Toward High Precision XCO_2 Retrievals from TanSat Observations: Retrieval Improvement and Validation against TCCON Measurements

D. Yang[1,2,3]*, H. Boesch[1,4], Y. Liu[2,3], P. Somkuti[1,4,5], Z. Cai[2], X. Chen[2], A. Di Noia[1,4], C. Lin[6], N. Lu[7], D. Lyu[2], R. J. Parker[1,4], L. Tian[8], M. Wang[3], A. Webb[1,4], L. Yao[2], Z. Yin[8], Y. Zheng[6], N. M. Deutscher[9], D. W. T. Griffith[9], F. Hase[10], R. Kivi[11], I. Morino[12], J. Notholt[13], H. Ohyama[12], D. F. Pollard[14], K. Shiomi[15], R. Sussmann[10], Y. Té[16], V. A. Velazco[9], T. Warneke[13], D. Wunch[17]

[1]Earth Observation Science, School of Physics and Astronomy, University of Leicester, UK.

[2]Institute of Atmospheric Physics, Chinese Academy of Sciences, China.

[3]Shanghai Advanced Research Institute, Chinese Academy of Sciences, Shanghai, China.

[4]National Centre for Earth Observation, University of Leicester, UK.

[5]Colorado State University, Fort Collins, CO, USA.

[6]Changchun Institute of Optics, Fine Mechanics and Physics, China.

[7]National Satellite Meteorological Center, China Meteorological Administration, China

[8]Shanghai Engineering Center for Microsatellites, China.

[9]Centre for Atmospheric Chemistry, School of Earth, Atmospheric and Life Sciences, University of Wollongong, NSW, 2522, Australia.

[10]Karlsruhe Institute of Technology, IMK-IFU, Garmisch-Partenkirchen, Germany.

[11]Space and Earth Observation Centre, Finnish Meteorological Institute, Finland.

[12]National Institute for Environmental Studies (NIES), 16-2 Onogawa, Tsukuba, Ibaraki 305-8506, Japan.

[13]Institute of Environmental Physics (IUP), University of Bremen, 28334 Bremen, Germany.

[14]National Institute of Water and Atmospheric Research Ltd (NIWA), Lauder, New Zealand.

[15]Japan Aerospace Exploration Agency, Japan

[16]Laboratoire d'Etudes du Rayonnement et de la Matière en Astrophysique et Atmosphères (LERMA-IPSL), Sorbonne Université, CNRS, Observatoire de Paris, PSL Université, 75005 Paris, France.

[17]University of Toronto, Canada.

Corresponding author: Dongxu Yang (yangdx@mail.iap.ac.cn)

Key Points:

- First using O_2 A and 1.61 um CO_2 band approaching TanSat XCO_2 retreival
- Development a method on radiometric correction of TanSat L1B data in O_2 A and 1.61 um CO_2
- Validation of new TanSat retrieval against TCCON and recived significant improved results compare to previously retrieval

Abstract

TanSat is the 1st Chinese carbon dioxide (CO_2) measurement satellite, launched in 2016. In this study, the University of Leicester Full Physics (UoL-FP) algorithm is implemented for TanSat nadir mode XCO_2 retrievals. We develop a spectrum correction method to reduce the retrieval errors by the online fitting of an 8^{th} order Fourier series. The model and a priori is developed by analyzing the solar calibration measurement. This correction provides a significant improvement to the O_2 A band retrieval. Accordingly, we extend the previous TanSat single CO_2 weak band retrieval to a combined O_2 A and CO_2 weak band retrieval. A Genetic Algorithm (GA) has been applied to determine the threshold values of post-screening filters. In total, 18.3% of the retrieved data is identified as high quality compared to the original measurements. The same quality control parameters have been used in a footprint independent multiple linear regression bias correction due to the stronger correlation with the XCO_2 retrieval error. Twenty sites of the Total Column Carbon Observing Network (TCCON) have been selected to validate our new approach for the TanSat XCO_2 retrieval. We show that our new approach produces a significant improvement on the XCO_2 retrieval accuracy and precision when compared to TCCON with an average bias and RMSE of -0.08 ppm and 1.47 ppm, respectively. The methods used in this study can help to improve the XCO_2 retrieval from TanSat and subsequently the Level-2 data production, and hence will be applied in the TanSat operational XCO_2 processing.

1 Introduction

Carbon Dioxide (CO_2) has been recognized as the most important greenhouse gas causing climate change and global warming due to the rise in anthropogenic emissions since the industrial revolution. Accurate measurement of atmospheric CO_2 and accordingly reducing the uncertainties of CO_2 fluxes is a key requirement for meeting the "measurable, reportable and verifiable" mitigation commitments of the United Nations Framework Convention on Climate Change (UNFCCC) (https://unfccc.int/resource/docs/2007/cop13/eng/06a01.pdf) that is aimed at avoiding disastrous consequences caused by climate change. The existing in-situ surface CO_2 measurement networks provide a wealth of accurate data related to the global carbon cycle. Unfortunately, the sparse coverage of such networks is still a major limitation for global carbon cycle research and large uncertainties in our quantitative understanding of regional carbon fluxes remain.

The new generation of satellites with state-of-the-art near infrared (NIR) and shortwave infrared (SWIR) hyperspectral spectrometers bring us a step closer towards global mapping of CO_2 with sufficient accuracy, precision and coverage for reliable flux estimates on regional scales. Specifically, NIR/SWIR spectroscopy provides a means for measurements of the total column-averaged dry air CO_2 mole fraction (XCO_2) which captures the CO_2 signals in the lower troposphere including the boundary layer that can then be used to improve our knowledge on CO_2 surface fluxes (Kuang et al., 2002).

The European Space Agency (ESA) SCanning Imaging Absorption SpectroMeter for Atmospheric CHartographY (SCIAMACHY) (Bovensmann et al., 1999) on board the ENVIronmental SATellite (ENVISAT) that launched in 2002 and operated until 2012, was the first space-borne instrument to provide SWIR hyperspectral measurements of CO_2 (Buchwitz et al., 2005; Bösch et al., 2006; Reuter et al., 2011; Heymann et al., 2015), followed by the Greenhouse Gases Observing Satellite (GOSAT) from Japan (Kuze et al., 2009) and the Orbiting Carbon Observatory-2 (OCO-2) from the U.S. (Crisp et al., 2008), launched in 2009 and 2014,

respectively. These missions have significantly contributed to the effort to obtain global CO_2 measurements from space (Yokota et al., 2009; Eldering et al., 2017a) and subsequently for carbon flux studies (Maksyutov et al., 2013; Yang et al., 2017; Feng et al., 2017; Eldering et al., 2017b; Hakkarainen et al., 2019). When validated against measurements from the Total Carbon Column Observing Network (TCCON) (Wunch et al., 2011a), XCO_2 derived from GOSAT and OCO-2 has an accuracy of better than 2 part per million (ppm) (Cogan et al., 2012; Butz et al., 2011; O'Dell et al., 2012; Crisp et al., 2012; Oshchepkov et al., 2013; Yoshida et al., 2013; Yang et al., 2015; Kim et al., 2016; Buchwitz et al., 2017a; Wu et al., 2018), thanks to the high performance of these instruments and long-term calibration efforts (Yoshida et al., 2013; Kuze et al., 2014; Frankenberg et al., 2015; Rosenberg et al., 2016; Crisp et al., 2017). Further advances are expected from recently launched missions such as GOSAT-2 (Nakajima et al., 2019) and OCO-3 (Eldering et al., 2019) launched in 2018 and 2019 respectively, and future missions including MicroCarb (Bertaux et al., 2019) and CO2M (Kuhlmann et al., 2019).

China plays an important role in the global carbon budget as a major source of anthropogenic carbon (Le Quéré et al., 2018) due to its rapid economic development but also as a region of increased carbon sequestration thanks to a number of reforestation projects (Chen et al., 2019). In China, a series of ambitious projects on mitigation of carbon emission was kicked-off in the last 10 years, which include the first Chinese greenhouse gas monitoring satellite mission (TanSat), which is supported by the Ministry of Science and Technology of China, the Chinese Academy of Sciences, and the China Meteorological Administration. The TanSat mission was initiated in 2011 and successfully launched on 22 Dec 2016. TanSat started acquiring and archiving data operationally in March 2017 (Chen et al., 2012; Ran et al., 2019).

TanSat carries two instruments on-board: the Atmospheric Carbon Dioxide Grating Spectrometer (ACGS) and the Cloud and Aerosol Polarimetry Imager (CAPI). ACGS is a state-of-the-art hyperspectral grating spectrometer aimed at allowing XCO_2 retrievals (Liu et al., 2013; Wang et al., 2014; Lin et al., 2017) by measuring backscattered sunlight in three NIR/SWIR bands: the oxygen (O_2) A-band (758–778 nm with ~0.04 nm spectral resolution), the CO_2 weak band (1594–1624 nm with ~0.125 nm spectral resolution) and the CO_2 strong band (2042–2082 nm with ~0.16 nm spectral resolution). CO_2 column information is largely drawn from the weak CO_2 band which includes a series of strong but not saturated CO_2 lines which respond to even small variations in atmospheric CO_2 (Kuang et al., 2003; Bösch et al., 2006). The O_2 A band hyperspectral measurement contains information on aerosol and cloud scattering both in total scattering amount and scattering vertical distribution (Heidinger and Stephens, 2000; Corradini and Cervino, 2006; Geddes and Bösch, 2015). The strong CO_2 band provides information on aerosol and cloud scattering at longer wavelengths, in conjunction with information on CO_2 and water vapor (Wu et al., 2019). The CAPI is a multi-band imager which provides radiance measurements in five bands (365-408nm, 660-685nm, 862-877nm, 1360-1390nm, 1628-1654nm) from UV to NIR. In order to achieve more information on aerosol size, which has a significant impact on the wavelength dependence of aerosol optical properties, CAPI includes two polarization channels (660-685nm and 1628-1654nm) to measure the Stokes parameters (Chen et al., 2017a).

TanSat flies in a sun-synchronous low Earth orbit (LEO) with an equator crossing time around 13:30 local time. The satellite operates in three measurement modes: nadir (ND), glint (GL) and target (TG). ND provides the routine measurements over land which are obtained by tracking the principle plane (the principal plane is spanned by the vectors from the sun to the

surface footprint and from the surface point to the observer). We only use ND observations in this study. The swath width of TanSat measurements is ~18km across the satellite track and contains 9 footprints each with a size of 2 km × 2 km in nadir (Lin et al., 2017; Zhang et al., 2019). Nadir and glint mode alternate orbit-by-orbit, and the TanSat nadir model ground track is typically between two OCO-2 tracks, which provides potential future opportunities for combined usage of both data products for increased spatial coverage.

The first global XCO_2 dataset from TanSat observations for the first half year of operations has been produced with the Institute of Atmospheric Physics Carbon dioxide retrieval Algorithm for Satellite measurement (IAPCAS) (Yang et al., 2018), using the CO_2 weak band only. The nadir XCO_2 dataset has been evaluated using eight TCCON sites to show an average precision of 2.11 ppm for the TanSat retrievals (Liu et al., 2018).

In this study, we apply the University of Leicester Full Physics (UoL-FP) algorithm to the TanSat XCO_2 retrieval using a joint CO_2 weak band and O_2 A band retrieval, after adopting a new approach to radiometrically correct TanSat spectra. This new radiometric correction is introduced in Section 2 together with the UoL-FP TanSat retrieval approach. The applied quality filtering and bias correction methods are discussed in Section 3. Section 4 gives the results of the comparisons of the TanSat XCO_2 retrievals to ground-based observations from the TCCON network and Section 5 provides summary and outlook.

2 UoL-FP TanSat XCO_2 Retrieval Description

2.1 Introduction of the UoL-FP algorithm

The UoL-FP is an XCO_2 retrieval algorithm based on the Optimal Estimation Method (OEM) that has been originally developed for the NASA Orbiting Carbon Observatory (OCO) mission (Bösch et al., 2006), and has been extensively used for XCO_2 retrievals from GOSAT (Cogan et al., 2012). Besides XCO_2, the UoL-FP has been used to retrieve methane (CH_4) (Parker et al., 2011, 2015), HDO and H_2O (Boesch et al., 2015) and Solar Induce chlorophyll Fluorescence (Somkuti et al., 2020). UoL-FP is also one of the algorithms used to generate the Essential Climate Variables (ECV) XCO_2 and XCH_4 from GOSAT for the ESA Climate Change Initiative (CCI) (Dils et al., 2014; Buchwitz et al., 2017a) and subsequent European Commission Copernicus Climate Change Service (C3S) (Buchwitz et al., 2017b).

Several publications have already introduced the UoL-FP algorithm and its applications in detail (Parker et al., 2011; Cogan et al., 2012; Boesch et al., 2015; Somkuti et al., 2018) and here we only provide a brief overview. Basically, the retrieval aims to resolve an optimization problem to find the best estimate of a state vector $\hat{\boldsymbol{x}}$ by minimizing the difference between a measured and a simulated spectrum taking into consideration additional constraints on the measurement errors and state vector a priori uncertainties. This state vector gives all retrieved parameters and includes atmospheric, surface and instrument variables. Full physics refers to the fact that the algorithm generates the simulated spectrum via an accurate multiple-scattering Radiative Transfer (RT) model. As many processes involved in the transfer of light through the atmosphere respond in a non-linear way to changes in state vector elements, an iterative Levenberg-Marquardt (L-M) inversion scheme is used in the retrieval,

$$\boldsymbol{x}_{i+1} = \boldsymbol{x}_i + [(1+\lambda)\boldsymbol{S}_a^{-1} + \boldsymbol{K}_i^T\boldsymbol{S}_\varepsilon^{-1}\boldsymbol{K}_i)\,]^{-1}[\boldsymbol{K}_i^T S_\varepsilon^{-1}(\boldsymbol{y} - \boldsymbol{F}(x_i)) - \boldsymbol{S}_a^{-1}(\boldsymbol{x}_i - \boldsymbol{x}_a)], \quad (1)$$

where $\boldsymbol{x_a}$ is the a priori of the state vector. $\boldsymbol{S_a}$ and $\boldsymbol{S_\varepsilon}$ indicate the covariances of the state vector and the measurement respectively. The weighting function (known as Jacobians) $\boldsymbol{K}$ gives the linear change of the spectrum per change in state vector $\partial \boldsymbol{y}/\partial \boldsymbol{x}$. The update step of the state vector's i^{th} iteration from $\boldsymbol{x}_i$ to $\boldsymbol{x}_{i+1}$ can be adjusted by the L-M factor λ.

The forward model, which includes a vector RT model, is one of the most essential parts of the retrieval algorithm. In UoL-FP, the radiative transfer model LIDORT is used, which is a linearized discrete ordinate radiative transfer model that generates radiances and Jacobians simultaneously (Spurr et al., 2001; Spurr and Christi, 2014). According to the instrument design, ACGS/TanSat only measures one direction of polarized light instead of the total intensity; hence we need to compute the Stokes vector {*I*, *Q*, *U*, *V*} (Stokes, 1852; Liou, 2002; Mishchenko et al., 2004) in the forward simulation. Since multiple scattering is depolarizing, it is reasonable to expect that the polarization could be accounted for by a low-order scattering approximation. For a relatively clear atmosphere (e.g. aerosol optical depth <0.3), retaining only the second order of scattering components for *Q* and *U* is generally sufficient (Natraj and Spurr, 2007a). A 2-orders of scattering (2OS) model (Natraj and Spurr, 2007a) is applied in our RT model to extend the scalar LIDORT to vector simulation capability (Somkuti et al., 2017; Somkuti, 2018). The Low Stream Interpolation (LSI) method is used to speed-up the RT calculations (O'Dell, 2010).

The atmosphere is discretized into 20 layers from the surface up to 0.1 hPa allowing 10 sublayers within each layer to reduce interpolation errors from non-linear changes of the gas absorption cross sections with pressure and temperature. Absorption cross sections for O_2, CO_2, H_2O and CH_4 are considered in the RT computation. We use the NASA ACOS/OCO-2 v5.0 absorption coefficients (ABSCO) that have been extensively used in ACOS GOSAT and OCO-2 retrievals (Benner et al., 2016; Devi et al., 2016; Drouin et al., 2017).

UoL-FP employs two aerosol retrieval types representing large and small aerosol particles. The optical depth profile and optical properties for each type are inferred from the aerosol data from the Copernicus Atmosphere Monitoring Service (CAMS) (https://atmosphere.copernicus.eu/). There are five basic aerosol types, including sea salt, dust, organic matter, black carbon and sulphate, provided by CAMS. Beyond the type, we also consider hydrophobic and hydrophilic effects in computing optical properties, as organic matter and black carbon are separated into hydrophobic and hydrophilic particles, whereas sea salt and sulphate are treated as hydrophilic, and dust as hydrophobic only. For aerosol concentration and vertical distribution, we use the CAMS model data as a climatology which is created from the CAMS data for the years 2014-2016. In addition, a high-altitude cirrus retrieval type is included using the ice particle model of Baum et al., (2014). The surface reflectance is assumed to be Lambertian and is described by a mean albedo and its wavelength dependent slope for each band.

2.2 Polarization

As shown by the sensitivity study of Natraj et al. (2007b) and Bai et al. (2017), significant errors could be introduced in the O_2 A band and CO_2 weak band RT computation of radiances when using intensity instead of a combination of Stokes components, which can then cause significant errors in the CO_2 retrieval due to the wrong surface pressure and aerosol contributions (Butz et al., 2009; Geddes and Boesch, 2015).

In physical terms, the light measured by the instrument can be represented by a linear combination of Stokes components $\{I, Q, U, V\}$,

$$L = a_1 I + a_2 Q + a_3 U + a_4 V, \quad (2)$$

As the circular polarization component V is very weak in a realistic atmosphere, we ignore this parameter by setting a_4 to 0. The Stokes parameter coefficients a_1, a_2 and a_3 are determined by the satellite position, measurement geometry and the pointing direction of the polarizer, which are provided in the L1B data product. More detail on the definition of the polarization angle is given in Appendix A.

2.3 Radiometric Correction Approach of TanSat Spectra

2.3.1 TanSat solar calibration measurement

TanSat has multiple on-orbit calibration strategies, including solar, dark field and white light lamp calibration, which are helpful to monitor the instrument status and performance. TanSat switches to solar calibration measurements when it flies over the ascending end of each orbit until almost in darkness and regularly takes ~7 minutes (~1260 frames) of solar measurements. The solar calibration procedure provides direct measurements of the solar spectrum which has no contamination from the Earth's atmosphere and surface and thus without uncertainties from the radiative transfer of light. During the on-orbit test phase, solar calibration is performed once every two orbits, and then changes to daily during operational observations.

Ideally, we can imagine the solar calibration as a model for a measurement essentially without atmospheric extinction and scattering (there is a well-calibrated solar diffuser used in the solar calibration, which involves a reflection). Therefore, one can use solar calibration measurements to verify the radiometric and spectral calibration applied to a spectrum if an accurate solar model is used and the reflection from the diffuser is well known (Wang et al., 2018).

Here, we use the new solar line model created by G. C. Toon (2014) (https://mark4sun.jpl.nasa.gov/toon/solar/solar_spectrum.html) combined with the UoL-FP solar continuum model that is obtained from a polynomial fitting of SOLar SPECtrometer (SOLSPEC) measurements (Meftah et al., 2018). The solar line model has been extensively used and verified with GOSAT and OCO-2 retrievals (Uchino et al., 2012). This solar model combines the solar lines and solar continuum, and hence can be directly used for solar model fitting without any further calculation.

TanSat observes the Sun though a reflective diffuser before the relay optical system, hence the wavelength-dependent diffuser reflectance needs to be compensated for, otherwise it could cause an extra pattern in the measured solar spectrum. In this study, we use a wavelength-dependent Bidirectional Reflectance Distribution Function (BRDF) that has been characterized pre-flight in the laboratory (Wang et al., 2014) without any corrections for a time dependent degradation as the instrument performance is stable. The solar irradiance is non-polarized and we use a factor of 0.5 to adjust for the linear polarizer.

To avoid contamination of terrestrial absorption from the upper atmosphere when the satellite observes the limb measurement region, only measurements with boresight solar zenith angle (defined angle between line of sight and solar) between 5º and 6º are used in fitting. Note that TanSat rotates by a 5º pitch angle (away from the Earth) to avoid damage of CAPI from strong incident light (Chen et al., 2017a, 2017b).

2.3.2 Solar calibration analysis and radiometric corrections

Monitoring the solar calibration spectrum shows a stable instrument performance during the first year of TanSat operations (Figure 1, the CO_2 weak band is also stable but this is not shown here). The mean normalized solar measurement in the 750 nm region shows little change (mean normalized spectrum changes < 0.1%) for spectra acquired every 3 months, except for a very small difference for solar lines, which is probably caused by small instrument performance changes, e.g. instrument line shape (ILS). The figure shows that spectra acquired for the different cross-track footprints show some differences to each other in the continuum, which means each footprint has to be analyzed separately.

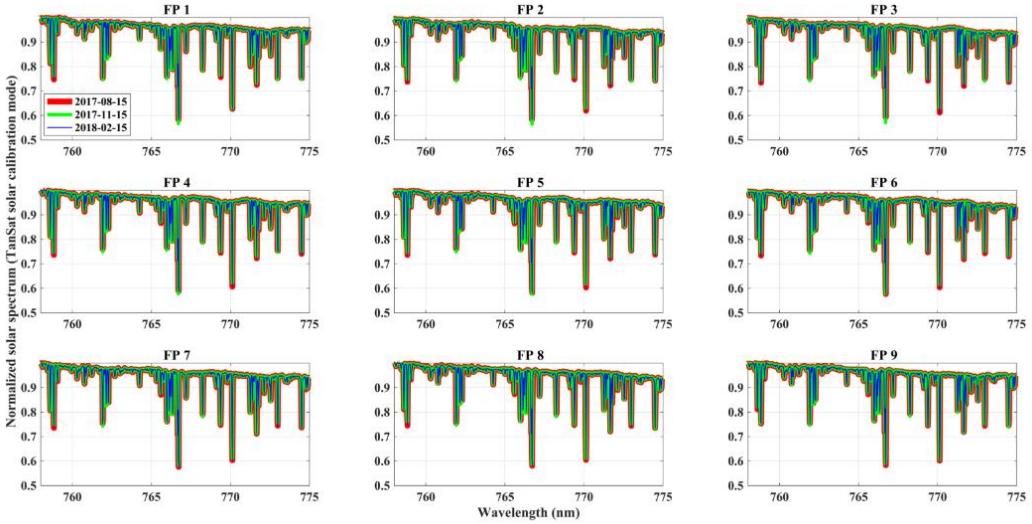

Figure 1. Mean spectrum of the normalized solar calibration measurement of the O_2 A band. The average is calculated for each calibration measurement in an orbit. Red, green and blue indicate the mean measurement taken for 2017/08/15, 2017/11/15 and 2018/02/15 respectively. Sub-figures represent the footprints (FP1-9) across the swath. The date is shown in the legend of FP1 sub-figure.

In order to further investigate the solar calibration spectrum, we developed a fitting tool for the solar spectrum based on the solar model corrected for the variable Earth-Sun distance and the Doppler shift effect (in wavelength) due to Earth's rotation and revolution. The fitting tool retrieves the wavelength dependence continuum correction in different forms (e.g. polynomial or Fourier series) as well as a wavelength shift and stretch simultaneously. The fitting tool uses the Gaussian-Newton method without the constraint of measurement noise.

The fitting algorithm is fast and hence we fit all individual solar spectra between March 2017 and May 2018 without averaging or manual spectrum selection. In total, 181,232 retrievals have been carried out and we find consistent fitting residuals for each cross-track footprint, which is expected considering the stable solar calibration measurement (Figure 1). Averaged relative fitting residuals for the NIR band are given in Figure 2. A considerable and consistent pattern with a mean RMSE of 0.48% remained in the fitting residual when we only adjust a wavelength-independent scaling factor to the continuum, which needs to be corrected to avoid errors in the XCO_2 retrieval. Thus, a method is needed that can compensate for these structures and improve the fitting residual. However, this pattern is not a simple linear or quadratic function

with wavelength and a more complex model is needed. We found that the relative residual did not change much with changes in intensity of the incident light introduced by instrument degradation and Earth-Sun distance changes and hence we adopt a correction model based on multiplicative continuum scaling rather than a model using additive offsets.

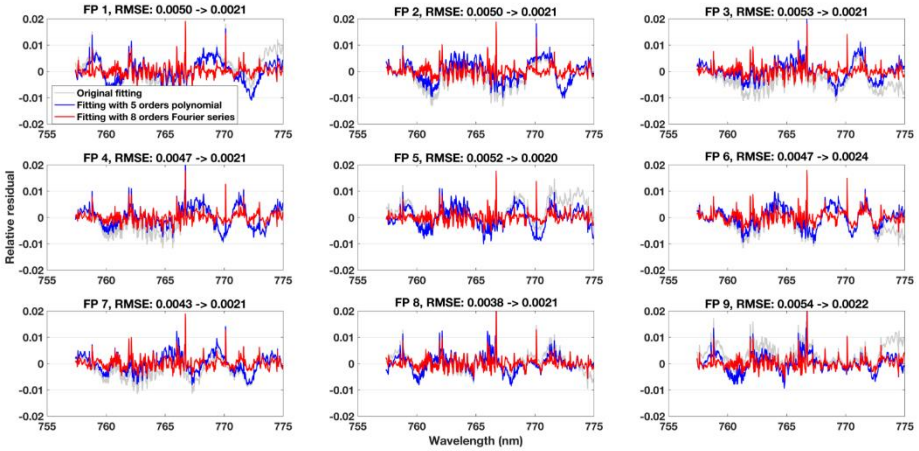

Figure 2. Mean fitting residuals of the solar calibration measurements. The average is calculated from the fitting residuals of 181,232 individual spectra (for each footprint, the quantity is ~20,136) during Mar. 2017 – May 2018. The blue and red lines show the mean residual obtained by fitting a 5th order polynomials or a 8th order Fourier series model, respectively (see the detail in text section 2.3). The light grey line is the fitting residual when fitting a scale factor without any wavelength dependent corrections. Sub-figures indicate the 9 footprints (FP1-FP9) across the swath. The RMSE shown in the title of each sub-figure shows the improvement of the Fourier series model compared to using the wavelength-independent scale factor.

A common approach is to use a polynomial as a function of wavelength to scale the continuum (L_{inc}). We have found that a 5th order polynomial is the best choice,

$$\mathrm{L} = \left(\sum_{i=0}^{5} a_i \cdot \Delta\lambda^i\right) \cdot L_{inc} + D(\lambda), \quad (3)$$

where a_i is the coefficient of the i^{th} order of the polynomial component for a wavelength relative ($\Delta\lambda = \lambda - \lambda_0$, λ_0 is the reference wavelength, in this study, we use first wavelength grid point of each band) to a reference of 750 nm. $D(\lambda)$ represents and additive offset assumed to be a linear function given by a slope and a constant. As can be seen in Figure 2, this polynomial approach leads only to minor improvements in the fitting residual. Increasing the order of the polynomial did not lead to significant improvements compared to the 5th order.

An alternative approach that should have a better capability to capture the oscillating nature of the fitting residuals is to use a Fourier series:

$$L = (c + \sum_{i=1}^{8}(a_i \cdot \cos(i \cdot \omega \cdot \Delta\lambda) + b_i \cdot \sin(i \cdot \omega \cdot \Delta\lambda))) \cdot L_{inc} + D(\lambda), \quad (4)$$

where a_i and b_i are coefficients of i^{th} order of the Fourier series cosine and sine components with c being the zero-order coefficient, and ω the scaling coefficient for frequency. $\Delta\lambda$ and $D(\lambda)$ are the same as for the polynomial model above. As a result of using Fourier series model, we find a significant improvement in the fitting residual with a mean RMSE of 0.21% (Figure 2) compared to 0.48% for the polynomial approach. Note that, (1) the fitting residual

311 near solar lines is still larger than the measurement noise, and (2) a gas absorption structure is
312 visible in the residuals, which is also the case in fit residuals for the CO_2 weak band. This could
313 be correlated with stray light from Earth or preflight calibration (e.g. ILS and radiometric
314 calibration), but the exact reason is unclear and needs to be further investigated in the future. In
315 principal, the continuum correction corrects the dominant effects in the fitting residual but some
316 features still remain visible in the residuals. For example, larger residual structures appear for
317 solar lines. Considering that this affects only a small number of pixels throughout a spectral
318 band, it can be expected that impact on the retrieval is limited and we do not attempt to further
319 correct these features.

320 The applied continuum correction in our study is basically a correction of the radiometric
321 gain and no other corrections are applied. This is because (1) errors of the continuum are more
322 obvious and they are stable in the case of TanSat, (2) solar lines are not sufficiently deep to
323 constrain potential non-linearity corrections, (3) the ILS cannot be easily re-analyzed from
324 space-based data, especially since TanSat did not provide solar calibration measurement for the
325 entire dayside which would scan the ILS due to the Doppler shift (Sun et al., 2017).

326 ### 2.4. O_2 A band surface pressure retrieval for cloud screening

327 The O_2 A band is important in the XCO_2 retrieval because, (1) it allows cloud screening
328 based on apparent surface pressure to remove measurements that are highly contaminated by
329 thick cloud, and (2) to provide information on aerosol and thin clouds in a joint O_2 A and CO_2
330 band retrieval to reduce errors that are otherwise introduced by light path modification from
331 scattering.

332 Table 1. List of state vector elements and descriptions for the UoL-FP/TanSat retrieval

Acronym	Description	N	A priori	A priori error (1σ)
CO_2 profile	Carbon dioxide (CO_2) concentration on each layer surface	21	LMDZ MACC-II CO_2 model	
H2O scale	Scaling for water vapor (H_2O) profile	1	ECMWF interim 0.75º	$\sqrt[2]{0.1}$
Temperature shift	Shift for temperature profile	1	ECMWF interim 0.75º	$\sqrt[2]{10}$ K
P surf	Surface pressure	1	ECMWF interim 0.75º	2 hPa
Albedo B1	Surface albedo of oxygen (O_2) A band	1	Estimate from spectrum continuum level	1
Albedo B1S	Surface albedo wavelength dependence slop of O_2 A band	1	0	0.0042
Albedo B2	Surface albedo of CO_2 weak band	1	Estimate from spectrum continuum level	1
Albedo B2S	Surface albedo wavelength dependence slope of CO_2 weak band	1	0	0.01
Aerosol M1	The profile of the most dominant aerosol type	21	Copernicus Atmosphere Monitoring Service (CAMS)	
Aerosol M2	The profile of the 2nd dominant aerosol type	21	Copernicus Atmosphere Monitoring Service (CAMS)	
Cirrus	The profile of cirrus	21		
Zeroff B1	The zero offset of O_2 A band	1	0	1% of continuum (SIF signal)
Zeroff B1S	The zero offset wavelength dependence slope of O_2 A band	1	0	1

Zeroff B2	The zero offset of CO_2 weak band	1	0	10% of continuum level
Zeroff B2S	The zero offset wavelength dependence slope of CO_2 weak band	1	0	1
Continuum B1	Fourier series correction coefficients on continuum of O_2 A band, 1 frequency scale and 16 coefficients of trigonometric function	17	Fitting results from long term solar calibration measurement	Fixed with experimental value
Continuum B2	Fourier series correction coefficients on continuum of CO_2 band, 1 frequency scale and 16 coefficients of trigonometric function	17	Fitting results from long term solar calibration measurement	Fixed with experimental value
Dispersion B1	Polynomials on wavelength grid for dispersion of O_2 A band	6	L1B data, dispersion coefficients O_2 A	Fixed with experimental value
Dispersion B2	Polynomials on wavelength grid for dispersion of CO_2 weak band	6	L1B data, dispersion coefficients CO_2 weak	Fixed with experimental value

An Oxygen A-Band (ABO2) cloud-screening algorithm is used as the cloud screening algorithm based on the analysis of a small number of micro-windows in the O_2 A band. The ABO2 algorithm has been applied to GOSAT and OCO-2 cloud screening (Taylor et al., 2012) and verified against MODIS and CALIOP measurement (Taylor et al., 2016). Unfortunately, the continuum correction described above cannot easily be applied to a retrieval that uses narrow micro-windows, and benefit from the fast RT model. Hence, we adopt an apparent surface pressure retrieval (assuming aerosol-free conditions) based on a fast RT model for the whole range of the O_2 A band (0.757-0.772 μm) that covers a multitude of O_2 lines and continuum for both sides of the band. This retrieval includes surface pressure, temperature profile offset, Lambertian surface albedo and its wavelength dependence slope, wavelength stretch and the coefficients of the Fourier series continuum model in the NIR as retrieved parameters. The outputs are used later as a priori values for a subsequent XCO_2 retrieval (Table 1). A priori surface pressure is taken from the European Centre for Medium-Range Weather Forecasts (ECMWF) ERA-Interim 0.75º × 0.75º reanalysis data product (Dee et al., 2011), and is interpolated to the location of the sounding and corrected for elevation differences using the U.S. Geological Survey's (USGS) EROS Data Center Shuttle Radar Topography Mission Global 30 Arc-Second Elevation (SRTM30) digital elevation model (DEM) (ftp://edcsgs9.cr.usgs.gov/pub/data/srtm/SRTM30).

We found that the frequency coefficient ($\boldsymbol{\omega}$ in equation (8)) of the Fourier series cannot be well-fitted in the O_2 surface pressure retrieval if the first guess is not very close to the true value due to non-linear effects. Since we observe that the structure in the fitting residuals of the solar spectra changes little with time we can obtain a good first guess from the fitting of solar calibration measurement value, not only for $\boldsymbol{\omega}$ but also for other coefficients. A set of 20,000 high quality solar calibration soundings (RMSE < 0.21%, which is the mean RMSE of all retrievals) including all 9 footprints have been selected for this calculation.

Another highly non-linear parameter is the stretch of the wavelength grid. The update of TanSat L1B data from version 1 to version 2 significantly improved the wavelength calibration, but additional corrections are still necessary. The solar calibration fitting also simultaneously provides wavelength stretch coefficients which we then use in the O_2 A band surface pressure retrieval.

The impact of the Fourier continuum correction on the surface pressure retrieval is significant (Figure 3). We found that (1) the retrieval without the correction has a large bias and scatter, and (2) there are large differences between the cross-track footprints in the retrieval even after continuum correction (with a constant gain factor). Therefore, the application of the apparent surface pressure retrieval for cloud screening without continuum correction is problematic due to the different scatter and bias for different footprints which could mean that a large number of clear-sky measurement will be screened out for some footprints. As is shown for the case given in Figure 3, the surface pressure values for the retrieval with continuum correction mostly fall in the range between -10 to 0 hPa which satisfies the commonly used criterion for cloud screening of -20 to +20 hPa. In contrast, the retrieval without correction spreads between -10 to -30 hPa.

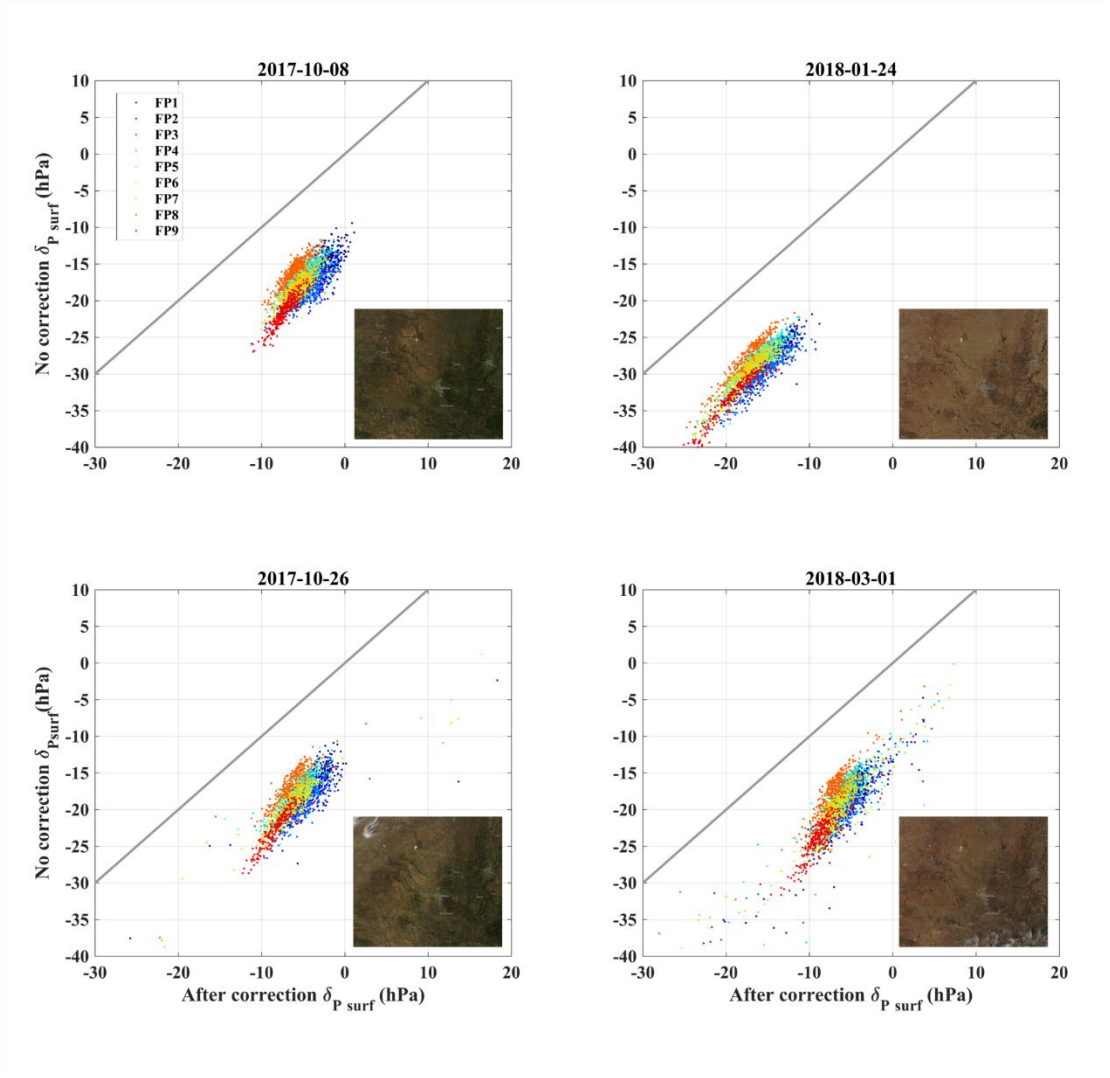

Figure 3. Comparison of surface pressure changes from O_2 A band retrievals with and without Fourier series continuum correction for selected clear-sky conditions around TCCON/Lamont (USA) site. The changes are calculated by subtracting the a priori (height corrected surface pressure from ECMWF interim) from the retrieved apparent surface pressure. The sub-figures show the selected cases on 2017-10-08, 2017-10-26, 2018-01-24 and 2018-03-01. Corresponding

images from MODIS/Aqua (https://worldview.earthdata.nasa.gov) are shown in the bottom right corner of each corresponding sub-figure. The color represents the footprints (FP1-FP9) across the swath as indicated in the upper left sub-figure. The retrieved apparent surface pressure with continuum correction still shows a small bias to the a priori, because the retrieval has been carried out without any consideration of aerosol and cloud scattering (see the detail in text section 2.4).

A ±20 hPa threshold for cloud screening is reasonable and relatively loose, which means more data is allowed to pass the cloud screening. The benefit is obvious: we do not include aerosol and cloud scattering corrections in the apparent surface pressure retrieval, so that there will always be a small bias in the retrieved surface pressure which can become more significant for very dark surfaces (Figure 3).

2.5 TanSat XCO_2 two-band retrieval

The information on CO_2 volume mixing ratio (VMR) comes to a large extent from the 1.6 μm CO_2 weak band, which means that a retrieval based on only the weak band can provide a relatively accurate result if the measurement scene is not impacted by aerosols or if perfect knowledge on aerosols would be available. Unfortunately, both are not possible for real scenarios. The preliminary TanSat XCO_2 retrieval (version 1.0) has been created using the CO_2 weak band only (Yang et al., 2018; Liu et al., 2018). An extremely stringent filter has been applied in post screening, and hence screens out a large number of retrievals. The data produced with this approach provides good information on the global CO_2 distribution and trend but does not yield enough quantity or accuracy for reliable surface flux inversions.

The O_2 A band surface pressure retrieval with our newly developed continuum correction shows reliable results; hence we can extend the CO_2 retrieval with confidence to use the O_2 A and CO_2 weak band together to improve the retrieval accuracy. We have applied a two-band retrieval to produce XCO_2 data from soundings that are recognized as clear-sky scenes by the cloud screening. The retrieval uses a state vector that includes a CO_2 profile, scale factors for temperature and water vapor, surface pressure, surface albedo and spectral slope. In addition, we fit an additive zero offset and its wavelength dependent slope in both the O_2 A and CO_2 weak bands. For the CO_2 weak band, the same correction model as for the O_2 A band is applied. The full state vector is given in Table 1.

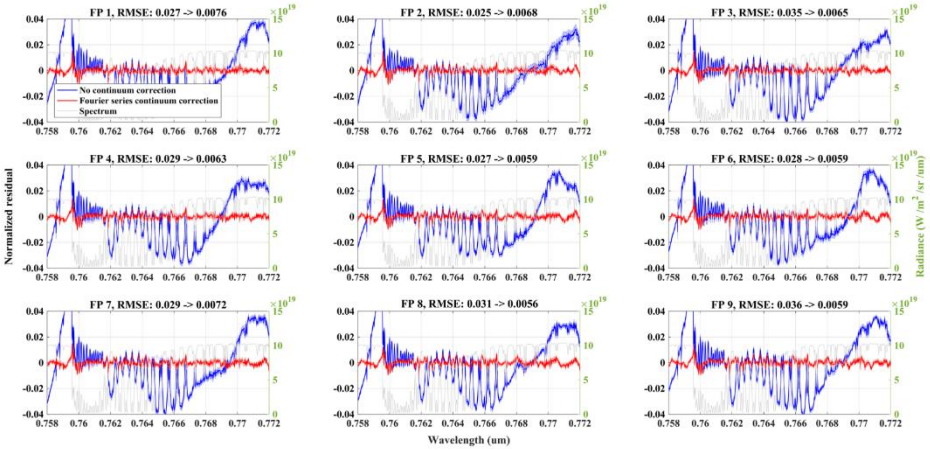

Figure 4. Mean O_2 A band fitting residuals (normalized by continuum level) from the two bands retrieval with (red) and without (blue) Fourier series continuum correction. The average is computed from selected clear-sky measurements on 08/10/2017 around the TCCON/Lamont site. The corresponding MODIS/Aqua RGB image is shown in Figure 3. The light grey line (right y-axis) shows the measurement spectrum as reference. The red and blue shading indicates the continuum normalized standard deviation (SD) for the retrieval with and without Fourier series continuum correction respectively.

Significant improvements to the O_2 A band residual have been found when using the Fourier series correction. The Standard Deviation (SD) of the normalized residual indicates that the fitting is stable (Figure 4). Notice that the residuals of the 9 footprints show a slight difference even with the correction, e.g. at the long-wavelength end that contain few atmospheric absorption features, which means this correction cannot perfectly compensate all of the spectral patterns. The residual still contains structures related to O_2 lines, which is because (1) the spectroscopy is not perfect (Connor et al., 2016), and (2) the observed residual features for the solar calibration fitting which we discussed in section 2.3. The improvement for the CO_2 weak band is also large (Figure 5) and the RMSE is reduced by almost half for the retrieval with the continuum correction. Similar to the O_2 A band, minor residual patterns remain which need to be investigated in the future.

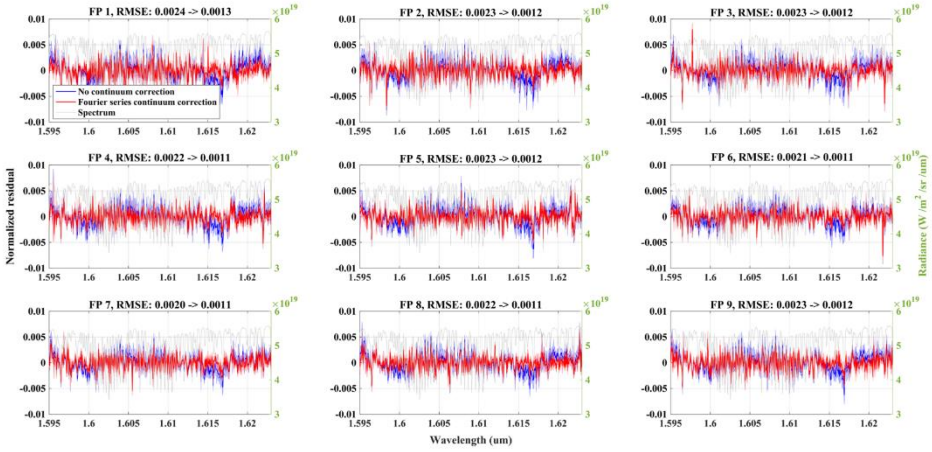

Figure 5. Same as figure 4, but for the CO_2 weak band.

3 Quality control and bias correction

3.1 Dataset

To remove outliers and to correct for small biases, a quality control filter and bias correction is applied using a reference dataset that is reliable enough for indicating retrieval errors. The Total Carbon Column Observing Network (TCCON) provides accurate measurement of XCO2 (Wunch et al., 2011a, 2015), and has been used for GOSAT and OCO-2 filtering and bias correction (Wunch et al., 2011b, 2017; Oshchepkov et al., 2013; Yoshida et al., 2013; Kim et al., 2016; Wu et al., 2018). The OCO-2 team has also developed additional methods for filtering and bias correction for their version 8 product, including small area approximation, multi-model median and southern hemisphere approximation (O'Dell et al., 2018). In this study

441 we only focus on the retrieval around TCCON sites, and hence only the data of 20 TCCON sites
442 (Table 2) has been selected as a reference dataset for bias/filtering and validation in this study.

443 Table 2. The TCCON site list used in the validation study and the site validation statistics

Site	Date range	Validation		
		N	bias (ppm)	RMSE (ppm)
Bialystok, Poland (Deutscher et al., 2019)	20090301-20180427	2	0.78	0.93
Bremen, Germany (Notholt et al., 2019)	20070115-20180420	1	-0.29	0.29
Burgos, Philippines (Morino et al., 2018c; Velazco et al., 2017)	20170303-20180427	2	0.27	1.10
Darwin, Australia (Griffith et al., 2014b)	20050828-20180308	12	0.29	1.36
East Trout Lake, Canada (Wunch et al., 2018)	20161007-20181231	19	0.21	1.12
Edwards, USA (Iraci et al., 2016)	20130720-20181231	3	1.36	1.39
Garmisch, Germany (Sussmann and Rettinger, 2018a)	20070716-20181220	5	0.24	1.18
JPL, USA (Wennberg et al., 2014)	20110519-20180514	20	-1.12	1.39
Karlsruhe, Germany. (Hase et al., 2015)	20100419-20190121	6	0.33	1.67
Lamont, USA (Wennberg et al., 2016)	20080706-20181231	17	0.37	0.76
Lauder, New Zealand (Sherlock et al., 2014)	20100202-20181031	9	1.19	1.40
Orléans, France (Warneke et al., 2019)	20090829-20180405	2	1.40	1.83
Paris, France. (Te et al., 2014)	20140923-20180126	4	0.048	0.62
Park Falls, USA (Wennberg et al., 2017)	20040602-20181229	15	0.41	1.20
Pasadena, USA (Wennberg et al., 2015)	20120920-20181231	19	-1.41	1.84
Rikubetsu, Japan (Morino et al., 2018b)	20131116-20180423	4	-0.85	1.12
Sodankylä Finland (Kivi et al., 2014)	20090516-20181030	9	1.17 (0.35)*	2.83 (1.25)*
Saga, Japan (Shiomi et al., 2014)	20110728-20181021	13	-0.92	1.53
Tsukuba, Japan (Morino et al., 2018a)	20110804-20180427	7	-1.04	1.62
Wollongong, Australia (Griffith et al., 2014a)	20080626-20180425	5	0.90	1.23
Zugspitze, Germany (Sussmann and Rettinger, 2018b)	20150424-20181218	-	-	-

444 * Large bias point removed

445 The quality control (filtering), bias correction and validation is based on an inter-
446 comparison of UoL-FP/TanSat retrievals against TCCON measurement, and those two results
447 have been obtained from two different instruments and retrieval algorithms with different

averaging kernels and a priori assumptions. A method for removing smoothing error differences caused by the different instruments and retrieval algorithm (Rodger and Connor, 2003) has been used in GOSAT (Cogan et al., 2006) and OCO-2 validation (O'Dell et al., 2018; Wunch et al., 2017). In practice, the application of this correction has led to small changes of <0.3 ppm (Nguyen et al., 2014; O'Dell et al., 2018). In this study, we directly compare the TanSat retrieval with TCCON measurement without attempting to remove differences in smoothing errors.

A colocation criterion of ±3° in latitude/longitude for matching up the TanSat soundings and TCCON measurements is applied. Other spatial co-location methods exist which allow to increase the quantity of matched up data, which is important for GOSAT (Wunch et al., 2011b; Guerlet et al., 2013), but since TanSat has a much higher data density compared to GOSAT, the spatial criterion above does provide a sufficient number of soundings. All TCCON measurements within ±1 hour of the satellite overpass are averaged to provide a reference dataset for each overpass, and only overpasses with more than 20 TCCON measurements within the 2 hour period are used. In this study, we found 396,068 co-located TanSat retrievals across the 20 TCCON sites.

Using the TCCON average as reference, we define the individual retrieval error as the XCO2 difference between each TanSat retrieval ($\widehat{\boldsymbol{C}}_{\boldsymbol{TanSat}}$) and the TCCON average ($\overline{\boldsymbol{C}}_{\boldsymbol{TCCON}}$) for a satellite overpass: $\boldsymbol{\delta}_{\boldsymbol{xco_2}} = \widehat{\boldsymbol{C}}_{\boldsymbol{TanSat}} - \overline{\boldsymbol{C}}_{\boldsymbol{TCCON}}$. Hence, there will be hundreds of individual TanSat retrievals for each average TCCON value for an overpass, except when heavily contaminated by clouds.

3.2 Semi-autonomous sounding selection

3.2.1 Method

The retrieval algorithm adjusts a number of parameters related to the atmosphere, surface and instrument and due to the complexity and non-linearity of the retrieval problem and limitations of the forward model to perfectly model the real behavior of the instrument, some retrievals will converge to a false solution or a local minimum. Therefore, as a first step, a quality control filter is applied to the retrievals to screen out such outliers before using the XCO_2 retrievals for bias correction and validation.

The main goal of a filter design is to efficiently screen out the poor retrievals (defined by large individual errors $\boldsymbol{\delta}_{\boldsymbol{xco_2}}$) while keeping a maximum number of high-quality retrievals. Assuming that large errors are introduced by an imperfect forward model and/or measurement, we expect errors to correlate with other parameters used in the retrieval that are adjusted together with CO_2.

The quality control normally comes with two basic questions, namely how to select the filter parameters and how to decide the best threshold values.

To answer the first question, we select 37 candidate filters and calculate the correlations with $\boldsymbol{\delta}_{\boldsymbol{xco_2}}$ (only the top 8 have been listed in Table 3). The significance of each candidate filter is sorted with respect to the correlation coefficient, and the candidates with the lowest impact on $\boldsymbol{\delta}_{\boldsymbol{xco_2}}$ (correlation coefficient < 0.3) are abandoned at the beginning. We select the candidates according to their rank, which means the first candidates have the highest priority to be chosen. Actually, it is not possible to decide on the best choice of filters (we call the number of filters

complexity from hereon in) before the analysis on the performance for all possible complexity options has been carried out.

Table 3. Selected filter used in quality control and corresponding lower and upper thresholds

Name	Description	Lower boundary	Upper boundary
Grad CO2	The retrieval changes of layer CO_2 gradient between 700hPa and the surface	-4.34	21.47
Delta Psurf	The retrieval changes on surface pressure from a priori	-4.45	1.99
Continuum B1C3	Continuum correction coefficients of cos(x) of O_2 A band	-0.76	0.60
Zeroff B2S	Zero offset wavelength dependence slope of CO_2 band	-0.14	0.017
AlbedoB2	Surface albedo of CO_2 weak band	0.033	0.33
Zeroff B1S	Zero offset wavelength dependence slope of O_2 A band	-	-
H2Oscale	Scale factor of H_2O	-	-
Continuum B1C4	Continuum correction coefficients of sin(x) of O_2 A band	-	-

The solution to the latter question is often approached in an empirical manner. Physical basis methods have been developed and applied in the NASA Atmospheric CO_2 Observations from Space (ACOS) OCO-2 retrieval for versions V7 (Wunch et al., 2011a), and in subsequent versions V8 (O'Dell et al., 2018) and V9 (Kiel et al., 2019). In this study, we hope to use a method that is less driven by empirical decisions. A machine learning Genetic Algorithm (GA) method has been developed and applied to OCO-2 to generate warn level data (Mandrake et al., 2013). The algorithm optimizes complexity (how many filters are selected), threshold value and transparency (how many data points pass the filter, represented by percent) simultaneously. Mandrake (2013) use one additional species of gene to control the filter selection and optimize this selection simultaneously, which treats each candidate equally, and then causes the filter combination for each complexity to be different. In this study, the filter has been selected according to the rank of the $\boldsymbol{\delta_{xco_2}}$ correlation, which means that the selection of the filter combination is carried out after the complexity is fixed. For each GA, runs with different complexity from 2 to 8 are carried out and shown in Figure 6. We optimize the threshold values of all filters for each GA run with chosen complexity with different transparency simultaneously (see more details of the GA that is used in this paper in Appendix B).

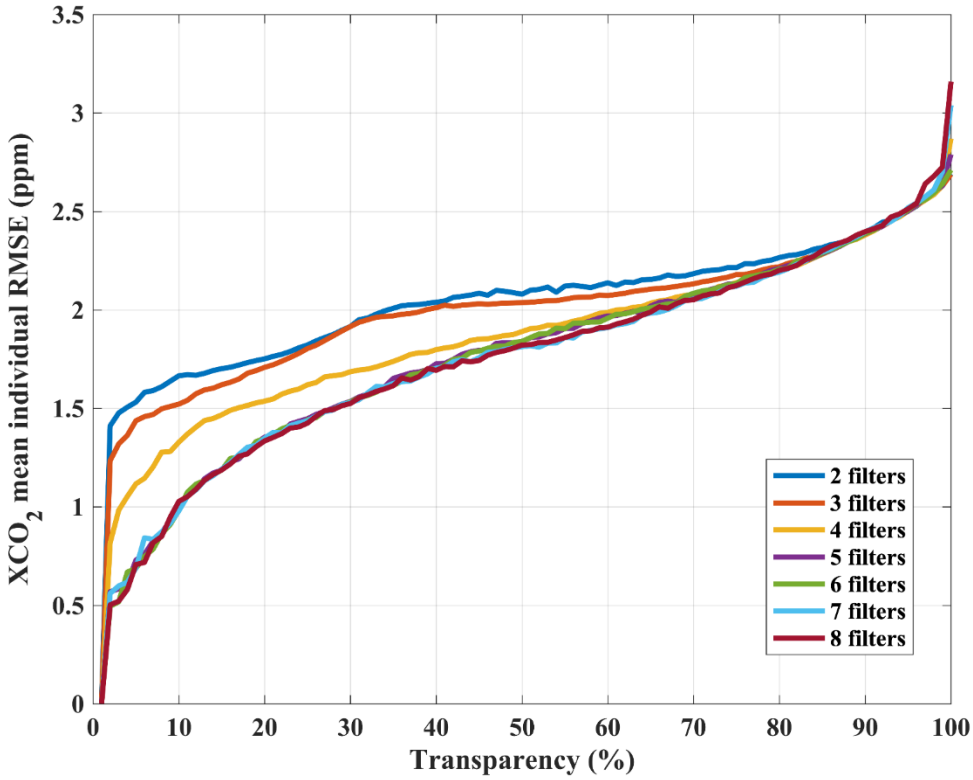

Figure 6. The optimal Target run Genetic Algorithm (GA) profile (Pareto-optimal trade-off curves) for the selection of filter complexity and transparency. The XCO_2 mean individual RMSE is a total RMSE calculated from all retrievals which pass the optimized threshold value (good retrieval). The color indicates the number of parameters (complexity) that are used in the GA run. For each complexity, the filter is determined and fixed (see section 3.2.1 for details and table 3 for filter definitions). Transparency has a 1% interval through 0-100%. No significant additional reduction in the RMSE was seen when using more than 5 filters. In this study, we cut-off at 2 ppm with 5 filters, the transparency is 64.3%.

Subjective selection of transparency and complexity is required at the end of the applied GA. The transparency for different complexities against RMSE is shown in Figure 6. It needs to be noted that for each complexity the filter is fixed, which means the complexity +1 represents the addition of an extra filter (Table 3). Few improvements on filter selection have been found when the complexity is larger than 4 for the determination of OCO-2 warn levels (Mandrake et al., 2013). We also found similar results for a complexity of 5 when compared to the complexity runs with values of 2-8 (Figure 6). The advantages of multi-feature combinations for more than 4 features appears only when the fraction of filtered-out data is close to 50%, and there are few advantages if the fraction of filtered-out data becomes less than 30% (transparency > 70%). For a target of 2 ppm RMSE, we select a transparency between 64-65% (64.3%) with five filters.

3.2.2 Application of the Filter

Retaining as much data as possible for a given requirement of RSME is an advantage of the GA method. However, GA can give an optimal solution in a mathematical sense, which may

not be physically reasonable; namely the filter thresholds are unrealistic. For contrasting the results obtained with GA against manual selection, we also applied an empirical selection of filter thresholds for the five first filters applied in GA (Table 3). The results are compared in Figure 7 using the bin-error plot method which is similar to that used by O'Dell (2018) for OCO-2 post screening. In general, we find that the thresholds from GA and the empirical method (manual selection) have similar effects on the filtering. However, to achieve a similar RSME with the empirical selection of thresholds, we reduce the amount filtered data by an additional 13.5% compared to GA.

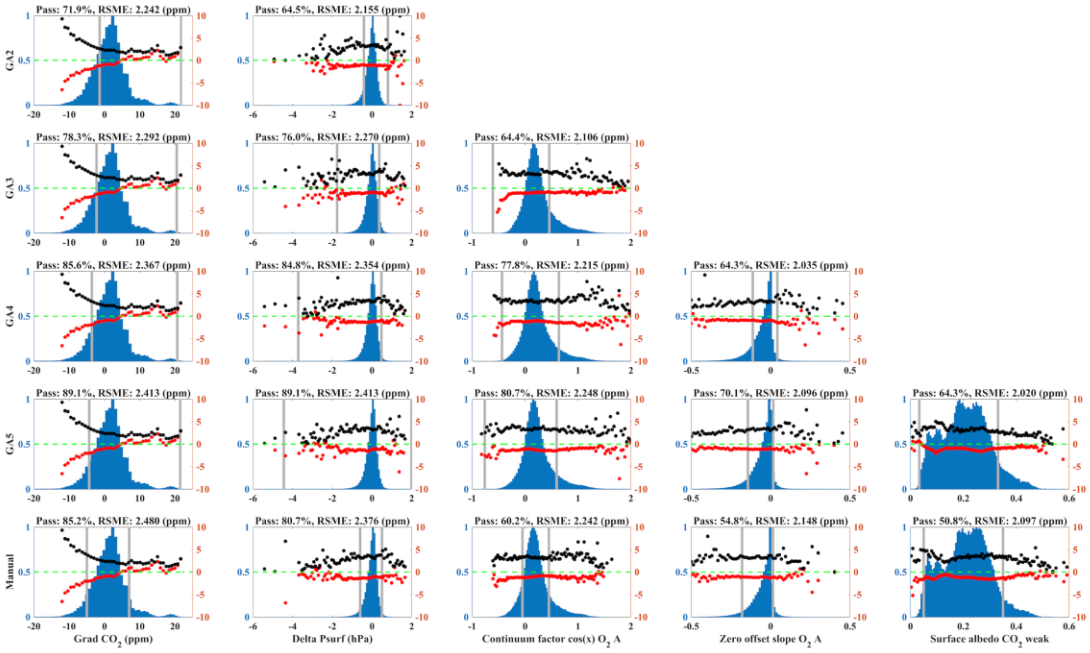

Figure 7. The performance of the GA with 2-5 filters (rows 1-4) and a manual filter threshold selection with 5 filters (row 5). The histograms (left y-axis) indicate the counts for each filter bin, and red and black solid points show the bias and RMSE for each bin (right y-axis). The grey line is the upper and lower threshold used for each filter. The filters are applied from left to right (columns 1-5) sequentially.

Some filters, e.g. Grad CO_2, Delta Psurf, and AlbedoB2, are parameters which have a specific physical meaning. Although these parameters are constrained during the retrieval, unreasonable results still occasionally occur. The GA that has been applied in this study has no capability of judging if a threshold has a reasonable value or not. Therefore, in practice, we strongly recommend applying additional physical filters to remove unreliable retrievals, if needed. However, in this study we only use GA screening.

In summary, our retrieval convergence rate is 94.8% of the cloud-free measurement (~30% of the original measurements for a 20hPa threshold of the apparent surface pressure from the O_2 A band) and 64.3% have been recognized as good retrievals by GA. In total, we keep ~18.3% of all nadir measurements, and subsequently apply a bias correction to them.

3.3 Bias correction

The bias correction is applied after the quality control. Biases in retrieved XCO_2 can be introduced by shortcomings in the retrieval algorithm (e.g. parameterized models and radiative transfer), the measurements (e.g. stray light or calibration) and the databases which are used (e.g. gas absorption and solar model). The former two will typically lead to variable biases that will depend on other parameters while the latter one more likely induces a global bias. Commonly, the parameter-dependent bias dominates and can be corrected (Wunch et al., 2011b) using a linear combination of identified bias parameters, $\Delta_{XCO2} = \sum_{i=1}^{n} a_i \cdot p_i + b$, with a_i being the linear coefficient for the ith parameter p_i, and b an offset.

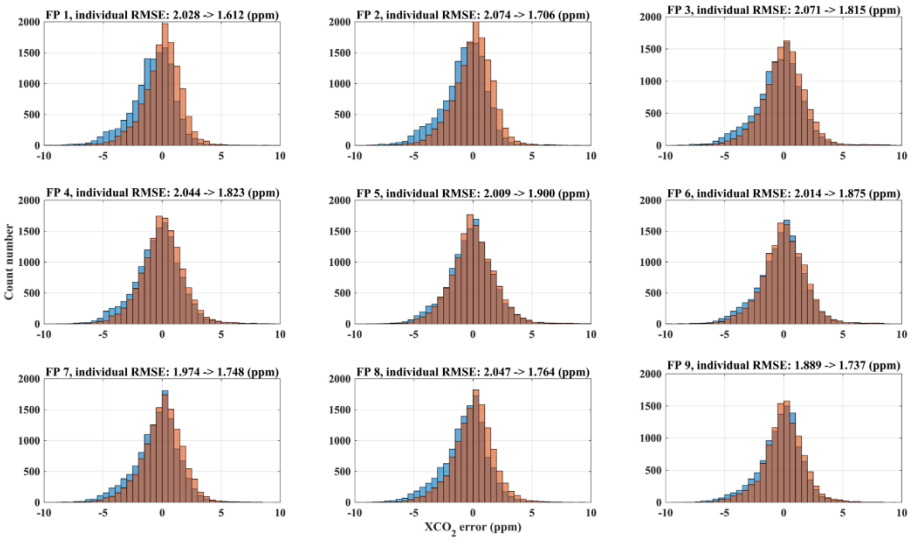

Figure 8. The individual XCO_2 error (UoL-FP/TanSat against TCCON) change with and without parameter bias correction. The orange and blue histograms indicate the XCO_2 individual error distribution with and without bias correction. The improvement of RMSE with and without bias correction for each footprint (FP 1-9) across the swath is shown in the titles.

For the bias correction, we use the filters that have already been applied in the quality control as these five parameters are most strongly correlated with the error in XCO_2. A multiple linear regression is applied to find the coefficients $\boldsymbol{a_i}$ for each across-track footprint. A further small improvement in RMSE is found with increasing the number of parameters up to 16 (SI section 1). The most significant improvement appears for the first three parameters and further parameters have a smaller effect. Improvement becomes less significant when using more than 12 parameters. Here, we use five parameters to avoid over-optimizing the bias correction. The number and selection of filter parameters is somewhat subjective and has been made to agree with the quality control.

The effect of the bias correction is shown in Figure 8. The largest improvements in RMSE are found for footprint 1, 2, 7, and 8. Besides a parameter-related bias, there can also be a footprint dependent bias as has been shown for OCO-2. Typically, a stable and constant bias mainly relates to instrument effects (O'Dell et al., 2018; Wu et al., 2018). TanSat has 9 footprints in a swath across ~18 km on the ground on average. In our retrieval, we also investigate the cross-footprint bias by subtracting the mean value of a swath when all 9 across-track footprints

are available. After the independent parameter bias correction, no obvious cross-track footprint bias is found and the average bias is less than 0.1 ppm but with a large (>0.3 ppm) standard deviation (SD). This result is also true, when we select the swaths for low variation of surface albedo across the swath (SD<0.01) to guarantee that the scene is comparable through all footprints (only 532 swaths satisfy this criteria). This result indicates that parameter bias correction, if carried out for each footprint independently, can reduce the across-track bias.

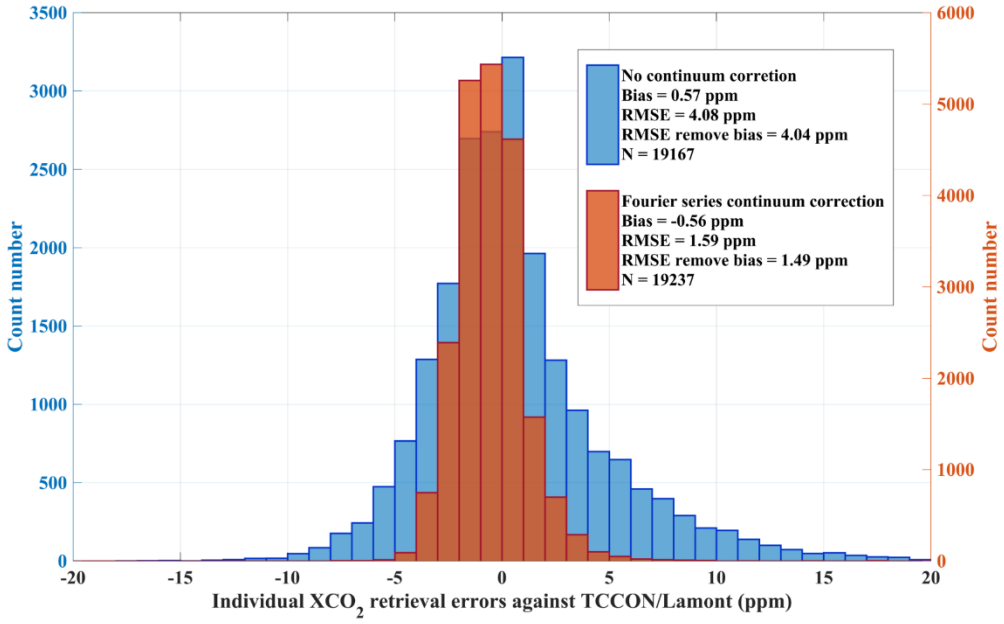

Figure 9. Histogram of individual XCO_2 retrieval errors (difference between TanSat and TCCON XCO_2) for TCCON/Lamont for clear-sky measurements. The orange (right y-axis) and blue (left y-axis) give the results for the two-band retrieval with (orange) and without (blue) Fourier series continuum correction. All data statistics in this figure passed quality control, but there is no bias correction applied. The RMSE remove bias shown in text box of this figure gives the RMSE calculated from each individual error after subtracting the overall bias.

4 XCO_2 Retrievals over TCCON sites

4.1 The discussions on two-bands retrieval

The Fourier series approach, introduced in section 2.3, has been instrumental in allowing a two-bands retrieval that uses the O_2 A band together with the weak CO_2 band. As shown in section 2.3, the continuum correction for the O_2 A band in the XCO_2 retrieval leads to a significantly improved fitting residual for all 9 footprints. We have also found that the apparent surface pressure retrieval (sect.2.4) from the O_2 A band yields reliable results when adopting the continuum correction. In this study, we retrieve XCO_2 from TanSat nadir measurements from March 2017 to May 2018 around 20 TCCON sites (Table 2, and see detail in sect.3.1) by using the setup that has been introduced in section 2.5. The continuum correction effect (as discussed in sect.2.3) on XCO_2 is shown in Figure 9 for TanSat retrievals around the TCCON/Lamont (USA) site using all TanSat retrievals that pass the quality control but without bias correction. The RMSE decreases from 4.08 ppm to 1.59 ppm, while the bias changes from 0.57 ppm to -0.56

ppm. We also found that more retrievals fail to converge if no continuum correction is applied, meaning that the continuum correction also improved the retrieval robustness. In summary, the continuum correction and usage of the O_2 A band shows a significant improvement on the retrieval precision. We also compared our new approach with the TanSat XCO_2 retrieval from a single CO_2 band only (Figure 10). The single CO_2 weak band retrieval has been carried out with UoL-FP/TanSat retrieving the CO_2 profile, surface albedo and wavelength stretch, assuming no aerosol and cloud scattering in the atmosphere. This is the same strategy used by the IAPCAS (the Institute of Atmospheric Physics Carbon dioxide retrieval Algorithm for Satellite remote sensing) retrieval to generate preliminary TanSat XCO_2 data (Yang et al., 2018; Liu et al., 2018). The improvement is significant both on the bias and RMSE. The bias is reduced by ~1ppm, while the RMSE decreases from 3.43 to 1.59 ppm.

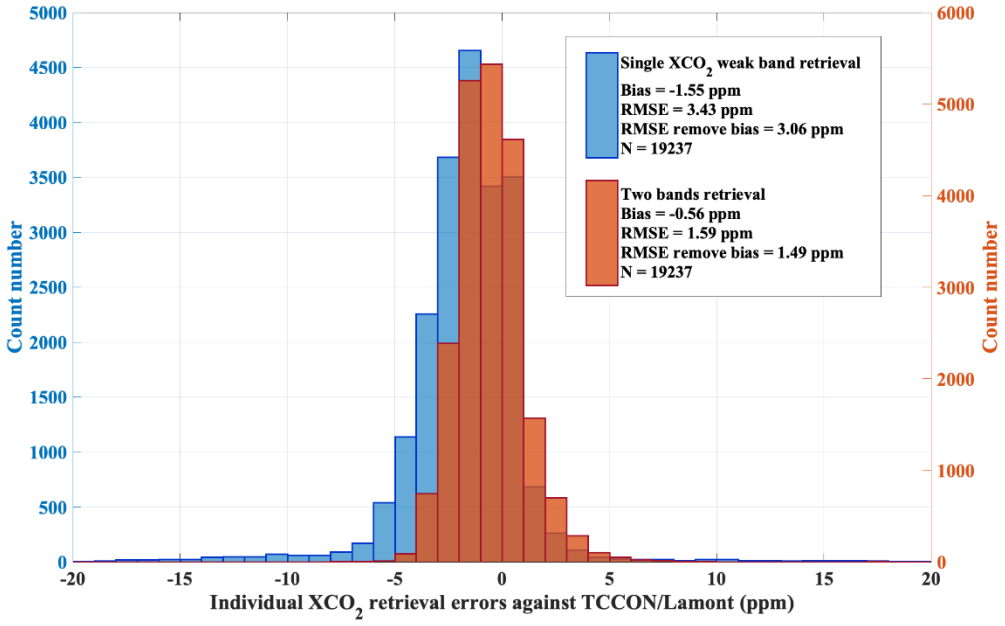

Figure 10. Same as figure 9, but for a comparison between a single CO_2 weak band retrieval (on continuum correction) and the two-band retrieval with Fourier series continuum correction.

4.2 Validation against TCCON measurement

To compare TanSat retrievals to TCCON measurements, we average the quality-controlled, bias-corrected TanSat retrievals per overpass including all across-track footprints for a swath, as no obvious across-track footprint bias has been observed. Only overpasses with more than 50 soundings are used. Figure 11 shows the comparisons of the TanSat XCO_2 average per overpass compared to the TCCON retrievals averaged over ±1 hour of the overpass time. From the 174 data pairs found, we can infer a daily mean bias of 0.08 ppm and a RSME of 1.47 ppm, which are parameters often used to characterize retrieval performance (Cogan et al., 2012; O'Dell et al., 2018; Wu et al., 2018). The linear regression has a slope of 0.83 and R^2 of 0.77. However, these statistics can be misleading, as sites with large numbers of overpasses will have a larger weight than sites with fewer overpasses, and therefore this average is dominated by the few sites with a large number of overpasses. Consequently, the figure also displays the overall

mean of the mean RMSE per site with the individual values for mean bias and RMSE given in (Table 2).

Overall, we find that mean biases are small but show a noticeable variation between sites which is partly due to the low number of data points at some sites. Considering only the seven sites with more than 10 overpasses, we find that JPL (USA), Pasadena (USA) and Saga (Japan) show a large (~1ppm) negative bias. JPL (USA) and Pasadena (USA), for example, have a strong impact from the nearby city of Los Angles (USA), and it is suggested not to include these sites for bias correction (O'Dell et al., 2018). The linear regression with a slope of 0.83 and R^2 of 0.77 can be improved to 0.84 and 0.92 respectively by removing measurements at Pasadena (USA), JPL (USA), Tsukuba (Japan) and Saga (Japan) (see SI section 2).

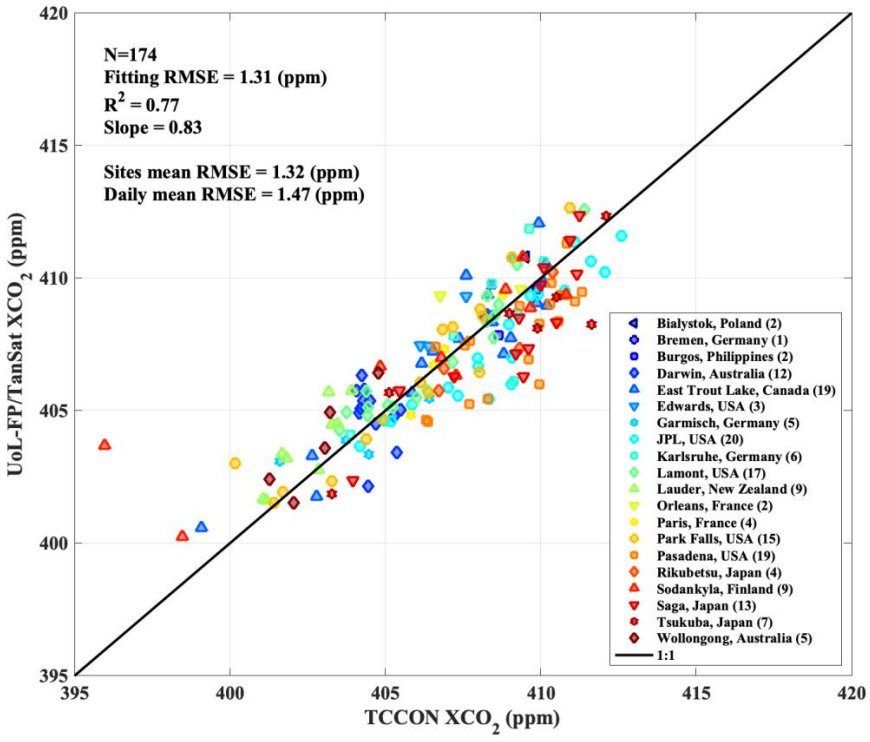

Figure 11. Validation of UoL-FP/TanSat XCO_2 retrievals against measurements from 20 TCCON sites. Each symbol represents the mean of one overpass (see detail in text section 3.1 for colocation criteria) for TanSat (only shown if the available quantity N>50) and the TCCON average during the overpass (only show if the available quantity N>20). The total number of overpasses per site is given in the legend. In total 174 daily data couples are involved in this validation. Statistics are shown in the upper-left corner of the figure. The daily mean RMSE is the total RMSE computed from each overpass mean and the site mean RMSE is computed by averaging the RMSE of each site. The black line indicates the 1:1 line as reference. The slope, R^2 and fitting RMSE are the statistics from a linear regression weighted by bi-square.

For the other four sites, Lamont, Park Falls, Darwin and East Trout Lake, we observe small biases of a few tenths of a ppm only. For Lauder (New Zealand) and Sodankyla (Finland) we also observe large biases, but the number of data points are small. In the case of Sodankyla,

we find a very large outlier for one day (24 July 2017) with a large difference to TCCON (~+8 ppm) and to the a priori value (~+ 5 ppm). Checking cloud information from the RGB Image and MODIS cloud mask did not show cloud contamination within the satellite field of view (FOV). However, if we remove this day, the bias for this site reduces to 0.35 ppm and the RMSE to 1.25 ppm which indicates that the statistics given in Table 3 can be impacted by single outliers. It should be noted that the linear regression will be impacted by the quantity of overpasses used. In this study, we only use 15 months of TanSat nadir measurements, leading to few overpasses for many sites. Including longer time series in the future will help to increase the robustness of the results.

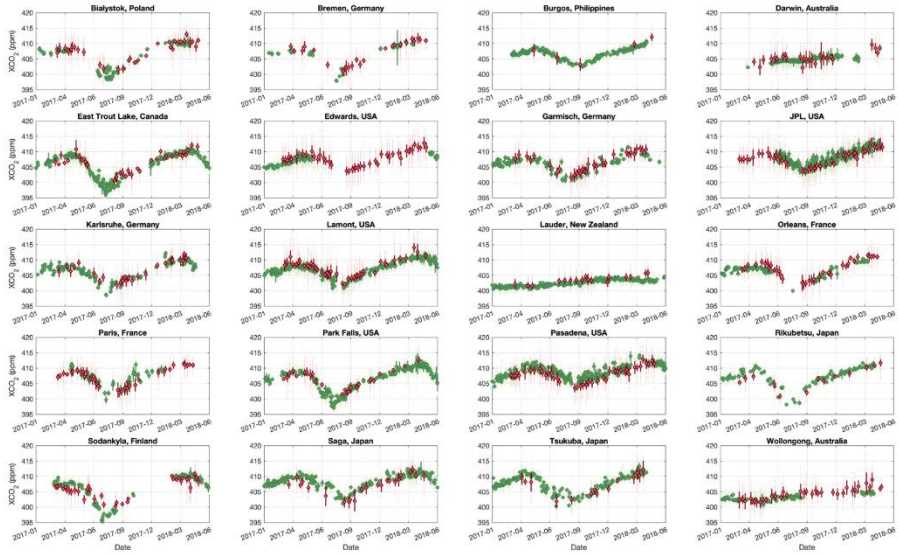

Figure 12. XCO_2 time series from 2017 to 2018 for each TCCON site used in this study (site name is shown in the title of each sub-figure). The green and red large solid circles represent TCCON and TanSat overpass mean with error bar indicating the standard deviation. The individual measurements are shown in light green and light red small points for TCCON and TanSat overpasses, respectively. The TCCON measurements are only shown when the quantity is greater than 20 and TanSat measurements are only shown when the quantity is greater than 50. This figure shows all available TCCON and TanSat data, not only TanSat-TCCON couples, hence the data quantity shown in this figure is larger than is used for the validation.

4.3 Temporal trend and variations

XCO_2 has typically a detrended, seasonal amplitude of ~5-8 ppm from peak to trough (Lindqvist et al., 2015) in the Northern Hemisphere and roughly an annual growth of ~2 ppm globally, which is also seen from TCCON measurement from 2017 to 2018 (Figure 12). This behavior is well captured in the TanSat retrievals for both the Northern and Southern Hemisphere. For example, for Edwards (USA) the satellite measurements show a very clear seasonal variation with a peak to peak detrended variation of ~6 ppm. A +1.36 ppm bias and RMSE of 1.39 ppm is found for this site but this is derived from 3 co-located data-pairs only. The negative biases found for JPL (USA) and Pasadena (USA) are clearly visible in the time series. Higher values from TanSat compared to TCCON are observed toward the end of the time

series for Southern Hemisphere sites Darwin (Australia), Lauder (New Zealand) and Wollongong (Australia). A potential seasonal dependence of biases will not necessarily be corrected by the applied bias correction as it does not include parameters like solar zenith angle (or airmass) or time. To further investigate potential seasonal biases, Figure 13 shows a time series of the observed biases for sites with more than nine co-located overpasses. Although, we find some variations of biases; for example, Saga (Japan) indicates a negative bias in the spring of 2017 and a positive bias in the spring of 2018, we do not observe a general correlation of biases with season across all sites.

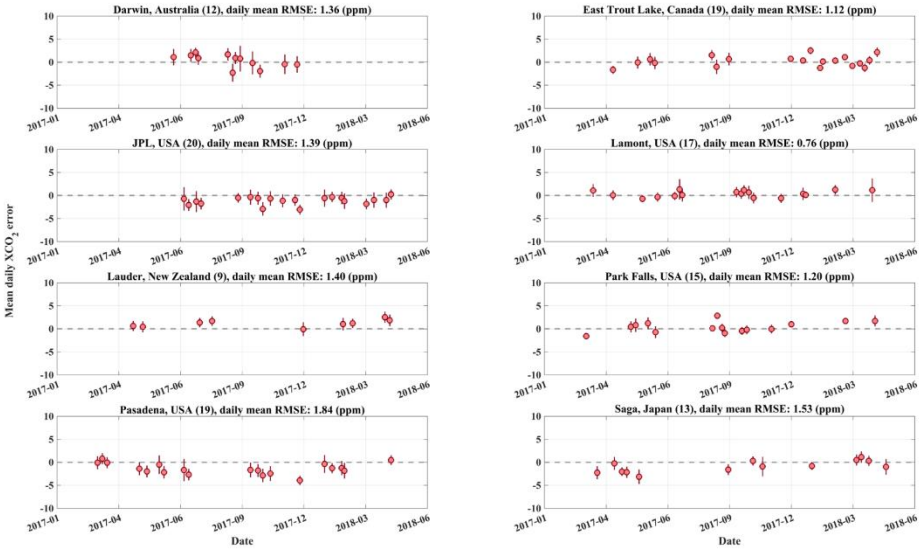

Figure 13. Seasonal variation of the XCO_2 overpass mean error (difference between TanSat and TCCON XCO_2). The data couples shown in this figure are the same as used in Figure 11. Only sites with more than 9 overpasses data points (the number of data points for each site is shown in parentheses of each sub-figure caption) have been shown in this figure. The error bars indicate the standard deviation. The dashed grey line shows the zero-error line as reference. The site name and RMSE are given in the title for each sub-figure.

5 Summary and outlook

In this study, the UoL-FP algorithm has been implemented to retrieve XCO_2 from TanSat nadir mode observations. A major obstacle towards high-quality retrieval of XCO_2 from TanSat has been spectral artifacts in the fitting residual of the O_2 A band. By analyzing the solar calibration measurement, we found that this pattern is stable with time and that it can be effectively eliminated by applying an 8-orders Fourier series model. This correction significantly improves the fitting residual, and accordingly reduces the XCO_2 retrieval RMSE against measurements from the TCCON site at Lamont (USA) from 4.08 to 1.59 ppm.

A data-driven quality control and bias correction strategy has been applied to improve the data quality of the XCO_2 retrievals. Based on the correlation of TanSat-TCCON XCO_2 individual error against different retrieval and forward model parameters, including Grad CO_2, Delta Psurf, Continuum B1C3, Zeroff B2S and AlbedoB2 (see description in Table 3), a Genetic Algorithm (GA) has been used to select quality filters for 64.3% of transparency and ~2 ppm of

RMSE. This leads to an overall retrieval throughput of ~18.3% (good retrievals). Compared to the manual selection of filters this achieves 13.5% more soundings for comparable RMSE. We apply a multiple linear regression for the same parameters as selected by the GA to derive a footprint dependent bias correction. After applying the bias correction, no obvious difference in remaining bias between cross-track footprints is found.

The validation involves 20 TCCON sites with co-location criteria within ±3º latitude/longitude and ±1 hour in time. The mean RMSE is found to be 1.47 ppm and the mean of the RMSE per site is 1.32 ppm. Typically, we find biases of a few tenths of a ppm for individual TCCON sites but larger biases (~1 ppm) are observed for some sites, especially in the proximity of major cities such as for JPL (USA), Pasadena (USA) and Tsukuba (Japan).

Previous TanSat retrievals have adopted an approach using a single (weak) CO_2 band retrieval only. To improve this approach, the development of a 2-band retrieval in this paper combining the O_2 A band with the weak CO_2 band represents a step forward in reliable TanSat retrieval as demonstrated by the improvement of the RMSE against TCCON from 3.43 to 1.59 ppm, which was found in a retrieval study of the Lamont TCCON site.

The methods used in this study, such as continuum correction, can help to improve the XCO_2 retrieval from TanSat and subsequently the L2 data product, and hence will be applied in the IAPCAS XCO_2 retrievals which are used for the operational processing of TanSat data. There are differences in models and retrieval strategy between the UoL-FP/TanSat and IAPCAS/TanSat retrieval. In future studies we will investigate the impact of the differences of the two algorithms and their advantages and disadvantages. In this study we introduced an improved TanSat retrieval over land based on a two-band retrieval. The inclusion of the strong CO_2 band still needs to be investigated in the future and can be expected to further improve the retrieval performance by providing more information on water vapor and temperature, as well as the wavelength dependency of aerosols and clouds in the long-wave end of the SWIR. Further studies on target mode observations would be helpful to improve the retrievals. Furthermore, the retrieval of glint mode observations over both ocean and land surfaces will need to be studied, allowing us to extend the coverage of the TanSat XCO_2 dataset.

Appendix A: the polarization angle of TanSat measurement

To consider polarization effects in RT and the retrieval, the polarization angle needs to be known as accurately as possible. Errors in calculated radiances that could arise from a miscalculation of the polarization angle can be as large as 20% in deep absorption lines in the O_2 A band (Figure A1), leading to large errors in the XCO_2 retrieval. The results shown in Figure A1 have been calculated for an aerosol-free scene with a solar zenith angle of 30º and a Lambertian albedo of 0.2 assuming the 1976 U.S standard atmosphere. However, the impact will become even worse for larger solar zenith angles; for example, Beijing, China (with latitude 40º) the SZA increases to a value larger than 60º in winter for the TanSat overpass time. The polarization impact in the CO_2 weak and strong bands is lower in aerosol-free scenes as the impact of Rayleigh scattering is greatly reduced, but an accurate calculation is still required in the presence of polarizing surfaces or large polarizing particles (cirrus).

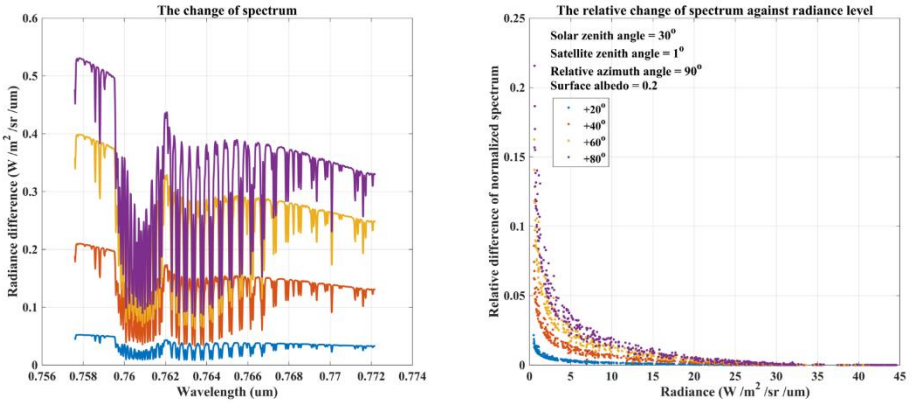

Figure A1. A simulation on the impact of the polarization angle on a simulated O_2 A Band spectrum. The left sub-figure shows the changes of the spectrum when the polarization angle is 20º, 40º, 60º and 80º larger than the true angle. The color index is marked in the text box of the right sub-figure. The right sub-figure shows the relative change to the normalized spectrum against radiance level. Notice that the change becomes larger when the radiance decreases. This effect introduces significant errors in deep absorption lines where important information for the retrieval comes from. The simulated scene has a surface albedo of 0.2 and solar zenith angle of 30º in nadir measurement mode. The calculated polarization angle is 86.3º as calculated by the method described in section 2.1.2.

Commonly, RT simulations are carried out in the local meridian plane (a plane defined by the unit vector of the local normal of a footprint and the vector from the footprint to the satellite). Unfortunately, the polarizer direction is the reference to the principal plane due to TanSat in-flight tracking strategy (tracking the principal plane). The change of the Stokes vectors due to the rotation of the reference plane can be described by an angle defined between the principal plane and the local meridian plane (Liou, 2002; Mishchenko et al., 2004), namely the polarization angle (σ),

$$\cos\sigma = \frac{\cos\theta_{sun} - \cos\theta_{obs}\cdot\cos\Theta}{\sin\theta_{obs}\cdot\sin\Theta}, \quad \text{(A-1)}$$

in which θ_{sun} and θ_{obs} are the footprint solar zenith and observation zenith angles, while Θ is the scattering angle in the scattering plane, defined by the satellite line of sight and the solar direction:

$$\cos\theta_{obs}\cos\theta_{sun} + \sin\theta_{obs}\sin\theta_{sun}\cos(\Delta\varphi), \quad \text{(A-2)}$$

with the relative azimuth angle $\Delta\varphi$ defined by the footprint satellite azimuth and the solar azimuth angle.

In practice, due to the instrument optical design and integration strategy, the polarizer direction is parallel to (or in) the principal plane. This could lead to the light that passed the polarizer to be approximately zero (in ideal conditions) meaning that the measurement will become almost negligible when the measurement geometry meets the Brewster angle condition (solar zenith angle approaching to Brewster angle). For the purpose of avoiding this effect and increasing the received signal, the TanSat spacecraft is rotated by a +45º yaw angle during flight. Therefore, the actual polarization angle is $\sigma + 45°$. Then the Stokes parameter coefficients are given by,

$$a_1 = {}^1/_2, \text{ (A-3)}$$

$$a_2 = {}^1/_2 \cos 2(\sigma + 45°), \text{ (A-4)}$$

$$a_3 = {}^1/_2 \sin 2(\sigma + 45°), \text{ (A-5)}$$

In summary, the Stokes vectors $\{\boldsymbol{I}, \boldsymbol{Q}, \boldsymbol{U}\}$ at the top of atmosphere are firstly computed by the forward model and then combined to a simulation of the signal received by the instrument according to Eq. (2) in main text.

Appendix B: The Genetic algorithm

Genetic algorithms (Holland, 1975; Goldberg, 1989) belong to the class of derivative-free (Jacobi-free) optimization methods (Rios and Sahinidis, 2013), which also includes other popular methods such as simulated annealing (Kirkpatrick, 1983) and particle swarm optimization (Kennedy and Eberhart, 1995). As opposed to traditional gradient-based methods, where a cost function is minimized by updating an initial guess along a descent direction identified by computing the derivatives of the cost function with respect to its parameters (Nocedal and Wright, 1999), in derivative-free methods the optimization is carried out through successive evaluations of the cost function itself, until a definite convergence criterion is met. An advantage of derivative-free optimization methods compared to gradient-based methods is that they are more capable of escaping local minima of the cost function to be optimized. Their disadvantage is that they may require a higher number of iterations before reaching the minimum. For this reason, they are particularly advantageous when the cost function is relatively inexpensive to evaluate.

Optimization in genetic algorithms is based on concepts drawn from the process of evolution taking place in biology. If by $\mathbf{f}(\boldsymbol{x})$ we denote a cost function to be optimized with respect to a vector x, a genetic algorithm first generates a dataset (population) $\{x_i\}$ of candidate solutions, and then randomly updates the dataset by applying two operations to its elements:

(1) Mutation: one or more component of an element x are randomly perturbed

(2) Cross-over: a new element of the population is generated by combining two existing elements

All the elements in the population are then ranked with respect to a fitness function which measures the goodness of a certain element as a solution of the optimization problem, and a selection is performed by retaining a user-defined percentage of elements which obtain the best ranking. Typically, the fitness function is the cost function itself. As a result of this selection process, a new population is created, and the random operations of mutation and cross-over – followed by a selection – are applied to the new population. The process described above is repeated until a convergence criterion is met.

In practice, each filter there are two threshold values that need to be decided: upper boundary and lower boundary (gene). Therefore, we define two segments (upper boundaries and lower boundaries) for each individual respectively, and the combined segments as a whole, containing filters with two boundaries, and we run GA for each complexity with determined filters. The evolution is completed by multi reproductions that typically include cross-over and mutation. Cross-over exchanges the gene segment among individuals that win the fitness selection. The purpose of cross-over is to combine the advantaged genes and to improve the

fitness of a new generation of individuals. However, we always hope to change advantaged genes by at least a small distance towards the optimized solution. Therefore, only mutations are applied in reproduction, which has been found to be sufficient (Mandrake et al., 2013). Threshold values for each gene are initiated randomly with a quantity of 100,000 individuals ranging between the maximum and minimum values in the training dataset. For each reproduction, only 15 individuals pass the fitness selection and replace the population of the last reproduction. Therefore, the mutation is only coming from these 15 individuals that contain advantaged genes. We let the GA mutate the gene 100 times for each individual by selecting gene and according number randomly, and then 150,000 individuals have reproduced after each mutation. A convergence criterion has been chosen to stop the GA run when the RSME of the retrieval error reduces to less than 0.2% over the previous ten iterations. This criterion is looser than the one that has been introduced by Mandrake (2013), but in practice, we found this works sufficiently well.

Acknowledgments, Samples, and Data

This work is supported by the National Key R&D Program of China (No. 2016YFA0600203), The Key Research Program of the Chinese Academy of Sciences (ZDRW-ZS-2019-1), International Partnership Program of Chinese Academy of Sciences (Grant No. GJHZ201903), and the National Natural Science Foundation of China (No. 41905029). ESA Climate Change Initiative CCI+ (GhG theme), Earthnet Data Assessment Pilot (EDAP) project and ESA-MOST Dragon-4 programme (ID 32301). HB, AW and RJP are supported by the UK NERC National Centre for Earth Observation (NCEO grant number: nceo020005).

The TanSat L1B data service has been provided by the International Reanalysis Cooperation on Carbon Satellites Data (IRCSD) and the Cooperation on the Analysis of carbon SAtellites data (CASA). We also thank FENGYU Satellite Data Center of National Satellite Meteorological Center who providing TanSat L1B data service. This research used the ALICE High Performance Computing Facility at the University of Leicester.

The authors thank TanSat mission, and the supporting from everyone that working with TanSat mission is highly appreciate.

The Paris TCCON site has received fundings from Sorbonne Université, the French research center CNRS, the French space agency CNES and Région Île-de-France. The TCCON stations at Tsukuba, Rikubetsu, and Burgos are supported in part by the GOSAT series project. Local support for Burgos is provided by the Energy Development Corporation (EDC, Philippines). Darwin and Wollongong TCCON operations are supported in part via Australian Research Council Grants (DP0879468, DP110103118, DP140101552, DP160101598, LE0668470 and DE140100178). Local support for Darwin is provided by the Australian Bureau of Meteorology. We also acknowledge NASA support for Darwin. NMD is supported by an Australian Research Council Future Fellowship, FT180100327. VAV is supported by a UOW/SMAH PEPA grant. We thank all TCCON PIs for providing excellent measurement and validation dataset for this study.

The TanSat/UoL-FP XCO_2 retrieval dataset that produced, introduced and investigated in this study is one of dataset of ESA Climate Change Initiative CCI+ (GhG theme) and will be available to the public on the ESA CCI data service website: http://cci.esa.int/data

References

Bai, W., Zhang, P., Zhang, W., Ma, G., Qi, C., & Liu. H. (2018), CO_2 column-retrieval errors arising from neglecting polarization in forward modelling of 1.6 μm band measurements. *Science Bulletin,* 63(8), 766–776. doi:10.1360/N972017-00853

Baum, B. A., Yang, P., Heymsfield, A. J., Bansemer, A., Cole, B. H., Merrelli, A., Schmitt, C., & Wang. C. (2014),Ice cloud single-scattering property models with the full phase matrix at wavelengths from 0.2 to 100 um. *Journal of Quantitative Spectroscopy & Radiative Transfer*. 146, 123–139. doi:10.1016/j.jqsrt.2014.02.029

Benner, D. C., Devi, V. M., Sung, K., Brown, L. R., Miller, C. E., Payne, V. H., Drouin, B. J., Yu, S., Crawford, T. J., Mantz, A. W., Smith, M. A. H., & Gamache, R. R. (2016),Line parameters including temperature dependences of air- and self-broadened line shapes of 12C16O2: 2.06 micron region. *J Mol Spectrosc*, 326, 21–47. doi:10.1016/j.jms.2016.02.012

Boesch, H., Deutscher, N. M., Warneke, T., Byckling, K., Cogan, A. J., Griffith, D. W. T., Notholt, J., Parker, R. J., & Wang, Z. (2013),HDO/H2O ratio retrievals from GOSAT. *Atmos. Meas. Tech.*, 6, 599–612. doi:10.5194/amt-6-599-2013

Bovensmann, H., Burrows, J. P., Buchwitz, M., Frericka, J., Noëla, S., Rozanova, V. V., Chanceb, K. V., & Goedec, A. P. H. (1999),SCIAMACHY: mission objectives and measurement modes. *J Atmos Sci*, 56, 127–50. doi:10.1175/1520-0469(1999)056<0127:SMOAMM>2.0.CO;2

Bösch, H., Toon, G. C., Sen, B., Washenfelder, R. A., Wennberg, P. O. Buchwit, M., de Beek, R., Burrows, J. P., Crisp, D., Christi, M., Connor, B. J., Natraj, V., & Yung, Y. L. (2006),Space-based near-infrared CO2 measurements: Testing the Orbiting Carbon Observatory retrieval algorithm and validation concept using SCIAMACHY observations over Park Falls, Wisconsin. *J. Geophys. Res.*, 111, D23302, doi:10.1029/2006JD007080

Buchwitz, M., de Beek, R., Burrows, J. P., Bovensmann, H., Warneke, T., Notholt, J., Meirink, J. F., Goede, A. P. H., Bergamaschi, P., Körner, S., Heimann, M., & Schulz, A. (2005),Atmospheric methane and carbon dioxide from SCIAMACHY satellite data: initial comparison with chemistry and transport models. *Atmos. Chem. Phys.*, 5, 941–962. doi:10.5194/acp-5- 941-2005.

Buchwitz, M., Dils, B., Boesch, H., Brunner, D., Butz, A., Crevoisier, C., Detmers, R., Frankenberg, C., Hasekamp, O., Hewson, W., Laeng, A., Noël, S., Notholt, J., Parker, R., Reuter, M., Schneising, O., Somkuti, P., Sundström, A., & De Wachter, E. (2017a) ESA Climate Change Initiative (CCI) Product Validation and Intercomparison Report (PVIR) for the Essential Climate Variable (ECV) Greenhouse Gases (GHG) for data set Climate Research Data Package No. 4 (CRDP# 4),Technical Note, 4, 253. http://www.esa-ghg-cci.org/?q=node/95

Buchwitz, M., Reuter, M., Schneising, O., Hewson, W., Detmers, R., Boesch, H., Hasekamp, O., Aben, I., Bovensmann, H., Burrows, J., Butz, A., Chevallier, F., Dils, B., Frankenberg, C., Heymann, J., Lichtenberg, G., Maziè re, M. D., Notholt, J., Parker, R., Warneke, T., Zehner, C., Griffith, D. W. T., Deutscher, N. M., Kuze, A., Suto, H., & Wunch, D. (2017b),Global satellite observations of column averaged carbon dioxide and methane:

The GHG-CCI XCO_2 and XCH_4 CRDP3 data set. *Remote Sens. Environ.*, 203, 276–295. doi:10.1016/j.rse.2016.12.027.

Butz, A., Hasekamp, O. P., Frankenberg, C., & Aben, I. (2009),Retrievals of atmospheric CO_2 from simulated space-borne measurements of backscattered near-infrared sunlight: accounting for aerosol effects. *Appl. Opt.*, 48(18), 3322–3336. doi:10.1364/AO.48.003322

Butz, A., Guerlet, S., Hasekamp, O., Schepers, D., Galli, A., Aben, I., Frankenberg, C., Hartmann, J.-M., Tran, H., Kuze, A., Keppel-Aleks, G., Toon, G., Wunch, D., Wennberg, P., Deutscher, N., Griffith, D., Macatangay, R., Messerschmidt, J., Notholt, J., & Warneke, T. (2011),Toward accurate CO_2 and CH_4 observations from GOSAT. *Geophys. Res. Lett.*, 38, L14812. doi:10.1029/2011GL047888

Chen, C., Park, T., Wang, X., Piao, S., & Xu, B., Chaturvedi, R. K., Fuchs, R., Brovkin, V., Ciais, P., Fensholt, R., Tømmervik, H., Bala, G., Zhu, Z., Nemani, R. R., Myneni, R. B. (2019),China and India lead in greening of the world through land-use management. *Nature sustainability*, 2, 122–129. doi:10.1038/s41893-019-0220-7

Chen, X., Yang, D., Cai, Z., Liu, Y., & Spurr, R. J. D. (2017a),Aerosol Retrieval Sensitivity and Error Analysis for the Cloud and Aerosol Polarimetric Imager on Board TanSat: The Effect of Multi-Angle Measurement. *Remote Sens.*, 9(2), 183. doi:10.3390/rs9020183

Chen, X., Wang, J., Liu, Y., Xu, X., Cai, Z., Yang, D., Yan, C., & Feng., L. (2017b),Angular dependence of aerosol information content in CAPI/TanSat observation over land: Effect of polarization and synergy with A-train satellites. *Remote. Sens. Environ.*, 196(2017), 163–177, 2017b. doi:10.1016/j.rse.2017.05.007

Chen W., Zhang, Y., Yin, Z., Zheng, Y., Yan, C., Yang, Z., & Liu, Y. (2012),The TanSat Mission: Global CO_2 Observation and Monitoring. Proceedings of the 63rd IAC (International Astronautical Congress), Naples, Italy

Crisp, D., Miller, C. E., & DeCola, P. L. (2008),NASA Orbiting Carbon Observatory: measuring the column averaged carbon dioxide mole fraction from space. *Journal of Applied Remote Sensing*, 2(1), 023508. doi:10.1117/1.2898457

Crisp, D., Fisher, B. M., O'Dell, C., Frankenberg, C., Basilio, R., Bösch, H., Brown, L. R., Castano, R., Connor, B., Deutscher, N. M., Eldering, A., Griffith, D., Gunson, M., Kuze, A., Man- drake, L., McDuffie, J., Messerschmidt, J., Miller, C. E., Morino, I., Natraj, V., Notholt, J., O'Brien, D. M., Oyafuso, F., Polonsky, I., Robinson, J., Salawitch, R., Sherlock, V., Smyth, M., Suto, H., Taylor, T. E., Thompson, D. R., Wennberg, P. O., Wunch, D., & Yung, Y. L. (2012),The ACOS CO_2 retrieval algorithm – Part II: Global XCO_2 data characterization. *Atmos. Meas. Tech.*, 5, 687–707. doi:10.5194/amt-5-687-2012

Crisp, D., Pollock, H. R., Rosenberg, R., Chapsky, L., Lee, R. A. M., Oyafuso, F. A., Frankenberg, C., O'Dell, C., Bruegge, C. J., Doran, G. B., Eldering, A., Fisher, B. M., Fu, D., Gunson, M. R., Mandrake, L., Osterman, G. B., Schwandner, F. M., Sun, K., Taylor, T. E., Wennberg, P. O., Wunch., D. (2017),The on-orbit performance of the Orbiting Carbon Observatory-2 (OCO-2) instrument and its radiometrically calibrated products. *Atmos. Meas. Tech.*, 10, 59–81. doi:10.5194/amt-10-59-2017

Cogan, A. J., Boesch, H, Parker, R. J., Feng, L., Palmer, P. I., Blavier, J.-F. L., Deutscher, N. M., Macatangay, R., Notholt, J., Roehl, C., Warneke, T., & Wunch, D. (2012),Atmospheric carbon dioxide retrieved from the Greenhouse gases Observing SATellite (GOSAT): Comparison with ground-based TCCON observations and GEOS-Chem model calculations. *J. Geophys. Res.*,117, D21301, doi:10.1029/2012JD018087

Connor, B., Bösch, H., McDuffie, J., Taylor, T., Fu, D., Frankenberg, C., O'Dell, C., Payne, V. H., Gunson, M., Pollock, R., Hobbs, J., Oyafuso, F., and Jiang, Y. (2016),Quantification of uncertainties in OCO-2 measurements of XCO_2: simulations and linear error analysis. *Atmos. Meas. Tech.*, 9(10), 5227–5238. doi:10.5194/amt-9-5227-2016.

Corradini, S., & Cervino, M. (2006),Aerosol extinction coefficient profile retrieval in the oxygen A-band considering multiple scattering atmosphere. Test case: SCIAMACHY nadir simulated measurements. *J. Quant. Spectrosc. Ra.*, 97(3), 354–380. doi: 10.1016/j.jqsrt.2005.05.061

Dils, B., Buchwitz, M., Reuter, M., Schneising, O., Boesch, H., Parker, R., Guerlet, S., Aben, I., Blumenstock, T., Burrows, J. P., Butz, A., Deutscher, N. M., Frankenberg, C., Hase, F., Hasekamp, O. P., Heymann, J., De Mazière, M., Notholt, J., Sussmann, R., Warneke, T., Griffith, D., Sherlock, V., and Wunch, D. (2014),The Greenhouse Gas Climate Change Initiative (GHG-CCI): comparative validation of GHG-CCI SCIAMACHY/ENVISAT and TANSO-FTS/GOSAT CO_2 and CH_4 retrieval algorithm products with measurements from the TCCON. *Atmos. Meas. Tech.*, 7, 1723–1744. doi:10.5194/amt-7-1723-2014

Deutscher, N. M., Notholt, J., Messerschmidt, J., Weinzierl, C., Warneke, T., Petri, C., Grupe, P., and Katrynski, K. (2019),TCCON data from Bialystok (PL), Release GGG2014R2. TCCON data archive, hosted by CaltechDATA. doi:10.14291/tccon.ggg2014.bialystok01.R2, 2019

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., & Bechtold, P. (2011),The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. Quarterly Journal of the Royal Meteorological Society, 137(656), 553–597. doi:10.1002/qj.828

Devi, V. M., Benner, D. C., Sung, K., Brown, L. R., Crawford, T. J., Miller, C. E., Drouin, B. J., Payne, V. H., Yu, S., Smith, M. A. H., Mantz, A. W., & Gamache, R. R. (2016),Line parameters including temperature dependences of self- and foreign-broadened line shapes of 12C16O2: 1.6 micron region. *J. Quant. Spectrosc. Ra.*, 177, 117-144. doi:10.1016/j.jqsrt.2015.12.020

Drouin B. J., Benner, D. C., Brown, L. R., Cich, M., Crawford, T., Devi, V. M., Guillaume, A., Hodges, J. T., Mlawer, E. J., Robichaud, D., Oyafuso, F., Payne, V. H., Sung, K., Wishnow, E., & Yu, S. (2017),Multispectrum analysis of the oxygen A-band. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 186, 118–138. doi:10.1016/j.jqsrt.2016.03.037

Eldering, A., O'Dell, C. W., Wennberg, P. O., Crisp, D., Gunson, M. R., Viatte, C., Avis, C., Braverman, A., Castano, R., Chang, A., Chapsky, L., Cheng, C., Connor, B., Dang, L., Doran, G., Fisher, B., Frankenberg, C., Fu, D., Granat, R., Hobbs, J., Lee, R. A. M., Mandrake, L., McDuffie, J., Miller, C. E., Myers, V., Natraj, V., O'Brien, D., Osterman, G. B., Oyafuso, F., Payne, V. H., Pol- lock, H. R., Polonsky, I., Roehl, C. M., Rosenberg,

R., Schwand- ner, F., Smyth, M., Tang, V., Taylor, T. E., To, C., Wunch, D., and Yoshimizu, J. (2017a),The Orbiting Carbon Observatory-2: first 18 months of science data products. *Atmos. Meas. Tech.*, 10, 549–563. doi:10.5194/amt-10-549-2017.

Eldering, A., Wennberg, P. O., Crisp, D., Schimel, D. S., Gun- son, M. R., Chatterjee, A., Liu, J., Schwandner, F. M., Sun, Y., O'Dell, C. W., Frankenberg, C., Taylor, T., Fisher, B., Oster- man, G. B., Wunch, D., Hakkarainen, J., Tamminen, J., & Weir, B. (2017b),The Orbiting Carbon Observatory-2 early science investigations of regional carbon dioxide fluxes. *Science*, 358, eaam5745. doi:10.1126/science.aam5745

Eldering, A., Taylor, T. E., O'Dell, C. W., & Pavlick, R. (2019) The OCO-3 mission: measurement objectives and expected performance based on 1 year of simulated data. *Atmos. Meas. Tech.*, 12, 2341–2370. doi:10.5194/amt-12-2341-2019

Feng, L., Palmer, P. I., Bösch, H., Parker, R. J., Webb, A. J., Correia, C. S. C., Deutscher, N. M., Domingues, L. G., Feist, D. G., Gatti, L. V., Gloor, E., Hase, F., Kivi, R., Liu, Y., Miller, J. B., Morino, I., Sussmann, R., Strong, K., Uchino, O., Wang, J., & Zahn, A. (2017),Consistent regional fluxes of CH_4 and CO_2 inferred from GOSAT proxy $XCH_4 : XCO_2$ retrievals, 2010–2014. *Atmos. Chem. Phys.*, 17, 4781-4797. doi:10.5194/acp-17-4781-2017.

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., & Wunch, D. (2015),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

Geddes, A., & Bösch, H. (2015),Tropospheric aerosol profile information from high-resolution oxygen A-band measurements from space. *Atmos. Meas. Tech.*, 8, 859–874. doi:10.5194/amt-8-859-2015

Goldberg, D. E. (1989),Genetic Algorithms in Search, Optimization, and Machine Learning. Addison-Wesley

Griffith, D. W., Velazco, V. A., Deutscher, N. M., Murphy, C., Jones, N. B., Wilson, S. R. Macatangay, R. C., Kettlewell, G. C., Buchholz, R. R., and Riggenbach, M. O. (2014a),TCCON data from Wollongong, (AU), Release GGG2014R0. TCCON data archive, hosted by Caltech-DATA. doi:10.14291/tccon.ggg2014.wollongong01.R0/1149291

Griffith, D.W. T., Deutscher, N. M., Velazco, V. A.,Wennberg, P. O., Yavin, Y., Keppel-Aleks, G.,Washenfelder, R. A., Toon, G. C., Blavier, J.-F., Murphy, C., Jones, N. B., Kettlewell, G. C., Connor, B. J., Macatangay, R. C., Roehl, C., Ryczek, M., Glowacki, J., Culgan, T., and Bryant, G. W. (2014b),TCCON data from Darwin (AU), Release GGG2014R0. TCCON data archive, hosted by CaltechDATA. doi:10.14291/tccon.ggg2014.darwin01.R0/1149290

Guerlet, S., Butz, A., Schepers, D., Basu, S., Hasekamp, O. P., Kuze, A., Yokota, T., Blavier, J.-F., Deutscher, N. M., Griffith, D. W., Hase, F., Kyro, E., Morino, I., Sherlock, V., Sussmann, R., Galli, A., and Aben, I. (2013),Impact of aerosol and thin cirrus on retrieving and validating XCO_2 from GOSAT shortwave in frared measurements. J. Geophys. Res.-Atmos., 118, 4887–4905. doi:10.1002/jgrd.50332

Hakkarainen, J., Ialongo, I., & Tamminen, J. (2016),Direct space-based observations of anthropogenic CO_2 emission areas from OCO-2. *Geophys. Res. Lett.*, 43(21), 11400–11406. doi:10.1002/2016GL070885s

Hase, F., Blumenstock, T., Dohe, S., Groß, J., & Kiel, M. (2015),TCCON data from Karlsruhe (DE), Release GGG2014R1. TCCON data archive, hosted by CaltechDATA. doi:10.14291/tccon.ggg2014.karlsruhe01.R1/1182416

Heidinger, A. K., & Stephens, G. L. (2000),Molecular Line Absorption in a Scattering Atmosphere. Part II: Application to Remote Sensing in the O2 A band. J. Atmos. Sci., 57(10), 1615–1634. doi:10.1175/1520-0469(2000)057<1615:MLAIAS>2.0.CO;2

Heymann, J., Reuter, M., Hilker, M., Buchwitz, M., Schneising, O., Bovensmann, H., Burrows, J. P., Kuze, A., Suto, H., Deutscher, N. M., Dubey, M. K., Griffith, D. W. T., Hase, F., Kawakami, S., Kivi, R., Morino, I., Petri, C., Roehl, C., Schneider, M., Sherlock, V., Sussmann, R., Velazco, V. A., Warneke, T., & Wunch, D. (2015),Consistent satellite XCO2 retrievals from SCIAMACHY and GOSAT using the BESD algorithm. *Atmos. Meas. Tech.*, 8(7), 2961-2980, doi:10.5194/amt-8-2961-2015

Holland, J. (1975),Adaptation in Natural and Artificial Systems. Michigan University Press

Iraci, L., Podolske, J., Hillyard, P., Roehl, C., Wennberg, P. O., Blavier, J.-F., Landeros, J., Allen, N., Wunch, D., Zavaleta, J., Quigley, E., Osterman, G., Albertson, R., Dunwoody, K., & Boyden, H. (2016),TCCON data from Edwards (US), Release GGG2014.R1. TCCON data archive, hosted by CaltechDATA. doi:10.14291/tccon.ggg2014.edwards01.R1/1255068

Kennedy, J., & Eberhart, R. (1995),*Particle swarm optimization*. Paper presented at Proceedings of ICNN'95 - International Conference on Neural Networks, Perth, Australia. doi:10.1109/ICNN.1995.488968.

Kiel, M., O'Dell, C. W., Fisher, B., Eldering, A., Nassar, R., MacDonald, C. G., & Wennberg, P. O. (2019),How bias correction goes wrong: measurement of XCO_2 affected by erroneous surface pressure estimates. *Atmos. Meas. Tech.*, 12, 2241–2259. doi:10.5194/amt-12-2241-2019

Kim, W., Kim, J., Jung, Y., Boesch, H., Lee, H., Lee, S., Goo, T.-Y., Jeong, U., Kim, M., Cho, C.-H., & Ou, M.-L. (2016),Retrieving XCO_2 from GOSAT FTS over East Asia Using Simultaneous Aerosol Information from CAI. *Remote Sens*, 8(12), 994. doi:10.3390/rs8120994

Kirkpatrick, S., Gelatt Jr., C. D., & Vecchi, M. P. (1983),Optimization by simulated annealing. Science, 220, 671-680. doi:10.1126/science.220.4598.671

Kivi, R., Heikkinen, P., & Kyrö, E. (2014),TCCON data from Sodankylä (FI), Release GGG2014R0. TCCON data archive, hosted by CaltechDATA. doi:10.14291/tccon.ggg2014.sodankyla01.R0/1149280

Kuang, Z., Margolis, J., Toon, G., Crisp, D., & Yung, Y. (2002),Spaceborne measurements of atmospheric CO_2 by high-resolution NIR spectrometry of reflected sunlight: An introductory study. *Geo. Res. Lett.*, 29(15),doi:10.1029/2001GL014298

Kuze, A., Suto, H., Nakajima, M., & Hamazaki, T. (2009),Thermal and near infrared sensor for carbon observation Fourier-transform spectrometer on the Greenhouse Gases Observing Satellite for greenhouse gases monitoring. *Appl. Opt.*, 48(35), 6716-6733. doi:10.1364/AO.48.006716

Kuze, A., Taylor, T. E., Kataoka, F., Bruegge, C. J., Crisp, D., Harada, M., Helmlinger, M., Inoue, M., Kawakami, S., Kikuchi, N., Mitomi, Y., Murooka, J., Naitoh, M., O'Brien, D. M., O'Dell, C. W., Ohyama, H., Pollock, H., Schwandner, F. M., Shiomi, K., Suto, H., Takeda, T., Tanaka, T., Urabe, T., Yokota, T., & Yoshida, Y. (2014),Long-term Vicarious Calibration of GOSAT ShortWave Sensors: Techniques for Error Reduction and New Estimates of Radiometric Degradation Factors. *IEEE T. Geosci. Remote*, 52, 3991–4004. doi:10.1109/TGRS.2013.2278696

Le Quéré, C., Andrew, R. M., Friedlingstein, P., Sitch, S., Hauck, J., Pongratz, J., Pickers, P. A., Korsbakken, J. I., Peters, G. P., Canadell, J. G., Arneth, A., Arora, V. K., Barbero, L., Bastos, A., Bopp, L., Chevallier, F., Chini, L. P., Ciais, P., Doney, S. C., Gkritzalis, T., Goll, D. S., Harris, I., Haverd, V., Hoffman, F. M., Hoppema, M., Houghton, R. A., Hurtt, G., Ilyina, T., Jain, A. K., Johannessen, T., Jones, C. D., Kato, E., Keeling, R. F., Goldewijk, K. K., Landschützer, P., Lefèvre, N., Lienert, S., Liu, Z., Lombardozzi, D., Metzl, N., Munro, D. R., Nabel, J. E. M. S., Nakaoka, S., Neill, C., Olsen, A., Ono, T., Patra, P., Peregon, A., Peters, W., Peylin, P., Pfeil, B., Pierrot, D., Poulter, B., Rehder, G., Resplandy, L., Robertson, E., Rocher, M., Rödenbeck, C., Schuster, U., Schwinger, J., Séférian, R., Skjelvan, I., Steinhoff, T., Sutton, A., Tans, P. P., Tian, H., Tilbrook, B., Tubiello, F. N., van der Laan-Luijkx, I. T., van der Werf, G. R., Viovy, N., Walker, A. P., Wiltshire, A. J., Wright, R., Zaehle, S., & Zheng, B. (2018),Global Carbon Budget 2018. *Earth Syst. Sci. Data*, 10, 2141–2194. doi:10.5194/essd-10-2141-2018

Lin, C., Li, C., Wang, L., Bi, Y., & Zheng, Y. (2017),Preflight spectral calibration of hyperspectral carbon dioxide spectrometer of TanSat (in Chinese),Opt. Precis. Eng. 25(8), 2064–2075

Lindqvist, H., O'Dell, C. W., Basu, S., Boesch, H., Chevallier, F., Deutscher, N., Feng, L., Fisher, B., Hase, F., Inoue, M., Kivi, R., Morino, I., Palmer, P. I., Parker, R., Schneider, M., Sussmann, R., & Yoshida, Y. (2015),Does GOSAT capture the true seasonal cycle of carbon dioxide? *Atmos. Chem. Phys.*, 15, 13023–13040. doi:10.5194/acp-15-13023-2015

Liou, K. (2002),An Introduction to Atmospheric Radiation (Int'l Geophysics Series Vol 84) 2nd edition. Academic press

Liu, Y., Cai, Z., Yang, D., Duan, M., Lyu D. (2013),Optimization of the instrument configuration for TanSat CO2 spectrometer. *Sci Bull*, 58(27), 2787–2789

Liu, Y., Wang, J., Yao, L., Chen, X., Cai, Z., Yang, D., Yin, Z., Gu, S., Tian, L., Lu, N., & Lyu D. (2018),The TanSat mission: preliminary global observations. *Science Bulletin*, 63(2018), 1200-1207, 2018. doi:10.1016/j.scib.2018.08.004

Kuhlmann, G., Broquet, G., Marshall, J., Clément, V., Löscher, A., Meijer, Y., and Brunner, D.: Detectability of CO_2 emission plumes of cities and power plants with the Copernicus Anthropogenic CO_2 Monitoring (CO2M) mission. (2019),*Atmos. Meas. Tech.*, 12, 6695–6719, doi:10.5194/amt-12-6695-2019

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., & Yokota, T. (2013),Regional CO_2 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

Mandrake, L., Frankenberg, C., O'Dell, C. W., Osterman, G., Wennberg, P., & Wunch, D. (2013),Semi-autonomous sounding selection for OCO-2. *Atmos. Meas. Tech.*, 6, 2851–2864. doi:10.5194/amt-6-2851-2013

Meftah, M., Damé, L., Bolsée, D., Hauchecorne, A., Pereira, N., Sluse, D., Cessateur, G., Irbah, A., Bureau, J., Weber, M., Bramstedt, K., Hilbig, T., Thiéblemont, R., Marchand, M., Lefèvre, F., Sarkissian, A., & Bekki, S. (2018),SOLAR-ISS: A new reference spectrum based on SOLAR/SOLSPEC observations. A&A, 611(2018), A1. doi:10.1051/0004-6361/201731316

Mishchenko, M., Travis, L. D., & Lacis, A. A. (2004),Scattering, Absorption, and Emission of Light by Small Particles. Cambridge University Press

Nocedal, J., & Wright, S. W. (1999),*Numerical Optimization* (Springer Series in Operations Research and Financial Engineering),Springer

Notholt, J., Petri, C., Warneke, T., Deutscher, N., Buschmann, M., Weinzierl, C., Macatangay, R., & Gruppe, P. (2019),TCCON data from Bremen (DE), Release GGG2014R1. TCCON data archive, hosted by CaltechDATA. doi:10.14291/tccon.ggg2014.bremen01.R1, 2019

Morino, I., Matsuzaki, T., & Shishime, A. (2018a),TCCON data from Tsukuba, Ibaraki (JP) 125HR, Release GGG2014R2. TCCON data archive, hosted by CaltechDATA. doi:10.14291/tccon.ggg2014.tsukuba02.R2

Morino, I., Yokozeki, N., Matzuzaki, T., & Horikawa, M. (2018b),TCCON data from Rikubetsu (JP), Release GGG2014R1. TCCON data archive, CaltechDATA. doi:10.14291/tccon.ggg2014.rikubetsu01.R1/1242265

Morino, I., Velazco, V. A., Hori, A., Uchino, O., & Griffith, D. W. T. (2018c),TCCON data from Burgos, Ilocos Norte (PH),Release GGG2014.R0. TCCON data archive, hosted by CaltechDATA. doi:10.14291/tccon.ggg2014.burgos01.R0

Nakajima, M., Hashimoto, M., Sakai, M., Suto, H., Shiomi, K., Imai, H., Crawford, L., & Davis, G. S. (2019),Results of the commissioning phase of the mission instruments on GOSAT-2. Proc. SPIE 11151, Sensors, Systems, and Next-Generation Satellites XXIII, 1115103. doi:10.1117/12.2533898

Natraj, V., & Spurr, R. J. D. (2007a),A fast linearized pseudo-spherical two orders of scattering model to account for polarization in vertically inhomogeneous scattering–absorbing media. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 107(2): 263–293. doi:10.1016/j.jqsrt.2007.02.011

Natraja, V., Spurrb, R. J.D., Boesch, H., Jiang, Y., & Yung. Y. L. (2007b),Evaluation of errors from neglecting polarization in the forward modeling of O_2 A band measurements from space, with relevance to CO_2 column retrieval from polarization-sensitive instruments. *Journal of Quantitative Spectroscopy & Radiative Transfer*, 103(2), 245–259. doi:10.1016/j.jqsrt.2006.02.073

Nguyen, H., Osterman, G., Wunch, D., O'Dell, C., Mandrake, L., Wennberg, P., Fisher, B., & Castano, R. (2014),A method for collocating satellite XCO_2 data to ground-based data and its application to ACOS-GOSAT and TCCON. *Atmos. Meas. Tech.*, 7, 2631–2644. doi:10.5194/amt-7-2631-2014

O'Dell, C. W. (2010),Acceleration of Multiple-Scattering, Hyperspectral Radiative Transfer Calculations via Low-Streams Interpolation. *Journal of Geophysical Research – Atmospheres*, 115, D10. doi:10.1029/2009JD012803

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., & Wunch, D. (2012),The ACOS CO_2 retrieval algorithm – Part 1: Description and validation against synthetic observations. *Atmos. Meas. Tech.*, 5, 99–121. doi:10.5194/amt-5-99-2012

O'Dell, C. W., Eldering, A., Wennberg, P. O., Crisp, D., Gunson, M. R., Fisher, B., Frankenberg, C., Kiel, M., Lindqvist, H., Mandrake, L., Merrelli, A., Natraj, V., Nelson, R. R., Osterman, G. B., Payne, V. H., Taylor, T. E., Wunch, D., Drouin, B. J., Oyafuso, F., Chang, A., McDuffie, J., Smyth, M., Baker, D. F., Basu, S., Chevallier, F., Crowell, S. M. R., Feng, L., Palmer, P. I., Dubey, M., García, O. E., Griffith, D. W. T., Hase, F., Iraci, L. T., Kivi, R., Morino, I., Notholt, J., Ohyama, H., Petri, C., Roehl, C. M., Sha, M. K., Strong, K., Sussmann, R., Te, Y., Uchino, O., & Velazco, V. A. (2018),Improved retrievals of carbon dioxide from Orbiting Carbon Observatory-2 with the version 8 ACOS algorithm, Atmos. Meas. Tech., 11, 6539–6576. doi:10.5194/amt-11-6539-2018

Spurr, R.J.D., Kurosu, T.P., & Chance, K.V. (2001),A linearized discrete ordinate radiative transfer model for atmospheric remote-sensing retrieval. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 68(6), 689-735. doi:10.1016/S0022-4073(00)00055-8

Spurr, R.J.D., & Christi M. (2014),On the generation of atmospheric property Jacobians from the (V)LIDORT linearized radiative transfer models. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 142(2014), 119-105. doi:10.1016/j.jqsrt.2014.03.011

Oshchepkov, S., Bril, A., Yokota, T., Morino, I., Yoshida, Y., Matsunaga, T., Belikov, D., Wunch, D., Wennberg, P., Toon, G., O'Dell, C., Butz, A., Guerlet, S., Cogan, A., Boesch, H., Eguchi, N., Deutscher, N., Griffith, D., Macatangay, R., Notholt, J., Sussmann, R., Rettinger, M., Sherlock, V., Robinson, J., Kyrö, E., Heikkinen, P., Feist, D. G., Nagahama, T., Kadygrov, N., Maksyutov, S., Uchino, O., & Watanabe, H. (2013),Effects of atmospheric light scattering on spectroscopic observations of greenhouse gases from space. Part 2: Algorithm intercomparison in the GOSAT data processing for CO_2 retrievals over TCCON sites. *J. Geophys. Res.-Atmos.*, 118, 1493–1512. doi:10.1002/jgrd.50146

Parker, R., Boesch, H., Cogan, A., Fraser, A., Feng, L., Palmer, P. I., Messerschmidt, J., Deutscher, N., Griffith, D. W. T., Notholt, J., Wennberg, P. O., & Wunch, D. (2011),Methane observations from the Greenhouse Gases Observing SATellite: comparison to ground-based TCCON data and model calculations. *Geophys. Res. Lett.*, 38, L15807. doi:10.1029/2011GL047871

Parker, R. J., Boesch, H., Byckling, K., Webb, A. J., Palmer, P. I., Feng, L., Bergamaschi, P., Chevallier, F., Notholt, J., Deutscher, N., Warneke, T., Hase, F., Sussmann, R., Kawakami, S., Kivi, R., Griffith, D. W. T., & Velazco, V. (2015),Assessing 5 years of GOSAT Proxy XCH4 data and associated uncertainties. *Atmos. Meas. Tech.*, 8, 4785–4801. doi:10.5194/amt-8-4785-2015

Ran. Y., & Li, X. (2019),TanSat: a new star in global carbon monitoring from China. *Sci Bull*, 64(5), 284-285. doi:10.1016/j.scib.2019.01.019

Rios, L. M., & Sahinidis, N. V. (2013),Derivative-free optimization: a review of algorithms and comparison of software implementations. *J. Glob. Optim.*, 56, 1247-1293. doi:10.1007/s10898-012-9951-y

Reuter, M., Bovensmann, H., Buchwitz, M., Burrows, J. P., Con- nor, B. J., Deutscher, N. M., Griffith, D. W. T., Heymann, J., Keppel-Aleks, G., Messerschmidt, J., Notholt, J., Petri, C., Robinson, J., Schneising, O., Sherlock, V., Velazco, V., Warneke, T., Wennberg, P. O., & Wunch, D. (2011),Retrieval of atmospheric CO_2 with enhanced accuracy and precision from SCIAMACHY: Validation with FTS measurements and comparison with model results. *J. Geophys. Res.-Atmos.*, 116, D04301. doi:10.1029/2010JD015047

Rodgers, C. D., & B. J. Connor. (2003),Intercomparison of remote sounding instruments. *J. Geophys. Res.*, 108(D3), 4116. doi:10.1029/2002JD002299.

Rosenberg, R., Maxwell, S., Johnson, B. C., Chapsky, L., Lee, R. A. M., & R. Pollock. (2016),Preflight Radiometric Calibration of Orbiting Carbon Observatory 2. *IEEE T Geosci Remote*, 55(5), 2499-2508. doi:10.1109/TGRS.2016.2634023

Sherlock, V., Connor, B., Robinson, J., Shiona, H., Smale, D., & Pollard, D. F. (2014),TCCON data from Lauder (NZ), 125HR, Release GGG2014R0. TCCON data archive, hosted by CaltechDATA. doi:10.14291/tccon.ggg2014.lauder02.R0/1149298

Shiomi, K., Kawakami, S., Ohyama, H., Arai, K., Okumura, H., Taura, C., Fukamachi, T., and Sakashita, M. (2014),TCCON data from Saga (JP), Release GGG2014R0. TCCON data archive, hosted by CaltechDATA. doi:10.14291/tccon.ggg2014.saga01.R0/1149283

Somkuti, P., Boesch, H., Natraj V., & Kopparla, P. (2017),Application of a PCA-Based Fast Radiative Transfer Model to XCO_2 Retrievals in the Shortwave Infrared. *J. Geophys. Res. Atmos.*, 122, 10477–10496. doi:10.1002/2017JD027013

Somkuti, P. (2018),*Novel Methods for Atmospheric Carbon Dioxide Retrieval from the JAXA GOSAT and NASA OCO-2 Satellites*, (Doctoral dissertation),University of Leicester, UK

Somkuti, P., Bösch, H., Feng, L. Palmer, P.I., Parker, R. & Quaife, T. (2020), A new space-borne perspective of crop productivity variations over the US Corn Belt. *Agricultural and Forest Meteorology*, 281(2020), 107826. doi:10.1016/j.agrformet.2019.107826

Stokes, G. G. (1852),On the composition and resolution of streams of polarized light from different sources. Transactions of the Cambridge Philosophical Society, 9, 399

Sun, K., Liu, X., Nowlan, C. R., Cai, Z., Chance, K., Frankenberg, C., Lee, R. A. M., Pollock, R., Rosenberg, R., & Crisp, D. (2017),Characterization of the OCO-2 instrument line shape functions using on-orbit solar measurements. Atmos. Meas. Tech., 10, 939-953. doi:10.5194/amt-10-939-2017

Sussmann, R., & Rettinger, M. (2018a),TCCON data from Garmisch (DE), Release GGG2014R2. TCCON data archive, hosted by Caltech-DATA. doi:10.14291/tccon.ggg2014.garmisch01.R2

Sussmann, R., & Rettinger, M. (2018b),TCCON data from Zugspitze (DE), Release GGG2014R1. TCCON data archive, hosted by CaltechDATA. doi:10.14291/tccon.ggg2014.zugspitze01.R1

Taylor, T. E., O'Dell, C. W., O'Brien, D. M., Kikuchi, N., Yokota, T., Nakajima, T. Y., Ishida, H., Crisp, D., & Nakajima, T. (2012),Comparison of cloud-screening methods applied to GOSAT near-infrared spectra. *IEEE T. Geosci. Remote*, 50, 295–309. doi:10.1109/TGRS.2011.2160270

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., & Gunson, M. R. (2016),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

Té, Y., Jeseck, P., & Janssen, C. (2014),TCCON data from Paris (FR), Release GGG2014R0. TCCON data archive, hosted by CaltechDATA. doi:10.14291/tccon.ggg2014.paris01.R0/1149279

Toon, G. C. (2014),Solar line list for GGG2014, TCCON data archive. hosted by the Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee, U.S.A. doi:10. 14291/tccon.ggg2014.solar.R0/1221658

Velazco V. A., Morino, I., Uchino, O., Hori, A., Kiel, M., Bukosa, B., Deutscher, N. M., Sakai, T., Nagai, T., Bagtasa, G., Izumi, T., Yoshida, Y., & Griffith, D. W. T. (2017),TCCON Philippines: First Measurement Results, Satellite Data and Model Comparisons in Southeast Asia. *Remote Sens*, 9(12), 1228. doi:10.3390/rs9121228

Wang, Q., Yang, Z. D., & Bi, Y. M. (2014),Paper presented at Spectral parameters and signal-to-noise ratio requirement for CO_2 hyper spectral remote sensor. SPIE Asia-Pacific Remote Sensing, Beijing, China

Wang., L., Lin, C., Ji, Z., Zheng, Y., & Bi, Y. (2018),Preflight diffuser's calibration of carbon dioxide spectrometer of TanSat. Optics and Precision Engineering, 26(8), 1967-1976

Warneke, T., Messerschmidt, J., Notholt, J.,Weinzierl, 5 C., Deutscher, N. M., Petri, C., Grupe, P., Vuillemin, C., Truong, F., Schmidt, M., Ramonet, M., & Parmentier, E. (2019),TCCON data from Orléans (FR), Release GGG2014R1. TCCON data archive, hosted by CaltechDATA. doi:10.14291/tccon.ggg2014.orleans01.R1

Wennberg, P. O., Roehl, C. M., Blavier, J.-F., Wunch, D., Landeros, J., & Allen, N. T. (2014),TCCON data from Jet Propulsion Laboratory (US), 2011, Release GGG2014.R0 (Version GGG2014.R0) [Data set]. CaltechDATA. doi:10.14291/tccon.ggg2014.jpl02.r0/1149297

Wennberg, P. O.,Wunch, D., Roehl, C. M., Blavier, J.-F., Toon, G. C., & Allen, N. T. (2015),TCCON data from Caltech (US), Release GGG2014R1. TCCON data archive, hosted by CaltechDATA. doi:10.14291/tccon.ggg2014.pasadena01.R1/1182415

Wennberg, P. O., Wunch, D., Roehl, C. M., Blavier, J.-F., Toon, G. C., Allen, N. T., Dowell, P., Teske, K., Martin, C., & Martin, J. (2016),TCCON data from Lamont (US), Release GGG2014R1. TCCON data archive, hosted by CaltechDATA. doi:10.14291/tccon.ggg2014.lamont01.R1/1255070

Wennberg, P. O., Roehl, C. M., Wunch, D., Toon, G. C., Blavier, J.-F., Washenfelder, R., Keppel-Aleks, G., Allen, N. T., & Ayers, J. (2017),TCCON data from Park Falls (US), Release GGG2014R1. TCCON data archive, hosted by CaltechDATA. doi:10.14291/tccon.ggg2014.parkfalls01.R1

Wu, L., Hasekamp, O., Hu, H., Landgraf, J., Butz, A., aan de Brugh, J., Aben, I., Pollard, D. F., Griffith, D. W. T., Feist, D. G., Koshelev, D., Hase, F., Toon, G. C., Ohyama, H., Morino, I., Notholt, J., Shiomi, K., Iraci, L., Schneider, M., de Mazière, M., Sussmann, R., Kivi, R., Warneke, T., Goo, T.-Y., & Té, Y. (2018),Carbon dioxide retrieval from OCO-2 satellite observations using the RemoTeC algorithm and validation with TCCON measurements. *Atmos. Meas. Tech.*, 11, 3111-3130. doi:10.5194/amt-11-3111-2018

Wu, L., Hasekamp, O., Hu, H., aan de Brugh, J., Landgraf, J., Butz, A., & Aben, I. (2019),Full-physics carbon dioxide retrievals from the Orbiting Carbon Observatory-2 (OCO-2) satellite by only using the 2.06 μm band. *Atmos. Meas. Tech.*, 12, 6049–6058. doi:10.5194/amt-12-6049-2019

Wunch, D., Toon, G. C., Blavier, J.-F. L., Washenfelder, R. A., Notholt, J., Connor, B. J., Griffith, D. W. T., Sherlock, V., & Wennberg, P. O. (2011a),The Total Carbon Column Observing Network. *Phil. Trans. R. Soc. A*, 369, 2087–2112. doi:10.1098/rsta.2010.0240

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. (2011b),A method for evaluating bias in global measurements of CO_2 total columns from space. *Atmos. Chem. Phys.*, 11, 12317–12337. doi:10.5194/acp-11-12317-2011

Wunch, D., Toon, G. C., Sherlock, V., Deutscher, N. M., Liu, C., Feist, D. G., & Wennberg, P. O. (2015),The Total Carbon Column Observing Network's GGG2014 Data Version, Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA, available at: https://data.caltech.edu/records/249

Wunch, D., Wennberg, P. O., Osterman, G., Fisher, B., Naylor, B., Roehl, C. M., O'Dell, C., Mandrake, L., Viatte, C., Kiel, M., Griffith, D. W. T., Deutscher, N. M., Velazco, V. A., Notholt, J., Warneke, T., Petri, C., De Maziere, M., Sha, M. K., Sussmann, R., Rettinger, M., Pollard, D., Robinson, J., Morino, I., Uchino, O., Hase, F., Blumenstock, T., Feist, D. G., Arnold, S. G., Strong, K., Mendonca, J., Kivi, R., Heikkinen, P., Iraci, L., Podolske, J., Hillyard, P. W., Kawakami, S., Dubey, M. K., Parker, H. A., Sepulveda, E., García, O. E., Te, Y., Jeseck, P., Gunson, M. R., Crisp, D., & Eldering, A. (2017),Comparisons of the Orbiting Carbon Observatory-2 (OCO-2) XCO2 measurements with TCCON. *Atmos. Meas. Tech.*, 10, 2209–2238. doi:10.5194/amt-10-2209-2017

Wunch, D., Mendonca, J., Colebatch, O., Allen, N. T., Blavier, 5 J.-F., Roche, S., Hedelius, J., Neufeld, G., Springett, S., Worthy, D., Kessler, R., & Strong, K. (2018),TCCON data from East Trout Lake, SK (CA), Release GGG2014.R1. TCCON data archive, hosted by CaltechDATA. doi:10.14291/tccon.ggg2014.easttroutlake01.R1

Yang, D., Liu, Y., Cai, Z., Deng, J., Wang, J., & Chen, X. (2015),An Advanced Carbon Dioxide Retrieval Algorithm for Satellite Measurements and its Application to GOSAT Observations. *Chin. Sci. Bull.*, 60(23), 2063–2066. doi:10.1007/s11434-015-0953-2

Yang, D., Zhang, H., Liu, Y., Chen B., Cai, Z., & Lyu, D. (2017),Monitoring carbon dioxide from space: Retrieval algorithm and flux inversion based on GOSAT data and using CarbonTracker-China. *Adv. Atmos. Sci.*, 34(8), 965–976. doi:10.1007/s00376-017-6221-4.

Yang, D., Liu, Y., Cai, Z., Chen, X., Yao, L., & Lu, D. (2018),First Global Carbon Dioxide Maps Produced from TanSat Measurements. *Adv. Atmos. Sci.*, 35, 621–623. doi:10.1007/s00376- 018-7312-6

Yokota, T., Yoshida, Y., Eguchi, N., Ota, Y., Tanaka, T., Watanabe, H., & Maksyutov, S. (2009),Global concentrations of CO_2 and CH_4 retrieved from GOSAT: First preliminary results. Scientific Online Letters on the Atmosphere, 5(1), 160-163. doi:10.2151/sola.2009-041

Yoshida, Y., Kikuchi, N., Morino, I., Uchino, O., Oshchepkov, S., Bril, A., Saeki, T., Schutgens, N., Toon, G. C., Wunch, D., Roehl, C. M., Wennberg, P. O., Griffith, D. W. T., Deutscher, N. M., Warneke, T., Notholt, J., Robinson, J., Sherlock, V., Connor, B., Rettinger, M., Sussmann, R., Ahonen, P., Heikkinen, P., Kyrö, E., Mendonca, J., Strong, K., Hase, F., Dohe, S., & Yokota, T. (2013),Improvement of the retrieval algorithm for GOSAT SWIR XCO_2 and XCH_4 and their validation using TCCON data. *Atmos. Meas. Tech.*, 6, 1533-1547. doi:10.5194/amt-6-1533-2013

Uchino, N. Kikuchi, T. Sakai, I. Morino, Y. Yoshida, T. Nagai, A. Shimizu, T. Shibata, A. Yamazaki, A. Uchiyama, N. Kikuchi, S. Oshchepkov, A. Bril1, & T. Yokota. (2012),Influence of aerosols and thin cirrus clouds on the GOSAT-observed CO_2: a case study over Tsukuba. *Atmos. Chem. Phys.*, 12, 3393–3404

Velazco, V., Morino, I., Uchino, O., Hori, A., Kiel, M., Bukosa, B., Deutscher, N., Sakai, T., Nagai, T., Bagtasa, G., Izumi, T., Yoshida, Y., & Griffith, D. (2017),TCCON Philippines: First Measurement Results, Satellite Data and Model Comparisons in Southeast Asia. *Remote Sens.*, 9(12), 1228. doi:10.3390/rs9121228

Zhang, H., Zheng, Y., Li, S., Lin, C., Li, C., Yuan, J., & Li, Y. (2019),Geometric correction for TanSat atmospheric carbon dioxide grating spectrometer. *Sensors and Actuators A-Physical*, 293, 62-69. doi:10.1016/j.sna.2019.04.034

Figure 1. Mean spectrum of the normalized solar calibration measurement of the O2 A band. The average is calculated for each calibration measurement in an orbit. Red, green and blue indicate the mean measurement taken for 2017/08/15, 2017/11/15 and 2018/02/15 respectively. Sub-figures represent the footprints (FP1-9) across the swath. The date is shown in the legend of FP1 sub-figure.

Figure 2. Mean fitting residuals of the solar calibration measurements. The average is calculated from the fitting residuals of 181,232 individual spectra (for each footprint, the quantity is ~20,136) during Mar. 2017 – May 2018. The blue and red lines show the mean residual obtained by fitting a 5th order polynomials or a 8th order Fourier series model, respectively (see the detail in text section 2.3). The light grey line is the fitting residual when fitting a scale factor without any wavelength dependent corrections. Sub-figures indicate the 9 footprints (FP1-FP9) across the swath. The RMSE shown in the title of each sub-figure shows the improvement of the Fourier series model compared to using the wavelength-independent scale factor.

Figure 3. Comparison of surface pressure changes from O2 A band retrievals with and without Fourier series continuum correction for selected clear-sky conditions around TCCON/Lamont (USA) site. The changes are calculated by subtracting the a priori (height corrected surface pressure from ECMWF interim) from the retrieved apparent surface pressure. The sub-figures show the selected cases on 2017-10-08, 2017-10-26, 2018-01-24 and 2018-03-01. Corresponding images from MODIS/Aqua (https://worldview.earthdata.nasa.gov) are shown in the bottom right corner of each corresponding sub-figure. The color represents the footprints (FP1-FP9) across the swath as indicated in the upper left sub-figure. The retrieved apparent surface pressure with continuum correction still shows a small bias to the a priori, because the retrieval has been carried out without any consideration of aerosol and cloud scattering (see the detail in text section 2.4).

Figure 4. Mean O2 A band fitting residuals (normalized by continuum level) from the two bands retrieval with (red) and without (blue) Fourier series continuum correction. The average is computed from selected clear-sky measurements on 08/10/2017 around the TCCON/Lamont site. The corresponding MODIS/Aqua RGB image is shown in Figure 3. The light grey line (right y-axis) shows the measurement spectrum as reference. The red and blue shading indicates the continuum normalized standard deviation (SD) for the retrieval with and without Fourier series continuum correction respectively.

Figure 5. Same as figure 4, but for the CO2 weak band.

Figure 6. The optimal Target run Genetic Algorithm (GA) profile (Pareto-optimal trade-off curves) for the selection of filter complexity and transparency. The XCO2 mean individual RMSE is a total RMSE calculated from all retrievals which pass the optimized threshold value (good retrieval). The color indicates the number of parameters (complexity) that are used in the GA run. For each complexity, the filter is determined and fixed (see section 3.2.1 for details and table 3 for filter definitions). Transparency has a 1% interval through 0-100%. No significant additional reduction in the RMSE was seen when using more than 5 filters. In this study, we cut-off at 2 ppm with 5 filters, the transparency is 64.3%.

Figure 7. The performance of the GA with 2-5 filters (rows 1-4) and a manual filter threshold selection with 5 filters (row 5). The histograms (left y-axis) indicate the counts for each filter bin, and red and black solid points show the bias and RMSE for each bin (right y-axis). The grey line is the upper and lower threshold used for each filter. The filters are applied from left to right (columns 1-5) sequentially.

Figure 8. The individual XCO2 error (UoL-FP/TanSat against TCCON) change with and without parameter bias correction. The orange and blue histograms indicate the XCO2 individual error distribution with and without bias correction. The improvement of RMSE with and without bias correction for each footprint (FP 1-9) across the swath is shown in the titles.

Figure 9. Histogram of individual XCO2 retrieval errors (difference between TanSat and TCCON XCO2) for TCCON/Lamont for clear-sky measurements. The orange (right y-axis) and blue (left y-axis) give the results for the two-band retrieval with (orange) and without (blue) Fourier series continuum correction. All data statistics in this figure passed quality control, but there is no bias correction applied. The RMSE remove bias shown in text box of this figure gives the RMSE calculated from each individual error after subtracting the overall bias.

Figure 10. Same as figure 9, but for a comparison between a single CO2 weak band retrieval (on continuum correction) and the two-band retrieval with Fourier series continuum correction.

Figure 11. Validation of UoL-FP/TanSat XCO2 retrievals against measurements from 20 TCCON sites. Each symbol represents the mean of one overpass (see detail in text section 3.1 for colocation criteria) for TanSat (only shown if the available quantity N>50) and the TCCON average during the overpass (only show if the available quantity N>20). The total number of overpasses per site is given in the legend. In total 174 daily data couples are involved in this validation. Statistics are shown in the upper-left corner of the figure. The daily mean RMSE is the total RMSE computed from each overpass mean and the site mean RMSE is computed by averaging the RMSE of each site. The black line indicates the 1:1 line as reference. The slope, R2 and fitting RMSE are the statistics from a linear regression weighted by bi-square.

Figure 12. XCO2 time series from 2017 to 2018 for each TCCON site used in this study (site name is shown in the title of each sub-figure). The green and red large solid circles represent TCCON and TanSat overpass mean with error bar indicating the standard deviation. The individual measurements are shown in light green and light red small points for TCCON and TanSat overpasses, respectively. The TCCON measurements are only shown when the quantity is greater than 20 and TanSat measurements are only shown when the quantity is greater than 50. This figure shows all available TCCON and TanSat data, not only TanSat-TCCON couples, hence the data quantity shown in this figure is larger than is used for the validation.

Figure 13. Seasonal variation of the XCO2 overpass mean error (difference between TanSat and TCCON XCO2). The data couples shown in this figure are the same as used in Figure 11. Only sites with more than 9 overpasses data points (the number of data points for each site is shown in parentheses of each sub-figure caption) have been shown in this figure. The error bars indicate the standard deviation. The dashed grey line shows the zero-error line as reference. The site name and RMSE are given in the title for each sub-figure.

Figure A1. A simulation on the impact of the polarization angle on a simulated O2 A Band spectrum. The left sub-figure shows the changes of the spectrum when the polarization angle is 20º, 40º, 60º and 80º larger than the true angle. The color index is marked in the text box of the right sub-figure. The right sub-figure shows the relative change to the normalized spectrum against radiance level. Notice that the change becomes larger when the radiance decreases. This effect introduces significant errors in deep absorption lines where important information for the retrieval comes from. The simulated scene has a surface albedo of 0.2 and solar zenith angle of 30º in nadir measurement mode. The calculated polarization angle is 86.3º as calculated by the method described in section 2.1.2.

Table 1. List of state vector elements and descriptions for the UoL-FP/TanSat retrieval

Table 2. The TCCON site list used in the validation study and the site validation statistics

Table 3. Selected filter used in quality control and corresponding lower and upper thresholds