name in the text.
19. Line 189: Lateral exchange from where to where?
20. Line 197-198: This should be mentioned before, when presenting the "nested model".
21. Line 213: Dot is missing.
22. Line 326: How long is the embargo?
23. Line 462: New technologies like?
24. Figure 6: Where does this checkered pattern come from? It would be helpful to have the Gemini-UK locations plotted here as well.
25. Abstract line 10: these ... network. Shouldn't it be this network?
26. Line 46: you could consider to cite also Bayesian inversion studies that incorporate EM27/SUN measurements, e.g.:  
<https://acp.copernicus.org/articles/21/13131/2021/>  
<https://acp.copernicus.org/articles/23/6897/2023/>
27. Line 388: what is considered an observation here? 15min averages or 1min averages? What is your single measurement sample frequency of the EM27/SUN?

28. Line 421: You mention TCCON data is shown but it is only 3 rows later that you explain where it is shown. This is confusing.

29. Line 428: two times "within"

30. Appendix A2: For readers to get more context it could be helpful to briefly mention existing enclosure designs, e.g. <https://amt.copernicus.org/articles/11/2173/2018/>