Response to Reviewer 2

Original text is in black. Our responses are in blue text. References to the revised text take the following format: RP78,L54-55 (Revised Page 78, Lines 54-55). The reference are for the manuscript with changes tracked.

According to my understanding, this manuscript addresses two major topics. The first is how adjustments to surface fluxes (posterior minus prior) manifest themselves in the atmosphere. This is done by performing inversions for the first two years of GOSAT data using a variational GEOS-Chem system, and propagating the posterior and prior fluxes through a transport model. Along the way, the authors perform some evaluation of their inverse results, such as comparison to TCCON and HIPPO. The second is how that manifestation varies if a different higher resolution online atmospheric transport model is used. In my opinion, the authors spend too much time on the first topic and not enough on the second, which makes the work not significant enough for a journal like Atmospheric Chemistry and Physics. If this focus were reversed, or the first topic were explored further (explained below), it would make for a much more interesting and scientifically significant paper. The authors perform inversions of GOSAT and in situ data for two years, and look at the fluxes and resultant atmospheric CO_2 fields in the first two years of GOSAT, primarily focusing on 2010. They use a variational inversion technique using the GEOS-Chem transport model. Their conclusions are very similar to previously published literature, such as Houweling et al (2015), Basu et al (2013), Chevallier et al (2014), which they cite. In fact, a very similar (if not identical) set of inversions was already submitted by some of the co-authors to an intercomparison of GOSAT inversions published by Houweling et al (2015). As far as I can tell, there is nothing new or unique about their inversion or analysis compared to the multitude of GOSAT inversions already published for 2010, and this part of the work does not add to the body of existing knowledge about GOSAT retrievals and derived fluxes in and around 2010. GOSAT has been up for eight years now, and retrievals of column CO_2 from GOSAT exist for the majority of that period. I do not understand why the authors have limited their study to the first couple of years of GOSAT data. If the authors want to publish a GOSAT inversion study that would be of value to the scientific community, I would recommend performing a longer term study, such as (say) the inter-annual variability of fluxes as seen by GOSAT, or the longer term trends in atmospheric CO2 and CO2 fluxes as seen by GOSAT. The current inversion study, focused on 2010 (with some padding on either side), is of limited interest. The second thread in their work, however, is more interesting. They perform forward runs with two different models of atmospheric transport driven by the same fluxes and look at the difference in the “flux signal” in the atmosphere. The non-GEOS-Chem model is the higher resolution GEM-MACH-GHG, a fairly new addition to this community (Polavarapu et al, 2016). Not only did they transport CO_2 with GEM-MACH-GHG, they also perturbed the transport with analysis errors from the meteorological assimilation system, thereby simulating the impact of uncertainties in the met fields on CO_2 variations. They derive a “baseline” CO_2 variation from this error propagation, contending that variations smaller than this detected by an observing system cannot be reliably ascribed to fluxes. This, to my knowledge, is fairly unique in the tracer transport community, and provides a recipe for deriving transport errors in CO_2 space. Such errors can be used, e.g., if GEM-MACH-GHG or a derived offline model is used for trace gas inversions. This technique may also be valid for deriving “baseline” transport errors for an offline model if an ensemble is run for the parent model with greenhouse gases (e.g., GEOS5 for GEOS-Chem).

If the authors would like to revise their manuscript and make it scientifically significant enough for this journal, I can offer two different suggestions. Either they need to extend their GOSAT analysis to 5+ years and address questions such as long term trends and interannual variability of CO_2 fluxes. Or they need to more or less excise the GOSAT inversions and focus on the performance of GEM-MACH-GHG in simulating atmospheric CO_2 and its meteorological errors. For the first choice, I would suggest questions such as:

1. Do GOSAT retrievals estimate a stronger European sink consistently over time, as first suggested by Reuter et al (2014) with SCIAMACHY and a single year of GOSAT data?
2. Do GOSAT retrievals require a stronger northern hemisphere uptake consistently, as noted by Houweling et al (2015) for one year?
3. According to GOSAT, which region contributes most to the interannual variability of atmospheric CO_2, the Tropics or semi-arid ecosystems? This has been an ongoing debate in the atmospheric carbon community, see e.g., Baker et al (2006), Poulter et al (2014) and Ahlström et al (2015).
4. Are there persistent differences between GOSAT and surface data inversions across multiple years?

These are just some suggestions, and I'm sure the authors can think of many such questions to address with a multi-year GOSAT inversion. On the other hand, if the authors choose to focus on GEM-MACH-GHG, then that would make for a very interesting paper as well. The authors have already addressed some of the interesting questions that arise from using a high resolution online model for CO_2 transport. Some additional questions could be:

1. Are high frequency variations of CO_2 near the surface better represented by the higher resolution model? If yes, we could potentially move to assimilating more data from surface measurement sites in the future with online models such as GEM-MACH-GHG.
2. Can one construct a "look up table" for the baseline transport-driven errors using GEM-MACH-GHG, varying (say) by region and season? How do those errors differ between surface and total column measurements? I'm looking for something like Figure 17, but much finer grained than three zonal bands. At the very least, ocean sites, coastal sites and continental sites should be separated. Similarly, for total column measurements, ocean and land soundings should be separated.
3. If inversions were performed using errors from step 2, versus more traditional prescription of errors, how do the fluxes change?
4. Is it true that transport errors matter less in assimilating a total column than assimilating surface sites or a vertical profile? This was first suggested by Rayner & O'Brien (2001), but to my knowledge never explicitly demonstrated. The crucial thing to compare here would be the size of the transport error and the size of the flux signal, since that is small as well in the total column.
5. Is there any covariance between transport error and CO_2 variation, especially along weather fronts? This has also been a topic of much debate, especially whether assimilating high frequency CO_2 measurements can improve weather forecasts (Engelen et al, 2001), and whether CO_2 inversions need to assimilate met observations. Again, these are just some suggestions, and I'm sure the authors could come up with an interesting set of questions relevant to the atmospheric CO_2 community that they could answer with GEM-MACH-GHG.

Response: We are grateful to the Reviewer for the comments and suggestions. It is clear from the comments that the organization of our original manuscript led to some confusion about the focus of the manuscript. Specially, the main focus of the manuscript is on the new diagnostic associated with the posterior atmospheric CO_2 adjustment. We felt that it was necessary to provide a thorough evaluation of the posterior fluxes so that the reader would be able to interpret the main results of the work. As we noted in our manuscript, our fluxes are for the most part documented in the literature. Nevertheless, it would be impossible for us to proceed without describing them (at least briefly) here for two reasons. (1) Modeling and data assimilation systems are constantly changing and improving. Because of the passage of time, the inversions done here are close to but not identical to those shown in Deng et al. (2014, 2016). It is therefore important to assure the reader that our inversion results are understandable and reasonable. (2) The flux inversion estimates under study here were described in two separate papers (Deng et al. 2014, 2016) and the two inversions results were never presented together. Because the main part of the paper compares the impact of these two posterior fluxes on atmospheric distributions, it is useful to see them in a side-by-side comparison. Furthermore, we felt that it would be unreasonable to expect the reader to be familiar with both Deng et al. papers to be able to interpret the results presented in the manuscript. However, we acknowledge the Reviewer's concern and have shortened the discussion of the flux inversion results in section 3.1 to only those remarks necessary for understanding the later sections. Specifically, we stick to only a discussion of zonally averaged fluxes. Figure S6 replaces Figures 5 which was moved to the supplemental material. Figure 6 was deleted. We also now explicitly state that the reasons for very briefly describing the posterior fluxes.

Another reason we felt that it was necessary to go into detail about the evaluation of the flux inversions is because our diagnostic is based on atmospheric adjustments from a prior distribution, therefore, when we compare adjustments from multiple posterior fluxes, we cannot determine which is more "correct". However, we can infer which is "better" by comparing posterior CO_2 fields to measurements. Thus, the comparisons to observations is not done to validate the fluxes, but rather to inform the discussion of the posterior atmospheric adjustments resulting from the fluxes. The importance of these comparisons is recognized by Reviewer 1 since his/her comments primarily focus on improving the discussion related to validation of the posterior CO_2 distributions. In retrospect, the narrative in the original manuscript may have been deficient in this regard. However, in the revised manuscript, the focus of our work is squarely on the new diagnostic and its potential value for learning about flux inversion results. This involved rewriting the abstract and revising portions of the introduction and discussion sections as well as adding explanatory sentences throughout the results section. Finally, we appreciate the enthusiastic comments from the Reviewer regarding our new approach of comparing posterior atmospheric adjustments due to fluxes with those due to meteorological uncertainty.

In summary, as a result of the Reviewer's comments, we have rewritten portions of the manuscript to improve clarity and to focus on our main points: the introduction of a new diagnostic which is useful for studying flux inversion model results.

Other comments

1. The prior fluxes in Figure 5 look strange. It is rare for me to see a terrestrial flux prior that is positive in the annual aggregate over North America, and Boreal and Temperate Eurasia. Where do those priors come from?

Response: As mentioned in Deng et al. (2014), the priors come from the Boreal Ecosystem Productivity Simulator (BEPS) model (Chen et al. 1999) driven by NCEP reanalysis data and remotely sensed leaf area index. Deng and Chen (2012) scaled the hourly GPP and Respiration (RSP) to generate a set of GPP and RSP that are annually neutral for each grid box as priors for atmospheric inversion modeling. The positive priors mentioned here is/are caused by adding the biomass burning emissions. The addition of biofuel and biomass burning sources to the natural fluxes was noted in the caption for Figure 5. However, this figure was moved to the supplementary section (Figure S6) in the revised manuscript.

Chen, J. M., Liu, J., Cihlar, J., and Goulden, M. L.: Daily canopy photosynthesis model through temporal and spatial scaling for remote sensing applications, Ecol. Model., 124, 99–119, 1999.

2. In Figure 7, I would prefer to see time averages, say over a week or month, instead of snapshots. Snapshots often display misleading variations that do not matter for what the authors are considering. This comment only holds, of course, if an equivalent of Figure 7 still exists in the revised manuscript.

Response: In general, time averages are typically preferred for the reasons the reviewer presents. However, in our manuscript, we need snapshots and not time averages. Figure 7 is an encapsulation of the full time animations provided in the supplemental material. The evolution of the flux signal reveals interesting differences and hints at transport pathways that are explored in the section on adjoint sensitivity (and Figure 12). The time evolution of the animations (which the snapshots summarize) also corresponds to the time evolution of flux signals seen in later figures (Figures 13-18). Furthermore, the qualitative consistency of differences between the two types of flux signals across the two years, given that these are snapshots, is additional useful information.

3. I was surprised to see no data providers as co-authors in an inverse modeling paper. It is usual in this field to offer co-authorship to data providers, which they may or may not accept. In fact the ObsPack fair use policy explicitly states:

> *"Your use of this data product implies an agreement to contact each contributing laboratory to discuss the nature of the work and the appropriate level of acknowledgment. If this product is essential to the work, or if an important result or conclusion depends on this product, co-authorship may be appropriate. This should be discussed with each data provider at an early stage in the work. Contacting the data providers is not optional; if you use this data product, you must contact the data providers."*

Were the data providers contacted, at the very least to let them know that an inversion study using their data was about to be submitted? If not, that is a significant oversight that needs to be corrected.

Response: We are well aware of protocols for data usage and we fully appreciate the dedication and effort (by our close colleagues and by all measurement scientists) required to make high quality measurements. For that reason, we always try to acknowledge the expertise and effort required to make measurements and to provide them to researchers. Here are some specific details, regarding this manuscript.

a) TCCON. Some of our authors are co-located with TCCON PI Kim Strong at the University of Toronto. As with previous work, we contacted all TCCON PIs whose data appears in the tables and Figures. We also had discussions with Kim Strong about Eureka data, but we had no requests from any TCCON PIs for co-authorship. This is the same procedure and same result we got in our previous article (Polavarapu et al. 2016, ACP). The references for all 14 sites were cited in the text and all PIs were thanked in the acknowledgements.

b) HIPPO. We contacted Steve Wofsy by email (21 March 2017) and received no reply. This was also the case with our previous publication. (Also, in the case of the Deng et al. (2016) manuscript, Steve Wofsy declined co-authorship.) In the absence of a reply for this manuscript, we referenced his papers and the dataset DOI and added an acknowledgement. Note that the data usage protocol found at https://www.eol.ucar.edu/system/files/HIPPO_Full_Data_Policy_lah_20170915_1.pdf requests only the following: (1) Acknowledge with references and use the DOI number, (2) Acknowledge NSF and NOAA in the acknowledgements, (3) add "HIPPO, HIAPER Pole-to-Pole Observations, National Science Foundation, NSF, NSF/NCAR Gulfstream-V (GV)" to keywords. Note that we had done all of these in the original manuscript.

c) NOAA aircraft profiles. We contacted Colm Sweeney by email (8 Nov 2017) and provided the manuscript, figures and supplemental material and received no response. In the absence of a response, we cited his publications, added an acknowledgement to him, and submitted the manuscript on 29 Dec 2017. NOAA aircraft profiles were obtained from ObsPack2013 which was acknowledged as noted in (d).

d) NOAA surface measurements. For ObsPack, the contact in the datafiles mentioned is Ken Masarie. But we heard from NOAA colleagues that Ken had retired. Therefore, we contacted Arlyn Andrews (8 Nov 2017). Again, we provided the manuscript, figures and supplemental information. We did not get a reply. In the absence of a response, we cited Conway and Tans (2012), Conway et al. (2011), Masarie et al. (2014) and the DOI (http://dx.doi.org/10.3334/OBSPACK/1001).

4. I have a problem with the terminology "flux signal", even though the authors made the explicit caveat that this "signal" by definition depends on the inverse model and the prior. The term "flux signal" makes it sound like it's an inherent property of the observations, which it is not. I would recommend using a different term, such as "CO_2 adjustment" or "mole fraction update".

Response: Reviewer 1 had exactly the same concern and suggested an alternative expression which we like: "Posterior Atmospheric Adjustment". We adopted this new terminology and its acronym (PAA) in the revised manuscript.

References

A. Ahlström, M. R. Raupach, G. Schurgers, B. Smith, A. Arneth, M. Jung, M. Reichstein, J. G. Canadell, P. Friedlingstein, A. K. Jain, E. Kato, B. Poulter, S. Sitch, B. D. Stocker, N. Viovy, Y. P. Wang, A. Wiltshire, S. Zaehle, and N. Zeng, "The dominant role of semi-arid ecosystems in the trend and variability of the land CO2 sink," Science (80-.)., vol. 348, no. 6237, 2015.

D. F. Baker, R. M. Law, K. R. Gurney, P. Rayner, P. Peylin, A. S. Denning, P. Bousquet, L. Bruhwiler, Y.-H. Chen, P. Ciais, I. Y. Fung, M. Heimann, J. John, T. Maki, S. Maksyutov, K. Masarie, M. Prather, B. Pak, S. Taguchi, and Z. Zhu, "TransCom 3 inversion intercomparison: Impact of transport model errors on the interannual variability of regional CO2 fluxes, 1988-2003," Glob. Biogeochem. Cycles, vol. 20, no. 1, p. GB1002, Jan. 2006.

S. Basu, S. Guerlet, A. Butz, S. Houweling, O. Hasekamp, I. Aben, P. Krummel, P. Steele, R. Langenfelds, M. Torn, S. Biraud, B. Stephens, A. Andrews, and D. Worthy, "Global CO2 fluxes estimated from GOSAT retrievals of total column CO2," Atmos. Chem. Phys., vol. 13, pp. 8695–8717, 2013.

F. Chevallier, P. I. Palmer, L. Feng, H. Boesch, C. W. O'Dell, and P. Bousquet, "Toward robust and consistent regional CO2 flux estimates from in situ and spaceborne measurements of atmospheric CO2," Geophys. Res. Lett., vol. 41, no. 3, pp. 1065–1070, 2014.

R. J. Engelen, G. L. Stephens, and A. S. Denning, "The effect of CO2 variability on the retrieval of atmospheric temperatures," Geophys. Res. Lett., vol. 28, no. 17, pp. 3259–3262, 2001.

S. Houweling, D. Baker, S. Basu, H. Boesch, A. Butz, F. Chevallier, F. Deng, E. J. Dlugokencky, L. Feng, A. Ganshin, O. Hasekamp, D. Jones, S. Maksyutov, J. Marshall, T. Oda, C. W. O'Dell, S. Oshchepkov, P. I. Palmer, P. Peylin, Z. Poussi, F. Reum, H. Takagi, Y. Yoshida, and R. Zhuravlev, "An intercomparison of inverse models for estimating sources and sinks of CO2 using GOSAT measurements," J. Geophys. Res. Atmos., vol. 120, no. 10, pp. 5253–5266, 2015.

S. M. Polavarapu, M. Neish, M. Tanguay, C. Girard, J. de Grandpré, K. Semeniuk, S. Gravel, S. Ren, S. Roche, D. Chan, and K. Strong, “Greenhouse gas simulations with a coupled meteorological and transport model: the predictability of CO2,” Atmos. Chem. Phys., vol. 16, no. 18, pp. 12005–12038, 2016.

B. Poulter, D. Frank, P. Ciais, R. B. Myneni, N. Andela, J. Bi, G. Broquet, J. G. Canadell, F. Chevallier, Y. Y. Liu, S. W. Running, S. Sitch, and G. R. van der Werf, “Contribution of semi-arid ecosystems to interannual variability of the global carbon cycle.,” Nature, vol. 509, no. 7502, pp. 600–3, May 2014.

P. J. Rayner and D. M. O’Brien, “The utility of remotely sensed CO2 concentration data in surface source inversions,” Geophys. Res. Lett., vol. 28, no. 1, pp. 175–178, 2001.

M. Reuter, M. Buchwitz, M. Hilker, J. Heymann, O. Schneising, D. Pillai, H. Bovensmann, J. P. Burrows, H. Bösch, R. Parker, A. Butz, O. Hasekamp, C. W. O’Dell, Y. Yoshida, C. Gerbig, T. Nehrkorn, N. M. Deutscher, T. Warneke, J. Notholt, F. Hase, R. Kivi, R. Sussmann, T. Machida, H. Matsueda, and Y. Sawa, “Satellite-inferred European carbon sink larger than expected,” Atmos. Chem. Phys., vol. 14, no. 24, pp. 13739–13753, 2014.