National Aeronautics and Space Administration

Goddard Earth Science Data Information and Services Center (GES DISC)

README Document for Orbiting Carbon Observatory Products

Last Revised 2/18/2021

Prepared By:

Andrey Savtchenko

Name
GES DISC
GSFC Code 610.2

Date 2/18/2021

Name
External Organization

Reviewed By:

Goddard Earth Sciences Data and Information Services Center (GES DISC)
http://disc.gsfc.nasa.gov
NASA Goddard Space Flight Center
Code 610.2
Greenbelt, MD 20771 USA

Reviewer Name

GES DISC

GSFC Code 610.2

Date

Goddard Space Flight Center
Greenbelt, Maryland

Revision History

Revision Date	***Changes***	***Author***
6/3/2015	Document Created	Andrey Savtchenko
12/02/2017	Document updated	Andrey Savtchenko
8/11/2020	Updates for V10	Andrey Savtchenko

Table of Contents

1 Introduction

This README file attempts to address basic most important information on the OCO products, regardless of the data version.

In early August 2020, the OCO project finalized the delivery of the data products output from Algorithm Build 10. This included the most important "Lite" products with bias-corrected xCO2 full physics retrieval and Solar-Induced chlorophyll Fluorescence (SIF).

The new Algorithm Build 10 includes changes in the L2 retrieval algorithm, leading to an improved data set. The list of algorithm and calibration updates in Algorithm Build 10 includes changes to:
• L1B calibration, to account for radiometric degradation and bad samples;
• Trace gas spectroscopic parameters used in the retrieval algorithm (ABSCO V5.1), including to the O_2 A-band scaling;
• The solar continuum model;
• *a priori* information for aerosols and CO_2;
• The treatment of the surface albedo; and
• The constraints used in the Solar-Induced Fluorescence (SIF) retrievals

Documentation, including the new OCO-2 and OCO-3 User Guide and the Data Quality Statement, are posted on the OCO product page:
https://disc.gsfc.nasa.gov/information/documents?title=OCO-2 Documents

An important thing to remember is that the best quality products are output from the so called **"retrospective"** processing flow that exploits the best knowledge of the instrument's calibration. These data are stored as separate collection version recognized by the letter "r" in the collection version, e.g. "10r". It is these **"retrospective"** data that are recommended for use in science research.

The "**forward**" version is considered near real-time, is of lower quality, and is rolled out from the Goddard DAAC archive as the corresponding new "**retrospective**" data becomes available. The availability of the "**retrospective**" data normally has 1 month latency.

Brief overview of OCO-2 mission, retrieval concept, description of the content of the current and planned for public distribution OCO-2 data products, naming conventions, key data fields recommendations for data analysis, and tools to view and search the data products are all provided in the current User Guide from the above Documents link.

OCO-2 employs a dedicated spacecraft with a single instrument. It was launched on July 2, 2014, into a near-polar orbit on an expendable launch vehicle. OCO-2 joined the A-Train formation of satellites on August 3, 2014. The OCO-2 instrument incorporates three high-resolution spectrometers that make coincident measurements of reflected sunlight in the near-infrared CO2 near 1.61 and 2.06 μm, and in molecular oxygen (O2) A-Band at 0.76 μm.

Concerning the spectral domain, high spectral resolving power ($\lambda/\delta\lambda > 20{,}000$) is needed to resolve the CO2 and O2 lines from the adjacent continuum to maximize the sensitivity to small (< 0.3%) variations in total column CO2.

In the spatial domain, the OCO-2 instrument have a narrow swath - it collects 8 soundings over its 0.8-degree wide swath every 0.333 seconds, yielding surface footprints with along-track dimensions < 2.25 km and cross-track dimensions that vary from 0.1 to 1.3 km at nadir.

1.1 Dataset Description

The full suite of OCO-2 products is given in Table 1. These are all data collections currently released by the OCO-2 Science Team.

Table 1

Product	Long Name	Level
OCO2_Att	OCO-2 spacecraft attitude data	0
OCO2_Eph	OCO-2 spacecraft ephemerides	0
OCO2_L1aIn_Sample	Collated, parsed, OCO-2 Science or Calibration Data	1A
OCO2_L1aIn_Pixel	Collated, parsed, OCO-2 Calibration Data	1A
OCO2_L1B_Science	Calibrated, geolocated OCO-2 science spectra	1B
OCO2_L1B_Calibration	Calibrated, geolocated OCO-2 calibration spectA	1B
OCO2_L2_IMAPDOAS	Geolocated retrieved values of XCO2 and fluorescence generated by the IMAP-DOAS algorithm	2
OCO2_L2_Diagnostic	Geolocated XCO2 retrieval results plus algorithm diagnostic information	2
OCO2_L2_Standard	Geolocated XCO2 retrieval results	2
OCO2_L2_Lite_FP	OCO-2 Level 2 bias-corrected XCO2 and other select fields from the full-physics retrieval aggregated as daily files	2
CO2_L2_Lite_SIF	OCO-2 Level 2 bias-corrected solar-induced fluorescence and other select fields from the IMAP-DOAS algorithm aggregated as daily files	2
OCO2_L2_Met	OCO-2 Level 2 meteorological parameters interpolated from global assimilation model for each sounding	2
OCO2_L2_ABand	OCO-2 Level 2 spatially ordered geolocated retrievals using the A-band retrieval algorithm	2
OCO2_L2_CO2Prior	CO-2 Level 2 CO2 prior based on CO2 monthly flask record, global meteorology, and age of air	2

OCO-2 Project also released Absorption Coefficients for researchers willing to run OCO-2 algorithms.

1.2 Algorithm Background and Data Disclaimer

In preparation to work with the final CO2 retrievals from OCO-2, it would be most informative for users to start with the User Guide. This location is also the centralized place that refers to all available documentation approved for release by the OCO-2 Project. We strongly recommend

Users to familiarize with the Data Quality Disclaimer and the Algorithm Theoretical Basis Documents.

1.3 Mailing List

A mailing list is established for everyone interested in updates on a monthly basis. We will notify users if there is new documentation, important announcements about the dataset, etc. We encourage everyone who downloads OCO-2 data to sign up to the oco2_updates email list. To subscribe, send an email to **sympa at list.jpl.nasa.gov** with the subject:
subscribe oco2_updates

2 Data Organization

2.1 Three basic groups of OCO-2 data collections

Publicly available OCO-2 data are organized in Attitude and Ephemeris, Level 1, and Level 2 final retrieval of column CO2 (see Table 1).

2.2 File Naming Convention

The file naming for OCO-2 products products consists from the following elements that are described in Table 2.
oco2_[*ProductId*][*Mode*]_[*Orbit*][*ModeCounter*]_[*AcquisitionDate*]_[*ShortBuildId*]_[*ProductionDateTime*].h5
Users should carefully observe the "Mode" of observation, as the platform and the instrument operates in various of them. For Users who expect to only work with the final xCO2 retrieval, the Glint and Nadir (GL, ND) modes are perhaps the most important.

Table 2

Field	Description	Format	Selection
ProductId	A mnemonic indicating a file type.	String	L1bSc - Level 1B Science product
			L1bCl - Level 1B Calibration product
Mode	The acquisition Mode associated with the data.	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

Field	Description	Format	Selection
			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
			SB - Stand-by
Orbit	The Orbit on which the associated data were acquired. If the Orbit number is less than 10,000, zeros are prepended to the number to ensure that the field is five digits long.	nnnnn	Actual Orbit number for data acquired dur
ModeCounter	This field indicates how many times an acquisition Mode occurs in an Orbit. If a mode occurs only once, ModeCounter is set to "a".	Single character	a, b, c, …
AcquisitionDate	The date (UTC) the data were acquired.	yymmdd	
ShortBuildId	The identification of the related software build	B*stuu*	s = ID of major build cycle t = ID of scheduled build within a major build cycle uu = ID of incremental or patch build
ProductionDateTime	The date and time (UTC) that the file was produced.	yymmddhhmmss	

The granularity of the Lite files is one file per day, that includes best quality sounding for the day. The data files are in NetCDF-4 format which is compatible with HDF5, and the file names have the format:
oco2_LtSIF_<yymmdd>_B<build id>_<production date>.nc4
oco2_LtCO2_<yymmdd>_B<build id>_<production date>.nc4

where the string <yymmdd> reflects the year, month and the day of the data.

2.3 File Format and Structure

All OCO-2 products, except the Lite files, from **Table 1** are in plain HDF5 format. The Lite files are in NetCDF-4 format which is built upon HDF5 and thus is compatible with it. The HDF5 formatting allows conspicuous grouping of datasets that are related by some criteria under folders. The latter can be easily viewed by means of the executable utility "h5dump" with option "-n". Below, a liberally truncated sample is given (replace for the current build id and production time):

h5dump –n oco2_L1bScND_02412a_141214_B<build id>_<production date>.h5

```
group      /Dimensions
 group      /Dimensions/AncFile
……
dataset    /FootprintGeometry/footprint_latitude
……
 dataset    /FrameConfiguration/color_slice_position_o2
……
 dataset    /FrameTemperatures/temp_smooth_fpa_o2
……
dataset    /Metadata/AcquisitionMode
……
dataset    /SoundingGeometry/sounding_latitude
……
 group      /SoundingMeasurements
 dataset    /SoundingMeasurements/radiance_o2
 dataset    /SoundingMeasurements/radiance_strong_co2
 dataset    /SoundingMeasurements/radiance_weak_co2
```

Full description of groups and datasets within groups can be found in the documentation of the corresponding data type.

2.4 Key Science Data Fields

There is a very large number of variables in every OCO-2 HDF5 file, and every data type would have its key data fields. For instance, in the calibrated L1B spectra (product mnemonic "L1bSc" in the file name) the key group of variables can be assumed to be "/SoundingMeasurements" (see above). In Level 2 Full Standard Retrieval, "L2Std" in the file name, the key variables are under /RetrievalResults, but users should pay as much attention to quality fields under /PreprocessingResults and /SpectralParameters. For a description of key variable, users should first refer to the Data User Guide.

2.5 Science Area

NASA successfully launched its first spacecraft dedicated to studying atmospheric carbon dioxide at 2:56 a.m. PDT (5:56 a.m. PDT) on Wednesday, July 2, 2014. Orbiting Carbon

Observatory-2 (OCO-2) is NASA's first dedicated Earth remote sensing satellite to study atmospheric carbon dioxide from space. OCO-2 is an exploratory science mission designed to collect space-based global measurements of atmospheric CO_2 with the precision, resolution, and coverage needed to characterize sources and sinks (fluxes) on regional scales (≥1000km). OCO-2 will also be able to quantify CO_2 variability over the seasonal cycles year after year. This mission will also validate a space-based measurement approach and analysis concept that could be used for future systematic CO_2 monitoring missions.

3 Data Contents

3.1 Dimensions

Some sense of dimensions can be acquired from the group /Dimensions. Although all dimensions are described in that group, admittedly it is not an easy task to relate that dimensions with the data fields they belong to. Still, the value of this group is that each dimension is given a short description.

3.2 Attributes

All Metadata attributes in OCO-2 files are under the group "/Metadata". It contains a large number of Metadata objects. The easiest way to print a filtered list is e.g.:

h5dump –n oco2_L1bScND_02412a_141214_B<build id>_<production date>.h5|grep Metadata

An example of important object there is "/Metadata/BuildId", that identifies the algorithm version used in the processing.

4 Options for Reading the Data

4.1 Command Line Utilities

The HDF Group lists a number of HDF5 command line tools on their website:
https://www.hdfgroup.org/hdf-tools/
Among them, the one that is the most useful for previewing HDF5 content is **h5dump**. It is the "must-have" utility for quick previews of the HDF5 files structure and contents. For instance use to print the algorithm version:

h5dump -d "/Metadata/BuildId" oco2_L1bScND_02412a_141214_B<build id>_<production date>.h5

4.2 Tools for simple visualization and file content view

It should be noted that OCO-2 is a sounder with a very narrow swath, as opposed to imagers like MODIS. Simple L1B swath images, while technically possible, will make little sense for superficial viewing. Just go give a flavor, in L1B swaths, every pixel contains 1016 spectral "colors".

4.2.1 HDFView

Users involved with sensor data debugging and spectroscopy may find **HDFView** very useful. Among the strengths of the tool is that OCO-2 files content can be viewed in a very friendly fashion – all groups are presented as folders where user can easily drill down the file hierarchy. Quick line plots, as cross sections in the spectral or spatial domain, are very easy to view. Numerical data sets can be viewed as spreadsheets, and as images (multi-dimensional datasets). Simple data manipulations are also possible.

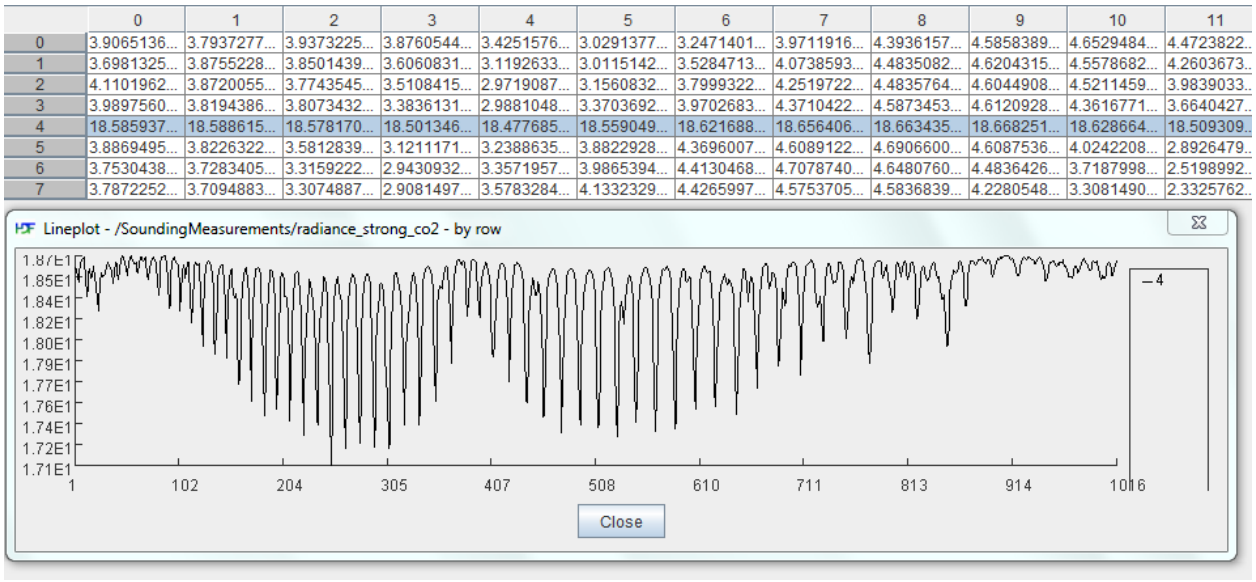

Figure 1. Screen snapshot of HDFView spreadsheet of the 8 OCO-2 footprints, and plot of OCO-2 strong CO2 band (centered at 2.06 μm) spectrum from footprint 4 ($\log_{10}$ is taken).

For download and more information, follow this link:
http://www.hdfgroup.org/hdf-java-html/hdfview/index.html

4.2.2 Commercial

IDL and **MatLAB**, are two major commercial data languages, coming with their libraries that fully support all HDF formats.
Although created for the ACOS Task, the recipes listed at the GES DISC website will give some guidance on how to open and read HDF5 files, and even build some gridding code:

https://docserver.gesdisc.eosdis.nasa.gov/public/project/OCO/acos-data-handling-recipes.html

Since HDF5 files created under the ACOS Task (using GOSAT/TANSO-FTS radiances), and OCO-2, have very similar structure, the recipes are providing simple examples of basic HDF5 functions.

4.2.3 Brief code recipes

The best source of code recipes, including for reading of the OCO data files, in different programming languages is given in the "HDF Zoo":

https://hdfeos.org/zoo/index_openGESDISC_Examples.php

5 Data Services

5.1 Goddard DAAC Unified Interface

The Goddard DAAC transitioned to a content-based web interface. All content, including the datasets, is searchable by appropriate keywords. For example, a good way to search and find the information pages on all available Lite products is:

https://disc.gsfc.nasa.gov/datasets?keywords=oco-2%20lite

The search should produce links to product landing pages, from where all additional information on data access methods and documentation can be found.

5.2 On-line access

Note that authenticated data access has been mandated in 2016. Instruction for registration and data download strategies are provided here:
https://disc.gsfc.nasa.gov/data-access
Since all data are on-line, users can do global recursive downloads wget. The top data directory is:
https://oco2.gesdisc.eosdis.nasa.gov/data/OCO2_DATA/

5.3 OPeNDAP

OPeNDAP stands for "Open-source Project for a Network Data Access Protocol". OPeNDAP is a framework that simplifies all aspects of scientific data networking. It provides simple means for parameter and spatial subset. In the case of OCO-2 Level 1 and 2 data, the simple spatial subset can be materialized by array indexes, not geographical coordinates. In the most simplistic case, OPeNDAP can be used to convert data from HDF5 to NetCDF3, ASCII, and plain binary. The data directory hierarchy, as served by OPeNDAP, can be viewed in any browser:
https://oco2.gesdisc.eosdis.nasa.gov/opendap/
In this case OPeNDAP will be convenient to preview file contents and in particular variables names, dimensions sizes, and quick print of reasonably small variables to the screen.

5.4 Summary Access

To summarize, the new Goddard DAAC interface allows to search and find information by content, like "dataset", "project", "data release", etc. For instance, all datasets currently publicly available from the OCO project can be found using:

https://disc.gsfc.nasa.gov/datasets?keywords=oco&page=1&project=OCO-2

6 More Information

- Detailed Science Team documentation on all OCO-2 products can be found at:

https://disc.gsfc.nasa.gov/information/documents?title=OCO-2 Documents

- GES DISC is also summarizing essential information for every OCO-2 product in product pages that can be accessed from:

  https://disc.gsfc.nasa.gov/datasets?project=OCO

- OCO-2 Science Team at JPL is maintaining an excellent website, where the richest information, from the sensor/spacecraft operations, to science perspectives of CO2 observations, can be found:

http://oco.jpl.nasa.gov/

For further assistance, please use this contact information:

Email: gsfc-dl-help-disc at mail.nasa.gov

Voice: 301-614-5224

Fax: 301-614-5268

Mailing Address:

Code 610.2
Goddard Earth Sciences Data and Information Services Center
NASA Goddard Space Flight Center
Greenbelt, Maryland 20771, U.S.A

7 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.

8 References

Exhaustive list of references can be found in the User Guide.