National Aeronautics and Space Administration
Goddard Earth Science Data Information and Services Center (GES DISC)

README Document for OMPS_NPP_NMHCHO_L2 and OMPS_N20_NMHCHO_L2

Last Revised September 14, 2022

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

Prepared By:

Caroline Nowlan

Name
Center for Astrophysics | Harvard & Smithsonian
60 Garden St., Cambridge, MA, 02138

Gonzalo González Abad

Name
Center for Astrophysics | Harvard & Smithsonian
60 Garden St., Cambridge, MA, 02138

14 September 2022

Date

Reviewed By:

James E. Johnson

Reviewer Name
GES DISC
GSFC Code 610.2

Date

Goddard Space Flight Center
Greenbelt, Maryland

Revision History

Revision Date	Changes	Author
14 September 2022	Initial version	Caroline Nowlan and Gonzalo González Abad

Table of Contents

1.0 Introduction 6

1.1 Dataset and Instrument Description 6

1.1.1 OMPS Instrument Description 6

1.1.2 OMPS NMHCHO dataset 6

1.1.3 Algorithm Background 7

1.2 Data Disclaimer 7

1.3.1 Data Citation and Acknowledgement 8

1.3.2 Contact Information 8

1.3 Quality Issues 8

2.0 Data Organization 8

2.1 File Naming Convention 9

2.2 File Format and Structure 9

2.3 Key Science Data Fields 10

3.0 Data Contents 11

3.1 Global Metadata 11

3.2 Variable Data Attributes 13

3.3 Dimensions 14

4.0 Product 14

4.1 Data Fields 14

4.1.1 key_science_data 14

4.1.2 geolocation 15

4.1.3 qa_statistics 15

4.1.4 support_data 16

4.1.5 uncertainty_budget 17

4.2 Quality Control 18

5.0 Options for Reading the Data 19

5.1 Command Line Utilities 19
5.2 Tools/Programming 19
6.0 GES DISC Data Services 20
7.0 Acknowledgments 20
8.0 References 21
Appendix A. List of Acronyms 22

1.0 Introduction

This document provides information for using Version 1.0 of the OMPS formaldehyde (HCHO) products derived from the Ozone Mapping and Profiler Suite Nadir Mapper (OMPS-NM) Level 1B spectra. The OMPS-NM HCHO products consist of orbital swath total vertical column densities of formaldehyde. The vertical columns are accompanied by support data consisting of uncertainty estimates, geolocation, quality flags and statistics, vertically resolved scattering weights, a priori formaldehyde profiles and ancillary data. This document aims to provide a basic summary and guide for using the data products. For detailed information on the algorithm, please see the Algorithm Theoretical Basis Document (ATBD).

1.1 Dataset and Instrument Description

1.1.1 OMPS Instrument Description

OMPS was launched on the Suomi National Polar-orbiting Partnership (SNPP) satellite on 28 October 2011, and on the JPSS-1 satellite (now known as NOAA-20) on 18 November 2017. OMPS/SNPP consists of the full OMPS suite of three instruments: 1) the OMPS nadir mapper (OMPS-NM), 2) the OMPS profile mapper (OMPS-NP) and 3) the OMPS limb profiler (OMPS-LP). OMPS/NOAA-20 consists only of the nadir package (OMPS-NM and OMPS-NP). The HCHO products described in this document are derived from Level 1B geolocated and calibrated radiance spectra from the OMPS-NM instruments.

The OMPS-NM instruments are hyperspectral nadir viewing spectrometers measuring backscattered light with a spectral resolution of ~1 nm (FWHM). OMPS/SNPP has a spectral range of 300 – 380 nm while OMPS/NOAA-20 measures from 300 – 420 nm. The instruments use 2-D CCD array detectors in pushbroom geometry to observe the two-dimensional field below the satellite's orbit over a swath width of ~2800 km. A detailed description of the instrument and the on-orbit performance of OMPS/SNPP can be found Flynn et al. (2014), Pan et al. (2017), and Seftor et al. (2014).

With 14 or 15 orbits per day, OMPS-NM provides daily global coverage of trace gas columns in the early afternoon local time, with NOAA-20's orbit behind that of SNPP by 50 minutes. Both satellites have a local equatorial crossing time of approximately 13:30. At nadir, OMPS/SNPP observations have a spatial resolution of 50 km × 50 km, with a coarser resolution at the edges of the swath. Most OMPS/NOAA-20 observations have a nadir spatial resolution of 17 km × 17 km up to 13 February 2019. On that date (orbit number 6419), the nominal spatial resolution was changed to 12 km × 17 km. Some high-resolution observations were also collected during the early months of the OMPS/NOAA-20 mission.

1.1.2 OMPS NMHCHO dataset

OMPS HCHO consists of two products: 1) HCHO from OMPS on Suomi NPP (OMPS_NPP_NMHCHO_L2) and 2) HCHO from OMPS on NOAA-20 (OMPS_N20_NMHCHO_L2). OMPS/SNPP Level 2 HCHO data are available from January 2012 onwards, while OMPS/NOAA-20 data begin in January 2018. These products are processed with the same algorithm and

nearly identical ancillary inputs. Each product contains the total vertical column density (VCD) HCHO, VCD uncertainty and quality flags, as well as support data used in the VCD derivation. The files also contain vertically-resolved scattering weights which describe the vertical sensitivity of the retrieval to different layers in the atmosphere.

1.1.3 Algorithm Background

The OMPS HCHO retrieval uses a three-step approach. First, we retrieve a differential slant column density, ΔSCD, of a trace gas in a nadir spectrum as compared to a clean reference spectrum using the OMPS L1B spectra (Jaross, 2017b). This is done by fitting a modeled spectrum to match the measured radiance spectrum of backscattered light in each ground pixel observation. For HCHO, we use a clean nadir reference spectrum determined from measurements over the Pacific. In the second step, we calculate an air mass factor (AMF) which describes the path of light through the atmosphere. The AMF is primarily a function of viewing geometry, aerosol and molecular scattering, surface reflectance, trace gas profile shape, cloud fraction and height, and of the strong absorption features of ozone in the UV which inhibit the penetration of photons close to the surface. Third, the retrieved SCD is corrected by adding the estimated background column SCD_{Ref} (determined from a chemical transport model) in the "clean" reference and for any remaining small latitude-dependent biases, SCD_B. These biases can sometimes occur at high latitudes due to unresolved calibration and spectroscopic issues. The final VCD is determined using:

$$VCD = \frac{\Delta SCD + SCD_{Ref} + SCD_B}{AMF}. \quad (1)$$

The AMF is calculated for each individual ground pixel using the VLIDORT radiative transfer model (Spurr, 2006) and climatological trace gas profiles from a 2018 GEOS-Chem global chemical transport model simulation at $0.5^\circ \times 0.5^\circ$ resolution at the time of overpass. The surface reflectance for the AMF uses an observation-geometry dependent bidirectional reflectance distribution function (BRDF) from MODIS (Schaaf et al., 2002; Wang et al., 2018), extended to the UV with surface albedo EOFs (Zoogman et al., 2016) and a SCIAMACHY surface albedo database (Tilstra et al., 2017). The surface reflectance over water is approximated using a Cox-Munk slope distribution.

While an independent cloud fraction and pressure retrieval is publicly available for OMPS/SNPP (Vasilkov et al., 2014), such a product does not currently exist for OMPS/NOAA-20. In order to be consistent between OMPS/SNPP and OMPS/NOAA-20 HCHO, the AMF calculation uses a cloud fraction derived from the OMPS-NM total ozone product reflectivity (Jaross, 2017a). The cloud pressure is from the total ozone product, which provides a cloud pressure climatology determined from OMI.

1.2 Data Disclaimer

Formaldehyde retrievals are provided for orbits that have valid publicly-available Level 1B calibrated radiances and total ozone files. As of June 2022, OMPS/SNPP L1B radiances and total

ozone were available through the GES DISC and OMPS/NOAA-20 radiances and total ozone were available through the OMPS website (https://ozoneaq.gsfc.nasa.gov/data/omps).

The OMPS instrument collects spectra on the dark side of the Earth for calibration purposes with a typical frequency of once per week. These orbit files will be twice as large in size as the nominal files. While the sunlit part of the orbit will contain valid data, the dark side of the orbit does not contain useful data. These ground pixels can be discarded using the main quality flag or a solar zenith angle (SZA) filter < 90°.

1.3.1 Data Citation and Acknowledgement

When using these data in your publication please cite the data product:

Gonzalo González Abad (2022), OMPS-NPP L2 NM Formaldehyde (HCHO) Total Column swath orbital, Greenbelt, MD, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC), Accessed:[Data Access Date], 10.5067/IIM1GHT07QA8

Gonzalo González Abad (2022), OMPS-N20 L2 NM Formaldehyde (HCHO) Total Column swath orbital, Greenbelt, MD, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC), Accessed:[Data Access Date], 10.5067/CIYXT9A4I2F4

References to the OMPS HCHO retrieval paper (Nowlan et al., 2022, in preparation) and the OMPS HCHO validation paper (Kwon et al., 2022, in preparation) should also be included in your publication.

1.3.2 Contact Information

Data inquiries should be addressed via email to Dr. Caroline Nowlan (cnowlan@cfa.harvard.edu) or Dr. Gonzalo González Abad (ggonzale@cfa.harvard.edu).

1.3 Quality Issues

Some orbits may be missing due to missing radiance files, missing total ozone files, satellite maneuvers or outages, invalid reference spectra, or other reasons. On occasion, Level 2 files may be available but do not contain valid data (usually appearing as NaN values) due to one of the above reasons or because of invalid geolocation or other input data.

2.0 Data Organization

Each OMPS HCHO file contains Level 2 swath data for a single orbit. There are typically 14 to 15 orbits per day.

Most OMPS/SNPP files have 36 pixels across the track and 400 pixels along the track.

Most OMPS/NOAA-20 files have either 104 pixels across the track and 1201 pixels along the track (prior to orbit 6419 on 13 February 2019), or 140 pixels across the track and 1201 pixels along the track (orbit 6419 and later).

2.1 File Naming Convention

OMPS HCHO files are named following the schema:
<Sensor>-<Platform>_<Product>-<Level>_<Version>_<BeginDateTime>-o<OrbitNumber>_<ProductionDateTime>.<suffix>

Where:
<Sensor> = OMPS
<Platform> = NPP or N20
<Product> = NMHCHO
<Level> = L2
<Version> = version number, e.g., v1.0
<BeginDateTime> = the start date and time in format <YYYYmMMDDtHHMMSS>
<OrbitNumber> = 6-character string representing orbit number (e.g., 005961)
<ProductionDateTime> = date and time of data production in format <YYYYmMMDDtHHMMSS>
<suffix> = nc

Filename example:
OMPS-N20_NMHCHO-L2_v1.0_2019m0112t111052-o005961_2022m0517t211821.nc

2.2 File Format and Structure

The OMPS_NPP_NMHCHO_L2 and OMPS_N20_NMHCHO_L2 files are in netCDF (version 4) format. NetCDF is a set of software libraries and self-describing, machine-independent data formats that support the creation, access and sharing of array-oriented datasets. It was developed by UCAR/Unidata (http://doi.org/10.5065/D6H70CW6) https://www.unidata.ucar.edu/software/netcdf/.

The information is divided into five main groups:

1. **key_science_data**: the HCHO column, uncertainty and main data quality flag
2. **geolocation**: information on observation time, latitude, longitude, viewing and solar angles, time and terrain height at observation surface location
3. **qa_statistics**: fit convergence statistics and flags, RMS fitting residuals
4. **support_data**: support data used in the VCD calculation, including fitted slant column, air mass factor, cloud and surface information. This group also contains the vertically-resolved scattering weights.
5. **uncertainty_budget**: uncertainty estimates in key parameters

Orbital metadata are included as global keyword:value pairs.

Each orbital swath contains HCHO observations on a two-dimensional swath of ground pixels. These ground locations have dimensions *along_track* and *cross_track*. Fields with a vertical dimension additionally have a dimension *vertical_layer*. Fields that describe the latitude and longitude bounds of each ground pixel have the additional dimension *corner*. The dimension *vertical_level* is also given in the file. This defines the layer edges, where *vertical_level=vertical_layer+1*.

2.3 Key Science Data Fields

The variables included in the key_science_data group are the ones most likely to be used by users. These include the column_amount, column_amount_uncertainty and main_data_quality_flag. Other key data includes the scattering_weights in the support_data group. These can be used to recalculate AMFs with user-supplied profiles (Palmer et al., 2001) through:

$$AMF = \int_z \ w(z)S(z)dz \quad (2)$$

where *w(z)* is the scattering weight at layer *z* (provided in the file) and *S(z)* is the shape factor of the a priori profile, which is the partial column of HCHO in the layer normalized by the total column of HCHO (a new a priori profile can also be provided by the user).

Profile pressures can be reconstructed using the surface_pressure and its attributes eta_a and eta_b. The bottom boundary (pressure level) of a layer *i* is defined by:

$$p(i) = \ eta\_a(i) \ + \ surface\_pressure \ * \ eta\_b(i) \quad (3)$$

with the top pressure level of the layer defined as:

$$p(i+1) = \ eta\_a(i+1) \ + \ surface\_pressure \ * \ eta\_b(i+1). \quad (4)$$

Other key data fields for reconstruction of the column_amount (VCD) in Equation 1 include the support_group fields amf (AMF), fitted_slant_column_amount (ΔSCD), ref_sector_correction (SCD_{Ref}) and bias_correction (SCD_B).

3.0 Data Contents

3.1 Global Metadata

In addition to arrays containing geophysical quantities, support variables, and dimension scales, global metadata are also stored in the file. Some metadata are required by standard conventions, some are included to meet data provenance requirements and others as a convenience to users of the OMPS_NPP_NMHCHO_L2 and OMPS_N20_NMHCHO_L2 products. A summary of metadata global attributes present in all files is shown in Table 1.

Global Attribute	**Description**	**Type**
InputOriginalFile	Comma-separated list of input files used to generate the HCHO product (1. Level 1B radiance file, 2. Level 1B radiance reference file, 3. Level 1B irradiance file – this is not used for HCHO, 4. source of cloud information)	string
ContactPersonEmail	Email address of the responsible person	string
ContactPersonName	Name of the responsible person	string
ContactPersonRole	Role of responsible person	string
contributor_name	Names of contributors	string
contributor_role	Roles of contributors	string
creator_email	Email of person principally responsible for creating this data	string
creator_name	Name of person principally responsible for creating this data	string
creator_role	Role of person principally responsible for creating this data	string
Conventions	CF metadata convention	string
DataSetQuality	Description of the data quality	string
Format	Format of data (netCDF-4)	string
id	Dataset identifier (same as ShortName)	string
IndentifierProductDOI	Product DOI identifier	string
IdentifierProductionDOIAuthority	http://dx.doi.org/	string
institution	Name of institution responsible for originating data	string
instrument	Name of contributing instrument	string
instrument_vocabulary	Controlled vocabulary for names in “instrument” attribute	string
keywords	Comma-separated list of key words and phrases describing dataset	string
keywords_vocabulary	Controlled vocabulary for names/phrases in “keywords” attribute	string
LongName	Descriptive OMHCHO product name	string
metadata_link	URL to complete metadata	string
ObservationArea	Spatial coverage of the OMHCHO dataset	string
platform	Name of platform supporting sensor	string
platform_vocabulary	Controlled vocabulary for name in “platform” attribute	string
processing_level	Level of data processing	string
ProductGenerationAlgorithm	Algorithm software used to generate the file	string

ProductGenerationAlgorithmVersion	Version of the OMHCHO algorithm	string
program	Overarching program of which dataset is a part	string
project	Comma-separated list of projects responsible for originating data	string
references	References describing data and production	string
ShortName	Abbreviated name of the product	string
source	Instrument origin of the product	string
summary	Summary of dataset	string
title	Short phrase describing dataset	string
VersionID	ECS collection identifier	32-bit integer
license	License restriction ("Freely Distributed")	string
publisher_name	Name of group responsible for publishing data file	string
publisher_email	Email of group responsible for publishing data file	string
publisher_type	Type of publisher responsible for data file	string
publisher_url	Name of group responsible for publishing data file	string
publisher_institution	Institution responsible for publishing data file	string
OrbitNumber	Orbit number	32-bit integer
DayofYear	Day of year	32-bit integer
time_coverage_start	Time of first data point in dataset (format YYYY-MM-DDThh:mm:ss.sssZ)	string
time_coverage_end	Time of last data point in dataset (format YYYY-MM-DDThh:mm:ss.sssZ)	string
RangeBeginningDate	Date of first data point in dataset (format YYYY-MM-DD)	string
RangeBeginningTime	UTC time of first data point in dataset (format hh:mm:ss.sssZ)	string
RangeEndingDate	Date of last data point in dataset (format YYYY-MM-DD)	string
RangeEndingTime	UTC time of last data point in dataset (format hh:mm:ss.sssZ)	string
EquatorCrossingDate	Date of equator crossing time (format YYYY-MM-DD)	string
EquatorCrossingTime	UTC time of equator crossing time (format hh:mm:ss.sss)	string
EquatorCrossingLongitude	Equator crossing longitude (degrees_east)	32-bit floating point
GranuleID	OMPS HCHO product filename	string
ProductionDateTime	Date and time of production (format YYYY-MM-DDThh:mm:ss.sssZ)	string
geospatial_bounds	Polygons describing data's 2D geospatial extent. Coordinates are latitude (degrees_north) and longitude (degrees_east).	string
history	Character array with line for each invocation of program that modified dataset	string

Table 1. File metadata

3.2 Variable Data Attributes

Table 2 shows a list of common data field metadata stored as attributes (keyword:values) for each variable. Not all metadata fields apply to all data fields. Table 3 gives attributes that only apply to specific data fields.

Data Field Attribute	Description	Type
_FillValue	Fill value or missing value	data field type
long_name	Data field long name	string
comment	Additional description about the data field	string
units	Data field units (geophysical units or "1" for fraction)	string
valid_min	Data field minimum value. Values below valid_min should likely be discarded.	data field type
valid_max	Data field maximum value. Values above valid_max should likely be discarded.	data field type
coordinates	Dimension coordinates of the data field	string
_ChunkSizes	netCDF4 chunking size	unsigned integer

Table 2. Common data field attributes

Data Field Attribute	Data Fields Using Attribute	Description	Type
standard_name	geolocation/latitude geolocation/longitude geolocation/time	Standard name for data field	string
bounds	geolocation/latitude geolocation/longitude	Bounds of data field	string
flag_values	key_science_data/main_data_quality_flag qa_statistics/fit_convergence_flag	Possible flag values	data field type
flag_masks	support_data/glint_flag	Possible mask values	data field type
flag_meanings	key_science_data/main_data_quality_flag qa_statistics/fit_convergence_flag support_data/glint_flag	Definition of flag values	string
wavelength	support_data/amf	Wavelength at which AMF was calculated	64-bit floating point
wavelength_unit	support_data/amf	Unit of wavelength attribute	string
radiance_reference_granule	support_data/ref_sector_correction	Level 2 file used in radiance reference correction	string
bias_granule	support_data/bias_correction	List of Level 2 files used in bias correction	string
eta_a	support_data/surface_pressure	Eta level A at pressure layer edge, with dimensions vertical_level	64-bit floating point

eta_b	support_data/surface_pressure	Eta level B at pressure layer edge, with dimensions vertical_level	64-bit floating point

Table 3. Specific data field attributes

3.3 Dimensions

Table 4 lists dimensions associated with the data fields. Not all data fields use all dimensions. Only nominal dimensions are given (in particular, the along_track dimension varies occasionally).

Global Attribute	Description	Nominal Dimensions
along_track	Number of ground pixels along the satellite track	400 (NPP) 1201 (NOAA-20)
cross_track	Number of ground pixels across the satellite track	36 (NPP) 104 (NOAA-20, Orbits 1 – 6418) 140 (NOAA-20, Orbits 6419 – present)
corner	Number of corners in latitude and longitude bounds	4
vertical_layer	Number of layers in data fields with vertical information	47
vertical_level	Number of levels in data fields with vertical information	48

Table 4. Global file dimensions

4.0 Product

4.1 Data Fields

The following tables list the data fields in each group.

4.1.1 key_science_data

Data Field Name	Description	Type	Dimensions	Units
column_amount	HCHO column amount	64-bit floating-point	along_track, cross_track	molecules/cm^2
column_uncertainty	HCHO column amount uncertainty. This is derived from the random uncertainty in the slant column spectral fit.	64-bit floating-point	along_track, cross_track	molecules/cm^2

main_data_quality_flag	main data quality flag	16-bit integer	along_track, cross_track	none

Table 5: Data fields in *key_science_data* group

4.1.2 geolocation

Data Field Name	Description	Type	Dimensions	Units
latitude	Latitude at pixel center	32-bit floating-point	along_track, cross_track	degrees north
latitude_bounds	Latitude at pixel corners (SW,SE,NE,NW)	32-bit floating-point	along_track, cross_track, corner	degrees north
longitude	Longitude at pixel center	32-bit floating-point	along_track, cross_track	degrees east
longitude_bounds	Longitude at pixel corners (SW,SE,NE,NW)	32-bit floating-point	along_track, cross_track, corner	degrees east
solar_zenith_angle	Solar zenith angle at pixel center	32-bit floating-point	along_track, cross_track	degrees
solar_azimuth_angle	Solar azimuth angle at pixel center	32-bit floating-point	along_track, cross_track	degrees
relative_azimuth_angle	Relative azimuth angle at pixel center	32-bit floating-point	along_track, cross_track	degrees
terrain_height	Terrain height	16-bit integer	along_track, cross_track	m
time	Exposure start time in seconds since 1993-01-01T00:00:00Z	64-bit floating-point	along_track	seconds
viewing_zenith_angle	Viewing zenith angle at pixel center	32-bit floating-point	along_track, cross_track	degrees
viewing_azimuth_angle	Viewing azimuth angle at pixel center	32-bit floating-point	along_track, cross_track	degrees

Table 6: Data fields in *geolocation* group

4.1.3 qa_statistics

Data Field Name	Description	Type	Dimensions	Units
fit_convergence_flag	Slant column fit convergence flag	16-bit integer	along_track, cross_track	none
fit_rms_residual	Normalized radiance fit RMS residual	64-bit floating-point	along_track, cross_track	none

num_good_input	Number of pixels for which slant column fitting is attempted	32-bit integer	1	none
percent_bad_output	Percent of num_good_input flagged as “bad” in main quality flag	32-bit floating-point	1	%
percent_good_output	Percent of num_good_input flagged as “good” in main quality flag	32-bit floating-point	1	%
percent_suspect_output	Percent of num_good_input flagged as “suspect” in main quality flag	32-bit floating-point	1	%

Table 7: Data fields in *qa_statistics* group

4.1.4 support_data

Data Field Name	Description	Type	Dimensions	Units
albedo	Geometry-dependent surface Lambertian-Equivalent Reflectivity. This is not used in the AMF calculation but is given to help user estimate effective surface reflectivity.	32-bit floating-point	along_track, cross_track	none
amf	Calculated air mass factor	32-bit floating-point	along_track, cross_track	none
bias_correction	Bias correction	32-bit floating-point	along_track, cross_track	molecules/cm^2
brdf_geo	Amplitude of Li-Sparse BRDF kernel	32-bit floating-point	along_track, cross_track	none
brdf_iso	Amplitude of isotropic BRDF kernel	32-bit floating-point	along_track, cross_track	none
brdf_vol	Amplitude of Ross-Thick BRDF kernel	32-bit floating-point	along_track, cross_track	none
cloud_fraction	Effective cloud fraction used in AMF computation	32-bit floating-point	along_track, cross_track	none
cloud_pressure	Cloud pressure used in AMF computation	32-bit floating-point	along_track, cross_track	hPa
fitted_slant_column_amount	Fitted slant column density	64-bit floating-point	along_track, cross_track	molecules/cm^2
fitted_slant_column_uncertainty	Fitted slant column density uncertainty	64-bit floating-point	along_track, cross_track	molecules/cm^2

gas_profile	A priori gas mixing ratio profile used in AMF calculation	32-bit floating-point	vertical_layer, along_track, cross_track	none
glint_flag	Flag for possible glint	byte	along_track, cross_track	none
ice_fraction	Sea ice fraction	32-bit floating-point	along_track, cross_track	none
land_fraction	Land fraction	32-bit floating-point	along_track, cross_track	none
meridional_wind	Meridional wind	32-bit floating-point	along_track, cross_track	m/s
ocean_salinity	Ocean salinity in Practical Salinity Units (PSU)	32-bit floating-point	along_track, cross_track	g/kg (1e-3)
ref_sector_correction	Reference sector background correction	32-bit floating-point	along_track, cross_track	molecules/cm^2
scattering_weights	Scattering weights	32-bit floating-point	vertical_layer, along_track, cross_track	none
snow_fraction	Snow fraction	32-bit floating-point	along_track, cross_track	none
surface_pressure	Surface pressure	32-bit floating-point	along_track, cross_track	hPa
temperature_profile	Temperature profile	32-bit floating-point	vertical_layer, along_track, cross_track	K
zonal_wind	Zonal wind	32-bit floating-point	along_track, cross_track	m/s

Table 8: Data fields in *support_data* group

4.1.5 uncertainty_budget

Data Field Name	Description	Type	Dimensions	Units
amf_total_uncert	Total AMF uncertainty. This field is empty in Version 1.	32-bit floating-point	along_track, cross_track	%
bias_uncertainty	Estimated uncertainty in bias correction, assuming AMF uncertainty of 50%	32-bit floating-point	along_track, cross_track	molecules/cm^2
ref_sector_uncertainty	Estimated uncertainty in reference background correction, assuming AMF uncertainty of 50%	32-bit floating-point	along_track, cross_track	molecules/cm^2

Table 9: Data fields in *uncertainty_budget* group

4.2 Quality Control

Most users should filter data by key_science_data/main_data_quality_flag. Pixels flagged as "bad" are almost certainly not useful. Pixels flagged as "suspect" may still be usable in some cases but should be used with caution.

<u>Pixels are flagged as "suspect" if any of the following criteria are met:</u>
snow_fraction > 0
ice_fraction > 0
column_amount + 2*column_amount_uncertainty < 0
$AMF_G > 4$

<u>Pixels are flagged as "bad" if any of the following criteria are met:</u>
|column_amount| > 2e17 molecules/cm^2
column_amount + 3*column_amount_uncertainty < 0
amf < 0.1
$AMF_G > 5$

Where the geometric air mass factor is:
AMF_G = sec(solar_zenith_angle) + sec(viewing_zenith_angle)

We recommend that most users further limit the use of HCHO data to SZA < 70° and cloud fractions < 0.4. At large SZA, the signal-to-noise is severely degraded and low radiance spectra typically show more biases in the retrieval. At higher cloud fractions, the systematic uncertainties become very large. In addition, we recommend users filter out ground pixels over ice and snow using the ice and snow fractions in the support data. The ice and snow fractions are not currently used for the AMF calculation, but are included for convenience in data interpretation. While slant column retrievals over these surfaces are usually valid, the cloud fraction retrievals may be unable to differentiate between a bright surface and clouds and will therefore be inaccurate.

5.0 Options for Reading the Data

5.1 Command Line Utilities

ncdump
With ncdump it is possible to generate CDL text representations of OMPS HCHO netCDF datasets. A full description of the capabilities of ncdump can be found here:
https://www.unidata.ucar.edu/software/netcdf/netcdf/ncdump.html
To quickly explore the structure of an OMPS HCHO file, outputting it to struc.txt, issue the following command:

ncdump –c <filename> > struc.txt

or if coordinate variable values (longitude and latitude) are not desired in the ouput:

ncdump –h <filename> > struct.txt

5.2 Tools/Programming

HDFView
HDFView is a Java-based graphical user interface created by the HDF Group, which can be used to browse the OMPS HCHO files. The utility allows users to view all objects in an HDF file hierarchy, which is represented as a tree structure.
HDFView documentation and downloads are available at
https://www.hdfgroup.org/downloads/hdfview/

Panoply
OMPS HCHO datasets are geo-referenced and can be visualized using Panoply, a Java based graphical user interface capable of plotting arrays from netCDF files. Further information and download options are available at
https://www.giss.nasa.gov/tools/panoply/

netCDF programing interfaces are available for major high-level languages including IDL, Matlab, R, and Python.

6.0 GES DISC Data Services

OMPS HCHO products are archived and distributed by the Goddard Earth Science Data & Information Services Center (GES DISC). The files can be directly downloaded from the GES DISC or search using NASA's EarthData web services which provides capabilities for spatial and temporal subsetting. GES DISC provides a list of tools that can read netCDF-4 files. To download GES DISC data you must (1) register in Earthdata Login and (2) be authorized for NASA GES DISC Data Access.

Registering and downloading data with Earthdata can be achieved here:
https://disc.gsfc.nasa.gov/

If you need assistance or wish to report a problem:
Email: gsfc-dl-help-disc@mail.nasa.gov
Voice: 301-614-5224
Fax: 301-614-5268
Address:
Goddard Earth Sciences Data and Information Services Center, NASA Goddard Space Flight Center Code 610.2 Greenbelt, MD 20771 USA

7.0 Acknowledgments

The OMPS HCHO products were developed under support from NASA programs Making Earth System Data Records for Use in Research Environments (80NSSC18M0091) and The Science of Terra, Aqua and Suomi-NPP (80NSSC18K0691), and NOAA's Climate Program Office's Atmospheric Chemistry, Carbon Cycle, and Climate program (NA18OAR4310108). Algorithm development was performed on the Smithsonian High Performance Cluster (SI/HPC), Smithsonian Institution (https://doi.org/10.25572/SIHPC).

We would like to thank the personnel at the Ozone SIPS for their efforts implementing the operational OMPS HCHO algorithm (Phil Durbin in particular) and GES DISC personnel (James Johnson in particular) for facilitating the public release of the OMPS HCHO datasets.

8.0 References

Flynn, L., Long, C., Wu, X., Evans, R., Beck, C. T., Petropavlovskikh, I., et al. (2014). Performance of the Ozone Mapping and Profiler Suite (OMPS) products. *Journal of Geophysical Research: Atmospheres*, *119*, 6181–6195. https://doi.org/10.1002/2013JD020467

Jaross, G. (2017a). OMPS-NPP L2 NM Ozone (O3) Total Column swath orbital V2. Greenbelt, MD, USA: Goddard Earth Sciences Data and Information Services Center (GES DISC). https://doi.org/10.5067/0WF4HAAZ0VHK

Jaross, G. (2017b). OMPS/NPP L1B NM Radiance EV Calibrated Geolocated Swath Orbital V2. Greenbelt, MD, USA: Goddard Earth Sciences Data and Information Services Center (GES DISC). https://doi.org/10.5067/DL081SQY7C89

Palmer, P. I., Jacob, D. J., Chance, K., Martin, R. V., Spurr, R. J. D., Kurosu, T. P., et al. (2001). Air mass factor formulation for spectroscopic measurements from satellites: Application to formaldehyde retrievals from the Global Ozone Monitoring Experiment. *Journal of Geophysical Research: Atmospheres*, *106*(D13), 14539–14550. https://doi.org/10.1029/2000JD900772

Pan, C., Weng, F., & Flynn, L. (2017). Spectral Performance and Calibration of the Suomi NPP OMPS Nadir Profiler Sensor. *Earth and Space Science*, *4*(12), 737–745. https://doi.org/10.1002/2017EA000336

Schaaf, C. B., Gao, F., Strahler, A. H., Lucht, W., Li, X., Tsang, T., et al. (2002). First operational BRDF, albedo nadir reflectance products from MODIS. *Remote Sensing of Environment*, *83*(1–2), 135–148. https://doi.org/10.1016/S0034-4257(02)00091-3

Seftor, C. J., Jaross, G., Kowitt, M., Haken, M., Li, J., & Flynn, L. E. (2014). Postlaunch performance of the Suomi National Polar-orbiting Partnership Ozone Mapping and Profiler Suite (OMPS) nadir sensors. *Journal of Geophysical Research: Atmospheres*, *119*(7), 4413–4428. https://doi.org/10.1002/2013JD020472

Spurr, R. J. D. (2006). VLIDORT: A linearized pseudo-spherical vector discrete ordinate radiative transfer code for forward model and retrieval studies in multilayer multiple scattering media. *Journal of Quantitative Spectroscopy and Radiative Transfer*, *102*(2), 316–342. https://doi.org/10.1016/j.jqsrt.2006.05.005

Tilstra, L. G., Tuinder, O. N. E., Wang, P., & Stammes, P. (2017). Surface reflectivity climatologies from UV to NIR determined from Earth observations by GOME-2 and SCIAMACHY. *Journal of Geophysical Research: Atmospheres*, *122*(7), 4084–4111. https://doi.org/10.1002/2016JD025940

Vasilkov, A., Joiner, J., & Seftor, C. (2014). First results from a rotational Raman scattering cloud

algorithm applied to the Suomi National polar-orbiting partnership (NPP) Ozone mapping and profiler suite (OMPS) Nadir Mapper. *Atmospheric Measurement Techniques*, *7*(9), 2897–2906. https://doi.org/10.5194/amt-7-2897-2014

Wang, Z., Schaaf, C. B., Sun, Q., Shuai, Y., & Román, M. O. (2018). Capturing rapid land surface dynamics with Collection V006 MODIS BRDF/NBAR/Albedo (MCD43) products. *Remote Sensing of Environment*, *207*(February), 50–64. https://doi.org/10.1016/j.rse.2018.02.001

Zoogman, P., Liu, X., Chance, K., Sun, Q., Schaaf, C., Mahr, T., & Wagner, T. (2016). A climatology of visible surface reflectance spectra. *Journal of Quantitative Spectroscopy and Radiative Transfer*, *180*, 39–46. https://doi.org/10.1016/j.jqsrt.2016.04.003

Appendix A. List of Acronyms

Acronym	Meaning
AMF	Air Mass Factor
ATBD	Algorithm Theoretical Basis Document
BRDF	Bidirectional reflectance distribution function
CDL	Common Data Language
CF	Climate and Forecast
DOI	Digital Object Identifier
ECS	EOSDIS Core System
EOF	Empirical Orthogonal Function
EOS	Earth Observing System
EOSDIS	EOS Data and Information System
FWHM	Full width at half maximum
GEOS-Chem	Goddard Earth Observing System chemical transport model
GES DISC	Goddard Earth Sciences Data and Information Services Center
HDF	Hierarchical Data Format
JPSS	Joint Polar Satellite System
L1B	Level-1B (calibrated radiances or irradiances)
L2	Level-2 (retrieved geophysical values)
LP	Limb Profiler
MEaSUREs	Making Earth Science Data Records for Use in Research Environments
MODIS	Moderate Resolution Imaging Spectroradiometer
N20	NOAA-20
NE	Northeast

NetCDF	Network Common Data Form
NM	Nadir Mapper
NOAA	National Oceanic and Atmospheric Administration
NP	Nadir Profiler
NW	Northwest
OMI	Ozone Monitoring Instrument
OMPS	Ozone Mapping and Profiling Suite
RMS	Root Mean Square
SCD	Slant Column Density
SCIAMACHY	SCanning Imaging Absorption spectroMeter for Atmospheric CHartographY
SE	Southeast
SIPS	Science Investigator-led Processing System
SNPP	Suomi National Polar-orbit Partnership
SW	Southwest
SZA	Solar Zenith Angle
UCAR	University Corporation for Atmospheric Research
UTC	Universal Time Coordinated
VCD	Vertical Column Density
VLIDORT	vector linearized discrete ordinate radiative transfer

Table 10. List of acronyms and abbreviations.