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.

Wang et al. simulated CO_2 mole fractions over Western Europe in the summer of 2018 using the WRF-Chem model combined with three different CO_2 emission inventories. The simulations were evaluated by comparing with ground-based in situ and column observations. They showed, by taking into account the sector-specific vertical profiles of emissions, the agreement between the simulations and the observations was significantly improved for sites near large emission sources.

The topic of this manuscript is important and relevant to the scope of Atmospheric Chemistry and Physics. In addition, the analysis method is appropriate, and the writing structure is well organized. However, it seems to lack novelty and contains few new scientific findings. I recommend clarifying what is novel and addressing the following concerns and questions.

We thank the reviewer for the constructive comments. We summarize the main novel aspects of this study as the following two points, both of which have also been incorporated into the revised conclusion:

- This study provides the first high-resolution WRF-Chem simulation over Belgium and its surrounding countries, and evaluates the model performance using observations from multiple ICOS and TCCON sites. The results provide a systematic assessment of the applicability of the WRF-GHG model at regional scale in the core region of Western Europe, offering a solid basis for related modeling and applied studies.
- This study highlights the necessity of considering the vertical distribution of anthropogenic emissions of CO_2 in simulations, especially near strong emission sources.

Specific comments

1. L56: Please provide a clear description of whether the signal is positive or negative.

We appreciate the reviewer's suggestion.

The original sentence has been revised to: *Using the ICOS atmospheric measurements,* *Ramonet et al. (2020) reported that a severe drought event in Europe in 2018 led to an atmospheric CO_2 signal of +1 to +2 ppm at most stations* ***in summer****.*

2. L72: Do the multi-source observational data refer to ICOS and TCCON data? If so, I do not think they were used for simulating CO_2 mole fractions.

Yes, the multi-source observational data refer to ICOS and TCCON data. We added "and evaluate", and the sentence has been revised as: "*It employs the Weather Research and Forecasting Greenhouse Gas model (WRF-GHG; Beck et al., 2011) and multi-source observational data to simulate* ***and evaluate*** *CO_2 mole fractions over Western Europe, with a focus on Belgium and surrounding regions during summer 2018.*"

3. L242–247: Please add a discussion of how the overestimation in inland areas and the underestimation in coastal regions affected the mean bias error in wind speed.

We appreciate the reviewer's suggestion. We have added some descriptions detailing the mean bias error (MBE) for inland and coastal stations. The corresponding sentence has been revised to: "*Besides, as shown in Fig. A1, the model tends to underestimate wind speed values significantly along coastal regions, with an MBE of -1.11 m/s across the four coastal stations, which is probably due to the coastal effects in the WRF simulation (Hahmann et al., 2015). In inland areas, the model tends to overestimate at most sites with dense vegetation cover. Such overestimation of wind speed has also been found in previous studies (Duan et al., 2018; Liu et al., 2022; Che et al., 2024). This bias may be attributed to the complex wind distribution in areas with rugged terrain, where the WRF model fails to adequately account for the additional resistance effects of vegetation on unresolved terrain, ultimately leading to an overestimation of wind speed. While an underestimation of wind speed is also observed at some inland stations, most of these stations are located at airports or in areas dominated by cropland. The land-use classification of the corresponding model grid cells is mostly identified as urban and built-up, which may lead to an overestimation of the surface roughness near the station and consequently an underestimation of wind speed. Similar findings were reported by Aylas et al. (2020), who found that WRF underestimated wind speed at airport stations by approximately -0.124 m/s, while the bias was reduced to -0.079 m/s after updating the land-use data.*" in lines 243-254 of the revised manuscript.

4. L253: It would be easier to read if the expressions "between the observations and the simulations" and "between the simulations and the observations" were made consistent throughout this manuscript.

We appreciate the reviewer's suggestion. All related instances in the manuscript have been uniformly revised to "*between simulations and observations*".

5. Figures 8 and 9: The differences between the WRF-GHG simulations and TCCON data in Figure 9 appear to be larger than those in Figure 8. Are the differences due to the fact that Figure 9 does not take into account smoothing using the column averaging kernel?

Yes. In Figure 8 (i.e., Figure 7 in the revised manuscript), we use the simulated results after AVK smoothing. In Figure 9 (i.e., Figure 8 in the revised manuscript), since we can't apply smoothing to the background field and the individual tracers, all simulated results shown in Figure 9 are therefore presented without AVK smoothing in order to maintain consistency. The difference of XCO_2 before and after applying the AVK smoothing is around 1.08 ± 0.44 ppm (MBE ± STD) at the Orléans site, 1.18 ± 0.34 ppm at the Paris site, and 1.10 ± 0.28 ppm at the Karlsruhe site, respectively.

6. 338–347: If the data period for Orleans was matched with that for Paris, would similar results be obtained? In other words, would the difference between Paris (an urban area) and Orleans (a suburban area) be reflected in the observed XCO_2?

We thank the reviewer for the constructive comment.
We selected the data points where the observation times at the Orléans and Paris sites overlap, resulting in a sample number of 94. The statistical metrics between the simulated and observed XCO_2 at the Paris and Orléans site are given in the Table A, and the time series of observed and simulated XCO_2 using TNO inventory at Paris (magenta) and Orléans (teal) TCCON sites are shown in Figure A.
It can be found that, for this selected time series, compared to the large differences caused by using different emission inventories at the Paris site (up to approximately 0.49 ± 0.33 ppm between TNO and EDGAR), differences are also present at the Orléans site but are smaller (up to approximately 0.1 ± 0.08 ppm between TNO and CAMS). The simulated XCO_2 at the Orléans site (a suburban area) shows a systematic overestimation (around 0.4 - 0.5 ppm), but its magnitude is obviously lower than that at the Paris site (an urban area, around 1.1 - 1.6 ppm). As the anthropogenic emissions around Orléans are relatively low (Figure 9 in revised manuscript), it indirectly suggests that, in addition to uncertainties in the background fields, the overestimation at the Paris site is to a large extent caused by uncertainties in the anthropogenic emissions.

Table A. The statistical metrics between the simulated XCO_2 using five different anthropogenic emission inputs and observed XCO_2 at Paris and Orléans sites. Here, the time series are selected based on the period overlapping between the Paris and Orléans site, resulting in N=94.

Paris

	EDGAR_S	*EDGAR_P*	*CAMS_S*	*CAMS_P*	*TNO_CAMS*
MBE	1.09	1.08	1.18	1.16	1.58
STD	0.84	0.84	0.87	0.86	0.96
RMSE	1.38	1.37	1.47	1.45	1.85
R	0.75	0.75	0.74	0.74	0.70
			Orléans		
MBE	0.43	0.42	0.41	0.41	0.52
STD	0.76	0.76	0.77	0.77	0.77
RMSE	0.87	0.87	0.88	0.87	0.93
R	0.75	0.75	0.74	0.74	0.74

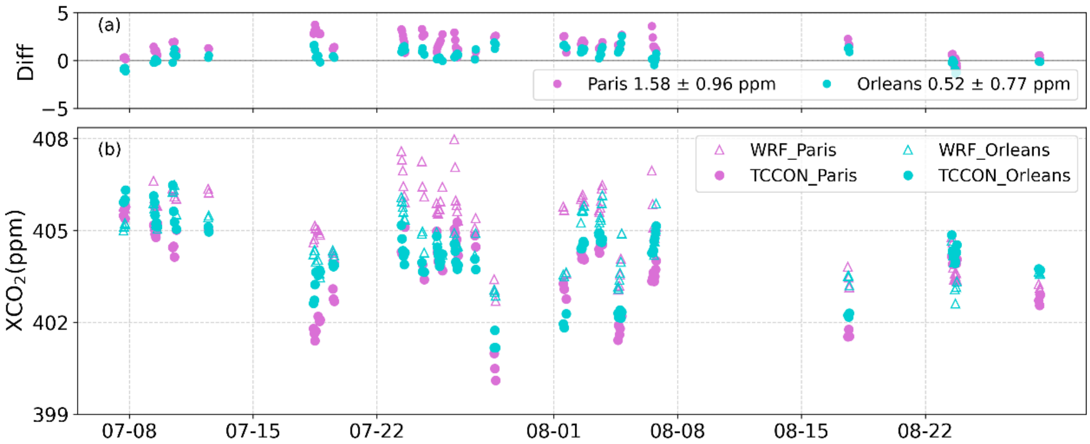

Figure A. Time series (local time) of observed and simulated XCO_2 using TNO inventory at Paris and Orléans TCCON sites (b), and their absolute differences (WRF-GHG - TCCON) (a).

7. L378: Please add an explanation of SNAP.

We appreciate the reviewer's suggestion. 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.

8. L382–383: In Figure 10, it appears that there are also large emission sources near sites other than KIT (i.e., CBW and SAC). What degree of closeness does "near" represent?

Yes, in addition to the KIT site, there are also strong emission sources near the CBW and SAC sites. Figure B shows the simulated near surface CO_2 mole fractions at the lowest model level at each site, which is located at approximately 25 m above the ground. It can be found that at the TRN and OPE sites, the five simulated results are almost identical, which is consistent with the relatively weak anthropogenic emissions in their vicinity. At the SAC and CBW sites, whether vertical emissions are taken into account indeed has an impact on the simulation results. Because the emission sources near the SAC site are comparatively weaker, the differences among the results there are also smaller. However, the sensitivity tests at the observation heights of SAC and CBW do not show significant biases. For SAC, this may be because the emission sources are not as strong. For CBW, this may be because the observation height used is relatively high (207 m above the ground). As shown by the results at different heights at the KIT site, the higher the observation height, the smaller the difference between simulations that consider only surface emissions and those that include vertical emissions.

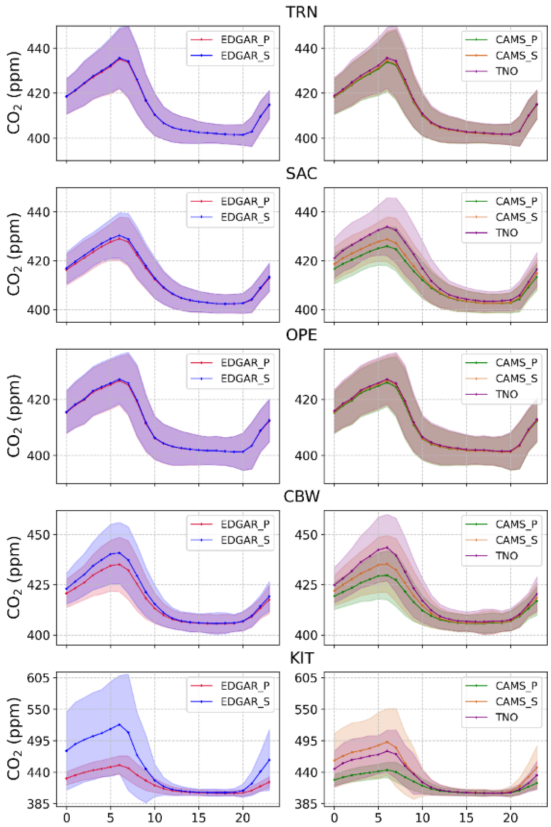

Figure B. Diurnal cycles (local time) of simulations with different anthropogenic emissions at five ICOS sites at the model lowest layer (approximately 25 m above the ground).

For the degree of closeness represented by "near," since the influence of emission sources on a specific site is jointly affected by multiple factors, including wind direction, wind speed, emission source strength, and the distance to the site, we are sorry that we cannot provide a precise quantitative value.

We have added "*as well as near the CBW and SAC sites, whereas there are almost no emission sources in the vicinity of the TRN and OPE sites. Figure A4 shows the diurnal cycles of the five near-surface CO_2 mole fractions at each ICOS site as simulated by WRF-GHG fractions at the lowest model level (approximately 25 m above ground level). It can be found that at the TRN and OPE sites, the five simulations show nearly identical patterns, consistent with the relatively weak anthropogenic emissions in their* *surrounding areas. In contrast, at the SAC and CBW sites, whether vertical emissions are taken into account does indeed affect the simulations. However, at the observation heights of these two sites, the impact is small and does not exhibit characteristics similar to those observed at the KIT site. For the SAC site, this may be due to the relatively weaker emission sources nearby, whereas for the CBW site, this may be related to the relatively high observation height used (207 m above ground level), as shown for the KIT site (Fig. 4), the discrepancies between sensitivity tests decrease with increasing observation height*" to lines 391-400 of the revised manuscript. In addition, Figure B has been incorporated into the revised manuscript as Figure A4.

9. L419: including -> includes

Done.

10. L431–434: The simulated XCO_2 values in July and August 2018 were higher than the observed XCO_2. What caused this overestimation by the model? In years other than 2018 when no drought occurred, will the simulated XCO_2 values be lower than the observed values?

This overestimation may be due to an overestimation of anthropogenic emissions, or an overestimation of biogenic fluxes, or a combination of both effects.

As our present study only focuses on simulations conducted for the period from June to August 2018, we are unfortunately unable to provide simulation results outside this time window. Therefore, we regret that we cannot currently elaborate on how the simulated XCO_2 would behave in years without drought conditions. Nevertheless, we appreciate the reviewer's insightful comment.

11. L454: analyze -> analyzing

Done.

12. L486–488: Please revise the sentence by adding a conjunction.

We appreciate the reviewer's suggestion. We have added the conjunction "as" to the sentence. The revised sentence is as follows: *At the same time, observational data play a critical role in model evaluation, **as** the availability of time-synchronized ground-based observations of CO_2 fluxes and concentrations can help assess the performance of the model, thereby enhancing the credibility and scientific interpretation of the simulation results.*

Reference

Aylas, Y.G.R., De Souza Campos Correa, W., Santiago, A.M., Reis Junior, N.C., Albuquerque, T.T.A., Santos, J.M., Moreira, D.M.: Influence of land use on the performance of the WRF model in a humid tropical climate, Theor. Appl. Climatol., 141, 201–214, https://doi.org/10.1007/s00704-020-03187-3, 2020.