



The potential of clear-sky carbon dioxide satellite retrievals

Robert R. Nelson¹, Christopher W. O'Dell², Thomas E. Taylor², Lukas Mandrake³, and Mike Smyth³

¹Colorado State University, Fort Collins, CO, USA

²Cooperative Institute for Research in the Atmosphere, Fort Collins, CO, USA

³Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA

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

Received: 2 November 2015 – Published in Atmos. Meas. Tech. Discuss.: 10 December 2015

Revised: 29 March 2016 – Accepted: 1 April 2016 – Published: 15 April 2016

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 retrievals 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 errors 0–20 % larger than the full-physics retrieval over land and errors roughly 20–35 % larger over ocean, depending on filtration level. In general, the clear-sky re-

trieval had X_{CO_2} root-mean-square errors (RMSEs) of less than 2.0 ppm, relative to Total Carbon Column Observing Network (TCCON) measurements and a suite of CO_2 models, when adequately filtered through the use of a custom genetic algorithm filtering system. These results imply that non-scattering X_{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. 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.

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_2 stays in the atmosphere. The remaining molecules are absorbed by the land and ocean, but where this absorption takes place is still highly uncertain, especially over land surfaces (Le Quééré et al., 2009). CO_2 inverse modeling systems, designed to answer important questions about Earth's carbon sources and sinks and their interaction with the atmosphere, are heavily dependent on the density and quality of CO_2 measurements (Rayner and O'Brien, 2001; Baker et al., 2010; Chevallier et al., 2007, 2009). Global coverage of 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.

Specifically, it has been shown that space-based measurements of the column-averaged dry-air mole fraction of carbon dioxide, or X_{CO_2} , need a precision of better than about 0.5% (~ 2 ppm) 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 CO_2 inverse modeling systems (Chevallier et al., 2007; Basu et al., 2013). Thus, it is critically important to minimize random errors and biases in satellite measurements of CO_2 in order to be able to correctly answer questions about the carbon cycle.

When retrieving 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 to accurately measure X_{CO_2} , 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} . Thus, 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 X_{CO_2}) 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 the 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 their 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, for certain retrieval algorithm setups, 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 scenes, this becomes an issue because the addition of a cloud and aerosol parameterization may be detrimental rather than beneficial. Additionally, Fig. 1 shows a comparison of retrieved full-physics optical depths from build 3.4 (B3.4) of the NASA Atmospheric CO_2 Observations from Space (ACOS) X_{CO_2} retrieval algorithm (O'Dell et al., 2012; Crisp et al., 2010) to optical depths mea-

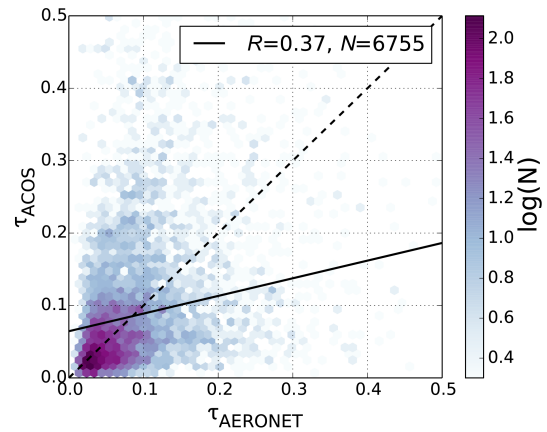


Figure 1. Heat map of AERONET aerosol optical depth compared to the retrieved ACOS B3.4 aerosol optical depth from GOSAT measurements. A linear fit is shown by the solid black line.

sured from the highly accurate Aerosol RObotic NETwork (AERONET; Holben et al., 1998), using coincidence criteria of ± 30 min and 0.1° . The aerosol optical depths retrieved by the full-physics retrieval were not well correlated with the AERONET measurement for a particular scene, which suggests that the full-physics retrieval often has difficulty obtaining information about clouds and aerosols.

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

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 per CPU core and OCO-2 collects about 10^6 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 1–2 orders of magnitude more data to be processed.

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. We use mature 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) 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 data sets used in this study. The fourth section describes the validation sources and 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_{CO_2} retrievals

Hyperspectral measurements of reflected sunlight in the near-infrared can be used to infer CO_2 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 X_{CO_2} to be accurately estimated. The instruments onboard GOSAT and OCO-2 make use of this method. Typically, a relatively weak CO_2 absorption band located in the near-infrared at approximately $1.6\ \mu\text{m}$ and a stronger CO_2 absorption band at $2.0\ \mu\text{m}$ are used in conjunction to estimate the average amount of CO_2 in the light path seen by the instrument's sensors. Of note, the $2.0\ \mu\text{m}$ band is used to gain information about CO_2 but is also more sensitive to aerosols than the $1.6\ \mu\text{m}$ band. Additionally, an oxygen absorption feature near $0.76\ \mu\text{m}$, known as the O_2 A-band, is often employed to help filter out clouds and aerosols (Taylor et al., 2012, 2016) (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} :

$$X_{\text{CO}_2} = \frac{\int_0^\infty N_{\text{CO}_2}(z) dz}{\int_0^\infty N_{\text{d}}(z) dz}, \quad (1)$$

where $N_{\text{CO}_2}(z)$ is the molecular number density of CO_2 at altitude z and $N_{\text{d}}(z)$ is the molecular number density of dry air at altitude z .

Values of X_{CO_2} are estimated by the ACOS retrieval algorithm using the measured radiances and a priori information to optimize a state vector (Rodgers, 2000). Complete

Table 1. Properties of the full-physics and clear-sky X_{CO_2} retrievals.

Type	Bands used (μm)	Clouds & aerosols retrieved	P_{sfc} retrieved
Full-physics	0.76, 1.6, 2.0	Yes	Yes (O_2 A-band)
Clear-sky	1.6, 2.0	No	No

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 represent 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_2 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 European Centre for Medium-Range Weather Forecasts Integrated Forecast System (ECMWF IFS; ECMWF, 2015) model fields, and the CO_2 profile priors are interpolated from the Total Carbon Column Observing Network (TCCON) a priori (Wunch et al., 2011a).

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

The clear-sky retrieval utilizes the CO_2 near-infrared bands at 1.6 and $2.0\ \mu\text{m}$ but does not include cloud or aerosol parameters in the state vector. Instead of using the O_2 A-band at $0.76\ \mu\text{m}$ to retrieve information about surface pressure, which is used to estimate N_{d} , the clear-sky retrieval simply uses the a priori surface pressure value. We included Rayleigh scattering by air molecules for the two near-infrared CO_2 bands, but these effects are likely negligible at such long wavelengths.

The full-physics retrieval uses the two near-infrared CO_2 bands as well as the O_2 A-band at $0.76\ \mu\text{m}$. The O_2 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). The Kahn 2b aerosol type is a mixture of coarse- and fine-mode dust, while the Kahn 3b aerosol type is a mixture of smaller carbonaceous aerosols. Both 2b and 3b also contain sulfate and sea salt components. Simulations have suggested that a combination of these four

scattering types is sufficient to approximately represent any type of scene observed by GOSAT or OCO-2 (O'Dell et al., 2012).

3 Data

The simulated OCO-2 data set consists of retrievals performed on approximately 44 000 synthetic measurements spanning 17–18 June 2012 and 19–20 December 2012 (a total 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. Glint viewing geometry over land was not examined in this study because GOSAT is unable to make measurements using that viewing method. The simulated radiances were generated by the Colorado State University (CSU) Orbit Simulator, which uses realistic meteorology, cloud and aerosol distributions, and surface properties (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 were used for the meteorological a priori, while ECMWF IFS model data were used to create the synthetic radiances. This intentional mismatch in meteorology mimics real-world inaccuracies when measuring a given scene. For both the full-physics and clear-sky retrievals performed on simulated measurements, the a priori surface pressure was taken from the NCEP reanalysis data. As discussed in Sect. 2, the full-physics retrieval then used the O₂ A-band to fine-tune the surface pressure estimate, while the clear-sky retrieval assumed the a priori to be correct. 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 Constellation (L'Ecuyer and Jiang, 2010). Land surfaces included albedos and bi-directional reflectance distribution functions (BRDFs) 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 the ECMWF IFS.

The GOSAT data set contained retrievals performed 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 data set if they had a corresponding X_{CO_2} validation source (see Sect. 4.1). Surface pressure a priori values were taken from the ECMWF IFS model, which has been shown to be accurate to within 1–2 hPa under

most conditions (Salstein et al., 2008; Crowell et al., 2015). As the clear-sky retrieval did not retrieve surface pressure, it was necessary to have an accurate a priori surface pressure estimate. To ensure this methodology was valid, we tested a version of the clear-sky retrieval that was allowed to retrieve surface pressure from the O₂ A-band and found the results to be consistent. Finally, OCO-2 and GOSAT retrievals that did not converge in the full-physics retrieval were not used for this study.

4 Methodology

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 used to further improve the quality of the data set by removing additional contaminated scenes.

4.1 X_{CO_2} validation sources

To evaluate the accuracy and precision 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 TCCON measurement co-located in time and space using the technique described in Guerlet et al. (2013) or the average of seven CO₂ models that assimilate ground-based and aircraft CO₂ measurements and agreed to 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 CO₂ models used include two from the University of Edinburgh (Feng et al., 2011), one from Le Laboratoire des Sciences du Climat et de l'Environnement (Chevallier et al., 2010), two from the National Institute for Environmental Studies (Maksyutov et al., 2013), the 2010 version of CarbonTracker (Peters et al., 2007), and the National Oceanic and Atmospheric Administration Parameterized Chemistry and Transport Model (NOAA PCTM; Kawa et al., 2004). Both methods of X_{CO_2} validation have limitations, but we believe they are still useful in evaluating the performance of the retrieval algorithms. Additionally, the results of this study were found to be largely insensitive to the method of validation used.

4.2 Pre-filtering

In order to remove measurements heavily contaminated by clouds or aerosols, the OCO-2 and GOSAT data sets were pre-filtered using the O₂ A-band Preprocessor (ABP) (Taylor et al., 2012; O'Dell et al., 2012). The ABP works by estimating surface pressure and albedo from the O₂ A-band assuming no clouds or aerosols are present. If the scene is contaminated by thick cloud or aerosol layers, the results may

have identifiably large residuals between the measured and modeled radiances as well as unrealistic surface pressure or albedo values. The ABP was run on all the measurements because it is extremely fast and computationally inexpensive. The use of the ABP removes most, but not all, scenes contaminated by clouds and aerosols (Taylor et al., 2012, 2016). Clear-sky measurements at high latitudes and over ice and snow, which typically give reduced signal to noise ratios, are also removed by our pre-filters.

Figure 2 shows the global distribution of the 44 000 pre-filtered synthetic OCO-2 measurements, while Fig. 3 shows the location of the 25 000 pre-filtered real GOSAT measurements used in this study, and whether the X_{CO_2} validation source (see Sect. 4.1) was a model consensus or a TCCON measurement.

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} retrievals 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 a genetic algorithm (DOGO; Mandrake et al., 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. Unpublished studies have shown that this approach yields similar results to the hand-tuned filters, but typically 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 data set, minimized the root mean square 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 RMSE includes both variance and bias.

The parameters allowed for selection by DOGO 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 CALIOP measurements (used to create the OCO-2 simulations). This was to ensure that DOGO did not “cheat” by having access to external information that would not generally be available for real instruments. Filtering was done for different “throughputs”, which equal the percent of data retained after post-filtering

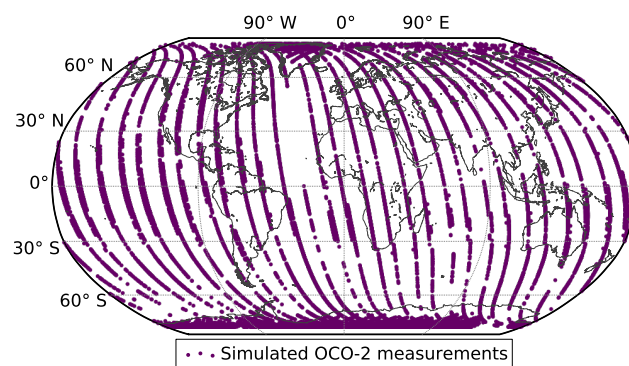


Figure 2. Global distribution of the 44 000 simulated OCO-2 measurements.

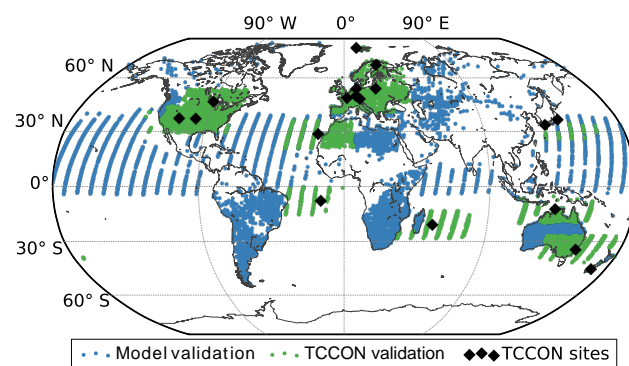


Figure 3. Global distribution and X_{CO_2} validation source of the 25 000 GOSAT measurements.

with DOGO at a given filtration level. 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} RMSE at a given throughput, while two rules select 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 the errors. DOGO was run separately for both clear-sky and full-physics retrievals as well as for land and ocean surfaces because it was hypothesized that a different set of four filtering parameters might be selected for each retrieval and surface type combination.

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_{CO_2} retrievals, similar to what is done for operational GOSAT retrievals (Wunch et al., 2011b; Guerlet et al., 2013). This was because the clear-sky GOSAT retrievals over ocean initially contained a large bias in X_{CO_2} , which initial testing showed was not entirely removed by DOGO. The X_{CO_2} data were adjusted using both a slope and an offset:

$$X_{\text{CO}_2} = X_{\text{CO}_2} + A + Bx, \quad (2)$$

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_{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 $2.0 \mu\text{m}$ CO_2 band to the $1.6 \mu\text{m}$ 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_{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 CO_2 bands.

5 X_{CO_2} retrieval comparison

In the previous section we described our use of pre-filtering with the ABP and post-filtering with DOGO to produce a data set 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_{CO_2} retrieval relative to the full-physics X_{CO_2} retrieval for simulated OCO-2 measurements and real GOSAT measurements. We evaluate the effectiveness of DOGO at reducing RMSEs, investigate the impact of clouds and aerosols on the OCO-2 simulations, and examine regional biases, scatter, and RMSEs.

5.1 Summary of OCO-2 error statistics

We begin by evaluating the performance of the clear-sky retrieval on simulated OCO-2 measurements. Figure 4 demonstrates the effectiveness of DOGO at reducing X_{CO_2} RMSEs 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 data sets begin to plateau at approximately 50–80% throughput as DOGO has already removed the obviously contaminated scenes, as discussed below and demonstrated in Figs. 5 and 6, and is now selecting the best of what remains. At very low throughputs, there are not enough measurements for the algorithm to function properly. Additionally, at such low throughputs, the RMSE approaches the error due to simulated instrument noise over ocean (~ 0.35 ppm). Interestingly, the RMSEs over land plateau around 0.80 ppm, while the sim-

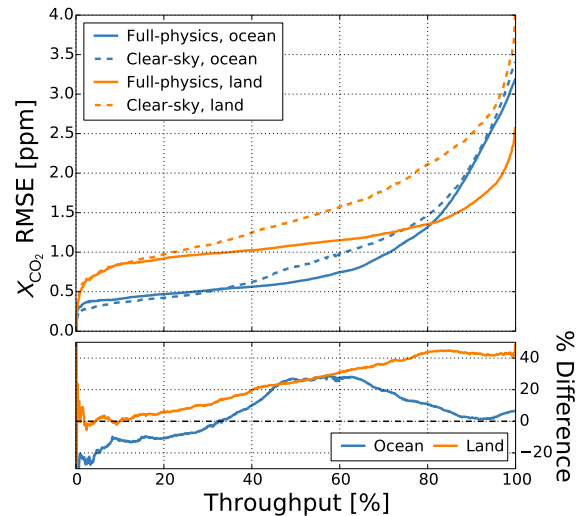


Figure 4. DOGO filters applied to full-physics (solid) and clear-sky (dashed) retrievals performed on simulated OCO-2 measurements for ocean (blue) and land (orange) surfaces. Four variables were chosen and optimized by the DOGO system. The x axis is throughput, which represents the percentage of data that remains after applying the DOGO filters. The y axis in the top panel is the X_{CO_2} RMSE. The y axis in the bottom panel is the percent difference between the clear-sky X_{CO_2} RMSE (E_1) and the full-physics X_{CO_2} RMSE (E_2): $(E_1 - E_2) / (0.5 \cdot (E_1 + E_2)) \cdot 100$. Positive values indicate smaller full-physics X_{CO_2} RMSEs, while negative values indicate smaller clear-sky X_{CO_2} RMSEs.

ulated noise limit is ~ 0.50 ppm. This suggests an underestimation of the posterior errors for OCO-2 simulations over land, a feature seen previously in both simulations (O'Dell et al., 2012) and real GOSAT data (Kulawik et al., 2016).

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} RMSE more than 0.25 ppm worse than the full-physics data set. The first 20% of scenes filtered out by DOGO (from 100 to 80% throughput) likely contain very thick clouds and aerosols, as discussed below, such that both the full-physics and clear-sky retrievals produce large RMSEs 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 for these highly contaminated scenes, even though it attempts to account for the presence of clouds and aerosols. This may suggest that in our simulations over ocean, the full-physics retrieval struggles to properly quantify the light path modifications of clouds and aerosols. However, the relative magnitudes and trends in the RMSE reduction can be sensitive to the pre-filtering setup and bias corrections. From 80 to 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

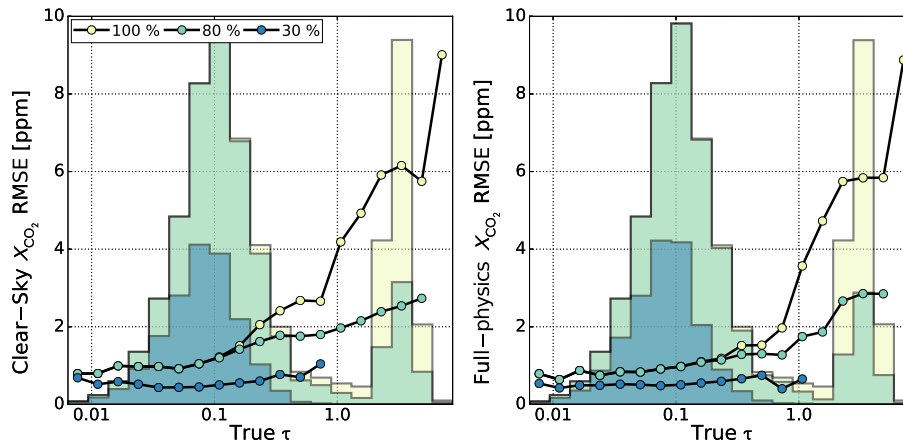


Figure 5. Clear-sky (left) and full-physics (right) retrieval X_{CO_2} RMSEs vs. the true total optical depth for simulated OCO-2 measurements over ocean. The black lines are binned averages of the X_{CO_2} RMSE for 100 % throughput (yellow markers), 80 % throughput (green markers), and 30 % throughput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

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 ocean scenes become pristine enough that the clear-sky retrieval has $\sim 10\%$ smaller X_{CO_2} RMSEs than the full-physics retrieval. It is likely that the full-physics retrieval struggles slightly compared to the clear-sky retrieval because it is 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 demonstrates that the clear-sky retrievals (dashed orange line) consistently have higher RMSEs than the full-physics retrievals (solid orange line) at high throughputs, upwards of a difference of 1.5 ppm, or $\sim 40\%$ larger. At these higher throughputs, most scenes still contaminated by clouds and aerosols remain and the full-physics 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} RMSEs 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 contaminated scenes.

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} RMSE vs. the true total optical depth (the sum of the true aerosol,

ice cloud, and water cloud optical depths from CALIOP, used to create the synthetic measurements) for clear-sky retrievals (left panel) and full-physics retrievals (right panel) over ocean. At 100 % throughput, there are a significant number of thick ($\tau > 1.0$) cloud and aerosol scenes present, along with a secondary peak of thinner cloud and 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} RMSEs for clear-sky and full-physics retrievals over ocean are nearly equivalent, which agrees with our analysis of the ocean retrievals in Fig. 4. For high optical depth scenes ($\tau > 1.0$) at 100 % throughput, the RMSE of the data is large (over 6 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) results in DOGO greatly reducing the number of these high optical depth scenes, which corresponds to the steep initial decline of the ocean retrieval RMSEs 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 information contained in the near-infrared measurements themselves to infer the amount of clouds and aerosols in a given scene. Specifically, the CO_2 and H_2O ratios were frequently selected by DOGO to filter the OCO-2 retrievals. The CO_2 and H_2O ratios were taken from the Iterative Maximum A-Posteriori Differential Optical Absorption Spectroscopy (IMAP-DOAS) algorithm (Frankenberg et al., 2005; Taylor et al., 2016) and are calculated using estimates of CO_2 and H_2O from both the 1.6 and 2.0 μm CO_2 bands independently using a fast, non-scattering algorithm. Deviations from unity in the ratio of the 1.6 to 2.0 μm band values, caused by the wavelength dependence of clouds and aerosols, allows for the identification of many contaminated scenes.

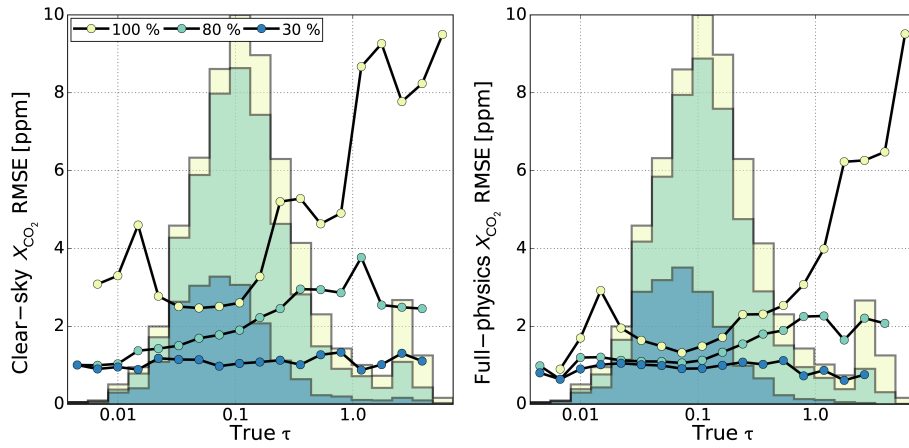


Figure 6. Clear-sky (left) and full-physics (right) retrieval X_{CO_2} RMSEs vs. the true total optical depth for simulated OCO-2 measurements over land. The black lines are binned averages of the X_{CO_2} RMSE for 100 % throughput (yellow markers), 80 % throughput (green markers), and 30 % throughput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

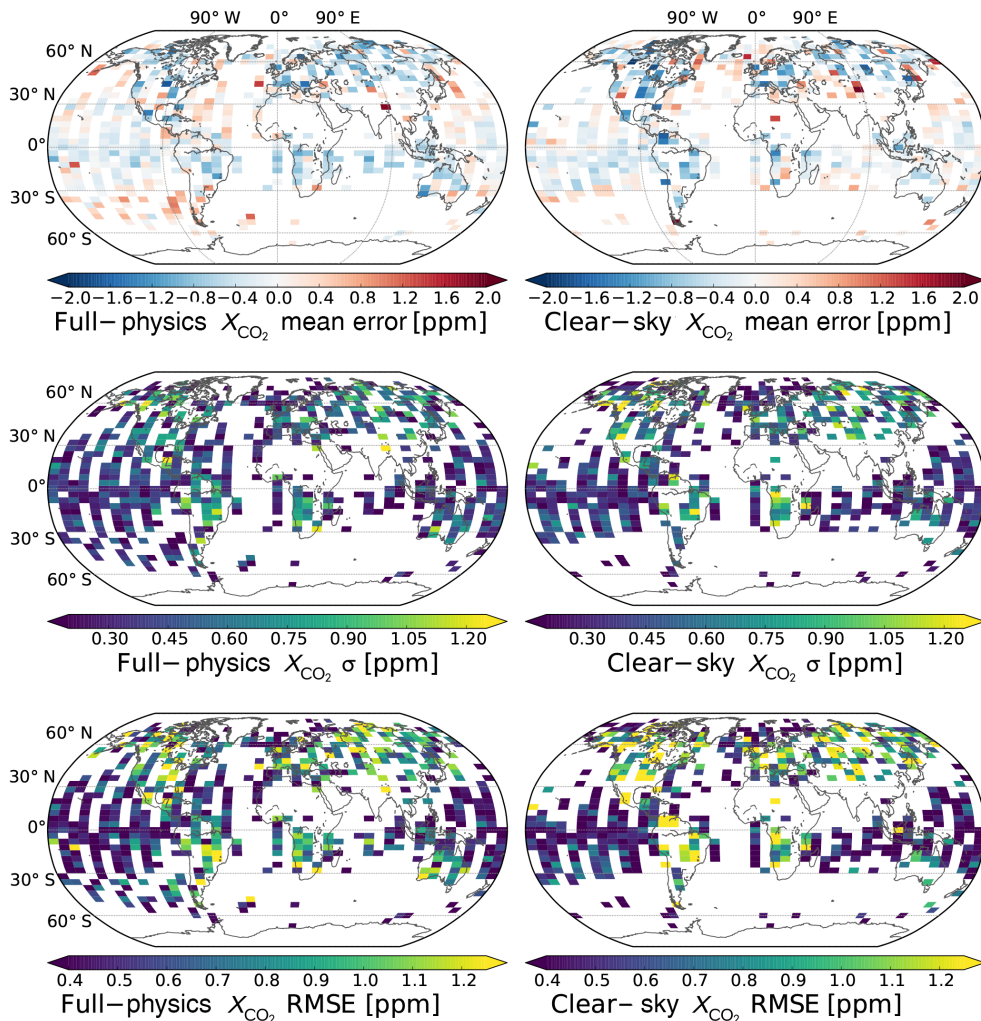


Figure 7. OCO-2 full-physics (left column) and clear-sky (right column) retrieval mean X_{CO_2} errors (top row), standard deviation of the X_{CO_2} errors (middle row), and X_{CO_2} RMSEs (bottom row) for $8^\circ \times 4^\circ$ (longitude \times latitude) bins for a throughput of 30 %.

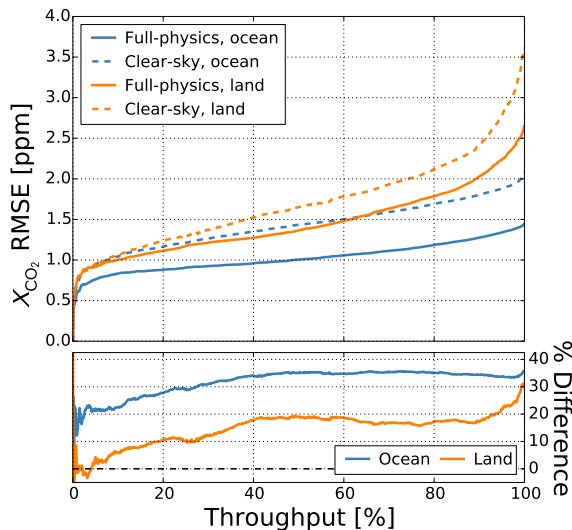


Figure 8. DOGO filters applied to full-physics (solid) and clear-sky (dashed) retrievals performed on GOSAT measurements for ocean (blue) and land (orange) surfaces. Four variables were chosen and optimized by the DOGO system. The x axis is throughput, which represents the percentage of data that remains after applying the DOGO filters. The y axis in the top panel is the X_{CO_2} RMSE. The y axis in the bottom panel is the percent difference between the clear-sky X_{CO_2} RMSE (E_1) and the full-physics X_{CO_2} RMSE (E_2): $(E_1 - E_2)/(0.5 \cdot (E_1 + E_2)) \cdot 100$. Positive values indicate smaller full-physics X_{CO_2} RMSEs, while negative values indicate smaller clear-sky X_{CO_2} RMSEs.

dP_{ABP} and parameters related to the $2.0 \mu\text{m}$ CO₂ band signal were also often selected as filters, as they relate to scattering and absorption by clouds and aerosols. At 80 % throughput (green lines and histograms), the full-physics retrieval has slightly smaller RMSEs 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 remain (blue lines and histograms), 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 X_{CO_2} RMSE. The same is true for precision and bias (not shown). One might think that even slightly contaminated scenes ($\tau \sim 0.1\text{--}0.3$) should have been removed for the 30 % throughput case, but DOGO's goal is to minimize the X_{CO_2} RMSE, not to remove scenes with high optical depths.

A similar analysis was performed for measurements over land, with the results displayed in Fig. 6. As was seen over ocean in Fig. 5, at larger throughputs higher optical depths correspond to larger RMSEs in the X_{CO_2} data. The clear-sky retrieval struggles with these high optical depth scenes,

but also has relatively large RMSEs (~ 3 ppm) for moderate to small optical depths. Interestingly, the initial number of high optical depth scenes over land is considerably smaller than over ocean. This indicates that the ABP may be more effective at removing high optical depth scenes over land, likely because scattering between cloud or aerosol layers and the surface makes these contaminants more identifiable. As the throughput decreases, the RMSE 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 RMSE is still optimally reduced by DOGO and only ~ 10 % worse than the full-physics retrieval. For the full-physics data set, DOGO chooses to remove nearly all of these thick cloud or aerosol scenes. This suggests that the clear-sky retrieval may be slightly 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 entire data set, spatial errors in the OCO-2 retrievals were analyzed to see if regional variability existed and if there were regions where the clear-sky retrieval had smaller or larger errors relative to the full-physics retrieval. The binned mean X_{CO_2} errors, standard deviation of the X_{CO_2} errors, and X_{CO_2} RMSEs for the OCO-2 simulations are shown in Fig. 7. Here we use a throughput of 30 %, where the globally averaged clear-sky X_{CO_2} RMSEs are approximately equivalent to the full-physics errors over ocean and slightly larger (0.1 ppm) over land (as shown in Fig. 4). The coverage of post-filtering to 30 %, shown in Fig. 7, is spatially dependent because of a preference to remove measurements over regions that persistently contain clouds or aerosols (e.g., the Sahara) or have low signal to noise ratios (e.g., high latitudes). The clear-sky and full-physics retrieval mean error spatial patterns are similar and both relatively small in magnitude, which indicates that they do not contain large regional biases and, more importantly, that the clear-sky retrieval does not have significantly larger biases than the full-physics retrieval. The scatter and the RMSEs both show limited regional variability over ocean and modest variability over land. The regional clear-sky retrieval errors are approximately the same magnitude as the full-physics retrieval errors over ocean and only marginally larger over land. Overall, these simulated results are promising because they demonstrate that the clear-sky X_{CO_2} retrieval has global and regional error statistics similar to the full-physics X_{CO_2} retrieval.

5.2 Summary of GOSAT error statistics

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 filtering is employed to remove low-quality scenes, including those contaminated by clouds

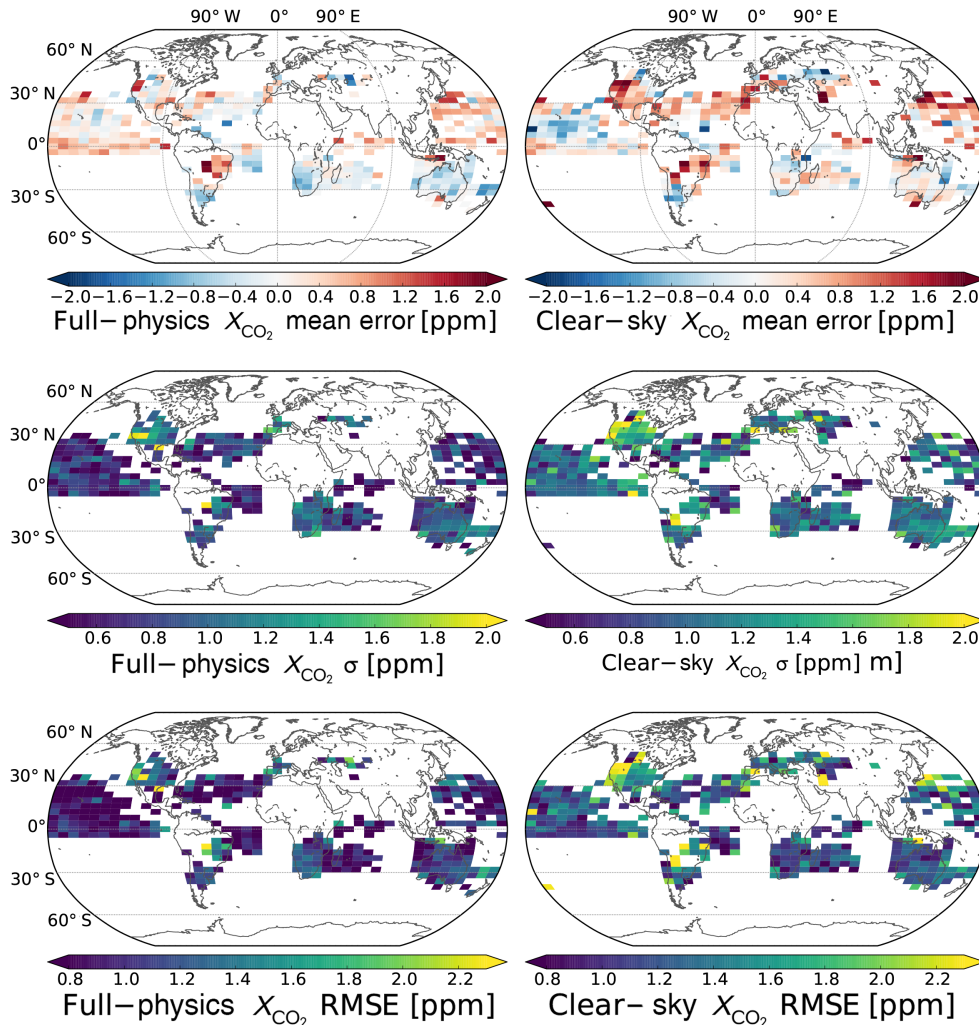


Figure 9. GOSAT full-physics (left column) and clear-sky (right column) retrieval mean X_{CO_2} errors (top row), standard deviation of the X_{CO_2} errors (middle row), and X_{CO_2} RMSEs (bottom row) for $8^\circ \times 4^\circ$ (longitude \times latitude) bins for a throughput of 30 %.

and aerosols. In this section, we explore whether this result is reproducible using real observations.

The effectiveness of applying DOGO to the pre-filtered GOSAT data sets is shown in Fig. 8. Initially, as in the OCO-2 simulations, there is a large reduction in the X_{CO_2} RMSE 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 scattering effects. Interestingly, the initial RMSEs over ocean are much smaller than in the simulations, and the initial reduction of error is modest. This may indicate that our pre-filtering and bias correction techniques are especially effective for retrievals over ocean. Regarding the parameters chosen by DOGO, the primary filters selected were dP_{ABP} and parameters related to the $2.0 \mu\text{m}$ CO_2 band signal. These parameters, as discussed in Sect. 5.1, relate to the light path modification effects of clouds and aerosols.

Over ocean surfaces, the clear-sky retrieval (dashed blue line) has larger X_{CO_2} RMSEs than the full-physics retrieval (solid blue line), even at very high levels of filtration (low throughputs). The clear-sky retrieval X_{CO_2} RMSEs over ocean range from ~ 1.0 to 2.0 ppm, depending on throughput. This error is about 0.5 – 1.0 ppm larger than the corresponding full-physics retrieval errors. As the throughput is initially decreased, the difference in error between the clear-sky and full-physics retrieval over ocean stays approximately constant. Once the throughput drops below $\sim 40\%$, however, the clear-sky errors begin to approach the full-physics errors. 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. However, even at low throughputs (less than 40%), the clear-sky retrievals still have X_{CO_2} RMSEs over ocean roughly 20 – 35% larger than those of the full-physics retrieval. This is in contrast to our simulation-based OCO-2 results, suggesting that addi-

tional unknown real-world mechanisms not included in the simulations may limit the performance 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.

Over land, the clear-sky retrieval (dashed orange line) has errors closer to the full-physics retrieval (solid orange line) compared to over ocean. For throughputs greater than $\sim 40\%$, the clear-sky retrieval typically has RMSEs $\sim 20\%$ larger than the full-physics retrieval. These errors are, as they were over ocean, roughly constant for the entire throughput range above $\sim 40\%$, which may suggest that even a moderate amount of contamination by clouds or aerosols prevents the clear-sky retrieval from performing as well as the full-physics retrieval for real GOSAT measurements. However, as the throughput is decreased further and most of the contaminated scenes are removed, the clear-sky errors become comparable to the full-physics errors. Unlike over ocean, this result qualitatively agrees with our simulated results over land and suggests that the data can be filtered well enough so that the clear-sky retrieval can perform as well as the full-physics retrieval.

The regional binned X_{CO_2} mean errors, standard deviation of the X_{CO_2} errors, and X_{CO_2} RMSEs are shown in Fig. 9. The data sets shown were post-filtered using DOGO with a throughput of 30%. The post-filtering, as was seen in the simulated OCO-2 data, has a preference to remove measurements in regions persistently contaminated by clouds or aerosols and measurements at high latitudes that had low signal to noise ratios. The Sahara, for example, is entirely devoid of measurements at a throughput of 30%, likely because of contamination by large dust particles. Over ocean, the clear-sky retrievals have mean X_{CO_2} errors similar to those of the full-physics retrieval. Regarding the standard deviation of the X_{CO_2} errors, the clear-sky retrieval generally has modestly larger values. Because of this, the clear-sky retrieval consistently also has modestly larger regional X_{CO_2} RMSEs than the full-physics retrievals. The results over land are more variable but are still in agreement with Fig. 8. The regional clear-sky biases may be slightly larger in a few places (e.g., North America), but this may partly be due to low number statistics. In general, the clear-sky mean errors, standard deviation of the errors, and RMSEs are comparable to their full-physics counterparts. The observed regional variability, compared to relatively uniform ocean error patterns, could be due to heterogeneous surface characteristics or cloud and aerosol compositions. Thus, we can not say with confidence that clear-sky retrievals perform better or worse for specific regions for real GOSAT data without further investigation, but in general the two retrievals are roughly equivalent over land when contaminated measurements are removed.

6 Conclusions

In this study we evaluated the performance of non-scattering, or “clear-sky”, X_{CO_2} retrievals performed on hyperspectral near-infrared measurements of reflected sunlight by comparing them to “full-physics” X_{CO_2} retrievals, which include scattering and absorption 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 applied, consistent with previous findings. However, with the application of pre- and post-filters 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 level.

For GOSAT measurements over land, the clear-sky retrieval has X_{CO_2} RMSEs 0–20% larger than the full-physics retrieval, when the data sets are filtered to remove scenes contaminated by clouds and aerosols. Over ocean, the clear-sky retrieval has X_{CO_2} RMSEs roughly 20–35% larger than the full-physics retrieval. The source of this extra error in the clear-sky retrieval applied to real GOSAT measurements, especially over ocean, 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. An alternative way to view this result is that the full-physics cloud and aerosol parameterization benefits GOSAT measurements over both land and ocean, but the improvement is more substantial over ocean.

For synthetic OCO-2 measurements and real GOSAT measurements over both land and ocean surfaces, the clear-sky retrieval has X_{CO_2} RMSEs less than 2.0 ppm when the data set is sufficiently filtered. In addition, clear-sky retrievals can be 1–2 orders of magnitude faster than full-physics retrievals, as scattering by clouds and aerosols can be completely ignored. For proposed sensors that collect an enormous volume of data, such as GeoCarb (Polonsky et al., 2014) and CarbonSat (Bovensmann et al., 2010), this could allow for significantly more data to be processed. Additionally, estimates of parameters from a clear-sky retrieval, such as surface albedo, could serve as a useful first guess for the full-physics retrieval.

Acknowledgements. The authors would like to thank Natalie Tourville, Scott Denning, and Chris Kummerow of Colorado State University for their contributions to this work along with the three reviewers for their helpful comments. Funding sources for this research include the NASA Jet Propulsion Laboratory (NASA JPL OCO-2 Subcontract #1439002) and the American Meteorological Society (AMS) Graduate Fellowship Program.

Edited by: H. Worden

References

- Aben, I., Hasekamp, O., and Hartmann, W.: Uncertainties in the space-based measurements of CO₂ columns due to scattering in the Earth's atmosphere, *J. Quant. Spectrosc. Ra.*, 104, 450–459, doi:10.1016/j.jqsrt.2006.09.013, 2007.
- Baker, D. F., Bösch, H., Doney, S. C., O'Brien, D., and Schimel, D. S.: Carbon source/sink information provided by column CO₂ measurements from the Orbiting Carbon Observatory, *Atmos. Chem. Phys.*, 10, 4145–4165, doi:10.5194/acp-10-4145-2010, 2010.
- 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 CO₂ fluxes estimated from GOSAT retrievals of total column CO₂, *Atmos. Chem. Phys.*, 13, 8695–8717, doi:10.5194/acp-13-8695-2013, 2013.
- Bovensmann, H., Buchwitz, M., Burrows, J. P., Reuter, M., Krings, T., Gerilowski, K., Schneising, O., Heymann, J., Tretner, A., and Erzinger, J.: A remote sensing technique for global monitoring of power plant CO₂ emissions from space and related applications, *Atmos. Meas. Tech.*, 3, 781–811, doi:10.5194/amt-3-781-2010, 2010.
- Butz, A., Hasekamp, O. P., Frankenberg, C., and Aben, I.: Retrievals of atmospheric CO₂ 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.
- 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.
- Chevallier, F., Bréon, F.-M., and Rayner, P. J.: Contribution of the Orbiting Carbon Observatory to the estimation of CO₂ sources and sinks: Theoretical study in a variational data assimilation framework, *J. Geophys. Res.*, 112, D09307, doi:10.1029/2006JD007375, 2007.
- Chevallier, F., Maksyutov, S., Bousquet, P., Bréon, F.-M., Saito, R., Tshida, Y., and Yokota, T.: On the accuracy of the CO₂ surface fluxes to be estimated from the GOSAT observations, *Geophys. Res. Lett.*, 36, L19807, doi:10.1029/2009GL040108, 2009.
- Chevallier, F., Ciais, P., Conway, T. J., Aalto, T., Anderson, B. E., Bousquet, P., Brunke, E. G., 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.: CO₂ surface fluxes at grid point scale estimated from a global 21 year reanalysis of atmospheric measurements, *J. Geophys. Res.*, 115, D21307, doi:10.1029/2010JD013887, 2010.
- 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.
- 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.
- 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.
- 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., Oya-fuso, 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.
- Crowell, S., Rayner, P., Zaccheo, S., and Moore, B.: Impacts of atmospheric state uncertainty on O₂ measurement requirements for the ASCENDS mission, *Atmos. Meas. Tech.*, 8, 2685–2697, doi:10.5194/amt-8-2685-2015, 2015.
- 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 CO₂ using ground-based, aircraft, and space-borne data, *Atmos. Chem. Phys.*, 11, 2789–2803, doi:10.5194/acp-11-2789-2011, 2011.
- Frankenberg, C., Platt, U., and Wagner, T.: Iterative maximum a posteriori (IMAP)-DOAS for retrieval of strongly absorbing trace gases: Model studies for CH₄ and CO₂ retrieval from near infrared spectra of SCIAMACHY onboard ENVISAT, *Atmos. Chem. Phys.*, 5, 9–22, doi:10.5194/acp-5-9-2005, 2005.
- Frankenberg, C., Pollock, R., Lee, R. A. M., Rosenberg, R., Blavier, J.-F., Crisp, D., O'Dell, C. W., 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.
- 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., Sussmann, R., Galli, A., and Aben, I.: Impact of aerosol and thin cirrus on retrieving and validating XCO₂ from GOSAT shortwave infrared measurements, *J. Geophys. Res.*, 118, 1–19, doi:10.1002/jgrd.50332, 2013.
- 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.
- IFS documentation, available at: <https://software.ecmwf.int/wiki/display/IFS/Official+IFS+Documentation>, last access: February 2016.
- 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.
- Kawa, S. R., Erickson III, D. J., Pawson, S., and Zhu, Z.: Global CO₂ transport simulations using meteorological data from the NASA data assimilation system, *J. Geophys. Res.*, 109, 18312–18329, doi:10.1029/2004JD004554, 2004.

- Kulawik, S., Wunch, D., O'Dell, C., Frankenberg, C., Reuter, M., Oda, T., Chevallier, F., Sherlock, V., Buchwitz, M., Osterman, G., Miller, C. E., Wennberg, P. O., Griffith, D., Morino, I., Dubey, M. K., Deutscher, N. M., Notholt, J., Hase, F., Warneke, T., Sussmann, R., Robinson, J., Strong, K., Schneider, M., De Mazzière, M., Shiomi, K., Feist, D. G., Iraci, L. T., and Wolf, J.: Consistent evaluation of ACOS-GOSAT, BESD-SCIAMACHY, CarbonTracker, and MACC through comparisons to TCCON, *Atmos. Meas. Tech.*, 9, 683–709, doi:10.5194/amt-9-683-2016, 2016.
- 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.
- Le Quééré, 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.
- 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 CO₂ flux estimates for 2009–2010 based on GOSAT and ground-based CO₂ observations, *Atmos. Chem. Phys.*, 13, 9351–9373, doi:10.5194/acp-13-9351-2013, 2013.
- 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.
- 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., Bösch, 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_{CO₂} data, *J. Geophys. Res.*, 112, D10314, doi:10.1029/2006JD007659, 2007.
- Nelson, R. R.: The Impact of Aerosols on Space-Based Retrievals of Carbon Dioxide, MSc thesis, Colorado State University, Fort Collins, USA, available at: <http://hdl.handle.net/10217/166883>, 2015.
- O'Brien, D. M. and Rayner, P. J.: Global observations of the carbon budget 2. CO₂ column from differential absorption of reflected sunlight in the 1.61 μm band of CO₂, *J. Geophys. Res.*, 107, 4354, doi:10.1029/2001JD000617, 2002.
- 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., Cooperative Institute for Research in the Atmosphere, Colorado State University, Fort Collins, USA, available at: ftp://ftp.cira.colostate.edu/ftp/TTaylor/publications/20090813_OCO_simulator.pdf, 2009.
- 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., Polonsky, I., Smyth, M., Taylor, T., Toon, G. C., Wennberg, P. O., and Wunch, D.: The ACOS CO₂ retrieval algorithm – Part 1: Description and validation against synthetic observations, *Atmos. Meas. Tech.*, 5, 99–121, doi:10.5194/amt-5-99-2012, 2012.
- Osterman, G., Eldering, A., Avis, C., O'Dell, C., Martinez, E., Crisp, D., Frankenberg, C., and Fisher, B.: ACOS Level 2 Standard Product Data User's Guide, v3.4, Jet Propulsion Laboratory, Pasadena, CA, USA, available at: https://co2.jpl.nasa.gov/static/docs/v3.4_DataUsersGuide-RevB_131028.pdf, 2013.
- 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 XCO₂ and XCH₄, Tech. rep., version 2, University of Leicester, Leicester, UK, 2011.
- 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.
- Polonsky, I. N., O'Brien, D. M., Kumer, J. B., O'Dell, C. W., and the geoCARB Team: Performance of a geostationary mission, geoCARB, to measure CO₂, CH₄ and CO column-averaged concentrations, *Atmos. Meas. Tech.*, 7, 959–981, doi:10.5194/amt-7-959-2014, 2014.
- Rayner, P. J. and O'Brien, D. M.: The utility of remotely sensed CO₂ concentration data in surface source inversions, *Geophys. Res. Lett.*, 28, 175–178, doi:10.1029/2000GL011912, 2001.
- 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 Optimal Estimation DOAS (BESD) Algorithm for the Retrieval of X_{CO₂}, Tech. rep., Institute of Environmental Physics (IUP) University of Bremen, Bremen, Germany, version 2, 2013.
- Rodgers, C. D.: *Inverse Methods for Atmospheric Sounding: Theory and Practice*, World Scientific, Singapore, 240 pp., 2000.
- 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.
- 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 spectra, *IEEE T. Geosci. Remote*, 50, 295–309, doi:10.1109/TGRS.2011.2160270, 2012.
- Taylor, T. E., O'Dell, C. W., Frankenberg, C., Partain, P. T., Cronk, H. Q., Savtchenko, A., Nelson, R. R., Rosenthal, E. J., Chang, A. Y., Fisher, B., Osterman, G. B., Pollock, R. H., Crisp, D., Eldering, A., and Gunson, M. R.: Orbiting Carbon Observatory-2 (OCO-2) cloud screening algorithms: validation against collocated MODIS and CALIOP data, *Atmos. Meas. Tech.*, 9, 973–989, doi:10.5194/amt-9-973-2016, 2016.
- 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.
- 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.
- Wunch, D., Wennberg, P. O., Toon, G. C., Connor, B. J., Fisher, B., Osterman, G. B., Frankenberg, C., Mandrake, L., O'Dell, C., Ahonen, P., Biraud, S. C., Castano, R., Cressie, N., Crisp, D., Deutscher, N. M., Eldering, A., Fisher, M. L., Griffith, D. W. T., Gunson, M., Heikkinen, P., Keppel-Aleks, G., Kyrö, E.,

Lindenmaier, R., Macatangay, R., Mendonca, J., Messerschmidt, J., Miller, C. E., Morino, I., Notholt, J., Oyafuso, F. A., Rettinger, M., Robinson, J., Roehl, C. M., Salawitch, R. J., Sherlock, V., Strong, K., Sussmann, R., Tanaka, T., Thompson, D. R., Uchino, O., Warneke, T., and Wofsy, S. C.: A method for evaluating bias in global measurements of CO₂ total columns from space, *Atmos. Chem. Phys.*, 11, 12317–12337, doi:10.5194/acp-11-12317-2011, 2011b.

Yokota, T., Yoshida, Y., Eguchi, N., Ota, Y., Tanaka, T., Watanabe, H., and Maksyutov, S.: Global concentrations of CO₂ and CH₄ retrieved from GOSAT: first preliminary results, *Scientific Online Letters on the Atmosphere*, 5, 160–163, doi:10.2151/sola.2009-041, 2009.