

**Answers to Referee #1 comments: *Review of Impact of physical parametrizations and initial conditions on simulated atmospheric transport and CO<sub>2</sub> mole fractions in the US Midwest***

We thank the referee for the helpful comments that will improve the manuscript. In the text below, we have tried our best to respond to all the general and specific comments provided by the reviewer.

**Comments to Author:**

Diaz-Isaac et al. studied the impact of transport errors on simulated CO<sub>2</sub> mole fractions in the US Midwest, which is very relevant to the goal of improving the estimate of surface CO<sub>2</sub> fluxes, since transport models are used to derive surface CO<sub>2</sub> fluxes in an inverse analysis. The authors tested a series of physical parameterizations and initial conditions and pointed out that most tested physical parameters and initial conditions have a significant impact on simulated CO<sub>2</sub>, either influencing the planetary boundary layer height (PBLH) that confirms the previous finding that the correct representation of PBLH is important for accurate CO<sub>2</sub> simulations or changing wind speed and direction.

The paper is well structured and clearly written. The reviewer suggests publication after the following concerns have been addressed.

**REF-C1:** One of the major conclusions that “all physics parameterization except for microphysics have a significant impact on both CO<sub>2</sub> mole fractions and meteorological variables” is based on the magnitude of the simulated CO<sub>2</sub> root mean square difference (RMSD). The authors mentioned that it was computed for each model ensemble member by varying only the type of physics parameterization. My understanding is that for the LSM scheme, multiple sets of ensemble members can be used for the computation, e.g. models nos. 1&7, nos. 2&8, nos. 3&9, nos. 22&40&43, nos.23&41&44, nos.24&42&45. Was the presented RMSD for LSM the mean of all different sets of ensemble members? For the calculation of the mean of the ensemble for day *i* in equation 1, are all 45 ensembles used or only the sets with varying one type of physics parameterization? This should be clarified.

**Author-C1:** Yes, to compute the RMSD we create multiple set of ensemble members and the result that we present is the mean of that different sets of ensemble members. This same process is used for the other physics parameterizations (i.e., PBL schemes, Cumulus, Microphysics) and Reanalysis data set. RMSD presented not only for the LSM, but for PBL schemes, Cumulus, Microphysics and Reanalysis was the mean of all the different sets of the ensemble members that we were able to generate. For the calculation of the mean of the ensemble for day *i* in equation 1, we only use the set of ensemble members for the different varying type of physics parameterization.

We clarify this in the methods Sections 2.9.a, as follow (see ***bold italic***):

P9/L24-L26: The RMSD was estimated for the different physics parameterization used (i.e., LSM, PBL schemes, CP, MP) and reanalysis. ***A different set of ensembles were created for each of the physics parameterization, where the model configuration remained identical except for the tested physics parameterization and the different set of members were to compute the ensemble mean.***

**REF-C2:** I agree with the concerns raised by reviewer #2 on the use of the 14 radiosounding sites. The vertical profiles of temperature and CO<sub>2</sub> mole fractions at multiple sites from the NOAA aircraft program and/or from other intensive campaigns could be looked into, at least for the PBLH.

**Author-C2:** We considered adding more observation to our evaluation but decided not to for several reasons. The radiosonde data provides the best spatial and temporal resolution that we can have for our project. Other data such as CO<sub>2</sub> mole fractions from the NOAA aircraft program (or other aircraft campaigns) will bring the limitation of the time and/or spatial coverage. Only five sites are within our WRF simulation domain and profiles are usually collected every two weeks (i.e. about 2 per site over our simulation period). Most aircraft campaigns do not sample the mixing depth but rather collect long transects over the continent (e.g. COBRA, ATom). In addition, we have no available intensive campaign in that part of the country and for the simulation period we selected. In the near future, the Atmospheric Carbon & Transport (ACT) project funded by NASA, that is currently performing multiple fields campaigns over the East and Midwest of United States for different seasons, will provide a significant spatial coverage to address transport model errors. The campaign has just finished its fourth deployment in Spring of 2018. Regarding surface stations, including them in our analysis would bias our results towards near-surface model errors, which we know are not representative of the whole Planetary Boundary Layer (PBL) (e.g. Hu et al., 2010; Rogers et al., 2013; Deng et al., 2017). CO<sub>2</sub> molecules are mixed over the entire PBL during daytime. Therefore, adding tens to hundreds of surface stations will not represent the actual model errors in the PBL. For these different reasons, we decided to focus on operational radiosondes launched at 00z, using mid-PBL wind measurements. Because we are aware of this limitation in our study and this is a concern from both reviewers we decided to add this to the last paragraph of our Discussion section, where we introduce the different limitation of this study (see ***bold/italic*** lines in the next paragraph).

P18/L1-L3: ***“We also note that models were compared only to rawinsonde data, the only type of observation that had both the temporal and vertical resolution needed to evaluate the models within the PBL. More observations with higher temporal, spatial and vertical resolution will be an asset for future evaluation of transport models, focusing on intensive campaigns over multiple seasons.”***

**REF-C3:** P5/L6: use longwave instead of long wave

**Author-C3:** Done

***“The members in our multi-physics ensemble all use the same radiation schemes (both longwave and shortwave) but the land surface, surface layer, boundary layer, cumulus, and microphysics schemes are varied for both the inner and the outer domain.”***

**REF-C4:** P12/L22, P13/L5, P15/L21: use the plural form “show” instead of “shows”

**Author-C4:** Done

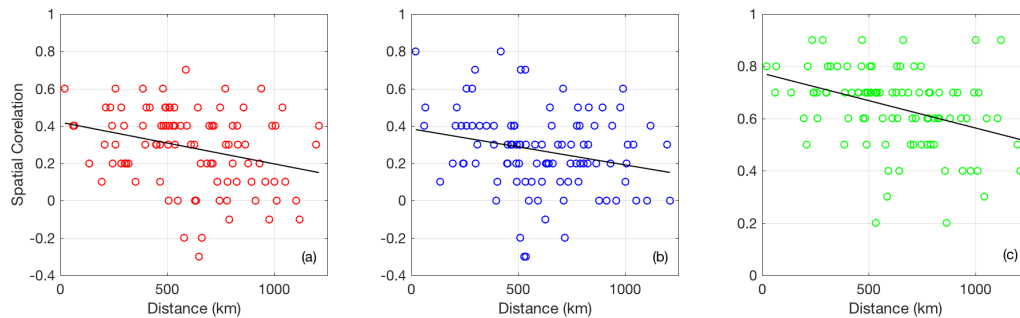
P12/L22: ***“Although the configurations that show the highest RMSE are not always the same across the different variables, these configurations share the same LSM (RUC).”***

P13/L5: ***“Similar to the regional RMSE (Figure 5), both LBF and MPX show that the LSM RUC leads to the highest RMSE for the three meteorological variables.”***

P15/L21: “Both wind speed (Figure 13a) and wind direction (Figure 13b) show low correlations, whereas PBLH (Figure 13c) shows consistently positive correlation with the CO<sub>2</sub> mole fraction errors across all sites.”

**REF-C5:** L15/22: “We did not find any relationship between error correlation and distance”. It would be convincing to show a scatter plot, error correlation vs. distance for each grid in Figure 13, or at least present the correlation of error correlation vs. distance.

**Author-C5:** Thanks for your suggestion. To address your comment, we decided to follow your suggestion and add a scatter plot with the error correlation vs. distance. We added the following figure (i.e., scatter plot) to the supplement section.



**Figure S1.** Tower and rawinsonde sites specific spatial correlation coefficient between ensemble mean MBE of (a) wind speed, (b) wind direction and (c) PBLH and ensemble mean MBE of DDA CO<sub>2</sub> mole fractions versus their distance. The abscissa shows the distance between the rawinsonde and tower sites, while the ordinate shows the spatial correlation.

The following was added to the line cited in the referee comment (see ***bold/italic*** line in the next paragraph):

P15/L19: “Both wind speed (Figure 13a) and wind direction (Figure 13b) shows low correlations, whereas PBLH (Figure 13c) show consistently positive correlation with the CO<sub>2</sub> mole fraction errors across all sites. ***We did not find any relationship between error correlation and distance (see Figure S1).***”

**REF-C6:** L16/11-12: provide evidence of PBL winds impacting the distribution and magnitude of the inverse CO<sub>2</sub> fluxes over the region.

**Author-C6:** In this paper we do not show the impact of PBL wind explicitly on the distribution and magnitude of the CO<sub>2</sub> fluxes over the region. However, studies like Lauvaux and Davis (2014) and Deng et al., (2017) show some of this impact. For example, Deng et al. (2017) show the impact that a Four Dimensional Data Assimilation (FDDA) technique has on the different atmospheric variables (e.g., wind speed, temperature) used as input in the atmospheric inversion. Figure 9 from Deng et al. (2017) shows how the influence function vary based on the type of observations assimilated into the model. This figure shows, the influence function when no observations are assimilated (a), only surface stations are assimilated (b), both surface stations and lidar winds are assimilated (c), surface stations, lidar and commercial aircraft data is also assimilated.

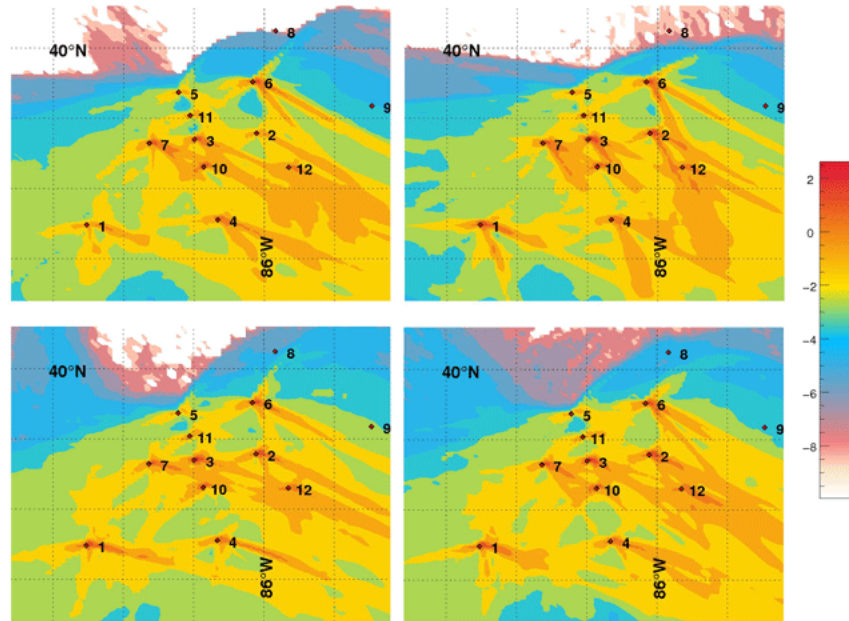


Figure 9. “Influence functions over the INFLUX 1-km resolution domain for 10 of the 12 tower locations of the INFLUX network using the Lagrangian Particle Dispersion Model, averaged for 26–30 October 2013 (corresponding to the observation time 17–22UTC) driven by the meteorological variables from the four different WRF configurations: NOFDFA (Upper left), FDDA\_WMO (Upper right), FDDA\_WMO\_Lidar (Lower left), and FDDA\_WMO\_Lidar\_ACARS (Lower right), in log scale (ppm hour m<sup>2</sup> g<sup>-1</sup>). Note that numbers 1–12 indicate the tower locations as detailed in Figure 8 and two towers were not operational during the time period Oct 26–30. DOI: <https://doi.org/10.1525/elementa.133.f9>” Note: This caption was copied from Deng et al., (2017) article.

To this line cited by the referee we added the following citations as an example of this (see ***bold/italic*** line):

P16/L7-L8: “The relationship between PBL winds and CO<sub>2</sub> mole fraction is dependent on the local spatial distribution of CO<sub>2</sub> surface fluxes and could easily show no clear correlation when averaged over time and space. ***However, we know errors in these two variables can impact the distribution and magnitude of the inverse CO<sub>2</sub> fluxes over the region (Deng et al., 2017, Lauvaux and Davis 2014).***”

New citation: Deng, Aijun, Thomas Lauvaux, Kenneth J. Davis, Brian J. Gaudet, Natasha Miles, Scott J. Richardson, Kai Wu et al. "Toward reduced transport errors in a high resolution urban CO<sub>2</sub> inversion system." *Elem Sci Anth* 5 (2017).

**REF-C7:** The reviewer found it difficult to obtain any meaningful information from the figures 7,10&14 where the results of all 45 model results are presented for three selected sites. The results should be first summarized before being presented, or simply be moved to the supplementary.

***Author-C7:*** Thanks for the comment, we add a summary of the results prior explaining what we found at each figure. Nest paragraphs shows what we added for the different figures as a summary in ***bold/italic***.

***Author-C7.1:*** Figure 7:

P12/L28-L30: “Figure 7 shows the monthly average RMSE of wind speed (Figure 7a-c), wind direction (Figure 7d-f), and PBLH (Figure 7g-i) for each model configuration at specific rawinsonde sites. ***We computed the RMSE for all the different sites (not shown) and we found***

*the highest RMSE in the model configurations that included RUC and Thermal diffusion as the LSM and at some sites when these LSMs are combined with YSU as a PBL scheme.”*

*Author-C7.2:* Figure 10:

*P14/L5-L8: “The MBE analysis was performed for all the sites (not shown), for this statistic we found that all the model configurations show a positive wind speed MBE (overestimation) for the majority of the rawinsonde sites, whereas, wind direction and PBLH shows both positive and negative (underestimation) MBE across the different configuration at the different rawinsonde sites. Some of the positive and negative biases are associated to specific LSMs and PBL schemes. Figure 10 show the MBE of three sites that are representative of regional patterns.”*

*Author-C7.3:* Figure 14

*P17/L31-L34: “The sensible heat flux was averaged from 1200 to 2300 UTC and we computed the MBE of the sensible heat flux for the eddy covariance stations close to the rawinsonde sites. The MBE was estimated at all the eddy covariance stations available over the region (not shown) and we found that the highest positive sensible heat MBE were found on simulations that used YSU as PBL scheme and RUC or Thermal Diffusion as LSM. Figure 14, shows PBLH MBE of two rawinsonde sites (Figure 14a,b) and the sensible heat MBE of two eddy covariance stations (Figure 14c,d) close to each of these rawinsonde sites.”*