



**HAL**  
open science

## **The Adaptable 4A Inversion (5AI): description and first XCO<sub>2</sub> retrievals from Orbiting Carbon Observatory-2 (OCO-2) observations**

Matthieu Dogniaux, Cyril Crevoisier, Raymond Armante, Virginie Capelle, Thibault Delahaye, Vincent Cassé, Martine de Mazière, Nicholas M Deutscher, Dietrich G Feist, Omaira E Garcia, et al.

### ► To cite this version:

Matthieu Dogniaux, Cyril Crevoisier, Raymond Armante, Virginie Capelle, Thibault Delahaye, et al.. The Adaptable 4A Inversion (5AI): description and first XCO<sub>2</sub> retrievals from Orbiting Carbon Observatory-2 (OCO-2) observations. Atmospheric Measurement Techniques, 2021, 14 (6), pp.4689 - 4706. 10.5194/amt-14-4689-2021 . hal-03288099

**HAL Id: hal-03288099**

**<https://hal.sorbonne-universite.fr/hal-03288099>**

Submitted on 16 Jul 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



# The Adaptable 4A Inversion (5AI): description and first $X_{\text{CO}_2}$ retrievals from Orbiting Carbon Observatory-2 (OCO-2) observations

Matthieu Dogniaux<sup>1</sup>, Cyril Crevoisier<sup>1</sup>, Raymond Armante<sup>1</sup>, Virginie Capelle<sup>1</sup>, Thibault Delahaye<sup>1</sup>, Vincent Cassé<sup>1</sup>, Martine De Mazière<sup>2</sup>, Nicholas M. Deutscher<sup>3,4</sup>, Dietrich G. Feist<sup>5,6,7</sup>, Omaira E. Garcia<sup>8</sup>, David W. T. Griffith<sup>3</sup>, Frank Hase<sup>9</sup>, Laura T. Iraci<sup>10</sup>, Rigel Kivi<sup>11</sup>, Isamu Morino<sup>12</sup>, Justus Notholt<sup>4</sup>, David F. Pollard<sup>13</sup>, Coleen M. Roehl<sup>14</sup>, Kei Shiomi<sup>15</sup>, Kimberly Strong<sup>16</sup>, Yao Té<sup>17</sup>, Voltaire A. Velasco<sup>3</sup>, and Thorsten Warneke<sup>4</sup>

<sup>1</sup>Laboratoire de Météorologie Dynamique/IPSL, CNRS, École polytechnique, Institut Polytechnique de Paris, Sorbonne Université, École Normale Supérieure, PSL Research University, 91120 Palaiseau, France

<sup>2</sup>Royal Belgian Institute for Space Aeronomy, Brussels, Belgium

<sup>3</sup>Centre for Atmospheric Chemistry, School of Earth, Atmospheric and Life Sciences, University of Wollongong, Wollongong, Australia

<sup>4</sup>University of Bremen, Bremen, Germany

<sup>5</sup>Max Planck Institute for Biogeochemistry, Jena, Germany

<sup>6</sup>Lehrstuhl für Physik der Atmosphäre, Ludwig-Maximilians-Universität München, Munich, Germany

<sup>7</sup>Deutsches Zentrum für Luft- und Raumfahrt, Institut für Physik der Atmosphäre, Oberpfaffenhofen, Germany

<sup>8</sup>Izaña Atmospheric Research Center (IARC), State Meteorological Agency of Spain (AEMET), Tenerife, Spain

<sup>9</sup>Institute of Meteorology and Climate Research (IMK-ASF), Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany

<sup>10</sup>NASA Ames Research Center, Moffett Field, CA, USA

<sup>11</sup>Finnish Meteorological Institute, Sodankylä, Finland

<sup>12</sup>National Institute for Environmental Studies (NIES), Tsukuba, Japan

<sup>13</sup>National Institute of Water and Atmospheric Research Ltd (NIWA), Lauder, New Zealand

<sup>14</sup>Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA, USA

<sup>15</sup>Japan Aerospace Exploration Agency (JAXA), Tsukuba, Japan

<sup>16</sup>Department of Physics, University of Toronto, Toronto, Canada

<sup>17</sup>Laboratoire d'Études du Rayonnement et de la Matière en Astrophysique et Atmosphères (LERMA-IPSL), Sorbonne Université, CNRS, Observatoire de Paris, PSL Université, 75005 Paris, France

**Correspondence:** Matthieu Dogniaux (matthieu.dogniaux@lmd.ipsl.fr)

Received: 5 October 2020 – Discussion started: 10 November 2020

Revised: 22 March 2021 – Accepted: 18 April 2021 – Published: 24 June 2021

**Abstract.** A better understanding of greenhouse gas surface sources and sinks is required in order to address the global challenge of climate change. Space-borne remote estimations of greenhouse gas atmospheric concentrations can offer the global coverage that is necessary to improve the constraint on their fluxes, thus enabling a better monitoring of anthropogenic emissions. In this work, we introduce the Adaptable 4A Inversion (5AI) inverse scheme that aims to retrieve geophysical parameters from any remote sensing observation. The algorithm is based on the Optimal Estimation algorithm,

relying on the Operational version of the Automated Atmospheric Absorption Atlas (4A/OP) radiative transfer forward model along with the Gestion et Étude des Informations Spectroscopiques Atmosphériques: Management and Study of Atmospheric Spectroscopic Information (GEISA) spectroscopic database. Here, the 5AI scheme is applied to retrieve the column-averaged dry air mole fraction of carbon dioxide ( $X_{\text{CO}_2}$ ) from a sample of measurements performed by the Orbiting Carbon Observatory-2 (OCO-2) mission. Those have been selected as a compromise between

coverage and the lowest aerosol content possible, so that the impact of scattering particles can be neglected, for computational time purposes. For air masses below 3.0, 5AI  $X_{\text{CO}_2}$  retrievals successfully capture the latitudinal variations of  $\text{CO}_2$  and its seasonal cycle and long-term increasing trend. Comparison with ground-based observations from the Total Carbon Column Observing Network (TCCON) yields a bias of  $1.30 \pm 1.32$  ppm (parts per million), which is comparable to the standard deviation of the Atmospheric  $\text{CO}_2$  Observations from Space (ACOS) official products over the same set of soundings. These nonscattering 5AI results, however, exhibit an average difference of about 3 ppm compared to ACOS results. We show that neglecting scattering particles for computational time purposes can explain most of this difference that can be fully corrected by adding to OCO-2 measurements an average calculated–observed spectral residual correction, which encompasses all the inverse setup and forward differences between 5AI and ACOS. These comparisons show the reliability of 5AI as an optimal estimation implementation that is easily adaptable to any instrument designed to retrieve column-averaged dry air mole fractions of greenhouse gases.

## 1 Introduction

The atmospheric concentration of carbon dioxide ( $\text{CO}_2$ ) has been rising for decades because of fossil fuel emissions and land use changes. However, large uncertainties still remain in the global carbon budget (e.g. Le Quéré et al., 2009). In order to address the global challenge of climate change, a better understanding of carbon sources and sinks is necessary, and remote space-borne estimations of  $\text{CO}_2$  columns can help to constrain these carbon fluxes in atmospheric inversion studies, thus reducing the remaining uncertainties (e.g. Rayner and O'Brien, 2001; Chevallier et al., 2007; Basu et al., 2013, 2018).

The column-averaged dry air mole fraction of  $\text{CO}_2$  ( $X_{\text{CO}_2}$ ) can be retrieved from thermal infrared (TIR) soundings, mostly sensitive to the midtroposphere (e.g. Chédin et al., 2003; Crevoisier et al., 2004, 2009a), and from near-infrared (NIR) and shortwave infrared (SWIR) measurements, which are sensitive to the whole atmospheric column and especially to levels close to the surface where carbon fluxes take place. Current NIR and SWIR satellite missions observing carbon dioxide include the Japanese Greenhouse gases Observing SATellites (GOSAT and GOSAT-2), NASA's Orbiting Carbon Observatory-2 and 3 (OCO-2 and OCO-3) and the Chinese mission TanSat. Over time, different algorithms have been developed to exploit their measurements; those rely on different inverse methods and use various hypotheses to address the fundamentally ill-posed problem of  $X_{\text{CO}_2}$  retrieval. These algorithms, notably, include the Japanese National Institute for Environmental Studies (NIES) algorithm (Yokota et al., 2009; Yoshida et

al., 2011, 2013), the Atmospheric  $\text{CO}_2$  Observations from Space (ACOS) algorithm (Bösch et al., 2006; Connor et al., 2008; O'Dell et al., 2012, 2018), the University of Leicester Full Physics (UoL-FP) retrieval algorithm from the University of Leicester (Parker et al., 2011), RemoTeC from the Netherlands Institute for Space Research (SRON; Butz et al., 2011; Wu et al., 2018) and the Fast atmospheric trace gAs retrieval (FOCAL) algorithm from the University of Bremen (Reuter et al., 2017a, b).

They also rely on different forward radiative transfer models to compute synthetic measurements and their partial derivatives. It was noted, for the Scanning Imaging Absorption Spectrometer for Atmospheric Cartography (SCIAMACHY) mission (Bovensmann et al., 1999), that NIR and SWIR measurements are also quite sensitive to the presence of scattering particles on the optical path, which can then substantially perturb  $X_{\text{CO}_2}$  retrievals if unaccounted for (Houweling et al., 2005). As exact multiple scattering calculations are too time consuming for operational  $X_{\text{CO}_2}$  retrievals, all the previously mentioned retrieval algorithms have radiative transfer models that implement various approximations to speed up forward modelling. Finally, radiative transfer fundamentally depends on spectroscopic databases that contain the parameters enabling to compute atmospheric gas absorption. The HITRAN spectroscopic database (latest version from 2016; Gordon et al., 2017) is widely used for greenhouse gas concentration retrievals, as are the absorption coefficient (ABSCO) atmospheric absorption tables for the ACOS algorithm (Drouin et al., 2017; Oyafuso et al., 2017).

The design of an  $X_{\text{CO}_2}$  retrieval algorithm, from the forward model and the spectroscopic parameters it uses to the choice of the adjusted quantities in the state vector, has a critical influence on the overall performance of the observing system (Rodgers, 2000). The systematic errors in retrieved  $X_{\text{CO}_2}$  and their standard deviations (the latter being also called single measurement precision) with regard to the true (but unknown) state of the atmosphere particularly impact the uncertainty reduction and bias in atmospheric  $\text{CO}_2$  flux inversion studies (e.g. Chevallier et al., 2007). Retrieved  $X_{\text{CO}_2}$  products are most often validated against columns with similar observation geometry, like the ground-based solar absorption spectrometry. The Total Carbon Column Observing Network (TCCON) is a network of ground stations that retrieve the column-averaged dry air mole fraction of  $\text{CO}_2$  and other species from NIR and SWIR spectra measured with Fourier transform spectrometers (FTSs) directly pointing at the Sun (Wunch et al., 2011b). The network currently consists of 27 stations all around the world, and its products constitute a truth proxy reference for the validation of space-borne retrievals of greenhouse gas atmospheric concentrations. For instance, TCCON data sets were used to validate SCIAMACHY (Reuter et al., 2011), GOSAT  $X_{\text{CO}_2}$  retrieved by the ACOS (Wunch et al., 2011a) and NIES algorithms (Inoue et al., 2016), and OCO-2  $X_{\text{CO}_2}$  produced by

ACOS (O'Dell et al., 2018; Wunch et al., 2017), RemoTeC (Wu et al., 2018), and FOCAL (Reuter et al., 2017b). These three last algorithms exhibit different biases compared to TCCON, depending on their respective forward modelling and bias correction strategies, i.e.  $0.30 \pm 1.04$ ,  $0.0 \pm 1.36$ , and  $0.67 \pm 1.34$  ppm (parts per million) for OCO-2 nadir land soundings, respectively.

In this paper, we present the Adaptable 4A Inversion (5AI) that implements the optimal estimation inverse method (Rodgers, 2000). 5AI relies on the OPERational version of the Automatized Atmospheric Absorption Atlas (4A/OP) radiative transfer model (Scott and Chédin, 1981; Tournier, 1995; Cheruy et al., 1995; <https://4aop.aeris-data.fr>, last access: 25 May 2021) and the GEISA (Gestion et Étude des Informations Spectroscopiques Atmosphériques: Management and Study of Spectroscopic Information) spectroscopic database (Jacquinet-Husson et al., 2016; <http://cds-espri.ipsl.fr/etherTypo/?id=950>, last access: 25 May 2021). The 5AI scheme is applied to retrieve  $X_{\text{CO}_2}$  from a sample of OCO-2 measurements that compromises between coverage and the lowest possible values of ACOS retrieved aerosol optical depth in order to avoid possible singular biases due to strong aerosol events. This sample selection comprises the (1) OCO-2 best flag target mode soundings between 2014 and 2018 and (2) a sample of 2 years of OCO-2 nadir measurements with a global land coverage. First, for computational time purposes, retrievals are performed without taking into account scattering particles. We then discuss how considering them and accounting for differences in the radiative transfer modelling and retrieval setups impact the 5AI results, which are compared to ACOS and TCCON reference data over identical sets of soundings.

This paper is organised as follows: Sect. 2 describes the 5AI retrieval scheme and its current features, as well as the 4A/OP radiative transfer model, the GEISA spectroscopic database, and the empirically corrected O<sub>2</sub> A-band absorption continuum on which it relies. Section 3 presents the OCO-2 and TCCON data selection. Section 4 presents the a posteriori filters used for this work and shows the 5AI  $X_{\text{CO}_2}$  target and nadir retrieval results which are compared to TCCON and ACOS (B8r version). 5AI results are discussed in Sect. 5. It shows how taking into account scattering particles in 5AI retrievals can impact the results and how systematic differences between different  $X_{\text{CO}_2}$  products can be accounted for by compensating them with a spectral residual adjustment  $X_{\text{CO}_2}$ . Finally, Sect. 6 highlights the conclusions of this work.

## 2 The 5AI retrieval scheme

As for any other retrieval scheme, 5AI aims at finding the estimate of atmospheric and surface parameters (for example, trace gas concentration, temperature profile, surface albedo, or scattering particle optical depth) that best fits hyperspec-

tral measurements made from space. This inverse problem can be expressed with the following equation:

$$\mathbf{y} = \mathbf{F}(\mathbf{x}) + \boldsymbol{\varepsilon}, \quad (1)$$

where  $\mathbf{y}$  is the measurement vector containing the radiances measured by the space instrument,  $\mathbf{x}$  is the state vector containing the geophysical parameters to be retrieved,  $\boldsymbol{\varepsilon}$  is the measurement noise, and, finally,  $\mathbf{F}$  is the forward radiative transfer model that describes the physics linking the geophysical parameters to be retrieved to the measured infrared radiances.

### 2.1 Forward modelling – 4A/OP and GEISA spectroscopic database

The 5AI retrieval scheme uses the OPERational version of the Automatized Atmospheric Absorption Atlas (4A/OP). 4A/OP is an accurate, line-by-line radiative transfer model that enables a fast computation of atmospheric transmittances based on atlases containing precomputed monochromatic optical thicknesses for reference atmospheres. Those are used to compute atmospheric transmittances for any input atmospheric profile and viewing configuration that enable one to solve the radiative transfer equation and yield radiances and their partial derivatives with respect to the input geophysical parameters at a pseudo-infinite spectral resolution ( $0.0005 \text{ cm}^{-1}$  best) or can be convolved with an instrument function. 4A/OP is the reference radiative transfer model for the Centre National d'Études Spatiales (CNES) IASI (Infrared Atmospheric Sounding Interferometer) level 1 calibration/validation and operational processing, and it is used for daily retrieval of midtropospheric columns of CO<sub>2</sub> (Crevoisier et al., 2009a) and CH<sub>4</sub> (Crevoisier et al., 2009b) from IASI. Moreover, 4A/OP has also been chosen by CNES as the reference radiative transfer model for the development of the New Generation of the IASI instrument (IASI-NG; Crevoisier et al., 2014) and the French NIR and SWIR CO<sub>2</sub> remote sensing MicroCarb mission (Pascal et al., 2017) and the French–German Methane Remote sensing Lidar Mission (MERLIN; Ehret et al., 2017).

Although originally developed for the thermal infrared spectral region, 4A/OP has been extended to the near- and shortwave infrared regions (NIR and SWIR). (1) The computation of the atlases of optical thickness was extended to the  $3000\text{--}13\,500 \text{ cm}^{-1}$  domain and takes into account line mixing and collision-induced absorption (CIA) in the O<sub>2</sub> A band (Tran and Hartmann, 2008), as well as line mixing and H<sub>2</sub>O broadening of CO<sub>2</sub> lines (Lamouroux et al., 2010). The absorption lines of CO<sub>2</sub> that we use in this work are, thus, identical to those included in HITRAN 2008. (2) The solar spectrum is a flexible input, and the Doppler shift of its lines is computed. (3) The radiative transfer model is now coupled with the LIDORT model (Spurr, 2002) for scalar multiple-scattering simulation performed with the discrete ordinates method and with VLIDORT (Spurr, 2006) if po-

larisation or bidirectional reflectance distribution functions (BRDFs) need to be taken into account. This coupling especially enables us to take into account Rayleigh scattering and, if necessary, scattering particles in NIR and SWIR forward modelling.

The 4A/OP radiative transfer model can be used with monochromatic optical thickness atlases computed from any spectroscopic database. For this present work, the atlases are computed using the GEISA 2015 (Gestion et Étude des Informations Spectroscopiques Atmosphériques: Management and Study of Spectroscopic Information) spectroscopic database. Being the base of many projects since the beginning in astronomical and astrophysical communities, GEISA has also been used since the 2000s for the preparation of several current and future spatial missions and has been chosen by CNES as the reference spectroscopic database for the definition of IASI-NG, MicroCarb, and MERLIN. Due to imperfections in the Tran and Hartmann (2008) line mixing and CIA models, an empirical correction to the absorption continuum in the O<sub>2</sub> A band, fitted from Park Falls TC-CON spectra, following the method described in Drouin et al. (2017), has been added. Finally, we use Toon (2015) as input solar spectra.

## 2.2 Inverse modelling in the 5AI retrieval scheme

### 2.2.1 Optimal estimation applied for X<sub>CO<sub>2</sub></sub> retrieval

The whole formalism of optimal estimation that enables us to find a satisfying solution to Eq. (1) may be found in Rodgers (2000). This subsection only outlines the key steps that are implemented in order to retrieve X<sub>CO<sub>2</sub></sub>.

Equation (1) includes  $\boldsymbol{\varepsilon}$ , the experimental noise of the measured radiances. Hence, it appears more appropriate to use a formalism that takes into account this measurement uncertainty and translates it into retrieval uncertainty; this is done by representing the state of the atmosphere  $\mathbf{x}$  and the measurement  $\mathbf{y}$  as random variables. Assuming Gaussian statistics, the inversion problem consists in finding the state vector which compromises between an a priori knowledge of the geophysical state parameters (most often brought by climatologies) and the information brought by the measurement, both weighted by their respective uncertainties. It finally boils down to the minimisation of the following  $\chi^2$  cost function:

$$\chi^2 = (\mathbf{y} - \mathbf{F}(\mathbf{x}))^T \mathbf{S}_e^{-1} (\mathbf{y} - \mathbf{F}(\mathbf{x})) + (\mathbf{x} - \mathbf{x}_a)^T \mathbf{S}_a^{-1} (\mathbf{x} - \mathbf{x}_a), \quad (2)$$

where  $\mathbf{x}_a$  is the a priori state vector, which is also, in most cases, chosen as the first guess for iterative retrievals. Assuming again Gaussian statistics,  $\mathbf{S}_a$  is the a priori state covariance matrix that represents the variability around the a priori state vector, and similarly,  $\mathbf{S}_e$  is the a priori measurement error covariance matrix that represents the noise model of the instrument. Moreover, as the forward model for this

retrieval is highly nonlinear, it is practical to use a local linear approximation, here expressed around the a priori state as follows:

$$\mathbf{F}(\mathbf{x}) = \mathbf{F}(\mathbf{x}_a) + \frac{\partial \mathbf{F}}{\partial \mathbf{x}} \Big|_{\mathbf{x}_a} (\mathbf{x} - \mathbf{x}_a). \quad (3)$$

The partial derivatives of the forward radiative transfer model  $\mathbf{F}$  (here 4A/OP) are expressed as a matrix, called the Jacobian matrix, and is denoted as  $\mathbf{K}$ .

All these assumptions enable the maximum posterior probability state  $\hat{\mathbf{x}}$  that minimises the cost function defined in Eq. (2) to be found. It can be computed by iteration, using the general approach, as follows:

$$\mathbf{x}_{i+1} = \mathbf{x}_i + \left[ (1 + \gamma) \mathbf{S}_a^{-1} + \mathbf{K}_i^T \mathbf{S}_e^{-1} \mathbf{K}_i \right]^{-1} \left( \mathbf{K}_i^T \mathbf{S}_e^{-1} (\mathbf{y} - \mathbf{F}(\mathbf{x}_i)) - \mathbf{S}_a^{-1} (\mathbf{x}_i - \mathbf{x}_a) \right), \quad (4)$$

where  $\gamma$  is a scaling factor that can be set to 0 (Gauss–Newton method) or whose value can be adapted along iterations in order to prevent divergence (Levenberg–Marquardt method, in which successful retrievals use decreasing  $\gamma$  values and, eventually, 0 for the final iteration).  $\mathbf{K}_i$  here denotes the forward radiative transfer Jacobian matrix, whose values are evaluated for the state vector  $\mathbf{x}_i$ .

A successful retrieval reduces the a priori uncertainty of the state vector described in  $\mathbf{S}_a$ . The a posteriori covariance matrix of the retrieved state vector  $\hat{\mathbf{S}}$ , whose diagonal elements give the posterior variance of the retrieved state vector elements, is expressed as follows:

$$\hat{\mathbf{S}} = \left[ \mathbf{S}_a^{-1} + \mathbf{K}^T \mathbf{S}_e^{-1} \mathbf{K} \right]^{-1}. \quad (5)$$

Finally, the sensitivity of the retrieval with respect to the true geophysical state  $\mathbf{x}_{\text{true}}$  is given by the averaging kernel matrix  $\mathbf{A}$  calculated according to the following:

$$\mathbf{A} = \frac{\partial \hat{\mathbf{x}}}{\partial \mathbf{x}_{\text{true}}} (\mathbf{x}_a) = \left[ \mathbf{S}_a^{-1} + \mathbf{K}^T \mathbf{S}_e^{-1} \mathbf{K} \right]^{-1} \mathbf{K}^T \mathbf{S}_e^{-1} \mathbf{K}. \quad (6)$$

In most cases, the CO<sub>2</sub> concentration is included in the state vector as a level or layer profile from which X<sub>CO<sub>2</sub></sub>, the retrieved column-averaged dry air mole fraction of CO<sub>2</sub>, is computed (e.g. O’Dell et al., 2012). If we note  $\hat{\mathbf{x}}_{\text{CO}_2}$ , the part of the retrieved state vector  $\hat{\mathbf{x}}$  containing the CO<sub>2</sub> profile, and  $\mathbf{A}_{\text{CO}_2}$  and  $\hat{\mathbf{S}}_{\text{CO}_2}$ , the corresponding square parts of  $\mathbf{A}$  and  $\hat{\mathbf{S}}$ , we have the following:

$$X_{\text{CO}_2} = \mathbf{h} \hat{\mathbf{x}}_{\text{CO}_2}, \quad (7)$$

$$\sigma_{X_{\text{CO}_2}} = \sqrt{\mathbf{h}^T \hat{\mathbf{S}}_{\text{CO}_2} \mathbf{h}}, \quad (8)$$

$$(\mathbf{a}_{\text{CO}_2})_j = \frac{\partial X_{\text{CO}_2}}{\partial \mathbf{x}_{\text{true}}} = \frac{(\mathbf{h}^T \mathbf{A}_{\text{CO}_2})_j}{h_j}, \quad (9)$$

where  $\mathbf{h}$  is the pressure weighting function.  $\sigma_{X_{\text{CO}_2}}$  denotes the posterior uncertainty of the retrieved X<sub>CO<sub>2</sub></sub>, and  $\mathbf{a}_{\text{CO}_2}$  is the

CO<sub>2</sub> column averaging kernel. This profile vector describes the vertical sensitivity of the retrieved column with regard to the true profile; it is essential to characterise retrieval results and to compare them to other products, as shown in Sect. 4.2.

## 2.2.2 5AI features and retrieval scheme setups for OCO-2

The 5AI retrieval scheme enables the retrieval of multiple geophysical variables from hyperspectral measurements. Those currently include trace gas concentration represented in the state vector as a concentration profile or a profile scaling factor, global temperature profile offset, surface temperature and pressure band-wise albedo, whose spectral dependence is modelled as a polynomial, and, finally, scattering particle layer-wise optical depth. For this work, the state vector includes the main geophysical parameters necessary to retrieve  $X_{\text{CO}_2}$  and is described in Table 1. The a priori values and their covariance are identical to those used in the ACOS B8r version (O'Dell et al., 2018) in order to ease the retrieval result comparison.

The OCO-2 spectrometer measures Earth-reflected near- and shortwave infrared (NIR and SWIR) sunlight in three distinct bands, namely the O<sub>2</sub> A band (0.7  $\mu\text{m}$ ), the weak CO<sub>2</sub> band (1.6  $\mu\text{m}$ ), and the strong CO<sub>2</sub> band (2.0  $\mu\text{m}$ ). In order to accurately model OCO-2 measurements, polarisation effects have to be taken into account. As 4A/OP coupling with (V)LIDORT is not optimal yet, forward calculations can reach unmanageable durations without some assumptions that allow faster radiative transfer simulations. Therefore, as explained in Sect. 1, we first restrict this work to the lowest scattering particle content possible (while compromising with coverage) so that only Rayleigh scattering needs to be taken into account in the O<sub>2</sub> A band (0.7  $\mu\text{m}$ ). This is done by using 4A/OP coupling with VLIDORT, and the ACOS Stokes coefficients are applied to yield the final scalar radiances. For CO<sub>2</sub> weak and strong bands, scattering and polarisation can be neglected with this low scattering particle content assumption, and only the Stokes coefficient, 0.5, for the  $I$  component of the electric field is applied to yield the final scalar radiances. As we neglect, for computation time purposes, the possible impact of scattering particles in forward calculations and in the state vector, the retrieval problem becomes more linear. Thus, we can also assume a slow variation in the Jacobian matrix along the retrieval iterations and, therefore, choose not to update it in order to save computational time. Hence, the partial derivatives of the radiative transfer model are evaluated once and for all around the a priori state. We performed a sensitivity test and assessed that this approximation does not significantly change the retrieval results (not shown).

## 3 Data

### 3.1 Data description

The OCO-2 satellite has three distinct observation modes. The nadir and glint modes are the nominal science observation modes; they constitute the vast majority of OCO-2 measurements. In addition, the target mode of the OCO-2 mission provides data for the validation of the retrievals. In target mode, the satellite tilts and aims at a validation target (most of them are TCCON stations) and scans its whereabouts several times during the overpass. These sessions thus provide OCO-2 data points closely collocated with validation targets (over areas that can be as small as 0.2° longitude  $\times$  0.2° latitude) and registered over a few minutes (Wunch et al., 2017).

OCO-2 high-resolution spectra are analysed by the ACOS team in order to retrieve  $X_{\text{CO}_2}$  and other geophysical parameters from them. The ACOS team provides two different  $X_{\text{CO}_2}$  values, i.e. raw and posterior bias-corrected  $X_{\text{CO}_2}$ . Raw  $X_{\text{CO}_2}$  is the direct output of the ACOS algorithm following the full physics retrieval; B8 retrospective (B8r) ACOS data release is used here (O'Dell et al., 2018). Posterior bias-corrected  $X_{\text{CO}_2}$  is an empirically corrected  $X_{\text{CO}_2}$  that has reduced averaged bias with regard to different truth proxies (O'Dell et al., 2018). In this work, 5AI results are compared with raw  $X_{\text{CO}_2}$  as we do not perform any empirical bias correction.

In addition, we compare  $X_{\text{CO}_2}$  retrieved from OCO-2 spectra to TCCON data. The TCCON network uses ground-based high-resolution Fourier transform spectrometers to measure NIR and SWIR spectra that enable the retrieval of the column-averaged dry air mole fractions of greenhouse gases. These retrievals are performed by GGG2014 software (Wunch et al., 2015), and their results are available on the TCCON Data Archive (<https://tccondata.org/>, last access: 25 May 2021).

### 3.2 Data selection

We intend to compare 5AI results with regard to TCCON against ACOS results for corresponding sets of soundings. First, we select all the OCO-2 target soundings between 2015 and 2018 with the best ACOS cloud, sounding quality, and outcome flags values. As a compromise between scattering particle content and coverage, we set an upper limit of 0.5 for the ACOS retrieved total aerosol optical depth (AOD). This sample set of OCO-2 target soundings includes 16 414 soundings, with a median ACOS retrieved total AOD of 0.05 and a 75 % percentile of 0.1.

For this study, we select the TCCON official products measured  $\pm 2$  h compared to OCO-2 overpass time and only keep the target sessions where there are at least five OCO-2 measurements passing 5AI posterior filters and five TCCON data points are available. This set includes 9449 TCCON individual retrieval results from 19 TCCON stations listed in Table 2.

**Table 1.** 5AI state vector composition for OCO-2 retrievals.

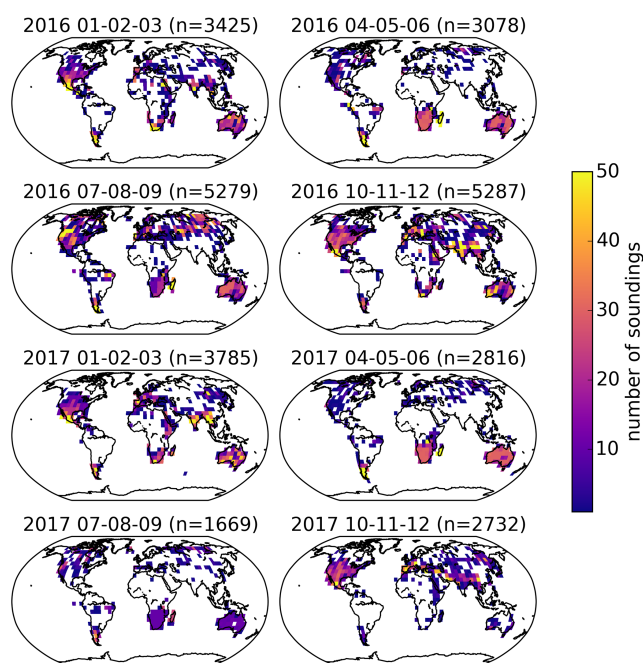
Variable name	Length	A priori value	A priori uncertainty ( $1\sigma$ )	Notes
H <sub>2</sub> O scaling factor	1	1.0	0.5 (same as ACOS)	–
CO <sub>2</sub> layer concentration	19 layers	ACOS a priori	ACOS prior covariance matrix	See prior covariance matrix in O'Dell et al. (2012)
Surface pressure	1	ACOS a priori	4.0 hPa (same as ACOS)	–
Temperature profile offset	1	ACOS a priori	5.0 K (same as ACOS)	–
Surface albedo (order 0 of albedo model)	3 bands	ACOS a priori	1.0 (same as ACOS)	Evaluated at 0.77, 1.615 and 2.06 $\mu\text{m}$ for O <sub>2</sub> and CO <sub>2</sub> weak and strong bands, respectively
Surface albedo slope (order 1 of albedo model)	3 bands	0.0	1.0/cm <sup>-1</sup>	–

**Table 2.** TCCON data used in this work.

TCCON station	Coordinates (latitude, longitude, and altitude)	No. of target sessions	Date range (first and last sessions)	Reference
Ascension Island	7.92° S, 14.33° W, 0.01 km	4	16 January 2015–15 January 2018	Feist et al. (2014)
Bialystok (Poland)	53.23° N, 23.03° E, 0.18 km	1	18 March 2015	Deutscher et al. (2019)
Bremen (Germany)	53.10° N, 8.8° E, 0.027 km	1	17 March 2016	Notholt et al. (2014)
Caltech (USA)	34.14° N, 118.13° W, 0.230 km	19	12 September 2014–16 September 2018	Wennberg et al. (2015)
Darwin (Australia)	12.424° S, 130.89° E, 0.03 km	8	15 May 2015–28 July 2017	Griffith et al. (2014a)
Edwards (USA)	34.96° N, 117.88° W, 0.700 km	1	22 August 2018	Iraci et al. (2016)
Eureka (Canada)	80.05° N, 86.42° W, 0.61 km	2	16 June–28 June 2015	Strong et al. (2019)
Izaña (Tenerife)	28.31° N, 16.50° W, 2.37 km	2	5 January–24 March 2018	Blumenstock et al. (2017)
Karlsruhe (Germany)	49.10° N, 8.44° E, 0.116 km	1	12 September 2016	Hase et al. (2015)
Lamont (USA)	36.60° N, 97.49° W, 0.32 km	9	10 February 2015–11 November 2016	Wennberg et al. (2016)
Lauder (New Zealand)	45.04° S, 169.68° E, 0.37 km	3	17 February 2015–6 February 2016	Sherlock et al. (2014)
Orléans (France)	47.97° N, 2.11° E, 0.13 km	1	8 April 2015	Warneke et al. (2019)
Paris (France)	48.85° N, 2.36° E, 0.06 km	1	25 August 2016	Té et al. (2014)
Park Falls (USA)	45.95° N, 90.27° W, 0.44 km	7	11 October 2014–21 April 2017	Wennberg et al. (2017)
Réunion Island	20.90° S, 55.49° E, 0.087 km	4	24 March–1 August 2015	De Mazière et al. (2017)
Saga (Japan)	33.24° N, 130.29° E, 0.007 km	5	31 July 2015–2 December 2017	Kawakami et al. (2014)
Sodankylä (Finland)	67.37° N, 26.63° E, 0.188 km	4	20 August 2015–17 July 2018	Kivi et al. (2014), Kivi and Heikkinen (2016)
Tsukuba (Japan)	36.05° N, 140.12° E, 0.03 km	6	14 November 2014–17 June 2017	Morino et al. (2018)
Wollongong (Australia)	34.40° S, 150.88° E, 0.03 km	13	23 September 2014–6 May 2018	Griffith et al. (2014b)

Besides target data, we also select a sample of OCO-2 nadir land soundings with a coverage as global as possible over the years 2016–2017 (all ACOS flags at their best value possible). For every month and 5° longitude  $\times$  5° latitude bins, we select 25 (10 for North America, southern

Africa, and Australia) soundings with low ACOS retrieved total AOD. For 2016 and 2017, this selection is done for a maximum ACOS retrieved total AOD of 0.035 and 0.045, respectively, yielding 17 069 soundings for 2016 and 11 002



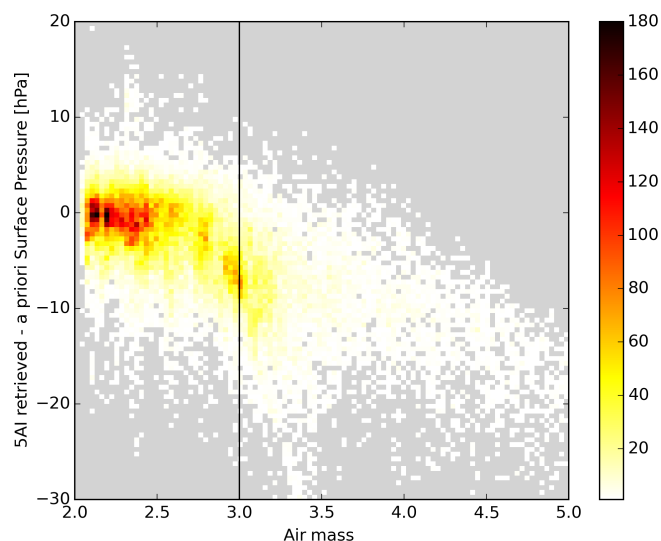
**Figure 1.** Spatial and temporal repartition of the sample of nadir OCO-2 soundings selected for 5AI retrievals, in seasonal and  $5^\circ \times 5^\circ$  square bins. The titles include  $n$ , the number of soundings for the corresponding panel. The low number of selected soundings in July–September 2017 is due to an identified OCO-2 data gap.

for 2017. Figure 1 shows the spatial and temporal distribution of these OCO-2 points.

## 4 Results

### 4.1 Post-filtering of retrieval results

We apply the a posteriori filters described in Table 3 to ensure the quality of the retrieved results. The surface pressure filter removes soundings for which it proved difficult to successfully model the optical path, suggesting scattering related errors leading to a large difference between the retrieved and prior surface pressure. The reduced  $\chi^2$  filter removes the worst spectral fits. In the end, 95 % of our selected soundings pass these first two filters. In addition, the blended albedo filter removes the 12 % fraction of target data representative of challenging snow or ice-covered surfaces (Wunch et al., 2011a). With the current retrieval setup, the difference between the 5AI retrieved surface pressure and its prior exhibit an air mass dependence, as shown in Fig. 2. For this present work, we filter out all soundings with an air mass above 3.0. Future studies will refine the 5AI forward and inverse setup in order to process hyperspectral infrared soundings with larger air masses. Results detailed in the following subsections are based on the 9605 target and 21 254 nadir OCO-2 soundings that passed all these filters.



**Figure 2.** Distribution of target and nadir 5AI retrievals passing surface pressure, blended albedo, and reduced  $\chi_r^2$  filters according to air mass and difference between retrieved and prior surface pressures. Grey areas denote bins for which no 5AI retrieval is available.

### 4.2 OCO-2 target retrieval results

For every target session, we consider a unique average of the available retrieval results from OCO-2 measurements and a unique average of the corresponding TCCON official products as performed in, for example, O’Dell et al. (2018) and Wu et al. (2018). As OCO-2 and TCCON  $X_{\text{CO}_2}$  vertical sensitivities described by their averaging kernels are not exactly identical, we take into account the averaging kernel correction of TCCON data, as performed by the ACOS team (O’Dell et al., 2018) and described by Eq. (10) (Nguyen et al., 2014) as follows:

$$X_{\text{OCO-2, TCCON}} = X_{\text{a priori}} + \left( \frac{\hat{X}_{\text{TCCON}}}{X_{\text{a priori}}} - 1 \right) \sum_j h_j(\text{aCO}_2)_j \mathbf{x}_{\text{a priori}, j}. \quad (10)$$

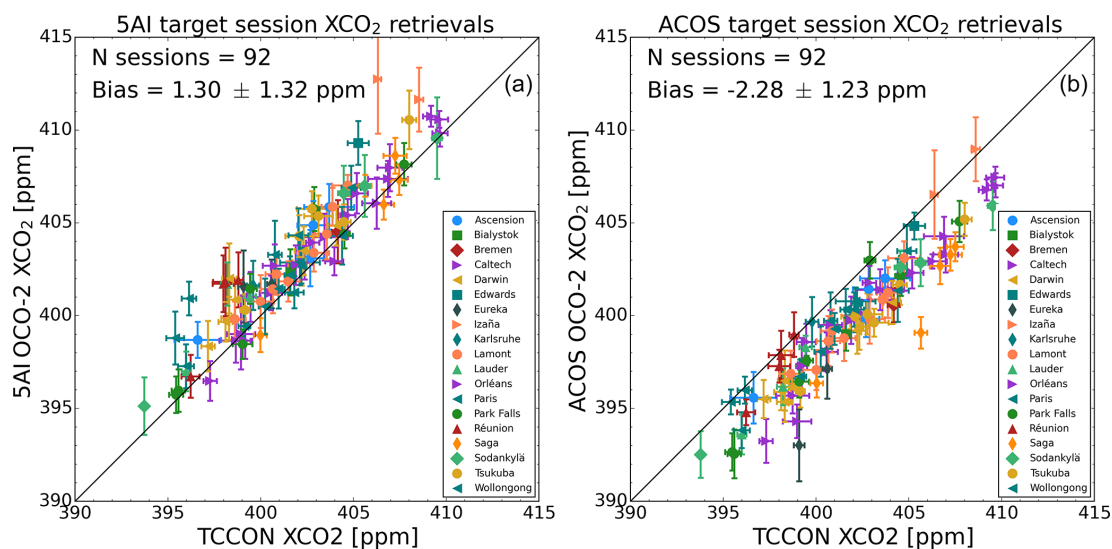
$X_{\text{OCO-2, TCCON}}$  is the column-averaged dry air mole fraction of  $\text{CO}_2$  that would have been retrieved from the OCO-2 measurement if the collocated TCCON retrieval was the true state of the atmosphere.  $X_{\text{a priori}}$  is the a priori column-averaged dry air mole fraction of  $\text{CO}_2$ , considered to be very similar between 5AI (or ACOS) and GGG2014.  $\hat{X}_{\text{TCCON}}$  is the TCCON retrieved column-averaged dry air mole fraction of  $\text{CO}_2$ .  $h$  is the pressure weighting function vector defined previously.  $(\text{aCO}_2)_j$  is the  $\text{CO}_2$  column averaging kernel vector defined in Eq. (9), and  $\mathbf{x}_{\text{a priori}}$  is the a priori  $\text{CO}_2$  concentration profile vector. The effect of this correction yields a positive shift of the bias with regard to TCCON of about 0.2 ppm for the set of target sessions considered in this work.

Following post-filtering, Fig. 3 shows 5AI raw results compared to the TCCON official product over 92 target ses-



**Table 3.** Filters applied on 5AI retrieval results for this work.

Variable name	Minimum value	Maximum value	Definition and reference	OCO-2 mode
Retrieved surface pressure	$P_{\text{nlev}-1}$	–	The atmosphere is discretised in 20 levels bounding 19 layers. We do not allow the surface pressure, $P_{\text{nlev}}$ , to be lower than its preceding pressure level.	Nadir and target
Reduced $\chi^2$	–	7.0	Overall goodness of the spectral fit (e.g. Wu et al., 2018).	Nadir and target
Blended albedo	–	0.8	$2.4 \times \text{O}_2$ A-band albedo + $1.13 \times \text{CO}_2$ strong band albedo (Wunch et al., 2011a, 2017)	Target
Air mass	–	3.0	$\frac{1}{\cos(\text{SZA})} + \frac{1}{\cos(\text{VZA})}$ , with SZA, the solar zenith angle, and VZA, the viewing zenith angle (Wunch et al., 2011a)	Nadir and target

**Figure 3.** The 5AI (a) and raw ACOS B8r (b) OCO-2 target  $X_{\text{CO}_2}$  retrieval results compared to the TCCON official  $X_{\text{CO}_2}$  product. Individual sounding results are averaged for every target session. Markers show the session average for OCO-2 and TCCON  $X_{\text{CO}_2}$ , and error bars show the standard deviations.

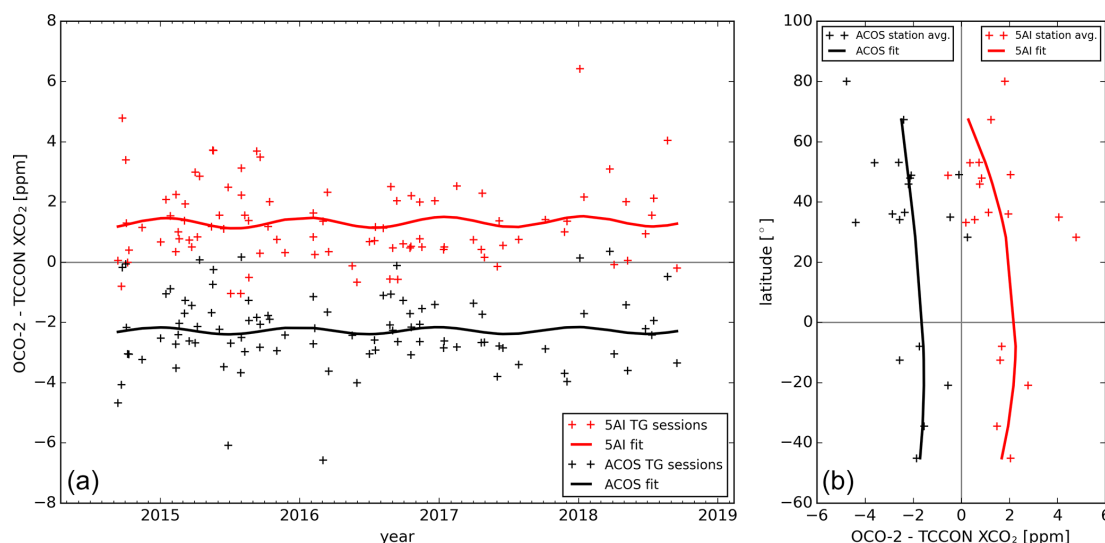
sions. The mean systematic  $X_{\text{CO}_2}$  bias (5AI – TCCON) is 1.30 ppm, and its standard deviation is 1.32 ppm. The ACOS raw  $X_{\text{CO}_2}$  and TCCON  $X_{\text{CO}_2}$  comparison for the corresponding set of OCO-2 soundings is also presented in Fig. 3. The bias compared to TCCON is  $-2.28$  ppm, and its standard deviation is 1.23 ppm. This difference in bias compared to TCCON may be greatly influenced by forward modelling and retrieval setup differences between 5AI and ACOS, as detailed later in this work. Bias-corrected FOCAL and RemoTeC  $X_{\text{CO}_2}$  retrieval results compared to the ACOS official product exhibit similar differences in bias standard deviations (Reuter et al., 2017b; Wu et al., 2018).

Temporal and latitudinal fits of 5AI and ACOS  $X_{\text{CO}_2}$  biases, compared to TCCON, are displayed in Fig. 4. Temporal biases are fitted with a first-order polynomial added to a cosine and exhibit a quasi-null slope with a  $\sim 0.4$  ppm am-

plitude of yearly oscillation in both 5AI and ACOS cases. Latitudinal bias fits performed with all the available target sessions, except those from Eureka, show that 5AI bias compared to TCCON appears to be larger in the Southern Hemisphere than in the Northern Hemisphere, but its behaviour is quite parallel to ACOS, except at higher latitudes where 5AI and ACOS become closer. The Eureka station (latitude  $80^\circ$  N) has been removed from those fits, as satellite retrievals and validation are known to be challenging at these latitudes (O'Dell et al., 2018).

### 4.3 OCO-2 nadir retrieval results

In this subsection, raw 5AI retrieved  $X_{\text{CO}_2}$  is compared to the ACOS raw product on a sample of OCO-2 nadir soundings, as described in Sect. 3 and displayed in Fig. 1. The nadir viewing configuration is the nominal science mode of the

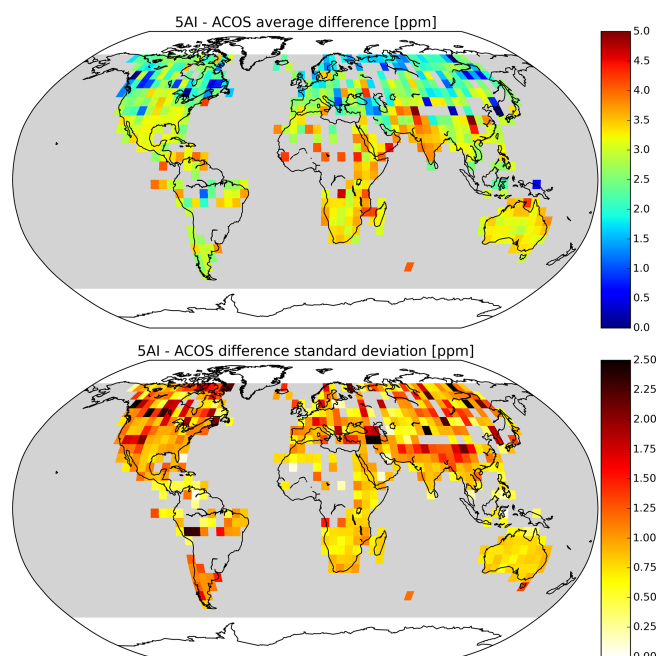


**Figure 4.** The 5AI and raw ACOS B8r OCO-2 target  $X_{CO_2}$  bias compared to TCCON as a function of time (a) and latitude (b). Crosses show the individual session averages in panel (a) and individual station averages in panel (b). The full lines show the polynomial fits of this bias for all target sessions.

OCO-2 mission and allows comparisons at a larger spatial scale than the one offered by the target mode dedicated to validation.

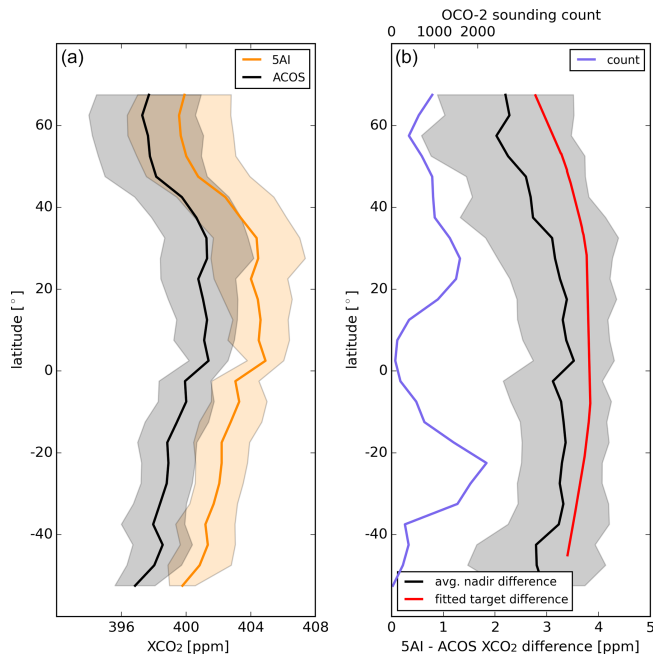
Figure 5 shows the average and associated standard deviation of the difference between 5AI- and ACOS-retrieved raw  $X_{CO_2}$ . The overall 5AI–ACOS difference is about 3 ppm, with a latitudinal dependency; it is lower above midlatitudes in the Northern Hemisphere. 5AI differences to ACOS also exhibit features over India or the Sahara that are places which are often associated with strong aerosol events; those may be due to the neglecting of scattering parameters in the 5AI retrievals. The standard deviation is mainly correlated with topography; it is higher in the vicinity of mountain chains and lower in flatter areas. As we do not take into account topography in the sampling strategy of the processed OCO-2 nadir soundings, its greater variability in mountainous areas can result in a greater variability in the retrieved surface pressure, which is strongly correlated with retrieved  $X_{CO_2}$ . As for the highest standard deviations in South America, they may be caused by the South Atlantic Anomaly which they are close to (Crisp et al., 2017).

As seen in Fig. 6, latitudinal variations in raw 5AI retrieved  $X_{CO_2}$  are consistent with those of ACOS, with a difference between the two products being almost constant, except above midlatitudes in the Northern Hemisphere where the differences are smaller. In addition, the comparison between 5AI and ACOS in the nadir mode is consistent with the results obtained for target sessions. Indeed, the raw 5AI–ACOS target difference lies within  $\pm 1\sigma$  of nadir results, with  $\sigma$  being the standard deviation of the 5AI–ACOS difference. Figure 7 details the temporal variations in the retrieved  $X_{CO_2}$ . The global long-term increase in the atmospheric concentra-

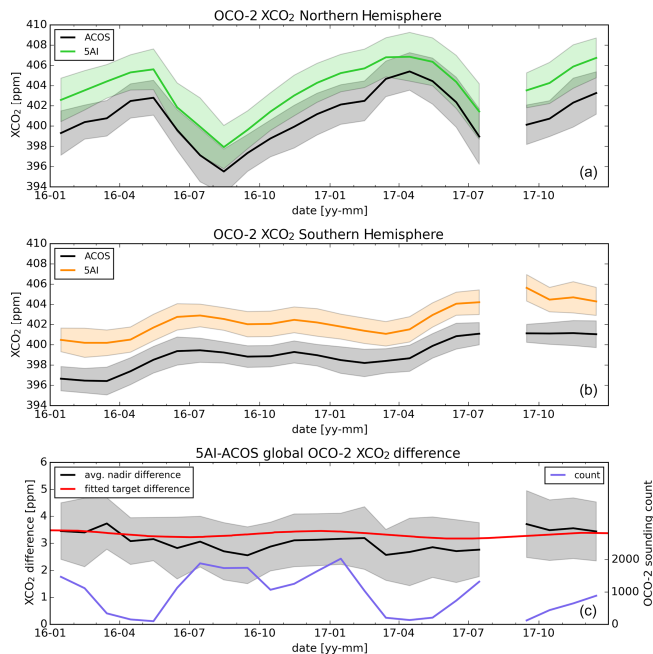


**Figure 5.** Spatial repartition of 5AI difference with the raw ACOS B8r (average and standard deviation) on  $5^\circ \times 5^\circ$  square bins for the nadir data selection.

tion of  $CO_2$  can be observed in both hemispheres and the seasonal cycle, which is stronger in the Northern Hemisphere, where most of the vegetation respiration and photosynthesis happen. The temporal variations in the 5AI–ACOS  $X_{CO_2}$  retrieval differences in the nadir mode are also consistent with those presented in the target mode.



**Figure 6.** Latitudinal variation in 5AI and the raw ACOS B8r retrieved  $X_{\text{CO}_2}$  (a) and their difference (b). Panel (b) compares the 5AI–ACOS average difference for nadir soundings and 5AI–ACOS difference fitted on target sessions (bottom axis). The number of available nadir soundings is also shown in panel (b) (top axis).



**Figure 7.** Temporal variation in 5AI and the raw ACOS B8r retrieved  $X_{\text{CO}_2}$  in the Northern Hemisphere (a), Southern Hemisphere (b), and the global difference (c). Panel (c) compares 5AI–ACOS difference for nadir and target OCO-2.

## 5 Discussion

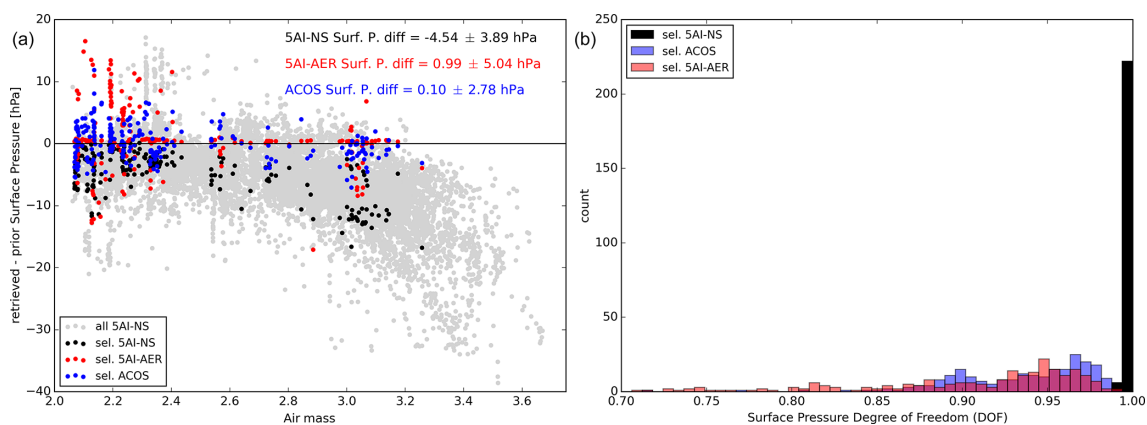
### 5.1 Sensitivity of raw retrieval results to scattering particles

Of the main forward and inverse differences between 5AI and ACOS, one is the accounting of scattering particles on the optical path. Indeed, ACOS considers five Gaussian-shaped vertical profiles of different scattering particle types for which it retrieves three parameters (O’Dell et al., 2018), while, for computational time purposes, none is considered in the previously presented 5AI results (hereafter denoted 5AI-NS for no scattering). In order to assess the sensitivity of 5AI results to this neglecting of scattering particles, we propose performing some 5AI  $X_{\text{CO}_2}$  retrievals from OCO-2 soundings while taking into account some aerosol parameters in the forward modelling and state vector.

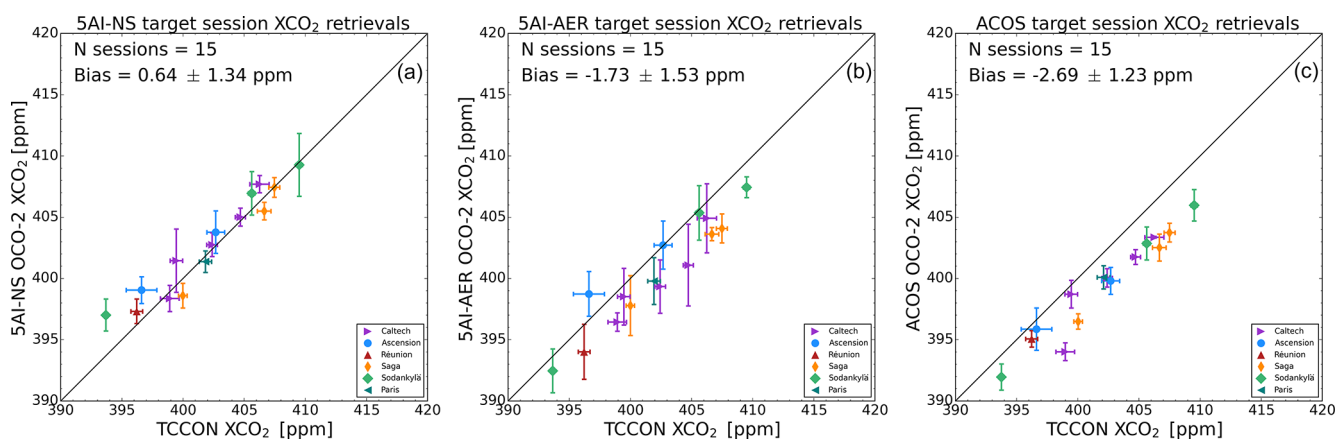
Several adaptations of the 5AI setup are required for this sensitivity test (hereafter denoted 5AI-AER for aerosols). First, we consider here two fixed-height fixed-width aerosol layers. The first one, representative of coarse mode minerals, is located between about 800 and 900 hPa, and the second, representative of fine mode soot, is between about 900 and 1013 hPa. Only the two layer-wise optical depths are retrieved (defined at 755 nm, as ACOS), each with an a priori value of 0.025 and an a priori uncertainty of 0.15. Otherwise, the state vector and its a priori, described in Table 1, remain unchanged. Regarding forward modelling, we still rely on 4A/OP coupling with VLIDORT for the  $\text{O}_2$  A-band calculations, and we now use 4A/OP coupling with LIDORT for  $\text{CO}_2$  weak and strong band calculations (thus still neglecting polarisation effects in these bands). Finally, as the retrieval problem becomes less linear when considering scattering particle parameters, we update the Jacobian matrix with every iteration.

With these adaptations, 5AI retrievals are about 12 times slower than when not accounting for scattering particles. Considering the increase in computation time, this sensitivity test can only be performed for a small subsample of the data. We choose to focus here on 15 OCO-2 target sessions (out of the 92 presented in Sect. 4) that have available AERONET (Aerosol RObotic NETwork, version 3; AOD level 2.0) optical depths acquired  $\pm 2$  h compared to OCO-2 overpass (Holben et al., 1998; Eck et al., 1999; Giles et al., 2019), thus enabling us to also discuss total retrieved aerosol optical depths. A total of 445 OCO-2 soundings have been processed, and 228 remain after filtering according to the quality of the spectral fit (reduced  $\chi^2 < 7.0$ ).

Figure 8 shows how taking into account scattering particles in the state vector impacts the retrieval of surface pressure. The air mass dependence, exhibited in 5AI-NS results and shown in Fig. 2, appears to be reduced or even removed for the 5AI-AER results. Indeed, neglecting scattering particles results in neglecting the backscattered photons that leads to forward a priori synthetic measurements being less intense



**Figure 8.** Retrieved surface pressure air mass dependence (a) for all 5AI-NS target OCO-2 soundings (light grey), 5AI-NS soundings selected in the small subsample that passed all filters (black), and the corresponding 5AI-AER (red) and ACOS (blue). Distributions of surface pressure degrees of freedom are shown (b) for 5AI-NS (black), 5AI-AER (red) and ACOS (blue).

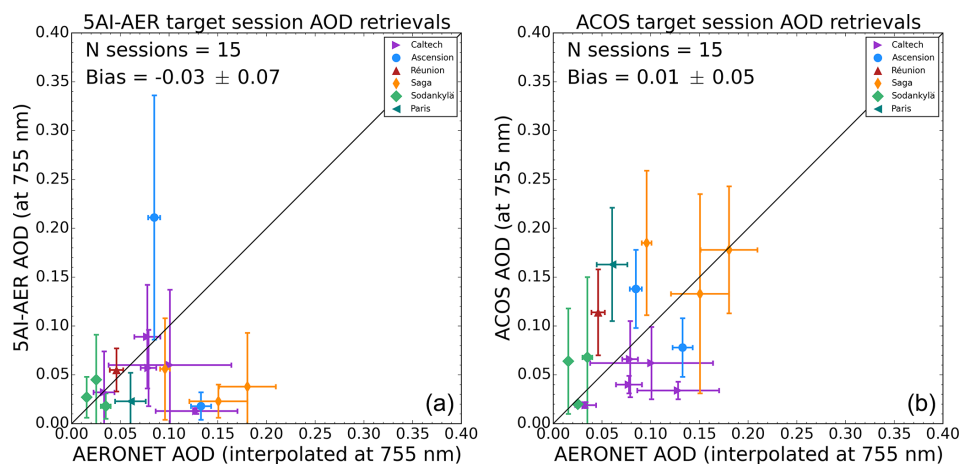


**Figure 9.** The 5AI-NS (a), 5AI-AER (b), and raw ACOS B8r (c) OCO-2 target  $X_{\text{CO}_2}$  retrieval results compared to the TCCON official  $X_{\text{CO}_2}$  product. Individual sounding results are averaged for every target session. Markers show the session average for OCO-2 and TCCON  $X_{\text{CO}_2}$ , and error bars show the standard deviations.

than those actually measured. This difference is seen by the retrieval scheme as an a priori overestimation of the amount of O<sub>2</sub> along the optical path; thus, it is seen as an overestimation of surface pressure which is then reduced. Hence, the  $-5$  hPa surface pressure average bias of 5AI-NS results, compared to the a priori surface pressure in Fig. 8, as opposed to the 1 hPa bias obtained with 5AI-AER for this small subsample of OCO-2 target soundings. Besides, the fraction of measured backscattered photons increases with air mass, leading to the air mass dependence of 5AI-NS results, as shown in Fig. 8. Furthermore, adding scattering particle parameters in the retrieval state vector interferes with surface pressure retrieval as the scattering particle and surface pressure information carried by the O<sub>2</sub> A band is entangled. As it can be seen in Fig. 8b, this leads to lower degrees of freedom for surface pressure compared to retrievals performed without scattering particle parameters in the state vector. 5AI-AER surface pressure degrees of freedom have a distribution

that is more similar to ACOS's than to the 5AI-NS surface pressure degrees of freedom. When scattering particle parameters are included in the state vector, this consequently leads to a stronger pull-back of the retrieved surface pressure towards the a priori value, also helping to reduce or even remove the air mass dependence for surface pressure. Thus, for reasons related to both radiative transfer and the retrieval methodology, taking into account scattering particles modifies the average difference between retrieved and a priori surface pressure and helps to remove the air mass dependence seen in 5AI-NS results.

Figure 9 shows  $X_{\text{CO}_2}$  retrieved from OCO-2 measurements for these 15 target sessions by the initial 5AI-NS setup (Fig. 9a), this adapted 5AI-AER setup (Fig. 9b) and ACOS in the B8r raw results (Fig. 9c). The impact of taking into account scattering particles in the retrievals directly translates from surface pressure to  $X_{\text{CO}_2}$ ; it appears that the difference of about 3 ppm exhibited in 5AI-NS results compared



**Figure 10.** The 5AI-AER (a) and ACOS B8r (b) OCO-2 target total AOD retrieval results compared to AERONET AOD interpolated at 755 nm. Individual sounding results are averaged for every target session. Markers show the session average for OCO-2 and AERONET AOD, and error bars show the standard deviations.

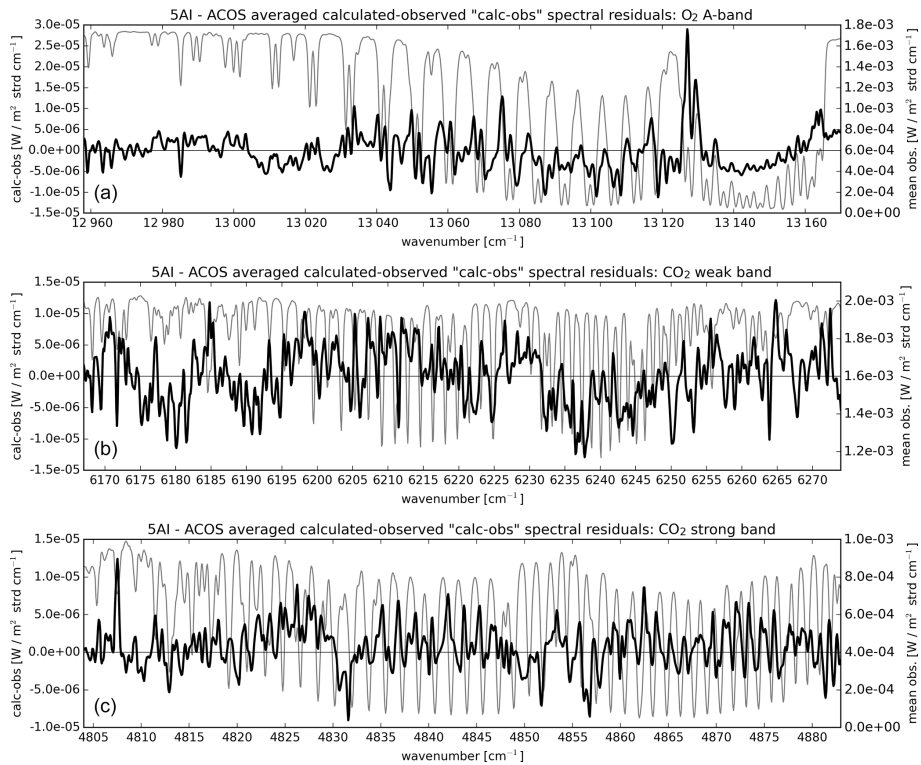
to ACOS is reduced to a difference of close to 1 ppm in 5AI-AER results. This shows that taking scattering particle parameters into account can indeed explain much of the differences between 5AI-NS results and ACOS. Regarding the retrieved optical depths, Fig. 10 shows 5AI-AER and ACOS retrieved total AOD compared to AERONET reference data interpolated at 755 nm. 5AI-AER exhibits a higher average difference to AERONET than ACOS, but both retrieval algorithms exhibit a considerable scatter of their results compared to AERONET. Efforts to optimise 4A/OP coupling with (V)LIDORT are underway so that more OCO-2 data can be processed. Once those are completed, a dedicated study will help to tune the 5AI scattering particle setup better (varying aerosol types, impact of cirrus clouds, varying layer altitudes, etc.).

## 5.2 Sensitivity of raw retrieval results to inverse and forward modelling

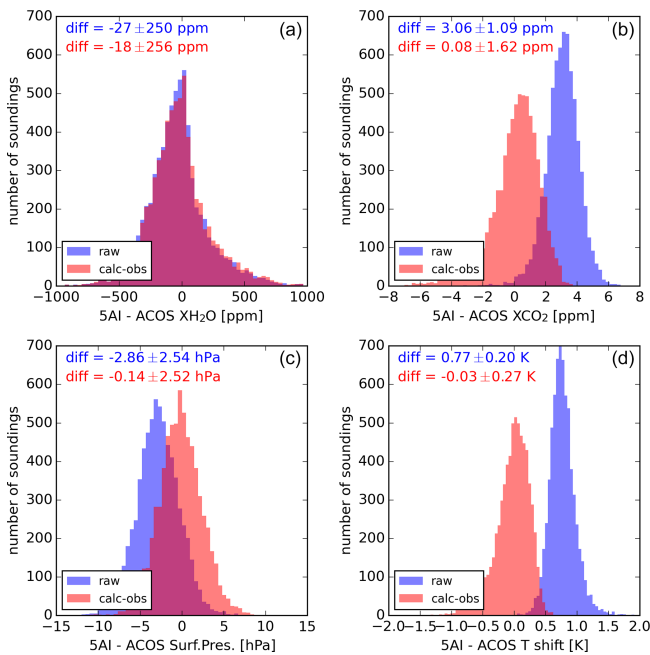
A difference of about 3 ppm is found between 5AI and ACOS raw  $X_{\text{CO}_2}$  retrieved from OCO-2 for both nadir and target observations. In Sect. 5.1 we show that neglecting scattering particles for computational time purposes can explain most of this difference. However, the 5AI-AER retrieval setup does not exactly reproduce ACOS setup, as state vector, forward radiative transfer, and spectroscopic parameter differences remain. All those can be encompassed and accounted for by using an average calculated–observed spectral residual analysis (hereafter calc–obs). It consists in calculating a spectrum (convolved to OCO-2 instrument line shape) based on the ACOS retrieval results (posterior pressure grid, temperature,  $\text{H}_2\text{O}$ , and  $\text{CO}_2$  profiles as well as albedo and albedo slope) and comparing it to the corresponding OCO-2 observation. Possible background differences are also compensated by scaling the OCO-2 spectrum so that its transparent

spectral windows fit those of the calculated 4A/OP spectrum. Such comparisons must be performed and averaged over a spatially and temporally unbiased data set with a homogeneous viewing geometry in order to cancel out possible dependences. Thus, it is performed here for a randomly chosen half of the nadir OCO-2 points with an air mass below 3.0 selected in 2016 (6790 in total). Figure 11 shows the resulting averaged calc–obs spectral residuals and the corresponding average OCO-2 measurement. Differences are principally located in the  $0.7\ \mu\text{m}$   $\text{O}_2$  absorption band but also in the  $1.6$  and  $2.0\ \mu\text{m}$   $\text{CO}_2$  absorption bands. They are due to the differences in inverse setup and in radiative transfer models between ACOS and 5AI (impact of aerosols, parameterisation of continua, spectroscopy, etc.).

In order to compare 5AI retrievals with ACOS products while attenuating the impact of the forward and inverse modelling differences, the obtained averaged calc–obs residual is added to every OCO-2 measurements within the complementary half of the selected nadir soundings from 2016 (6799 in total). We then apply the 5AI inverse scheme on this new data set. Fig. 12 compares the distributions of 5AI–ACOS retrieval results obtained with and without the calc–obs adjustment. The systematic differences between 5AI and ACOS results for  $X_{\text{H}_2\text{O}}$ ,  $X_{\text{CO}_2}$ , surface pressure, and global temperature profile shift are fully removed when adding the spectral residual adjustment to OCO-2 measurements (remaining differences are negligible compared to standard deviations). This  $X_{\text{CO}_2}$  shows that 5AI can, on average, reproduce ACOS results when all their respective differences are compensated with a calc–obs adjustment. However, it impacts the standard deviations of 5AI–ACOS differences. Indeed, only ACOS raw results that relate to the 5AI state vector parameters have been used to compute the calculated spectrum used in this calc–obs analysis, and other ACOS parameters, such as scattering particles for instance, have not been considered.



**Figure 11.** 5AI-ACOS average calc-obs spectral residuals in the O<sub>2</sub> A band (a), CO<sub>2</sub> weak band (b), and CO<sub>2</sub> strong band (c) appear in thick black lines (left axis). A typical spectrum for the three bands is shown in thin grey lines (right axis).



**Figure 12.** 5AI raw ACOS B8r difference distributions for X<sub>H<sub>2</sub>O</sub> (a), X<sub>CO<sub>2</sub></sub> (b), surface pressure (c), and temperature profile global shift (d) shown without applying the average calc-obs spectral residual correction (in blue) and with the correction (in red).

Their impact may be attenuated by the background difference correction, which, if disabled, leads to a similar standard deviation of 5AI-ACOS differences in both with and without calc-obs cases. However, without the background compensation, the average difference between 5AI-ACOS is only reduced to 1.9 ppm for X<sub>CO<sub>2</sub></sub> (not shown). This exemplifies how highly challenging the sounding-to-sounding intercomparison of retrieval results remains and highlights how forward modelling and retrieval setup design impact X<sub>CO<sub>2</sub></sub> retrieval results.

## 6 Conclusions

In this work, we have introduced the 5AI inverse scheme; it implements the optimal estimation algorithm and uses the 4A/OP radiative transfer model with the GEISA spectroscopic database and an empirically corrected absorption continuum in the O<sub>2</sub> A band. We have applied the 5AI inverse scheme to retrieve X<sub>CO<sub>2</sub></sub> from a sample of ~44k OCO-2 soundings that compromises between coverage and the lowest ACOS-retrieved total AOD. We neglected the impact of scattering particles for computational time purposes and obtained a global averaged uncorrected bias compared to TC-CON of 1.30 ppm, with a standard deviation of 1.32 ppm for air masses below 3.0. These results are comparable in standard deviation with those obtained by ACOS on the corre-

sponding set of OCO-2 soundings. Moreover, we showed that, similarly to ACOS, 5AI  $X_{\text{CO}_2}$  retrievals satisfactorily capture the global increasing trend of atmospheric  $\text{CO}_2$ , its seasonal cycle, and its latitudinal variations, and that 5AI results are consistent between OCO-2 nadir and target modes. Although 5AI exhibits a difference of about 3 ppm compared to ACOS, we showed that neglecting scattering particles can explain most of it. Indeed, 5AI–ACOS average difference is reduced to 1 ppm when accounting for the optical depths of two coarse- and fine-mode aerosol layers in 5AI state vector, respectively. This is in part due to how taking into account scattering particles impacts the retrieval of surface pressure, which becomes closer to ACOS. The air mass dependence of the 5AI retrieved surface pressure is also reduced. Finally, we showed that 5AI can on average reproduce ACOS results when adding to OCO-2 measurements an average calc–obs spectral residual correction. It encompasses all the inverse and forward differences between 5AI and ACOS, and, thus, underlies the critical sensitivity of retrieval results to the inverse setup design and forward modelling.

For favourable conditions (all the best ACOS flags and lowest ACOS retrieved total AOD possible), we showed that 5AI is a reliable implementation of the optimal estimation algorithm, whose results can be compared to ACOS raw products. Efforts are underway in order to optimise and increase the speed of 4A/OP coupling with (V)LIDORT. Finally, the implementation of the 5AI retrieval scheme is intended to be compatible with 4A/OP structure, so that the code can be easily adapted to any current or future greenhouse gas monitoring instrument, from TCCON, EM27/SUN (e.g. Gisi et al., 2012; Hase et al., 2016), and OCO-2 to MicroCarb or Copernicus  $\text{CO}_2$  Monitoring ( $\text{CO}_2\text{M}$ ; Meijer and Team, 2019), and can even be applied to research concepts such as the one proposed in the European Commission Horizon 2020 project of SCARBO (Space CARBON Observatory; Brooker, 2018).

**Data availability.** For this work, we use the B8r of OCO-2 data that were produced by the OCO-2 project at the Jet Propulsion Laboratory, California Institute of Technology, and obtained from the OCO-2 data archive maintained at the NASA Goddard Earth Science Data and Information Services Center (NASA GES-DISC). All TCCON references are given in Table 2. AERONET data are available from the AERONET website ([https://aeronet.gsfc.nasa.gov/new\\_web/download\\_all\\_v3\\_aod.html](https://aeronet.gsfc.nasa.gov/new_web/download_all_v3_aod.html), last access: 25 May 2021, Stutz et al., 2021). 5AI retrieval results presented in this work are available, upon emailed request, from Matthieu Dogniaux ([matthieu.dogniaux@lmd.ipsl.fr](mailto:matthieu.dogniaux@lmd.ipsl.fr)).

**Author contributions.** MD developed 5AI and tested it on the OCO-2 soundings under the supervision of CC, with input and support from CC, RA, VirC, TD, and VinC. MDM, NMD, DGF, OEG, DWTG, FH, LTI, RK, IM, JN, DFP, CMR, KS, KS, YT, VAV, and

TW provided the TCCON data. MD wrote this article, with feedback from all co-authors.

**Competing interests.** The authors declare that they have no conflict of interest.

**Acknowledgements.** The authors would like to thank the ACOS team and NASA for OCO-2 data availability. We thank the AERONET team for the whole database availability and especially Jochen Stutz, Brent Holben, Toshihiko Takemura, Rigel Kivi, Veijo Aaltonen, Francois Ravetta, and Jacques Pelon for their effort in establishing and maintaining Caltech, Ascension Island, Saint-Denis, Réunion, Fukuoka, Sodankylä, and Paris AERONET sites. Finally, we would like to thank the two anonymous referees whose comments and advice significantly improved the quality of this paper.

The TCCON site at Réunion is operated by the Royal Belgian Institute for Space Aeronomy, with financial support, since 2014, from the EU project ICOS-INWIRE and the ministerial decree for ICOS (grant nos. FR/35/IC1 to FR/35/C5) and local activities supported by LACy/UMR8105 at Université de La Réunion. The TCCON stations at Tsukuba and Burgos are supported, in part, by the GOSAT series project. Local support for Burgos is provided by the Energy Development Corporation (EDC, Philippines). The Paris TCCON site has received funding from Sorbonne Université, the French research centre CNRS, the French space agency CNES, and Région Île-de-France. The Ascension Island TCCON station has been supported by the European Space Agency (ESA; grant no. 4000120088/17/I-EF) and by the German Bundesministerium für Wirtschaft und Energie (BMWi; grant nos. 50EE1711C and 50EE1711E). We thank the ESA Ariane Tracking Station at North East Bay, Ascension Island, for hosting and local support. Nicholas M. Deutscher is funded by an ARC Future Fellowship (grant no. FT180100327). Darwin and Wollongong TCCON stations are supported by ARC (grant nos. DP160100598, LE0668470, DP140101552, DP110103118, and DP0879468) and NASA (grant nos. NAG5-12247 and NNG05-GD07G).

**Financial support.** This work has received funding from CNES and CNRS. Matthieu Dogniaux has been funded by Airbus Defence and Space in the framework of a scientific collaboration with École polytechnique.

**Review statement.** This paper was edited by Abhishek Chatterjee and reviewed by two anonymous referees.

## References

- Basu, S., Guerlet, S., Butz, A., Houweling, S., Hasekamp, O., Aben, I., Krummel, P., Steele, P., Langenfelds, R., Torn, M., Biraud, S., Stephens, B., Andrews, A., and Worthy, D.: Global  $\text{CO}_2$  fluxes estimated from GOSAT retrievals of total column  $\text{CO}_2$ , *Atmos. Chem. Phys.*, 13, 8695–8717, <https://doi.org/10.5194/acp-13-8695-2013>, 2013.

- Basu, S., Baker, D. F., Chevallier, F., Patra, P. K., Liu, J., and Miller, J. B.: The impact of transport model differences on CO<sub>2</sub> surface flux estimates from OCO-2 retrievals of column average CO<sub>2</sub>, *Atmos. Chem. Phys.*, 18, 7189–7215, <https://doi.org/10.5194/acp-18-7189-2018>, 2018.
- Blumenstock, T., Hase, F., Schneider, M., García, O. E., and Sepúlveda, E.: TCCON data from Izana (ES), Release GGG2014.R1 (Version R1), CaltechDATA [data set], <https://doi.org/10.14291/TCCON.GGG2014.IZANA01.R1>, 2017.
- Bösch, H., Toon, G. C., Sen, B., Washenfelder, R. A., Wennberg, P. O., Buchwitz, M., de Beek, R., Burrows, J. P., Crisp, D., Christi, M., Connor, B. J., Natraj, V., and Yung, Y. L.: Space-based near-infrared CO<sub>2</sub> measurements: Testing the Orbiting Carbon Observatory retrieval algorithm and validation concept using SCIAMACHY observations over Park Falls, Wisconsin, *J. Geophys. Res.*, 111, D23302, <https://doi.org/10.1029/2006JD007080>, 2006.
- Bovensmann, H., Burrows, J. P., Buchwitz, M., Frerick, J., Noël, S., Rozanov, V. V., Chance, K. V., and Goede, A. P. H.: SCIAMACHY: Mission objectives and measurement modes, *J. Atmos. Sci.*, 56, 127–150, [https://doi.org/10.1175/1520-0469\(1999\)056<0127:SMOAMM>2.0.CO;2](https://doi.org/10.1175/1520-0469(1999)056<0127:SMOAMM>2.0.CO;2), 1999.
- Brooker, L.: CONSTELLATION OF SMALL SATELLITES FOR THE MONITORING OF GREENHOUSE GASES, in: 69th International Astronautical Congress (IAC), Bremen, Germany, 1–5 October 2018.
- 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., and Warneke, T.: Toward accurate CO<sub>2</sub> and CH<sub>4</sub> observations from GOSAT, *Geophys. Res. Lett.*, 38, L14812, <https://doi.org/10.1029/2011GL047888>, 2011.
- Chédin, A., Serrar, S., Scott, N. A., Crevoisier, C., and Armante, R.: First global measurement of midtropospheric CO<sub>2</sub> from NOAA polar satellites: Tropical zone, *J. Geophys. Res.*, 108, 4581, <https://doi.org/10.1029/2003JD003439>, 2003.
- Cheruy, F., Scott, N. A., Armante, R., Tournier, B., and Chedin, A.: Contribution to the development of radiative transfer models for high spectral resolution observations in the infrared, *J. Quant. Spectrosc. Ra.*, 53, 597–611, [https://doi.org/10.1016/0022-4073\(95\)00026-H](https://doi.org/10.1016/0022-4073(95)00026-H), 1995.
- Chevallier, F., Bréon, F.-M., and Rayner, P. J.: Contribution of the Orbiting Carbon Observatory to the estimation of CO<sub>2</sub> sources and sinks: Theoretical study in a variational data assimilation framework, *J. Geophys. Res.*, 112, D09307, <https://doi.org/10.1029/2006JD007375>, 2007.
- Connor, B. J., Boesch, H., Toon, G., Sen, B., Miller, C., and Crisp, D.: Orbiting Carbon Observatory: Inverse method and prospective error analysis, *J. Geophys. Res.*, 113, D05305, <https://doi.org/10.1029/2006JD008336>, 2008.
- Crevoisier, C., Heilliette, S., Chédin, A., Serrar, S., Armante, R., and Scott, N. A.: Midtropospheric CO<sub>2</sub> concentration retrieval from AIRS observations in the tropics, *Geophys. Res. Lett.*, 31, L17106, <https://doi.org/10.1029/2004GL020141>, 2004.
- Crevoisier, C., Chédin, A., Matsueda, H., Machida, T., Armante, R., and Scott, N. A.: First year of upper tropospheric integrated content of CO<sub>2</sub> from IASI hyperspectral infrared observations, *Atmos. Chem. Phys.*, 9, 4797–4810, <https://doi.org/10.5194/acp-9-4797-2009>, 2009a.
- Crevoisier, C., Nobileau, D., Fiore, A. M., Armante, R., Chédin, A., and Scott, N. A.: Tropospheric methane in the tropics – first year from IASI hyperspectral infrared observations, *Atmos. Chem. Phys.*, 9, 6337–6350, <https://doi.org/10.5194/acp-9-6337-2009>, 2009b.
- Crevoisier, C., Clerbaux, C., Guidard, V., Phulpin, T., Armante, R., Barret, B., Camy-Peyret, C., Chaboureaud, J.-P., Coheur, P.-F., Crépeau, L., Dufour, G., Labonnote, L., Lavanant, L., Hadji-Lazaro, J., Herbin, H., Jacquinet-Husson, N., Payan, S., Péquignot, E., Pierangelo, C., Sellitto, P., and Stubenrauch, C.: Towards IASI-New Generation (IASI-NG): impact of improved spectral resolution and radiometric noise on the retrieval of thermodynamic, chemistry and climate variables, *Atmos. Meas. Tech.*, 7, 4367–4385, <https://doi.org/10.5194/amt-7-4367-2014>, 2014.
- Crisp, D., Pollock, H. R., Rosenberg, R., Chapsky, L., Lee, R. A. M., Oyafuso, F. A., Frankenberg, C., O'Dell, C. W., 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., and Wunch, D.: The on-orbit performance of the Orbiting Carbon Observatory-2 (OCO-2) instrument and its radiometrically calibrated products, *Atmos. Meas. Tech.*, 10, 59–81, <https://doi.org/10.5194/amt-10-59-2017>, 2017.
- De Mazière, M., Sha, M. K., Desmet, F., Hermans, C., Scolas, F., Kumps, N., Metzger, J.-M., Duflot, V., and Cammas, J.-P.: TCCON data from Réunion Island (RE), Release GGG2014.R1 (Version R1), CaltechDATA [data set], <https://doi.org/10.14291/TCCON.GGG2014.REUNION01.R1>, 2017.
- Deutscher, N. M., Notholt, J., Messerschmidt, J., Weinzierl, C., Warneke, T., Petri, C., and Grupe, P.: TCCON data from Bialystok (PL), Release GGG2014.R2 (Version R2), CaltechDATA [data set], <https://doi.org/10.14291/TCCON.GGG2014.BIALYSTOK01.R2>, 2019.
- Drouin, B. J., Benner, D. C., Brown, L. R., Cich, M. J., Crawford, T. J., Devi, V. M., Guillaume, A., Hodges, J. T., Mlawer, E. J., Robichaud, D. J., Oyafuso, F., Payne, V. H., Sung, K., Wishnow, E. H., and Yu, S.: Multispectrum analysis of the oxygen A-band, *J. Quant. Spectrosc. Ra.*, 186, 118–138, <https://doi.org/10.1016/j.jqsrt.2016.03.037>, 2017.
- Eck, T. F., Holben, B. N., Reid, J. S., Dubovik, O., Smirnov, A., O'Neill, N. T., Slutsker, I., and Kinne, S.: Wavelength dependence of the optical depth of biomass burning, urban, and desert dust aerosols, *J. Geophys. Res.*, 104, 31333–31349, <https://doi.org/10.1029/1999JD900923>, 1999.
- Ehret, G., Bousquet, P., Pierangelo, C., Alpers, M., Millet, B., Abshire, J. B., Bovensmann, H., Burrows, J. P., Chevallier, F., Ciais, P., Crevoisier, C., Fix, A., Flamant, P., Frankenberg, C., Gilbert, F., Heim, B., Heimann, M., Houweling, S., Hubberten, H. W., Jöckel, P., Law, K., Löw, A., Marshall, J., Agusti-Panareda, A., Payan, S., Prigent, C., Rairoux, P., Sachs, T., Scholze, M., and Wirth, M.: MERLIN: A French-German Space Lidar Mission Dedicated to Atmospheric Methane, *Remote Sens.*, 9, 1052, <https://doi.org/10.3390/rs9101052>, 2017.
- Feist, D. G., Arnold, S. G., John, N. and Geibel, M. C.: TCCON data from Ascension Island (SH), Release GGG2014.R0 (Version GGG2014.R0), CaltechDATA [data set], <https://doi.org/10.14291/TCCON.GGG2014.ASCENSION01.R0/1149285>, 2014.



- Giles, D. M., Sinyuk, A., Sorokin, M. G., Schafer, J. S., Smirnov, A., Slutsker, I., Eck, T. F., Holben, B. N., Lewis, J. R., Campbell, J. R., Welton, E. J., Korkin, S. V., and Lyapustin, A. I.: Advancements in the Aerosol Robotic Network (AERONET) Version 3 database – automated near-real-time quality control algorithm with improved cloud screening for Sun photometer aerosol optical depth (AOD) measurements, *Atmos. Meas. Tech.*, 12, 169–209, <https://doi.org/10.5194/amt-12-169-2019>, 2019.
- Gisi, M., Hase, F., Dohe, S., Blumenstock, T., Simon, A., and Keens, A.: XCO<sub>2</sub>-measurements with a tabletop FTS using solar absorption spectroscopy, *Atmos. Meas. Tech.*, 5, 2969–2980, <https://doi.org/10.5194/amt-5-2969-2012>, 2012.
- Gordon, I. E., Rothman, L. S., Hill, C., Kochanov, R. V., Tan, Y., Bernath, P. F., Birk, M., Boudon, V., Campargue, A., Chance, K. V., Drouin, B. J., Flaud, J.-M., Gamache, R. R., Hodges, J. T., Jacquemart, D., Perevalov, V. I., Perrin, A., Shine, K. P., Smith, M.-A. H., Tennyson, J., Toon, G. C., Tran, H., Tyuterev, V. G., Barbe, A., Császár, A. G., Devi, V. M., Furtenbacher, T., Harrison, J. J., Hartmann, J.-M., Jolly, A., Johnson, T. J., Karmann, T., Kleiner, I., Kyuberis, A. A., Loos, J., Lyulin, O. M., Massie, S. T., Mikhailenko, S. N., Moazzen-Ahmadi, N., Müller, H. S. P., Naumenko, O. V., Nikitin, A. V., Polyansky, O. L., Rey, M., Rotger, M., Sharpe, S. W., Sung, K., Starikova, E., Tashkun, S. A., Auwera, J., Vander, Wagner, G., Wilzewski, J., Wcislo, P., Yu, S., and Zak, E. J.: The HITRAN2016 molecular spectroscopic database, *J. Quant. Spectrosc. Ra.*, 203, 3–69, <https://doi.org/10.1016/j.jqsrt.2017.06.038>, 2017.
- 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., Paton-Walsh, C., Jones, N. B., Kettlewell, G. C., Connor, B. J., Macatangay, R. C., Roehl, C., Ryzek, M., Glowacki, J., Culgan, T., and Bryant, G. W.: TCCON data from Darwin (AU), Release GGG2014.R0 (Version GGG2014.R0), CaltechDATA [data set], <https://doi.org/10.14291/TCCON.GGG2014.DARWIN01.R0/1149290>, 2014a.
- Griffith, D. W. T., Velazco, V. A., Deutscher, N. M., Paton-Walsh, C., Jones, N. B., Wilson, S. R., Macatangay, R. C., Kettlewell, G. C., Buchholz, R. R., and Rigenbach, M. O.: TCCON data from Wollongong (AU), Release GGG2014.R0 (Version GGG2014.R0), CaltechDATA [data set], <https://doi.org/10.14291/TCCON.GGG2014.WOLLONGONG01.R0/1149291>, 2014b.
- Hase, F., Blumenstock, T., Dohe, S., Groß, J., and Kiel, M.: TCCON data from Karlsruhe (DE), Release GGG2014.R1 (Version GGG2014.R1), CaltechDATA [data set], <https://doi.org/10.14291/TCCON.GGG2014.KARLSRUHE01.R1/1182416>, 2015.
- Hase, F., Frey, M., Kiel, M., Blumenstock, T., Harig, R., Keens, A., and Orphal, J.: Addition of a channel for XCO observations to a portable FTIR spectrometer for greenhouse gas measurements, *Atmos. Meas. Tech.*, 9, 2303–2313, <https://doi.org/10.5194/amt-9-2303-2016>, 2016.
- Holben, B. N., Eck, T. F., Slutsker, I., Tanré, D., Buis, J. P., Setzer, A., Vermote, E., Reagan, J. A., Kaufman, Y. J., Nakajima, T., Lavenue, F., Jankowiak, I., and Smirnov, A.: AERONET—A Federated Instrument Network and Data Archive for Aerosol Characterization, *Remote Sens. Environ.*, 66, 1–16, [https://doi.org/10.1016/S0034-4257\(98\)00031-5](https://doi.org/10.1016/S0034-4257(98)00031-5), 1998.
- Houweling, S., Hartmann, W., Aben, I., Schrijver, H., Skidmore, J., Roelofs, G.-J., and Breon, F.-M.: Evidence of systematic errors in SCIAMACHY-observed CO<sub>2</sub> due to aerosols, *Atmos. Chem. Phys.*, 5, 3003–3013, <https://doi.org/10.5194/acp-5-3003-2005>, 2005.
- Inoue, M., Morino, I., Uchino, O., Nakatsuru, T., Yoshida, Y., Yokota, T., Wunch, D., Wennberg, P. O., Roehl, C. M., Griffith, D. W. T., Velazco, V. A., Deutscher, N. M., Warneke, T., Notholt, J., Robinson, J., Sherlock, V., Hase, F., Blumenstock, T., Rettinger, M., Sussmann, R., Kyrö, E., Kivi, R., Shiomi, K., Kawakami, S., De Mazière, M., Arnold, S. G., Feist, D. G., Barrow, E. A., Barney, J., Dubey, M., Schneider, M., Iraci, L. T., Podolske, J. R., Hillyard, P. W., Machida, T., Sawa, Y., Tsuboi, K., Matsueda, H., Sweeney, C., Tans, P. P., Andrews, A. E., Biraud, S. C., Fukuyama, Y., Pittman, J. V., Kort, E. A., and Tanaka, T.: Bias corrections of GOSAT SWIR XCO<sub>2</sub> and XCH<sub>4</sub> with TCCON data and their evaluation using aircraft measurement data, *Atmos. Meas. Tech.*, 9, 3491–3512, <https://doi.org/10.5194/amt-9-3491-2016>, 2016.
- Iraci, L. T., Podolske, J. R., Hillyard, P. W., Roehl, C., Wennberg, P. O., Blavier, J.-F., Landeros, J., Allen, N., Wunch, D., Zavaleta, J., Quigley, E., Osterman, G. B., Albertson, R., Dunwoody, K., and Boyden, H.: TCCON data from Edwards (US), Release GGG2014.R1 (Version GGG2014.R1), CaltechDATA [data set], <https://doi.org/10.14291/TCCON.GGG2014.EDWARDS01.R1/1255068>, 2016.
- Jacquinet-Husson, N., Armante, R., Scott, N. A., Chédin, A., Crépeau, L., Boutammine, C., Bouhdaoui, A., Crevoisier, C., Capelle, V., Boonne, C., Poulet-Crovisier, N., Barbe, A., Benner, D. C., Boudon, V., Brown, L. R., Buldyreva, J., Campargue, A., Coudert, L. H., Devi, V. M., Down, M. J., Drouin, B. J., Fayt, A., Fittschen, C., Flaud, J.-M., Gamache, R. R., Harrison, J. J., Hill, C., Hodnebrog, Ø., Hu, S.-M., Jacquemart, D., Jolly, A., Jiménez, E., Lavrentieva, N. N., Liu, A.-W., Lodi, L., Lyulin, O. M., Massie, S. T., Mikhailenko, S., Müller, H. S. P., Naumenko, O. V., Nikitin, A., Nielsen, C. J., Orphal, J., Perevalov, V. I., Perrin, A., Polovtseva, E., Predoi-Cross, A., Rotger, M., Ruth, A. A., Yu, S. S., Sung, K., Tashkun, S. A., Tennyson, J., Tyuterev, V. G., Auwera, J. V., Voronin, B. A., and Makie, A.: The 2015 edition of the GEISA spectroscopic database, *J. Mol. Spectrosc.*, 327, 31–72, <https://doi.org/10.1016/j.jms.2016.06.007>, 2016.
- Kawakami, S., Ohyama, H., Arai, K., Okumura, H., Taura, C., Fukamachi, T., and Sakashita, M.: TCCON data from Saga (JP), Release GGG2014.R0 (Version GGG2014.R0), CaltechDATA [data set], <https://doi.org/10.14291/TCCON.GGG2014.SAGA01.R0/1149283>, 2014.
- Kivi, R. and Heikkinen, P.: Fourier transform spectrometer measurements of column CO<sub>2</sub> at Sodankylä, Finland, *Geosci. Instrum. Method. Data Syst.*, 5, 271–279, <https://doi.org/10.5194/gi-5-271-2016>, 2016.
- Kivi, R., Heikkinen, P., and Kyrö, E.: TCCON data from Sodankylä (FI), Release GGG2014.R0 (Version GGG2014.R0), CaltechDATA [data set], <https://doi.org/10.14291/TCCON.GGG2014.SODANKYLA01.R0/1149280>, 2014.
- Lamouroux, J., Tran, H., Laraia, A. L., Gamache, R. R., Rothman, L. S., Gordon, I. E., and Hartmann, J.-M.: Updated database plus software for line-mixing in CO<sub>2</sub> infrared spectra and their test using laboratory spectra in the 1.5–2.3 μm region, *J. Quant. Spectrosc. Ra.*, 111, 2321–2331, <https://doi.org/10.1016/j.jqsrt.2010.03.006>, 2010.

- Le Quéré, C., Raupach, M., Canadell, J., Marland, G., Bopp, L., Ciais, P., Conway, T., Doney, S., Feely, R., Foster, P., Friedlingstein, P., Gurney, K., Houghton, R., House, J., Huntingford, C., Levy, P., Lomas, M., Majkut, J., Metzl, N., and Woodward, I.: Trends in the sources and sinks of carbon dioxide, *Nat. Geosci.*, 2, 831–836, <https://doi.org/10.1038/ngeo689>, 2009.
- Meijer, Y. and Team, C.: Copernicus CO<sub>2</sub> Monitoring Mission Requirements Document, ESA, available at: [https://esamultimedia.esa.int/docs/EarthObservation/CO2M\\_MRD\\_v2.0\\_Issued20190927.pdf](https://esamultimedia.esa.int/docs/EarthObservation/CO2M_MRD_v2.0_Issued20190927.pdf) (last access: 25 May 2021), 2019.
- Morino, I., Matsuzaki, T., and Horikawa, M.: TCCON data from Tsukuba (JP), 125HR, Release GGG2014.R2 (Version R2), CaltechDATA [data set], <https://doi.org/10.14291/TCCON.GGG2014.TSUKUBA02.R2>, 2018.
- Nguyen, H., Osterman, G., Wunch, D., O'Dell, C., Mandrake, L., Wennberg, P., Fisher, B., and Castano, R.: A method for collocating satellite X<sub>CO<sub>2</sub></sub> data to ground-based data and its application to ACOS-GOSAT and TCCON, *Atmos. Meas. Tech.*, 7, 2631–2644, <https://doi.org/10.5194/amt-7-2631-2014>, 2014.
- Notholt, J., Petri, C., Warneke, T., Deutscher, N. M., Palm, M., Buschmann, M., Weinzierl, C., Macatangay, R. C., and Grupe, P.: TCCON data from Bremen (DE), Release GGG2014.R0 (Version GGG2014.R0), CaltechDATA [data set], <https://doi.org/10.14291/TCCON.GGG2014.BREMEN01.R0/1149275>, 2014.
- O'Dell, C. W., Connor, B., Bösch, H., O'Brien, D., Frankenberg, C., Castano, R., Christi, M., Eldering, D., Fisher, B., Gunson, M., McDuffie, J., Miller, C. E., Natraj, V., Oyafuso, F., Polonsky, I., Smyth, M., Taylor, T., Toon, G. C., Wennberg, P. O., and Wunch, D.: The ACOS CO<sub>2</sub> retrieval algorithm – Part 1: Description and validation against synthetic observations, *Atmos. Meas. Tech.*, 5, 99–121, <https://doi.org/10.5194/amt-5-99-2012>, 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., and Velasco, V. A.: Improved retrievals of carbon dioxide from Orbiting Carbon Observatory-2 with the version 8 ACOS algorithm, *Atmos. Meas. Tech.*, 11, 6539–6576, <https://doi.org/10.5194/amt-11-6539-2018>, 2018.
- Oyafuso, F., Payne, V. H., Drouin, B. J., Devi, V. M., Benner, D. C., Sung, K., Yu, S., Gordon, I. E., Kochanov, R., Tan, Y., Crisp, D., Mlawer, E. J., and Guillaume, A.: High accuracy absorption coefficients for the Orbiting Carbon Observatory-2 (OCO-2) mission: Validation of updated carbon dioxide cross-sections using atmospheric spectra, *J. Quant. Spectrosc. Ra.*, 203, 213–223, <https://doi.org/10.1016/j.jqsrt.2017.06.012>, 2017.
- 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., and Wunch, D.: Methane observations from the Greenhouse Gases Observing SATellite: Comparison to ground-based TCCON data and model calculations, *Geophys. Res. Lett.*, 38, L15807, <https://doi.org/10.1029/2011GL047871>, 2011.
- Pascal, V., Buil, C., Loesel, J., Tauziede, L., Jouglet, D., and Buisson, F.: An improved microcarb dispersive instrumental concept for the measurement of greenhouse gases concentration in the atmosphere, in: International Conference on Space Optics – ICSO 2014, edited by: Sodnik, Z., Cugny, B., and Karafolas, N., SPIE, 10563, 1028–1036, 2017.
- Rayner, P. J. and O'Brien, D. M.: The utility of remotely sensed CO<sub>2</sub> concentration data in surface source inversions, *Geophys. Res. Lett.*, 28, 175–178, <https://doi.org/10.1029/2000GL011912>, 2001.
- Reuter, M., Bovensmann, H., Buchwitz, M., Burrows, J. P., Connor, 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., Velasco, V., Warneke, T., Wennberg, P. O., and Wunch, D.: Retrieval of atmospheric CO<sub>2</sub> with enhanced accuracy and precision from SCIAMACHY: Validation with FTS measurements and comparison with model results, *J. Geophys. Res.*, 116, D04301, <https://doi.org/10.1029/2010JD015047>, 2011.
- Reuter, M., Buchwitz, M., Schneising, O., Noël, S., Rozanov, V., Bovensmann, H., and Burrows, J.: A Fast Atmospheric Trace Gas Retrieval for Hyperspectral Instruments Approximating Multiple Scattering–Part 1: Radiative Transfer and a Potential OCO-2 XCO<sub>2</sub> Retrieval Setup, *Remote Sens.*, 9, 1159, <https://doi.org/10.3390/rs9111159>, 2017a.
- Reuter, M., Buchwitz, M., Schneising, O., Noël, S., Bovensmann, H., and Burrows, J.: A Fast Atmospheric Trace Gas Retrieval for Hyperspectral Instruments Approximating Multiple Scattering–Part 2: Application to XCO<sub>2</sub> Retrievals from OCO-2, *Remote Sens.*, 9, 1102, <https://doi.org/10.3390/rs9111102>, 2017b.
- Rodgers, C. D.: Inverse Methods for Atmospheric Sounding, World Scientific, Singapore, 2000.
- Scott, N. A. and Chédin, A.: A Fast Line-by-Line Method for Atmospheric Absorption Computations: The Automated Atmospheric Absorption Atlas, *J. Appl. Meteorol.*, 20, 802–812, 1981.
- Sherlock, V., Connor, B., Robinson, J., Shiona, H., Smale, D., and Pollard, D. F.: TCCON data from Lauder (NZ), 125HR, Release GGG2014.R0 (Version GGG2014.R0), CaltechDATA [data set], <https://doi.org/10.14291/TCCON.GGG2014.LAUDER02.R0/1149298>, 2014.
- Spurr, R. J. D.: Simultaneous derivation of intensities and weighting functions in a general pseudo-spherical discrete ordinate radiative transfer treatment, *J. Quant. Spectrosc. Ra.*, 75, 129–175, [https://doi.org/10.1016/S0022-4073\(01\)00245-X](https://doi.org/10.1016/S0022-4073(01)00245-X), 2002.
- Spurr, R. J. D.: VLIDORT: A linearized pseudo-spherical vector discrete ordinate radiative transfer code for forward model and retrieval studies in multilayer multiple scattering media, *J. Quant. Spectrosc. Ra.*, 102, 316–342, <https://doi.org/10.1016/j.jqsrt.2006.05.005>, 2006.
- Strong, K., Roche, S., Franklin, J. E., Mendonca, J., Lutsch, E., Weaver, D., Fogal, P. F., Drummond, J. R., Batchelor, R., and Lindenmaier, R.: TCCON data from Eureka (CA), Release GGG2014.R3 (Version R3), CaltechDATA [data set], <https://doi.org/10.14291/TCCON.GGG2014.EUREKA01.R3>, 2019.
- Stutz, J., Holben, B., Takemura, T., Kivi, R., Aaltonen, V., Ravetta, F., and Pelon, J.: AERONET Version 3: AOD Level 2.0, NASA [data set], available at: [https://aeronet.gsfc.nasa.gov/new\\_web/download\\_all\\_v3\\_aod.html](https://aeronet.gsfc.nasa.gov/new_web/download_all_v3_aod.html), last access: 25 May 2021.

- Té, Y., Jeseck, P., and Janssen, C.: TCCON data from Paris (FR), Release GGG2014.R0 (Version GGG2014.R0), CaltechDATA [data set], <https://doi.org/10.14291/TCCON.GGG2014.PARIS01.R0/1149279>, 2014.
- Toon, G. C.: Solar Line List for the TCCON 2014 Data Release (Version GGG2014.R0), CaltechDATA [data set], <https://doi.org/10.14291/TCCON.GGG2014.SOLAR.R0/1221658>, 2015.
- Tournier, B.: STRANSAC-93 et 4A-93: STRANSAC-93 et 4A-93: Développement et validation des nouvelles versions des codes de transfert radiatif pour application au projet IASI, Internal Rep. LMD, No. 201, LMD/CNRS, Ecole Polytechnique, PALAISEAU, France, 1995.
- Tran, H. and Hartmann, J.-M.: An improved O<sub>2</sub> A band absorption model and its consequences for retrievals of photon paths and surface pressures, *J. Geophys. Res.*, 113, D18104, <https://doi.org/10.1029/2008JD010011>, 2008.
- Warneke, T., Messerschmidt, J., Notholt, J., Weinzierl, C., Deutscher, N. M., Petri, C., and Grupe, P.: TCCON data from Orléans (FR), Release GGG2014.R1 (Version R1), CaltechDATA [data set], <https://doi.org/10.14291/TCCON.GGG2014.ORLEANS01.R1>, 2019.
- Wennberg, P. O., Wunch, D., Roehl, C. M., Blavier, J.-F., Toon, G. C., and Allen, N. T.: TCCON data from Caltech (US), Release GGG2014.R1 (Version GGG2014.R1), CaltechDATA [data set], <https://doi.org/10.14291/TCCON.GGG2014.PASADENA01.R1/1182415>, 2015.
- Wennberg, P. O., Wunch, D., Roehl, C. M., Blavier, J.-F., Toon, G. C., and Allen, N. T.: TCCON data from Lamont (US), Release GGG2014.R1 (Version GGG2014.R1), CaltechDATA [data set], <https://doi.org/10.14291/TCCON.GGG2014.LAMONT01.R1/1255070>, 2016.
- Wennberg, P. O., Roehl, C. M., Wunch, D., Toon, G. C., Blavier, J.-F., Washenfelder, R., Keppel-Aleks, G., Allen, N. T., and Ayers, J.: TCCON data from Park Falls (US), Release GGG2014.R1 (Version GGG2014.R1), CaltechDATA [data set], <https://doi.org/10.14291/TCCON.GGG2014.PARKFALLS01.R1>, 2017.
- Wu, L., Hasekamp, O., Hu, H., Landgraf, J., Butz, A., van 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 Mazzière, M., Sussmann, R., Kivi, R., Warneke, T., Goo, T.-Y., and Té, Y.: Carbon dioxide retrieval from OCO-2 satellite observations using the RemoTeC algorithm and validation with TCCON measurements, *Atmos. Meas. Tech.*, 11, 3111–3130, <https://doi.org/10.5194/amt-11-3111-2018>, 2018.
- Wunch, D., Wennberg, P. O., Toon, G. C., Connor, B. J., Fisher, B., Osterman, G. B., Frankenberg, C., Mandrake, L., O'Dell, C., Ahonen, P., Biraud, S. C., Castano, R., Cressie, N., Crisp, D., Deutscher, N. M., Eldering, A., Fisher, M. L., Griffith, D. W. T., Gunson, M., Heikkinen, P., Keppel-Aleks, G., Kyrö, E., Lindenmaier, R., Macatangay, R., Mendonca, J., Messerschmidt, J., Miller, C. E., Morino, I., Notholt, J., Oyafuso, F. A., Rettinger, M., Robinson, J., Roehl, C. M., Salawitch, R. J., Sherlock, V., Strong, K., Sussmann, R., Tanaka, T., Thompson, D. R., Uchino, O., Warneke, T., and Wofsy, S. C.: A method for evaluating bias in global measurements of CO<sub>2</sub> total columns from space, *Atmos. Chem. Phys.*, 11, 12317–12337, <https://doi.org/10.5194/acp-11-12317-2011>, 2011a.
- Wunch, D., Toon, G. C., Blavier, J.-F. L., Washenfelder, R. A., Notholt, J., Connor, B. J., Griffith, D. W. T., Sherlock, V., and Wennberg, P. O.: The Total Carbon Column Observing Network, *Philos. T. Roy. Soc. A*, 369, 2087–2112, <https://doi.org/10.1098/rsta.2010.0240>, 2011b.
- Wunch, D., Toon, G. C., Sherlock, V., Deutscher, N. M., Liu, C., Feist, D. G., and Wennberg, P. O.: The Total Carbon Column Observing Network's GGG2014 Data Version, Tech. rep., California Institute of Technology, Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA, <https://doi.org/10.14291/TCCON.GGG2014.DOCUMENTATION.R0/1221662>, 2015.
- 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., Velasco, V. A., Notholt, J., Warneke, T., Petri, C., De Mazzière, 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., and Eldering, A.: Comparisons of the Orbiting Carbon Observatory-2 (OCO-2) X<sub>CO<sub>2</sub></sub> measurements with TCCON, *Atmos. Meas. Tech.*, 10, 2209–2238, <https://doi.org/10.5194/amt-10-2209-2017>, 2017.
- Yokota, T., Yoshida, Y., Eguchi, N., Ota, Y., Tanaka, T., Watanabe, H., and Maksyutov, S.: Global Concentrations of CO<sub>2</sub> and CH<sub>4</sub> Retrieved from GOSAT: First Preliminary Results, *SOLA*, 5, 160–163, <https://doi.org/10.2151/sola.2009-041>, 2009.
- Yoshida, Y., Ota, Y., Eguchi, N., Kikuchi, N., Nobuta, K., Tran, H., Morino, I., and Yokota, T.: Retrieval algorithm for CO<sub>2</sub> and CH<sub>4</sub> column abundances from short-wavelength infrared spectral observations by the Greenhouse gases observing satellite, *Atmos. Meas. Tech.*, 4, 717–734, <https://doi.org/10.5194/amt-4-717-2011>, 2011.
- 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., and Yokota, T.: Improvement of the retrieval algorithm for GOSAT SWIR XCO<sub>2</sub> and XCH<sub>4</sub> and their validation using TCCON data, *Atmos. Meas. Tech.*, 6, 1533–1547, <https://doi.org/10.5194/amt-6-1533-2013>, 2013.