The authors would like to thank Referee #2, Dr. Maximilian Reuter, 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.

Introduction: You only discuss full physics algorithms and their 0th order approximation, i.e., scattering is entirely neglected. What about the 1st order approximation using a light path proxy (e.g., O2 as used by Schneising et al. (2012)) to account for scattering. In principle, also a light path proxy method could be set up relying only on clear-sky RT simulations.

Multiple configurations could certainly have been studied, but we limited our work to investigating the use of a retrieval that entirely neglects scattering. We agree that simplified scattering schemes would be valuable to investigate in the future. Regarding an O_2 proxy method, potential complications that may arise include how the light path modification varies from 0.76 to 2.06 μm. Additionally, surface properties often change significantly over that same wavelength range.

Section 4.2 and 4.3: As non-scattering retrievals are not per se inaccurate but only if they are confronted with aerosol and/or cloud contaminated scenes, I expect that the used pre-filtering and DOGO technique critically influence the results presented in this study. With this in view, I have the impression that the manuscript would benefit from more details in the description of the pre-filtering and DOGO technique. i) Which parameters have been used for pre-filtering and DOGO?

Additional information regarding pre-filtering has been added to section 4.2:

> "The ABP works by estimating surface pressure and albedo from the \chem{O_2} A-band assuming no clouds or aerosols are present. If the scene is contaminated by thick clouds or aerosols, the results may have identifiably large residuals between the measured and modeled radiances as well as unrealistic surface pressure and albedo values."

Additionally, while addressing the reviewer responses, we decided to only pre-filter GOSAT using the ABP and to incorporate the IDP results (i.e., the CO_2 and H_2O ratios) into the post-filtering process. This was the methodology we had used for the OCO-2 data and decided it made more sense to let the genetic algorithm choose to apply the IDP parameters to the GOSAT data than to manually pre-filter with them.

A brief description of the parameters DOGO was allowed to select is present beginning on P13049 L20. Additional information has also been added to section 4.3:

"In this study a~genetic algorithm (DOGO; \citealp{mandrake_2013}), which is an optimization tool designed to mimic natural selection to explore high dimensional parameter spaces, was employed to find optimal post-filters for retrievals performed on synthetic OCO-2 measurements and real GOSAT measurements."

ii) What is the throughput of the pre-filtering for GOSAT and OCO2?

Regarding the throughput, pre-filtering generally removes 50-70% of soundings. The exact number depends on which preprocessors are used and precisely how the filters are set, and hence is somewhat arbitrary.

iii) Which are the most important parameters?

Regarding the parameters DOGO selected to filter upon, they vary depending on the retrieval and surface type. We do mention (P13052 L28), however, that for simulated OCO-2 data the IDP variables (the CO_2 and H_2O ratios) are frequently selected. We have also added the following to section 5.1:

"dP_{ABP} and parameters related to the 2.0\,\unit{{\mu}m} \chem{CO_2} band signal were also often selected as filters, as they relate to scattering and absorption by clouds and aerosols."

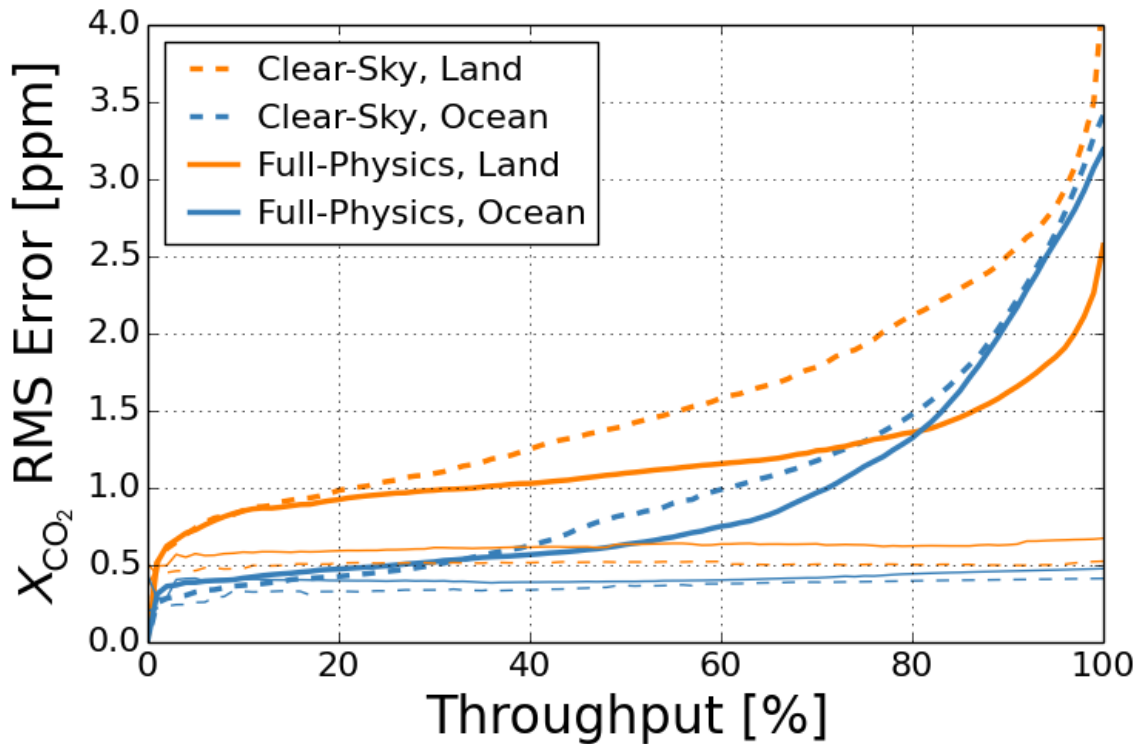Regarding GOSAT, we have added the following to section 5.2:

"Regarding the parameters chosen by DOGO, the primary filters selected were dP_{ABP} and parameters related to the 2.0\,\unit{{\mu}m} \chem{CO_2} band signal. These parameters, as discussed in Sect.~\ref{subsec:oco2errorstatistics}, relate to the light path modification effects of clouds and aerosols."

iv) Why do you assume that DOGO filters out heavily contaminated scenes first (e.g., P13050 L25)? For ocean, Fig.5 may show this but Fig.6 shows that this is not so clear for land. Is this the reason why the FP (full physics) land algorithm outperforms the CS (clear-sky) land algorithm (with simulations)?

We address why DOGO is less effective at removing contaminated scenes over land (Fig. 6) and hypothesize why on P13053 L20-25. As DOGO still optimally reduced the RMS error, we can't say for certain that its failure to remove contaminated scenes leads to a worse performance from the clear-sky retrieval.

Fig.4: i) Please add an estimate for OCO2's RMS errors expected from noise (land and ocean). I expect that it is something in the order of 0.5ppm - 1.0ppm. This value can be used as lowermost estimate when DOGO starts removing "good" soundings.

Below, we have added corresponding thin lines to figure 4 to show the noise as a function of throughput. As you had hypothesized, when the genetic algorithm nears low throughputs, it roughly approaches the limit of random noise over ocean. Over land, however, neither retrieval version is able to reach the approximate noise limit. This suggests that we're underestimating the posterior error in the OCO-2 simulations over land. We believe the sharp decline in RMS error that occurs at extremely low throughputs is when the genetic algorithm begins to filter out random noise and remove "good" soundings.

We have added the following to section 5.1:

> "Additionally, at such low throughputs, the RMS error approaches the error due to simulated instrument noise over ocean (~ 0.35\,\unit{ppm}). Interestingly, the RMS errors over land plateau around 0.80 ppm while the simulated noise limit is ~ 0.50\,\unit{ppm}. This suggests an underestimation of the posterior errors for OCO-2 simulations over land, a~feature seen previously in both simulations \citep{odell_2012} and real GOSAT data \citep{kulawik_2015}."

ii) Please indicate the DOGO throughput used by ACOS for OCO2 operationally. If these are smaller than 15%, CS should be considered as real option to be used in stead of FP, which is not the case for throughputs greater than about 50%. iii) I don't know how large the operationally used throughput values are but if they are larger than 15%, you should note that over land, the differences are quite substantial: if you would consider 1.5ppm as acceptable, FP gains 30% in terms of throughput; for a throughput of 90%, the variance roughly triples.

Approximately 20-30% of all OCO-2 measurements are currently processed. The measurements not processed are those that are certainly contaminated by clouds or aerosols, are over ice or snow, or are over water in nadir mode.

Fig.8: i) As for Fig.4, please add an estimate for GOSAT's RMS errors expected from noise (land and ocean).

We have repeated the previous calculations for the real GOSAT retrievals. Here, the noise appears to be considerably lower than the RMS errors for most retrieval versions except for the full-physics retrieval over ocean. This may be due to larger uncertainty in the noise estimates for real measurements.
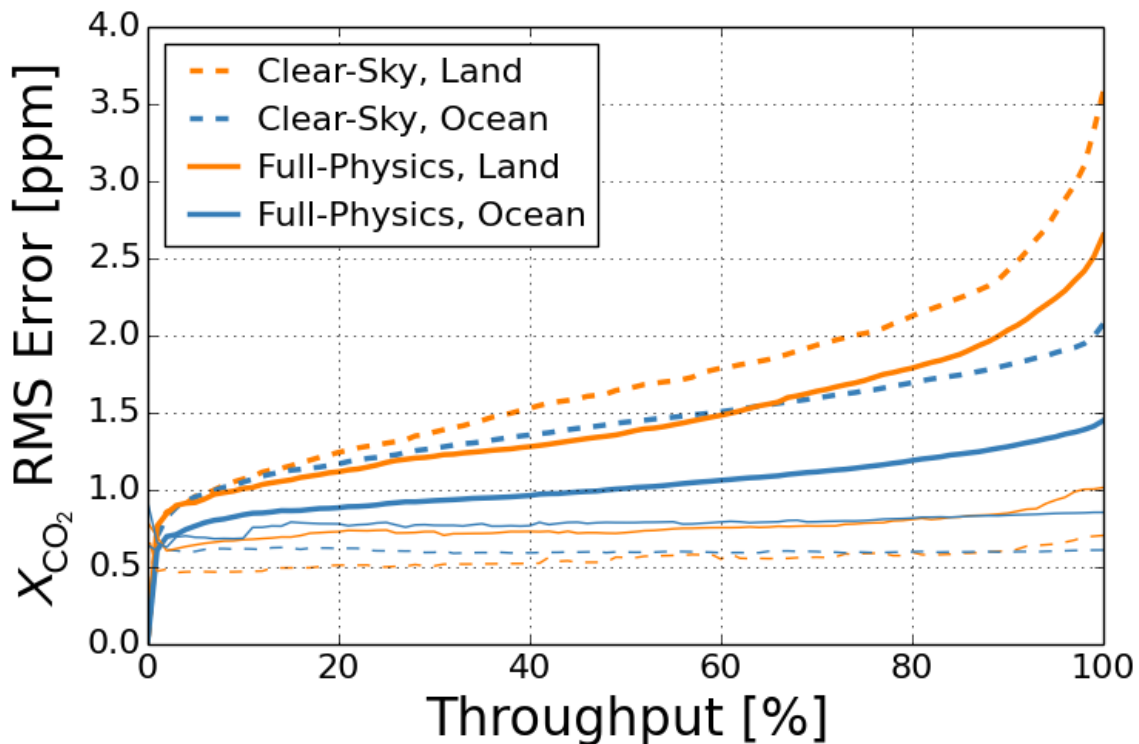
ii) Please indicate the DOGO throughput used by ACOS for GOSAT operationally.

All GOSAT data is processed because the data volume is low enough. We then use DOGO to determine the highest quality retrievals.

Fig.7, Fig.9: In contrast to the FP retrieval, the CS retrieval assumes that the surface pressure is perfectly known from ECMWF. Because of this and because of the additional state vector elements for the scattering properties in the FP retrieval, I would expect, that the FP retrieval diagnoses somewhat larger a posteriori noise for XCO2. If you find similar RMS values this could indicate that biases are larger for CS than for FP. For the application of surface flux inversion, regional biases are much more severe. Therefore, please add maps as Fig.7 and 9 but for scatter and bias.

As seen in the figure above, the posterior noise is indeed larger for the full-physics retrievals. The reviewer is therefore correct in hypothesizing that the GOSAT biases are larger for clear-sky compared to full-physics and thus it is misleading to simply analyze the RMS error alone. Because of this, we have made a slight change to the methodology to include a bias correction for the GOSAT data. Additionally, we have added maps of the bias, scatter, and RMS error for both OCO-2 and GOSAT.

Regarding the OCO-2 simulations (Fig. 7), the differences in RMS error are quite small when filtered to a throughput of 30% and the differences are not driven primarily by the bias or scatter (see figures below). Because of this, we continue to feel that it is unnecessary/unhelpful to apply a bias correction to the OCO-2 data.
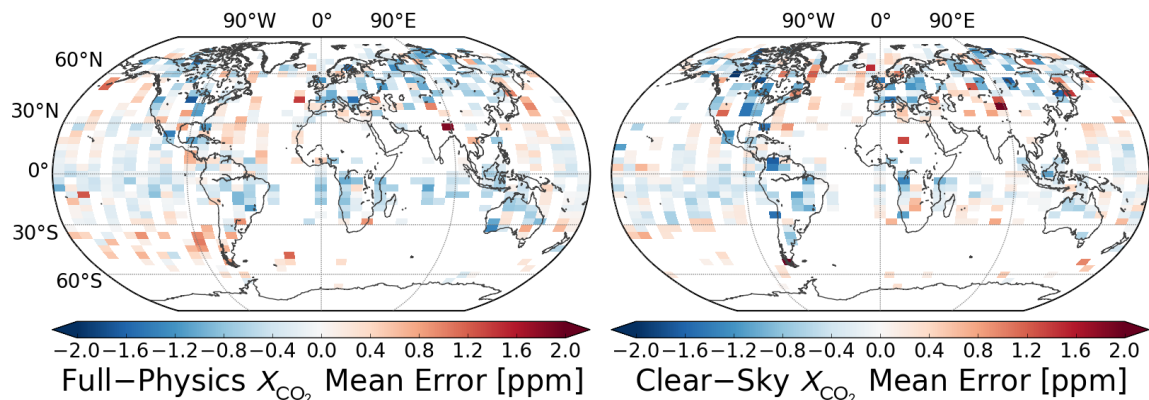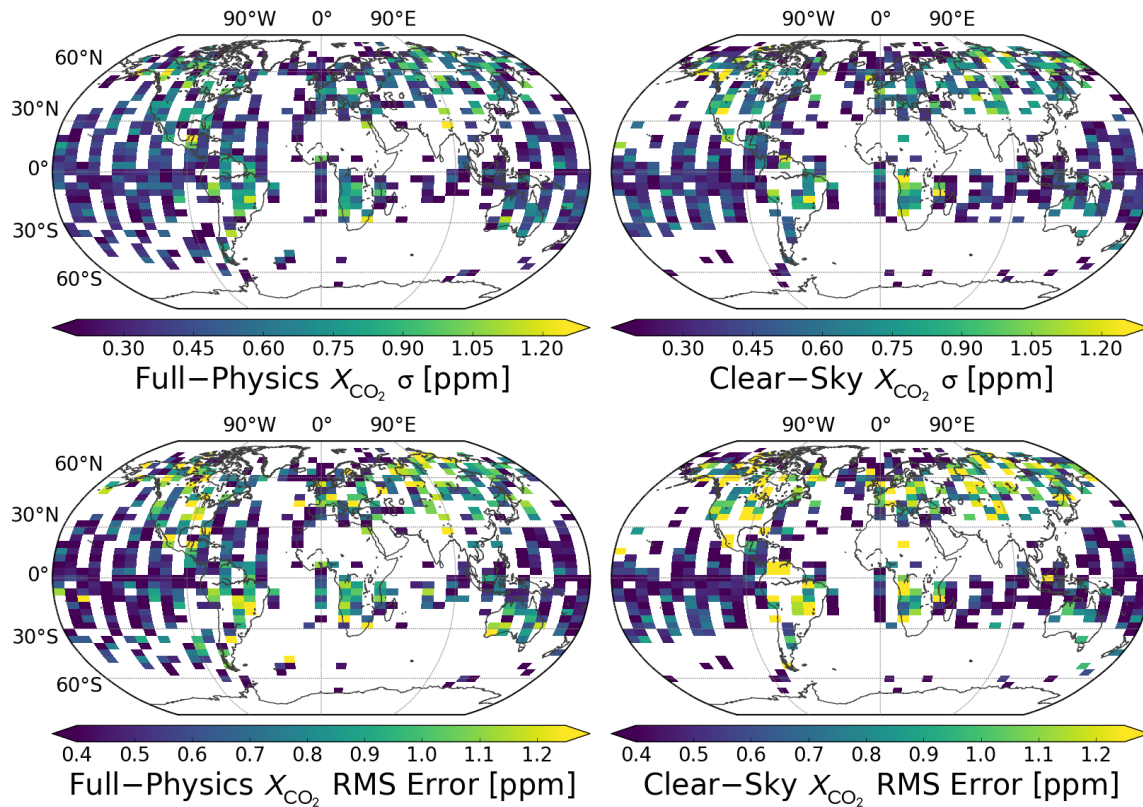
Regarding GOSAT (Fig. 9), however, the larger RMS errors over ocean in the clear-sky retrieval (and somewhat in the full-physics retrieval) are driven by a significant bias:
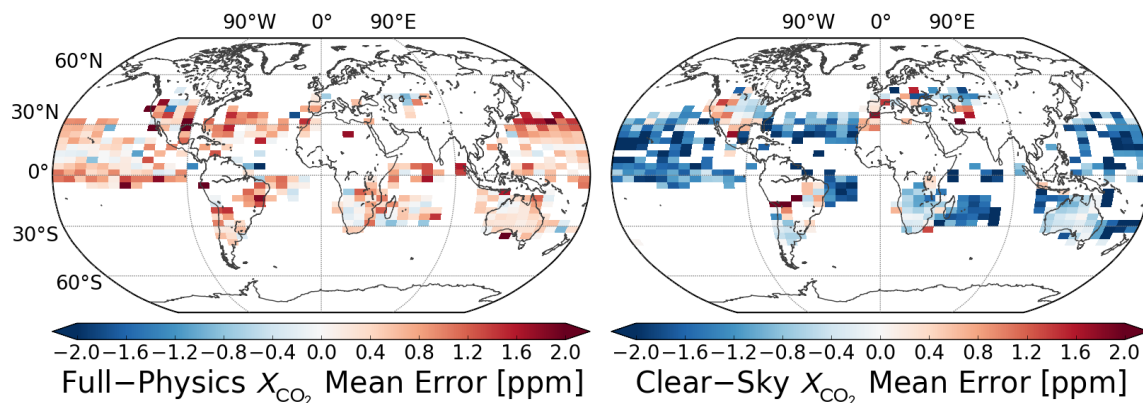
Because of this, we have modified our methodology to include a simple bias correction applied to the GOSAT data after pre-filtering but before post-filtering via DOGO. The data was adjusted using both a slope and an offset:
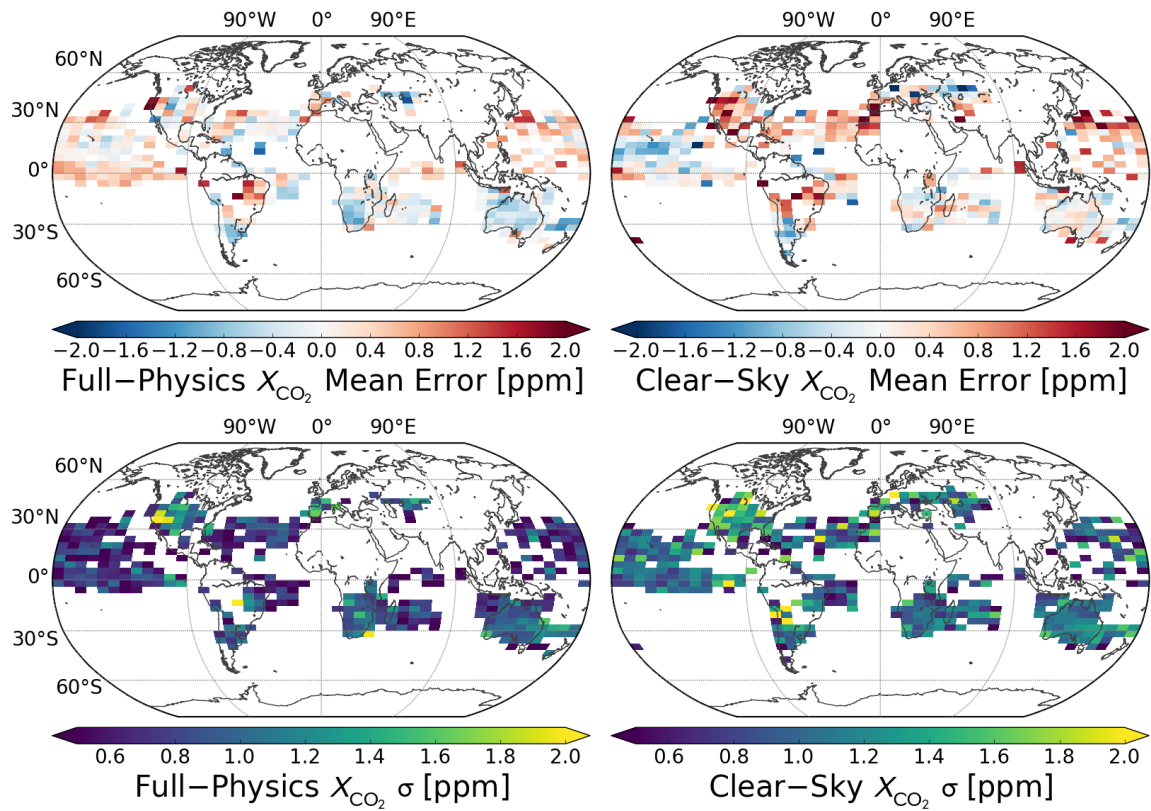
$$X_{CO_2} = X_{CO_2} + A + Bx$$

where A is an offset, x is the variable chosen for bias correction, and B is the slope of a best-fit line through x with respect to the X_{CO2} error. The difference between the ABP-estimated surface pressure and the prior surface pressure (dP_{ABO2}) was used to adjust the slope of the clear-sky retrieval over land, clear-sky retrieval over ocean, and full-physics retrieval over land. The signal ratio of the 2.0 μm CO_2 band to the 1.6 μm CO_2 band was used for the full-physics retrieval over ocean. These parameters were selected because they showed the highest correlation with the X_{CO2} error. The offset was simply the global mean X_{CO2} error. We have added this bias correction information to section 4.3:

> "Prior to post-filtering the data with DOGO, we applied a~simple bias correction, unique to each retrieval and surface type, to the GOSAT $X_{\chem{CO_2}}$ retrievals, similar to what is done for operational GOSAT retrievals \citep{wunch_2011c,guerlet_2013}. This was because the clear-sky GOSAT retrievals over ocean initially contained a~large bias in $X_{\chem{CO_2}}$, which initial testing showed was not entirely removed by DOGO. The $X_{\chem{CO_2}}$ data was adjusted using both a~slope and an offset:
> \begin{align}
> \label{eq:bias_equation}
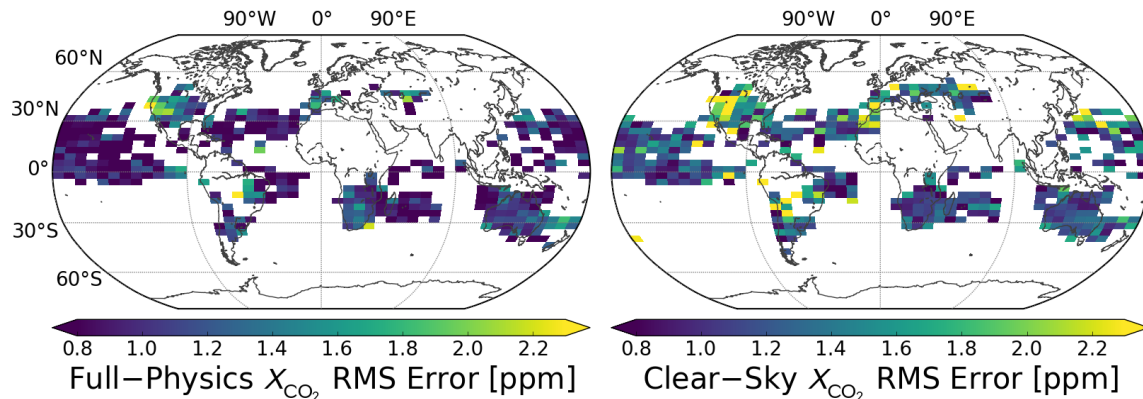> X_{\chem{CO_2}} = X_{\chem{CO_2}} + A + Bx
> \end{align}
> where A is an offset, x is a~single variable chosen for bias correction, and B is the slope of a~best-fit line through x with respect to the $X_{\chem{CO_2}}$ error. We use a~single correction parameter, rather than 2-4 as employed operationally, because the qualitative results of this study are unchanged when employing more bias-correction parameters. The difference between the ABP-estimated surface pressure and the a~priori surface pressure (dP_{ABP}) was selected for the bias correction for the clear-sky retrieval over land, clear-sky retrieval over ocean, and full-physics retrieval over land. The signal ratio of the 2.0\,\unit{{\mu}m} \chem{CO_2} band to the 1.6\,\unit{{\mu}m} \chem{CO_2} band (SR32) was used for the full-physics retrieval over ocean. These parameters were chosen because they showed the highest correlation with the $X_{\chem{CO_2}}$ error. Both dP_{ABP} and SR32 relate to light path modification caused by clouds and aerosols. dP_{ABP} is typically negative and the magnitude is correlated with the amount of light path shortening due to clouds and aerosols. The signal ratio is influenced by the wavelength dependence of clouds and aerosols in the two near-infrared \chem{CO_2} bands."

As we're now applying a bias correction before running the genetic algorithm, figure 8 has been updated appropriately (note the different scale). We now see that both the clear-sky and full-physics ocean retrieval RMS errors are lower than before, as expected, as the initial bias was reduced. The bias correction has only a small influence on the land retrievals, except at high thresholds.
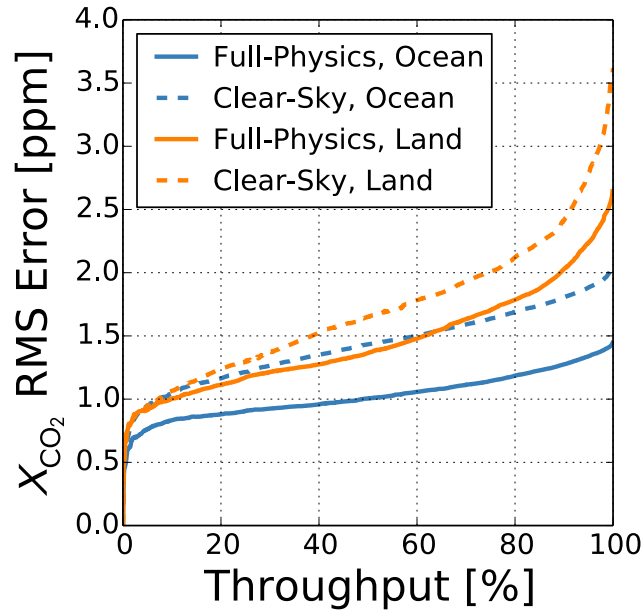
As demonstrated below in the updates for figure 9, the biases over ocean are, as expected, significantly reduced for both retrieval types.

Full−Physics X_{CO_2} Mean Error [ppm]

Clear−Sky X_{CO_2} Mean Error [ppm]

Full−Physics X_{CO_2} σ [ppm]

Clear−Sky X_{CO_2} σ [ppm]

Full−Physics X_{CO_2} RMS Error [ppm]   **Clear−Sky X_{CO_2} RMS Error [ppm]**

As the clear-sky X_{CO2} RMS errors over ocean are still about ∼20-35% larger than the full-physics RMS errors, we believe that the addition of this bias correction does not fundamentally change the conclusions of the paper. However, we have updated the abstract and conclusions to reflect our refined methodology:

> "For real GOSAT data, the clear-sky retrieval had errors 0-20\,{{\%}} larger than the full-physics retrieval over land and roughly 20-35\,{{\%}} larger over ocean, depending on filtration level."

> "For GOSAT measurements over land, the clear-sky retrieval has $X_{\chem{CO_2}}$ RMS errors 0-20\,{{\%}} larger than the full-physics retrieval, when the datasets are filtered to remove scenes contaminated by clouds and aerosols. Over ocean, the clear-sky retrieval has $X_{\chem{CO_2}}$ RMS errors roughly 20-35\,{{\%}} larger than the full-physics retrieval."

P13040 L25,"RMS errors of less than 2.0ppm": i) Relative to which truth? Please specify.

The following has been added to the abstract:

> "In general, the clear-sky retrieval had $X_{\chem{CO_2}}$ root-mean-square (RMS) errors of less than 2.0\,\unit{ppm}, relative to Total Carbon Column Observing Network (TCCON) measurements and a∼suite of \chem{CO_2} models, when adequately filtered through the use of a∼custom genetic algorithm filtering system."

ii) Please add for comparison the corresponding RMS value for the full physics algorithm using the same sample. Is the difference significant?

We feel the comparison of RMS values for clear-sky vs. full-physics should remain in the body of the paper, as it depends on many details including data type, surface

type, and throughput level. A clear-sky RMS error value of less than 2.0 ppm is an approximate but fair assessment of the clear-sky retrieval's performance. The full-physics retrieval does as well as or better than the clear-sky retrieval but, again, it depends on a variety of factors, which are given in detail in the main text.

P13040 L27,"These results imply that...": To my knowledge you primarily find in the literature conclusions on errors of "non-scattering" retrievals confronted with cloud/aerosol contaminated scenes (for clear sky scenes, you would not expect any problems). I.e., it is well known that issues of clear-sky algorithms arise from inabilities of the used filtering method. Therefore, I would suggest to reformulate the sentence to something like: "Our results imply that filtering methods can be found so that ...".

We agree and have modified the abstract:

> "These results imply that non-scattering $X_{\chem{CO_2}}$ retrievals are potentially more useful than previous literature suggests, as the filtering methods we employ are able to remove measurements in which scattering can cause significant errors."

P13041 L3,"certain applications": Which are? Another potential application could be to use a FP (full physics) algorithm only for anchor points at nadir center pixels and use the fast CS (clear-sky) algorithm for the rest of the swath.

We believe the statement in the abstract should be left broad to encompass several possibilities. Additionally, we discuss a couple of these applications in the conclusion (i.e., future missions that make more measurements than can be processed with the full-physics retrieval, using the clear-sky retrieval as a first-guess for the full-physics retrieval)

P13041 L18,"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 L24,"errors and biases": Do you mean "random errors and biases"?

Yes. We have made the following change in section 1:

"to minimize random errors and biases…"

P13042 L22,"non-linearity of the forward model": What is the problem with non-linearity of the forward model? As long as the non-linearity is moderate and the cost function does not have local minima, an iterative algorithm should be able to find the cost function's minimum. Please avoid citing not peer-reviewed literature or at least specify a web link.

It has been demonstrated that the non-linearity in the full-physics algorithm due to scattering by clouds and aerosols is significant and that multiple local minima often exist in the cost function. Thus, the retrieval algorithm is often unable to find the optimal solution. A web link has been provided to the relevant master's thesis (Nelson, 2015):

> "available at: \url{http://hdl.handle.net/10217/166883}"

P13042 L23,"full-physics retrievals may incur biases": In principle this is possible but it probably strongly depends on the specific retrieval set-up and likely on the used constraints for the a priori scattering information. For example, Reuter et al. (2010) have not found strong indications for issues with SCIAMACHY retrievals under cloud free conditions.

We agree that it may be a unique problem of ACOS. Further work is needed to determine if the same issues are present in other retrieval algorithm. We have added the following to section 1:

> "it has been shown that, for certain retrieval algorithm set-ups, these ``full-physics'' retrievals may incur biases…"

P13043 L18,"10 minutes": Is this for one CPU core?

Yes. We have added the following to section 1:

> "roughly 10\,min per measurement per CPU core…"

P13043 L25,"our hypothesis": Please specify what exactly your hypothesis is.

We have added the following clarification to section 1.

> "We begin by testing our hypothesis that clear-sky retrievals may perform as well as full-physics retrievals for sufficiently clear scenes by evaluating

synthetic OCO-2 measurements. We then extend the analysis by testing our hypothesis on real GOSAT measurements."

P13043 L28,"Mandrake, 2015": Please avoid "grey" literature or at least provide a web link.

We have removed the citation.

P13045 L24,"a priori surface pressure from ECMWF": OCO2 footprints are small compared to the ECMWF grid boxes. Do you inter/extrapolate the meteorological profiles to the OCO2 surface height?

We account for potential surface pressure uncertainties 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.

P13045 L26,"We included Rayleigh scattering": Usually, this would mean that you account for multiple scattering in the RT so that you lose any gain in computational efficiency? How have you included Rayleigh without performing multiple scattering RT simulations?

Multiple scattering RT was performed, using the efficient approach described in O'Dell (2010). For future clear-sky retrievals, the Rayleigh scattering calculation could likely be massively accelerated, either using the O'Dell (2010) approach, or a simple single-scattering approximation (or both).

P13046 L19,"Gaussian noise consistent with...": Please explain why you consider it useful to add Gaussian noise. It is well known how Gaussian noise propagates in retrieval noise. The drawback I see here is, that systematic errors (in which you are probably much more interested than in the noise) are potentially less visible because hidden in the noise.

For this study, we wanted the simulated measurements to be as realistic as possible. Additionally, we investigated noiseless simulations and found qualitatively similar results.

P13049 L7,"Osterman et al., 2013": Please provide a web link in the references.

A web link has been provided to the relevant documentation:

> "available at:
> \url{https://co2.jpl.nasa.gov/static/docs/v3.4_DataUsersGuide-
> RevB_131028.pdf}

P13050 L25,"obviously contaminated": Please explain why you find it obvious that DOGO removes the (cloud/aerosol) contaminated scenes. If it is because of Fig.5, please add a reference to the figure. From section 4.3 (description of DOGO), I read that DOGO simply optimizes the RMS relative to an assumed truth by (smart) rejection of outlying soundings. DOGO does not care about the reason for the outliers. See also P13053 L13: "... DOGOs goal is ... not to remove scenes with high optical depths".

Correct. It is in reference to figures 5 and 6 and the corresponding discussion. We have added the following clarification to section 5.1:

> "obviously contaminated scenes, as discussed below and demonstrated in Figs.~\ref{fig:genetic_algorithm_removing_contaminated_ocean_scenes} and~\ref{fig:genetic_algorithm_removing_contaminated_land_scenes}, and is now selecting the best of what remains."

P13051 L5-L12,"The first 20% ...": i) Same as last point; please explain why you think that DOGO eliminates clouds at first. If it is because of Fig.5, please add a reference to the figure.

Correct. It is, again, in reference to figures 5 and 6 and the corresponding discussion. We have added the following clarification to section 5.1:

> "likely contain very thick clouds and aerosols, as discussed below,..."

ii) Which post-processing filters have been used for the FP and the CS algorithm? Reuter et al. (2010) found in simulations poor convergence for scenes with thick clouds. Is the convergence behaviour similar for the FP and the CS retrieval version for the first 20 %? Do you exclude non converging scenes?

We excluded retrievals that did not converge. Most clear-sky retrievals converged, due to the simplicity of the algorithm, while a small number of full-physics retrievals failed to converge. We have added the following to section 3:

> "Finally, OCO-2 and GOSAT retrievals that did not converge were not used for this study."

P13051 L16-L18,"Below 30% throughput...": The same argument should hold for land observations but for land the FP performs better than the CS algorithm. Why?

Full-physics performs about as well as clear-sky for retrievals over land below roughly 20% throughput (Fig. 4). The reason why clear-sky doesn't perform any better than full-physics, like it does over ocean, may be complex scattering between the surface and the small amount of clouds/aerosols remaining.

P13052 L13,"Yellow corresponds ...": Move to figure caption.

We have removed the sentence from section 5.1.

P13053 L9, "as the RMS errors are nearly identical": RMS can be the same for totally different precisions σ and biases Δ: σ = 2, Δ = 0 ⇒ RMS = 2; σ = 0, Δ = 2 ⇒ RMS = 2. Recommendation: rephrase the sentence to something like "...performs about as well as the FP retrieval in terms of RMS. The same is true for precision and bias (not shown)."

We agree and have modified section 5.1:

> "performs about as well as the full-physics retrieval in terms of $X_{\chem{CO_2}}$ RMS error. The same is true for precision and bias (not shown)."

Fig.5: DOGO does not remove the same data points for FP and CS maybe because DOGO uses retrieved quantities as input parameters. Which?

As previously mentioned, the CO_2 and H_2O ratios (from the IDP) and the difference between the ABP-estimated surface pressure and the surface pressure from the retrieval algorithm (dP_{ABP}) are often selected, as they both relate to light path modification caused by clouds and aerosols. As the retrieved surface pressure is fixed for clear-sky but variable for full-physics, this may result in differences between how useful DOGO finds dP_{ABP}. However, we assign a tight uncertainty to the prior surface pressure and thus the difference between clear-sky and full-physics retrieved surface pressure is small so DOGO is given approximately the same information from dP_{ABP} from clear-sky compared to full-physics.

P13054 L9, "... has regional scatter and bias similar ...": Not shown here because you only show RMS comparisons.

We agree and, because of the other additions discussed above, have modified section 5.1:

> "Overall, these simulated results are promising because they demonstrate that the clear-sky $X_{\chem{CO_2}}$ retrieval has global and regional error statistics similar to the full-physics $X_{\chem{CO_2}}$ retrieval."

P13054 L12, "We have shown that ...": i) Fig.4 shows that FP performs always better than CS for throughputs larger than only 15%. Additionally, Fig.7 shows that FP is superior in most regions of the world for a throughput of 30%. If the operationally used throughput is considerably larger than 15% I would recommend to rephrase the sentence in order to not "over-sell" the results.

When this work was done, the throughput was ~10-15%. Now it's higher (20-30%), thanks to improved computational resources. However, in the future there may still be GHG missions with data processing limitations (e.g., CarbonSat), thus this work is still relevant in that sense.

ii) You are not filtering for clouds. DOGO just optimizes the RMS. Fig.6 shows that by far not only cloudy scenes are rejected.

We have modified the beginning of section 5.2:

> "when filtering is employed to remove low quality scenes, including those contaminated by clouds and aerosols."

P13055 L9-L12, "At 100% throughput, it still ...": Another possible explanation could be better a priori CO2 profiles over ocean so that shielding by thick clouds has little effect.

We are unaware of a difference in a priori CO_2 profile accuracy over land and ocean.

P13057 L12, "...more measurements than OCO-2": CarbonSat would be such a mission. Please cite Bovensmann et al. (2010).

We have added the following to the conclusions:

> "such as GeoCarb \citep{polonsky_2014} and CarbonSat \citep{bovensmann_2010}."

References:

Crisp, D., Bösch, H., Brown, L., Castano, R., Christi, M., Connor, B., Frankenberg, C., McDuffie, J., Miller, C. E., Natraj, V., O'Dell, C., O'Brien, D., Polonsky, I., Oyafuso, F., Thompson, D., Toon, G., and Spurr, R.: OCO (Orbiting Carbon Observatory)-2 Level 2 Full Physics Retrieval Algorithm Theoretical Basis, Tech. Rep. OCO D-65488, Tech. rep., NASA Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, version 1.0 Rev 4, available at: http://disc.sci.gsfc.nasa.gov/acdisc/documentation/OCO-2_L2_FP_ATBD_v1_rev4_ Nov10.pdf (last access: January 2012), 2010.

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.