We want to thank Jocelyn Turnbull for the review of our manuscript and the helpful suggestions for improvements. Our replies are marked in blue.

This paper uses real atmospheric observations of ^{14}C (from flasks) and CO (in situ) that infer ffCO2 values, and convolves them in an atmospheric inversion framework to test the ability of the observations to constrain ffCO2 in an urban area. They demonstrate that even though the ^{14}C-based ffCO2 observations are more precise, the low sampling density is insufficient to produce robust inversion results. In contrast, they show that CO-based ffCO2 values with larger uncertainties but far denser observations can produce robust inversion results. The paper also acknowledges some of the challenges for urban-scale inversions and presents some interesting variations on the atmospheric inversion framework to deal with these. This includes allowing the inversion to scale only total emissions rather than allowing the inversion to induce spatial variability, and fixing the large point source emissions which are expected to be quite well known.

This is a very nice study and I thoroughly enjoyed reading the paper. I have a few suggestions for clarification in the text, and my main comment is that the Discussion and Conclusion sections are overly long and largely repeat what is already stated in the Results section. I recommend shortening and combining the Discussion and Conclusionsections. I recommend acceptance with these minor revisions.

You are right, we tried to shorten the Discussion and Conclusion sections by combining them in our new manuscript version.

Specific comments:

Language and grammar: Please check throughout for language and grammar. I noticed a number of minor errors that should be corrected.

We went through the text and corrected several language and grammar errors.

Page 5, line 140. As we aim to investigate…. This sentence is confusing and should be rephrased.

Done (p. 6, l. 169ff).

Page 6 Line 155, last sentence of paragraph is confusing. Suggest rephrasing to: Times for the hourly-integrated ΔffCO2 observations are reported as the start of the hour e.g….

Page 6. Line 170-174. Is there a possibility of inducing bias by excluding the two sigma outliers? It is likely that the outliers represent some specific atmospheric conditions such as low wind speeds, which might also imply cold inversions that could have different emissions than other meteorological conditions.

We think that the 2σ filter is a quite soft selection criterion, as it only excludes less than 5% of the data. However, you are right that it excludes more events during winter compared to summer. In the case of the ΔCO-based $\Delta ffCO_2$ record, only 18% of the excluded data are from the summer half-year. The remaining 82% of the excluded data are from the winter half-year and especially from synoptic events, which are hard to represent with the model. However, the 2σ filtering leads to only two weeks (one week in January and one week in February 2020) with less than 21 hourly afternoon observations per week. Moreover, there are only 21 days within this two-year period from which no observations are used in the inversion. Therefore, we conclude that the potential of inducing a bias with the 2σ filter is probably small.

Page 10. Lines 246-252. It is not clear from what is written how the authors can be confident that the a-posteriori flux variabilities must indicate the inversion is wrong, versus the priors being wrong. Please clarify.

We also made an inversion run for which we use the monthly varying emissions from TNO for the area source prior emissions (Fig. 1b below, black dashed line). The a-posteriori flux variability is very similar to the results of the inversion run with the temporally flat prior emissions (Fig. 1a below, black dashed line). From this, we conclude that the large a-posteriori flux variabilities are not caused by wrong, i.e. the temporally flat prior emissions.
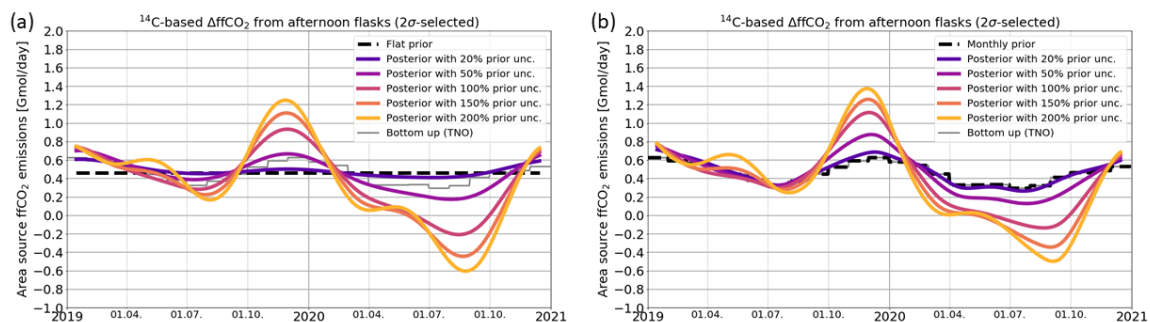

Figure 1: Results of the ^{14}C-based $\Delta ffCO_2$ inversion with (a) flat and (b) monthly varying prior emissions. Fig. 1a is the same as Fig. 3a in the manuscript.

Page 11. Lines 293-295. Does this aggregation of observations apply to the following section 3.2? Or to the previous section? Please clarify, and if this

aggregation applies to section 3.2, I suggest moving this text into the start of that section.

 The averaging belongs to the following section 3.2. We shifted this information to that section (p. 14, l. 384ff).

Page 12. Lines 303-323. This section awkwardly splits between two paragraphs of discussion in the main text, and referring to figures that are only presented in the appendix/supplementary material. I suggest substantially reducing the text in the main document and including the longer discussion in the supplementary material. Alternatively, move the figures into the main document to match the text (although this will make the paper even longer).

 We substantially reduced the discussion about the appropriate averaging interval in the main text (p. 14, l. 386 – p.15, l. 396) and refer to a more detailed discussion in the Appendix C in the manuscript (p. 26-27).

Page 13. Line 356-357. This sentence needs grammatical correction.

 Done (p. 16, l. 433ff).

Page 18. Lines 500-502. This analysis required guessing what the seasonal cycle might be, which is reasonable for this study. But would it be realistic to construct a seasonal cycle in the $\Delta CO/\Delta ffCO2$ ratio by estimating the seasonal contributions of each ffCO2 sector and it's characteristic ratio? I'm not suggesting this needs to be done for this study, but it would be a useful recommendation in the conclusions if indeed it is feasible.

 In the companion paper (Maier et al., 2023a, https://doi.org/10.5194/egusphere-2023-1237) we show that there can be large differences between ^{14}C-based and inventory-based $\Delta CO/\Delta ffCO_2$ ratios, both at an urban and a remote station. Therefore, we recommend to validate the bottom-up ratios by observations before using them to calculate a continuous $\Delta ffCO_2$ record. We included this information in the manuscript (p. 21, l. 601-605).

Page 16 – 20. As noted in my general comments, the discussion and conclusions largely repeat each other, and repeat much of what was said in the results section. I suggest substantially shortening the discussion section and merging with the conclusions section.

 We tried to shorten the discussion section and merged it with the conclusions section.