Atmos. Meas. Tech. Discuss., 8, 13039–13072, 2015 www.atmos-meas-tech-discuss.net/8/13039/2015/ doi:10.5194/amtd-8-13039-2015 © Author(s) 2015. CC Attribution 3.0 License.



This discussion paper is/has been under review for the journal Atmospheric Measurement Techniques (AMT). Please refer to the corresponding final paper in AMT if available.

# The potential of clear-sky carbon dioxide satellite retrievals

# R. R. Nelson<sup>1</sup>, C. W. O'Dell<sup>2</sup>, T. E. Taylor<sup>2</sup>, L. Mandrake<sup>3</sup>, and M. Smyth<sup>3</sup>

<sup>1</sup>Colorado State University, Fort Collins, CO, USA

<sup>2</sup>Cooperative Institute for Research in the Atmosphere, Fort Collins, CO, USA

<sup>3</sup>Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA

Received: 2 November 2015 - Accepted: 30 November 2015 - Published: 10 December 2015

Correspondence to: R. R. Nelson (rrnelson@atmos.colostate.edu)

Published by Copernicus Publications on behalf of the European Geosciences Union.

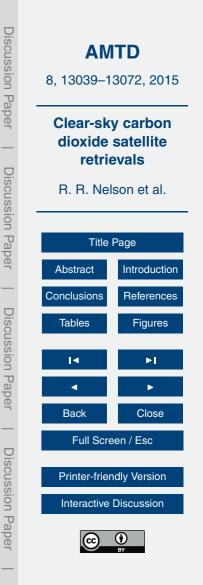
Discussion Pa		<b>AMTD</b> 8, 13039–13072, 2015		
ner   D	Clear-sky carbon dioxide satellite retrievals			
)iscussion	R. R. Nelson et al.			
D	Title	Page		
Đ	Abstract	Introduction		
_	Conclusions	References		
	Tables	Figures		
non		▶1		
Pan	•	•		
	Back	Close		
_	Full Screen / Esc			
SSUDSIC	Printer-friendly Version			
ion F	Interactive Discussion			
Daner	CC D			

## Abstract

Since the launch of the Greenhouse Gases Observing Satellite (GOSAT) in 2009, retrieval algorithms designed to infer the column-averaged dry-air mole fraction of carbon dioxide ( $X_{CO_2}$ ) from hyperspectral near-infrared observations of reflected sunlight have been greatly improved. They now generally include the scattering effects of clouds and aerosols, as early work found that absorption-only retrievals, which neglected these effects, often incurred unacceptably large errors, even for scenes with optically thin cloud or aerosol layers. However, these "full-physics" retrievals tend to be computationally expensive and may incur biases from trying to deduce the properties of clouds and aerosols when there are none present. Additionally, algorithms are now available that can quickly and effectively identify and remove most scenes in which cloud or aerosol scattering plays a significant role.

In this work, we test the hypothesis that non-scattering, or "clear-sky", retrievals may perform as well as full-physics retrievals for sufficiently clear scenes. Clear-sky re-

- <sup>15</sup> trievals could potentially avoid errors and biases brought about by trying to infer properties of clouds and aerosols when none are present. Clear-sky retrievals are also desirable because they are orders of magnitude faster than full-physics retrievals. Here we use a simplified version of the Atmospheric Carbon Observations from Space (ACOS)  $X_{CO_2}$  retrieval algorithm that does not include the scattering and absorption effects of clouds or aerosols. It was found that for simulated Orbiting Carbon Observatory-2 (OCO-2) measurements, the clear-sky retrieval had errors comparable to those of the full-physics retrieval. For real GOSAT data, the clear-sky retrieval had nearly indistinguishable error characteristics over land, but roughly 30–60 % larger errors over ocean, depending on filtration level, compared to the full-physics retrieval. In general,
- <sup>25</sup> the clear-sky retrieval had  $X_{CO_2}$  root-mean-square (RMS) errors of less than 2.0 ppm when adequately filtered through the use of the Data Ordering through Genetic Optimization (DOGO) system. These results imply that non-scattering  $X_{CO_2}$  retrievals are potentially much more accurate than previous literature suggests, when employing fil-

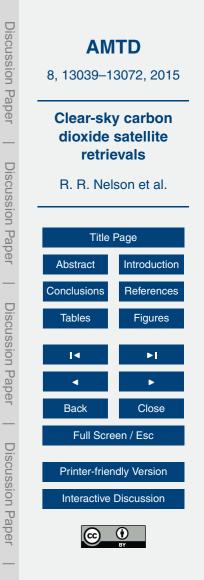


tering methods to remove measurements in which scattering can cause significant errors. Additionally, the computational benefits of non-scattering retrievals means they may be useful for certain applications that require large amounts of data but have less stringent error requirements.

#### 5 1 Introduction

Recently, space-based instruments such as the Greenhouse Gases Observing Satellite (GOSAT; Yokota et al., 2009) and the Orbiting Carbon Observatory-2 (OCO-2; Crisp et al., 2008) have been launched with the goal of providing accurate global measurements of greenhouse gas concentrations, including carbon dioxide  $(CO_2)$ . Only approximately half of the anthropogenically emitted CO<sub>2</sub> stays in the atmosphere. The 10 remaining molecules are absorbed by the land and ocean, but where this absorption takes place is still highly uncertain, especially over land surfaces (Le Quéré et al., 2009). Carbon flux models, designed to answer important guestions about Earth's carbon sources and sinks and their interaction with the atmosphere, are heavily dependent on the density and quality of CO<sub>2</sub> measurements (Rayner and O'Brien, 2001; Baker 15 et al., 2010; Chevallier et al., 2007, 2009). Global coverage of CO<sub>2</sub> measurements will improve the accuracy of their results, but only if the space-based measurements are of sufficiently high accuracy themselves. Specifically, it has been shown that a precision of better than about 0.5% (~ 2 ppm for CO<sub>2</sub>) for space-based measurements is

<sup>20</sup> needed to gain more information about the carbon cycle compared to only having access to ground-based measurements (Miller et al., 2007). In terms of a bias between the measured  $CO_2$  and the true amount present in the atmosphere, even a regional bias of a few tenths of a ppm may be detrimental to carbon flux models (Chevallier et al., 2007; Basu et al., 2013). Thus, it is critically important to minimize errors and biases in satellite measurements of  $CO_2$  in order to be able to correctly answer questions about the carbon cycle.



When retrieving the column-averaged dry-air mole fraction of carbon dioxide, or  $X_{CO_2}$ , from space-based instruments, one of the primary issues is the presence of clouds and aerosols. These contaminants can introduce large errors into a retrieval because they tend to modify the light path in ways that are difficult to quantify. In order

- <sup>5</sup> to accurately measure the number of molecules of  $CO_2$  in the column of air, the length of the light path must be known. If clouds and aerosols are present they can scatter the reflected sunlight in multiple directions, which can drastically alter the length of the light path seen by the sensor and result in significant errors when calculating  $X_{CO_2}$ . Neglecting scattering when measuring scenes containing clouds and aerosols can lead
- to substantial errors in  $X_{CO_2}$ . These errors are often in excess of 1 % (~ 4 ppm of CO<sub>2</sub>) and can be tens of ppm for scenes with thick cloud or aerosol layers (O'Brien and Rayner, 2002; Aben et al., 2007; Butz et al., 2009).

A common method used to avoid these large  $X_{CO_2}$  errors caused by light path modification is to parameterize clouds and aerosols explicitly within the  $X_{CO_2}$  retrieval. This

- often includes adding one or more scattering particle types to the retrieval along with variables that describe their optical and/or physical properties (e.g. optical depth, height of scattering layer, single scatter albedo) (Butz et al., 2009; Yokota et al., 2009; Crisp et al., 2010; Reuter et al., 2013; Parker et al., 2011). These particle types and corresponding properties are intended to represent typical clouds and aerosols found in
- the atmosphere. However, adding cloud and aerosol parameters to the retrieval algorithm can result in new issues such as creating an under-constrained problem or inducing nonlinearity in the forward model (Nelson, 2015). Further, it has been shown that these "full-physics" retrievals may incur biases from attempting to account for clouds and aerosols when none are present (O'Dell et al., 2012). For ideal, extremely clear
- <sup>25</sup> scenes, this becomes an issue because the addition of a cloud and aerosol parameterization may be detrimental rather than beneficial. Additionally, a comparison of retrieved full-physics optical depths from build 3.4 (B3.4) of the NASA Atmospheric CO<sub>2</sub> Observations from Space (ACOS)  $X_{CO_2}$  retrieval algorithm (O'Dell et al., 2012; Crisp et al., 2010) to optical depths measured from the highly accurate AErosol RObotic

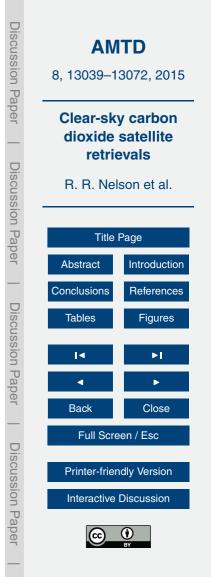


NETwork (AERONET; Holben et al., 1998), shown in Fig. 1, revealed that the aerosol optical depths retrieved by the full-physics retrieval were not well correlated with the AERONET measurement for a particular scene.

The aforementioned problems associated with the full-physics retrieval algorithm motivated a study of a simplified non-scattering, or "clear-sky", retrieval to test the hypotheses that it could provide comparably accurate  $X_{CO_2}$  measurements, given appropriate filtering of scenes contaminated by clouds and aerosols. These clear-sky retrievals are simple and highly linear because they assume no scattering or absorption effects caused by clouds or aerosols. Thus, clear-sky retrievals may avoid introducing unwanted biases when clouds and aerosols are not present. Recent work by Butz et al. (2013) has demonstrated that, for simulated measurements over ocean, a clear-sky retrieval can theoretically be used when scenes containing significant light path per-

turbations are removed. Correspondingly, this approach is now used in the operational RemoTeC retrieval (Guerlet et al., 2013; Butz et al., 2009).

- <sup>15</sup> Clear-sky retrievals are also desirable because of their high computational efficiency relative to full-physics retrievals. This is primarily because of the computational expense associated with calculating scattering from clouds and aerosols. The current operational ACOS algorithm takes roughly 10 min per measurement and OCO-2 collects about 10<sup>6</sup> measurements per day. This restricts the number of measurements
- able to be fully processed. The use of a clear-sky retrieval would thus, with current computational limits, allow for approximately one to two orders of magnitude more data to be processed. Additionally, the use of clear-sky retrievals would make it possible to perform repeated tests on large sets of data in significantly less time than if full-physics retrievals were used.
- We begin by testing our hypothesis on simulated OCO-2 measurements then extend our analysis to real GOSAT measurements. We use various pre-filtering techniques to remove scenes obviously containing clouds and aerosols and employ the Data Ordering through Genetic Optimization (DOGO) system (Mandrake et al., 2013; Mandrake, 2015) to filter out additional contaminated measurements and improve the quality of



the data. Global and regional statistics are calculated for retrievals over both land and ocean surfaces.

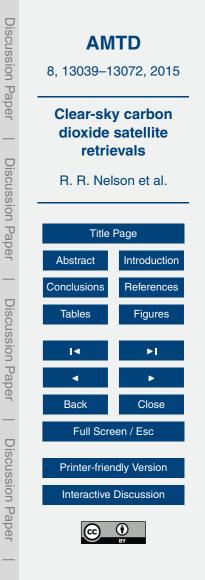
Section two gives details on the full-physics and clear-sky  $X_{CO_2}$  retrievals. The third section discusses the simulated OCO-2 and real GOSAT datasets used in this study.

<sup>5</sup> The fourth section describes the pre- and post-filtering techniques used to remove scenes containing clouds and aerosols. Section five contains a comparison of clear-sky  $X_{CO_2}$  retrievals to full-physics  $X_{CO_2}$  retrievals. The sixth section summarizes the study's results and draws conclusions about the utility of clear-sky  $X_{CO_2}$  retrievals.

# 2 Full-physics vs. clear-sky X<sub>CO2</sub> retrievals

Hyperspectral measurements of reflected sunlight in the near-infrared can be used to infer CO<sub>2</sub> concentrations from space by analyzing molecular absorption. The geometry of the light path must be known in conjunction with the magnitude of the absorption in order for CO<sub>2</sub> to be accurately estimated. The instruments onboard GOSAT and OCO-2 make use of this method. Typically, a relatively weak CO<sub>2</sub> absorption band <sup>15</sup> located in the near-infrared at approximately 1.6 µm and a stronger CO<sub>2</sub> absorption band at 2.0 µm are used in conjunction to estimate the average amount of CO<sub>2</sub> in the light path seen by the instrument's sensors. Additionally, an oxygen absorption feature near 0.76 µm, known as the O<sub>2</sub> A-band, is often employed to help filter out clouds and aerosols (Taylor et al., 2012, 2015) (see Sect. 4.2) and to retrieve surface pressure, which acts as a proxy for light path length.

Because current methods for passively measuring  $CO_2$  are unable to give much information about the vertical distribution of  $CO_2$  (Connor et al., 2008), a column-averaged value is typically the final product retrieved from the measurement. This value is specifically known as the column-averaged dry-air mole fraction of carbon dioxide,



or  $X_{CO_2}$ :

20

25

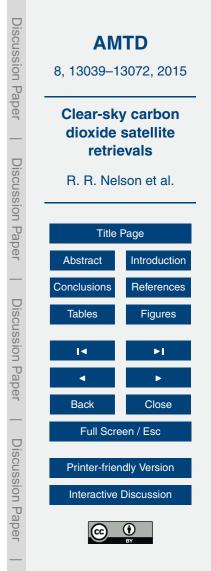
$$X_{\rm CO_2} = \frac{\int_0^\infty N_{\rm CO_2}(z) dz}{\int_0^\infty N_{\rm d}(z) dz}$$

where  $N_{CO_2}(z)$  is the molecular number density of  $CO_2$  with respect to dry air at altitude z and  $N_d(z)$  is the molecular number density of dry air at altitude z.

- <sup>5</sup> Values of  $X_{CO_2}$  are estimated by the ACOS retrieval algorithm using a priori information along with measured radiances to optimize a state vector (Rodgers, 2000). Complete details of the full-physics retrieval algorithm can be found in the ACOS retrieval Algorithm Theoretical Basis Document (Crisp et al., 2010). The parameters selected for inclusion in the state vector are sensitive to the measured radiances and often rep-
- resent physical quantities. Details on many of the elements can be found in O'Dell et al. (2012). Of note, the ACOS retrieval algorithm build used in this study (B3.4) contains up to 46 state vector elements, 20 of which are a vertical CO<sub>2</sub> profile that is used to calculate  $X_{CO_2}$ . The a priori state vector and its corresponding error covariance matrix are derived from multiple sources. The meteorological priors are taken from the Eu-
- ropean Centre for Medium-Range Weather Forecasts (ECMWF) and the CO<sub>2</sub> prior is estimated using zonally-averaged seasonal cycles coupled with a typical atmospheric growth rate.

In this work, we performed full-physics and clear-sky  $X_{CO_2}$  retrievals on both simulated OCO-2 and real GOSAT measurements. Some details of both retrievals are given in Table 1.

The clear-sky retrieval utilizes the CO<sub>2</sub> near-infrared bands at 1.6 and 2.0  $\mu$ m but does not include cloud or aerosol parameters in the state vector. Instead of using the O<sub>2</sub> A-band to retrieve surface pressure, which is used to estimate N<sub>d</sub>, the clear-sky retrieval uses the a priori surface pressure from ECMWF, which has been shown to be accurate to within 1–2 hPa under most conditions (Salstein et al., 2008; Crowell et al., 2015). We included Rayleigh scattering by air molecules for the two near-infrared CO<sub>2</sub> bands, but these effects are likely negligible at such long wavelengths.



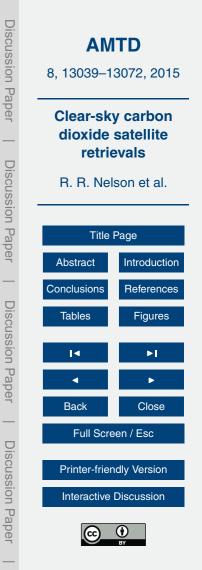
(1)

The full-physics retrieval uses the two near-infrared CO<sub>2</sub> bands as well as the O<sub>2</sub> A-band at 0.76 μm. The O<sub>2</sub> A-band is more sensitive to small cloud and aerosol particles and therefore its inclusion can improve the measurement of cloud and aerosol parameters in the full-physics retrieval. The ACOS B3.4 retrieval parameterizes scattering effects by including four unique cloud and aerosol types in its state vector (O'Dell et al., 2012), which are assumed to have a vertical Gaussian distribution and are assigned a magnitude, width, and height. Two of the four types are a generic water cloud and ice cloud. The remaining two types are the Kahn 2b and 3b aerosol types (Kahn et al., 2001). Simulations suggested that a combination of these four scattering types or OCO-2 (O'Dell et al., 2012).

#### 3 Data

The simulated OCO-2 dataset consists of retrievals performed on approximately 44 000 synthetic measurements spanning 17–18 June 2012 and 19–20 December 2012 (a to tal of 58 orbits), providing a full range of solar and satellite geometries. Scenes over land used nadir viewing geometry, while those over ocean used glint viewing geometry. The simulated radiances were generated by the Colorado State University (CSU) Orbit Simulator, which uses realistic surface, meteorology, and cloud and aerosol distributions (O'Brien et al., 2009). Gaussian noise consistent with the actual OCO-2 instrument noise (Frankenberg et al., 2015) was added to the synthetic measurements to make the retrievals as realistic as possible. For this study, National Centers for Environmental Prediction (NCEP) reanalysis data was used for the retrieval a priori while ECMWF short-term (0–9 h) forecast data was used to create the synthetic radiances. This intentional mismatch in meteorology mimics real-world inaccuracies when mea-

<sup>25</sup> suring a given scene. The vertical profiles of clouds and aerosols used to create the synthetic measurements were derived from the Cloud–Aerosol Lidar with Orthogonal Polarization (CALIOP) instrument onboard the Cloud–Aerosol Lidar and Infrared



Pathfinder Satellite Observations (CALIPSO; Winker et al., 2009), which currently flies in approximately the same polar orbit as OCO-2 as part of the Afternoon-Train (L'Ecuyer and Jiang, 2010). Land surfaces included albedos and bi-directional reflectance distribution functions (BRDF) from the Moderate-resolution Imaging Spectroradiometer (MODIS), while ocean surfaces were modeled as specular reflectors (Cox and Munk, 1954) with a foam component based on wind speed taken from ECMWF.

The GOSAT dataset contained retrievals on 25 000 real measurements made from April 2009 to December 2012. We included both ocean and land scenes and attempted to represent the majority of surface types across the globe without being regionally biased. Measurements were only included in the dataset if they had a corresponding  $X_{CO_2}$  validation source (see Sect. 4.1).

10

15

In this section we discuss methods of characterizing the  $X_{CO_2}$  errors in the retrievals, the pre-filtering used to initially remove heavily contaminated measurements, and post-filtering through the use of the DOGO system to further improve the quality of the dataset by removing additional contaminated scenes.

# 4.1 X<sub>CO2</sub> validation sources

To evaluate the accuracy of the  $X_{CO_2}$  retrievals, a "true"  $X_{CO_2}$  value was needed. For the OCO-2 retrievals, the truth was known because the measurements were synthetically created. For the GOSAT retrievals, we considered the truth to be either a Total Carbon Column Observing Network (TCCON) measurement (Wunch et al., 2011a) co-located in time and space using the technique described in Guerlet et al. (2013) or the average of seven CO<sub>2</sub> models that assimilate ground-based and aircraft CO<sub>2</sub> measurements and were required to agree within 1.0 ppm for a given GOSAT measurement location and time. This ensured we had sufficient ocean validation because TCCON stations



are mostly concentrated on large land masses. The  $CO_2$  models used include two from the University of Edinburgh, one from Le Laboratoire des Sciences du Climat et de l'Environnement, two from the National Institute for Environmental Studies (NIES), the 2010 version of CarbonTracker, and one from David Baker of the National Oceanic

- <sup>5</sup> and Atmospheric Administration (NOAA) (Feng et al., 2011; Chevallier et al., 2010; Maksyutov et al., 2013; Peters et al., 2007; Baker et al., 2010). Both methods of  $X_{CO_2}$ validation have limitations but we believe they are still useful in evaluating retrieval algorithm performance. Figure 2 shows the global distribution of the 25 000 GOSAT measurements used in this study and whether the true  $X_{CO_2}$  was a model consensus or a TCCON measurement. Figure 3 shows the location of the 44 000 simulated OCO-2
- <sup>10</sup> or a TCCON measurement. Figure 3 shows the location of the 44 000 simulated OCO-2 measurements.

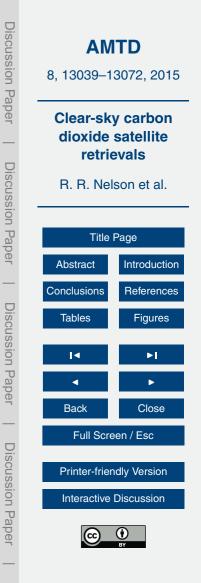
#### 4.2 Pre-filtering

In order to remove measurements heavily contaminated by clouds or aerosols, the OCO-2 and GOSAT datasets were pre-filtered using the O<sub>2</sub> A-band preprocessor (ABP) (Taylor et al., 2012; O'Dell et al., 2012). The ABP was run on all the measurements because it is extremely fast and computationally inexpensive. To further identify contaminated scenes, the Iterative Maximum A-Posteriori Differential Optical Absorption Spectroscopy (IMAP-DOAS) preprocessor (IDP) was also applied to the GOSAT measurements (Frankenberg et al., 2005; Taylor et al., 2015). The IDP estimates CO<sub>2</sub>

and H<sub>2</sub>O from both the strong and weak CO<sub>2</sub> bands independently using a fast, non-scattering algorithm. Deviations from unity in the ratio of the weak to strong band values allows for the identification of many scenes containing clouds or aerosols.

Because we did not have IDP results available for our OCO-2 simulations at the time of this study, similar parameters to those used in the preprocessor were included in the

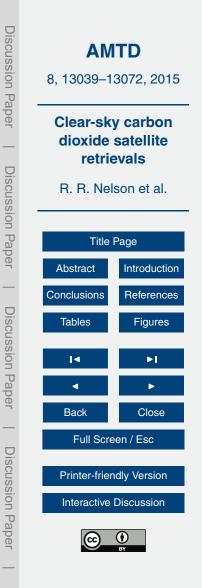
DOGO filter-generating process (see Sect. 4.3) so that the OCO-2 dataset could also benefit from this screening technique. The use of the ABP and IDP removes most, but not all, scenes contaminated by clouds and aerosols (Taylor et al., 2015).



## 4.3 DOGO: Data Ordering through Genetic Optimization

While the pre-filtering techniques we employed are effective at removing many scenes containing clouds and aerosols, other tools are needed to identify very clear scenes that are nearly free of cloud and aerosol contamination on which we believe  $X_{CO_2}$  re-

- trievals will be most accurate. One method used to filter ACOS B3.4 data was a suite of approximately 18 tests to identify scenes of the highest quality, which tend to be the most clear (Osterman et al., 2013). In this study, the same genetic algorithm used to create OCO-2 Warn Levels (Mandrake et al., 2013; Mandrake, 2015) was employed to find optimal post-filters for both simulated OCO-2 and real GOSAT data. Unpublished
- studies have shown that this approach yields similar results to the 18-parameter handtuned filter, but with far fewer filtering parameters necessary. Additionally, while the hand-tuned procedure requires trial and error to find the best possible filters, DOGO is automated and can quickly find an optimum filter set with minimal hand-tuning required. For this work, DOGO attempted to find variables that, when used to filter a dataset, min-
- <sup>15</sup> imized the root-mean-square (RMS) of the  $X_{CO_2}$  error, where the  $X_{CO_2}$  error is defined as the difference between the retrieved  $X_{CO_2}$  and the true  $X_{CO_2}$  (described in Sect. 4.1). We chose this parameter because it instructs DOGO to remove outliers as well as reduce the overall bias, as the formula for RMS error includes both variance and bias. We also investigated applying a custom bias correction to the data prior to determining an
- <sup>20</sup> optimal filter set, but found similar results. The parameters allowed for selection were all derived from the near-infrared measurements themselves, e.g. band signal levels, signal to noise ratios, retrieved surface pressure. The algorithm was not allowed to select certain variables, such as the validation  $X_{CO_2}$  or CALIPSO measurements (used to create the OCO-2 simulations). This was to ensure that DOGO did not "cheat" by hav-
- <sup>25</sup> ing access to external information. Filtering was done for different "throughputs", which equal the percent of data retained after filtering. The DOGO system can also use more than one "rule", or filtering variable, when determining an optimal filtering strategy. That is, one rule selects the single most effective parameter at minimizing the  $X_{CO_2}$  RMS er-



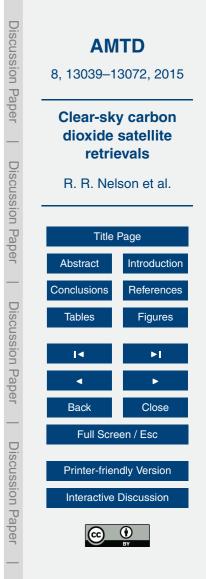
ror while two rules selects the best combination of two parameters in reducing the error. A larger number of rules results in a greater reduction in error for a given throughput, but typically only two to five rules are needed to maximize this reduction (Mandrake et al., 2013). In this study we used four rules and found that increasing the number of rules did not further reduce errors. DOGO was run separately for both clear-sky and full-physics retrievals as well as for land and ocean surfaces because it was hypothe-sized that different filtering parameters might be selected for each retrieval and surface type combination.

# 5 X<sub>CO2</sub> retrieval comparison

- In the previous section we described our use of pre-filtering with the ABP and IDP and post-filtering with DOGO to produce a dataset with biases and outliers minimized to the greatest extent possible. Here we apply and evaluate our technique and examine the performance of the clear-sky X<sub>CO2</sub> retrieval compared to the full-physics X<sub>CO2</sub> retrieval for simulated OCO-2 and real GOSAT data. We evaluate the effectiveness of DOGO at reducing RMS errors, investigate the impact of cloud and aerosol optical depths on the OCO-2 simulations, and examine regional biases in the datasets by binning retrievals into the standard Atmospheric Tracer Transport Model Intercomparison
  - Project-3 (TransCom-3) regions (Gurney et al., 2003).

# 5.1 Summary of OCO-2 error statistics

<sup>20</sup> We begin by evaluating the performance of the clear-sky retrieval on OCO-2 simulations. Figure 4 demonstrates the effectiveness of DOGO at reducing  $X_{CO_2}$  RMS errors as a function of throughput. The initial reduction in error at high throughputs is dramatic, as the algorithm is easily able to identify and remove highly contaminated scenes. All datasets begin to plateau at approximately 50–80 % throughput as DOGO has already <sup>25</sup> removed the obviously contaminated scenes and is now selecting the best of what re-



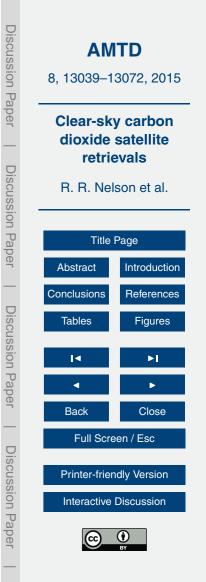
mains. At very low throughputs, there are not enough measurements for the algorithm to function properly.

Over ocean, the clear-sky retrieval (dashed blue line) performs nearly as well as or better than the full-physics retrieval (solid blue line) at all throughputs, never having an  $X_{CO_2}$  RMS error more than 0.25 ppm worse than the full-physics dataset. The first 20 % of scenes filtered out by DOGO (from 100–80 % throughput) likely contain very thick clouds and aerosols such that both the full-physics and clear-sky retrievals produce large RMS errors because neither is able to account for such severe light path modifications. It is interesting that the full-physics retrieval performs just as poorly as the clear-sky retrieval, even though in principle it accounts for the presence of clouds and aerosols. This suggests that, over ocean, the full-physics retrieval struggles to properly quantify the light path modifications of clouds and aerosols. From 80–30 % throughput,

there are still many scenes containing a non-trivial amount of contamination due to clouds and aerosols. Thus, the full-physics retrieval, which has state vector elements designed to handle these scenarios, outperforms the clear-sky retrieval, which is unable to account for any scattering or absorption by even thin clouds and aerosols. Below 30 % throughput, however, the scenes become pristine enough that the clear-sky retrieval has ~ 10 % smaller  $X_{CO_2}$  RMS errors than the full-physics retrieval. It's likely that

the full-physics retrieval struggles slightly compared to the clear-sky retrieval because
 it's trying to parameterize nonexistent clouds and aerosols and thus has too many degrees of freedom. This result agrees with Butz et al. (2013), who found that simulated measurements containing light path perturbations caused by clouds and aerosols can be identified and removed, thus allowing a clear-sky retrieval to perform well.

Over land, Fig. 4 shows that the clear-sky retrievals (dashed orange line) consistently have higher RMS errors than the full-physics retrievals (solid orange line) at high throughputs, upwards of a difference of 1.5 ppm, or ~ 50 % larger. At these higher throughputs, most scenes contaminated by clouds and aerosols remain and the fullphysics retrieval performs better, consistent with expectations. The clear-sky retrieval is unable to account for the complex multiple-scattering effects caused by cloud and

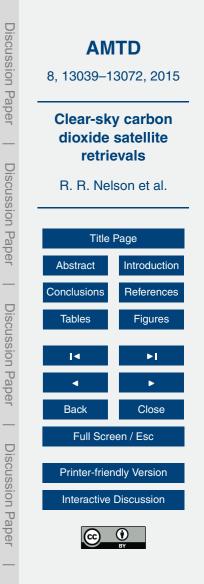


aerosol layers as well as their interaction with the surface. However, the  $X_{CO_2}$  RMS errors become more comparable at lower throughputs, with the clear-sky retrieval error coming within a tenth of a ppm of the full-physics retrieval error at a throughput of 30%. This is consistent with our hypothesis that when scenes contaminated by clouds

or aerosols are removed, the clear-sky retrieval can perform as well as the full-physics retrieval. However, to demonstrate this explicitly, we must show that DOGO is indeed filtering out those scenes containing appreciable amounts of clouds and aerosols.

Because we know the true profiles of clouds and aerosols used to create these simulated OCO-2 measurements, this is straightforward. Figure 5 shows the binned  $X_{CO_2}$ 

- <sup>10</sup> RMS error vs. the true total optical depth (the sum of the true aerosol, ice cloud, and water cloud optical depths from CALIPSO, used to create the synthetic measurements) for clear-sky retrievals (left panel) and full-physics retrievals (right panel) over ocean. Yellow corresponds to 100 % throughput, green to 80 % throughput, and blue to 30 % throughput. At 100 % throughput, there are a significant number of thick ( $\tau > 1.0$ ) cloud
- <sup>15</sup> or aerosol scenes present, along with a secondary peak of thinner cloud or aerosol scenes. These thick scenes are primarily water clouds near the surface that the ABP was unable to identify and remove. In general, the 100, 80, and 30 % throughput  $X_{CO_2}$  RMS errors for clear-sky and full-physics retrievals over ocean are nearly equivalent, which agrees with our analysis of the ocean retrievals plotted in Fig. 4. For high optical
- <sup>20</sup> depth scenes at 100% throughput, the RMS error of the data is large (over 8 ppm for both retrieval types). This indicates that, as hypothesized, both retrieval types have large errors for scenes containing thick cloud or aerosol layers. Going from 100% throughput (yellow) to 80% throughput (green), DOGO greatly reduces the number of these high optical depth scenes, which corresponds to the steep initial decline of
- the ocean retrieval RMS errors in Fig. 4. This is impressive because, as was explained in Sect. 4.3, DOGO is not allowed to use the true optical depth as a filter, indicating that it is using other parameters to infer the amount of clouds and aerosols in a given scene. For example, the IDP filters are often selected by DOGO, which indicates that they are functioning as anticipated and properly identifying scenes containing clouds



and aerosols. At 80% throughput, the full-physics retrieval has slightly smaller RMS errors than the clear-sky retrieval for scenes containing a moderate amount of clouds or aerosols ( $0.1 < \tau < 1.0$ ). This supports our hypothesis that the full-physics retrieval's parameterization of clouds and aerosols is helpful for these types of scenes and that

- the clear-sky retrieval struggles because it is unable to account for the light path modifications caused by these contaminants. When only 30% of the ocean data remains (blue histogram), primarily low optical depth scenes ( $\tau < 0.3$ ) remain and the clear-sky retrieval performs about as well as the full-physics retrieval in terms of both precision and bias, as the RMS errors are nearly identical. Additionally, the precision and bias
- were analyzed separately (not shown) and found to be similar. One might think that even slightly contaminated scenes ( $\tau \sim 0.1-0.3$ ) should have been removed for the 30 % throughput case, but DOGO's goal is to minimize the  $X_{CO_2}$  RMS error, not simply to remove scenes with high optical depths.
- A similar analysis was performed for land scenes, with the results displayed in Fig. 6. As was seen over ocean in Fig. 5, at larger throughputs, higher optical depths correspond to larger RMS errors in the  $X_{CO_2}$  data. The clear-sky retrieval struggles with these high optical depth scenes, but also has relatively large RMS errors (~ 3 ppm) for moderate to small optical depths. As the throughput decreases, the RMS error becomes smaller and more uniform over the entire range of optical depths for both retrieval types. Interestingly, at a throughput of 30 % some high optical depth scenes ( $\tau > 1.0$ ) for the clear-sky retrieval over land still remain. However, the RMS error is still optimally reduced by DOGO and only ~ 10 % worse than the full-physics retrieval. For the full-physics dataset, DOGO chooses to remove nearly all of these thick cloud or aerosol scenes. This suggests that the clear-sky retrieval is less sensitive to some high optical depth scenes over land, perhaps due to complex light path cancellation effects.

In addition to the statistical analysis of the full dataset, spatial errors in the OCO-2 retrieval datasets were analyzed to see if regional variability existed and if there were regions where the clear-sky retrieval had relatively small errors compared to the full-physics retrieval. These regional  $X_{CO_2}$  RMS errors for the simulated OCO-2 datasets



are shown in Fig. 7. Here we use a throughput of 30 %, where the globally averaged clear-sky RMS errors are approximately equivalent to the full-physics errors over ocean and slightly larger (0.1 ppm) over land (as shown in Fig. 4). For both retrieval types there is limited regional variability in RMS error over ocean and modest variability (a few

tenths of a ppm) over land. The mean TransCom RMS errors are nearly indistinguishable over ocean while the clear-sky retrieval has errors less than 30 % worse compared to the full-physics retrieval for most land regions. The magnitude of these differences is usually a fraction of a ppm. These simulated results are promising because they demonstrate that the clear-sky retrieval has regional scatter and bias similar to the full-physics retrieval.

# 5.2 Summary of GOSAT error statistics

15

20

We have shown that clear-sky retrievals can be as accurate as full-physics retrievals for OCO-2 simulations over both land and ocean surfaces when scenes contaminated by clouds and aerosols are appropriately filtered. In this section, we explore whether this result is reproducible with real observations.

The effectiveness of applying DOGO to the pre-filtered GOSAT datasets is shown in Fig. 8. Initially, as in the OCO-2 simulations, there's a large reduction in the  $X_{CO_2}$  RMS error as the throughput is decreased. Based on our results from the OCO-2 simulations, this is likely because DOGO is identifying and filtering out highly contaminated scenes that have large  $X_{CO_2}$  errors due to complex multiple-scattering effects.

Over ocean surfaces, the clear-sky retrieval (dashed blue line) has larger errors than the full-physics retrieval (solid blue line), even at very high levels of filtration (low throughputs). The clear-sky retrieval  $X_{CO_2}$  RMS errors over ocean range from ~ 1.5– 5.0 ppm, depending on throughput. This error is about 0.5–3.0 ppm larger than the cor-

responding full-physics retrieval errors. As the throughput is decreased, the difference in error between the clear-sky and full-physics retrievals over ocean steadily decreases. This qualitatively agrees with our simulated OCO-2 results in that the clear-sky retrieval performs better as contaminated scenes are preferentially removed by DOGO. How-



ever, even at low throughputs (less than 60%), the clear-sky retrieval dataset still has RMS errors roughly 30–60% larger than those of the full-physics retrieval. This is in contrast to our simulation-based OCO-2 results, suggesting that additional unknown real-world mechanisms not included in the simulations may limit the ability of the clear-

- sky retrieval on real measurements over ocean surfaces. It is also possible that, despite promising results from our simulation-based tests, our filtering technique is unable to sufficiently remove scenes contaminated by clouds and aerosols over ocean surfaces. Another striking feature in this result is how effective the full-physics retrieval performs over ocean, even for highly contaminated scenes. At 100 % throughput, it still has RMS
- errors less than 2 ppm, implying that the more homogeneous nature of the ocean surface allows for the full-physics algorithm to successfully capture the influence of cloud and aerosol scattering effects.

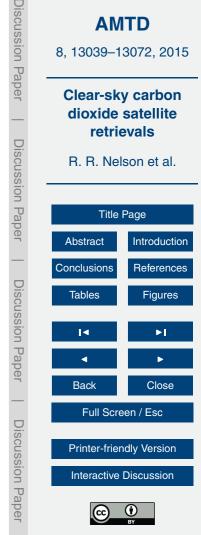
Over land, the clear-sky retrieval (dashed orange line) has errors similar to the fullphysics retrieval (solid orange line) when only modestly filtered by DOGO (throughput of ~ 80 %), which qualitatively reproduces our simulation-based results. The slopes of the initial decreases in RMS error over land are steeper than those for ocean. This may suggest that DOGO is more effective at initially removing contaminated or low quality measurements over land surfaces, despite having access to and using similar filter-

ing variables. This may be due to photons scattering multiple times between cloud or aerosol layers and the surface and producing a more distinguishable signal for DOGO to filter on than for ocean surfaces, where photons are much less likely to scatter off the surface multiple times (due to glint angle geometry).

The  $X_{CO_2}$  RMS errors for the TransCom regions are shown in Fig. 9. The left panel shows the RMS error for the GOSAT full-physics retrieval dataset while the right panel shows the same but for the GOSAT clear-sky dataset. The datasets shown were post-filtered using DOGO with a throughput of 30%. Over ocean, the clear-sky retrievals

25

(right) consistently have regional  $X_{CO_2}$  RMS errors 40–60 % larger than the full-physics retrievals (left). The results over land surfaces are more variable. While the global mean difference in RMS error over land at 30 % throughput is 0.1 ppm, or ~ 10 %, this differ-



ence, when regionally analyzed, can be positive or negative. Certain regions, such as Northern Africa, have lower full-physics retrieval RMS errors but others, such as South America, have slightly lower clear-sky retrieval RMS errors. This regional variability could be due to heterogeneous surface characteristics or cloud and aerosol compositions compared to the relatively uniform ocean. Thus, we can not say with confidence

tions compared to the relatively uniform ocean. Thus, we can not say with confidence that clear-sky retrievals perform better or worse over land surfaces for real GOSAT data without further investigation, but in general the two retrievals are surprisingly equivalent over land.

#### 6 Conclusions

- <sup>10</sup> In this study we evaluated the performance of non-scattering, or "clear-sky",  $X_{CO_2}$  retrievals from hyperspectral near-infrared measurements of reflected sunlight by comparing them to "full-physics"  $X_{CO_2}$  retrievals, which include scattering by clouds and aerosols. From our statistical analysis, we conclude that clear-sky  $X_{CO_2}$  retrievals typically do not perform as well as full-physics  $X_{CO_2}$  retrievals when no filtering is ap-<sup>15</sup> plied, consistent with previous findings. However, with the application of pre- and postfilters to remove low quality measurements contaminated by clouds and aerosols using only information contained in the near-infrared measurements themselves, our OCO-
- 2 simulation-based tests demonstrate that clear-sky retrievals are of similar or only slightly reduced quality compared to the full-physics retrieval, depending on filtration
   <sup>20</sup> level. This finding holds for real GOSAT observations over land, but not over ocean
- surfaces. For GOSAT measurements over land, the clear-sky retrieval has errors similar to those of the full-physics retrieval. For GOSAT measurements over ocean surfaces, the clear-sky retrieval has  $X_{CO_2}$  RMS errors 30–60 % larger than those of the full-physics retrieval when the datasets are modestly filtered. The source of this extra <sup>25</sup> error in the clear-sky retrieval applied to real GOSAT ocean measurements is unclear
- at this point and requires further study. Analysis of real OCO-2 measurements, which were unavailable during the time of this work, may help answer this question.



For OCO-2 simulations and real GOSAT measurements over both land and ocean surfaces, the clear-sky retrieval has  $X_{CO_2}$  RMS errors less than 2.0 ppm when the dataset is sufficiently filtered. In addition, clear-sky retrievals can be one to two orders of magnitude faster than full-physics retrievals, as scattering by clouds and aerosols can be completely ignored. For sensors that collect enormous volumes of data, such as OCO-2, this could allow for significantly more data to be processed. Thus, clear-sky  $X_{CO_2}$  retrievals may be useful for certain applications that require a larger number of retrievals but have less stringent error requirements. Further, estimates of parameters, such as surface albedo, from a clear-sky retrieval could serve as a useful first-guess for the full-physics retrieval. The faster processing and more accurate first-guess values could be of even greater utility to future satellite missions that will likely make even more measurements than OCO-2.

Acknowledgements. The authors would like to thank Natalie Tourville, Scott Denning, and Chris Kummerow of Colorado State University for their contributions to this work. Funding sources for this research include the NASA Jet Propulsion Laboratory (NASA JPL OCO-2 Subcontract #1 439 002) and the American Meteorological Society (AMS) Graduate Fellowship Program.

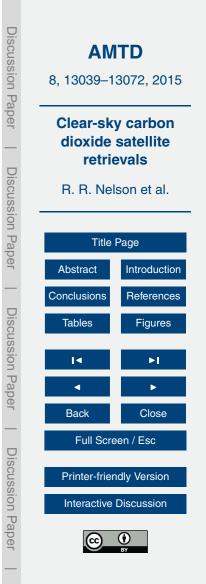
#### References

15

20

25

- Aben, I., Hasekamp, O., and Hartmann, W.: Uncertainties in the space-based measurements of CO<sub>2</sub> columns due to scattering in the Earth's atmosphere, J. Quant. Spectrosc. Ra., 104, 450–459, doi:10.1016/j.jgsrt.2006.09.013, 2007. 13042
- Baker, D. F., Bösch, H., Doney, S. C., O'Brien, D., and Schimel, D. S.: Carbon source/sink information provided by column CO<sub>2</sub> measurements from the Orbiting Carbon Observatory, Atmos. Chem. Phys., 10, 4145–4165, doi:10.5194/acp-10-4145-2010, 2010. 13041, 13048
  Basu, S., Guerlet, S., Butz, A., Houweling, S., Hasekamp, O., Aben, I., Krummel, P., Steele, P., Langenfelds, R., Torn, M., Biraud, S., Stephens, B., Andrews, A., and Worthy, D.: Global CO<sub>2</sub> fluxes estimated from GOSAT retrievals of total column CO<sub>2</sub>, Atmos. Chem. Phys., 13, 8695–8717, doi:10.5194/acp-13-8695-2013, 2013. 13041



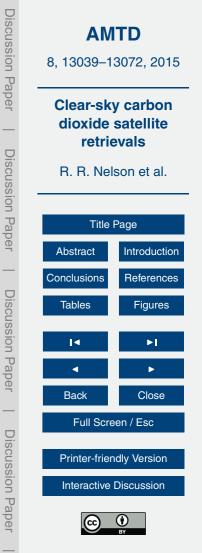
- Butz, A., Hasekamp, O. P., Frankenberg, C., and Aben, I.: Retrievals of atmospheric CO<sub>2</sub> from simulated space-borne measurements of backscattered near-infrared sunlight: accounting for aerosol effects, Appl. Optics, 48, 3322–3336, doi:10.1364/AO.48.003322, 2009. 13042, 13043
- <sup>5</sup> Butz, A., Guerlet, S., Hasekamp, O. P., Kuze, A., and Suto, H.: Using ocean-glint scattered sunlight as a diagnostic tool for satellite remote sensing of greenhouse gases, Atmos. Meas. Tech., 6, 2509–2520, doi:10.5194/amt-6-2509-2013, 2013. 13043, 13051
  - Chevallier, F., Bréon, F.-M., and Rayner, P. J.: Contribution of the Orbiting Carbon Observatory to the estimation of CO<sub>2</sub> sources and sinks: Theoretical study in a variational data assimila-
- tion framework, J. Geophys. Res., 112, D09307, doi:10.1029/2006JD007375, 2007. 13041
   Chevallier, F., Maksyutov, S., Bousquet, P., Bréon, F.-M., Saito, R., Toshida, Y., and Yokota, T.:
   On the accuracy of the CO<sub>2</sub> surface fluxes to be estimated from the GOSAT observations, Geophys. Res. Lett., 36, L19807, doi:10.1029/2009GL040108, 2009. 13041
  - Chevallier, F., Ciais, P., Conway, T. J., Aalto, T., Anderson, B. E., Bousquet, P., Brunke, E. G.,
- <sup>15</sup> Ciattaglia, L., Esaki, Y., Fröhlich, M., Gomez, A., Gomez-Pelaez, A. J., Haszpra, L., Krummel, P. B., Langenfelds, R. L., Leuenberger, M., Machida, T., Maignan, F., Matsueda, H., Morguí, J. A., Mukai, H., Nakazawa, T., Peylin, P., Ramonet, M., Rivier, L., Sawa, Y., Schmidt, M., Steele, L. P., Vay, S. A., Vermeulen, A. T., Wofsy, S., and Worthy, D.: CO<sub>2</sub> surface fluxes at grid point scale estimated from a global 21 year reanalysis of atmospheric
- <sup>20</sup> measurements, J. Geophys. Res., 115, D21307, doi:10.1029/2010JD013887, 2010. 13048 Connor, B. J., Bösch, H., Toon, G., Sen, B., Miller, C., and Crisp, D.: Orbiting Carbon Observatory: inverse method and prospective error analysis, J. Geophys. Res., 113, D05305, doi:10.1029/2006JD008336, 2008. 13044

Cox, C. and Munk, W.: Measurement of the roughness of the sea surface from photographs of

the sun's glitter, J. Opt. Soc. Am., 44, 838–850, doi:10.1364/JOSA.44.000838, 1954. 13047 Crisp, D., Miller, C. E., and DeCola, P. L.: NASA Orbiting Carbon Observatory: measuring the column averaged carbon dioxide mole fraction from space, J. Appl. Remote Sens., 2, 023508, doi:10.1117/1.2898457, 2008. 13041

Crisp, D., Bösch, H., Brown, L., Castano, R., Christi, M., Connor, B., Frankenberg, C., Mc-

<sup>30</sup> Duffie, 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. 13042, 13045

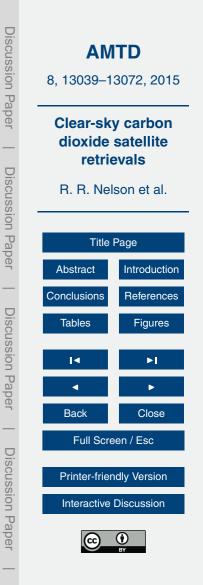
- Crowell, S., Rayner, P., Zaccheo, S., and Moore, B.: Impacts of atmospheric state uncertainty on O<sub>2</sub> measurement requirements for the ASCENDS mission, Atmos. Meas. Tech., 8, 2685– 2697, doi:10.5194/amt.8-2685-2015, 2015, 13045
- <sup>5</sup> 2697, doi:10.5194/amt-8-2685-2015, 2015. 13045
- Feng, L., Palmer, P. I., Yang, Y., Yantosca, R. M., Kawa, S. R., Paris, J.-D., Matsueda, H., and Machida, T.: Evaluating a 3-D transport model of atmospheric CO<sub>2</sub> using ground-based, aircraft, and space-borne data, Atmos. Chem. Phys., 11, 2789–2803, doi:10.5194/acp-11-2789-2011, 2011. 13048
- <sup>10</sup> Frankenberg, C., Platt, U., and Wagner, T.: Iterative maximum a posteriori (IMAP)-DOAS for retrieval of strongly absorbing trace gases: Model studies for CH<sub>4</sub> and CO<sub>2</sub> retrieval from near infrared spectra of SCIAMACHY onboard ENVISAT, Atmos. Chem. Phys., 5, 9–22, doi:10.5194/acp-5-9-2005, 2005. 13048

Frankenberg, C., Pollock, R., Lee, R. A. M., Rosenberg, R., Blavier, J.-F., Crisp, D., O'Dell, C. W.,

- <sup>15</sup> Osterman, G. B., Roehl, C., Wennberg, P. O., and Wunch, D.: The Orbiting Carbon Observatory (OCO-2): spectrometer performance evaluation using pre-launch direct sun measurements, Atmos. Meas. Tech., 8, 301–313, doi:10.5194/amt-8-301-2015, 2015. 13046
  - Guerlet, S., Butz, A., Schepers, D., Basu, S., Hasekamp, O. P., Kuze, A., Yokota, T., Blavier, J.-F., Deutscher, N. M., Griffith, D. W. T., Hase, F., Kyro, E., Morino, I., Sherlock, V., Suss-
- mann, R., Galli, A., and Aben, I.: Impact of aerosol and thin cirrus on retrieving and validating XCO<sub>2</sub> from GOSAT shortwave infrared measurements, J. Geophys. Res., 118, 1–19, doi:10.1002/jgrd.50332, 2013. 13043, 13047
  - Gurney, K. R., Law, R. M., Denning, A. S., Rayner, P. J., Baker, D., Bousquet, P., Bruhwiler, L., Chen, Y.-H., Ciais, P., Fan, S., Fung, I. Y., Gloor, M., Heimann, M., Higuchi, K., John, J.,
- Kowalczyk, E., Maki, T., Maksyutov, S., Peylin, P., Prather, M., Pak, B. C., Sarmiento, J., Taguchi, S., Takahashi, T., and Yuen, C.-W.: TransCom 3 CO<sub>2</sub> inversion intercomparison: 1. Annual mean control results and sensitivity to transport and prior flux information, Tellus B, 55, 555–579, doi:10.1034/j.1600-0889.2003.00049.x, 2003. 13050

Holben, B. N., Eck, T. F., Slutsker, I., Tanré, D., Buis, J., Setzer, A., Vermote, E., Reagan, J. A.,

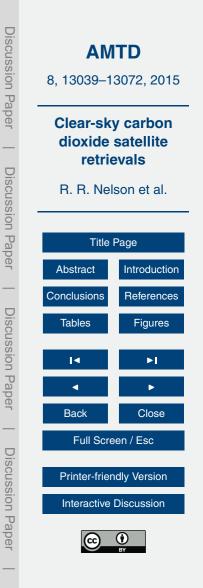
Kaufman, Y. J., Nakajima, T., Lavenu, F., Jankowiak, I., and Smirnov, A.: AERONET – a federated instrument network and data archive for aerosol characterization, Remote Sens. Environ., 66, 1–16, doi:10.1016/S0034-4257(98)00031-5, 1998. 13043



- Kahn, R., Banerjee, P., and McDonald, D.: Sensitivity of multiangle imaging to natural mixtures of aerosols over ocean, J. Geophys. Res., 106, 18219–18238, doi:10.1029/2000JD900497, 2001. 13046
- L'Ecuyer, T. S. and Jiang, J. H.: Touring the atmosphere aboard the A-Train, Phys. Today, 63, 36–41, doi:10.1063/1.3463626, 2010. 13047
- Le Quéré, C., Raupach, M. R., Canadell, J. G., and Marland, G.: Trends in the sources and sinks of carbon dioxide, Nat. Geosci., 2, 831–836, doi:10.1038/ngeo689, 2009. 13041

5

- Maksyutov, S., Takagi, H., Valsala, V. K., Saito, M., Oda, T., Saeki, T., Belikov, D. A., Saito, R., Ito, A., Yoshida, Y., Morino, I., Uchino, O., Andres, R. J., and Yokota, T.: Regional CO<sub>2</sub> flux
- estimates for 2009–2010 based on GOSAT and ground-based  $CO_2$  observations, Atmos. Chem. Phys., 13, 9351–9373, doi:10.5194/acp-13-9351-2013, 2013. 13048
  - Mandrake, L.: DOGO Warn Levels: An Upgrade to Quality Flags for the OCO-2 Data Products, AGU Fall Meeting, 17 December 2015, San Francisco, CA, USA, 2015. 13043, 13049
- Mandrake, L., Frankenberg, C., O'Dell, C. W., Osterman, G., Wennberg, P., and Wunch, D.:
   Semi-autonomous sounding selection for OCO-2, Atmos. Meas. Tech., 6, 2851–2864, doi:10.5194/amt-6-2851-2013, 2013. 13043, 13049, 13050
  - Miller, C. E., Crisp, D., DeCola, P. L., Olsen, S. C., Randerson, J. T., Michalak, A. M., Alkhaled, A., Rayner, P., Jacob, D. J., Suntharalingam, P., Jones, D. B. A., Denning, A. S., Nicholls, M. E., Doney, S. C., Pawson, S., Böosch, H., Connor, B. J., Fung, I. Y., O'Brien, D.,
- Salawitch, R. J., Sander, S. P., Sen, B., Tans, P., Toon, G. C., Wennberg, P. O., Wofsy, S. C., Yung, Y. L., and Law, R. M.: Precision requirements for space-based X<sub>CO<sub>2</sub></sub> data, J. Geophys. Res., 112, D10314, doi:10.1029/2006JD007659, 2007. 13041
  - Nelson, R. R.: The Impact of Aerosols on Space-Based Retrievals of Carbon Dioxide, MSc thesis, Colorado State University, Fort Collins, USA, 2015. 13042
- O'Brien, D. M. and Rayner, P. J.: Global observations of the carbon budget 2. CO<sub>2</sub> column from differential absorption of reflected sunlight in the 1.61 µm band of CO<sub>2</sub>, J. Geophys. Res., 107, ACH 6–1, doi:10.1029/2001JD000617, 2002. 13042
  - O'Brien, D. M., Polonsky, I., O'Dell, C., and Carheden, A.: Orbiting Carbon Observatory (OCO), Algorithm Theoretical Basis Document: The OCO Simulator ISSN 0737-5352-85, Tech. rep.,
- <sup>30</sup> Cooperative Institute for Research in the Atmosphere, Colorado State University, Fort Collins, USA, 2009. 13046
  - O'Dell, C. W., Connor, B., Bösch, H., O'Brien, D., Frankenberg, C., Castano, R., Christi, M., Eldering, D., Fisher, B., Gunson, M., McDuffie, J., Miller, C. E., Natraj, V., Oyafuso, F., Polon-



sky, I., Smyth, M., Taylor, T., Toon, G. C., Wennberg, P. O., and Wunch, D.: The ACOS CO<sub>2</sub> retrieval algorithm – Part 1: Description and validation against synthetic observations, Atmos. Meas. Tech., 5, 99–121, doi:10.5194/amt-5-99-2012, 2012. 13042, 13045, 13046, 13048 Osterman, G., Eldering, A., Avis, C., O'Dell, C., Martinez, E., Crisp, D., Frankenberg, C., and

- 5 Fisher, B.: ACOS Level 2 Standard Product Data User's Guide, v3.4, Jet Propulsion Laboratory, Pasadena, CA, USA, 2013. 13049
  - Parker, R., Boösch, H., and Cogan, A.: Algorithm Theoretical Basis Document Version 2 (ATBDv2) The University of Leicester Full-Physics Retrieval Algorithm for the retrieval of XCO<sub>2</sub> and XCH<sub>4</sub>, Tech. rep., version 2, University of Leicester, Leicester, UK, 2011. 13042
- Peters, W., Jacobson, A. R., Sweeney, C., Andrews, A. E., Conway, T. J., Masarie, K., Miller, J. B., Bruhwiler, L. M. P., Petron, G., Hirsch, A. I., Worthy, D. E. J., van der Werf, G. R., Randerson, J. T., Wennberg, P. O., Krol, M. C., and Tans, P. P.: An atmospheric perspective on North American carbon dioxide exchange: CarbonTracker, P. Natl. Acad. Sci. USA, 104, 18925–18930, doi:10.1073/pnas.0708986104, 2007. 13048
- Rayner, P. J. and O'Brien, D. M.: The utility of remotely sensed CO<sub>2</sub> concentration data in surface source inversions, Geophys. Res. Lett., 28, 175–178, doi:10.1029/2000GL011912, 2001. 13041

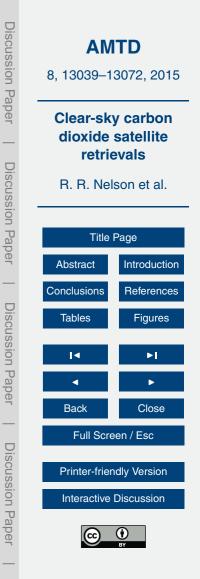
Reuter, M., Bovensmann, H., Buchwitz, M., Burrows, J. P., Heymann, J., Hilker, M., and Schneising, O.: Algorithm Theoretical Basis Document Version 2 (ATBDv2) – The Bremen

<sup>20</sup> Optimal Estimation DOAS (BESD) Algorithm for the Retrieval of  $X_{CO_2}$ , Tech. rep., Institute of Environmental Physics (IUP) University of Bremen, Bremen, Germany, version 2, 2013. 13042

Rodgers, C. D.: Inverse Methods for Atmospheric Sounding: Theory and Practice, World Scientific, Singapore, 2000. 13045

- Salstein, D. A., Ponte, R. M., and Cady-Pereira, K.: Uncertainties in atmospheric surface pressure fields from global analyses, J. Geophys. Res., 113, D14107, doi:10.1029/2007JD009531, 2008. 13045
  - Taylor, T. E., O'Dell, C. W., O'Brien, D. M., Kikuchi, N., Nakajima, T. Y. T. Y., Ishida, H., Crisp, D., and Nakajima, T.: Comparison of cloud-screening methods applied to GOSAT near-infrared
- <sup>30</sup> spectra, IEEE T. Geosci. Remote, 50, 295–309, doi:10.1109/TGRS.2011.2160270, 2012. 13044, 13048

Taylor, T. E., Savtchenko, A., Osterman, G., O'Dell, C. W., Nelson, R. R., Pollock, R. H., Frankenberg, C., Rosenthal, E. J., Crisp, D., Partain, P., Chang, A., Eldering, A., Cronk, H. Q.,



Fisher, B., and Gunson, M. R.: Orbiting Carbon Observatory-2 (OCO-2) cloud screening; validation against collocated MODIS and CALIOP data, Atmos. Meas. Tech., accepted, 2015. 13044, 13048

Winker, D. M., Vaughan, M. A., Omar, A., Hu, Y., and Powell, K. A.: Overview of the CALIPSO
 Mission and CALIOP data processing algorithms, J. Atmos. Ocean. Tech., 26, 2310–2323, doi:10.1175/2009JTECHA1281.1, 2009. 13047

Wunch, D., Toon, G. C., Blavier, J. F., Washenfelder, R. A., Notholt, J., Connor, B. J., Griffith, D. W. T., Sherlock, V., and Wennberg, P. O.: The Total Carbon Column Observing Network, Philos. T. Roy. Soc. A, 369, 2087–2112, doi:10.1098/rsta.2010.0240, 2011a. 13047

Yokota, T., Yoshida, Y., Eguchi, N., Ota, Y., Tanaka, T., Watanabe, H., and Maksyutov, S.: Global concentrations of CO<sub>2</sub> and CH<sub>4</sub> retrieved from GOSAT: first preliminary results, Scientific Online Letters on the Atmosphere, 5, 160–163, doi:10.2151/sola.2009-041, 2009. 13041, 13042

	<b>AMTD</b> 8, 13039–13072, 2015			
dioxide	Clear-sky carbon dioxide satellite retrievals			
R. R. Ne	lson et al.			
Title	Page			
Abstract	Introduction			
Conclusions	References			
Tables	Figures			
I	۶I			
	•			
Back	Close			
Full Scre	Full Screen / Esc			
Printer-frier	Printer-friendly Version			
Interactive	Interactive Discussion			

**Discussion** Paper

**Discussion** Paper

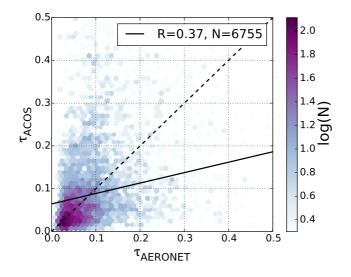
**Discussion Paper** 

Discussion Paper

e 1. Properties of the full-physics and clear-sky X <sub>CO2</sub> retrievals.TypeBands Used (μm)Clouds & Aerosols ParameterizedP <sub>sfc</sub> SourceFull-Physics0.76, 1.6, 2.0YesO2 A-bandClear-Sky1.6, 2.0NoECMWF				
Full-Physics         0.76, 1.6, 2.0         Yes         O <sub>2</sub> A-band				
Full-Physics0.76, 1.6, 2.0YesO2 A-band				
			-	-
	Туре	Bands Used (µm)	Clouds & Aerosols Parameterized	-

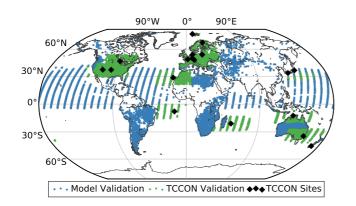


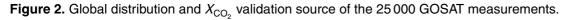
**Discussion** Paper



**Figure 1.** Heat-map of AERONET aerosol optical depth compared to the retrieved ACOS B3.4 aerosol optical depth from GOSAT measurements. A linear fit is shown by the solid black line.







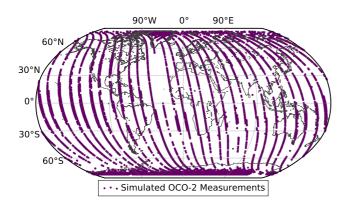
<b>AMTD</b> 8, 13039–13072, 2015				
Clear-sky carbon dioxide satellite retrievals				
R. R. Ne	lson et al.			
	Deser			
I Itle	Page			
Abstract	Introduction			
Conclusions	Conclusions References			
Tables	Tables Figures			
14	▶1			
•	•			
Back	Close			
Full Screen / Esc				
Printer-friendly Version				
Interactive	Interactive Discussion			
BY BY				

**Discussion** Paper

**Discussion** Paper

**Discussion** Paper

**Discussion** Paper





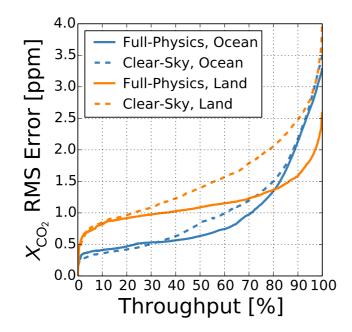
	AMTD				
8, 13039–1	3072, 2015				
dioxide	Clear-sky carbon dioxide satellite retrievals				
R. R. Nel	son et al.				
Title	Page				
	J.				
Abstract	Introduction				
Conclusions	Conclusions References				
Tables	Tables Figures				
14	►I				
•	•				
Back	Close				
Full Screen / Esc					
Printer-frier	Printer-friendly Version				
Interactive	Interactive Discussion				
BV					

**Discussion Paper** 

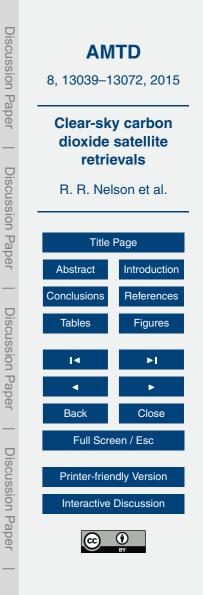
**Discussion** Paper

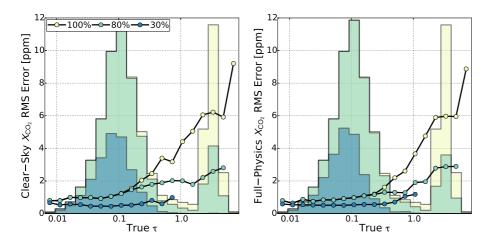
**Discussion Paper** 

**Discussion** Paper



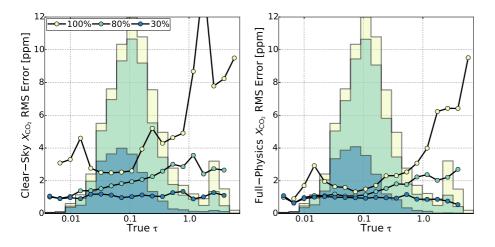
**Figure 4.** DOGO filters applied to retrievals on simulated OCO-2 measurements for ocean (blue) and land (orange) surfaces for both full-physics (solid) and clear-sky (dashed) retrievals. Four variables were chosen and optimized by the DOGO system. The *x* axis is throughput, which represents the percentage of data that remains after applying the filter. The *y* axis is the  $X_{CO_2}$  RMS error.



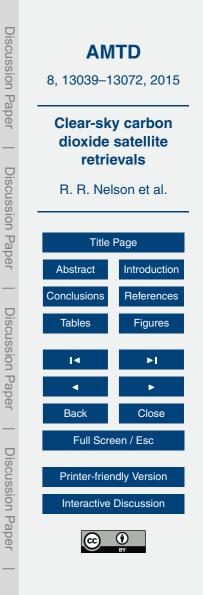


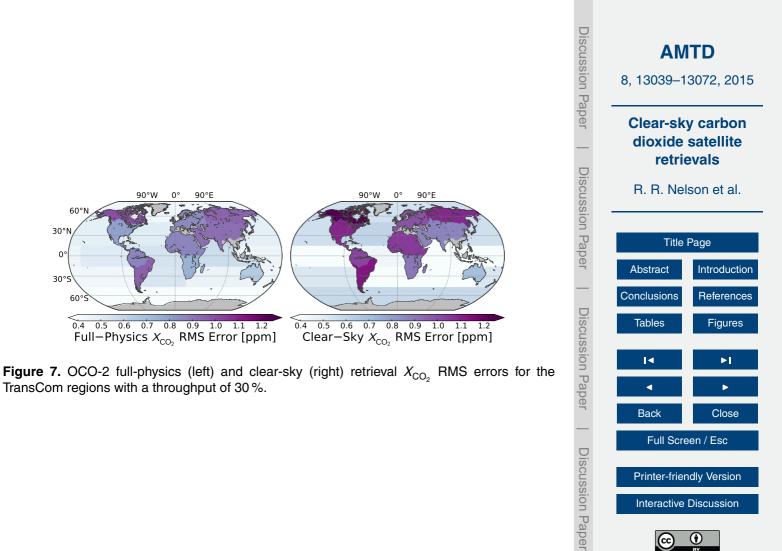
**Figure 5.** Clear-sky (left) and full-physics (right)  $X_{CO_2}$  retrieval RMS errors vs. the true total optical depth for simulated OCO-2 measurements over ocean. The black lines are binned averages of the  $X_{CO_2}$  RMS error for 100 % throughput (yellow markers), 80 % throughput (green markers), and 30 % throughput (blue markers). The histograms represent the relative amount of data for each throughput.



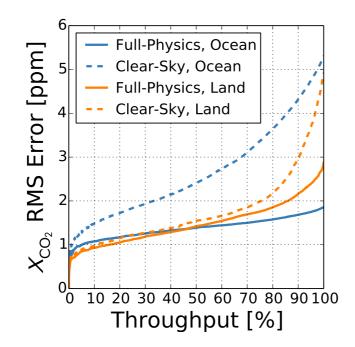


**Figure 6.** Clear-sky (left) and full-physics (right)  $X_{CO_2}$  retrieval RMS errors vs. the true total optical depth for simulated OCO-2 measurements over land. The black lines are binned averages of the  $X_{CO_2}$  RMS error for 100% throughput (yellow markers), 80% throughput (green markers), and 30% throughput (blue markers). The histograms represent the relative amount of data for each throughput.



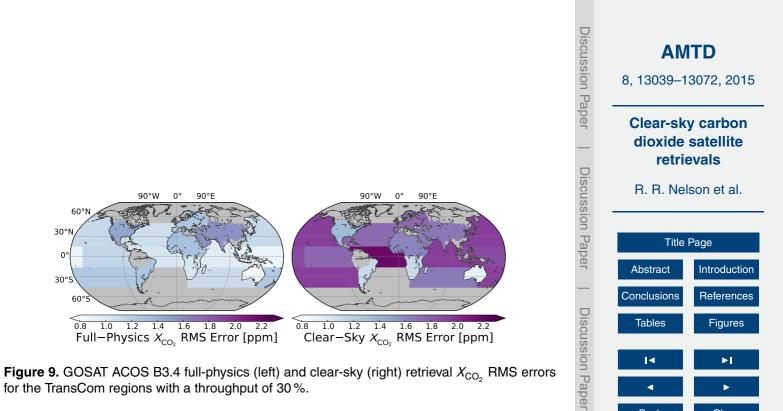






**Figure 8.** DOGO filters applied to retrievals on GOSAT measurements for ocean (blue) and land (orange) surfaces for both full-physics (solid) and clear-sky (dashed) retrievals. Four variables were chosen and optimized by the DOGO system. The *x* axis is throughput, which represents the percentage of data that remains after applying the filter. The *y* axis is the  $X_{CO_2}$  RMS error.





Close

Back

**Discussion** Paper

Full Screen / Esc

Printer-friendly Version Interactive Discussion