

Alberti et al. in this manuscript titled “Monitoring and modeling seasonally varying anthropogenic and biogenic CO₂ over a large tropical metropolitan area” attempted to investigate the atmospheric CO₂ dynamics in the Metropolitan Area of São Paulo, Brazil with using WRF-GHG/WRF-Chem model output after performing model validation with measurements of two sites (IAG and PDJ) located in the study domain. Although this is an important study for this region, there are a lot of concerns that appeared in the present manuscript and failed to demonstrate the scientific novelty with the model output. For instance, there is a methodological issue in the model setup (nested domain, chem IC/BCs, emissions, etc.) and I guess large uncertainties in the model output compared to the measured concentration are associated with this methodological issue (authors also acknowledged this issue but didn’t put any attempt to overcome this issue). Also model output is not well analyzed to support the major conclusions made in this study. Therefore, it needs more work before accepting this manuscript for the ACP.

Major issues:

1) Model setup

Since authors in this study attempted to perform WRF-Chem/GHG simulation in a small domain particularly in urban settings, they didn’t configure 3km resolution from a nested domain. If you think configuring nested domains is computationally expensive, you might perform a shorter period (Feb and August only). Nested domain configuration will allow you to perform model evaluation in two different resolutions. In Line#259: How did you establish this statement that model resolution is actually causing the mismatch between simulated and observed concentration at the PDJ site? Chemical IC/BCs should be well representative. If you think chem Carbon Tracker based IC/BCs are good enough for your model setup it needs well justification, but as you mentioned in line# 381-382 why you didn’t give a try to see what benefit you will achieve if you use other chem IC/BCs such as CAMS which has more finer resolution than CT? It needs more detailed information for the VEIN and EDGAR products while creating anthropogenic emissions. What is the spatial and temporal resolution and year of representation? Did you combine and sum both VEIN and EDGAR products to get total anthropogenic emissions? At this moment it’s not clear how these two inventories are configured in the model. You can also provide a map showing the differences of VEIN and EDGAR products. Are there any emission scaling factors being implemented in the inventory for diurnal cycle? It’s not clear how diurnal variation is imposed in the model. You might need to impose a temporal anthropogenic emission variation following Nassar et al. (2013). Also look into the following paper where they imposed temporal emission factor

Callewaert, S., Brioude, J., Langerock, B., Dufлот, V., Fonteyn, D., Müller, J.-F., Metzger, J.-M., Hermans, C., Kumps, N., Ramonet, M., Lopez, M., Mahieu, E., and De Mazière, M.: Analysis of CO₂, CH₄, and CO surface and column concentrations observed at Réunion Island by assessing WRF-Chem simulations, *Atmos. Chem. Phys.*, 22, 7763–7792, <https://doi.org/10.5194/acp-22-7763-2022>, 2022. I saw you referring to Gourджи et al. (2022) in a couple of areas but I’m surprised why you didn’t use their approach to optimized your VPRM model. In section 2.2.2

(CO₂ fluxes data) you should provide more detailed information about the Eddy flux sites used in the VPRM optimization process (perhaps in the supplementary or appendix).

2) Analyzing model outputs

Meteorological output such as wind speed and direction should be presented with a wind rose figure. Wind rose figures will help to interpret the sources of CO₂ in a site. In many places you are talking about atmospheric stability but there is no analysis to support atmospheric stability-based conclusions (such as section 3.3.1). You can perform PBL height related CO₂ profiles at two sites (IAG and PDJ). I'm not convinced with the conclusion made in this paragraph (line#345 to 354 – source contribution), because: Not sure the hour format in Fig. 8. Is this UTC or local hour? If midday has less vehicles how do you expect late night to have high vehicle emissions and cause high correlation between CO₂ and CO? Again, it's important to provide some evidence of the diurnal cycle in the VEIN or EDGAR product. In line#236: Do you have any information on the seasonal variation of traffic emission? Seasonal wind rose plot will help you to establish the wind speed related statement written in this line. Abstract is not well written and there is only one conclusion in the abstract without any quantitative evidence.

3) Figures

Most of the figures in the current manuscript are not publication standard. Figure captions are not well described. For instance, four panels in Fig. 1 but not clear which one is for what; no panel number for WS and WD panels in Fig. 2. Please elaborate captions in all figures and make clear for all legends.

Minor comments:

I'm confused with line#86 and #139 because in #86 it is indicated meteorological variables will come from wrf outputs but in line#139 it is written VPRM model is driven by the meteorological measurements of the sites. Please make it clear.

How did you decide 5 days would be good enough for spin up? Please give a justification.

Table 2, better to write full name of the variables (T2, Wd, Ws etc.) in the table footnote. I think it's more relevant to classify PDJ as Sub-urban park and IAG as University campus/urban park.

Section 2.2.2 (CO₂ fluxes data) should be merged with section 2.1.2 or vice versa.

NEE should have a standard unit system (PgC/yr or TgC/yr) and be consistent throughout the manuscript. Currently, two different units are introduced for NEE in Fig. 3 and Fig. 4.

Line#221: Is this statement correct? I see barely green colored in the August map (Fig. 4b).

Line:#223: only vegetation? Line#227 to line#230: Please use quantitative comparison between your study and Raju et al. (2023). Therefore, I suggested following the methodology of Gourdji et al. (2022) while using Eddy flux data for optimizing VPRM.

Be consistent with Pico do Jaragua and PDJ.

Spatial map: Fig 6: Adding a third row with the differences between first and 2nd row would be great.

Line 276: coastal region southwest of MASP -> southeast??

Line 312: how do you confirm this is related to vehicular emission?

Line 317-318: what do you mean by profile? Did you perform any altitude related comparison? Also, I don't see the simulated concentrations are consistent with the observed data at Fig. 7c.

Line# 149: Osterman et al. -> year??

Line# 151: I think satellite measurement time was also matched with the model data during interpolation. Please mention this information. But, if measurement time was not matched please do that.

Line#177:178: Acronym should be consistent throughout the manuscript. For instance, TM, WS and WD in these lines are different than Table 2.