Readme Document for Level 2 Ozone Monitoring Instrument (OMI) Formaldehyde Vertical Column Density Research Product based on Principal Component Analysis Retrieval Technique

Can Li, University of Maryland & Code 614, NASA Goddard Space Flight Center
Sergey Marchenko, Science Systems and Applications, Inc.
Nickolay A. Kroktov, Code 614, NASA Goddard Space Flight Center
Joanna Joiner, Code 614, NASA Goddard Space Flight Center

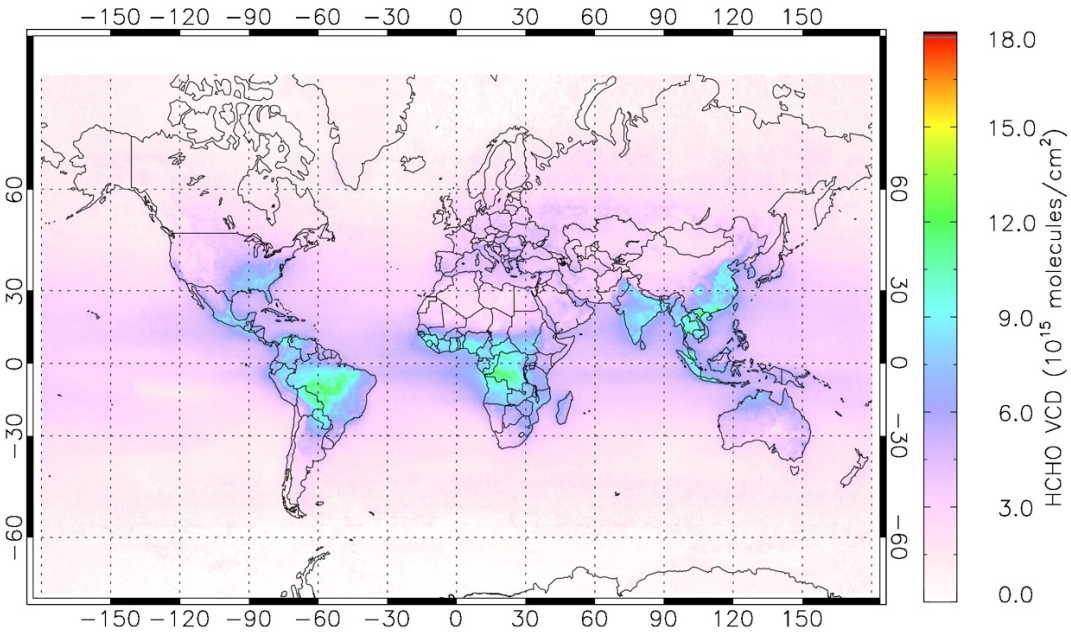

Figure 1. Annual mean OMI HCHO vertical column densities (VCDs) for 2007 retrieved with the PCA-based spectral fitting algorithm. Regions with relatively large sources of volatile organic compounds (VOCs) around the world can be found in the map, including the southeastern U.S., eastern China, Southeast Asia, India, South America, and Africa.

Email: can.li@nasa.gov, Tel: 301-614-5616

1. Overview

This document provides a brief introduction to the level 2 (L2) Aura Ozone Monitoring Instrument (OMI) formaldehyde (HCHO) vertical column density (VCD) research product. The product is produced with the NASA Goddard Space Flight Center (GSFC) principal component analysis (PCA) spectral fitting algorithm for the period of October 2004 to December 2018.

Detailed description of the PCA-based HCHO retrieval algorithm can be found in Li et al. (2015) and this document is intended as a quick reference for interested data users. Also please note that only limited validation effort has been afforded to the research product, which is offered on a best effort basis.

2. Product description

Each level 2 (L2) product file, also referred to as a granule, covers the sunlit portion of an OMI orbit with an approximately 2600 km wide swath. Each swath contains 60 positions or rows across the ground track of OMI (cross-track), with each row containing about 1600 pixels along the ground track (along-track). During normal operations, 14 or 15 granules are produced daily, providing contiguous coverage of the globe.

Since 25 June 2007 signal suppression (anomaly) has been observed in Level 1B (L1B) Earth radiance data for OMI rows 53-54 (0-based). This anomaly, also known as the OMI row anomaly, has since expanded and now affects approximately half of the OMI rows. HCHO retrievals were not attempted for pixels deemed to be affected by the OMI row anomaly in L1B data. For more information about OMI row anomaly, please refer to the KNMI website (http://projects.knmi.nl/omi/research/product/rowanomaly-background.php)

The product files are in HDF-5 format, and are named following the pattern below:
OMI-Aura_L2-OMHCHO_dailyret_VCD_EarthRef_YYYYMMDD-oXXXXX_nPC10_v001-yyyymmddthhmmss.h5 ,
where YYYY, MM, DD stand for the year, month, and day of the OMI radiance measurements, respectively, and XXXXX is the orbit number. The current product version is v001 and the production time is specified at the end of the filename in the format of yyyymmddthhmmss.

The most relevant parameters or data fields for data users in the product files are:

- **SCD_HCHO**: slant column density (SCD) of HCHO produced by spectral fitting using HCHO cross sections. SCD represents the estimated number of HCHO molecules along the light path between the Sun and OMI at the top of the atmosphere (TOA). The unit is molecules/cm^2, and the fill value is -1×10^{30} molecules/cm^2.

- **HCHO**: vertical column density (VCD) of HCHO, representing the estimated total number of HCHO molecules within the vertical atmospheric column between the surface and TOA over a unit area. It has the same unit and fill value as SCD_HCHO.

- **HCHO_Corrected**: vertical column density (VCD) of HCHO, but corrected for biases between the retrieved HCHO VCDs and monthly mean HCHO simulated with the GMI (Global Modeling Initiative) chemical transport model over the remote East Pacific (165-180°W). It has the same unit and fill value as SCD_HCHO.

- **AMF**: air mass factor (AMF) at 340 nm that is used to convert HCHO SCD to VCD. AMF is estimated from radiative transfer calculations and depends on various factors, particularly the *a priori* profiles of HCHO. AMF is unitless and has a fill value of -1×10^{30}.

- **CloudRadianceFraction**: cloud radiance fraction (CRF) at 340 nm is provided for each pixel. It is unitless, has a valid range of 0-1.0, and a fill value of -1×10^{30}.

- **ScatteringWeight**: scattering weight (or vertically resolved box-AMF) at 340 nm at 72 vertical layers (as defined by LayerBottomPressure and a TOA pressure of 0.01 hPa) is provided for each OMI pixel. Scattering weight represents the sensitivity of TOA radiances to perturbations in the amount of HCHO within each vertical layer.

- **LayerBottomPressure**: pressure at the lower edge of each of the 72 vertical layers for which scattering weight is provided. The pressure at the top of the highest vertical layer or the top of the atmosphere is 0.01 hPa.

- **GMILayerWeight**: GMI model based *a priori* profile used in the retrieval of HCHO VCD for each pixel. The profile is normalized (*i.e.*, the sum of GMILayerWeight from all 72 layers is one) and the value represents the fraction of HCHO molecules each layer contributes to the entire atmospheric column.

3. Recommendations for data filtering

We recommend that data users use HCHO_Corrected, and exclude pixels near the edge of the OMI swath (rows 0 and 59, 0-based), with relatively large cloud fraction (CRF > 0.5), or at relatively high solar zenith angles (SZA > 70°).

For best data quality, use only data from pixels near the center of the swath (rows 4-54, 0-based) with SZA < 65°, CRF < 0.3, and AMF > 0.3. Retrievals for OMI pixels from the descending node of the Aura satellite should not be used. Pixels with AlgorithmFlag_SnowIce = 2 should not be used.

Additionally, for certain days the quality of HCHO retrievals is lower (for example, due to the lack of coverage over the Pacific that is used as a reference in retrievals) and these days should be excluded from data analysis. For a list of such days, please refer to a text file named days_to_exclude.txt.

4. Use your own *a priori* profiles

AMF is calculated from the scattering weight, $m(z)$, and the *a priori* profile of HCHO, $n_{HCHO}(z)$:

$$AMF = \int_0^{TOA} m(z)\, n_{HCHO}(z) dz. \quad (1)$$

Note that $n_{HCHO}(z)$ is normalized against the VCD and represents the fraction of HCHO molecules contributed by layer z to the overall HCHO molecules within the entire atmospheric column from $z = 0$ to TOA.

To make use of user-supplied *a priori* profiles, it is recommended that those profiles (in mixing ratio) first be interpolated to the layers as specified by LayerBottomPressure. The interpolated mixing ratio can then be used to calculate the partial vertical column density, as well as the normalized fraction of each layer, $n_{HCHO}(z)$. Users can then make their own estimates of AMF using Eq. (1), and convert the SCD of HCHO to VCD, following Eq. (2):

$$VCD = \frac{SCD}{AMF}. \quad (2)$$

5. Data access and support

The level 2 OMI HCHO research product based on the PCA retrieval technique is available, free of charge, to data users at NASA Aura Validation Data Center (AVDC: https://avdc.gsfc.nasa.gov). Please note that AVDC also supports a number of other data products from Aura and other satellites, and bandwidth available for data downloading may be limited. Users who need to acquire large volumes of data are encouraged to first contact AVDC at Michael.M.Yan@nasa.gov or david.e.larko@nasa.gov.

Limited support for the research product is provided on a best effort basis. Please address questions or comments to the principal investigator of the product, Dr. Can Li (can.li@nasa.gov).

6. Acknowledgements

The level 2 OMI HCHO research product is produced with support from NASA Satellite Calibration Interconsistency Studies Program (Grant number: NNX16AT07G), managed by Dr. Lucia Tsaoussi. We also thank Mr. Philip Durbin at NASA Ozone SIPS for support in producing and archiving the dataset.

7. References

Li, C., J. Joiner, N. A. Krotkov, and L. Dunlap (2015), A new method for global retrievals of HCHO total columns from the Suomi National Polar-orbiting Partnership Ozone Mapping and Profiler Suite, *Geophys. Res. Lett.*, 42, 2515-2522. doi:10.1002/2015GL063204.