Geosci. Model Dev. Discuss.,
https://doi.org/10.5194/gmd-2019-283-RC1, 2019

Geoscientific Model Development Discussions

Interactive comment on "Optimizing a dynamic fossil fuel CO_2 emission model with CTDAS (v1.0) for an urban area using atmospheric observations of CO_2, CO, NO_x, and SO_2" *by* Ingrid Super et al.

Anonymous Referee #1

Received and published: 16 December 2019

This manuscript presents a modelling framework to optimize fossil fuel CO2 emissions using a data assimilation system and atmospheric observations. The prior emissions are estimated using a dynamic CO2 emission model, which allows constraining physically relevant parameters. The manuscript provides a novel approach, that can overcome some current limitations in urban-scale inversions such as source attribution, definition of the prior emissions and its uncertainties, and the sensitivity to errors in atmospheric transport.

The paper is well written and clear and a very good contribution for GMD. Results are presented in a detailed way and the conclusions are well-reasoned. My only ma-

jor comment has to do with some of the subsections of the methods sections, which sometimes are not presented in sufficient detail and/or remain a little bit too general.

1. Section 2.1.1 and Table A1: How is the "E/A" term derived from the IEA statistics (L175)? According to Table A1, "E/A" values are derived from CBS and KNMI (description of these acronyms should be provided). To the best of my knowledge, the information that IEA reports is primary energy consumption by sector and fuel, which is equivalent to the "E" term of equation 1. Should not it be more efficient to directly use the "E" information provided by IEA instead of deriving it from the expression A * (E/A) proposed in equation 1? I find difficult to understand what is the added value of having to compile the "A" and "E/A" terms instead of directly using "E". Also, when describing "A" some examples are used such as "vehicle kilometers driven" (L161), but according to Table A1, the units used for the term "E/A" in road traffic cars and HDV are "PJ/mln€ ́. Should not it be "PJ/km"? The "A" terms and corresponding sources of information should also be provided in Table A1.

2. Section 2.1.3: This section remains too general, especially when compared to the previous one, where the temporal disaggregation methodology is presented in a detailed way for each sector. It is not clear to the reader the specific datasets/methods that are being used to spatially distribute the emissions for each sector. More details should be provided (perhaps the spatial proxies used should also be summarized in Table A1). Later on, in the manuscript, the authors say that the spatial distribution is assumed to be well-known (L346) and therefore this element is not considered when performing the uncertainty analysis. This sentence however seems at odds with a previous statement, which says that "their uncertainty increases rapidly when disaggregating them towards finer spatiotemporal resolutions" (L52). Considering the special increase in the emissions uncertainty that the introduction of spatial disaggregation generally causes, the non-inclusion of this element in the uncertainty analysis should be better justified (i.e. better discussed why the spatial proxies applied in this study can be assumed to be well-known).

In addition to these major comments, I list several doubts and minor comments mostly related to suggestions to improve the presentation of the work:

L94: Change (Andres et al., 2016) (Super et al., 2019) for (Andres et al., 2016; Super et al., 2019)

L104: I think that the concept of "near real-time" is too strong. For instance, this would imply that traffic emissions are estimated based on near-real time data collected from traffic counts and, therefore, that congestion situations or traffic accidents are considered when calculating the dynamic emissions. A similar thing would apply to power plants (e.g. emissions are derived from near-real time collected data on the activity of each individual facility).

L120: Replace "inverse part" for "inverse modelling part"

Table 1: Could you provide a reference to the CO2 contribution shares that are shown in Table 1?

L206: Some European studies have suggested the use of 15.5°C as the value for defining the threshold temperature when calculating the HDD (e.g. https://rmets.onlinelibrary.wiley.com/doi/epdf/10.1002/joc.3959). According to the results shown in Figure 3 (left), the parametrization proposed for households (18°C) is underestimating most of the observed peaks in winter, while it overestimates the ones observed during spring/summer. On the contrary, the parametrization proposed for glasshouse (15°C) reproduces much better the winter peaks. Do you think that reducing the value of Tb for the household parametrization could allow improving the reproduction of winter peaks? (this is just a suggestion, does not need to be added in the revised manuscript)

L220: Are you referring to the MACC-III fixed temporal profile? Please specify

L239: Add a reference to the ENTSO-E database (e.g. https://www.sciencedirect.com/science/article/pii/S0306261918306068)

L244: The correlations presented between power generation and meteorological variables are rather low. This implies that the proposed parametrization for this sector is not so well correlated with observed activity data such as it is for other sectors (e.g. households or road transport). Considering the importance of this sector to the total CO2 emissions, perhaps it would be interesting to discuss how these parametrizations could be improved in future works.

L260: Could you also provide the R2 value for daily data?

Table 1 / L272 / Figure 9: The industrial sector is the largest contributor to total CO2, but at the same time is the only sector that has not been split between subcategories. Is there a specific reason for that? Should not a split between e.g. type of industries would help to provide better temporal parametrizations or reduce the uncertainty of the emission factors for this sector?

Figure 5 (left): It looks like the activity data (red line) is missing for the last day

Figure 9: According to this figure, the uncertainty of the time profile "T" is larger in the household sector than in the power plant sector. Nevertheless, the correlation between the temporal parametrization and true activity data reported for the household sector is higher than the one reported for the power plant sector. Is there a specific explanation for that?

Section 3.1: I assumed that the meteorological-dependent time profiles were calculated using the WRF model, but perhaps it should be clarified at some point in this section.

Section 5: In the introduction section the authors pose three research questions that want to answer with this study. It would be interesting to rewrite the conclusions section so that it provide concise and clear statements that directly answers each one of these research questions (i.e. include a bullet list with a statement per question). This structure may facilitate the reader to link the posed questions with the outcomes of the work.

Interactive comment on Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2019-283, 2019.