Reply to reviewer #2

The paper presents the use of a model for atmospheric transport of CO2 in an urban area to estimate CO2 emissions from a set of locally aggregated sources within the model domain of a case study. It explains how the quality of the estimates depends on the configuration of a CO2 sensor network, and on the precision of the individual sensors in that network. This information can be used to optimize sensor networks in urban areas.

The paper is well structured, and the science reported is worth publishing. I had some difficulty with the English from time to time, though. Also, the explanations and the line of thought of some sections were difficult to follow. I added many small comments to the manuscript where I suggest alternative formulations or ask for clarification. The more substantial comments are repeated below for clarity.

Overall, I think the paper does not need any reworking of the work on which the reporting is based, but the text and, to a limited extent, the figures, will need some rewriting and editing. When doing so, please make the captions of the table and the figures more explanatory so they can be read and understood stand-alone. I therefore recommend minor revisions.

We thank Gerrit de Rooji for his encouraging conclusion and his helpful comments. We have adjusted the manuscript based on the suggestions made here and in the annotated manuscript. We also proof-read the manuscript to improve language. We answer the general comments here and additionally reply to the - in our opinion- most important remarks in the annotated manuscript at the end of this reply.

General comments

Many acronyms appear in the paper. Please collect them in a list for easy reference. This will also resolve the issue that not all of them are explained on first use.

We have added a list of acronyms.

You sometimes switch between simple past tense and simple present tense within a paragraph for no obvious reason. Please go over the paper to ensure consistency.

We went over the paper and consistently put it in simple present tense.

Please explain how the term ‘state’ is defined. The term appears frequently, but it is not always clear what exactly is meant by it. I believe it means the CO2 emissions (in what units?) by each emission group, but I am not sure.

You are right with the interpretation. We added a sentence for clarification.

You do not discuss the effect of CO2 transfers across boundary of the modelled domain. Do these fluxes need to be taken into consideration?

This is a shortcoming of our study. In principle, transported emissions have to be considered as a so called “background”. In our study, we do not consider any background implicitly assuming that the background is known in the entire domain. While there are ways of estimating the background, e.g. by using upwind stations in

all major wind direction or by modelling the transported emissions using a mesoscale model, there are always uncertainties associated with background estimation, which we do not account for. It would also require an analysis of how large the systematic biases of background uncertainty are for our region. While this extends the scope of this paper, the framework is able to estimate the effect on background biases and we plan to elaborate on this in future. We have broadened the discussion on this limitation in the Discussion section.

Please explain what TNO data are.

TNO stands for „Nederlandse Organisatie voor Toegepast Natuurwetenschappelijk Onderzoek“. In English, this is the Dutch Organisation for Applied Scientific Research. The institute provides European emission inventories on high resolution. We have added the explanation and also added a Table of Abbreviations.

The sensor network optimization does not include the possibility of installing more precise sensors in areas with large CO2 emissions and cheaper sensors in areas with low emissions, but it seems to me an approach worth exploring. It would allow you, for instance, to minimize the absolute measurement error of the entire network. Or do the agencies/departments operating such networks gravitate toward networks with sensors of a single type?

This is an interesting remark. In principle, it is possible to purchase sensors of different precisions. This has been done for example in Zurich, Switzerland (Emmenegger et al., 2018). However, to our knowledge in most cases, a network consists of sensors of the same type as it reduces the maintenance efforts.
It is also not clear which of the sensors should actually be installed at which position. It might even be beneficial to install the sensors with lower precision close to large CO2 emissions as the signal to noise ratio is still large. However, the outcome will likely depend on the magnitude of the signals and the precision of the sensors. Therefore, using the established framework to analyse this for a given setting might be desirable if a mixed network is planned.
We have not included this analysis in the manuscript as it is rather specific and we would prefer to do this analysis only after adding background and biogenic CO2 emissions to the setting.

Specific comments

Table 1 could use some more explanation it its caption, for instance about the last column and its units, and why there are two categories for Road Transport diesel.

We have added some more explanation in the table caption. As you observed correctly, the names of some GNFR sectors along with the ratios were incorrect, i.e. we missed the “Other combustion emissions” (sector C). We have corrected the table in the manuscript (see also comment to Referee #1). All calculation used the correct values.

Figure 2:

Is the color scale well chosen?. In the top row everything and in the prior column everything is zero.

I do not understand why there are white spaces in row 2. Should not the entire area be covered with pixels?

We have adjusted the color bar. We originally thought it might be useful to have the same color bar for row 1 and 2, but we agree that it hampers visualization of the effects. White spaces are pixels where there is no combustion emission as there are no houses in this square or no emissions reported due to data protection policy. We have added a remark in the figure caption.

L. 263-265:

I think it would be good to discuss the potential effect on the optimal sensor network this simplification (ignoring background concentrations and biogenic CO2 sources) might have. When a network is to be implemented in a real-life situation, there is no way to exclude certain sources - the measured CO2 concentrations will be influenced by all existing sources and sinks.

We agree. In most situations it is vital to account for background and biogenic fluxes. Ignoring background and biogenic sources may actually hamper the result. Therefore, we seek to include these fluxes in the future. We seek to be as transparent as possible with this limitation. Therefore, we have further elaborated on this in the discussion section.

Figure 3 and later figures:

The horizontal axis only states ‘State’, but it is a bar graphs. Do the bars represent emission groups?

Also: please include more tick marks, and have them on all sides.
Done. Yes, state refers to the emission groups. We have changed the figure labels and added the tick marks as well as grid lines in Figure 3 and 4.

Figure 5:

You need to explain a bit how to read and use this figure. Also explain that this figure only applies to a particular configuration in a particular location.

I suppose this figure only becomes useful once sensor prices and installation/maintenance costs are available. You can then create a similar map with the total cost of a set of y sensors with noise x. A given budget will identify which squares on the map can be afforded. You can then go Fig. 5 and pick from this subset the square with the highest relative improvement.

Absolutely right. This figure becomes more useful once sensor (+ maintenance) costs for sensors with different precision are known. We have added an explanation on how this figure gets useful in the revised manuscript following your arguments.

L. 403-404:

Including spatial correlation lengths in the future does not logically follow from the effectiveness of having temporal correlations in your model. Wind directions and velocities vary strongly, which will affect spatial correlation lengths. The ticking of the clock and the daily cycle of basically everything vary considerably less.

Formulated more informally: Temporal correlations of CO2 emissions in Heidelberg are driven for a large part by the heavily synchronized time schedules of humans in a developed society. The wind is not bound by such constraints.

Yes, you are right. It does not follow automatically that a spatial correlation is beneficial just because a temporal correlation is beneficial. The usefulness of spatial correlation needs to be analysed. We deleted the sentence in the revised manuscript.

Comments from the annotated manuscript, which we would like to comment on additionally:

p. 4 Line 99: Does this mean you need to update the model every time a building is built, demolished, or modified?

In principle yes! Especially if there are changes near by a CO2 measurement station.

p.4. Line 117: Referring to "The total concentration enhancement field is obtained

as a linear combination of the concentration fields for each emission group."

Not simply the sum? How do you conserve mass if it is not a sum?

The sum of the emissions of every hour. As hourly emissions were scaled, we referred to it as linear combination. But we changed to sum for clarity.

p. 4 Line 122: Not only the groups and their substructures, but also their emissions over time, right?
Yes, we have corrected this in the manuscript.

p.6, Line 147: I thought the hourly steady states applied to the wind field only. Somewhere you lost me in the train of thought, apparently.

It equally applies for the concentration. For each hour, the emissions of that hour are transported using the steady-state wind field of that hour. So, we also chain the concentration fields hourly. The procedure is described in detail in Berchet et al. (2017). For clarity, we have added a sentence in the model description.

p. 6, Line 153: The emissions have a lower bound of zero but no upper bound. Is a symmetric distribution acceptable? Negative emissions can occur if drawn from a Gaussian distribution.
Defining and quantifying the prior uncertainties is very difficult. In most cases, emission inventories do not even publish an uncertainty estimate along with their best estimate. Often, a normal distribution for emission estimates is assumed as described in Solazzo et al. (2021). This assumption makes it easier to process the data and account for uncertainty in the inversion process.

Even though in principle, it is correct that the uncertainties are not distributed normally, we consider the impact on our conclusions negligible.

p.9, Line 201: From this sentence, I conclude that it includes traffic emissions, correct? Do you assume that fuel imports (by cars that filled up elsewhere before arriving in Heidelberg) cancel exports by cars that filled up in the city and then left)?
Actually, combustion emissions refer to heating emissions and does not comprise traffic emissions. We made this explicit in the revised manuscript. Traffic emissions however only account for the emissions caused by driving (traffic) in the GRAL domain.

p.9 Line 215 You have the city-wide fuel consumption data, as well as that for several districts. Would it therefore not be better to subtract from the city-wide consumption the known district consumptions, and divide the remaining consumption over the masked districts? This fuel is consumed and will affect the observed CO2 concentrations in real life. Also, having this additional CO2 source in your data set (albeit without or with an approximate spatial resolution) probably will affect the optimum location of sensors,
Actually it is not whole districts, which are masked, but only some 100m x 100m squares within the domain, which have been masked. We expect the overall masked emissions to be small, but we do not have city-wide fuel consumption data from the same data source to compare to. However, we only use the combustion emissions as a realistic truth to test the monitoring network. As in every OSSE, we are aware that the true emission pattern is close to, but not equal to the truth in this OSSE. We made no changes in the revised manuscript.

p.10 Table 1: What is the difference between E and F2? The categories Fugitives, Solvents, and Off Road are not intuitively clear. I suppose that categories with zero or tiny emissions are simply not represented in the study area. Is that correct?
As mentioned in the answer to reviewer #1, we have made a mistake in Table 1, which we corrected for in the revised manuscript. The categories A-L are the emission sectors as defined by Gridded Nomenclature for Reporting (GNFR) sectors. We added GNFR in the Table of Abbreviations and elaborate in the figure caption. Categories with zero emissions are not reported in the study area. Categories with tiny emissions, hardly appear in the study area.

p.18 L356 Do you think a similar analysis of real-life data is possible? It could give interesting results.

One could do a similar analysis based on the posterior emissions determined. However, these will not be independent of the prior. Alternatively, there might exist data sets, e.g. hourly traffic counting or energy consumption data, which could actually inform on the expected temporal correlation of the underlying emission data. However, we have not added a comment in the manuscript as it strongly depends on which data is available.

p. 5 Line 125 This (refers to Gaussian noise in Equation 1) in itself is a model of the measurement errors, is it not?

Yes, it is an assumption commonly made that the noise is Gaussian. However, we have not made any further comments as it is widely excepted and used.

p.20, L.377 More generally, I think the results of simulation experiments as those reported here will always be site-specific. That is not a problem though, because I cannot see any benefit to the configuration of sensor networks or to the development of emission-reducing measures at different sites by comparing these sites quantitatively.

This is correct. Quantitatively, we do not expect same results.

References in the Reply:

Berchet, A., Zink, K., Oettl, D., Brunner, J., Emmenegger, L., and Brunner, D.: Evaluation of high-resolution GRAMM–GRAL (v15.12/v14.8) NOx simulations over the city of Zürich, Switzerland, Geosci. Model Dev., 10, 3441–3459, https://doi.org/10.5194/gmd-10-3441-2017, 2017.

Emmenegger, L., Müller, M., Berchet , A., Graf, P., Meyer , J. , Hüglin, C., Brunner, D.: A Low-cost CO2 Sensor Network at Regional and Urban Scale, 2018: https://meetingorganizer.copernicus.org/EGU2018/EGU2018-9452.pdf

Solazzo, E., Crippa, M., Guizzardi, D., Muntean, M., Choulga, M., and Janssens-Maenhout, G.: Uncertainties in the Emissions Database for Global Atmospheric Research (EDGAR) emission inventory of greenhouse gases, Atmos. Chem. Phys., 21, 5655–5683, https://doi.org/10.5194/acp-21-5655-2021, 2021.