Dear Reviewer, we appreciate your time and effort in acknowledging and thoroughly reviewing our manuscript. We are sincerely grateful for your constructive comments and insightful suggestions, which have encouraged and helped us to improve the manuscript. We have revised the manuscript carefully based on your comments.

In the responses below, your comments are provided in black text and our responses are provided in blue text.

General comments:

This manuscript simulates CO_2 concentrations in Belgium and surrounding areas and conducts sensitivity tests using different emission inventories and in situ and remote-sensing CO_2 observations. The topic is interesting and meaningful, but many statements and explanations in the manuscripts are not rigorous enough. I suggest more modifications and improvements before acceptance.

Special comments:

1. The title includes Western Europe, but this study only focuses on Belgium and the surrounding areas. Is it reasonable to use Western Europe in the title?

We appreciate the reviewer's suggestion. The title has been revised as "*WRF-Chem simulations of CO_2 over Belgium and surrounding countries assessed by ground-based measurements*".

2. This paper also mentioned that the drought could increase CO_2 I think it is necessary to include precipitation in the research period and compare it with the year before and after.

We appreciate the reviewer's suggestion. The sentence "*According to records from the Royal Meteorological Institute of Belgium (RMI) at the Uccle station, the total precipitation for the summer (June to August) of 2018 was 134.7 mm, which was substantially lower than the climatological mean for the summer period 1991–2020 (234.2 mm).*" has been added to lines 101-103 of the revised manuscript.

3. I think Figures 1 and 2 can be combined.

Done. The figure has been revised as follows.

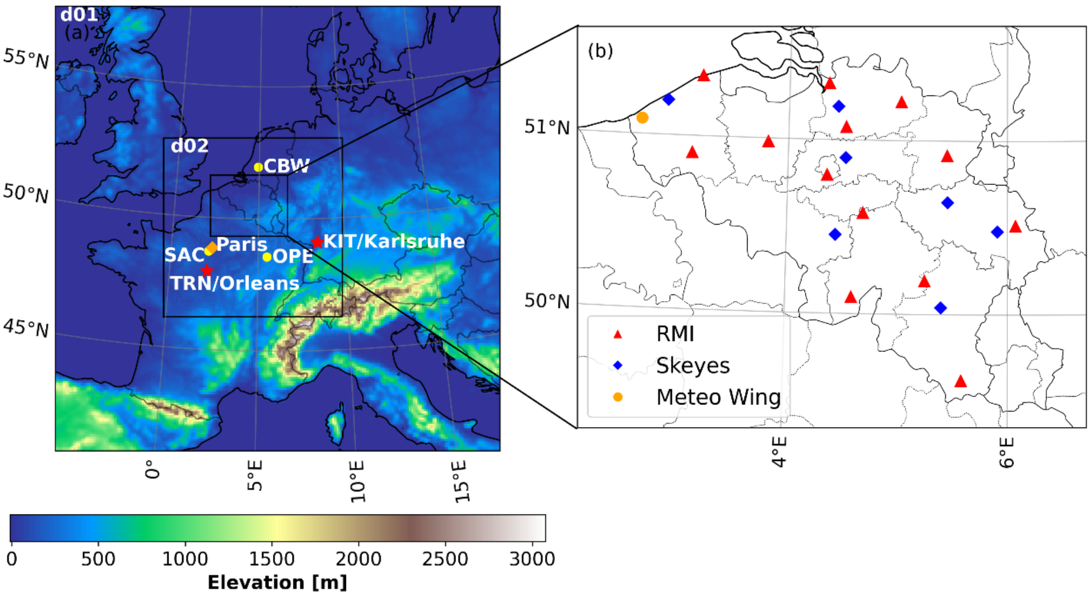

Figure 1. Terrain elevation map of the simulated domains, with horizontal resolutions of 9 km (d01) and 3 km (d02), showing ICOS (yellow dots), TCCON (orange diamond) and co-located (red stars) sites within d02 (a), and synoptic stations in Belgium for which data are available for our study period (b).

4. Lines 130-135: I suggest consistency in the parameters used in the equation and the text. For example, "T_{scale}" in Eq and "Tscale" in text. Maybe "T_{scale}" is more suitable. Also, other parameters such as Wscale, Pscale, Ts, Tmin, Tmax, …

Done.

5. Lines 158-159, Here are two downscaling methods used. What are the differences between them?

In these two methods, the applied time factors are different.

For the CAMS-REG-ANT emission inventory, we applied the temporal factors provided by CAMS-REG-TEMPO for downscaling, as this dataset is consistent with CAMS-REG-ANT in terms of sector classification, spatial resolution, and geographical coverage. In this method, sector-specific temporal factors were used for the downscaling. For example, emissions from the sector H_Aviation were assumed to have no temporal variability, therefore, the annual emissions were simply divided by the total number of hours in the year to obtain hourly emissions. While emissions from the sector A_PublicPower vary by grid cell on monthly, weekly (day-of-week), and hourly timescales, and these corresponding temporal factors were applied to downscale the annual emissions from the A_PublicPower sector. Similarly, for the remaining sectors, we also applied their respective temporal factors. Finally, emissions from each sector, after applying the corresponding temporal factors, were summed to obtain the total hourly emissions.

For the TNO emission inventory, we used the downscaling method proposed by Nassar et al. (2013). This method downscales the total anthropogenic emissions instead of applying sector-specific downscaling. In this method, anthropogenic emissions vary by grid cell according to the day of the week (Monday to Sunday) and on hourly timescales, after being interpolated onto the WRF grid, these time factors are used to downscale the total anthropogenic emissions.

6. Table 1. I suggest adding some words in the column of the aberrative. Also, the CBW attitude is 0?

In Table 1, it seems that we did not address the term "aberrative."

The CBW site is located in the Netherlands, and its elevation is 0 m; this information is obtained from the ICOS website.

7. Figures, the figures in the main text are not clear as the Figures in the Appendix.

We apologize for the inconvenience. All figures included in the manuscript comply with the ACP journal requirements and are provided in PNG format with a resolution of 300 dpi. Some figures may appear less clear because they contain a large amount of information and have large original sizes, which may reduce visual clarity when the figures are scaled. We have revised most of the figures in the main text to improve their clarity as much as possible.

8. Figure 3. What are the sunrise and sunset times in these ICOS stations? The highest temperature occurred at 18:00 local time, and the PBL reached its maximum height between 15:00-17:00 (line 319). It seems unreasonable.

We thank the reviewer for this comment. During the simulation period, according to NOAA Solar Calculator, the sunrise at the ICOS sites involved in our study occurs around 5:00 - 7:00 (local time), and sunset occurs around 20:00 - 22:00, and the sun reaches its highest position around 14:00 LT at most sites (energy saving time, UTC+2).

We have double-checked. As shown in Figure 2 of the revised manuscript, the simulated diurnal variation of temperature is consistent with observations. For the PBL, we extracted the corresponding PBL data from ERA5. Figure A presents the diurnal cycles of the PBL provided by ERA5 (blue) at each ICOS site. It can be seen that, at all sites, the diurnal variation of the PBL simulated by the WRF model is generally consistent with that from ERA5. Therefore, the consistency with observations and ERA5 data demonstrates the reliability of the results.

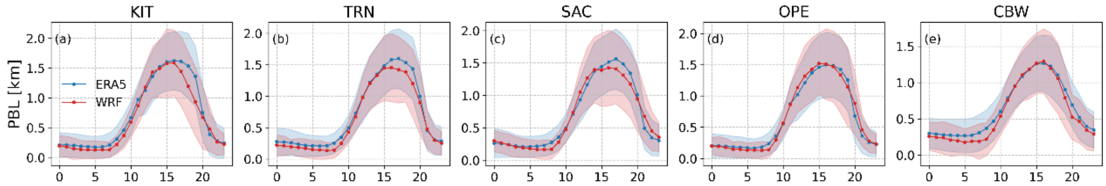

Figure A. The diurnal cycles (local time, UTC+2) of planetary boundary layer height simulated by WRF-GHG (red) and ERA5 (blue) at the KIT (a), the TRN (b), the SAC (c), the OPE (d) and the CBW (e) sites.

9. Figure 5. I think it is better to keep the y-axis of CO_2 concentrations the same across different emission inventories at the same height.

Done. The figure has been revised as follows.

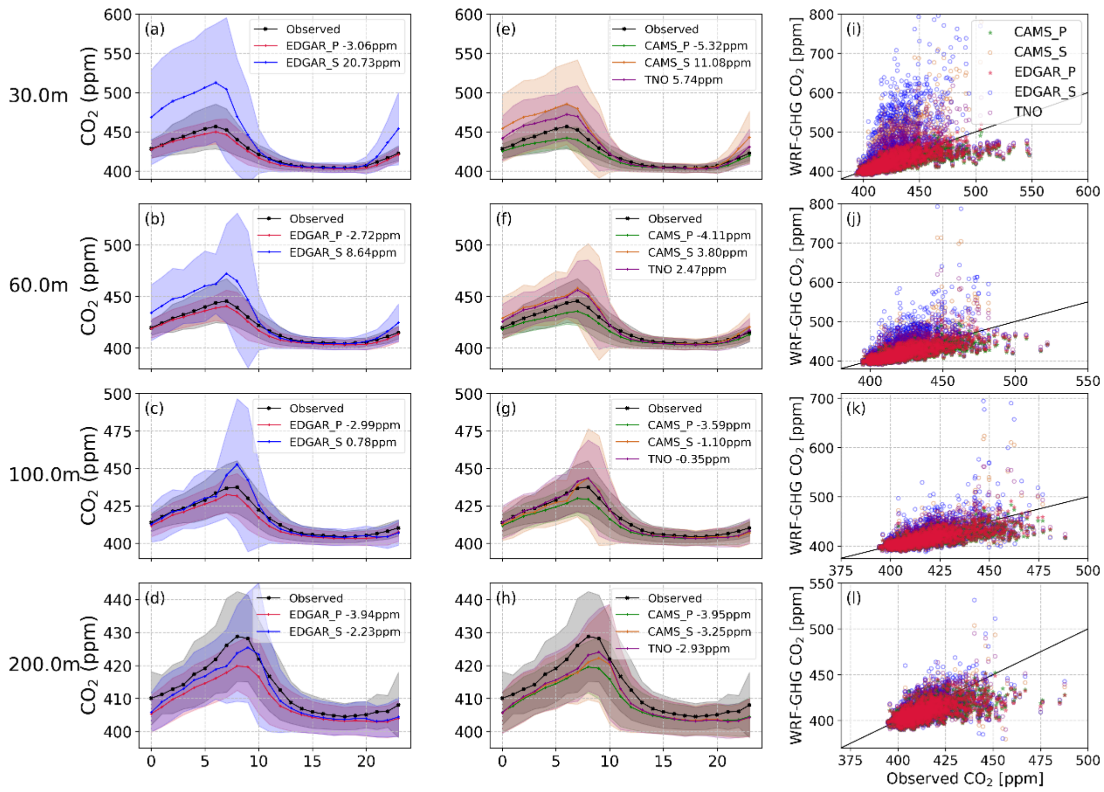

Figure 4. Diurnal cycles (local time) of simulations with different anthropogenic emissions and observations (a-h), where the values represent the MBE between each simulation and observations, along with scatterplots comparing each simulation to the observations (i-l) at different heights at the ICOS KIT site.

10. Figure 6. It seems there is no contribution from biomass burning in these figures. Why are biogenic contributions nearly negative from 10:00- 23:00 at all sites? Why are biogenic contributions only positive around 10:00?

The contributions from biomass burning and the ocean are both nearly zero; therefore, in Figure 6 (Figure 5 in revised manuscript), the biomass burning values are overlapped by the ocean contribution, making them appear not to be shown.

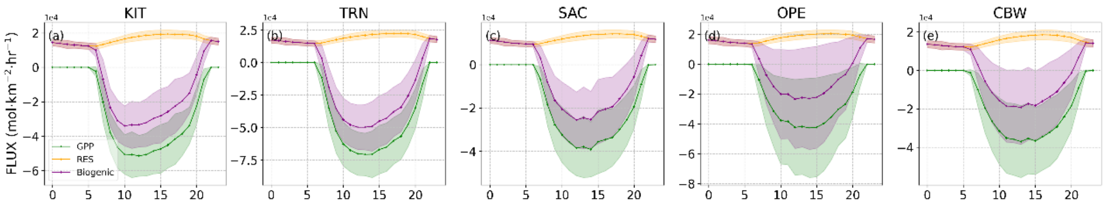

Figure B. Diurnal cycles of biogenic fluxes (GPP, RES, and Total) at five ICOS sites.

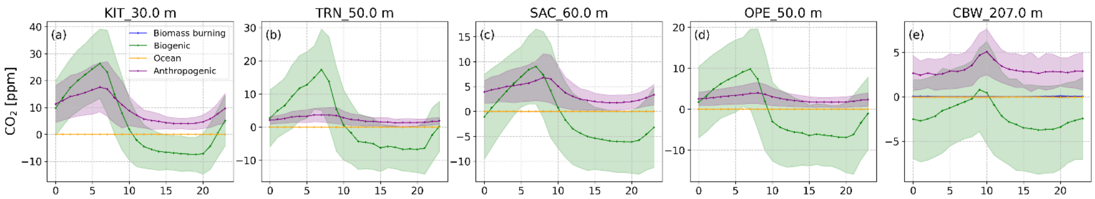

Figure 5. Diurnal cycles (local time) of simulated tracer contributions at 30 m at the KIT (a), 50 m at the TRN (b), 60 m at the SAC (c), 50 m at the OPE (d) and 207 m at the CBW (e) sites. Here, the anthropogenic emissions are based on EDGAR v2024, taking into account the vertical emission profiles.

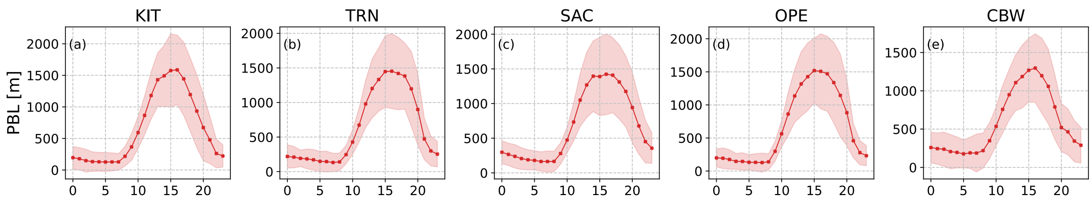

Figure A3. The diurnal cycles (local time) of planetary boundary layer height simulated by WRF-GHG at the KIT (a), the TRN (b), the SAC (c), the OPE (d) and the CBW (e) sites.

Across all sites, during the local time period of approximately 10:00 - 23:00, the biogenic contribution is nearly negative. This behavior is primarily attributed to the fact that daytime CO_2 uptake by vegetation through photosynthesis is substantially stronger than ecosystem respiration. As shown in the diurnal cycle of biogenic fluxes (Figure B), gross primary productivity (GPP) increases rapidly after sunrise (05:00 - 07:00) and acts as a strong negative flux. Its magnitude clearly exceeds that of respiration (RES) during the daytime, resulting in a net negative biogenic CO_2 flux, which indicates that terrestrial ecosystems act as a net sink of atmospheric CO_2. The period from sunrise to approximately 10:00 corresponds to the morning transition phase. During this period, photosynthesis gradually intensifies while respiration continues. Due to CO_2 accumulation caused by nighttime respiration under shallow PBL conditions, the biogenic contribution to near-surface CO_2 concentrations during this period remains positive. However, as photosynthesis strengthens, the overall biogenic contribution to near-surface CO_2 concentrations exhibits a decreasing trend.

Around 10:00 - 11:00, photosynthesis is strong at all sites, and the total biogenic flux reaches a strong CO_2 uptake (Figure B), and simultaneously, the PBL height starts rising rapidly (Figure A3). At this time, photosynthetic CO_2 uptake largely offsets the CO_2 accumulated from nighttime respiration. Subsequently, as the biogenic flux remains in a net uptake regime throughout the daytime, the biogenic contribution to near-surface CO_2 concentrations remains stably negative from daytime until approximately 20:00.

After 20:00, as photosynthesis weakens and ecosystem respiration becomes dominant again, the total biogenic flux gradually shifts to positive values. Meanwhile, the height of PBL decreases, facilitating the accumulation of respiration-derived CO_2 near the surface. Consequently, the biogenic contribution to near-surface CO_2 concentrations transitions from negative values toward zero and largely compensates for the daytime uptake by around 23:00.

We apologize if we may not have fully understood the reviewer's question. We infer that the last question may be: *Why are the biogenic contributions* ***at the CBW site*** *only positive around 10:00, whereas other sites show a different pattern?* For the CBW site, the biogenic contribution is only positive around 9:00-10:00, in contrast to the other sites where positive values most persist from 01:00 to 10:00. This difference may be attributed to the higher observation height at the CBW site (207 m above ground), whereas the observation heights at the other sites are around 50 m. At night, CO_2 released by respiration remains difficult to transport to the observation height of 207m because the nocturnal PBL is shallow and the atmospheric stratification is relatively stable (as shown in Figure A3(e), where the height of PBL at night is mostly below 200m). The positive value around 9:00-10:00 is likely caused by the onset of vertical mixing and the rise of PBL height, which transports CO_2 accumulated within the nocturnal boundary layer upward to the observation height. Before 9:00, the observation height is decoupled from the nocturnal boundary layer and is not influenced by nighttime accumulation and respiration, but rather by the remainder of the daytime uptake (residual layer), resulting in negative concentrations. After 10:00, as the PBL gradually deepens during the daytime and vertical mixing is enhanced, near-surface CO_2 is transported to higher altitudes. However, during this period, photosynthetic CO_2 uptake dominates, and thus the observed CO_2 concentration changes are predominantly negative.

11. Figure 7. It is better to give the slope and correlation coefficient in Figure 7c for the three TCCON sites.

We appreciate the reviewer's suggestion. We have revised the figure as follows.

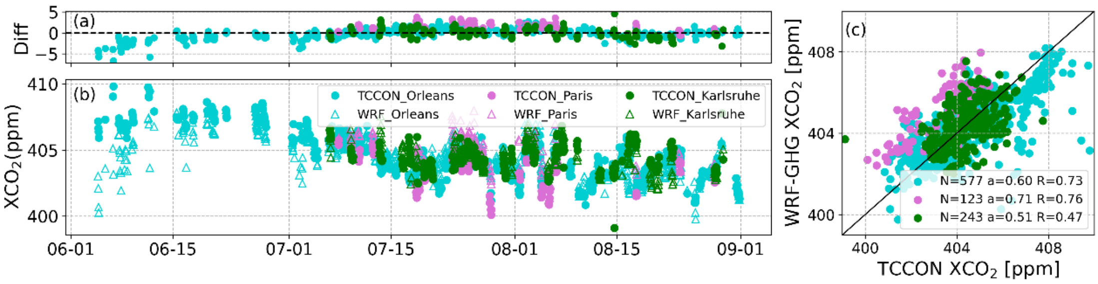

Figure 6. Time series (local time) of simulated XCO2 using TNO inventory and observed XCO2 at three TCCON sites (b), their absolute differences (WRF-GHG - TCCON) (a), and scatterplots (c). "N" represents the number of data pairs, "a" represents the slope of the linear regression curve and "R" represents the correlation coefficient.

12. Figure 8. Although the STD and RMSE increase from S to P, MBE becomes large. Which parameters are more critical to evaluate the model's sensitivity?

We thank the reviewer for this insightful comment. Indeed, at the KIT site, the MBE becomes larger when vertical emissions are considered (P) compared to the only surface emissions (S). However, the uncertainty associated with the TCCON XCO_2 measurements is approximately 0.5 ppm. The difference in MBE remains within the observational uncertainty.

Regarding the four statistical metrics (MBE, STD, RMSE, and R), we acknowledge that it's difficult to identify one as the most critical for evaluating model sensitivity. Each metric highlights different aspects of model performance, these metrics are generally interpreted together rather than individually. In this study, we find that except for MBE, the other three metrics exhibit slight improvements when vertical emissions are considered. Given that the change in MBE remains within the XCO_2 uncertainty range, we conclude that the overall improvement in XCO_2 simulation is present but relatively small. This is why we state in the manuscript that *"we observe a small improvement in the simulation of XCO_2, but not as notable as that observed for near-surface mole fractions."*

13. Line 378. What does SNAP mean in this paper?

SNAP stands for Selected Nomenclature for Air Pollutants. Following this classification, anthropogenic emissions are provided separately for different source categories. The sentence "*Anthropogenic sources are classified into 10 different categories according to the Selected Nomenclature for Air Pollutants (SNAP) in the study by Brunner et al. (2019) on vertical profiles.*" has been added to lines 169-170 of the revised manuscript.

14. Figure 11. There is a large gap between SNAP 1 for CAMS and EDGAR. Figure

10d also showed more emission sources than Figure 1a. Why did this gap occur? Were the emissions included in other emission sectors in CAMS?

The sectoral classifications of anthropogenic emissions differ between CAMS-REG-ANT and EDGAR. In CAMS-REG-ANT, anthropogenic emissions are categorized into 12 sectors (A - L) according to Gridded Nomenclature For Reporting (GNFR) categories, whereas EDGAR v2024 classifies anthropogenic emissions into 8 sectors.

In order to be able to apply vertical profiles to them, we mapped them to sectors classified according to the Selected Nomenclature for Air Pollutants (SNAP), which comprises a total of 10 sectors. However, due to the inconsistent criteria used for sectoral allocation in EDGAR and CAMS-REG-ANT, there may be some degree of cross-sectoral overlap between these datasets during the mapping process.

Besides, we also plotted the maps of the total annual anthropogenic emissions from CAMS-REG-ANT and EDGAR v2024 for 2018 (see Figure C), respectively. It can be seen that the annual emissions provided by EDGAR are significantly higher than those from CAMS.

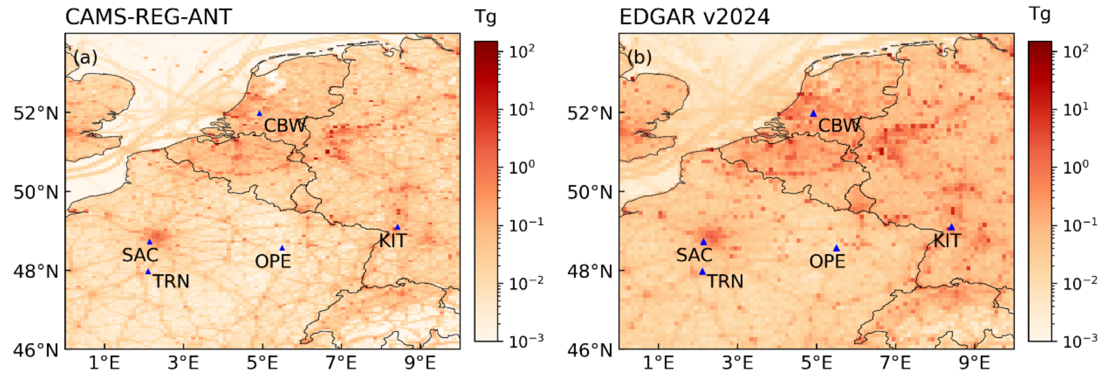

Figure C. Map of total anthropogenic emissions from CAMS-REG-ANT and EDGAR v2024 in 2018.

15. Figure 12. It seems that in late July and August, the land system was also active as a carbon source (Figure 12c), but anthropogenic emissions nearly disappeared from Figure 12b. Usually, drought can increase temperatures and the electricity demand for air conditioning, hence the anthropogenic emissions could increase in this period.

We thank the reviewer for this comment. Indeed, there is a carbon source near the surface in July and August for the biogenic tracer which can be seen in Fig. 12c (Figure 11 in revised manuscript), but this feature is less pronounced in Fig. 12b since the latter represents the column-averaged dry-air mole fractions of CO_2. As shown in Fig. 12c, the contribution of biogenic CO_2 decreases with increasing altitude and becomes negligible above approximately 2000 m.

And indeed, drought could increase electricity demand, but air conditioning remains relatively uncommon in Europe, especially in the countries involved in our study. In addition, Figure D shows the time series of monthly anthropogenic emissions from EDGAR v2024 obtained from ECCAD at the Paris (a), Karlsruhe (b), and Orléans (c) sites, respectively. The shaded area represents emissions during June - August of each year. Compared with non-drought years (e.g., 2016 and 2017), anthropogenic emissions in summer 2018 do not exhibit a noticeable increase and are even slightly lower than in those years. Based on these two considerations, the lack of a noticeable increase in simulated anthropogenic emissions during this period is reasonable.

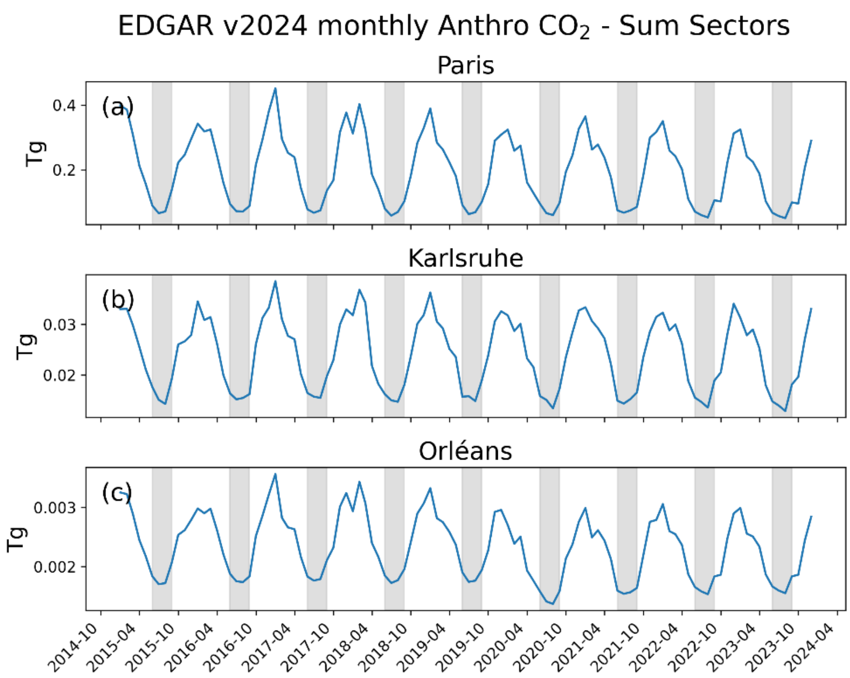

Figure D. Time series of monthly anthropogenic emissions from EDGAR v2024 at the Paris (a), Karlsruhe (b), and Orléans (c) sites.

16. Lines 438-439, please add ° before N and E for the GPS location. Also, add this to the GPS location in Table A3. What does "acid fen" mean here for FR_LGt?

We appreciate the reviewer's suggestion. We have added ° at the corresponding locations in the manuscript.

Here the "acid fen" refers to a type of wetland characterized by acidic conditions, typically with low pH levels. The ICOS description of this site is as follows: "*The La Guette station (FR-LGt) is a peatland located in Neuvy sur Barangeon (Sologne) at about 200 Km south of Paris and 80 Km south of the Université d''Orléans. It is an acid fen that is cut at the output by a road.*"