Response to Reviewer 1

We are very grateful that reviewer 1 took the time to carefully evaluate our submitted manuscript, and provide us with a valuable opportunity to reinforce the results of our analysis. We have worked to address all of the comments from reviewer 1, either in this response document or through revisions to the paper. In our efforts to address these comments, we believe that the revised manuscript is an improvement on the previous version.

1 General Comments

1. **Comment:** Many of the seasonal cycle fits shown in Fig. S4-S11 appear to be quite unphysical. Thus, it is unclear if the analysis is really capturing accurate SCA estimates. There should be uncertainty quantification in the SCA fits, perhaps using bootstrap resampling or another technique. Ideally, the analysis could also be performed fitting truncated Fourier series, to test the impact of the functional form on the results.

   Reply: While we are aware that fitting seasonal cycles to a truncated Fourier series has been established in previous literature, fits of this type do not have any more physical basis than the fits that we use here, and through extensive testing the fitting methods from Lindqvist et al. (2015) were found to be a better choice for this analysis. As stated in the cited literature (Lindqvist et al., 2015), these methods of seasonal cycle fitting offer improved results relative to fitting to a truncated Fourier series in situations where there are substantial gaps in the observed time-series. In particular, high latitude observations consistently have gaps in the winter-time. To get a reasonable fit to a truncated Fourier series we would need to artificially fill winter-time data, which is something we specifically want to avoid as it would detract from our data-driven approach. As discussed in our alternate reply to comment 2, below, we have added a new section to the supplement that considers how removing winter-time data from CAMS time-series effects the resulting SCA (using the methods from Lindqvist et al. (2015)). We do concede that there are a few zones in the Asian Tundra region that have seemingly unphysical results, and this was discussed in section 3.2 of the manuscript. We have now also changed the map layout of the Asian Tundra zones and removed the two northern-most zones in the Asian Tundra region because these zones have the largest winter gaps of any zones. The removal of these Asian Tundra zones does not seem to have changed any of the conclusions of this analysis as it pertains to the spatial distributions of X_{CO_2} seasonal cycle parameters or the roles of fluxes and atmospheric transport patterns. The Lindqvist et al. (2015) fitting methods yield fits to model-estimates that follow the data very closely, and it is our opinion that most of the seasonal fits to OCO-2 observations appear quite reasonable. This assessment of observed seasonal cycle fits is reinforced by the close agreement between spatial distributions of observed and model-derived SCA. We did evaluate some seasonal cycle fits to a truncated Fourier series with a linear trend, following methods used in Wunch et al. (2013) (see Fig. R1-1), but found these fits to be far more unrealistic than the fits to a skewed sine wave presented by Lindqvist et al. (2015) (see Fig. R1-2). Even for the seasonal cycles with no data gap in the winter, there are winter-time oscillations in the fit to the truncated Fourier series that do not seem to accurately represent variability in the data, and for sites with winter data gaps the fits to the truncated Fourier series are clearly unrealistic.

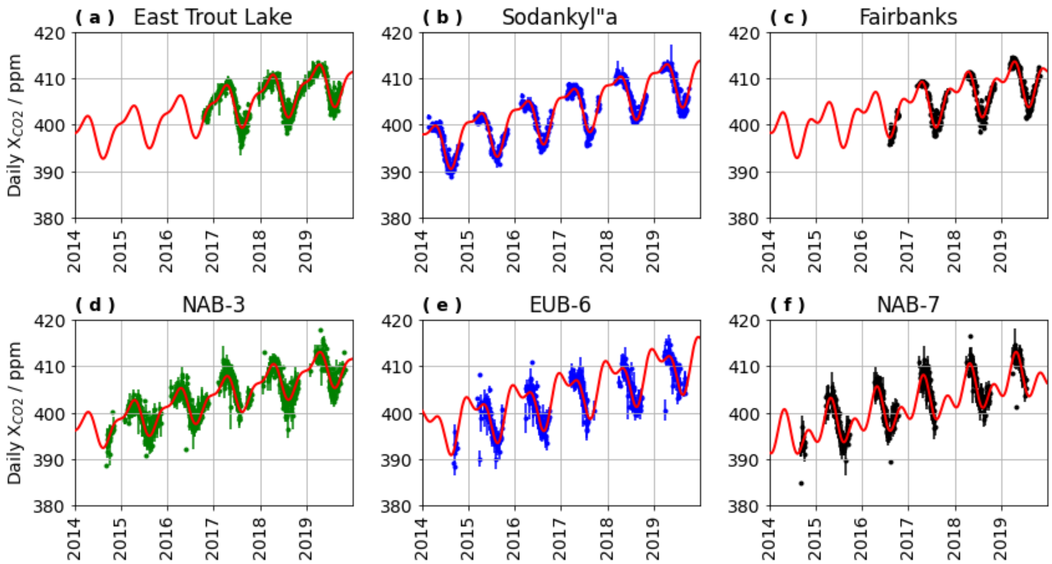

Figure R1-1. Time-series of ground-based observations (panels (a), (b), and (c)) and satellite-based observations for the zone containing the ground site (panels (d), (e), and (f)) at two TCCON sites East Trout Lake, Canada and Sodankylä, Finland, as well as Fairbanks Alaska, where EM27/SUN observations are collected. Also shown are seasonal cycle fits for these observations to a trended truncated Fourier series of the form: $f(t) = a_0 + a_1 t + \sum_{k=1}^{2} b_k \sin\left(2\pi k \frac{t}{365.25}\right) + c_k \cos\left(2\pi k \frac{t}{365.25}\right)$.

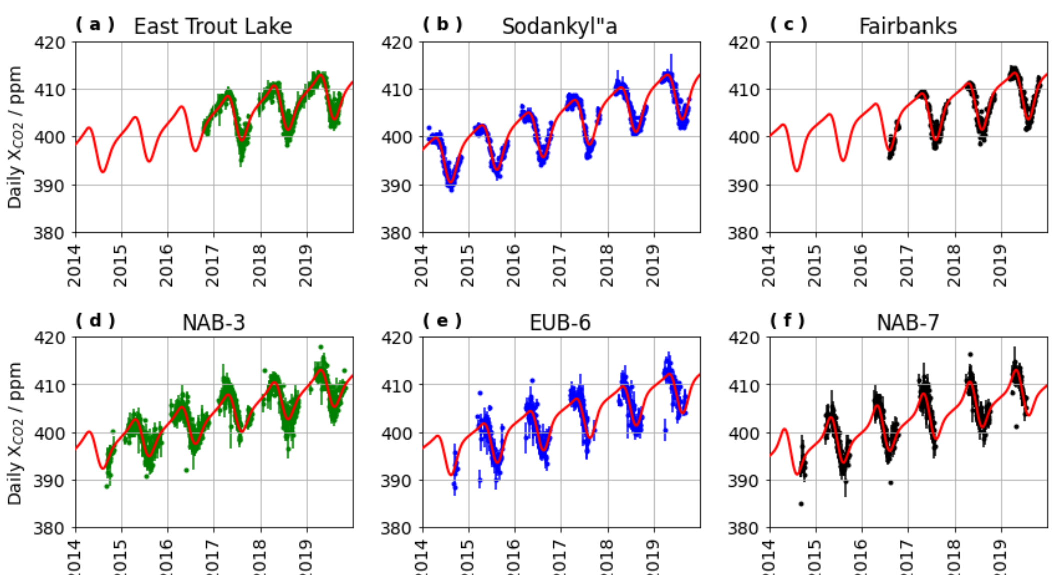

Figure R1-2. Time-series of ground-based observations (panels (a), (b), and (c)) and satellite-based observations for the zone containing the ground site (panels (d), (e), and (f)) at two TCCON sites East Trout Lake, Canada and Sodankylä, Finland, as well as Fairbanks Alaska, where EM27/SUN observations are collected. Also shown are seasonal cycle fits for these observations to a skewed sine wave equation, using methods described by Lindqvist et al. (2015), with the form: $f(t) = a_0 + a_1 t + a_2 \sin\left(\omega\left[t - a_3\right] + \arccos\left[a_4 \cos\left(\omega\left[t - a_5\right]\right)\right]\right)$.

In addition, we did provide uncertainty estimates for all parameters in the seasonal cycle fits, which translate to uncertainty in SCA and are derived using standard errors that are defined by daily standard deviation in X_{CO_2} for spatially and temporally averaged observations and model estimates, and fixed at 0.25 ppm for single-point daily model estimates near the ground sites. This method of estimating uncertainty is also following methods described in the cited literature (Lindqvist et al., 2015). We have added more details on this to the methods section of the paper. It is our understanding that boot-strapping methods are intended to be a work-around for situations in which statistical uncertainty is difficult or impossible to calculate, and it should not be warranted when we can directly calculate an uncertainty statistic.

2. **Comment:** I was not able to understand if the SCA analysis accounts for temporal sampling differences between OCO-2 and the model simulations. It appears that the models are sampled daily throughout the year. I think that it is quite likely that the SCA fits will be quite sensitive to the observational sampling. The sensitivity of SCA estimates to temporal sampling should be quantified.

   Reply: It is unclear to us whether this comment is referring to temporal sampling in terms of the time-of-day for measurements included in daily average X_{CO_2} or in terms of annual temporal sampling with respect to winter-time data gaps, so we address both interpretations in this reply.

   Time-of-day reply: OCO-2 is polar orbiting and as a result the OCO-2 observations are consistently within about an hour of local solar noon. The higher the latitude for northern latitudes, the closer the overpass is to local solar noon and the closer one gets to having measurements on a daily basis. NNG (Near Noon Ground-based) observations are used for the 5 ground sites in this analysis, which are the average of ground-based retrievals within $\pm$ 2 hours of local solar noon. The GEOS-Chem CO_2 simulation only provides daily estimates and higher temporal resolution is not available to us. For the CAMS estimates of X_{CO_2} values are provided every 3 hours and we took averages of every 24 hour day in UTC. To address the question of the effects of temporal sampling in the CAMS model estimates we compare the results of seasonal cycle fits to daily averages of X_{CO_2} with fits to X_{CO_2} for the time-point nearest local solar noon for European Boreal zone 6, containing Sodankylä (see Fig. R1-3), and North American Boreal zone 7, containing Fairbanks (see Fig. R1-4). Visually there is no perceivable difference in the seasonal fits and differences between the fits on a daily basis are smaller than estimates of instrument precision for satellite-based and ground-based observations. The overall differences in resulting SCA are less than 0.1 ppm and relative to a total range in SCA across northern high latitude regions of approximately 7 ppm, this shift is on the order of 1.4% of the overall variability in SCA across northern high latitudes.

   Annual reply: An alternative interpretation of this comment is not that it refers to temporal sampling in terms of time of day, but to temporal sampling as the number of days represented in the time-series (e.g. the winter data gap). This would refer to the way wintertime data gaps may effect OCO-2 or NNG SCA. To address this, a section has been added to the supplement and is mentioned in Sect. 3.2 of the revised manuscript. For this extended analysis we consider SCA calculated for CAMS time-series that include only days for which there are corresponding OCO-2 observations within a given zone. We found that the CAMS SCA from the clipped time-series are within +/- 0.5 ppm of CAMS SCA for the full time-series for all zones.

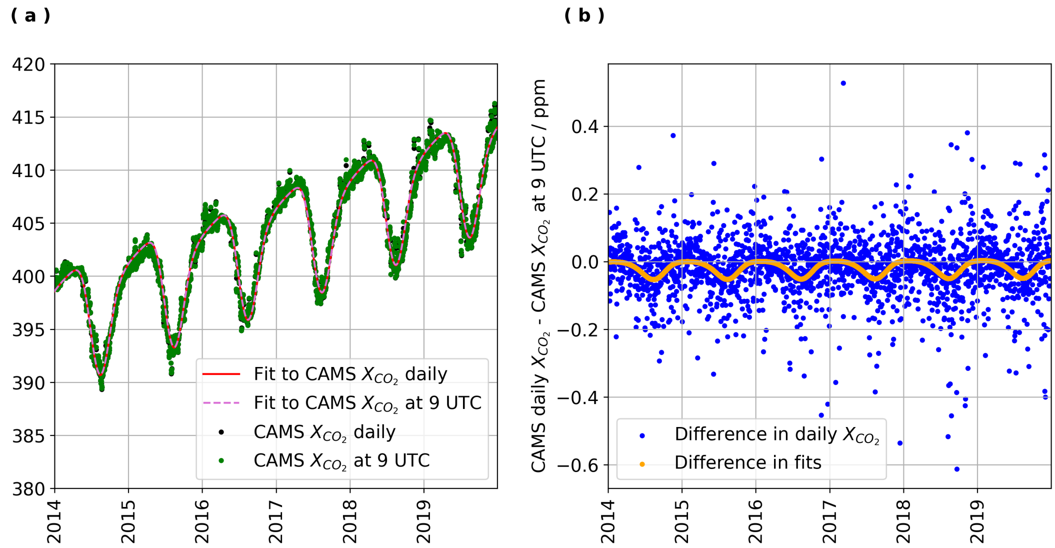

Figure R1-3. (a) CAMS 24-hour daily averages of X_{CO_2} and daily single point X_{CO_2} estimates near local solar noon (9:00 UTC in this case), averaged spatially for the zone containing Sodankylä, along with seasonal cycle fits for each set of daily X_{CO_2} estimates to a skewed sine wave (Lindqvist et al., 2015). (b) Daily residuals between the 24-hour daily averages and the single point estimates near local solar noon, as well as residuals between fits to each set of daily X_{CO_2} estimates. SCA$_{daily}$ = 10.76 ppm, SCA$_{9UTC}$ = 10.72 ppm, HDD$_{daily}$ = 174, HDD$_{9UTC}$ = 175.

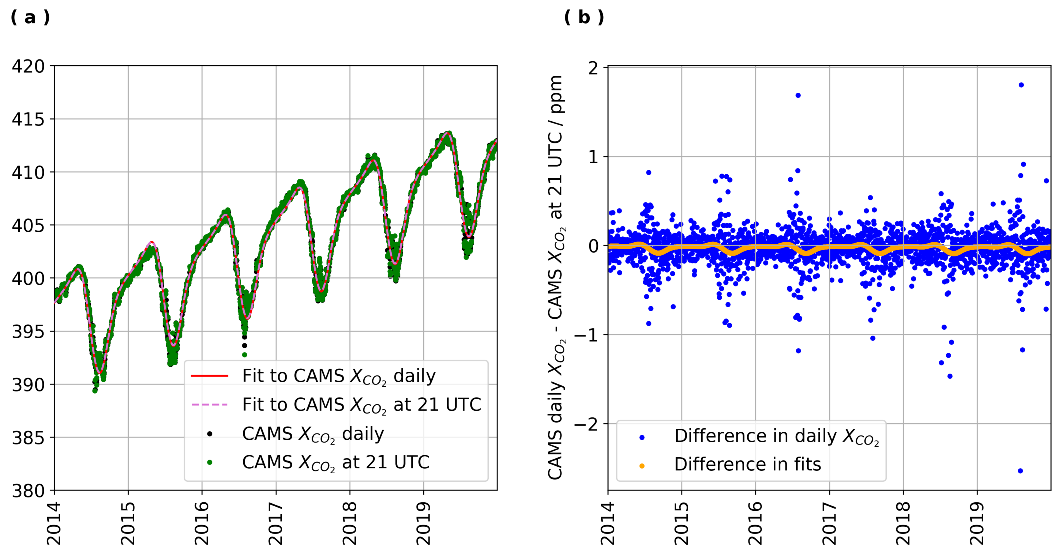

Figure R1-4. (a) CAMS 24-hour daily averages of X_{CO_2} and daily single point X_{CO_2} estimates near local solar noon (21:00 UTC in this case), averaged spatially for the zone containing Fairbanks, along with seasonal cycle fits for each set of daily X_{CO_2} estimates to a skewed sine wave (Lindqvist et al., 2015). (b) Daily residuals between the 24-hour daily averages and the single point estimates near local solar noon, as well as residuals between fits to each set of daily X_{CO_2} estimates. SCA$_{daily}$ = 10.57 ppm, SCA$_{21UTC}$ = 10.50 ppm, HDD$_{daily}$ = 174 = HDD$_{21UTC}$.

3. **Comment:** The contact tracer analysis was insufficiently described. I could not understand what this analysis was telling us. When are the tracers released? And what amount? How long is the simulation run? What precisely is being shown in Figure 10 (tracer was plotted at what time? for simulation starting on what day? that released what quantity of tracer? And released it over what spatiotemporal window?)?

   Reply: The requested details on the contact tracer analysis were added to Sect. 2.7 of the revised manuscript.

4. **Comment:** The manuscript is quite hard to follow in places. It would help to explicitly describe the subpanels in the figures, for example, "(a) Quantity Y versus quantity X with Z processing". Please also ensure that the main results from the figures are described in the text when the figure is referenced.

   Reply: We have attempted to clarify figure captions and descriptions in the text.

5. **Comment:** I think that the impact of this study would be improved if the OCO-2 retrievals with the standard bias correction and filtering were presented throughout the main text in addition to the high-latitude focused bias correction and filtering. As a potential user of these data, I am very interested in understanding the impact of differences in bias correction and filtering. From Fig. S1, it appears that differences are substantial. Furthermore, it would be of interest to determine if differences between different QC/bias-correction result in larger SCA differences than between the models.

   Reply: This analysis is intended to address science questions about the spatial distributions of seasonal cycles over northern high latitudes, and is not intended to be a technical evaluation of OCO-2 bias correction methods. The comparisons between B9 with the standard bias correction and the alternative bias correction were addressed in Jacobs et al. (2020). As a result, we feel that it is sufficient to demonstrate that the alternative bias correction improves agreement with ground-based and model-derived seasonal cycles (which we do in the supplemental materials), and a full analysis of OCO-2 retrievals using the standard B9 bias correction is outside the scope of this paper. It is not necessarily the case that using the standard B9 bias correction for data in this analysis would yield any additional useful insights. In addition, retrieval and bias correction methods for OCO-2 data are continuously being reassessed and modified for better accuracy and precision.

2 Specific Comments

Comment: P1L12: It is a little confusing to refer to a GEOS-Chem run with CT2019 fluxes as "GEOS-Chem". It would be better to use a specific acronym such as "GC-CT2019" to make clear that it is GEOS-Chem run with CT2019.
Reply: Changed this as requested. Note that the GEOS-Chem CO2 simulations also use other flux estimates from inventories, the CT2019 contribution is only for biospheric fluxes, but it still seems fine to label it as the reviewer suggests.

Comment: P1L16-17: This is only for >50N that is examined here. The meridional gradient is still greater from 0 ppm at the equator to >10ppm at the North pole.
Reply: Added specific latitude range that this is referring to.

Comment: P1L16-17: Reads strange to use "Longitudinal" and "meridional". I suggest using "zonal" instead of "longitudinal".
Reply: We have switched to using longitudinal and latitudinal, but "zonal" seems like it could get confused with references to the 5° latitude by 20° longitude zones that we define.

Comment: Sec 2.4: I do not see how XCO2 is calculated. Is an averaging kernel applied or is it the true XCO2?
Reply: There are no averaging kernels applied to any of the model estimates of XCO2 before calculating the seasonal cycle fits because we hope to take the seasonal cycles as separate independent products from each model. In general, averaging kernel

corrections are small enough that they would not be expected to cause substantial changes to XCO2 seasonal variability when evaluated over multiple years.

Comment: Sec 3.1: I found the main points of Sec 3.1 quite unclear. It would be helpful to walk the reader through the results. The section starts by stating that differences in spatial sampling may impact the results, but from reading the rest of the section I am not clear on the impact of spatial sampling on the results. It would be helpful to explicitly state the results of this comparison, and what are the implications for the analysis that follows. In particular, I'm having a hard time understanding what Fig. 2 and Fig. 3 are telling us (please explicitly state what the sub-panels are showing). What does Fig. 3 show us that Fig. 2 does not? And why are there TCCON symbols for the OCO-2 vs model comparison?
Reply: Figure 2 is comparing between observations and models by showing parameters from seasonal cycle fits to observations correlated against parameters from seasonal cycle fits to model estimates for different spatial scales focused on the 5 ground sites. In this case the single point model estimates nearest the ground site are matched with near-noon ground-based (NNG) observations, spatially averaged model estimates within the 5° latitude by 10° longitude coincidence region centered on the ground site location are matched with spatially averaged OCO-2 observations within the coincidence region, and spatially averaged model estimates within the 5° latitude by 20° longitude zone in which the ground site is located are matched with spatially averaged OCO-2 observations within that zone. Figure 3 is showing how parameters from seasonal cycle fits to spatially averaged OCO-2 observations or spatially averaged model estimates compare to parameters from seasonal cycle fits to NNG or single point model estimates nearest the ground site. Figure 3 is looking at the effects of spatial averaging within the separate contexts of observations or models, but does not compare between observed and model-derived seasonal cycles. To address the question, "And why are there TCCON symbols for the OCO-2 vs model comparison?", we are considering only data with relative proximity to the five ground sites, so symbols in Fig. 2 and 3 are labeled by the corresponding ground site. We have provided these details already in the captions of Fig. 2 and 3 in the submitted manuscript and it is unclear how we could present the information more clearly. We have attempted some refinement of the language in the figure captions.

Comment: P10L9-10: "Results in the supplement (see Fig. S30) indicate that SCA derived from clipped time-series of OCO-2 and CAMS (restricted to 2014-2016) were only marginally different from SCA derived for the full time (2014-2019)". Figure S30 does not really support this claim. The figure shows difference in SCA of 0.5 ppm and up to 10 days in HDD for TCCON sites. These do not seem marginal.
Reply: There were some points on the plots in Fig. S30 for the ground-based sites that should not have been included in the plots and were actually correlating CAMS and OCO-2 2014-2019 results against GC-CT2019 results. The supplemental figure has been corrected to exclude the ground-based sites. In addition, there has been some revision of the manuscript to provide more specific details when referring to these results.

Comment: P10L20: What does "systematic distribution" mean?
Reply: Text omitted.

Comment: P10L31: What does it mean for a gradient in North America to be "consistent" with a gradient in Eurasia?
Reply: Changed to "in the same direction as".

Comment: Sec. 4.1 & Sec 4.2: These are both results sections, and the methods for this analysis need to be fully described. I suggest moving Sec 4.1 and Fig. 8 to supplementary materials, as it is well known that terrestrial biosphere fluxes drive seasonal variations in XCO2.
Reply: Moved to the supplement.

Comment: Fig 4.: What exactly is being plotted? Is this an instantaneous field? And after a simulation of what length?
Reply: This figure shows maps of SCA and HDD from seasonal cycle fits for each $5^{c}irc$ latitude by 20° longitude zone represented by a color scale. The time periods for the seasonal cycle fits are given in the title of each panel of the figure and are also stated in the methods sections. We modified the caption of the figure in an attempt to make this more clear.

Comment: P13L32: I do not understand what is being correlated, and what is the correlation coefficient?
Reply: This has been revised in an attempt to provide more clarity, but was also changed as part of other reformatting of the paper.

Comment: P14L12-18: This seems out of place; this should be in the methods section.
Reply: The paper has been reformatted to include a methods section on this.

Comment: P14L20: Why are the CT2019 fluxes being referred to as “GEOS-Chem”. This analysis is only looking at CT2019 fluxes, GEOS-Chem is not used here.
Reply: The fluxes are a combination of CT2019 fluxes for biospheric exchange and emissions estimates from inventories as explained in section 2.4. They represent fluxes used in the GEOS-Chem CO2 simulation, which is why we refer to them as GEOS-Chem. This is stated on P14L9-13 in the first paragraph of section 4.3, but in response to your previous comment, we have moved that paragraph to the methods section. In addition, we hope that relabeling to GC-CT2019 will further help to resolve this issue.

Comment: P15L20-21: “The dominance of terrestrial biospheric exchange in the GEOS-Chem model is likely an intentional quality built into the model” – What does this mean?
Reply: Removed.

Comment: P15L24: “15 or 30 lifetimes”?
Reply: Edited to say "15 or 30 day lifetimes".

Response to Reviewer 2

We would like to express our gratitude to reviewer 2 for their insightful and constructive suggestions to reformat the paper and extend our analysis to additional model estimates in CarbonTracker2019. We have done our best to use these comments as a guide to composing a revised version of the paper that improves and expands on our original submission.

3 Comments and responses

1. **Comment:** My largest concern with the paper is that the authors are underutilized their contact tracing simulations. The contact tracing approach was not clearly explained, and results from these simulations were introduced only in the discussion section. This needs to be fleshed out more in the results section, with a mind to new or refined supporting figures. Fig. 10 could be modified to better guide the reader as to the similarities noted with Fig. 4. For example, could boxes be added in Fig. 10 around the zones whose amplitude is in the top Nth percentile in Fig. 4? Related to the contact tracer approach, it was unclear why the CarbonTracker posterior fluxes were run through GEOS-Chem rather than simply using the CarbonTracker posterior CO2 fields. Was this so the transport model was consistent for the CO2 simulation and the contact tracer simulation? If so, this should be explained in the text. However, the use of a single transport model is a major limitation to the authors' conclusion that accumulation of transported CO2 controls the seasonal cycle amplitude spatial patterns. It may be onerous to repeat the surface contact tracer analysis in another transport model, but could the authors combine information from the CarbonTracker (TM5) posterior and their own CarbonTracker/GEOS-Chem simulations to substantiate their conclusion?
Reply: We now show results with CarbonTracker2019 X_{CO_2}, as well as GEOS-Chem. The X_{CO_2} amplitudes are generally similar in GEOS-Chem and CarbonTracker/TM5, with some expected differences attributable to the different transport schemes used in GEOS-Chem and TM5. The similarity of X_{CO_2} SCA in GEOS-Chem and CarbonTracker/TM5 while using the same terrestrial and ocean fluxes suggests that large-scale atmospheric transport that influences SCA is similar in the two models. We are then able to diagnose the role of transport in GEOS-Chem using the novel surface contact tracers. Furthermore, we have made the results from our contact tracer simulations a more integral part of the paper by including dedicated methods, results, and discussion sections (see Sect. 2.7, 3.5, and 4.1 of the revised manuscript), which now provide more details on this component of the analysis.

2. **Comment:** A second major issue is that the key insights from the research could be better laid out. Is it that the column seems to show earlier drawdown than inversions constrained by surface observations? The different behavior in Siberia compared to other zones? What are the implications for a correlation between amplitude and HDD in terms of processes that affect the seasonality of net exchange? Perhaps reorganization of the discussion section, with the surface tracer analysis presented in the results, will make it easier for these points to come through in the discussion.
Reply: We have reformatted the paper, with the tracer analysis presented in the results and more streamlined discussion. We state in that the correlation between SCA and HDD would be interesting to consider in future studies, but do not know what is driving this relationship at this time.

3. **Comment:** I struggled with Section 3.2, since the section begins with a listing of material in the supplement. The authors should organize this information to provide a clear summary in the first few sentences of the paragraph/section, and then later refer to the supplementary figures for more details. In particular, the "unrealistic drop in wintertime values" might warrant a figure in the main text since the realism of the seasonal cycle fits is crucial for interpretation of the rest of the paper.
Reply: This section has been revised to make it more clear. We think that it is reasonable to refer the reader to the supplement if they are interested in the exact nature of how the winter data gaps affect seasonal cycle shape, rather than including an additional figure. Some details have also been added to the methods section on seasonal cycle fitting at the

behest of reviewer 1 that may be pertinent to this comment.

4. **Comment:** The term HDD was used before the concept was introduced or defined (L29 p7).
   Reply: Text has been revised to correct this.

5. **Comment:** The authors cite a meridional gradient in Fig. 12, which I didn't see convincingly in any panel. This is brought up in the discussion (p14), and I don't think removing this sentence would affect the authors' underlying argument. If it is left in the document, more support is required.
   Reply: This sentence was removed.

References

Jacobs, N., Simpson, W. R., Wunch, D., O'Dell, C. W., Osterman, G. B., Hase, F., Blumenstock, T., Tu, Q., Frey, M., Dubey, M. K., Parker, H. A., Kivi, R., and Heikkinen, P.: Quality controls, bias, and seasonality of CO2 columns in the boreal forest with Orbiting Carbon Observatory-2, Total Carbon Column Observing Network, and EM27/SUN measurements, Atmos. Meas. Tech., 13, 5033–5063, https://doi.org/10.5194/amt-13-5033-2020, 2020.

Lindqvist, H., O'Dell, C. W., Basu, S., Boesch, H., Chevallier, F., Deutscher, N., Feng, L., Fisher, B., Hase, F., Inoue, M., Kivi, R., Morino, I., Palmer, P. I., Parker, R., Schneider, M., Sussmann, R., and Yoshida, Y.: Does GOSAT capture the true seasonal cycle of carbon dioxide?, Atmos. Chem. Phys., 15, 13 023–13 040, https://doi.org/10.5194/acp-15-13023-2015, www.atmos-chem-phys.net/15/13023/2015/, 2015.

Wunch, D., Wennberg, P. O., Messerschmidt, J., Parazoo, N. C., Toon, G. C., Deutscher, N. M., Keppel-Aleks, G., Roehl, C. M., Randerson, J. T., Warneke, T., and Notholt, J.: The covariation of Northern and Hemisphere summertime CO_2 with surface temperature in boreal regions, Atmos. Chem. Phys., 13, 9447–9459, https://doi.org/10.5194/acp-13-9447-2013, www.atmos-chem-phys.net/13/9447/2013/, 2013.