OCO D-55208

# Orbiting Carbon Observatory–2 (OCO-2)

# Data Product User's Guide, Operational L1 and L2 Data Versions 7 and 7R

Version F
September 2, 2015
Data Releases 7 and 7R

National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Copyright 2015 California Institute of Technology. U.S. Government sponsorship acknowledged.

## Preparers

Gregory Osterman, JPL, California Institute of Technology
OCO-2 Science Validation Lead

Annmarie Eldering, JPL, California Institute of Technology
OCO-2 Deputy Project Scientist

Charles Avis, JPL, California Institute of Technology
ACOS, OCO-2 Science Data Operations System

Brian Chafin, JPL, California Institute of Technology
ACOS, OCO-2 Science Data Operations System

Christopher O'Dell, CSU
ACOS, OCO-2 Algorithm Team

Christian Frankenberg, JPL, California Institute of Technology
ACOS, OCO-2 Algorithm Team

Brendan Fisher, JPL, California Institute of Technology
ACOS, OCO-2 Validation Team

Lukas Mandrake, JPL, California Institute of Technology
ACOS, OCO-2 Validation Team

Debra Wunch, California Institute of Technology
ACOS, OCO-2 Validation Team

Robert Granat, JPL, California Institute of Technology
OCO-2 Algorithm Lead

David Crisp, JPL, California Institute of Technology
OCO-2 Science Team Lead

.

## Document History

| Revision Date | Changes | Author |
|---|---|---|
| A—11 April 2014 | Initial release, prelaunch release of simulated data | G. Osterman and team |
| B—16 June 2014 | Update to prerelease, correcting MaxMS values | G. Osterman and team |
| C—30 December 2014 | Updated for release with L1B data | R. Granat and team |
| D—30 March 2015 | Updated for version 6 release with initial L2 data | R. Granat and team |
| E—19 June 2015 | Updated for version 7 release | R. Granat and team |
| F – 02 September 2015 | Updated with information on lite files and warn levels | G. Osterman and team |

The research described in this document was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

Copyright 2015 California Institute of Technology. U.S. Government sponsorship acknowledged.

***TABLE OF CONTENTS***

**1** Introduction ..... 1
1.1 Document Overview ..... 1
1.2 Data Products ..... 1
1.3 Data Usage Policy ..... 2
2 Mission Overview ..... 3
2.1 Measurement Approach ..... 4
2.2 Instrument Characteristics ..... 8
2.3 OCO-2 Algorithms ..... 14
3 Overview of Data Products ..... 17
3.1 File Naming Conventions ..... 17
3.2 File Content ..... 19
4 OCO-2 L1BSc Data Products ..... 20
4.1 Key Data Fields ..... 20
5 OCO-2 L2 Data Products ..... 24
5.1 Data Description and User Alerts ..... 24
5.2 Key Data Fields for Standard and Diagnostic Files ..... 24
5.3 Key Data Fields Diagnostic Files ..... 26
6 ABO2 Preprocessor ..... 28
6.1 Prescreening of OCO-2 Soundings for Cloud and Aerosol ..... 28
6.2 Key Science Data Fields ..... 28
7 IMAP-DOAS Preprocessor ..... 31
7.1 Advanced Cloud and Aerosol Screening ..... 31
7.2 Retrievals of Solar-Induced Chlorophyll Fluorescence ..... 31
7.3 Key Science Data Fields ..... 31
8 Full Data Tables ..... 35
8.1 Metadata in all Product Files ..... 35
8.2 L2Std and L2Dia Data Tables ..... **Error! Bookmark not defined.**
8.3 L1B and DOAS Data Tables ..... 42
8.4 L1B Data Tables ..... 44
9 Tools and Data Services ..... 49
9.1 HDFView ..... 49
9.2 Panoply ..... 49
9.3 Mirador ..... 49
9.4 JPL CO2 Virtual Science Data Environment ..... 49
10 Contact Information ..... 50
11 Acknowledgements and References ..... 51
11.1 Acknowledgements ..... 51
11.2 Additional Resources ..... 51
11.3 References ..... 51
12 Acronyms ..... 55

***LIST OF FIGURES***

Figure 2-1. Nadir, glint, and target observations. (a) Nadir observations are acquired over the sunlit hemisphere at latitudes where the surface solar zenith angle is <85°. On all orbits except downlink orbits, as the Observatory passes over the northern terminator, it pitches up to point the instrument aperture at the Sun for solar radiometric calibrations. (b) Glint observations are made at latitudes where the solar zenith angle at the apparent glint spot is less than ~75°. (c) For target observations, the spacecraft scans the instrument across a stationary surface target as it flies overhead. ..... 4

Figure 2-2. (a) The principal plane is defined with respect to the Sun, surface footprint, and spacecraft. (b) In nadir mode, the spacecraft azimuth changes during the orbit to maintain the alignment of the spectrometer slits (which are roughly parallel with the axis of the solar panels) perpendicular to the principal plane [Crisp et al., 2007]. In glint mode, the spectrometer slits are rotated 30 degrees with respect to the principal plane to increase signal in highly polarized scenes. ..... 6

Figure 2-3. The OCO-2 instrument showing the major optical components and optical path (right) and images of spectra recorded by the FPA in the three spectral channels (left). ..... 8

Figure 2-4. (a) The illumination and readout scheme used for the OCO-2 focal plane arrays, showing the direction of spectral dispersion from bottom to top, and the spatial dimension from left to right. The ~160 illuminated pixels in the spatial dimension are summed into eight 20-pixel "super pixels" or "footprints." If one or more of the pixels in a footprint is "bad" (red pixels), it is eliminated from the sum. One of the 20 full-resolution "color slices" is also shown at the bottom. (b) Spatial layout of 8 cross-track footprints for nadir observations over Washington, D.C. Each footprint is shaped like a parallelogram, rather than a square, because of the rolling readout of the FPAs.. ..... 9

Figure 2-5. Images of the $O_2$ A band (top), 1.61 μm $CO_2$ band (middle) and the 2.06 μm $CO_2$ band (bottom) in "summed mode" taken from the OCO-2 first light frame over Papua New Guinea. The 8 spatially-summed footprints in each band are shown from bottom (footprint 1) to top (footprint 8). In this 0.333-second frame, footprint 1 was contaminated by a cloud and appears brighter than the rest. This frame clearly shows the rotation, or "clocking" of each FPA columns with respect to the dark $O_2$ and $CO_2$ absorption lines. The $O_2$ lines are tilted slightly counterclockwise, while the 1.61 and 2.06 μm lines are tilted clockwise. The columns where the clocking corrections are applied are highlighted with red circles. Soundings collected in regions with strong spatial gradients in illumination can produce radiance discontinuities at these wavelengths. ..... 11

Figure 2-6. The number of spectral samples in the $O_2$ A band that are not contaminated by 10-signma cosmic rays events clearly shows the geographical extent of the South Atlantic Anomaly (SAA). ..... 12

Figure 2-7. An example showing the impact of a radiation event on spectra. Note the three spikes with unusually high intensities. ..... 13

Figure 2-8. Median response of the $O_2$ spectrometer to solar observations - squares represent the roughly once per month lunar observations which agree with the solar observations to about ±1%. Corrections to the solar data include Sun-Earth distance and a 0.2% offset for nadir versus glint orbits due to slight changes in the spacecraft thermal environment. Corrections to the lunar data include Sun-Moon distance, Moon-Earth distance and corrections for albedo due nutation and libration effects. ..... 14

Figure 2-9. Level 2 full physics retrieval flow. ..... 16

Figure 3-1. An illustration of the flow through the OCO-2 data retrieval process highlighting some of the tasks that could be performed by the data user. .....**Error! Bookmark not defined.**
Figure 4-1. Folders in the L1BSc product. ..... 20
Figure 4-2. An example of the instrument line shapes. ..... 22
Figure 5-1. Folders contained in the L2Std product. ..... 24
Figure 5-2. Variables in the RetrievalHeader folder..... 24
Figure 5-3. Variables in the RetrievalGeometry folder. ..... 25
Figure 6-1. Screenshot of an HDFView look at the ABO2 preprocessor file. ..... 29
Figure 7-1. Screenshot of an HDFView look at the IMAP-DOAS preprocessor file..... 32

***LIST OF TABLES***

**Table 4-1.** Maximum measurable signal per band. ..... 21
**Table 8-1.** OCO-2 metadata in the Level 2 products. ..... 35
**Table 8-2.** AerosolResults data in Level 2 Standard Data file (L2Std) and Diagnostic file (L2Dia)..... 39
**Table 8-3.** AlbedoResults data in Level 2 Standard Data file (L2Std) and Diagnostic file (L2Dia).....**Error! Bookmark not defined.**
**Table 8-4.** DispersionResults fields in Level 2 Standard Data file (L2Std) and Diagnostic file (L2Dia).....**Error! Bookmark not defined.**
**Table 8-5.** L1BSc sounding reference in Level 2 Standard Data file (L2Std) and Diagnostic file (L2Dia).....**Error! Bookmark not defined.**
**Table 8-6.** L1BScSpectralParameters in Level 2 Standard Data file (L2Std) and Diagnostic file (L2Dia).....**Error! Bookmark not defined.**
**Table 8-7.** PreprocessingResults data in Level 2 Standard Data file (L2Std) and Diagnostic file (L2Dia).....**Error! Bookmark not defined.**
**Table 8-8.** RetrievalGeometry in Level 2 Standard Data file (L2Std) and Diagnostic file (L2Dia). **Error! Bookmark not defined.**
**Table 8-9.** RetrievalHeader data in Level 2 Standard Data file (L2Std) and Diagnostic file (L2Dia).....**Error! Bookmark not defined.**
**Table 8-10.** RetrievalResults data in Level 2 Standard Data file (L2Std) and Diagnostic file (L2Dia).....**Error! Bookmark not defined.**
Table 8-11. SpectralParameters data in Level 2 Standard Data file (L2Std) and Diagnostic file (L2Dia)..... 41
**Table 8-12.** Spacecraft position and orientation during observations. ..... 42
**Table 8-13.** Geometric location, atmospheric geometry, and surface conditions. ..... 43
**Table 8-14**. Calibrated radiance spectra. ..... 44
**Table 8-15.** Calibrated radiance values for the color slices. ..... 45
**Table 8-16.** Location and observational geometry for each focal plane and spatial footprint. .... 45
**Table 8-17.** Configuration of detectors and color slices. ..... 46
**Table 8-18.** Frame identification data. ..... 47
**Table 8-19.** Selected temperature data. ..... 47
**Table 8-20.** Instrument performance data. ..... 47

# 1 Introduction

## 1.1 Document Overview

This document provides a brief overview of the Orbiting Carbon Observatory-2 (OCO-2) mission and then discusses the content of the publically available OCO-2 data products. Section 2 provides an overview of the OCO-2 mission. Section 3 defines the naming conventions that are used throughout the data products. Section 4 discusses the key data fields in the L2 (Standard and Diagnostic) product. Section 5 discusses product characteristics and key data fields and provides recommendations for data analysis. Section 6 describes the IMAP-DOAS data product key fields. Section 7 focuses on the calibrated radiances found in the L1BSc product. Section 8 provides full tables of all of the fields in the data products. Section 9 lists tools to view and search the data products. Section 10 lists contact information for questions or issues with the OCO-2 data. Section 11 lists acknowledgements and relevant publications, and the last section lists the abbreviations and acronyms used in this document.

In addition, there are Level 1 and Level 2 algorithm theoretical basis documents (ATBD) that discuss the physics and algorithm details.

## 1.2 Data Products

The primary product delivered by the Orbiting Carbon Observatory-2 consists of spatially resolved estimates of the column-averaged dry air mole fraction. This quantity, called $X_{CO2}$ by members of the atmospheric carbon science community, quantifies the average concentration of carbon dioxide in a column of dry air extending from Earth's surface to the top of the atmosphere. Estimates of $X_{CO2}$ derived by taking the ratio of the column integrated number densities of carbon dioxide and molecular oxygen along the optical path between the Sun, the surface footprint, and the instrument, and then multiplying these results by the column-averaged oxygen concentration (0.20935). These carbon dioxide and oxygen number densities are estimated from high-resolution spectra of reflected sunlight, collected by the Observatory's instrument at wavelengths (colors) within the 0.765 nanometer molecular oxygen A band and two carbon dioxide bands centered at wavelengths near 1.61 and 2.06 microns. The Orbiting Carbon Observatory-2 mission will produce four primary data products for the user community that will provide comprehensive mission results and material for further research and investigation:

- Level 1B—full orbits or fractions of orbits of calibrated and geographically located spectral radiances from the spectral channels centered on the 0.765 micron molecular oxygen A band and the carbon dioxide bands at wavelengths near 1.61 and 2.06 microns. This product contains a unique record of every sounding the instrument collects while viewing Earth during a single spacecraft orbit—approximately 74,000 soundings. Each sounding consists of collocated (observing the same location) spectra from the three spectrometer channels
- Level 2—full orbits or fractions of orbits of geographically located estimates of the column-averaged dry air mole fraction of carbon dioxide (also called $X_{CO2}$) and several atmospheric and geophysical properties collected during each spacecraft orbit. This product typically includes more than 4,000 retrievals of the concentration of atmospheric carbon dioxide in cloud-free scenes, as well as surface pressure, surface albedo, aerosol

content and water vapor and temperature profiles. Estimates of the solar-induced chlorophyll fluorescence are provided in a preprocessor file (L2IDP).

- Level 3—gridded global maps of the atmospheric carbon dioxide concentration, generated monthly by members of the Orbiting Carbon Observatory-2 science team.

The NASA Goddard Earth Science Data and Information Services Center (GES DISC) will archive and distribute the mission's data products. Scientists expect to begin delivering calibrated spectral radiances about three months after the end of the spacecraft's in-orbit checkout. An exploratory atmospheric carbon dioxide concentration product is expected to be available within six months after completing in-orbit checkout.

## 1.3 Data Usage Policy

These data have been produced by the OCO-2 project, and are provided freely to the public. In order to improve our product and have continued support for this work, we need user feedback and user acknowledgement of data usage. Therefore, we request that when publishing using OCO-2 data, please acknowledge NASA and the OCO-2 project.

- Include OCO-2 as a keyword to facilitate subsequent searches of bibliographic databases if it is a significant part of the publication
- Include a bibliographic citation for OCO-2 data. The most relevant citations currently are Wunch et al. [2011a and 2011b], O'Dell et al. [2012], Crisp et al. [2012], and Frankenberg et al. [2014a and 2014b]
- Include the following acknowledgements: "These data 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."
- We recommend sending courtesy copies of publications to the OCO-2 Project Scientist, Michael.R.Gunson@jpl.nasa.gov, and Deputy Project Scientist, Annmarie.Eldering@jpl.nasa.gov

# 2 Mission Overview

Fossil fuel combustion, deforestation, and other human activities are currently adding almost 40 billion tons of carbon dioxide ($CO_2$) to the atmosphere each year [see Canadell et al., 2007; Le Quéré et al., 2014]. These $CO_2$ emissions are superimposed on active natural carbon cycle that emits more than 20 times as much $CO_2$ into the atmosphere each year as human activities, and then reabsorbs a comparable amount, along with about half of the human contributions. Existing ground-based measurements provide a strong global constraint on both human and natural $CO_2$ fluxes into the atmosphere [e.g., Conway et al., 2011]. However, a far more comprehensive measurement network is needed to identify and characterize the strongest natural sources and sinks, and to discriminate the human $CO_2$ emissions from the natural background. Such measurements are essential to any carbon management strategy.

One way to improve the spatial and temporal sampling of $CO_2$ is to retrieve precise, spatially resolved, global estimates of the column-averaged $CO_2$ dry air mole fraction ($X_{CO2}$) from space [see Rayner and O'Brien, 2001; Crisp et al., 2004]. Surface-weighted $X_{CO2}$ estimates can be retrieved from high-resolution spectroscopic observations of reflected sunlight in near infrared $CO_2$ and $O_2$ bands. This is a challenging space-based remote sensing observation because even the largest $CO_2$ sources and sinks produce changes in the background $X_{CO2}$ distribution no larger than 2%, and most are smaller 0.25% [see Crisp et al., 2012].

The European Space Agency EnviSat SCIAMACHY and Japanese Greenhouse Gases Observing Satellite (GOSAT) TANSO-FTS were the first satellite instruments designed to exploit this measurement approach. SCIAMACHY collected column-averaged $CO_2$ and methane ($X_{CH4}$) measurements over the sunlit hemisphere from 2002 to 2012 [Burrows et al., 1995; Buchwitz et al., 2007]. TANSO-FTS has been collecting $X_{CO2}$ and $X_{CH4}$ observations since April 2009 [Hamazaki et al., 2005; Kuze et al., 2009; Yoshida et al., 2011]. These data have provided an excellent proof of concept, and are beginning to yield new insights into the carbon cycle, but improvements in sensitivity, resolution, and coverage are still needed.

The Orbiting Carbon Observatory–2 (OCO-2) is NASA's first dedicated $CO_2$ monitoring satellite. OCO-2 is a "carbon copy" of the Orbiting Carbon Observatory (OCO), which was lost in 2009, when its launch vehicle malfunctioned and failed to reach orbit. A general description of the OCO mission is given in Crisp et al. [2004, 2008]. Most of the information is valid for OCO-2, but a few changes were needed to address known problems, replace obsolete hardware, or address operational changes. Changes that affect the OCO-2 data products are highlighted below.

Like OCO, the OCO-2 spacecraft carries and points a single instrument that incorporates 3, co-boresighted, imaging, grating spectrometers. This instrument collects high-resolution spectra of reflected sunlight in the molecular oxygen ($O_2$) A band, centered near 765 nm, and in the $CO_2$ bands centered near 1610 and 2060 nm. Each spectrometer collects 24 spectra per second, yielding about a million observations each day over the sunlit hemisphere. Coincident measurements from the three spectrometers are combined into "soundings" and analysed with a "full-physics" retrieval algorithm to yield estimates of $X_{CO2}$. Clouds and optically thick aerosols preclude observations of the full atmospheric column in many regions, but this approach is expected to yield over 100,000 full-column $X_{CO2}$ estimates each day.

OCO-2 was sucessfully launched from Vandenberg Air Force Base on a Delta-II 7320 launch vehicle at 2:56 AM PDT on July 2, 2014. The launch vehicle targeted an initial orbit about 15 km below the 705-km Afternoon Constellation (A-Train). Preliminary spacecraft checkout activities were completed during the first week of operations and the onboard propulsion system

has initiated a series of orbit-raising manuevers to insert the satellite at the head of the A-Train during the first week of August. This near-polar (98.2° inclination), Sun-synchronous orbit has a 98.8 minute period, a 16 day/233 orbit ground track repeat cycle and an equator crossing time near 1:36 PM. The OCO-2 orbit track is maintined ±20 km of nominal ground tracks of the CloudSat radar and CALIPSO lidar. Once in this orbit, the optical bench and focal plane array (FPA) detectors in the three-channel imaging grating spectrometer were cooled to its operating temperatures and a comprehensive series of characterization and calibration activities were initiated. The Observatory completed its in-orbit check-out phase in early October and is now routinely collecting almost 1 million soundings over the sunlit hemisphere each day. More than 10% of these measurements are sufficiently cloud free to yield precise estimates of the column-averaged $CO_2$ dry air mole fraction, $X_{CO2}$.

## 2.1 Measurement Approach

As noted above, the OCO-2 spacecraft carries and points a 3-channel imaging grating spectrometer designed to record high-resolution spectra of reflected sunlight in $O_2$ and $CO_2$ bands as it flies over the sunlit hemisphere. A more detailed description of the instrument is provided in the following section. For routine science operations, the instrument's bore sight is pointed to the local nadir or at the "glint spot," where sunlight is specularly reflected from the Earth's surface (Figure 2-1). Nadir observations provide the best spatial resolution and are expected to yield more useful $X_{CO2}$ soundings in partially cloudy regions or topographically round areas over land. Glint observations have much greater signal-to-noise ratios (SNR) over dark, specular surfaces and are expected to yield more useful sounding over ocean. The baseline operating strategy adopted early in the OCO-2 mission alternates between glint and nadir observations on consecutive 16-day ground-track repeat cycles, so that the entire sunlit hemisphere is sampled in both modes at 32-day intervals. OCO-2 can also target selected surface calibration and validation sites and collect thousands of observations as the spacecraft flies

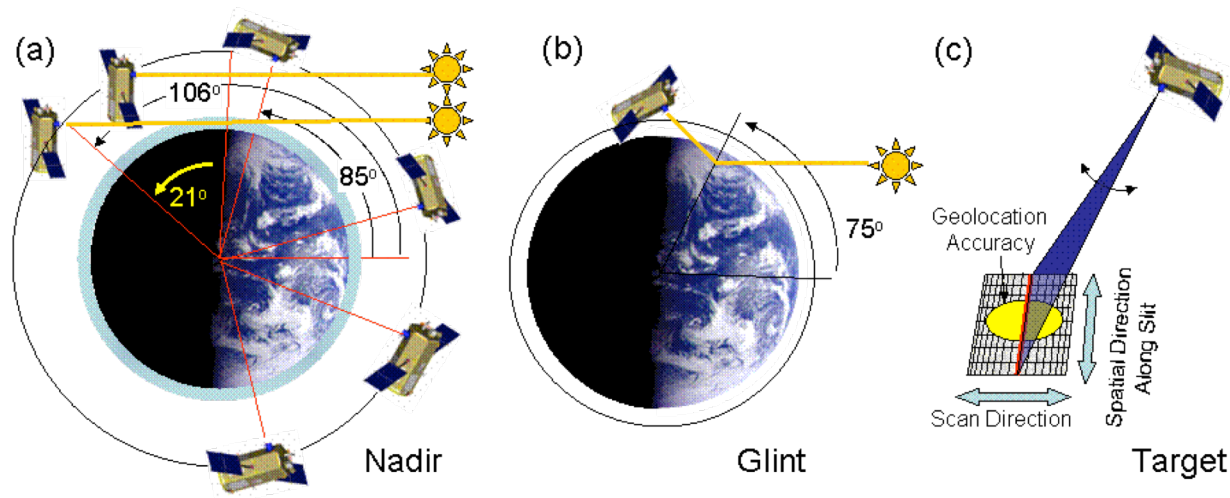



**Figure 2-1.** Nadir, glint, and target observations. (a) Nadir observations are acquired over the sunlit hemisphere at latitudes where the surface solar zenith angle is <85°. On all orbits except downlink orbits, as the Observatory passes over the northern terminator, it pitches up to point the instrument aperture at the Sun for solar radiometric calibrations. (b) Glint observations are made at latitudes where the solar zenith angle at the apparent glint spot is less than ~75°. (c) For target observations, the spacecraft scans the instrument across a stationary surface target as it flies overhead.

overhead. The instrument's rapid sampling, small (< 3 $km^2$) sounding footprint, and high sensitivity, combined with the spacecraft's ability to point the instrument's boresight toward the glint spot over the entire sunlit hemisphere, are expected to yield improved coverage of the ocean, partially cloudy regions, and high latitude continents than earlier missions.

The same rate of data sampling is used for nadir, glint, and target observations. The instrument collects 8 adjacent, spatially-resolved samples every 0.333 seconds (24 samples per second) along a narrow (0.8°) swath. At this data collection rate, it collects ~400 soundings per degree of latitude or ~1 million soundings each day over the sunlit hemisphere. Clouds, aerosols, and other factors will reduce the number of soundings available for $X_{CO2}$ retrievals, but the small sounding footprint ensures that some data will be sufficiently cloud free on regional scales at monthly intervals.

Nadir observations will be collected at all locations where the surface solar zenith angle is less than 85°. Glint soundings will be collected at all latitudes where the solar zenith angle of the glint spot is less than 75°. In this pointing mode, the instrument boresight is not pointed directly at the center of the glint spot, but at a point that is offset toward the local nadir, by an angle that increases with increasing solar zenith angle. Target observations are conducted over OCO-2 validation sites that are within 61° of the local spacecraft nadir along the orbit track and spacecraft viewing angles between +20° of the ground track. When the target is near the ground track, a single pass can last for up to 9 minutes, providing over 12000 soundings in the vicinity of the target. This large number of soundings reduces the impact of random errors and provides opportunities to identify spatial variability in the $X_{CO2}$ field near the target.

While OCO-2 carries an imaging grating spectrometer, it is not operated in a typical "pushbroom" fashion. Instead, for routine nadir and glint observations, the orientation of the spectrometers' 0.8° wide fields of view changes along the orbit track to optimize solar power input and instrument signal to noise. The instrument is installed in the spacecraft such that the spectrometer slits are aligned roughly parallel to axis of rotation of the solar panels. The solar panels produce maximum power when their rotation axis (and the spectrometers' fields of view) is oriented perpendicular to principal plane, which is defined by the Sun, the surface target, and the instrument aperture. (Figure 2-2a).

The Observatory collects nadir observations in the configuration described above. For these observations, as the Observatory ascends into daylight over the southern terminator, its solar panel axis and the spectrometer's field of view is oriented almost perpendicular the orbit track (Figure 2-2b) and the instrument collects data in a near-pushbroom fashion. The footprint dimensions are determined by the cross-track instantaneous field of view (0.1°) and the integration time (0.333 seconds). For nadir observations, this yields 4 to 8 cross-track footprints along the spectrometer slit with dimensions of 1.29 km by 2.25 km. As the Observatory proceeds northward along its orbit, it rotates counterclockwise about the nadir point, such that the solar panel axis and spectrometers' fields of view are aligned with the ground track just north of the sub-solar latitude. At this point, size and shape of each spatially resolved surface footprint is determined by the projected width of the slit (<0.03°) and the 0.333 second exposure time, yielding footprints that are ~0.4 km by 2.25 km. There is spatial overlap between footprints acquired in successive exposures by the spatial elements along the field of view. The principal plane azimuth rotation continues as the Observatory approaches the northern terminator, where the solar panel rotation axis and the spectrometers' fields of view are once again nearly perpendicular to the orbit track.

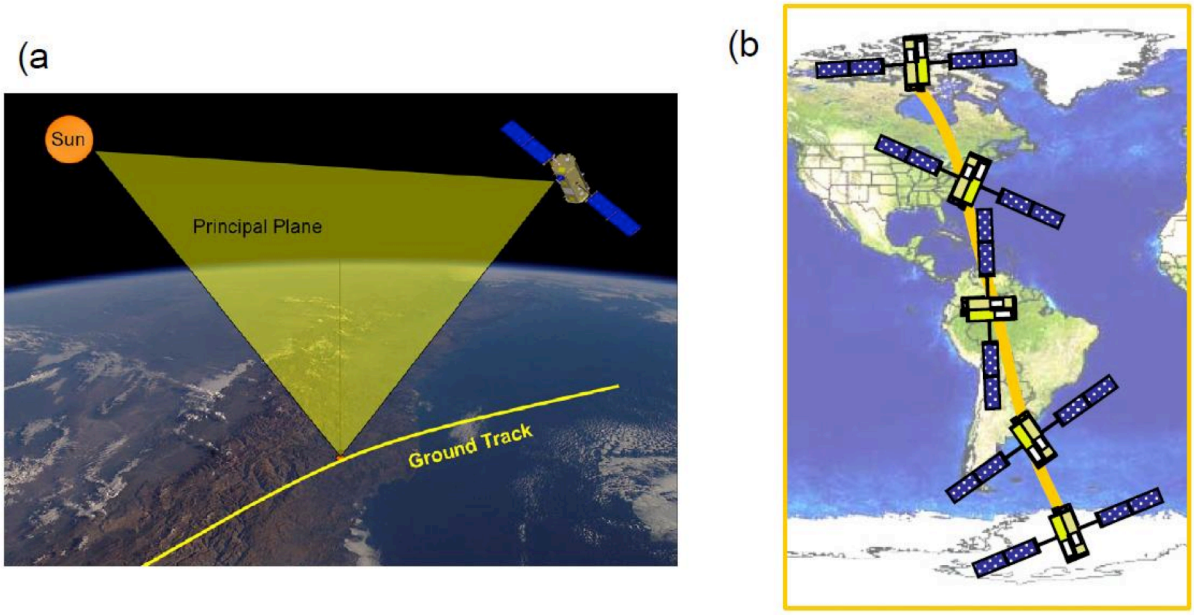
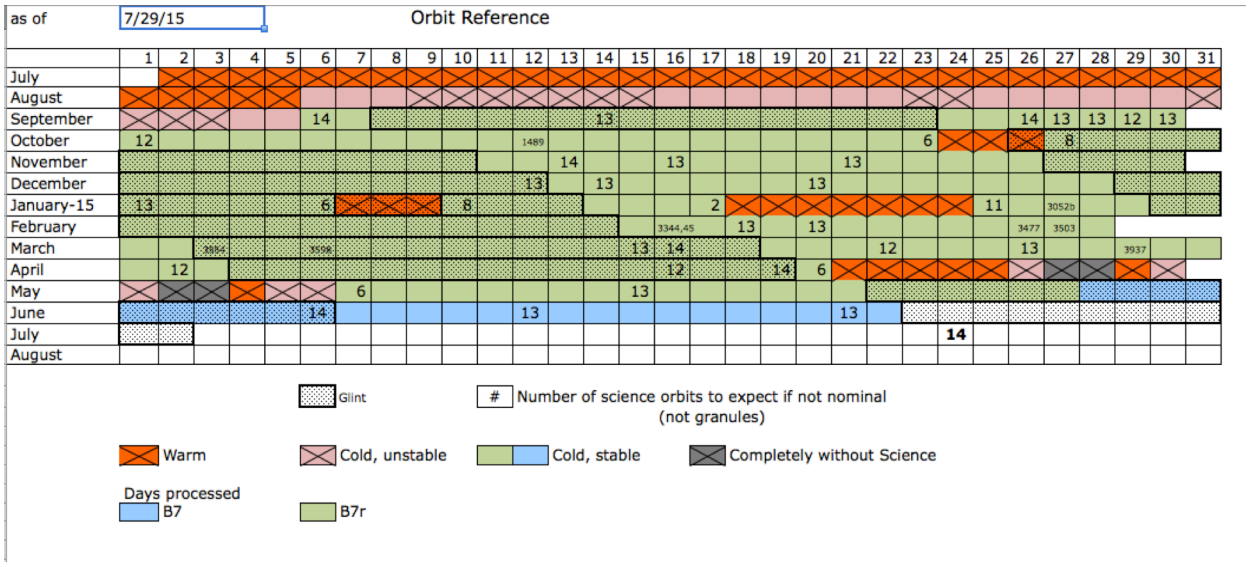



**Figure 2-2. (a)** The principal plane is defined with respect to the Sun, surface footprint, and spacecraft. **(b)** In nadir mode, the spacecraft azimuth changes during the orbit to maintain the alignment of the spectrometer slits (which are roughly parallel with the axis of the solar panels) perpendicular to the principal plane [Crisp et al., 2007]. In glint mode, the spectrometer slits are rotated 30 degrees with respect to the principal plane to increase signal in highly polarized scenes.

For glint observations, the orientation of the instrument's field of view had to be modified to improve the instrument signal to noise ratio (SNR) over the ocean at latitudes where the solar zenith angle is near the Brewster angle (~53°). This was implemented by rotating the field of view 30° clockwise around the bore sight vector (as seen from the spacecraft). This slightly reduces the power available to the solar panels, but yields adequate SNR at all solar zenith angles over the ocean. With this viewing geometry yields a broader cross-track swath south of the sub-solar latitude, and a narrower swath at high northern latitude.

Initially, the observatory was operated in 16 periods, where all observations within a period were in exclusively in nadir or glint. At times, this pattern was disrupted with avoidance manuevers and decontamination cycles. To reduce thermal changes of the observatory, and to gather ocean data more consistently over time, the measurement pattern as changed as of July 2nd, 2015. Beginning on July 2nd, 2015, we gather one orbit of madir, followed by an dorbit in glints, followed by nadir, glint and so on. There is one 16 day pattern that begins in nadir, and a second pattern that beginsin glint (there are an odd number of orbits in 16 days). So, a given orbit track (or path) is measured in nadir, and 16 days later in glint. The calendar below dociuments the mode of data collection over the mission to date.

Science and housekeeping data are transmitted to a NASA Near Earth Network station in Alaska once each day. The data are then transferred to the Earth Science Mission Operations center at the NASA Goddard Space Flight Center (GSFC) where the raw telemetry is converted to time-ordered raw radiance spectra (Level 0 products). This product is then delivered to the OCO-2 Science Data Operations System at the NASA Jet Propulsion Laboratory, where full




orbits are first processed to yield radiometrically calibrated, geolocated spectral radiances (Level 1B products) within the $O_2$ and $CO_2$ bands. The boresighted spectra for each coincident $CO_2/O_2$ sounding are then processed to estimate the column-averaged $CO_2$ dry air mole fraction, $X_{CO2}$ (Level 2 products). Other Level 2 data products to be retrieved from each sounding include the surface pressure, surface-weighted estimates of the column-averaged water vapor and atmospheric temperature, the vertical distribution and optical depth of optically thin clouds and aerosols, the $CO_2$ column averaging kernels, and a number of diagnostic products. Note that these secondary products are all retrieved in order to improve our estimate of $X_{CO2}$, and do not undergo their own validation. Therefore, they should be used with caution for science applications

Once the OCO-2 measurements are calibrated and geo-located, they are analyzed with a "full-physics" remote sensing retrieval algorithm to derive estimates of $X_{CO2}$ and other geophysical properties. This retrieval algorithm is a modified version of the one used to processes GOSAT TANSO-FTS data as part of the NASA Atmospheric $CO_2$ Observations from Space (ACOS) project [see O'Dell et al., 2011; Crisp et al., 2012]. Its principal features and the modifications needed to convert it from GOSAT to OCO-2 are described in Section 5.

To ensure that the $X_{CO2}$ products meet their stringent accuracy requirements, they are validated through comparisons with near-simultaneous measurements of $X_{CO2}$ acquired by ground-based Fourier Transform Spectrometers in the Total Carbon Column Observing Network (TCCON) [Washenfelder et al., 2006; Wunch et al., 2010]. This network currently includes over 20 stations, distributed over a range of latitudes ranging from Lauder New Zealand to Eureka, Canada, and is continuing to add new facilities. To relate TCCON measurements to the WMO $CO_2$ standard, aircraft observations have been collected over several stations, using in situ $CO_2$ measurement approaches used to define that standard. OCO-2 targets a TCCON site as often as once each day, acquiring thousands of measurements as it flies overhead. These measurements will be analyzed to reduce biases below 0.1% (0.3 ppm) at these sites. The space-based $X_{CO2}$ estimates will be further validated through comparisons with $CO_2$ and surface pressure measurements from ground based sites with the aid of data assimilation models to provide a more complete global assessment of measurement accuracy.

For the v7/v7r data, a set of 'lite' files have been created, which each contain a day of data, and have the warn levels and bias correction included. Please see the separate bias correction/lite file document for more details.

## 2.2 Instrument Characteristics

### 2.2.1 Instrument Optical Path

The OCO-2 instrument incorporates three co-boresighted, long-slit imaging grating spectrometers optimized for the $O_2$ A band at 0.765 microns (μm) and the $CO_2$ bands at 1.61 and 2.06 μm. The three spectrometers use similar optical designs and are integrated into a common structure to improve system rigidity and thermal stability. They share a common housing and a common F/1.8 Cassegrain telescope (Figure 2-3). Light entering the telescope is focused at a field stop and then recollimated before entering a relay optics assembly. There, it is directed to one of the three spectrometers by dichroic beam splitters, and then transmitted through a narrowband pre-disperser filter. The pre-disperser filter for each spectral range transmits light with wavelengths within ~1% of the central wavelength of the $CO_2$ or $O_2$ band of interest and rejects the rest. The light is then refocused on the spectrometer slits by a reverse Newtonian telescope.

Each spectrometer slit is about 3 mm long and about 25 μm wide. These long, narrow slits are aligned to produce co-boresighted fields of view that are ~0.0001 radians wide by ~0.0146 radians long. Because the diffraction gratings efficiently disperse only the light that is polarized in the direction perpendicular to the long axis of the slit, a linear polarizer was included in front of the slit to reject the unwanted polarization before it enters the spectrometer, where it could contribute to the scattered light background. Once the light traverses a spectrometer slit, it is collimated by a 2-element refractive collimator, dispersed by a gold-coated, reflective, planar, holographic diffraction grating, and then focused by a 2-element camera lens onto a 2-dimensional focal plane array (FPA), after traversing a second, narrowband filter. The narrowband filter just above the FPA is cooled to ~180 K to reject thermal emission from the instrument.

An onboard calibrator has been integrated into the telescope baffle assembly [Crisp et al., 2008]. This system consists of a calibration "propeller" that carries an aperture cover (lens cap) and a transmission diffuser. The cover is closed to protect the instrument aperture from external contamination during launch and orbit maintenance activities. It is also closed to acquire "dark frames" that are used to monitor the zero-level offset of the FPAs. The back side of the cover has a diffusively reflecting gold surface that can be illuminated by one of 3 tungsten lamps installed in the baffle assembly. The lamp "flat field" images are used to monitor the relative gain of the

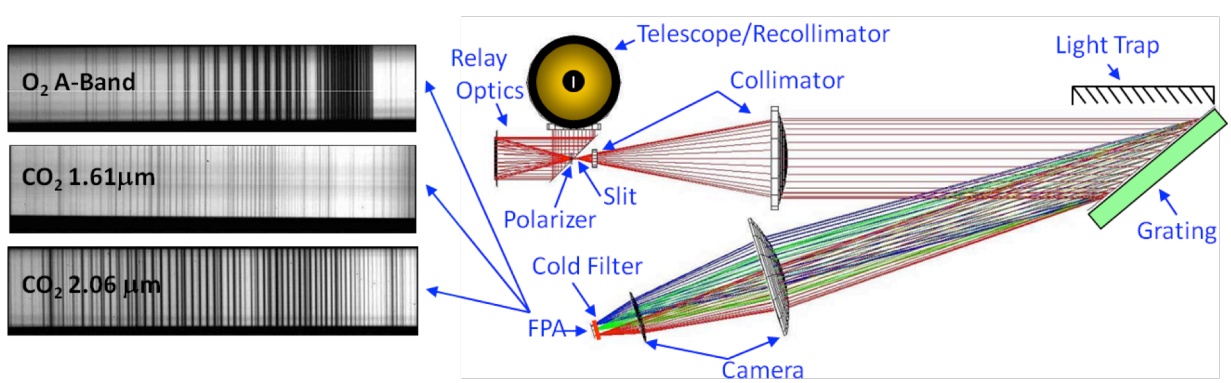



**Figure 2-3.** The OCO-2 instrument showing the major optical components and optical path (right) and images of spectra recorded by the FPA in the three spectral channels (left).

individual pixels on the FPAs. The calibration propeller is rotated 180 degrees from the closed position to place the transmission diffuser in front of the aperture to acquire observations of the Sun. Routine observations of the Sun are acquired just after the spacecraft crosses the northern terminator on all orbits except those that include downlinks. These measurements are used to monitor the absolute radiometric calibration of the instrument. The diffuser is also used to acquire solar spectra for full dayside orbits, which sample the full range of Doppler shifts (± ~7 km/sec) observed over the illuminated hemisphere. The calibration mechanism is rotated 90 degrees from either the closed or diffuser positions for normal science observations.

### 2.2.2 Focal Plane Array Detectors

The spectrometer optics produce a 2-dimensional image of a spectrum on a 1024 by 1024 pixel FPA with 18 μm pixels (Figure 2-3 and Figure 2-4). The grating disperses the spectrum onto 1016 of the 1024 FPA columns (4 columns are blanked out on each side of the FPA) in the direction perpendicular to the long axis of the slit. The full width at half-maximum (FWHM) of the slit image on the FPA is sampled by 2 to 3 pixels in the direction of dispersion. The 3 mm long slit limits spatial field of view to only ~190 pixels in the dimension orthogonal to the direction of dispersion. Science measurements are restricted to the center ~160 of these 190 pixels.

For normal science operations, the FPAs are continuously read out at 3 Hz. A "rolling readout" scheme has been adopted for reading out and resetting the FPAs, precluding the need for a physical shutter and gaps between the exposures. To reduce the downlink data volume and increase the signal to noise ratio, ~20 adjacent pixels in the FPA dimension parallel to the slit

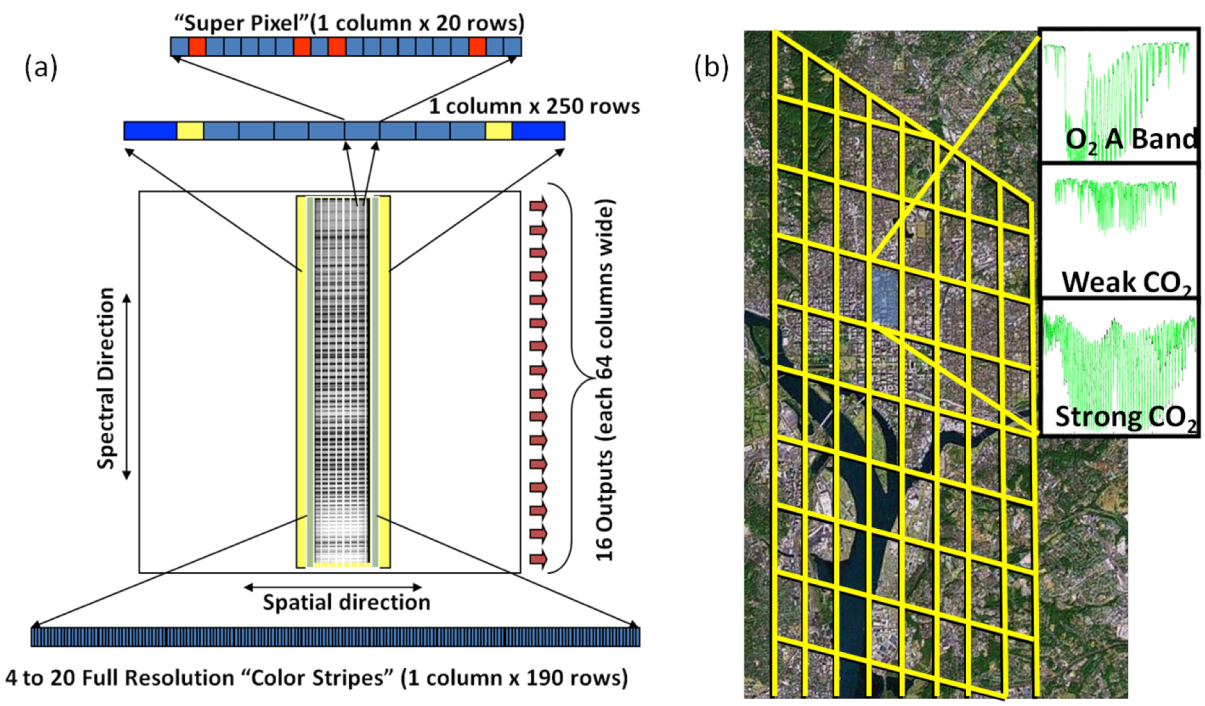



**Figure 2-4. (a)** The illumination and readout scheme used for the OCO-2 focal plane arrays, showing the direction of spectral dispersion from bottom to top, and the spatial dimension from left to right. The ~160 illuminated pixels in the spatial dimension are summed into eight 20-pixel "super pixels" or "footprints." If one or more of the pixels in a footprint is "bad" (red pixels), it is eliminated from the sum. One of the 20 full-resolution "color slices" is also shown at the bottom. **(b)** Spatial layout of 8 cross-track footprints for nadir observations over Washington, D.C. Each footprint is shaped like a parallelogram rather than a square because of the rolling readout of the FPAs.

(i.e., the "spatial direction" in Figure 2-4a) are summed on board to produce up to 8 spatially-averaged spectra along the slit, to produce "spectral sample" for a "summed footprint."

The along-slit angular field of view of each of these spatially averaged spectral samples is ~1.8 milliradians (0.1° or ~1.3 km at nadir from a 705 km orbit). The angular width of the narrow dimension of the slit is only 0.14 milliradians, but the focus of the entrance telescope was purposely blurred to increase the effective full width at half-maximum of each slit to ~0.6 milliradians to simplify the boresight alignment among the 3 spectrometer slits. Because it takes 0.333 seconds to scan across the active 220 rows of the array, the surface footprints of the 8 spatially-averaged spectra are read out at 0.03 second intervals, and are shaped liked parallelograms rather than squares, even when the slit is oriented orthogonal to the ground track (Figure 2-4b).

In addition to the 8 spatially-binned, 1016-element spectra, each spectrometer returns up to 20 columns from each FPA without any onboard spatial binning to sample the full along-slit spatial resolution. Each of these full-resolution "color slices" images a 220-pixel wide region of the FPA that includes the full length of the slit (~190 pixels) as well as a few pixels beyond the ends of the slit (Figure 2-4a). These full spatial resolution "color slices" are used to detect spatial variability within each of the spatially summed super pixels and to quantify the thermal emission and scattered light within the instrument. Their locations can be specified by commands from the ground..

In this instrument design, the spectrometer slits, the grooves on the diffraction gratings, and columns of the FPAs must be well aligned to ensure that a fixed series of rows on the FPA will sample the same angular field of view (or spatial footprint) throughout the spectral range recorded by the FPA. For the OCO-2 instrument, perfect alignment of the FPAs with the other optical components was not possible due to a physical obstruction discovered late in the instrument assembly process. The focal plane arrays are therefore slightly rotated (or "clocked") with respect to the slit and grating. Consequently, a given geographic position does not map onto a single row of pixels on a sensor, but instead varies (roughly linearly) with spectral position (i.e., column). To compensate for this, and record the same spatial information across the entire spectrum, the starting row index for each spectral sample can be adjusted in increments of one pixel. This corresponds to about 1/20th of a summed footprint (Figure 2-5). This approach introduces little error in spatially homogenous scenes, but can produce discontinuities in spectra of scenes with strong intensity variations near the edge of a footprint. An algorithm to correct for these discontinuities has been implemented and forms part of the calibration process (see the Level 1B Algorithm Theoretical Basis Dociment).

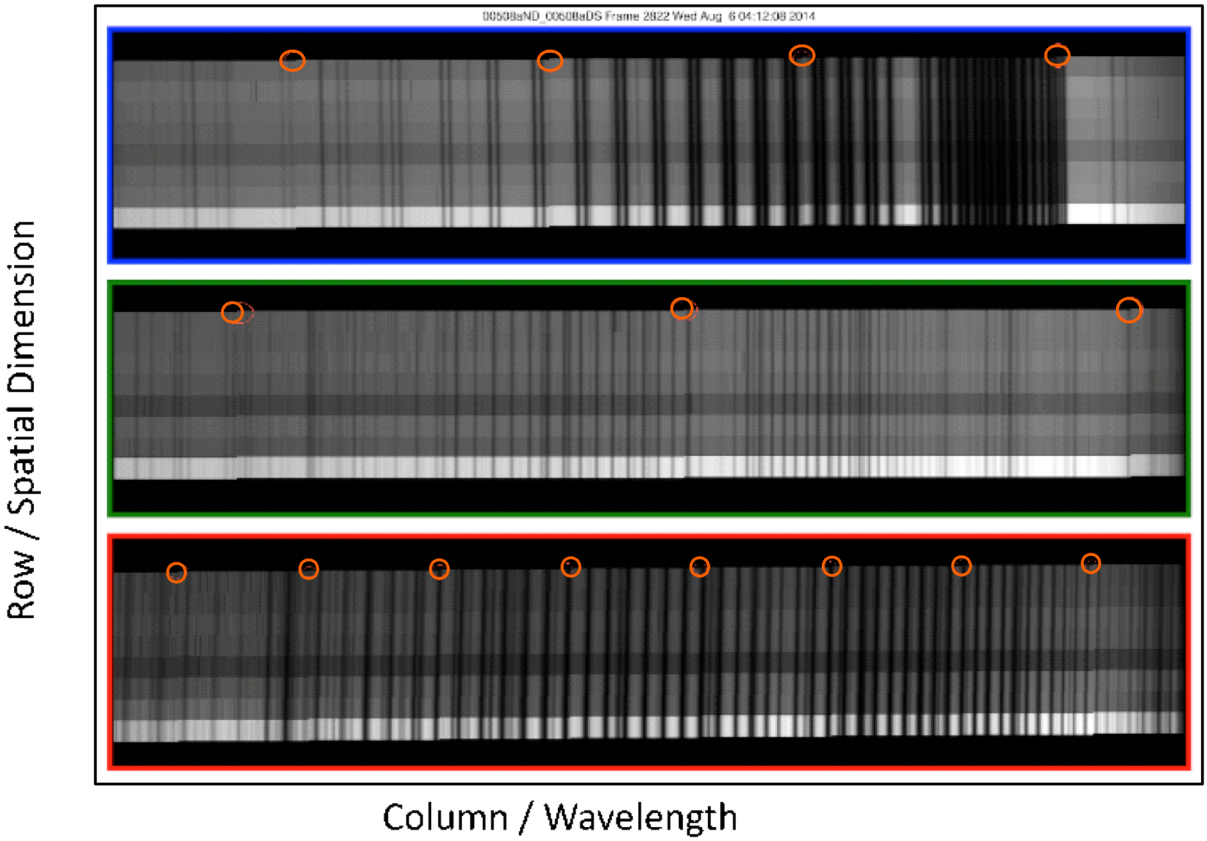



**Figure 2-5.** Images of the $O_2$ A band **(top)**, 1.61 μm $CO_2$ band **(middle)** and the 2.06 μm $CO_2$ band **(bottom)** in "summed mode" taken from the OCO-2 first light frame over Papua New Guinea. The 8 spatially-summed footprints in each band are shown from bottom (footprint 1) to top (footprint 8). In this 0.333-second frame, footprint 1 was contaminated by a cloud and appears brighter than the rest. This frame clearly shows the rotation, or "clocking" of each FPA columns with respect to the dark $O_2$ and $CO_2$ absorption lines. The $O_2$ lines are tilted slightly counterclockwise, while the 1.61 and 2.06 μm lines are tilted clockwise. The columns where the clocking corrections are applied are highlighted with red circles. Soundings collected in regions with strong spatial gradients in illumination can produce radiance discontinuities at these wavelengths.

### 2.2.3 Bad Pixels and Bad Samples

OCO-2 is flying flight-spare FPAs from the OCO mission that were delivered in 2006. A small fraction of the 220,000 pixels in the active area of these arrays are either dead or respond to light or thermal changes in a way that is not consistent with the majority of the other pixels. These pixels must be identified and excluded from the 20-pixel sums that are performed onboard, or they will contaminate the resulting 20-pixel "spectral samples" that constitute the spectra returned for the 8 footprints by each spectrometer. Spectral samples with too many bad pixels, or with other known issues (e.g., optical artifacts on the short-wavelength ends of all 3 bands) are marked as bad in the Level 1B product, so that they are not used in producing higher level products.

The bad pixel identification process was started during prelaunch testing, but must be updated on orbit because additional pixels degraded during the 2.5 year instrument storage period prior to launch, and others are damaged on-orbit by cosmic radiation and other effects. To

identify bad pixels, the calibration team routinely collects dark and lamp data using "single pixel" mode, which returns all 220,000 pixels in the active region of each array. (Single-pixel results cannot be returned routinely for science observations because it requires ~9 seconds to read out an FPA at full resolution.) These single-pixel data are then analyzed to identify bad pixels and add them to a "bad pixel map" that is uploaded to the spacecraft for use in the pixel-summing process. As the bad pixel maps improve, a larger fraction of the spectral samples in each spectrum can be used to generate Level 2 products.

### 2.2.4 Cosmic Ray Artifacts

Cosmic rays rarely produce permanent damage to the OCO-2 FPAs, but this radiation produces ion trails as they traverse the FPAs, producing spurious intensity spikes. Primary and secondary cosmic rays with a broad range of energies are occasionally seen just about anywhere along the orbit path but they are most common in the vicinity of the South Atlantic Anomaly (SAA), where up to 2% of an $O_2$ A band spectral range can be contaminated in individual spectra. The approximate geographical extent of the SAA as it impacts the OCO-2 mission can be seen in Figure 2-6. The characteristic spectral signature of cosmic rays is shown in Figure 2-7.

An algorithm has been developed to identify and screen individual spectral samples that have been contaminated by cosmic rays in the L1B data delivered to the science community. This algorithm is summarized in the OCO-2 L1B ATBD.

This algorithm compares each spectral sample for each footprint in each band to a synthetic spectrum generated through a least squares fitting process (see the L1B ATBD). Spectral samples with large positive residuals in these comparisons are assumed to be contaminated by cosmic rays. The primary products of this algorithm are noise-weighted fit residuals for each spectral sample, footprint, and band for each sounding. In the L1B HDF5 files, these weighted residuals are saved in the group SpikeEOF as rounded integer values, with all residuals less than three times the noise level (3 sigma) set to 0 to minimize file size. The variables used to record the weighted residuals in the L1B files are:

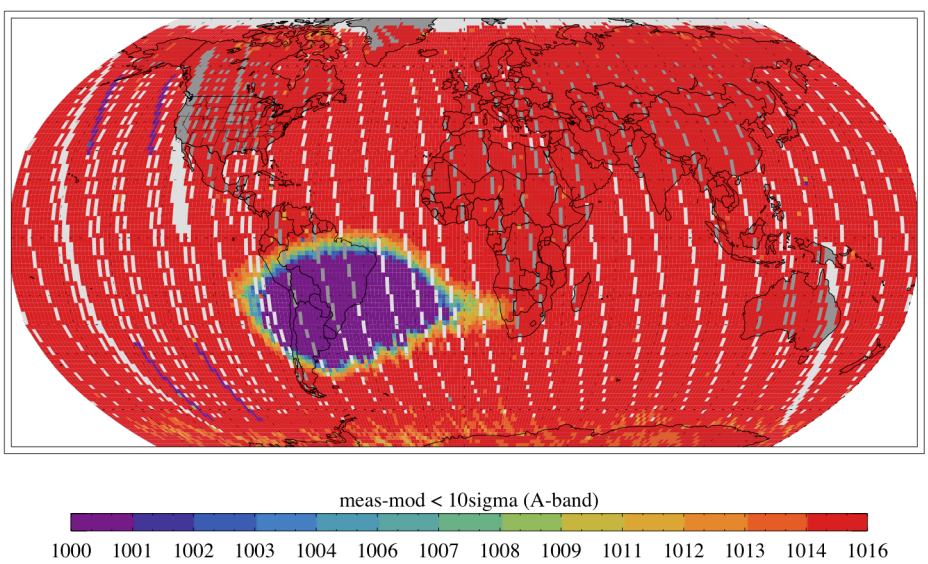



**Figure 2-6.** The number of spectral samples in the $O_2$ A band that are not contaminated by 10-sigma cosmic rays events clearly shows the geographical extent of the South Atlantic Anomaly (SAA).

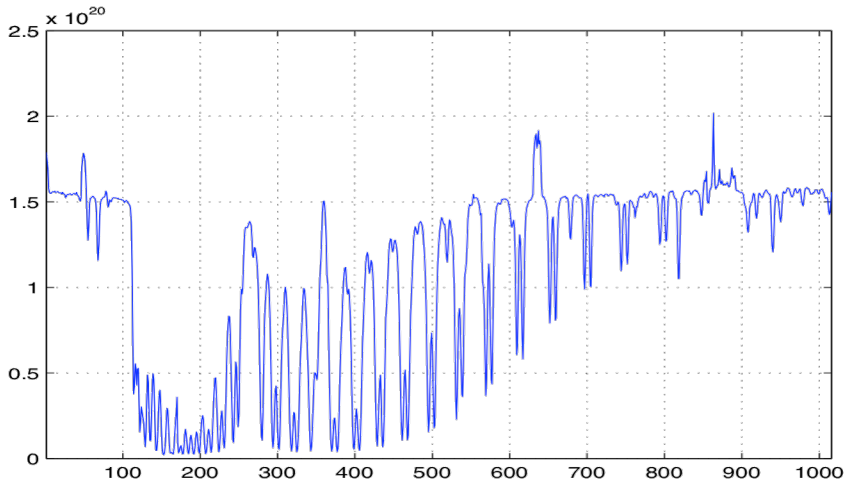



**Figure 2-7.** An example showing the impact of a radiation event on spectra. Note the three spikes with unusually high intensities.

SpikeEOF/spike_eof_weighted_residual_*band* (with *band* = o2, strong_co2, weak_co2)

The total number of spectral samples contaminated in by cosmic rays in any one spectrum are recorded in the variables:

SpikeEOF/spike_eof_bad_colors_*band* (with *band* = o2, strong_co2, weak_co2)

The OCO-2 L2 FP algorithm uses the weighted residuals in the L1B file to flag individual spectral samples as bad if their weighted residuals exceeds +6 sigma, and if the sounding falls within a geographic boundary box around the SAA (latitude −50 to 0, longitude −90 to 10 degrees). We use a one-sided threshold value as cosmic ray hits can only cause a positive anomaly in measured spectra. L1B users are advised adopt a similar approach for flagging spectral samples contaminated by cosmic rays before attempting any fit. We also advise users to apply this threshold only in the vicinity of the SAA.

### 2.2.5 Oxygen Spectrometer Sensitivity

Observations with OCO-2's onboard solar calibration system have indicated a degradation in the sensitivity of the oxygen spectrometer. This has been confirmed with the monthly lunar observations (see Figure 2-8). Since the lunar calibration uses the optical, detector and signal processing systems exactly like they are used for science observations, this indicates that the change is in the instrument rather than the solar diffuser.

Over the course of the mission, the relative response has dropped by as much as 15% from the prelaunch characterization. Each time the focal planes have been warmed up (known as a decontamination cycle), the loss of sensitivity has been almost, but not quite, completely reset to its prelaunch value. While these temperature cycles are called decontamination cycles, the project has not yet determined the root cause of this problem and is continuing to investigate several possibilities. It may very well turn out the loss of signal is not entirely or even partially due to contamination.

In the meantime, the loss of signal is being monitored with the ~6x per day solar calibration measurements and lunar calibrations at ~75% full (waxing) observations. These lead to adjustments in the calibration coefficients associated with:

- Conversion between the digital measurements returned by the instrument and the radiance reported in the L1B files
- Signal-to-noise coefficients reported in the L1B files and used to calculate the uncertainty for a given radiance

While the signal degradation is a smooth function in time, the OCO-2 data processing pipeline does not allow these parameters to include a temporal component. To compensate, the calibration coefficients will be delivered in small increments of time such that the curve will be simulated with a set of piecewise constant values. Initially, it is expected that the time increments will be on order of a few days during which time the change in sensitivity will be less than ±0.1%.

As the project more fully characterizes this degradation, later reprocessing of the L1A to L1B conversion will likely reduce any residual errors associated with this correction.

## 2.3 OCO-2 Algorithms

### 2.3.1 Level 1B Algorithm Overview

The OCO-2 Level 1B algorithms apply radiometric, spectroscopic, and geometric calibration corrections to the data number (DN) returned by the OCO-2 instrument to produce the calibrated, geo-located spectral radiances reported in the Level 1B product. The radiometric and spectroscopic calibration methods were initially developed using data collected during prelaunch testing [see Frankenberg et al., 2014 and references therein]. After launch, the radiometric

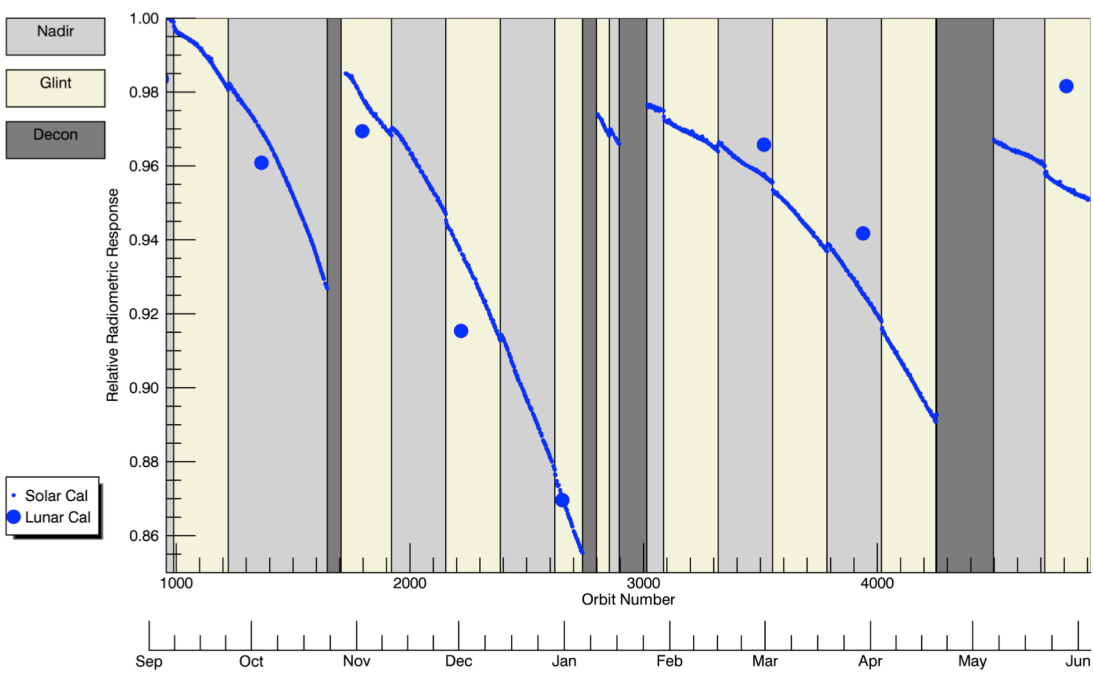



**Figure 2-8.** Median response of the $O_2$ spectrometer to solar observations - squares represent the roughly once per month lunar observations which agree with the solar observations to about ±1%. Corrections to the solar data include Sun-Earth distance and a 0.2% offset for nadir versus glint orbits due to slight changes in the spacecraft thermal environment. Corrections to the lunar data include Sun-Moon distance, Moon-Earth distance and corrections for albedo due nutation and libration effects.

calibration has been validated primarily by the solar calibrations described in section 2.2.5. Additional constraints are being established with lunar calibration using the USGS ROLO database [see Kieffer and Stone, 2005]. While the error bars are large for absolute calibrations (5-10%), the relative calibration confirms the trends seen in the solar calibration approach. The FPA zero-level offset is being validated via regular dark measurements. The spectral calibration was established primarily through prelaunch testing, but is being monitored using Solar Doppler observations and validated against synthetic atmospheric spectra generated by the OCO-2 Level 2 algorithm. No systematic changes have been found over the preflight calibration. The spatial calibration is being validated using observations of the Moon, with additional information provided by observing coastline crossings. Finally, there is a further algorithm that computes the polarization sensitivity of the instrument for each sounding. As stated above, the OCO-2 instrument has a linear polarizer in the optical path to each detector, such that only a single linear polarization is accepted for each spectral band. The polarizers for each spectral band are approximately parallel. To take into account the polarization sensitivity, one must know this polarization angle or associated Mueller matrix coefficients (referred to for OCO-2 as "stokes coefficients"). These algorithms, including the instrument polarization description, are described in detail in the OCO-2 Level 1B Algorithm Theoretical Basis Document.

### 2.3.2 Level 2 Algorithm Overview

The full physics (FP) $X_{CO2}$ retrieval algorithm was derived from the algorithm developed for the OCO. It was further refined in the time between the OCO and OCO-2 launches by use in producing the Atmospheric $CO_2$ Observations from Space data product. The algorithm is a Rodgers [2000]-type optimal estimation approach and has been described fully in O'Dell et al. [2012]. The retrieval algorithm consists of a forward model, an inverse method, and an error analysis step. The overall flow for the retrieval process is shown in Figure 2-9.

A forward radiative transfer model is used to generate synthetic spectra within the molecular $O_2$ A band at 0.76 μm and the weak and strong $CO_2$ bands centered near 1.61 and 2.06 μm, respectively. These synthetic spectra are then convolved with the OCO-2 instrument line shape and compared to the observed spectra in each of these spectral regions. An inverse model then modifies the assumed atmospheric state to improve the fit to the measured spectra, and the process is repeated until the convergence criteria are met. The forward radiative transfer model contains components simulating the solar spectrum, absorption by $CO_2$, $O_2$, $H_2O$, and other gases, scattering and absorption by clouds and aerosols, reflectance of the surface. Input to the forward model consists of meteorological conditions, surface properties, characteristics of the instrument, etc. The forward model returns simulated radiance spectra and the partial derivatives of those radiances with respect to properties of the atmospheric and surface state, also called Jacobians.

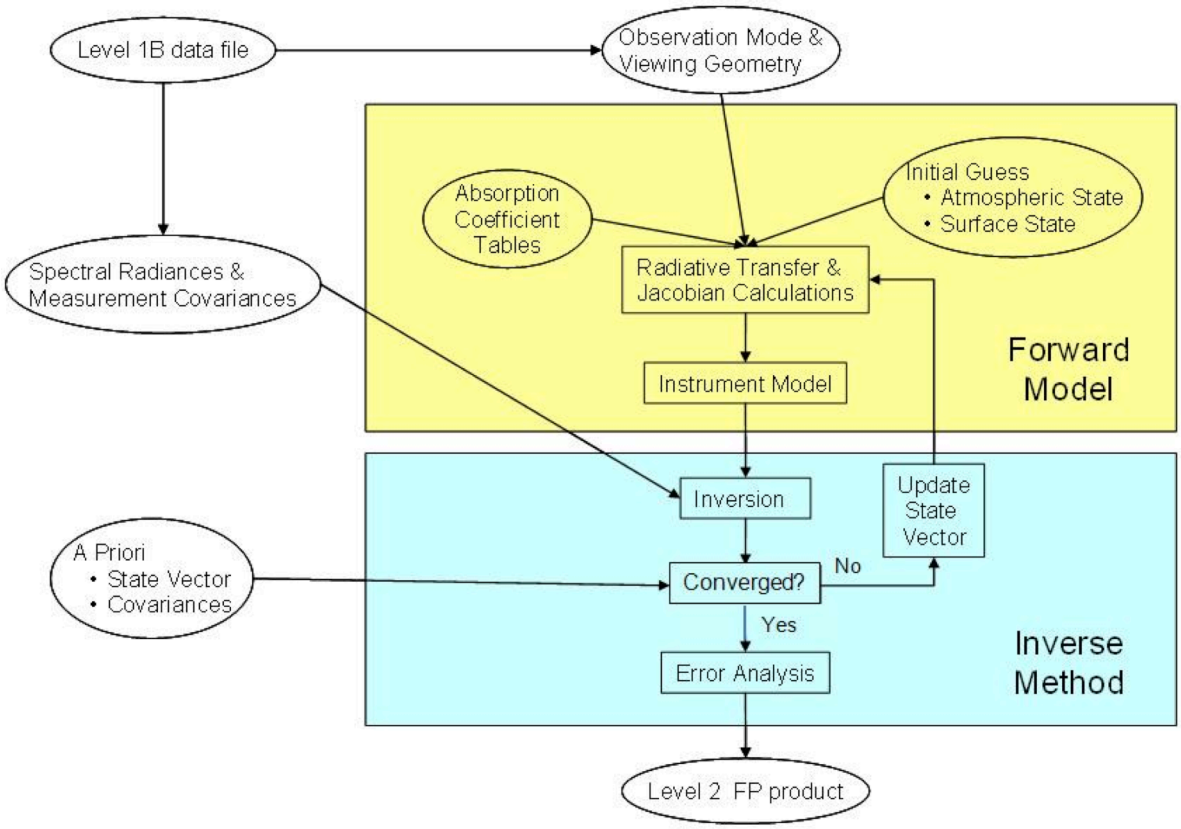



**Figure 2-9.** Level 2 full physics retrieval flow.

The residuals between the simulated and measured spectra are minimized by changing the properties of the atmospheric and surface state via the inverse method. This inversion uses the Jacobians to estimate the state changes needed to minimize the differences between the observed and simulated spectra.

Once the atmospheric state yielding the best match to the observed spectrum has been found, the algorithm then determines $X_{CO2}$, errors in $X_{CO2}$ from different sources (such as vertical smoothing, measurement noise, etc.), and the $X_{CO2}$ column averaging kernel. This is necessary because $X_{CO2}$ is not itself an element of the state vector. Rather, it is determined from the profile of $CO_2$, which is part of the state vector. It is formally given by the total number of $CO_2$ molecules in the column divided by the total number of dry air molecules in the column. This step is labeled "Error Analysis" in Figure 2-9.

# 3 Overview of Data Products

All OCO-2 product files are in Hierarchical Data Format-5 (HDF), developed at the National Center for Supercomputing Applications (http://www.hdfgroup.org). This format facilitates the creation of logical data structures by organizing the data product into folders and subfolders. Each file corresponds to one orbit contiguous mode set of data. Spectral radiances are contained in the SoundingMeasurements folder of the L1BSc product; retrieved $X_{CO2}$ values are contained in the RetrievalResults folder of the L2Std product. Official data products are stored at the NASA Goddard DAAC; user-customized Level 2 and Level 3 data products are available through the $CO_2$ web site (co2.jpl.nasa.gov). Sample tools for reading the data can be found on the $CO_2$ site.

We have lite files – at co2.jpl.nasa.gov (for XCO2 and SIF). There is a separate document for these. Note that they have the advantage of bing one file pr day, and the bias correction and warn levels are included.

Will get posted at the DAAC as soonas we can manage.

## 3.1 File Naming Conventions

The OCO-2 L1BSc files follow this convention:

oco2_L1bSc[Mode]_[Orbit][ModeCounter]_[AcquisitionDate]_[ShortBuildId]_[Production DateTime][Source].h5

The OCO-2 L2Std files follow this convention:

oco2_L2Std[Mode]_[Orbit][ModeCounter]_[AcquisitionDate]_[ShortBuildId]_[Production DateTime] [Source].h5

The OCO-2 L2Dia files follow this convention:

oco2_L2Dia[Mode]_[Orbit][ModeCounter]_[AcquisitionDate]_[ShortBuildId]_[Production DateTime] [Source].h5

The OCO-2 L2IDP files follow this convention:

oco2_L2IDP[Mode]_[Orbit][ModeCounter]_[AcquisitionDate]_[ShortBuildId]_[Production DateTime][Source].h5

For all files, the fields are defined as below.

- [Mode] is the acquisition mode as a two character string:
  - GL—Sample Glint
  - ND—Sample Nadir
  - TG—Sample Target
  - DS—Sample Dark Calibration
  - LS—Sample Lamp Calibration
  - SS—Sample Solar Calibration
  - BS—Sample Limb Calibration
  - NP—Single-Pixel Nadir
  - GP—Single-Pixel Glint

  - TP—Single-Pixel Target
  - DP—Single-Pixel Dark Calibration
  - LP—Single-Pixel Lamp Calibration
  - SP—Single-Pixel Solar Calibration
  - BP—Single-Pixel Limb Calibration
  - XS—Sample Transition
  - XP—Single-Pixel Transition
  - MS—Sample Lunar Calibration
  - MP—Single-Pixel Lunar Calibration
- [Orbit] is the five-digit orbit number
- [ModeCounter] is a letter (a, b, c, d) denoting the times an acquisition mode occurred in an orbit. If a mode occurs only once, ModeCounter is set to "a"
- [AcquisitionDate] is the UTC date (yymmdd) the data were acquired
- [ShortBuildId] identifies the L1B build version used (Bn.m.uu) where n is the major version, m is subversion number, and the uu is the incremental/patch number
- [ProductionDateTime] is the date and time the file was produced (yymmddhhmmss)
- [Source] (if present) identifies production sources different from the standard operations pipeline. This field will be missing from normal pipeline data

### 3.1.1 File Format and Structure

The OCO-2 product files contain data structures indexed by sounding (1 to N soundings/file) and are associated by the sounding_id variable in all products.

Variables are combined into groups by type (e.g., SoundingGeometry). Within each type, a variable has one or more values per sounding. Variables may be single-valued (e.g., sounding_altitude) or multi-valued (e.g., co2_profile).

The metadata of each variable describes the variable's attributes, such as dimensions, data representation, and units.

Note that many variables in the L2 products use _fph to denote full physics algorithm, _idp for IMAP-DOAS, and _abp for the $O_2$ A band preprocessor. For example, surface pressure is calculated in each of the algorithms, so it is reported with a tag on the variable to differentiate them.

### 3.1.2 Data Definition

The OCO-2 data products contain many variables with a variety of dimensions. The following list describes only the most important of the dimensions. Dimensions and data shapes are fully described within the HDF files.

- Retrieval—the number of retrievals reported (those soundings for which retrievals converged or were converging when the maximum number of iterations was reached)
- Band—the three bands of OCO-2 are $O_2$ A, weak $CO_2$, and strong $CO_2$
- Footprint—the eight footprints across the swatch are identified as 1 to 8
- Sample—the spectral element. Each band has 1016 spectral elements, although some are masked out in the L2 retrieval
- Sounding—one set of measurements (one footprint across three bands) that is the primary unit for retrievals

### 3.1.3 Global Attributes

In addition to variables and arrays of variables, global metadata is stored in the files. The granule-level metadata is described in Table 8-1. ECHO metadata and other metadata related to HDF version and production location can be found in the HDF file but are not discussed here.

## 3.2 File Content

**T**he relationship of the overall data flow and retrieval processes to the product files is illustrated in **Error! Reference source not found.**.

### 3.2.1 L2Std

Geolocated retrieved $CO_2$ column-averaged dry air mole fraction - physical model. This file contains the retrieved values for the state vector as well as geolocation information. A number of fields from the iterative maximum *a posteriori* differential optical absorption spectroscopy (IMAP-DOAS) and $O_2$ A-band (ABO2) cloud preprocessors are brought forward to the L2 products.

### 3.2.2 L2Dia

This file has the data that is in the L2Std, as well as averaging kernels, *a priori* covariance matrices, posterior covariance matrices, and measured and model radiances.

### 3.2.3 L2IDP

There are the results from the IMAP-DOAS process. The IMAP-DOAS process is used in cloud screening, but also provides solar induced fluorescence (SIF) measurements for a much larger set of data than contained in L2.

### 3.2.4 L1BSc

These are the radiance spectra that are the input to the L2 retrievals.

# 4 OCO-2 L1BSc Data Products

The OCO-2 L1BSc products contain the calibrated spectra that are used as input for the $X_{CO2}$ products (Figure 4-1). These files contain all acquired data, and consequently are quite voluminous. Additionally, complete geolocation information for all of the footprints as well as some instrument and calibration information can be found the L1BSc products.

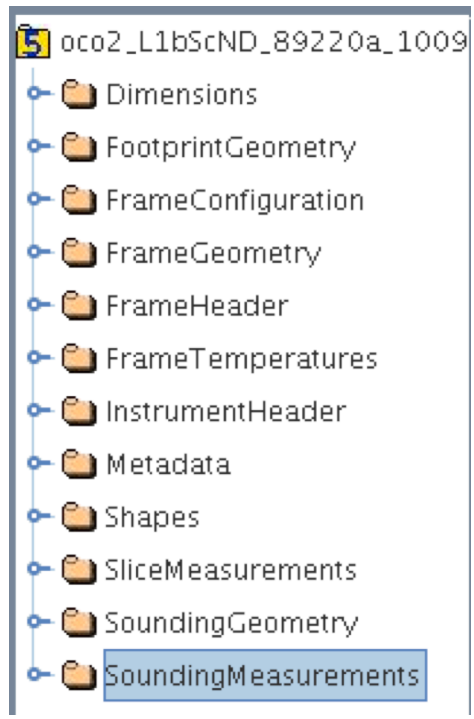



**Figure 4-1.** Folders in the L1BSc product.

## 4.1 Key Data Fields

### 4.1.1 FootprintGeometry

Geolocation uncertainty is <1 km (3σ) in Nadir and Glint modes. Uncertainty may be slightly larger in Target mode; data is still being collected to produce this estimation. Uncertainty estimates will be updated as the mission progresses and more data is obtained.

#### 4.1.1.1 Footprint_latitude

Note that these are for each band, unlike the retrieval_latitude, which is for a sounding.

#### 4.1.1.2 Footprint_longitude

Note that these are also for each band, using a −180 to +180 convention.

### 4.1.2 SoundingMeasurements

#### 4.1.2.1 Radiance_o2

This is reported for all 1016 columns in each of the eight footprints. Note that a field of the SNR coefficients will indicate if a sample should be used or not.

#### 4.1.2.2 Radiance_strong_co2

The radiance array for the strong $CO_2$ band.

#### 4.1.2.3 Radiance_weak_co2

Radiance array for the weak $CO_2$ band.

### 4.1.3 InstrumentHeader

#### 4.1.3.1 Dispersion_coef_samp

Coefficients that express the relationship between the spectral element index and its associated wavelength. Note that this grid does not account for the Doppler correction or dispersion adjustments that are applied in L2. These data are held in a 3 x 8 x 6 array for each band, footprint, and coefficient [Lee et al., 2015]. The coefficients are used as follows:

$$\lambda = \sum_{i=0}^{5} c_i \cdot column^i$$

where column refers to the column number in the L1BSc files (1 to 1016).

An example calculation of the wavelength grid is:

$$\lambda = 0.757633 + 1.75265\times10^{-5}\cdot column^1 \\ -2.91788\times10^{-9}\cdot column^2 + 3.29430\times10^{-13}\cdot column^3 \\ -2.72386\times10^{-16}\cdot column^4 + 7.66707\times10^{-20}\cdot column^5$$

#### 4.1.3.2 SNR_coef

The SNR coefficient contains three parameters. Two are used in the SNR calculation as described below. The third parameter is used to identify bad samples that should be excluded by the retrieval algorithms. The third entry can range between 0 and 15 and is the sum of the following byte codes:

- 0 = Good
- 1 = Radiometric Problem (e.g., jumping pixel that needs to be identified for future bad pixel map, linearity problems)
- 2 = Spatial Problem (e.g., low wavelength 90 columns)
- 4 = Spectral Problem (e.g., ILS not trustworthy—last couple of samples on the $CO_2$ bands where the mask shadows the FPA)
- 8 = Polarization Problem (no examples yet)

***Calculating noise equivalent radiance.*** The noise values are not stored directly in the file, but they can be calculated using a few fields in the file and the following formula for the noise equivalent radiance:

$$NEN = \frac{MaxMS}{100}\cdot\sqrt{\left|\frac{100\cdot N}{MaxMS}\right|\cdot C_{photon}^2 + C_{background}^2}$$

where $N$ is the radiance value (found in SoundingMeasurements), $MaxMS$ is the maximum measurable signal per band (see Table 4-1), $C_{photon}$ is the first coefficient of InstrumentHeader/snr_coef (zero-based indices [0,*,*,*]), and $C_{background}$ is the second coefficient of InstrumentHeader/snr_coef ([1,*,*,*]).

***Calculating SNR.*** The signal to noise ratio can be calculated using the following formula:

$$SNR = \sqrt{\frac{100\,N^2}{MaxMS * \left(C_{background}^2\frac{MaxMS}{100} + C_{photon}^2 N\right)}}$$

where $N$, $MaxMS$, $C_{photon}$, and $C_{background}$ are as defined above. The third entry of InstrumentHeader/snr_coef (zero-based indices [2,*,*,*]) is used to identify bad samples that should be excluded by retrieval algorithms. The third entry will be 1 if the sample is bad and 0 if it is good.

**Table 4-1.** Maximum measurable signal per band.

| Band | MaxMS value (photons/m2/sr/µm) |
|---|---|
| $O_2$ A band | $7.00*10^{20}$ |
| Weak $CO_2$ | $2.45*10^{20}$ |
| Strong $CO_2$ | $1.25*10^{20}$ |

#### 4.1.3.3 ils_relative_response and ils_delta_lambda

For each band, footprint, and spectral element index (3 x 8 x 1016), there are two 200-element lookup tables: ils_delta_lambda and ils_relative_response. The details of the ILS prelaunch determination are reported in Lee et al. [2015]. These curves describe the response of each spectral element of the instrument versus wavelength, and can be used to convolve high spectral resolution spectra for comparison with OCO-2 spectra. Figure 4-2 illustrates one of the instrument line shapes (ILS).

### 4.1.4 SliceMeasurements

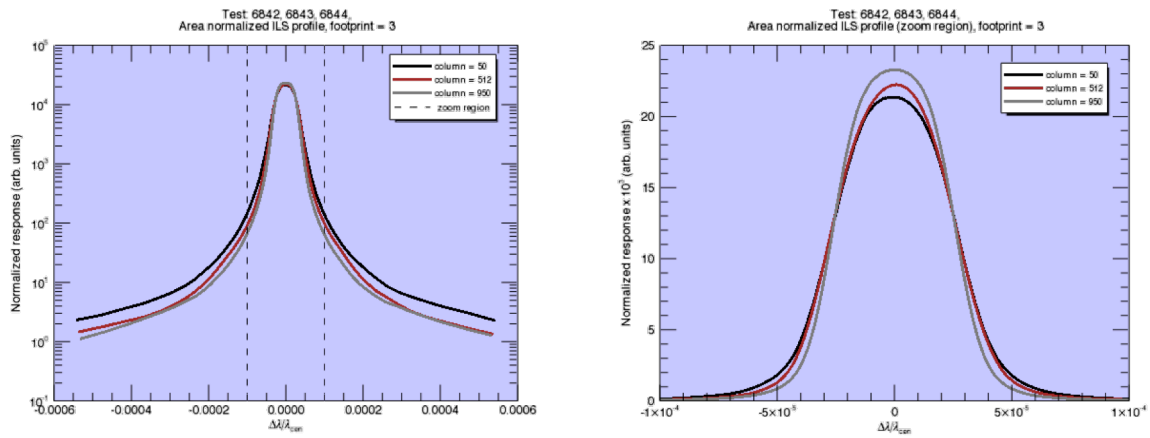



**Figure 4-2.** An example of the instrument line shapes.

The L1B file includes 19 wavelengths where the data are returned with the full spatial resolution of the focal plane (i.e., no on board averaging). This leads to spatial sampling of about 10 - 60 meters cross-track by about 2,200 meters along-track. The spatial resolution is lower in the cross track direction by a factor of 3 to 4 due to optical aberations induced in the telescope to aid in the correction of bad pixels. While the array contains 220 spatial elements, they are only valid over the range of the footprint limits for the given band. Note that the bad pixel map is not applied to the color slices, so there are a handful of invalid radiances—any point that shows a very high deviation from the average of its neighbors should be viewed with skepticism. The color slice radiance values are contained in the variables radiance_slice_o2, radiance_slice_weak_co2, and radiance slice_strong_co2.

Once the wavelengths for a given footprint are calculated, the one-based wavelength index for a color slice will be the same as the value of "color_slice_position_[band name]" – 3. (The "color_slice_position_[band name]" variables may be found in the "FrameConfiguration" folder.) Any index less than 1 or greater than 1016 addresses a part of the focal plane that is not light sensitive and should be ignored.

To determine which an element of the color slice corresponds to which footprint, the following steps should be applied to generate a 1024 x 220 x 8 x 3 binary array in which a value of 1 indicates that a pixel is in that footprint, and a 0 indicates that it is not.

- Load footprint_spatial_start_position & footprint_spatial_end_position (note that these variables are 1-based)
- Load clocking_offset_start and clocking_offset_interval (note that these variables are 0-based)
- For each footprint, calculate the number of clocking steps *s* by incrementing from 0 until

$4\,|2c_{start} + s\,c_{int}| \geq 1024$, where $c_{start}$ and $c_{int}$ are the clocking offset start and interval respectively

- The rows for each step are defined by the start and end positions, adjusted on each step
- For each step, step up if clockint is positive or down if it is negative
- First column is 0 on the first step, and 1 + the last column on subsequent steps
- Last column is lastcol = 4*(abs(2*clockstart[band,0])+i*clockint[band,0]), not to exceed 1023

Once the footprint and column of a data element in the color slices are calculated, the wavelength can be computed with the same approach outlined in 4.1.3.1.

# 5 OCO-2 L2 Data Products

## 5.1 Data Description and User Alerts

The data products include both Standard files and Diagnostic files. While much of the data in the two file types are the same, Section 5.3 describes a few of the fields that are uniquely in the L2Dia file.

Note that warn levels are currently not supported for the version 7 and 7R L2Std and L2Dia data; warn levels are included in the Lite files (daily files with bias correction and warn levels). Please see the anciallry documentation on these lite files, which includes all of the details of the development of warn levels and the bias correction.

A subset of fields is discussed here (Figure 5-1). The later section contains full data field tables.

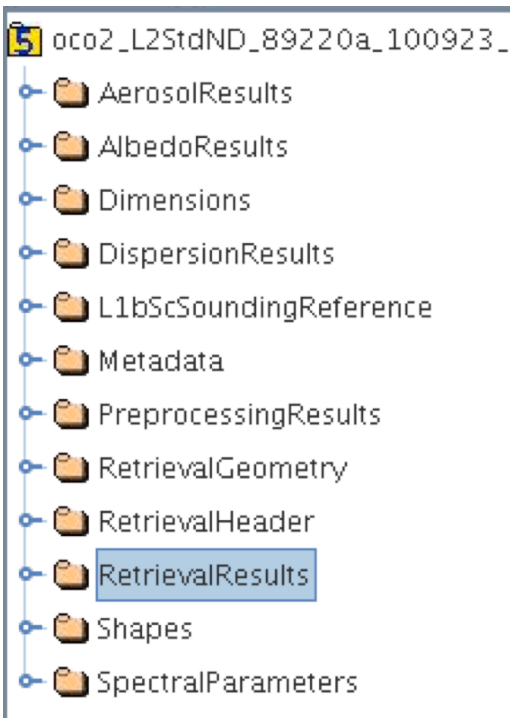



**Figure 5-1.** Folders contained in the L2Std product.

## 5.2 Key Data Fields for Standard and Diagnostic Files

The data and h5 folders that are included in the standard files are also included in the diagnostic files.

### 5.2.1 RetrievalHeader

Figure 5-2 shows the variables contained in the RetrievalHeader folder.

#### 5.2.1.1 Sounding_ID

This is the label for the sounding that will let you link the data across all of the product files. The sounding_id is the primary identification number for soundings across all OCO-2 products. The sounding_id is a composite of the time the data was acquired and the footprint number (f) in the form yyyymmddhhmmsssf.

#### 5.2.1.2 Warn_level

The warn levels are provided to assist the user in selecting data, as a more flexible approach than good/bad data quality flags. Warn levels range from 0 to 19. As discussed in Mandrake et al. [2013], low values indicate higher quality data. For each data release, we will include information about the standard deviation of the $X_{CO2}$ compared to validation data and how that changes with warn levels. We will also provide some recommendations about warn levels – how much data you will get by screening different ways, etc.

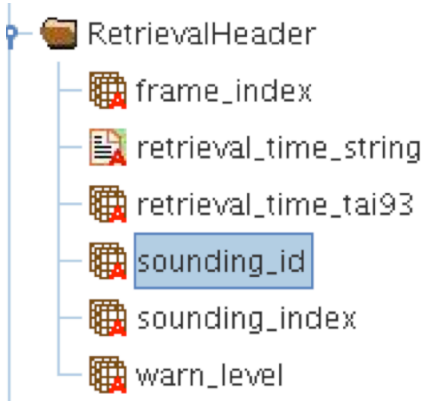



**Figure 5-2.** Variables in the RetrievalHeader folder.

Note that warn levels are currently not supported for the version 7 and 7R L2Std and L2Dia data; warn levels are included in the Lite files (daily files with bias correction and warn levels). Please see the anciallry documentation on

these lite files, which includes all of the details of the development of warn levels and the bias correction.

#### 5.2.1.3 Retrieval_time_string

This is the time as a string. An example of this field is 2010-09-23T18:36:04.334Z.

#### 5.2.1.4 Retrieval_time_tai93

This field is the time of the measurement, in seconds since Jan. 1, 1993. The TAI time corresponding to the example above is 559.420571E6.

### 5.2.2 RetrievalGeometry

Figure 5-3 shows the variables contained in the RetrievalGeometry folder.

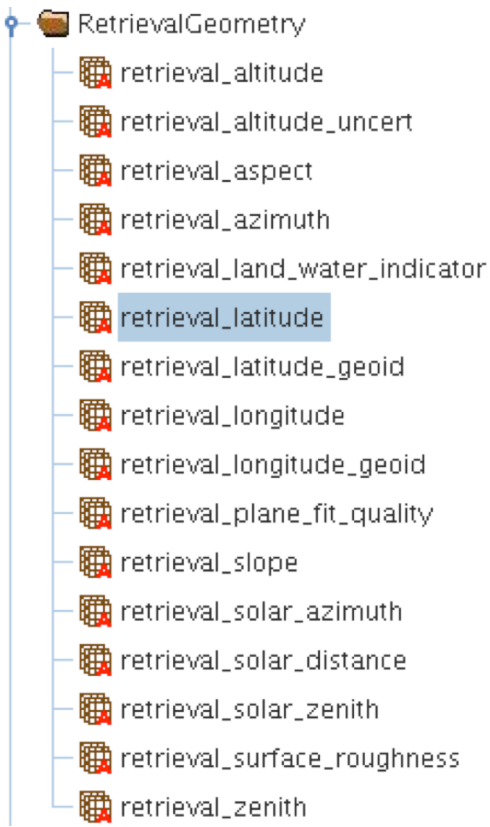



**Figure 5-3.** Variables in the RetrievalGeometry folder.

#### 5.2.2.1 Retrieval_latitude

This is the latitude of the sounding. These values range from −90 to +90.

#### 5.2.2.2 Retrieval_longitude

This is the longitude. We use the convention of −180 to +180 for longitude, with fractional degrees.

### 5.2.3 RetrievalResults

The retrieval results folder contains all of the elements of the state vector in the L2 full physics retrieval.

#### 5.2.3.1 XCO2

This variable expresses the column-averaged $CO_2$ dry air mole fraction for a sounding. Those soundings that did not converge will not be present. These values have units of mol/mol. This can easily be converted to ppm by multiplying by $10^6$.

#### 5.2.3.2 XCO2_uncert

This is an estimate of the uncertainty on the reported $X_{CO2}$. These values have units of mol/mol.

#### 5.2.3.3 XCO2_apriori

This is the initial guess for the $X_{CO2}$, in mol/mol.

#### 5.2.3.4 XCO2_avg_kernel

This is the averaging kernel. Note that the normalized averaging kernel (RetrievalResults/xco2_avg_kernel_norm) for a given pressure level is equal to the non-normalized value (retrieval-results/xco2_avg_kernel) divided by the pressure weighting function at that level. Note that levels are "layer boundaries" and have no thickness. See Appendix A of O'Dell et al. [2012] for details on how these quantities are defined.

#### 5.2.3.5 XCO2_avg_kernel_norm

See description in Section 5.2.3.4.

### 5.2.4 Declocking

A declocking algorithm has been implemented to correct for the slight misalignment of the FPAs (see section 2.2.2). The algorithm uses spatial information from selected color slices (also

described in section 2.2.2) to estimate a correction factor by which L1B spectra are multiplied to eliminate discontinuities. The selected color slices are first aggregated into "groups" for statistical purposes. For each group an estimate is made of a correction factor needed at each clocking boundary. Then, a correction function is determined from averages of these estimates, weighted by a column's distance to a group. A much more detailed description can be found in the Level 1B Algorithm Theoretical Basis Document (L1B ATBD). The variables described below are present in L1Sc and, with the exception of "radiance_jump_ratio_*", duplicated in the L2 product.

#### 5.2.4.1 declocking_color_indicator

This variable is a 1016-element integer array corresponding to the group associated with a given column. Groups are indexed starting with a value of 1. A value of 0 indicates that the column is not used in the correction algorithm.

#### 5.2.4.2 clocking_shift_color_indicator

This binary field specifies the location of the clocking jumps, i.e., the columns where a footprint's left and right boundaries are both incremented by 1.

#### 5.2.4.3 max_declocking_factor_o2/weak_co2/strong_co2

For each sounding, as described above, a correction function is determined by which the original L1B radiances are multiplied columnwise. This field indicated the value of the correction function at a column where the absolute value of the correction function minus unity is maximal. A value of 1.0 indicates that no correction has been performed. The fraction of soundings for whom this field lies outside the range [0.99, 1.01] (i.e., a correction of ≤1%) is on the order ~0.1. Corrections larger than several percent are very rare.

#### 5.2.4.4 radiance_jump_ratio_o2/weak_co2/strong_co2

The quantity described above (in 5.2.4.3) is determined from a weighted average of these fields, which is comprised of radiance ratios associated with each group and at each clocking jump. These fields are only saved in L1BSc. A more detailed description is provided in the L1B ATBD.

## 5.3 Key Data Fields Diagnostic Files

The Diagnostic files provide additional information that will be of use to scientists performing data assimilation.

### 5.3.1 RetrievalResults

The RetrievalResults folder for Diagnostic files has variables not included in the RetrievalResults folder for Standard products. These variables include the covariance matrices (*a priori* and *a posteriori*) and the averaging kernel matrix, as well as estimates of uncertainty from interference (interference smoothing uncertainty and Xco2_correlation_interf).

The measured and model radiances are also in this expanded file (along with a wavelength grid). The measured radiance matches what is found in the L1B file, and the modeled is the radiance calculated in the L2 full physics at the last retrieval step.

#### 5.3.1.1 Averaging_kernel_matrix

This matrix is the averaging kernel for all state vector elements.

#### 5.3.1.2 Aposteriori_covariance_matrix

This field is the *a posteriori* covariance matrix for all elements of the state vector. Note that there is also a separate variable for the co2_profile_covariance_matrix.

### 5.3.2 RetrievedStateVector

The RetrievedStateVector folder is unique to the Diagnostic files. It contains all of the state vector variables as a vector of the variables. All of these have been separately reported in the RetrievalResults folder as scalars.

# 6 ABO2 Preprocessor

The $O_2$ A band cloud screening algorithm was developed at Colorado State University (CSU) under the ACOS program. It employs a fast Bayesian retrieval to estimate surface pressure and surface albedo, assuming clear-sky conditions with only molecular Rayleigh scattering, from high-resolution spectra of the $O_2$ A band near 765 nm. The estimated surface pressure, surface albedo and the chi-squared goodness-of-fit statistic are used to flag scenes as cloudy, clear, or indeterminate [Taylor et al., 2011].

The basic method is that, absent clouds or aerosols, the surface pressure of a clear scene can be determined to within 2 – 5 hPa accuracy using the $O_2$ A band spectrum of reflected sunlight. This is because of the strong oxygen absorption features that are present in this band. When surface pressure is higher, the absorption features are deeper for a given observation geometry. When clouds or aerosols are present, they change the path lengths for most photons, either via shortening or lengthening, such that the retrieved surface pressure can be very different from the expected value based on a meteorological forecast, and also the $O_2$ A spectrum itself cannot be well-fit with a clear-sky assumption.

## 6.1 Prescreening of OCO-2 Soundings for Cloud and Aerosol

The primary objective of the ABO2 algorithm is to remove from the OCO-2 operational processing stream the soundings that are deemed too contaminated by clouds and /or aerosol for reliable $X_{CO2}$ retrieval in the computationally expensive L2 algorithm. The key screening parameters are interpreted internally via thresholds to provide a simple cloud flag. Furthermore, the key screening parameters can be used as inputs to a genetic algorithm used for individual sounding selection [Mandrake et al., 2013]. Details of the technique will be described in the Level 2 Algorithm Theoretical Basis Document (L2 ATBD) as well as an upcoming publication [Taylor et al., 2015, in prep]. As this is used as a preprocessor, the primary objective is to flag as cloudy scenes that are obviously cloudy or aerosol-contaminated. Therefore, thresholds are set rather loosely so scenes that have some cloud or aerosol contamination will sometimes pass the filter. This is by design. These scenes are sometimes useful for science, or may be rejected by additional pre- or postprocessor flags.

## 6.2 Key Science Data Fields

### 6.2.1 PreprocessingResults

The ABO2 data fields can be found in the standard L2 product (L2Std) PreprocessingResults folder, labeled with the _abp data field designation.

Figure 6-1 shows a screenshot of an example OCO-2 L2Sc product file as viewed by HDFView. The ABO2 data fields within the PreprocessingResults folder are highlighted and will be described in detail below.

#### 6.2.1.1 surface_pressure_delta_abp

The delta surface pressure is calculated as ECMWF forecast surface pressure – retrieved pressure – offset:

*surface_pressure_apriori_abp - surface_pressure_abp - surface_pressure_offset_abp*

in SI units (Pascals). This is the primary screening criterion for the ABO2 preprocessor, and represents the difference between the retrieved and meteorologically estimated surface pressure in the target field of view. Because the algorithm uses imperfect spectroscopy to retrieve surface pressure, there is a path length dependent offset term that is determined by analyzing clear scenes and subtracted from the retrieved value. This yields an unbiased estimate of the retrieved surface pressure. A threshold value can be set independently for nadir, glint, and target viewing modes. At the time of this writing, scenes with a difference greater than 25 hPa (2500 Pa) are flagged as cloudy, for all OCO-2 viewing modes.

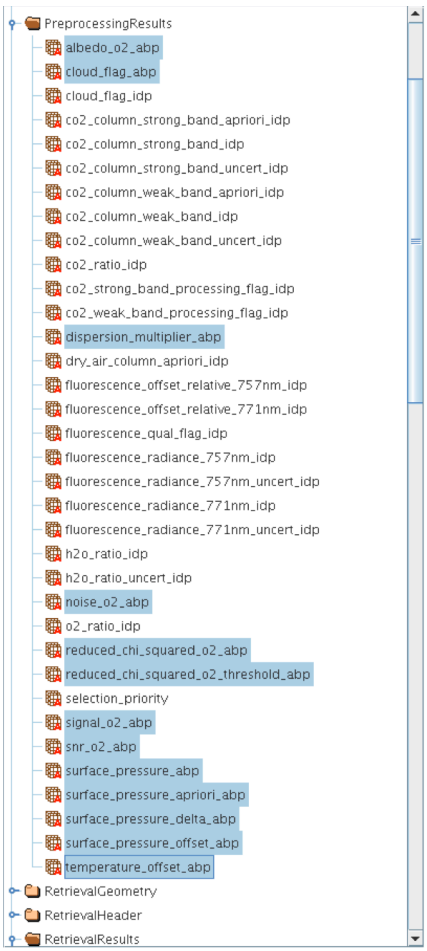



**Figure 6-1.** Screenshot of an HDFView look at the ABO2 preprocessor file.

#### 6.2.1.2 albedo_o2_abp

The retrieved surface albedo at 0.755 µm and 0.785 µm. Note that the retrieval assumes a perfect Lambertian surface albedo that varies linearly with wavelength. This assumption is currently made for all viewing modes; in glint mode, this means that the retrieved surface albedo can sometimes exceed unity. Over a dark ocean surface, the retrieved surface albedo is a good way to tell the presence of cloud: if the retrieved surface albedo is too large, a cloud or reflective aerosol layer is likely present. This test is not useful in glint mode or over land, and therefore is generally not used for OCO-2 prescreening.

#### 6.2.1.3 reduced_chi_squared_o2_abp

The reduced $\chi^2$ value of the spectral fit of the fast retrieval. Values greater than a specified threshold value are also indicative of clouds or aerosols present; as they are not accounted for in the retrieval, spectra containing them cannot be well fit. The threshold $\chi^2$ value is a parameterized function of SNR, as there are persistent spectral features (due to imperfect spectroscopy or instrument calibration) that scale with the signal level.

#### 6.2.1.4 dispersion_multiplier_abp

This parameter is also fit in the retrieval, and accounts for wavelength shifts in the spectra due primarily to the Earth-instrument Doppler shift. This is typically a maximum of ±7 km/sec, which corresponds roughly to shifts of ±0.3 cm-1 or ±0.018 nm (at maximum). This Doppler shift is easily fit for in the retrieval. Because the Doppler shift is formally a wavelength scaling, the ABO2 algorithm fits it as a scaling rather than a simple shift, though because of the narrowness of the fitted spectral region, it amounts to roughly the same thing.

#### 6.2.1.5 noise_o2_abp

The determined radiance noise value in the continuum based on the preflight calibration.

#### 6.2.1.6 reduced_chi_squared_o2_threshold_abp

The threshold reduced $\chi^2$ as described above. The logarithm of the reduced $\chi^2$ in clear scenes is assumed to be a piecewise linear function of SNR. This relationship is determined from clear scenes separately for nadir and glint modes.

#### 6.2.1.7 signal_o2_abp

The determined radiance mean value in the continuum.

#### 6.2.1.8 snr_o2_abp

The ratio of signal_o2_abp to noise_o2_abp.

#### 6.2.1.9 surface_presssure_abp

The retrieved value of surface pressure by the ABO2 algorithm, in units of Pa.

#### 6.2.1.10 surface_presssure_apriori_abp

The prior estimate of surface pressure in the target field of view, as determined from short-term ECMWF meteorological forecasts, and adjusted based on the mean elevation in the target field of view.

#### 6.2.1.11 surface_presssure_offset_abp

The empirically determined offset of the retrieved surface pressure for clear sky scenes. This offset is determined as a piecewise linear function of solar zenith angle, separately for nadir and glint modes, and is based on soundings that are determined to be approximately clear based on other methods (which include output from the IMAP-DOAS preprocessor as well as soundings collocated with the MODIS sensor aboard Aqua).

#### 6.2.1.12 temperature_offset_abp

The retrieved temperature offset, a simple additive offset to the meteorological estimate of the temperature profile. At the time of this writing this parameter is included in the state vector, but may be taken out to increase processing speed.

#### 6.2.1.13 cloud_flag_abp

The result of the surface pressure, $\chi^2$, and albedo tests to determine if a scene is likely cloudy (or heavily aerosol-laden). If a single test is failed, the cloud flag is set equal to 1, indicating the likely presence of cloud and/or aerosol. If all threshold checks are successfully passed, then the cloud flag is assigned a value of 0, indicating a sufficiently clear-sky scene. The cloud flag can also be set to 2 for "undetermined" cases. These are chiefly caused by solar zenith angle out of bounds or when viewing water surfaces in nadir observation mode (insufficient SNR).

# 7 IMAP-DOAS Preprocessor

The iterative maximum *a posteriori* differential optical absorption spectroscopy preprocessor is a non-scattering fast retrieval algorithm for optically thick absorbers [Frankenberg et al., 2005]. The preprocessor now serves two purposes for OCO-2: (1) retrieve vertical columns of trace gases in all bands individually for advanced cloud and aerosol screening and (2) retrieve solar induced chlorophyll fluorescence using an algorithm described in Frankenberg et al. [2011].

Mention lite files

## 7.1 Advanced Cloud and Aerosol Screening

For OCO-2, it is being used to derive vertical column densities of $H_2O$ and $CO_2$ independently in both $CO_2$ bands (1.6 and 2.0 µm) with the assumption of a non-scattering atmosphere. Given that Rayleigh scattering is very low in the near-infrared, this assumption holds true if neither aerosols nor clouds are present. In this case, both bands should yield an accurate and consistent result, i.e., the ratio of retrieved quantities in both $CO_2$ bands is approaching unity in clear sky conditions.

We found that the ratio of both $CO_2$ and $H_2O$ vertical column densities is deviating significantly from unity in the presence of aerosols and clouds [e.g., Mandrake et al., 2013]. The details of this technique will be described in the algorithm theoretical basis document as well as an upcoming publication.

## 7.2 Retrievals of Solar-Induced Chlorophyll Fluorescence

The possibility of retrieving solar induced chlorophyll fluorescence from high-resolution spectra in the vicinity of the $O_2$ A band have been shown in Frankenberg et al. [2011] and Joiner et al. [2011]. Here, we apply the retrieval algorithm described in Frankenberg et al. [2011], embedded in the IMAP-DOAS retrieval code. It is important to note that this is the dedicated fluorescence retrieval, as opposed to the fluorescence retrieval within the full-physics L2 code. Interference with scattering properties in this retrieval is thus minimized [e.g., Frankenberg et al., 2012].

## 7.3 Key Science Data Fields

Note: There are a few generic differences between the IMAP-DOAS data and the main L2 files based on the full-physics $X_{CO2}$ retrieval. First, all OCO-2 L1BSc data with a solar zenith angle smaller than 80 degrees and a valid quality flag are processed through the preprocessor. The other main difference is the structure of the IMAP L2 data fields as they are arranged in 2-dimensional data fields, with one dimension being time of readout (variable per orbit) and the other one representing each of the 8 OCO-2 footprints independently (i.e., the dimensions are n x 8, with n denoting the number of total focal plane array readouts per orbit).

Figure 7-1 shows a screenshot of the general structural overview of an IMAP-DOAS preprocessor L2 file using HDFView. The most important data folders are expanded and will be described in detail below.

### 7.3.1 SoundingGeometry

The SoundingGeometry folder includes all relevant location information for each footprint as well as information on viewing geometries as well as topography. It follows exactly the same definitions as the official full-physics L2 files and the reader is referred to the more detailed explanation in the respective document for the official $X_{CO2}$ L2 file.

### 7.3.2 DOASCO2

The DOASCO2 folder contains all relevant fields from the IMAP retrievals of $CO_2$ columns in both OCO-2 $CO_2$ bands. Definitions are as follows:

#### 7.3.2.1 DOASCO2/co2_column_strong_band_apriori_idp

*A priori* vertical column density of $CO_2$ in the strong (2.0 µm) band in molecules/m$^2$. Note that a constant volume mixing ratio (VMR) is assumed across the globe in this IMAP-DOAS version and that the *a priori* thus mostly depends on surface pressure.

#### 7.3.2.2 DOASCO2/co2_column_weak_band_apriori_idp

See above but for weak band (both are identical).

#### 7.3.2.3 DOASCO2/co2_column_strong_band_idp

Retrieved vertical column density of $CO_2$ in the strong (2.0 µm) band in molecules/m$^2$. Note that this is different from the FP L2 $CO_2$ vertical column as this variable here is only retrieved in one band and ignores scattering.

#### 7.3.2.4 DOASCO2/co2_column_weak_band_idp

See above but for the weak band.

#### 7.3.2.5 DOASCO2/co2_column_strong_band_uncert_idp

Uncertainty in retrieved vertical column density of $CO_2$ in the strong (2.0 µm) band in molecules/m$^2$.

#### 7.3.2.6 DOASCO2/co2_column_weak_band_uncert_idp

See above but for weak band.

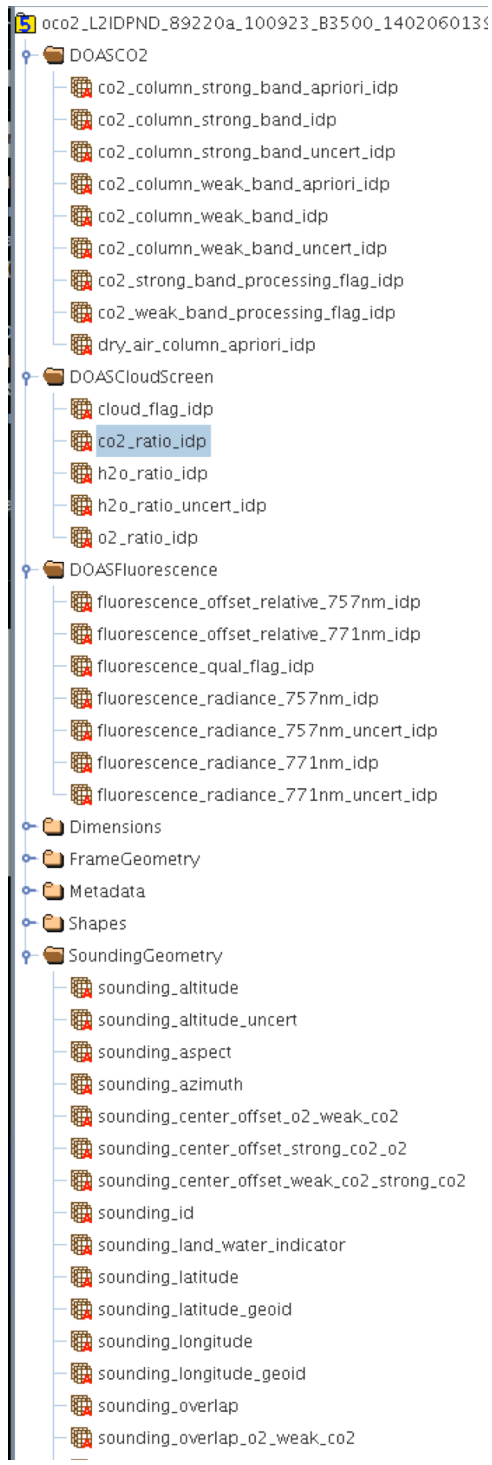



**Figure 7-1.** Screenshot of an HDFView look at the IMAP-DOAS preprocessor file.

#### 7.3.2.7 DOASCO2/co2_column_strong_band_processing_flag_idp

Processing flag for the strong $CO_2$ band retrieval. Every sounding will have a value: 0 (successfully processed), 1 (not converged) or 2 (not processed).

#### 7.3.2.8 DOASCO2/co2_column_weak_band_processing_flag_idp

See above but for weak band.

#### 7.3.2.9 DOASCO2/dry_air_column_apriori_idp

*A priori* total dry air column (mol/m$^2$) of the respective sounding (based purely on ECMWF input data and topography).

### 7.3.3 DOASCloudScreen

In this folder, all relevant variables for advanced cloud-screening are stored.

#### 7.3.3.1 DOASCloudScreen/cloud_flag_idp

Tentative cloud flag based purely used on the IMAP preprocessor. Values can be:

- -2 = unusable (not processed)
- -1 = not all retrievals converged
- 0 = clearly cloudy
- 1 = probably cloudy
- 2 = probably clear
- 3 = very clear

The values are based on simple threshold criteria and based on the ratios defined in the same folder. This flag is *not* being used as input for FP L2 sounding selection and is also not being validated.

#### 7.3.3.2 DOASCloudScreen/co2_ratio_idp

Ratio of retrieved $CO_2$ vertical column density (VCD) in weak and strong band ($VCD_{1.6\mu m}/VCD_{2.0\mu m}$). Please note that there might be a footprint dependence of this ratio, i.e., each OCO-2 footprint may have a different optimal ratio value. Eventually, we will document these subtleties but the users of early mission data should be aware of this fact, which applies to all variables discussed here.

#### 7.3.3.3 DOASCloudScreen/h2o_ratio_idp

Ratio of retrieved $H_2O$ vertical columns in weak and strong band ($VCD_{1.6\mu m}/VCD_{2.0\mu m}$).

#### 7.3.3.4 DOASCloudScreen/h2o_ratio_uncert_idp

1-sigma uncertainty of the ratio of retrieved $H_2O$ vertical columns in weak and strong band ($VCD_{1.6\mu m}/VCD_{2.0\mu m}$).

#### 7.3.3.5 DOASCloudScreen/o2_ratio_idp

Ratio of retrieved vs. expected $O_2$ column density based on the fluorescence 771 nm retrieval window, which includes some weak $O_2$ line used to derive an $O_2$ column. Values substantially lower than 1 indicated the presence of clouds.

### 7.3.4 DOASFluorescence

In this folder, all relevant variables for the fluorescence retrieval in the 757 nm and 771 nm window are saved. From experience with GOSAT and also looking at the general shape of the fluorescence emission spectrum, SIF at 771 nm is about a factor 1.7 lower than at 757 nm.

Please note that the primary retrieval fits the relative contribution of an additive offset (such as from fluorescence) in both retrieval windows. Over nonfluorescing targets, the average retrieval of this offset may deviate from 0 because of small uncertainties in instrument line-shape as well as the solar line-list. In relative units, this "bias" is expected to be constant across the globe but a constant bias in the relative contribution to of fluorescence will translate into a signal-dependent bias in the absolute value (in radiance units). We will characterize this bias as

soon as possible but users, esp. in the early mission phase, should be aware of that caveat and provide feedback of observed inconsistencies to the developer team.

#### 7.3.4.1 DOASFluorescence/fluorescence_offset_relative_757nm_idp

Fraction of continuum level radiance explained by an additive offset term in the 757 nm spectral window (unitless). In the absence of instrumental errors, this will be only caused by fluorescence. Rotational Raman scattering should be negligible over typical vegetated surface and moderate solar zenith angles (<65 degrees).

#### 7.3.4.2 DOASFluorescence/fluorescence_offset_relative_771nm_idp

Same as above but for the 771 nm window.

#### 7.3.4.3 DOASFluorescence/fluorescence_radiance_757nm_idp

Fluorescence term expressed in absolute radiance units (ph/s/$m^2$/sr/µm) in the 757 nm fit window. This is a derived quantity based on the fit of the relative contribution multiplied with the overall continuum level radiance.

#### 7.3.4.4 DOASFluorescence/fluorescence_radiance_771nm_idp

Same as above but for the 771 nm window.

#### 7.3.4.5 DOASFluorescence/fluorescence_radiance_757nm_corr_idp

Fluorescence term expressed in absolute radiance units (ph/s/$m^2$/sr/µm) in the 757 nm fit window. This is a derived quantity based on the fit of the relative contribution multiplied with the overall continuum level radiance. This version applies a first order correction of SIF as a function of signal level, as derived from early mission data. It is generally advised to only use SIF retrievals if the continuum level radiance in the $O_2$ A band is below 4e20 ph/s/$m^2$/sr/µm.

#### 7.3.4.6 DOASFluorescence/fluorescence_radiance_771nm_corr_idp

Same as above but for the 771 nm window.

#### 7.3.4.7 DOASFluorescence/fluorescence_radiance_757nm_uncert_idp

Estimated uncertainty (1-sigma) of the fluorescence term expressed in absolute radiance units (ph/s/$m^2$/sr/µm) in the 757 nm fit window.

#### 7.3.4.8 DOASFluorescence/fluorescence_radiance_771nm_uncert_idp

Same as above but for the 771 nm window.

#### 7.3.4.9 DOASFluorescence/fluorescence_qual_flag_idp

Quality flag for the fluorescence retrieval (0 = ok, 1 = didn not pass quality filter or unprocessed).

# 8 Full Data Tables

The tables below give a full listing of the variables in each folder/group and a short description including the data type. We suggest that users familiarize themselves with the data dimensions by browsing with HDFview or a similar tool. The most useful global attributes present in all files are presented in Section 8.1. Tables for L2 data products are in Section 8.2. The groups common to L1BSc and DOAS data products are described in Section 8.3. Finally, Section 8.4 provides a list of the data fields in each L1BSc group in the OCO-2 HDF data files.

## 8.1 Metadata in all Product Files

Metadata (see **Table 8-1** below) contains information about the orbit, including type, start and stop times, number of frames acquired, color slice locations that are common to all of the data in the file. Note that the metadata provided are slightly different for each OCO-2 product file, this table includes metadata provided in the L1B science (L1bSc), L2 standard (L2Std) and L2 diagnostic (L2Dia) files. Fields that are only included in the L1bSc files are shaded in blue and files that are only included in L2Dia files are shaded in red.

**Table 8-1.** OCO-2 metadata in the Level 2 products.

| Data Element | Type | Comments |
|---|---|---|
| AbscoCO2Scale | Float32 | Empirical scaling factors for CO2 ABSCO tables. Values should be different for the 1.6 micron and 2.06 micron bands and were chosen to provide agreement of retrieved XCO2 with TCCON XCO2. |
| AbscoH2OScale | Float32 | Empirical scaling factor for H2O ABSCO tables. Currently should be 1.0. |
| AbscoO2Scale | Float32 | Empirical scaling factor for O2 ABSCO tables. Values chosen to improve agreement between retrieved surface pressure and independent estimates from a numerical weather prediction model. |
| AcquisitionMode | String | The instrument mode in which the data in the product were collected. Valid values are:<br>'Glint'<br>'Nadir'<br>'Target'<br>'Sample Dark Calibration'<br>'Sample Lamp Calibration'<br>'Sample Solar/limb Calibration'<br>'Single-Pixel Dark Calibration'<br>'Single-Pixel Lamp Calibration'<br>'Single-Pixel Solar/limb Calibration' |
| ActualFrames | Int32 | Actual number of frames reported in this product |
| ActualGoodRetrievals | Signed32 | Actual number of reported retrievals with a "Good" quality flag. |
| ActualRetrievals | Signed32 | Actual number of retrievals reported in the product |
| ARPAncillaryDatasetDescriptor | String | The name of the Ancillary Radiometric Product file used to calibrate this file |

| Data Element | Type | Comments |
|---|---|---|
| AscendingEquatorCrossingDate | String | The date of the equator crossing of the spacecraft ground track in the ascending direction |
| AscendingEquatorCrossingLongitude | Float32 | The longitude of the equator crossing of the spacecraft ground track in the ascending direction |
| AscendingEquatorCrossingTime | String | The time of the equator crossing of the spacecraft ground track in the ascending direction |
| BuildId | VarLen Str | The ID of build which included the software that created this product |
| BadPixelMapVersionNum | UInt32 | Version number of corresponding Bad Pixel Map used to calibrate this data file |
| ColorSlicePositionO2 | Int16 | Absolute spectral position of each ABO2 color slice |
| ColorSlicePositionStrongCO2 | Int16 | Absolute spectral position of each SCO2 color slice |
| ColorSlicePositionWeakCO2 | Int16 | Absolute spectral position of each WCO2 color slice |
| DiffuserPosition | Float32 | The position of the solar diffuser at the beginning of the mode:<br>0—Lamp or Closed position<br>75—Open for science<br>150—Solar Calibration position |
| EphemerisType | String | The source of the spacecraft ephemeris data that were utilized to generate this data file |
| EquatorCrossingDate | String | The date of the equator crossing of the spacecraft ground track in the descending direction |
| EquatorCrossingLongitude | Float32 | The longitude of the equator crossing of the spacecraft ground track in the descending direction |
| EquatorCrossingTime | String | The time of the equator crossing of the spacecraft ground track in the descending direction |
| ExpectedFrames | Int32 | Nominal number of frames in this product |
| FirstSoundingId | Int64 | The ID of the first sounding in this file |
| GapStartTime | FixLen Str | The source of the spacecraft ephemeris data that were utilized to generate this data file. |
| GapStopTime | FixLen Str | The timestamp before which a nonexistent, unnecessary, spurious, questionable, or erroneous data segment ends. |
| GranulePointer | VarLen Str | NCSA HDF - A character string that describes the internal format of the data product. |
| HDFVersionId | VarLen Str | 5.x - A character string that identifies the version of the HDF (Hierarchical Data Format) software that was used to generate this data file. |
| InputPointer | VarLen Str | The ID of the first sounding in this file. |
| InstrumentShortName | VarLen Str | OCO-2 - The name of the instrument that collected the telemetry data. |
| InitialUnusedSpatialPixels | Int16 | Distance in spatial pixels of the start of first footprint from edge of FPA |

| Data Element | Type | Comments |
|---|---|---|
| L1BAlgorithmDescriptor | String | Identification of the algorithm and version used to generate this product – L1B Science files (L1bSc) |
| L2FullPhysicsAlgorithmDescriptor | VarLen Str | Identification of the algorithm and version used to generate this product"; |
| L2FullPhysicsDataVersion | VarLen Str | The source of the data used by the Full-physics algorithm: r01 - initial processing; r02, r03, etc. - reprocessed data. |
| L2FullPhysicsExeVersion | VarLen Str | The build version number of the Full-physics algorithm used to generate this product. |
| L2FullPhysicsInputPointer | VarLen Str | Pointer to one or more data granules and auxilliary files that provide the major input that was used to generate this product. |
| L2FullPhysicsProductionLocation | VarLen Str | Facility in which Full Physics code was run, typically: JPL, Pleiades |
| LastSoundingId | Int64 | The ID of the last sounding in this file |
| MaxMS | Float32 | The maximum measureable signal value to be used in radiance noise calculation – L1B Science files (L1bSc) |
| ModeCounter | String | The nth occurrence of this particular mode for this orbit, indicated by letter ('a', 'b', 'c', 'd', etc.) |
| OperationMode | String | The two-letter abbreviation of the AcquisitionMode: GL, ND, TG, DS, LS, SS, BS, NP, GP, TP, DP, LP, SP, BP, XS, XP, MS, MP, SB |
| OrbitEccentricity | Float32 | The eccentricity of the spacecraft orbital path |
| OrbitInclination | Float32 | The angle between the plane of the spacecraft orbital path and the Earth equatorial plane |
| OrbitParametersPointer | String | The data files that provided the orbit parameters used to generate this product |
| OrbitPeriod | Float32 | The time span between two consecutive descending node crossings |
| OrbitSemiMajorAxis | Float32 | The length of the semi-major axis of the spacecraft orbit |
| OrbitStartDate | String | The date of the equator crossing of the spacecraft nadir track in the descending direction |
| OrbitStartLongitude | Float32 | The longitude of the equator crossing of the spacecraft ground track in the descending direction |
| OrbitStartTime | String | The time of the equator crossing of the spacecraft ground track in the descending direction |
| PlatformLongName | VarLen Str | Orbiting Carbon Observatory 2 |
| PlatformShortName | VarLen Str | OCO-2 |
| PlatformType | VarLen Str | spacecraft - The type of platform associated with the instrument which acquires the accompanying data |
| ProcessingLevel | VarLen Str | Indicates data level (Level 0, Level 1A, Level 1B, Level 2) in this product |

| Data Element | Type | Comments |
|---|---|---|
| ProducerAgency | VarLen Str | NASA - Identification of the agency that provides the project funding. |
| ProducerInstitution | VarLen Str | JPL - Identification of the institution that provides project management. |
| ProductionDateTime | FixLen Str | The date and time at which the product was created (yyyy-mm-ddThh:mm:ss.mmmZ) |
| ProductionLocation | VarLen Str | Facility in which this file was produced, typically: 'Operations Pipeline', 'Operations Pipeline 2', 'Science Computing Facility', 'Test Pipeline', Test Pipeline 2' |
| ProductionLocationCode | FixLen Str | One-letter code indicating the ProductionLocation, typically: - Operations Pipelines (1) or 2, 's' - Science Computing Facility, 't' - Test Pipelines (1) or 2. |
| ProjectId | VarLen Str | OCO-2 - The project identification string |
| QAGranulePointer | VarLen Str | A pointer to the quality assessment product that was generated with this product. |
| RadianceConversionFactor | Float32 | Multiplicative factor used to convert W m-2 sr-1 um-1 to Ph sec-1 m-2 sr-1 um-1 at 760 nm (38.228x1017). |
| RangeBeginningDate | FixLen Str | The date on which the earliest data contained in the product were acquired (yyyy-mm-dd). |
| RangeBeginningTime | FixLen Str | The time at which the earliest data contained in the product were acquired (hh:mm:ss.mmmZ). |
| RangeEndingDate | FixLen Str | The date on which the latest data contained in the product were acquired (yyyy-mm-dd). |
| RangeEndingTime | FixLen Str | The time at which the latest data contained in the product were acquired (hh:mm:ss.mmmZ). |
| ReportedSoundings | Int8 | Indicates the inclusion of each footprint in the data:<br>0—not included<br>1—included |
| RetrievalIterationLimit | Signed32 | Maximum number of iterations allowed before the algorithm gives up and sets the outcome_flag to 3. |
| SciToFAColorOffset | Signed32 | The index of the first spectral pixel of arrays with FPAColor shape that appears in the first spectral element of arrays with SciColor shape. |
| ShortName | VarLen Str | The short name identifying the data type of this product. |
| SISName | VarLen Str | The name of the document describing the contents of the product. |
| SISVersion | VarLen Str | The version of the document describing the contents of the product. |
| SizeMBECSDataGranule | Float32 | The size of this data granule in Megabytes. |
| SpectralChannel | String | A description of the spectral channels used for the measurements |

| Data Element | Type | Comments |
|---|---|---|
| SpectralChannel | VarLen Str | A description of the spectral channels used for the measurements. |
| StartOrbitNumber | Signed32 | The first orbit on which data contained in the product were acquired. |
| StartPathNumber | Signed32 | The first WRS path on which data contained in the product was collected. |
| StopOrbitNumber | Signed32 | The last orbit on which data contained in the product were acquired. |
| StopPathNumber | Signed32 | The last WRS path on which data contained in the product was collected. |
| SciToFPAColorOffset | Int16 | The index of the first spectral pixel of arrays with FPAColor shape that appears in the first spectral element of arrays with SciColor shape – L1B Science files (L1bSc) |
| VMRO2 | Float32 | The volume mixing ratio of atmospheric O2 |

## 8.2 Data Tables for L2Std and L2Dia Files

The L2 data products in these tables are from the L2Std and L2Dia files. Fields only in the L2Dia files are highlighted in dark red.

### 8.2.1 Aerosol Results

The AerosolResults group (see **Table 8-2** below) contains detailed information about the data fields from the L2 retrievals.

**Table 8-2.** AerosolResults data in Level 2 Standard Data file (L2Std) and Diagnostic file (L2Dia).

| Data Element | Type | Comments |
|---|---|---|
| aerosol_1_aod | Float32 | Retrieved total column-integrated aerosol optical depth for aerosol type 1 |
| aerosol_1_aod_high | Float32 | Retrieved column-integrated aerosol optical depth of aerosol type 1 for pressure levels less than 50,000 Pa |
| aerosol_1_aod_low | Float32 | Retrieved column-integrated aerosol optical depth of aerosol type 1 for pressure levels greater than 80000 Pa |
| aerosol_1_aod_mid | Float32 | Retrieved column-integrated aerosol optical depth of aerosol type 1 for pressure levels between 50,000 and 80,000 Pa |
| aerosol_1_gaussian_log_param | Float32 | Retrieved Gaussian log parameters for aerosol type 1 [total log aod, center pressure/surf-pressure, pressure sigma/surf-pressure] |
| aerosol_1_gaussian_log_param _apriori | Float32 | *a priori* of retrieved gaussian log parameters for aerosol type 1 |
| aerosol_1_gaussian_log_param _uncert | Float32 | Uncertainty of retrieved gaussian log parameters for aerosol type 1 |
| aerosol_2_aod | Float32 | Retrieved total column-integrated aerosol optical depth for aerosol type 2 |

| | | |
|---|---|---|
| aerosol_2_aod_high | Float32 | Retrieved column-integrated aerosol optical depth of aerosol type 2 for pressure levels less than 50,000 Pa |
| aerosol_2_aod_low | Float32 | Retrieved column-integrated aerosol optical depth of aerosol type 2 for pressure levels greater than 80,000 Pa |
| aerosol_2_aod_mid | Float32 | Retrieved column-integrated aerosol optical depth of aerosol type 2 for pressure levels between 50,000 and 80,000 Pa |
| aerosol_2_gaussian_log_param | Float32 | Retrieved gaussian log parameters of aerosol type 2 [total log aod, center pressure/surf-pressure, pressure sigma/surf-pressure] |
| aerosol_2_gaussian_log_param _apriori | Float32 | a priori of retrieved gaussian log parameters for aerosol type 2 |
| aerosol_2_gaussian_log_param _uncert | Float32 | Uncertainty of retrieved gaussian log parameters for aerosol type 2 |
| aerosol_3_aod | Float32 | Retrieved total column-integrated aerosol optical depth for aerosol type 3 |
| aerosol_3_aod_high | Float32 | Retrieved column-integrated aerosol optical depth of aerosol type 3 for pressure levels less than 50,000 Pa |
| aerosol_3_aod_low | Float32 | Retrieved column-integrated aerosol optical depth of aerosol type 3 for pressure levels greater than 80,000 Pa |
| aerosol_3_aod_mid | Float32 | Retrieved column-integrated aerosol optical depth of aerosol type 3 for pressure levels between 50,000 and 80,000 Pa |
| aerosol_3_gaussian_log_param | Float32 | Retrieved gaussian log parameters for aerosol type 3 [total log aod, center pressure/surf-pressure, pressure sigma/surf-pressure] |
| aerosol_3_gaussian_log_param _apriori | Float32 | a priori of retrieved gaussian log parameters for aerosol type 3 |
| aerosol_3_gaussian_log_param _uncert | Float32 | Uncertainty of retrieved gaussian log parameters of aerosol type 3 |
| aerosol_4_aod | Float32 | Retrieved total column-integrated aerosol optical depth for aerosol type 4 |
| aerosol_4_aod_high | Float32 | Retrieved column-integrated aerosol optical depth of aerosol type 4 for pressure levels less than 50,000 Pa |
| aerosol_4_aod_mid | Float32 | Retrieved column-integrated aerosol optical depth of aerosol type 4 for pressure levels between 50,000 and 80,000 Pa |
| aerosol_4_aod_low | Float32 | Retrieved column-integrated aerosol optical depth of aerosol type 4 for pressure levels greater than 80,000 Pa |
| aerosol_4_gaussian_log_param | Float32 | Retrieved gaussian log parameters for aerosol type 4 [total log aod, center pressure/surf-pressure, pressure sigma/surf-pressure] |
| aerosol_4_gaussian_log_param _apriori | Float32 | a priori of retrieved gaussian log parameters for aerosol type 4 |
| aerosol_4_gaussian_log_param _uncert | Float32 | Uncertainty of retrieved gaussian log parameters for aerosol type 4 |
| aerosol_total_aod | Float32 | Retrieved total column-integrated aerosol optical depth for all aerosol types |

| | | |
|---|---|---|
| aerosol_total_aod_high | Float32 | Retrieved column-integrated aerosol optical depth for all aerosol types for pressure levels less than 50,000 Pa |
| aerosol_total_aod_low | Float32 | Retrieved column-integrated aerosol optical depth for all aerosol types for pressure levels greater than 80,000 Pa |
| aerosol_total_aod_mid | Float32 | Retrieved column-integrated aerosol optical depth for all aerosol types for pressure levels between 50,000 and 80,000 Pa |
| aerosol_types | String | Retrieved aerosol type |

### 8.2.2 Albedo Results

The AlbedoResults group (see Error! Reference source not found. below) contains detailed information about the albedo properties in the L2 retrievals.

Table 8-3. SpectralParameters data in Level 2 Standard Data file (L2Std) and Diagnostic file (L2Dia).

Note that the L2Std variables are generally a subset of data in L2Dia. Shapes are not noted on these tables but can be easily found inside the HDF files using the dimension and shape information.

### 8.2.3 L2 Spectral Parameters Group Fields

The SpectralParameters group (see **Error! Reference source not found.** below) contains the retreival results from the OCO-2 Level 2 data products.

Table 8-3. SpectralParameters data in Level 2 Standard Data file (L2Std) and Diagnostic file (L2Dia).

| Data Element | Type | Comments |
|---|---|---|
| measured_radiance | Float32 | The measured spectra mapped into the retrieval color grid – Diagnostic files only |
| measured_radiance_uncert | Float32 | Uncertainty of the measured spectra mapped into the retrieval color grid – Diagnostic files only |
| modeled_radiance | Float32 | The modeled spectra mapped into the retrieval color grid – Diagnostic files only |
| noise_o2_fph | Float32 | Aggregate noise in the ABO2 band |
| noise_strong_co2_fph | Float32 | Aggregate noise in the SCO2 band |
| noise_weak_co2_fph | Float32 | Aggregate noise in the WCO2 band |
| num_colors | Int32 | The number of spectral colors contributing to the retrieval result – Diagnostic files only |
| num_colors_per_band | Int32 | The number of spectral colors contributing to the retrieval result per spectral region – Diagnostic files only |
| reduced_chi_squared_o2_fph | Float32 | Reduced chi squared of spectral fit of the ABO2 band |
| reduced_chi_squared_strong_co2_fph | Float32 | Reduced chi squared of spectral fit of the SCO2 band |
| reduced_chi_squared_weak_co2_fph | Float32 | Reduced chi squared of spectral fit of the WCO2 band |
| relative_residual_mean_square_o2 | Float32 | Mean square of the residuals divided by the signal for the ABO2 band |

| Data Element | Type | Comments |
|---|---|---|
| relative_residual_mean_square _strong_co2 | Float32 | Mean square of the residuals divided by the signal for the SCO2 band |
| relative_residual_mean_square _weak_co2 | Float32 | Mean square of the residuals divided by the signal for the WCO2 band |
| residual_mean_square_o2 | Float32 | Mean of the squares of the residuals for the ABO2 band |
| residual_mean_square_strong_co2 | Float32 | Mean of the squares of the residuals for the SCO2 band |
| residual_mean_square_weak_co2 | Float32 | Mean of the squares of the residuals for the WCO2 band |
| sample_indexes | Int32 | The 1-based index of the detector sample associated with the radiance – Diagnostic files only |
| signal_o2_fph | Float32 | Aggregate signal in the ABO2 band |
| signal_strong_co2_fph | Float32 | Aggregate signal in the SCO2 band |
| signal_weak_co2_fph | Float32 | Aggregate signal in the WCO2 band |
| wavelength | Float32 | The wavelength represented by each point in the retrieval color grid – Diagnostic files only |

## 8.3 Selected L1B Data Tables from L1B Science Files (L1bSc)

The primary dimension of the data is the sounding. For each sounding, there are eight spatial footprints. For each footprint, there are three spectral bands: $O_2$ A band, weak $CO_2$, and strong $CO_2$.

### 8.3.1 FrameGeometry

The FrameGeometry group (see **Table 8-4** below) contains detailed information about the spacecraft position and orientation during the observations.

**Table 8-4.** Spacecraft position and orientation during observations.

| Data Element | Type | Comments |
|---|---|---|
| ground_track | Float32 | Ground track orientation relative to local north |
| pitch | Float32 | Interpolated spacecraft attitude at the frame time |
| relative_velocity | Float32 | Velocity of the spacecraft along the LOS; positive indicates spacecraft moving toward target location |
| roll | Float32 | Interpolated spacecraft attitude at the frame time |
| spacecraft_alt | Float32 | Altitude of the spacecraft above the reference ellipsoid at the frame time |
| spacecraft_lat | Float32 | Geodetic latitude of the spacecraft at the frame time |
| spacecraft_lon | Float32 | Longitude of the spacecraft at the frame time |
| spacecraft_position | Float32 | Interpolated spacecraft position at the frame time |
| spacecraft_velocity | Float32 | Interpolated spacecraft velocity at the frame time |
| yaw | Float32 | Interpolated spacecraft attitude at the frame time |

### 8.3.2 SoundingGeometry

The SoundingGeometry group (see **Table 8-5** below) contains detailed information about the geometric location, atmospheric geometry, and surface conditions for the sounding that combines the three spectrometers.

**Table 8-5.** Geometric location, atmospheric geometry, and surface conditions.

| Data Element | Type | Comments |
|---|---|---|
| sounding_altitude | Float32 | Altitude of the sounding based on Earth topography |
| sounding_altitude_uncert | Float32 | Uncertainty of the source Earth topography data |
| sounding_aspect | Float32 | Orientation of the surface slope relative to the local North |
| sounding_azimuth | Float32 | Angle between the LOS as defined from the sounding location to the spacecraft, and the sounding location local north direction |
| sounding_center_offset _o2_weak_co2 | Float32 | Distance between the ABO2 band footprint center and the $WCO_2$ band footprint center |
| sounding_center_offset _strong_co2_o2 | Float32 | Distance between the SCO2 band footprint center and the ABO2 band footprint center |
| sounding_center_offset _weak_co2_strong_co2 | Float32 | Distance between the WCO2 band footprint center and the SCO2 band footprint center |
| sounding_id | Int64 | Unique identifier for each sounding |
| sounding_land_fraction | Float32 | Percentage of land surface type within the sounding |
| sounding_land_water _indicator | Int8 | Surface type at the sounding location:<br>• 0—Land<br>• 1—Water<br>• 2—Unused<br>• 3—Mixed land water |
| sounding_latitude | Float32 | Geodetic latitude of the sounding based on Earth topography |
| sounding_latitude_geoid | Float32 | Geodetic latitude of the sounding based on standard geoid |
| sounding_longitude | Float32 | Longitude of the sounding based on Earth topography |
| sounding_longitude_geoid | Float32 | Longitude of the sounding based on standard geoid |
| sounding_overlap | Float32 | Area of intersection of all three band footprints relative to average area of all three band footprints |
| sounding_overlap _o2_weak_co2 | Float32 | Area of intersection of the footprints of ABO2 and WCO2 relative to the average area of the two footprint |
| sounding_overlap _strong_co2_o2 | Float32 | Area of intersection of the footprints of ABO2 and SCO2 relative to the average area of the two footprints |
| sounding_overlap _weak_co2_strong_co2 | Float32 | Area of intersection of the footprints of WCO2 and SCO2 relative to the average area of the two footprints |
| sounding_plane_fit_quality | Float32 | Goodness-of-fit of surface slope: the standard deviation of the points, to which the plane is fitted, with the expected values taken as the orthogonal projection of the points onto the plane |
| sounding_polarization_angle | Float32 | The angle between the accepted polarization axis of the instrument and the instrument reference plane for polarization, defined as the plane formed by the LOS and the ray from the sounding location to the local zenith |

| Data Element | Type | Comments |
|---|---|---|
| sounding_qual_flag | UInt64 | Bit flags indicating the quality of the data in sounding:<br>• 0—Good<br>• Non-zero—See **Error! Reference source not found.** |
| sounding_relative_velocity | Float32 | Velocity of the spacecraft along the LOS: positive indicates spacecraft moving toward sounding location |
| sounding_slant_path_diff _o2_weak_co2 | Float32 | Difference in slant path between ABO2 and WCO2 footprints |
| sounding_slant_path_diff _strong_co2_o2 | Float32 | Difference in slant path between SCO2 and ABO2 footprints |
| sounding_slant_path_diff _weak_co2_strong_co2 | Float32 | Difference in slant path between WCO2 and SCO2 footprints |
| sounding_slope | Float32 | Slope of a plane fit to points within the sounding |
| sounding_solar _relative_velocity | Float64 | Velocity of the Sun along the sounding location/Sun vector: negative indicates Sun moving toward sounding location |
| sounding_solar_azimuth | Float32 | Angle between the solar direction as defined from the sounding location to the Sun, and the sounding location local north direction |
| sounding_solar_distance | Float64 | Distance between sounding location and the Sun |
| sounding_solar_zenith | Float32 | Angle between the solar direction as defined from the sounding location to the Sun, and the sounding location local zenith direction |
| sounding_surface_roughness | Float32 | Standard deviation of the altitude within the sounding |
| sounding_time_string | String | Data acquisition time for the sounding based upon the three footprint times (yyyy-mm-ddThh:mm:ss.mmmZ) |
| sounding_time_tai93 | Float64 | Data acquisition time for the sounding based upon the three footprint times in seconds since Jan. 1, 1993 |
| sounding_zenith | Float32 | Angle between the LOS as defined from the sounding location to the spacecraft, and the sounding location local zenith direction |

## 8.4 L1B Data Tables

### 8.4.1 SoundingMeasurements

The SoundingMeasurements group (see **Table 8-6** below) contains the calibrated radiance spectra. The spectral data are stored as 32-bit floats whose units are photons/m$^2$/sr/µm.

**Table 8-6**. Calibrated radiance spectra.

| Data Element | Type | Comments |
|---|---|---|
| radiance_o2 | Float32 | Calibrated spectra for ABO2 |
| radiance_strong_co2 | Float32 | Calibrated spectra for SCO2 |
| radiance_weak_co2 | Float32 | Calibrated spectra for WCO2 |
| snr_o2_l1b | Float32 | The mean signal-to-noise ratio of the good samples in ABO2 falling between the 98th and 99th percentile for signal level |
| snr_strong_co2_l1b | Float32 | The mean signal-to-noise ratio of the good samples in SCO2 falling between the 98th and 99th percentile for signal level |

| Data Element | Type | Comments |
|---|---|---|
| snr_weak_co2_l1b | Float32 | The mean signal-to-noise ratio of the good samples in WCO2 falling between the 98th and 99th percentile for signal level |

### 8.4.2 SliceMeasurements

The SliceMeasurements group (see **Table 8-7** below) contains the calibrated radiance values for the color slices. The spectral slice data are stored as 32-bit floats whose units are photons/$m^2$/sr/µm.

**Table 8-7.** Calibrated radiance values for the color slices.

| Data Element | Type | Comments |
|---|---|---|
| radiance_slice_o2 | Float32 | Radiance values for all ABO2 color slice pixels |
| radiance_slice_strong_co2 | Float32 | Radiance values for all SCO2 color slice pixels |
| radiance_slice_weak_co2 | Float32 | Radiance values for all WCO2 color slice pixels |

### 8.4.3 FootprintGeometry

The FootprintGeometry group (see **Table 8-8** below) contains detailed information about the location and observational geometry for each focal plane and spatial footprint.

**Table 8-8.** Location and observational geometry for each focal plane and spatial footprint.

| Data Element | Type | Comments |
|---|---|---|
| footprint_time_tai93 | Float64 | Data acquisition time for the center of footprint in seconds since Jan. 1, 1993 |
| footprint_time_string | String | Data acquisition time for the center of footprint (yyyy-mm-ddThh:mm:ss.mmmZ) |
| footprint_o2_qual_flag | UInt16 | Bit flags indicating the quality of the ABO2 data in footprint:<br>• 0—Good<br>• Non-zero—See **Error! Reference source not found.** |
| footprint_weak_co2_qual_flag | UInt16 | Bit flags indicating the quality of the WCO2 data in footprint:<br>• 0—Good<br>• Non-zero—See **Error! Reference source not found.** |
| footprint_strong_co2_qual_flag | UInt16 | Bit flags indicating the quality of the SCO2 data in footprint:<br>• 0—Good<br>• Non-zero—See **Error! Reference source not found.** |
| footprint_latitude_geoid | Float32 | Geodetic latitude of the footprint based on standard geoid |
| footprint_longitude_geoid | Float32 | Longitude of the footprint based on standard geoid |
| footprint_latitude | Float32 | Geodetic latitude of the footprint center based on Earth topography |
| footprint_longitude | Float32 | Longitude of the footprint center based on Earth topography |
| footprint_altitude | Float32 | Altitude of the footprint center based on Earth topography |
| footprint_altitude_uncert | Float32 | Uncertainty of the source Earth topography data |
| footprint_slope | Float32 | Slope of a plane fit to points within the footprint |
| footprint_plane_fit_quality | Float32 | Goodness of fit - standard deviation of the points to which the plane is fitted, with the expected values taken as the orthogonal projection of the points onto the plane |
| footprint_aspect | Float32 | Orientation of the surface slope relative to the ground track |
| footprint_surface_roughness | Float32 | Standard deviation of the altitude within the footprint |

| Data Element | Type | Comments |
|---|---|---|
| footprint_solar_azimuth | Float32 | Angle between the solar direction as defined from the footprint location to the Sun, and the footprint location local north direction |
| footprint_solar_zenith | Float32 | Angle between the solar direction as defined from the footprint location to the Sun, and the footprint location local zenith direction |
| footprint_azimuth | Float32 | Angle between the LOS as defined from the footprint location to the spacecraft, and the footprint location local north direction |
| footprint_zenith | Float32 | Angle between the LOS as defined from the footprint location to the spacecraft, and the footprint location local zenith direction |
| footprint_vertex_longitude | Float32 | Longitude of the footprint vertices using Earth topography |
| footprint_vertex_latitude | Float32 | Geodetic latitude of the footprint vertices using Earth topography |
| footprint_vertex_altitude | Float32 | Altitude of the footprint vertices using Earth topography |
| footprint_stokes_coefficients | Float32 | Weighting factors applied to the Stokes parameters calculated by the radiative transfer code to compute the radiance |
| footprint_land_fraction | Float32 | Percentage of land surface type within the footprint |
| footprint_polarization_angle | Float32 | The angle between the accepted polarization axis of the instrument and the instrument reference plane for polarization, defined as the plane formed by the LOS and the ray from the footprint location to the local zenith |

### 8.4.4 FrameConfiguration

The FrameConfiguration group (see **Table 8-9** below) contains information about how the detectors and color slices were configured during the observations.

**Table 8-9.** Configuration of detectors and color slices.

| Data Element | Type | Comments |
|---|---|---|
| color_slice_position_o2 | Int16 | Absolute spectral position, in pixels, of each color slice in ABO2 band |
| color_slice_position_strong_co2 | Int16 | Absolute spectral position, in pixels, of each color slice in SCO2 band |
| color_slice_position_weak_co2 | Int16 | Absolute spectral position, in pixels, of each color slice in WCO2 band |
| footprint_spatial_end_position | Uint8 | Position of end of each footprint, in pixels, relative to initial_unused_pixels |
| footprint_spatial_start_position | Uint8 | Position of start of each footprint, in pixels, relative to initial_unused_pixels |
| initial_unused_pixels | Int16 | Distance in pixels of the start of first footprint from edge of FPA |

### 8.4.5 FrameHeader

The FrameHeader group (see **Table 8-10** below) contains the frame identification numbers, times (UTC and TAI), the L1B frame quality flag, and some details of the array clocking.

**Table 8-10.** Frame identification data.

| Data Element | Type | Comments |
|---|---|---|
| clocking_offset_interval | Int16 | The number of pixels between each successive pixel shift used to apply a clocking correction |
| clocking_offset_start | Int32 | The first pixel where the flight software applies a focal plane clocking correction - negative value indicates the clocking shift operates downward, positive value indicates clocking shift is upward |
| frame_id | Int64 | The mission-unique frame identifier |
| frame_qual_flag | UInt64 | Bit flags indicating the quality of the data in each frame:<br>• 0—Good<br>• Non-zero—See **Error! Reference source not found.** |
| frame_time_string | String | Time of telemetry frame (yyyy-mm-ddThh:mm:ss.mmmZ) |
| frame_time_tai93 | Float64 | Time of telemetry frame in seconds since Jan 1, 1993 |

### 8.4.6 FrameTemperatures

The FrameTemperatures group (see **Table 8-11** below) contains selected temperatures for each frame.

**Table 8-11.** Selected temperature data.

| Data Element | Type | Comments |
|---|---|---|
| temp_afe_electronics_enclosure | Float32 | Temperature of the AFE electronics enclosure |
| temp_fpa | Float32 | FPA temperatures extracted for this frame time from the temperature data history |
| temp_optical_bench_grating_mz | Float32 | Temperature of the optical bench grating |
| temp_relay_sco2_mz | Float32 | Temperature of the relay |
| temp_shroud_py_tz1 | Float32 | Temperature of the shroud |
| temp_smooth_fpa_o2 | Float32 | ABO2 FPA temperature after noise-reduction processing |
| temp_smooth_fpa_strong_co2 | Float32 | SCO2 FPA temperature after noise-reduction processing |
| temp_smooth_fpa_weak_co2 | Float32 | WCO2 FPA temperature after noise-reduction processing |
| temp_smooth_optical_bench_grating_mz | Float32 | Optics temperature for this frame after noise-reduction processing |
| temp_telescope | Float32 | Temperature of the telescope |

### 8.4.7 InstrumentHeader

The InstrumentHeader group (see **Table 8-12** below) contains information about the performance of the instrument.

**Table 8-12.** Instrument performance data.

| Data Element | Type | Comments |
|---|---|---|
| ils_delta_lambda | Float32 | Wavelength offset from peak response for sampled data |
| ils_relative_response | Float32 | The relative response defined at ils_delta_lamda |

| Data Element | Type | Comments |
|---|---|---|
| full_width_half_maximum | Float32 | The spectral response width at full-width-half-maximum, per pixel |
| measureable_signal_max_observed | Float32 | Maximum radiance measurable by each spectrometer without detector saturation |
| snr_coef | Float64 | Three coefficients of the noise model: photon component (proportional to the square root of signal), a constant background component and a bad sample flag where 0 - OK, 1 - ignore for retrievals |
| dispersion_coef_samp | Float64 | Coefficients that express the relationship between the spectral element index and its associated wavelength |
| residual_estimate | Float32 | Empirical estimate of the systematic residuals that cannot be removed by calibration |

# 9 Tools and Data Services

## 9.1 HDFView

HDFView is a Java based graphical user interface created by the HDF Group that can be used to browse all ACOS HDF products. The utility allows users to view all objects in an HDF file hierarchy, which is represented as a tree structure. HDFView can be downloaded or support found at: http://www.HDFgroup.org/HDF-java-html/HDFview/.

## 9.2 Panoply

Panoply is a cross-platform application that plots geo-gridded and other arrays from netCDF, HDF, GRIB, and other datasets. See http://www.giss.nasa.gov/tools/panoply/.

## 9.3 Mirador

Level 2 data from OCO-2 is available at the Goddard Earth Sciences Data and Information Services Center (GES DISC). The GES DISC provides basic temporal, advanced (event), and spatial searches through its search and download engine, Mirador (http://mirador.gsfc.nasa.gov). Mirador offers various download options that suit users with different preferences and different levels of technical skills. Users can start from a point where they don't know anything about these particular data, its location, size, format, etc., to quickly find what they need by just providing relevant keywords, like "OCO-2," or "CO2."

## 9.4 JPL CO2 Virtual Science Data Environment

The JPL $CO_2$ Virtual Science Data Environment (co2.jpl.nasa.gov) provides access to and information about satellite observations of carbon dioxide. This site is hosting the first version of daily lite files (released in September 2015). The site is allows users to create custom subsets of Level 2 data products and custom Level 3 gridded data product from OCO-2, ACOS and other satellite instruments. OCO-2 data products will be available from the site at the same schedule as official data releases to the NASA DAAC.

# 10 Contact Information

For the most up-to-date contact information, please refer to oco2.jpl.nasa.gov.

# 11 Acknowledgements and References

## 11.1 Acknowledgements

This research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

## 11.2 Additional Resources

There are a number of other project documents that the user should refer to as they work with the data

1. L1B ATBD—this algorithm theoretical basis document describes the process used to take the uncalibrated spectrum to calibrated radiance spectra
2. L2 ATBD—this ATBD steps through the physics and implementation of the Level 2 algorithm
3. ATBDs for IMAP-DOAS and ABO2—these ATBD documents describe the two methods of identifying potentially cloudy footprints, in what we refer to as the prescreening step. These data are then used for setting data quality and data selection levels
4. Published papers—in addition to the references in the next section, there are a number of published papers describing the algorithm, application to GOSAT, prescreening steps, etc. Please see the most up to date list of publications on oco2.jpl.nasa.gov

## 11.3 References

### 11.3.1 Links

- http://oco2.jpl.nasa.gov

Level 2 algorithm information

- ACOS Level 2 Algorithm Theoretical Basis Document, JPL D-65488

Releases and publications

- http://www.jaxa.jp/press/2009/02/20090209_ibuki_e.html

### 11.3.2 OCO-2 Mission

Canadell, J.G., C. Le Quere, M.R. Raupach, C.B. Field, E.T. Buitenhuis, P. Ciais, T.J. Conway, N.P. Gillett, R.A. Houghton, and G. Marlan (2007), Contributions to accelerating atmospheric CO2 growth from economic activity, carbon intensity, and efficiency of natural sinks, *Proceedings of the National Academy of Sciences*, 47, 18866-18870.

Conway, T.J., P.M. Lang, and K.A. Masarie (2011), Atmospheric carbon dioxide dry air mole fractions from the NOAA ESRL carbon cycle cooperative global air sampling network, 1968-2010, Version: 2011-10-14, http://www.esrl.noaa.gov/gmd/ccgg/trends/global.html

Buchwitz, M., Schneising, O., Burrows, J. P., Bovensmann, H., Reuter, M., and Notholt, J. (2007), First direct observation of the atmospheric CO2 year-to-year increase from space, *Atmos. Chem. Phys.*, 7, 4249–4256, doi:10.5194/acp-7-4249-2007,

Crisp, D., R.M. Atlas, F.-M. Breon, L.R. Brown, J.P. Burrows, P. Ciais, B.J. Connor, S.C. Doney, I.Y. Fung, D.J. Jacob, C.E. Miller, D. O'Brien, S. Pawson, J.T. Randerson, P. Rayner, R.J. Salawitch, S.P. Sander, B. Sen, G.L. Stephens, P.P. Tans, G.C. Toon, P.O. Wennberg, S.C. Wofsy, Y.L. Yung, Z. Kuang, B. Chudasama, G. Sprague, B. Weiss, R. Pollock, D. Kenyon, S. Schroll (2004), The Orbiting Carbon Observatory (OCO) mission, *Advances in Space Research* 34 700–709.

Crisp, D., C. E. Miller, and P. L. DeCola (2008), NASA Orbiting Carbon Observatory: Measuring the column averaged carbon dioxide mole fraction from space, *JARS*.

Frankenberg, C., Chris O'Dell, Joseph Berry, Luis Guanter, Joanna Joiner, Philipp Köhler, Randy Pollock, Thomas E. Taylor (2014a), Prospects for chlorophyll fluorescence remote sensing from the Orbiting Carbon Observatory-2, *Remote Sensing of Environment*, 147, 1-12.

Frankenberg, C., R. Pollock, R. A. M. Lee, R. Rosenberg, J.-F. Blavier, D. Crisp, C.W. O'Dell, G. B. Osterman, C. Roehl, P. O. Wennberg, and D. Wunch (2014b), The Orbiting Carbon Observatory (OCO-2): spectrometer performance evaluation using prelaunch direct Sun measurements, *Atmos. Meas. Tech.*, 7, 1–10. www.atmos-meas-tech.net/7/1/2014/doi:10.5194/amt-7-1-2014.

Hamazaki, T., Kaneko, Y., Kuze, A., and Kondo, K. (2005), Fourier transform spectrometer for greenhouse gases observing satellite (GOSAT), in: *Proceedings of SPIE*, vol. 5659, p. 73-80, doi: 10.1117/12.581198.

Kuze. A., Suto. H., Nakajima, M., and 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, 6716–6733, 15 doi:10.1364/AO.48.006716

H. H. Kieffer and T. C. Stone (2005), The spectral irradiance of the Moon. *Astronom. J.*, 129: 2887–2901.

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

Le Quéré, C., G. P. Peters, R. J. Andres, R. M. Andrew, T. A. Boden, P. Ciais, P. Friedlingstein, R. A. Houghton, G. Marland, R. Moriarty, S. Sitch, P. Tans, A. Arneth, A. Arvanitis, D. C. E. Bakker, L. Bopp, J. G. Canadell, L. P. Chini, S. C. Doney, A. Harper, I. Harris, J. I. House, A. K. Jain, S. D. Jones, E. Kato, R. F. Keeling, K. Klein Goldewijk, A. Körtzinger, C. Koven, N. Lefèvre, F. Maignan, A. Omar, T. Ono, G.-H. Park, B. Pfeil, B. Poulter, M. R. Raupach,*, P. Regnier, C. Rödenbeck, S. Saito, J. Schwinger, J. Segschneider, B. D. Stocker, T. Takahashi, B. Tilbrook, S. van Heuven, N. Viovy, R. Wanninkhof, A. Wiltshire, and S. Zaehle, (2014), Global carbon budget 2013, *Earth Syst. Sci. Data*, 6, 235–263, www.earth-syst-sci-data.net/6/235/2014/ doi:10.5194/essd-6-235-2014.

P.J. Rayner and D.M. O'Brien (2001), The utility of remotely sensed CO2 concentration data in surface source inversions, *Geophys. Res. Lett*. 28, 175-178.

Yoshida, Y., Ota, Y., Eguchi, N., Kikuchi, N., Nobuta, K., Tran, H., Morino, I., and Yokota, T.: (2011), Retrieval algorithm for $CO_2$ and $CH_4$ column abundances from short-wavelength infrared spectral observations by the Greenhouse Gases Observing Satellite, *Atmos. Meas. Tech.*, 4, 717–734, doi:10.5194/amt-4-717-2011.

### 11.3.3 Algorithms and Retrievals

Crisp, D., B. M. Fisher, C.W. O'Dell, C. Frankenberg, R. Basilio, H. Bösch, L. R. Brown, R. Castano, B. Connor, N. M. Deutscher, A. Eldering, D. Griffith, M. Gunson, A. Kuze, L. Mandrake, J. McDuffie, J. Messerschmidt, C. E. Miller, I. Morino, V. Natraj, J. Notholt, D. O'Brien, F. Oyafuso, I. Polonsky, J. Robinson, R. Salawitch, V. Sherlock, M. Smyth, H. Suto, T. Taylor, P. O. Wennberg, D. Wunch, and Y. L. Yung (2012), The ACOS XCO2 retrieval algorithm, Part 2: Global XCO2 data characterization. *Atmos Meas. Tech.*, 5, 687-707.

Frankenberg, C., Platt, U., and Wagner, T. (2005), Iterative maximum *a posteriori* (IMAP-) DOAS for retrieval of strongly absorbing trace gases: Model studies for CH4 and CO2 retrieval from near- infrared spectra of SCIAMACHY onboard ENVISAT, *Atmos. Chem. Phys.*, 5, 9–22.

Lee et al. (2015), Preflight Spectral Calibration of the Orbiting Carbon Observatory-2, in preparation.

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

O'Dell, C.W., B. Connor, H. Bösch, D. O'Brien, C. Frankenberg, R. Castano, M. Christi, D. Crisp, A. Eldering, B. Fisher, M. Gunson, J. McDuffie, C. E. Miller, V. Natraj, F. Oyafuso, I. Polonsky, M. Smyth, T. Taylor, G. C. Toon, P. O. Wennberg, and D. Wunch, (2012), The ACOS CO2 retrieval algorithm, Part 1: Description and validation against synthetic observations. *Atmos. Meas. Tech.*, 5, 99-121.

Taylor, T.E., C. O'Dell, D.M. O'Brien, N. Kikuchi, T. Yokota, T. Nakajima, H. Ishida, D. Crisp, and T. Nakajima (2011), Comparison of cloud screening methods applied to GOSAT near-infrared spectra, IEEE *Trans. Geosci. Rem. Sens*., doi: 10.1109/TGRS.2011.2160270.

Rodgers, C. (2000) Inverse Methods for Atmospheric Sounding: Theory and Practice. World Scientific Publishing Co Pte Ltd.

Taylor, T.E., C. O'Dell, D.M. O'Brien, N. Kikuchi, T. Yokota, T. Nakajima, H. Ishida, D. Crisp, and T. Nakajima (2011), Comparison of cloud screening methods applied to GOSAT near-infrared spectra, IEEE *Trans. Geosci. Rem. Sens*., doi: 10.1109/TGRS.2011.2160270.

Taylor, T.E., C.W. O'Dell, C. Frankenberg, P. Partian, H. W. Cronk, A. Savtchenko, H. R. Pollock, D. Crisp, A. Eldering, and M. Gunson, (2015), Orbiting Carbon Observatory-2 (OCO-2) aerosol and cloud screening; validation against collocated MODIS data, manuscript in preparation.

### 11.3.4 Chlorophyll Fluorescence

Frankenberg, C., Fisher, J., Worden, J., Badgley, G., Saatchi, S., Lee, J.-E., et al. (2011), New global observations of the terrestrial carbon cycle from GOSAT: Patterns of plant fluorescence with gross primary productivity. *Geophysical Research Letters*, 38(17), L17706.

Frankenberg, C., O'Dell, C., Guanter, L., & McDuffie, J. (2012), Remote sensing of near-infrared chlorophyll fluorescence from space in scattering atmospheres: implications for its

retrieval and interferences with atmospheric CO2 retrievals. *Atmospheric Measurement Techniques*, 5(8), 2081–2094, doi: 10.5194/amt-5-2081-2012.

Joiner, J., Y. Yoshida, A. P. Vasilkov, Y. Yoshida3 , L. A. Corp4 , and E. M. Middleton (2011), First observations of global and seasonal terrestrial chlorophyll fluorescence from space, *Biogeosciences*, 8, 637–651.

### 11.3.5 Validation

Papers related to validation of the ACOS data product, plans for OCO-2 data validation or the TCCON network:

Washenfelder et al. (2006), Carbondioxide column abundances at the Wisconsin Tall Tower site, *J. Geophys. Res*, 111, D22305, doi: 10.1029/2006JD007154.

Wunch et al. (2010), Calibration of the Total Carbon Column Observing Network using aircraft profile data, *Atmospheric Measurement Techniques*, 3, 1351–1362, doi: 10.5194/amt-3-1351-2010, http://www.atmos-meas-tech.net/3/1351/2010.

Wunch et al. (2011a), The Total Carbon Column Observing Network, *Phil. Trans. R. Soc. A*, 369, 2087–2112, doi: 10.1098/rsta.2010.0240.

Wunch et al. (2011b), A method for evaluating bias in global measurements of $CO_2$ total columns from Space, *Atmos. Chem. Phys. Discuss., 11,* 20899-20946.

# 12 Acronyms

| | |
|---|---|
| ABO2 | $O_2$ A band cloud screening algorithm |
| ACOS | Atmospheric $CO_2$ observations from space |
| AFE | Analog front-end electronics |
| aod | Aerosol optical depth |
| ARP | Ancillary radiometric product |
| ATBD | Algorithm theoretical basis document |
| $CO_2$ | Carbon dioxide |
| CSU | Colorado State University |
| DN | Data number |
| DOAS | Differential optical absorption spectroscopy |
| ECMWF | European Centre for Medium-Range Weather Forecasts |
| EOS | Earth Observing System |
| EOS A-Train | EOS Afternoon Constellation |
| FP | Full physics (algorithm) |
| FPA | Focal plane array (or assembly) |
| FWHM | Full width at half maximum |
| GES DISC | Goddard Earth Sciences Data and Information Services Center |
| GOSAT | Greenhouse Gases Observing Satellite |
| GSFC | Goddard Space Flight Center |
| $H_2O$ | Water |
| HDF | Hierarchical data format |
| ILS | Instrument line shape |
| IMAP-DOAS | Iterative maximum *a posteriori* differential optical absorption spectroscopy |
| L (0,1,..) | Level 0, 1, etc. data product |
| LOS | Line of sight |
| NASA | National Aeronautics and Space Administration |
| $O_2$ | Oxygen |
| OCO | Orbiting Carbon Observatory |
| OCO-2 | Orbiting Carbon Observatory-2 |
| SAA | South Atlantic Anomaly |
| SCO2 | Strong $CO_2$ band |
| SIF | Solar-induced chlorophyll fluorescence |
| SNR | Signal-to-noise ratio |
| TAI | International atomic time |
| TCCON | Total Carbon Column Observing Network |
| UTC | Coordinated universal time |
| VCD | Vertical column density |
| VMR | Volume mixing ratio |
| WCO2 | Weak $CO_2$ band |
| WMO | World Meteorological Organization |
| $X_{CH4}$ | Column-averaged dry air mole fraction of atmospheric methane |
| $X_{CO2}$ | Column-averaged dry air mole fraction of atmospheric carbon dioxide |