The authors would like to thank Anonymous Referee #3 for their comments on our manuscript entitled "The potential of clear-sky carbon dioxide satellite retrievals." Below, we have addressed their comments and made the necessary changes in the manuscript.

**1) The first major issue is that this is not a direct comparison between non-scattering (clear-sky) and scattering (FP) retrievals since the two approaches deal with surface pressure differently. It is retrieved in the FP retrievals by inclusion of the 0.76 micron band, but the clear-sky retrievals simply use the surface pressures from ECMWF. This difference is dismissed without much discussion or justification other than the fact that the O2 A band (0.76 micron) is the most sensitive of the three bands to cloud and aerosols.**

We tested allowing the clear-sky retrieval to use the 0.76 μm band to retrieve surface pressure and found the results to be consistent. Besides ECMWF being accurate to within 1-2 hPa, this is also partially due to the low uncertainty we assign to the ECMWF prior (1σ = 1 hPa). We have added the following to section 2:

> "To ensure this methodology was valid, we tested a~version of the clear-sky retrieval that was allowed to retrieve surface pressure from the \chem{O_2} A-band and found the results to be consistent."

**2) The second major issue is the very different result that was obtained between the tests with simulated OCO-2 data and real GOSAT data over the oceans. The real data showed much larger errors than the simulation, which is almost enough to bring the main conclusion into question. Fortunately, the authors are transparent about this difference and their inability to explain it (which is appreciated), but my guess is that it suggests that their representation of aerosols in the simulated data is not realistic enough. Not enough details of the treatment of aerosol in their simulator are provided to specifically critique their method, but Figure 9 suggests that the problem is worst over the equatorial Atlantic Ocean, a region with substantial Saharan dust aerosols, although this would not explain the larger errors over other ocean regions. I suspect that their simulated aerosol types or quantities differ substantially from the real aerosols encountered by GOSAT.**

At the request of another reviewer, we rethought our approach to analyzing the GOSAT retrievals and decided to apply a simple bias correction before post-filtering with DOGO. This resulted in lower GOSAT $X_{CO2}$ RMS errors (as expected) compared to before, which are more comparable to the simulated OCO-2 results. We have also updated figures 7 and 9, as you recommend later in this review, to provide a finer spatial resolution analysis of the regional error statistics. This again suggested that, for the simulations, the error statistics aren't especially regionally biased.

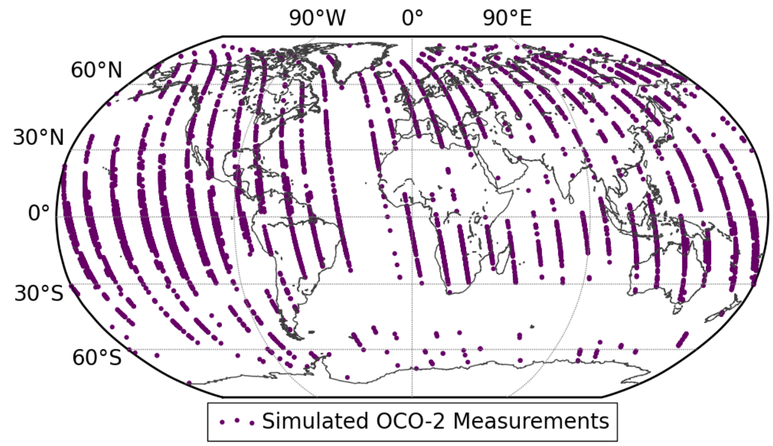
Regarding the aerosol setup in the simulator, the CALIPSO profiles used to generate the simulated radiances are actually quite realistic. I.e., from O'Dell (2010): 2.0 km vertical resolution, log-normal spherical particle distributions, realistic ice characteristics, six unique aerosol types, multiple cloud/aerosol layers.

**3) The third major issue is that nearly 50% of OCO-2 observations over land will be made in glint mode, yet the paper does not deal with these observations at all. Of course, the authors would not be able to test their method for these data with GOSAT, since it did not make land glint observations, but they still could have simulated land glint observations from OCO-2.**

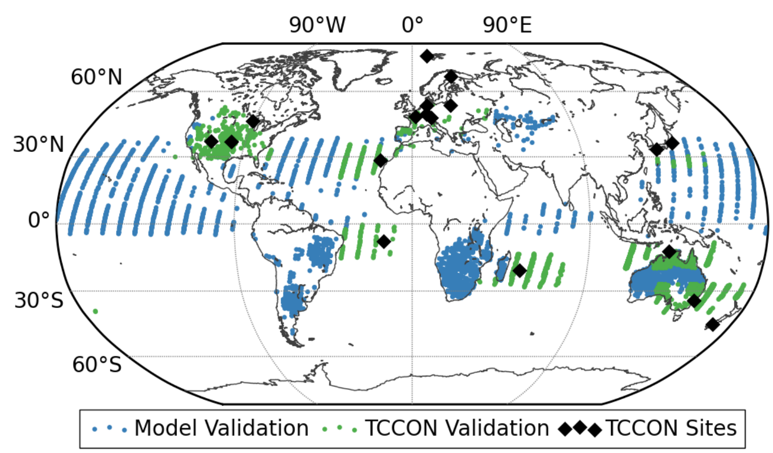
Besides not having corresponding GOSAT land-glint measurements, the errors for OCO-2 observations in glint mode over land may be worse than in nadir mode due to the longer slant paths through the atmosphere. We focused on a proof-of-concept approach here, which could be tested on the GOSAT data available at the time.

**4) The fourth major issue is that since pre-retrieval and post-retrieval filtering are so crucial to obtaining a good quality clear-sky retrieval dataset, the authors have not given enough detail about the post-filtered observational coverage. Post-filtered coverage is never shown in any of the figures. What kind of spatial coverage is obtained from a 30% throughput applied to observations that have already been filtered with the Aband preprocessor and the IMAP-DOAS preprocessor (used with GOSAT data)? Perhaps Figure 3 could be changed to show the unfiltered, pre-retrieval filtered, and 30% throughput post-retrieval filtered OCO-2 spatial coverage in different colors. Something similar could also be done for GOSAT. If the filtering is very spatially-dependent, then readers need to know this.**

In general, the filtering is not very spatially dependent. The figure below is the simulated OCO-2 data post-filtered to a throughput of 30%. Compared to figure 3 in the paper, a few regions have significantly lower data density, but this is to be expected. Soundings over the Sahara were likely removed due to aerosols (dust), soundings at very high latitudes removed because of low SNR caused by high zenith angles and being over snow/ice, and soundings over the southern ocean were probably removed because of clouds.




The spatial coverage in the clear-sky GOSAT data (pre-filtered and then post-filtered to a throughput of 30% via DOGO) is shown below. Overall, good global coverage is maintained, relative to figure 2 in the paper. Certain areas, such as the Sahara and Amazon, again have considerably fewer soundings, as they often contain clouds and aerosols. Very high latitudes also suffer in terms of data density, but that's again unsurprising as the SNR is lower and cloud/aerosol light path modifications can be more pronounced at higher zenith angles.




We have added the following to section 5.1:

> "The post-filtering to 30\,{{\%}} throughput is not spatially dependent, relative to Fig.~\ref{fig:OCO-2_worldmap}, except for a preference to remove measurements over regions that persistently contain clouds or aerosols (e.g., the Sahara) or have low signal to noise ratios (e.g., high latitudes)."

We have added the following to section 5.2:

> "The post-filtering, as was seen in the simulated OCO-2 data, tended to be spatially unbiased relative to the unfiltered dataset (Fig.

> \ref{fig:GOSAT_validation_worldmap}), except for a preference to remove measurements in regions persistently contaminated by clouds or aerosols and measurements at high latitudes that had low signal levels. The Sahara, for example, is entirely devoid of measurements at a throughput of 30\,{{\%}}, likely because of contamination by large dust particles."

**Minor issues**

**P13041, L13 – A better phrase than "Carbon flux models" would be "CO2 inverse modeling systems" since they are talking about a system combining a model and measurements.**

We agree and have modified section 1:

> "\chem{CO_2} inverse modeling systems…"

**P13041, L18 – Should change to "sufficiently high accuracy and precision".**

We agree and have modified section 1:

> "Global coverage of \chem{CO_2} measurements will improve the accuracy and precision of their results, but only if the space-based measurements are of sufficiently high accuracy and precision themselves."

**P13041, L19 – "0.5% (~2 ppm for CO2)" actually corresponds to XCO2 specifically and would not necessarily be the same for satellite CO2 profiles, so this should be changed to "XCO2".**

We agree and have modified section 1:

> "for $X_{\chem{CO_2}}$"

**P13042, L5 – The number of molecules of CO2 or the vertical column density can still be determined without detailed information about the light path. The light path is needed for any type of mole fraction like XCO2.**

We have modified section 1 to specify $X_{CO2}$ instead of molecules of $CO_2$:

> "In order to accurately measure $X_{\chem{CO_2}}$, the length of the light path must be known."

**P13044, L13 – Similar to above, the “CO2 mole fraction”.**

We have modified section 2 to specify $X_{CO2}$ instead of molecules of $CO_2$:

> “The geometry of the light path must be known in conjunction with the magnitude of the absorption in order for $X_{\chem{CO_2}}$ to be accurately estimated.”

**P13044, L21-23 – Another reason (perhaps of equal importance) is that it is theorized that the column-averaged quantity minimizes the impact of vertical transport errors in models, which become more significant at increasing horizontal resolutions.**

That is true, but the reason $X_{CO2}$ is retrieved is because there is little evidence that, with current passive sensor technology, any vertical $CO_2$ information can reliably be retrieved. If hyperspectral near-infrared measurements could give accurate, high-resolution profile information, we would certainly retrieve a profile. Then, if modelers wanted to use $X_{CO2}$ to minimize transport errors, they could calculate it themselves.

**P13045, L15 – It would be better to say “European Centre for Medium Range Forecasting (ECMWF) analyses”.**

We agree and have modified section 2:

> “the European Centre for Medium-Range Weather Forecasts Integrated Forecast System (ECMWF IFS; \citealp{ecmwf_2015}) model fields”

**P13045, L15 – Citations should be included for the ECMWF analyses and the CO2 prior (which sounds like the TCCON CO2 prior).**

We have added the following to section 2:

> “The meteorological priors are taken from the European Centre for Medium-Range Weather Forecasts Integrated Forecast System (ECMWF IFS; \citealp{ecmwf_2015}) model fields and the \chem{CO_2} profile priors are interpolated from the Total Carbon Column Observing Network (TCCON) a~priori \citep{wunch_2011a}.”

We have added the ECMWF IFS documentation to the references:

> “\bibitem[{ECMWF(2015)}]{ecmwf_2015}IFS documentation, available at:

\url{https://software.ecmwf.int/wiki/display/IFS/Official+IFS+Documentation} (last access: February~2016)."

**P13045, L25 – Are the ECMWF surface pressure uncertainties over complicated terrain at the OCO-2 pixel size known? These RMS differences would very likely exceed 1-2 hPa.**

We account for potential surface pressure uncertainties over complex terrain using a standard hypsometric adjustment (Crisp et al., 2010), which accounts for the altitude difference between the satellite field-of-view (FOV) and the model grid box. Because the digital elevation map used to find the mean altitude of the FOV is very accurate, we believe that the surface pressure errors are still small.

**P13046, L8 – The authors should expand on what is meant by "Kahn 2b and 3b", with a couple of words to avoid forcing the reader to consult this reference.**

We agree and have modified section 2:

> "The Kahn 2b aerosol type is a mixture of coarse- and fine-mode dust while the Kahn 3b aerosol type is a mixture of smaller carbonaceous aerosols. Both 2b and 3b also contain sulfate and sea salt components."

**P13047, L2 – Should replace "Afternoon-Train" with "Afternoon Constellation or A-Train".**

We agree and have modified section 3:

> "as part of the Afternoon Constellation \citep{lecuyer_2010}."

**P13048, L1-6 – It would be clearer to put each citation directly after the model description. Furthermore, referring to only one model by its contributor (David Baker) is inconsistent. It was the PCTM model.**

We agree and have modified section 4.1:

> "The \chem{CO_2} models used include two from the University of Edinburgh \citep{feng_2011}, one from Le Laboratoire des Sciences du Climat et de l'Environnement \citep{chevallier_2010}, two from the National Institute for Environmental Studies \citep{maksyutov_2013}, the 2010 version of CarbonTracker \citep{peters_2007}, and the National Oceanic and

> Atmospheric Administration Parameterized Chemistry and Transport Model (NOAA PCTM; \citealp{kawa_2004})."

**P13048, L26-27 – It is important to explicitly state that some clear-sky observations are also removed by the pre-filtering methods, specifically many of those over snow and ice covered surfaces.**

We agree and have modified section 4.2:

> "Clear-sky measurements at high-latitudes and over ice and snow, which typically have unacceptably low signal to noise ratios, are also removed by our pre-filters."

**P13050, L4 – it would be very useful if the author would specify the four rules.**

The four rules (or parameters) selected depend on the retrieval and surface type. The following clarification has been made in section 4.3:

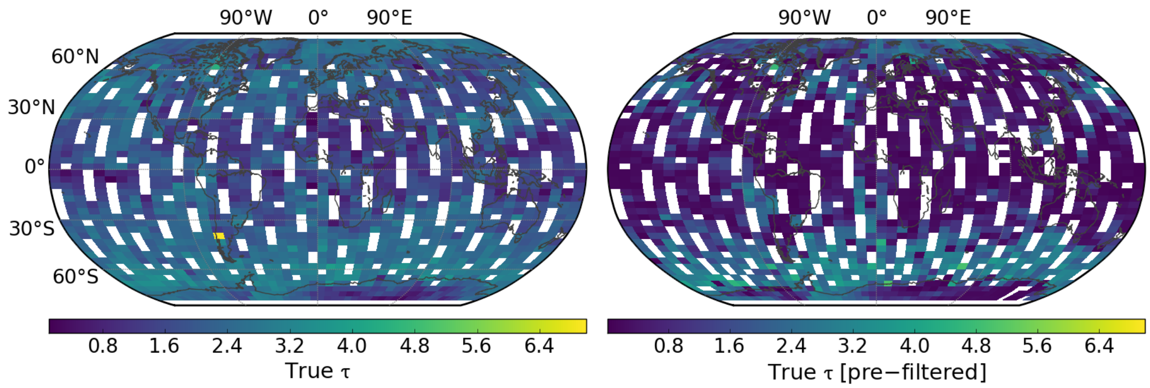
> "it was hypothesized that a~different set of four filtering parameters might be selected for each retrieval and surface type combination."

**P13051, L11-12 – Have tests ever been done for the full-physics algorithm where surface pressure is not retrieved, but rather fixed to ECMWF, to retrieve clouds and aerosols?**

To our knowledge, this has not been tested. However, we suspect that the results wouldn't change, as the clear-sky results are nearly insensitive to the method of assigning surface pressure (fixing at the prior or retrieving using the $O_2$ A-band). Additionally, our tight constraint (1 hPa) on the prior surface pressure significantly limits how far the retrieved surface pressure can vary and thus simply using ECMWF isn't typically a substantial change.

**Figure 6 – The histograms depict very little high optical depth data over land, relative to the data over the oceans. How realistic or representative of reality is this? A clarifying statement in the text would be desirable.**

The histograms in figures 5 and 6 show the data after pre-filtering. Thus, the ABP, in this setup, is better at removing simulated high optical depth scenes over land. The figure below shows the binned total optical depth of all the simulated scenes (left) and the pre-filtered data (right).




We have added the following to section 5.1:

> "Interestingly, the initial number of high optical depth scenes over land is considerably smaller than over ocean. This indicates that the ABP may be more effective at removing high optical depth scenes over land, likely because scattering between cloud or aerosol layers and the surface makes these contaminates more identifiable."

**Figures 7 and 9 - Presentation of the errors on the TransCom regions here is not the best choice in my opinion (which relates to major point #4). These regions mostly correspond to surface vegetation regions over land and simply latitude/longitude boxes for the oceans, with the number of regions chosen based on a plausible target number of degrees of freedom for an in situ CO2 surface flux inversion. Cloud and aerosol effects would have different spatial impacts and likely occur with smaller spatial scale features that are not represented here. (TransCom also ignores the Mediterranean Sea and both missions observe in this region.) It is not explicitly stated that the grey regions have no data. White is a better choice for no data, which would require revising the color bar scheme for low RMS error. More importantly though, showing the RMS errors at some grid box scale (perhaps something like 2 degrees) would be more informative.**

We agree and have updated figures 7 and 9 to show 8.0 x 4.0° bins instead of the TransCom regions. The conclusions remain the same, which provides evidence that the regional analysis is not especially sensitive to the method of spatial binning.

References:

O'Dell, C. W.: Acceleration of multiple-scattering, hyperspectral radiative transfer calculations via low-streams interpolation, J. Geophys. Res., 115, doi:10.1029/2009JD012803, 2010.