Optimising Urban Measurement Networks for CO2 Flux Estimation: A High-Resolution Observing System Simulation Experiment using GRAMM/GRAL

Reply to reviewer #1

General comments:

This study presents a study on the placement of a monitoring network based on a process model to optimize CO2 carbon fluxes in urban areas. The research addresses several critical aspects related to the precision and effectiveness of the proposed measurement network, including sensor quantity and quality, optimal sensor locations, the potential inclusion of carbon monoxide (CO) measurements, and the introduction of temporal correlations into prior emissions. A notable strength of the study is its practical relevance, aiming to inform decisions concerning sensor deployment in real-world urban settings, with Heidelberg, Germany, serving as a case study. This approach has the potential to guide similar efforts in other urban areas, making it of interest to both researchers and policymakers. However, there are a few aspects that could benefit from further attention or clarification in the manuscript. For instance, providing insights into potential challenges or limitations of the proposed approach, such as data availability and cost considerations, would be valuable for readers seeking to replicate or adapt the methodology. In summary, this study constitutes a valuable contribution to the field of urban CO2 flux estimation and measurement network design. With minor improvements in the clarity of methodology and the consideration of potential limitations, it has the potential to be a valuable reference for both researchers and practitioners involved in urban environmental monitoring and management.
We thank the reviewer for the helpful comments and for acknowledging the relevance of the system for researchers and policymakers. We revised the manuscript providing additional information on the uncertainties and limitations and answer to the specific comments below.

Specific comments:

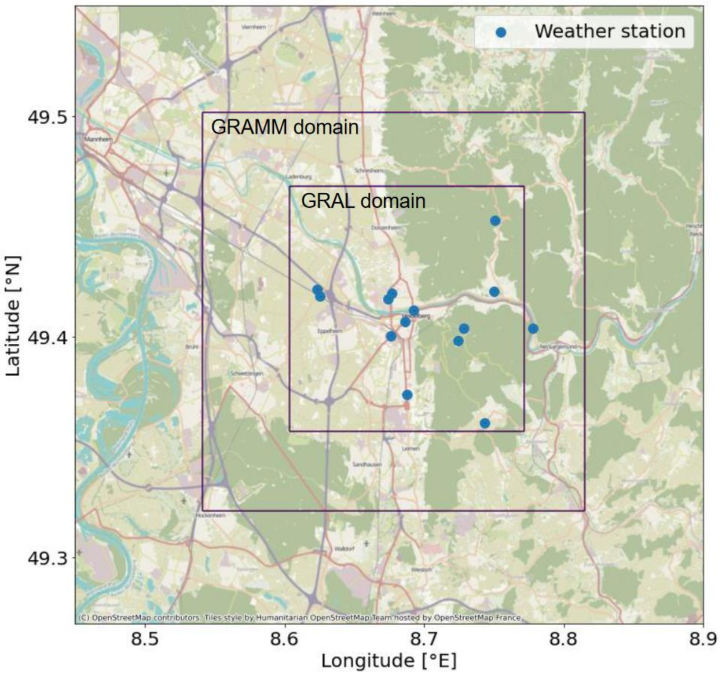
Page 2 and Line 36: Clarity on Observing System Simulation Experiments (OSSEs), When introducing Observing System Simulation Experiments (OSSEs), provide additional context for readers who may not be familiar with this term. Explain briefly how OSSEs work and their role in assessing monitoring networks.
OSSEs provide a controlled and consistent framework for assessing the performance of inversion methods used. In an OSSE emissions as well as atmospheric transport are known. The concentration is obtained by simulating the atmospheric transport of the emissions into the atmosphere. The concentration at selected sites can then be used in an inversion framework to estimate emissions. It is possible to e.g. add measurement uncertainty or model transport uncertainty to the concentration, or to change the prior emissions and evaluate the effect on the emission estimate by comparing to the known true emissions. Therefore, an OSSE enables isolating and analyzing various factors that contribute to uncertainties and errors in emission estimates. We have modified and extended the description and role of OSSEs in the revised manuscript.

Page 2: Consider providing definitions or explanations for key terminology used in the introduction, such as “pseudo observations” and “Jacobian” This will aid readers in understanding the technical aspects of your research.
Pseudo observations refers to the modelled concentration field sampled at selected measurement sites. We obtained the concentration by forward simulating the true emissions using the atmospheric model and adding a model-data mismatch error to

mimic measured concentration data. We have removed the terminology of pseudo observations and explicitly explained it. The Jacobian matrix is a linearization of the forward model representing the sensitivity of the observation to the states. We have added this description. We have read through the paper again and added an explanation where we thought it might be helpful or where it was specifically suggested by reviewer #2.

Line 100: Mentioning that the wind field resolution for GRAL is 2m with a total of 200 cells is informative, but you could briefly explain why this level of detail was selected and how it impacts the model's accuracy or performance.
GRAL has 2m vertical resolution and 10m horizontal resolution. These properties have been chosen following previous studies e.g. by Berchet et al. (2017). 10 meters is approximately the size of a street in an urban area, therefore enabling the simulation of street channeling. For Heidelberg, May et al. (2023) evaluate the model performance with these settings. The paper is under revision in Atmospheric research. May et al. (2023) show that for this resolution the urban meteorological fields can be simulated very well in Heidelberg with small biases and RMSEs in wind speed and wind direction. We do not have any information how a lower or higher resolution would impact the results in Heidelberg.

Consider optimizing Figure 1 by suggesting that the GRAL domain be displayed directly within the GRAMM domain. Also, shows the basic outline of the city of Heidelberg. Additionally, supplement the horizontal and vertical coordinate headings with the units of latitude and longitude, and include a legend.



The GRAL domain was already displayed in the GRAL domain. We have added coordinate headings with unit and included a legend. For clarity, we have decided to leave out the inlay showing emissions and we have not added city boundaries as the urban areas can already be identified looking at the underlying OpenStreetMap and additional boundaries would make the plot more difficult to read. However, note that Figure C1 displays the district borders in the GRAL domain.

Line 202: Explain the significance of using administrative districts as emission groups in your study. How does this choice impact the optimization process, and why were small districts and border districts aggregated?
The choice of state vector is important as it has political, as well as numerical implications. Jungmann et al. (2022) hypothise that providing CO_2 information on high resolution may provide policy makers with information, which may enhance ambition and ability for climate mitigation. At the same time, reliable and independent information of emissions on high resolution requires a large number of sensors to constrain the fluxes. Therefore, we chose the administrative districts as meaningful political unit, which can still be constrained with a realistic number of sensor nodes. We chose to aggregate smaller and border districts as they are very difficult to constrain as they contribute only weakly to an overall enhancement. In principle, other choices of emission groups are possible and may be explored in further analysis.

Line 216: TNO Abbreviation: Clarify what TNO represents (if it's an abbreviation).
TNO stands for the Netherlands Organisation for Applied Scientific Research. We have added this in the manuscript. We have also added a Table of Abbreviations to the Appendix.

Line 226: It may be helpful to add a brief explanation of Monte Carlo experiments and the analysis process to elucidate the concept for readers unfamiliar with Monte Carlo experiments.
Monte Carlo experiments are simulations that randomly sample a model variable, in our case sensor location, to estimate the probability of having a certain outcome, in our case of having a certain information content of the inversion. We have added a sentence in the manuscript.

Refine the information on Figure 2 map sheets and ensure that maps include a legend, latitude, longitude, and compass information.
Done.

The discussion of the uncertainty analysis of the overall model development is somewhat sparse and scattered. There is a need for additional integration of the discussion of uncertainty.
We have added a discussion on uncertainty in the new discussion section. We address some important sources of uncertainty including model transport and instrumentation errors, neglecting biogenic fluxes and transported background concentration, as well as the choice of state vector in the inversion framework itself.

Subfigures Numbering: Add numbers and subfigure titles to subfigures in some groupings. For example, Figures 2, 3, 4, and 6.
Done

Modify "a.)" in the figure captions to "(a)" to indicate this, and follow the same format for other subfigures.
Done

I read your paper published in 2015 "Vardag, S. N., Gerbig, C., Janssens-Maenhout, G., and Levin, I.: Estimation of continuous anthropogenic CO_2: a model-based evaluation of CO_2/CO, CO, $\delta^{13}C$-CO_2, and $\Delta^{14}C$-CO_2 tracer methods, Atmospheric Chemistry and Physics, 15, 12705-12729, https://doi.org/10.5194/acp-15-12705-2015, 2015." and comments from reviewer Jocelyn Turnbull. I found the comment that the study by Vardag et al.,(2015) suggests that in Europe, CO may not be as available as a tracer of fuel CO_2

as it is in other regions due to the low ratio of CO:CO2 emissions from European transportation. In contrast, in section 3.3 of this study, the tracer role of CO is emphasized, especially in distinguishing between different emission sources, such as transportation emissions. However, the fact that transportation emissions are generated from fuels, which includes emitted CO2. there is a need for further clarification or discussion as to what causes this discrepancy.
This is a very helpful remark. Most fossil sources emit CO as well as CO2 in a given ratio. We list the ratios used in a revised Table 1 as there have been a mistake in the previous table 1. We have realized that we used the incorrect names of the GNFR sectors and in few cases miscalculated the emission ratio. Note that the correct values were used in all calculations. We have corrected the table in the manuscript. As one can see, the emission ratio for CO depends very much on the GNFR sector. Even within the traffic emissions (F1-F3) there is a huge spread of different emission ratios. In general, the more distinct (e.g. the higher) the ratio CO/CO2 for a given sector compared to the other sectors, the better the tracer for the sector. Therefore, it is in accordance to Vardag et al. (2015), who state that the quality of CO as tracer for fossil fuel is deteriorating. However, it is still a valuable tracer for fossil fuel as was confirmed later e.g. by Maier et al. 2023 and Kim et al., 2023.

Note also that the study by Vardag et al. (2015) uses mean ratios of CO/CO2 (plus fixed diurnal cycle) and therefore does not consider a variation of CO/CO2 ratio dependent on the area of influence, i.e. on the footprint. This variation complicates a good estimation of fossil fuel CO2. However, in this study we conduct an actual inversion taking into account the area influencing the enhancement at every hour and therefore the emission ratio itself actually varies depending on the footprint. Therefore, the spatial variation of CO/CO2 can be taken into account in this inversion. In the manuscript we already discuss that the inversion result depends on the actual CO/CO2 ratio used. We have now picked this up in the discussion as well.

Note additionally, that we have updated Figure 8 as we previously used the CO/CO2 emission ratio of F1 instead of the weighted mean CO/CO2 ratio of F1-F3 to construct the CO concentration record for traffic as would be correct and as the text in the manuscript implies. Figure 8 now shows the estimation of total CO2, traffic CO2 and combustion CO2, when using the mean ratio for total traffic emissions (weighted mean of F1-F3 about 9 ppb/ppm) for the construction of the CO record. As expected, we see a slight deterioration of CO as tracer for traffic emissions, but no changes in the general picture. No changes in the text of the manuscript were made.

The conclusions section is quite long, and some of its content overlaps with the results and discussion. Consider optimizing the structure of the manuscript, and it is recommended to add a separate discussion section. The conclusions should be summarized insights based on the results of the entire study.
We agree that the manuscript could benefit from separating discussion and conclusion section. We have restructured the final chapter such that we have included a discussion section, in which we discuss the most important results and in which we add a discussion of uncertainty as requested above. The conclusion section now only contains the summarized insights of using GRAMM/GRAL for the inversion.

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.

Jungmann, M., Vardag, S.N., Kutzner, F. et al. Zooming-in for climate action—hyperlocal greenhouse gas data for mitigation action?. Clim Action 1, 8 (2022). https://doi.org/10.1007/s44168-022-00007-4

Kim, J., Miller, J. B., Miller, C. E., Lehman, S. J., Michel, S. E., Yadav, V., Rollins, N. E., and Berelson, W. M.: Quantification of fossil fuel CO2 from combined CO, δ13CO2 and Δ14CO2 observations, Atmos. Chem. Phys., 23, 14425–14436, https://doi.org/10.5194/acp-23-14425-2023, 2023.

Maier, F. M., Rödenbeck, C., Levin, I., Gerbig, C., Gachkivskyi, M., and Hammer, S.: Potential of 14C-based versus ΔCO-based ΔffCO2 observations to estimate urban fossil fuel CO2 (ffCO2) emissions, EGUsphere [preprint], https://doi.org/10.5194/egusphere-2023-1239, 2023.

May, M., Wald, S., Suter, I., Brunner, D., Vardag, S.N.: Evaluation of the GRAMM/GRAL model for high-resolution wind fields in Heidelberg, Germany, 2023 in review in Atmospheric Research

Vardag, S. N., Gerbig, C., Janssens-Maenhout, G., and Levin, I.: Estimation of continuous anthropogenic CO2: model-based evaluation of CO2, CO, δ13C(CO2) and Δ14C(CO2) tracer methods, Atmos. Chem. Phys., 15, 12705–12729, https://doi.org/10.5194/acp-15-12705-2015, 2015.

Reply to reviewer #2

The paper presents the use of a model for atmospheric transport of CO_2 in an urban area to estimate CO_2 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 CO_2 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 CO_2 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 CO_2 concentrations in real life. Also, having this additional CO_2 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.