Global Formaldehyde Products from the Ozone Mapping and Profiler Suite (OMPS) Nadir Mappers on Suomi NPP and NOAA-20

Caroline R. Nowlan[1], Gonzalo González Abad[1], Hyeong-Ahn Kwon[2], Zolal Ayazpour[3], Christopher Chan Miller[4], Kelly Chance[1], Heesung Chong[1], Xiong Liu[1], Ewan O'Sullivan[1], Huiqun Wang[1], Lei Zhu[5], Isabelle De Smedt[6], Glen Jaross[7], Colin Seftor[8], and Kang Sun[3]

[1]Center for Astrophysics — Harvard & Smithsonian
[2]University of Suwon
[3]University at Buffalo
[4]Harvard University
[5]Southern University of Science and Technology
[6]Belgian Institute for Space Aeronomy (BIRA-IASB)
[7]NASA Goddard Space Flight Center
[8]Science Systems and Applications, Inc.

February 22, 2024

Abstract

We describe new publicly-available, multi-year formaldehyde (HCHO) data records from the Ozone Mapping and Profiler Suite (OMPS) nadir mapper (NM) instruments on the Suomi NPP and NOAA-20 satellites. The OMPS-NM instruments measure backscattered UV light over the globe once per day, with spatial resolutions close to nadir of 50 × 50 km^2 (OMPS/Suomi-NPP) and 17 × 17 km^2 or 12 × 17 km^2 (OMPS/NOAA-20). After a preliminary instrument line shape and wavelength calibration using on-orbit observations, we use the backscatter measurements in a direct spectral fit of radiances, in combination with a nadir reference spectrum collected over a clean area, to determine slant columns of HCHO. The slant columns are converted to vertical columns using air mass factors derived through scene-by-scene radiative transfer calculations. Finally, a correction is applied to account for background HCHO in the reference spectrum, as well as any remaining high-latitude biases. We investigate the consistency of the OMPS products from Suomi NPP and NOAA-20 using long-term monthly means over 12 geographic regions, and also compare the products with publicly-available TROPOMI HCHO observations. OMPS/Suomi-NPP and OMPS/NOAA-20 monthly mean HCHO vertical columns are highly consistent (r = 0.98), with low proportional (2 %) and offset (2×10^{1}4 molecules cm-^{2}) biases. OMPS HCHO monthly means are also well-correlated with those from TROPOMI (r = 0.92), although they are consistently 10±16 % larger in polluted regions (columns >8×10^{1}5 molecules cm-^{2}). These differences result primarily from differences in air mass factors.

Global Formaldehyde Products from the Ozone Mapping and Profiler Suite (OMPS) Nadir Mappers on Suomi NPP and NOAA-20

Caroline R. Nowlan[1], Gonzalo González Abad[1], Hyeong-Ahn Kwon[1*], Zolal Ayazpour[2], Christopher Chan Miller[1,3], Kelly Chance[1], Heesung Chong[1], Xiong Liu[1], Ewan O'Sullivan[1], Huiqun Wang[1], Lei Zhu[4], Isabelle De Smedt[5], Glen Jaross[6], Colin Seftor[6,7], Kang Sun[2]

[1]Center for Astrophysics | Harvard & Smithsonian, Cambridge, MA, USA
[2]University at Buffalo, Buffalo, NY, USA
[3]Harvard University, Cambridge, MA, USA
[4]Southern University of Science and Technology, Shenzhen, China
[5]Royal Belgian Institute for Space Aeronomy, Brussels, Belgium
[6]NASA Goddard Space Flight Center, Greenbelt, MD, USA
[7]Science Systems and Applications, Inc., Lanham, MD, USA

Key Points:

- We describe new publicly-available formaldehyde products from the OMPS satellite instruments.
- OMPS HCHO data records retrieved from Suomi NPP and NOAA-20 are consistent.
- OMPS HCHO vertical column time series are compared with TROPOMI in 12 geographic regions.

*Now at the University of Suwon, Gyeonggi-do, Republic of Korea.

Corresponding author: Caroline Nowlan, cnowlan@cfa.harvard.edu

Abstract

We describe new publicly-available, multi-year formaldehyde (HCHO) data records from the Ozone Mapping and Profiler Suite (OMPS) nadir mapper (NM) instruments on the Suomi NPP and NOAA-20 satellites. The OMPS-NM instruments measure backscattered UV light over the globe once per day, with spatial resolutions close to nadir of 50 $\times$ 50 km^2 (OMPS/Suomi-NPP) and 17 $\times$ 17 km^2 or 12 $\times$ 17 km^2 (OMPS/NOAA-20). After a preliminary instrument line shape and wavelength calibration using on-orbit observations, we use the backscatter measurements in a direct spectral fit of radiances, in combination with a nadir reference spectrum collected over a clean area, to determine slant columns of HCHO. The slant columns are converted to vertical columns using air mass factors derived through scene-by-scene radiative transfer calculations. Finally, a correction is applied to account for background HCHO in the reference spectrum, as well as any remaining high-latitude biases. We investigate the consistency of the OMPS products from Suomi NPP and NOAA-20 using long-term monthly means over 12 geographic regions, and also compare the products with publicly-available TROPOMI HCHO observations. OMPS/Suomi-NPP and OMPS/NOAA-20 monthly mean HCHO vertical columns are highly consistent ($r = 0.98$), with low proportional (2 %) and offset (2×10^{14} molecules cm^{-2}) biases. OMPS HCHO monthly means are also well-correlated with those from TROPOMI ($r = 0.92$), although they are consistently 10$\pm$16 % larger in polluted regions (columns $> 8 \times 10^{15}$ molecules cm^{-2}). These differences result primarily from differences in air mass factors.

1 Introduction

Formaldehyde (HCHO) is important in atmospheric chemistry and outdoor air quality through its role in atmospheric oxidation and the production of ozone and secondary organic aerosols. The oxidation of non-methane volatile organic compounds (NMVOCs) from biomass burning, anthropogenic sources and biogenic emissions results in local and regional HCHO enhancements, while methane oxidation is largely responsible for HCHO in the global background atmosphere. A smaller amount of direct HCHO emission also occurs through industrial activity and biomass burning. Spaceborne remote sensing instruments can be used to map the global distribution of HCHO using characteristic absorption features in the ultraviolet region of the electromagnetic spectrum.

The first HCHO observations from space were made by the Global Ozone Monitoring Experiment (GOME) (1995–2011) (Thomas et al., 1998; Chance et al., 2000). Multi-year HCHO products have since been produced from GOME (De Smedt et al., 2008), the SCanning Imaging Absorption SpectroMeter for Atmospheric CHartographY (SCIAMACHY) (2002–2012) (De Smedt et al., 2012), GOME-2 (2006–2021/2012–present/2018–present) (De Smedt et al., 2012), the Ozone Monitoring Instrument (OMI) (2004–present) (De Smedt et al., 2015; González Abad et al., 2015), the Ozone Mapping and Profiler Suite (OMPS) on Suomi NPP (2011–present) (C. Li et al., 2015; González Abad et al., 2016; Su et al., 2019), the TROPOspheric Monitoring Instrument (Sentinel-5P/TROPOMI) (2017–present) (De Smedt et al., 2018, 2021) and the Environmental Trace Gases Monitoring Instrument (EMI) (2018–present) (Su et al., 2022). The next-generation geostationary air quality instruments Geostationary Environment Monitoring Spectrometer (GEMS; East Asia; launch 2020) (Kwon et al., 2019; Kim et al., 2020), Tropospheric Emissions: Monitoring of POllution (TEMPO; North America; 2023) (Zoogman et al., 2017) and Sentinel-4 (Europe and North Africa; 2024) (Ingmann et al., 2012), and the future Sentinel-5 low Earth orbit missions all include HCHO as a baseline data product.

Formaldehyde is removed relatively quickly from the atmosphere through photolysis and oxidation by the hydroxyl radical OH, with a resulting lifetime of a few hours. This high reactivity allows satellite measurements of HCHO to be used in combination with atmospheric chemistry models to provide top-down constraints on NMVOC emissions (e.g., Palmer et al., 2006; Barkley et al., 2008; Millet et al., 2008; Stavrakou et al., 2009; Marais et al., 2012; Bauwens et al., 2016; Kaiser et al., 2018). Satellite measurements of HCHO can also be used to examine secondary organic aerosol formation (Veefkind et al., 2011; Marais et al., 2016), tropospheric ozone production (Jin et al., 2017), the oxidative capacity of the atmosphere (Valin et al., 2016; Wolfe et al., 2019) and the health impacts of ambient HCHO (Zhu et al., 2017; Su et al., 2019).

OMPS is a suite of three instruments that are included on the Joint Polar Satellite System (JPSS) (Goldberg et al., 2013). The primary goal of OMPS is to provide ozone observations for use in near real time applications and for continuity of the long-term data record of global ozone (Flynn et al., 2014; Sofieva et al., 2017). The full OMPS suite consists 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 was launched on 28 October 2011 on the joint NASA/NOAA Suomi NPP (National Polar-

orbiting Partnership) satellite and on 18 November 2017 on the NOAA-20 (JPSS-1) satellite. OMPS on Suomi NPP consists of the full suite, while only the nadir package (OMPS-NM and OMPS-NP) flies on NOAA-20. Hereafter, we refer to the two OMPS-NM instruments currently in orbit as OMPS/SNPP and OMPS/NOAA-20. The Suomi NPP and NOAA-20 satellites are both in afternoon sun-synchronous orbits with daylight equatorial crossing times of approximately 13:30 local time. NOAA-20's orbit is 50 minutes behind that of Suomi NPP. Future JPSS satellites (JPSS-2, 2022; JPSS-3, 2028; JPSS-4, 2032) will each carry an OMPS.

In this paper, we describe new multi-year, publicly-available HCHO products developed by the Smithsonian Astrophysical Observatory (SAO) for the OMPS/SNPP and OMPS/NOAA-20 nadir mapper instruments (González Abad, 2022a, 2022b). Previous studies have demonstrated HCHO retrievals from OMPS (C. Li et al., 2015; González Abad et al., 2016; Su et al., 2019), but these efforts have been limited to specific timeframes and to the OMPS/SNPP instrument only. A limited data set for OMPS produced for the Korea-United States Air Quality (KORUS-AQ) campaign timeframe (May–June 2016), based on the retrieval described in González Abad et al. (2016), has also been used to derive emissions in Asia during KORUS-AQ (Souri et al., 2020; Choi et al., 2022).

The new OMPS products extend and augment long-term global data records of HCHO. After 2012, the OMPS/SNPP HCHO product provides global coverage that is missing in the widely-used OMI HCHO product (González Abad et al., 2015) due to an instrument row anomaly (this is particularly important before the launch of Sentinel-5P/TROPOMI in late 2017). Furthermore, with future Sentinel-5 instruments planned for morning orbits, OMPS is currently the only planned UV hyperspectral instrument in afternoon orbit post-TROPOMI, and hence, after TROPOMI decommissioning, the only instrument capable of continuing the afternoon HCHO data record that began with OMI in 2004.

The paper is organized as follows. Section 2 describes the OMPS instruments in more detail and the data products used in the analysis. Section 3 describes the HCHO retrieval algorithm and associated uncertainties. Section 4 presents comparisons of HCHO columns between the two OMPS instruments and with TROPOMI, which is the state-of-the-art low Earth orbit instrument. Section 5 summarizes the retrieval framework and presents directions for future improvements. We validate the OMPS HCHO retrievals with ground-based measurements in a separate companion paper (Kwon et al., 2022).

2 Data products

2.1 OMPS

2.1.1 The OMPS nadir mappers

The HCHO measurements in this study are derived from the OMPS nadir mapper (OMPS-NM). Detailed descriptions of OMPS and the on-orbit performance of OMPS/SNPP can be found in several previous papers (Flynn et al., 2014; Jaross et al., 2014; Seftor et al., 2014; Pan et al., 2017).

Table 1 summarizes several features of the OMPS nadir mappers relevant to HCHO retrievals. The nadir mapper uses a grating spectrometer to observe backscattered radiation with a two-dimensional CCD-array detector with 340 detector pixels in the spectral dimension and 740 pixels in the across-track dimension. In the CCD's spectral dimension, 196 of the 340 pixels are illuminated in OMPS/SNPP at wavelengths 300–380 nm. The wavelength range was extended for OMPS/NOAA-20, which has 294 pixels illuminated from 298–420 nm. The spectra in both instruments are sampled every 0.42 nm at a spectral resolution of about 1 nm at full-width at half maximum.

The full across-track OMPS field-of-view is 110°, resulting in a swath of about 2800 km at the Earth's surface. Of the 740 pixels in the across-track dimension, 708 are illuminated. Pixel measurements are binned together spatially and temporally by the instrument before being sent to the ground to achieve a lower data rate. This results in 36 across-track and 400 along-track nominal ground pixels per orbit for OMPS/SNPP, each with a spatial resolution at the ground of about 50 $\times$ 50 km^2 close to nadir, with larger ground pixels at the edge of the swath. The two center across-track positions are rebinned differently and have spatial resolutions of 30 $\times$ 50 km^2 and 20 $\times$ 50 km^2. The bin sizes were reduced on OMPS/NOAA-20 to achieve higher along-track and across-track resolution. Early in the mission, most OMPS/NOAA-20 observations were rebinned to 104 pixels across-track and 1201 along-track for a spatial resolution of 17 $\times$ 17 km^2 at nadir. On 13 February 2019, the rebinning scheme was modified to produce 140 rebinned across-track ground pixels, with a corresponding spatial resolution of 12 $\times$ 17 km^2 at nadir. The OMPS/NOAA-20 rebinning scheme provides enhanced spatial resolution over that of OMPS/SNPP, but also results in lower signal-to-noise in the rebinned spectra.

Table 1. Characteristics of the OMPS Nadir Mapper Instruments on Suomi NPP and NOAA-20.

Characteristic	OMPS SNPP	OMPS NOAA-20
Launch	28 October 2011	18 November 2017
Spectral coverage	300–380 nm	298–420 nm
Spectral resolution	1 nm	1 nm
Spectral sampling	0.42 nm	0.42 nm
Spatial resolution at nadir	50 × 50 km^2	17 × 17 km^2 (launch–13/02/2019)
(across-track × along-track)		12 × 17 km^2 (13/02/2019–present)
Nominal across-track ground pixels	36	104 (launch–13/02/2019)
		140 (13/02/2019–present)
Nominal along-track ground pixels	400	1201

2.1.2 OMPS radiance products

Table 2 summarizes the OMPS data products used in the HCHO retrievals. We produce OMPS/SNPP HCHO retrievals using the NASA OMPS Nadir Mapper Earth View (NMEV) Version 2.0 Level 1B (calibrated and geolocated) radiances, available through NASA's Goddard Earth Sciences Data and Information Services Center (GES DISC) (Jaross, 2017a). In addition to the radiances for one orbit, each Level 1B file contains static solar irradiance spectra, derived from direct measurements of the Sun through a diffuser near the beginning of the mission. We produce OMPS/NOAA-20 products using Version 1.3 OMPS/NOAA-20 radiances, which are produced by the NASA OMPS team using a similar algorithm with instrument-specific calibration. The OMPS/NOAA-20 products are currently available through the OMPS website (https://ozoneaq.gsfc.nasa.gov/omps/). In this paper, we present measurements from February 2012 for OMPS/SNPP and February 2018 for OMPS/NOAA-20, when regular daily or near-daily global measurements became available, through December 2020.

2.1.3 OMPS cloud products

Tropospheric trace gas retrievals require information on the cloud amount (cloud fraction) and height (optical centroid pressure) over each ground pixel. A publicly-available

Table 2. OMPS data products used in HCHO retrieval.

	OMPS/SNPP		OMPS/NOAA-20	
Input Source	Product Name	Version	Product Name	Version
Level 1B radiances	OMPS_NPP_NMEV_L1B	v2.0	OMPS_N20_NMEV_L1B	v1.3
Total ozone	OMPS_NPP_NMTO3_L2	v2.1	OMPS_N20_NMTO3_L2	v1.3

cloud product (NMCLDRR) that provides cloud fractions and pressures determined from rotational Raman scattering currently exists for OMPS/SNPP (Joiner, 2020; Vasilkov et al., 2014), but not for OMPS/NOAA-20. As a result, for consistency between the two instruments, we take cloud information from the total ozone products (NMTO3) which are available for both OMPS/SNPP and OMPS/NOAA-20. For OMPS/SNPP, we use the Version 2.1 total ozone product OMPS_NPP_NMTO3_L2 available from the NASA GES DISC (Jaross, 2017b), and for OMPS/NOAA-20, the Version 1.3 product OMPS_N20_NMTO3_L2 available from the OMPS website. The OMPS total ozone product is processed using the Total Ozone Mapping Spectrometer (TOMS) Version 8.5 algorithm, which is also used to produce OMI (Bhartia & Wellemeyer, 2002) and TOMS total ozone.

The OMPS cloud fractions in the total ozone product are determined from Level 1B radiances at 331 nm using a Mixed Lambertian Reflectivity (MLER) model (Ahmad et al., 2004), where the surface is assumed to have a constant global reflectivity of 0.15. In contrast, the rotational Raman cloud fraction retrieval at 354.1 nm uses location-dependent and much smaller reflectivities of approximately 0.02–0.08 (over snow-free and ice-free surfaces), as do most recent trace gas retrievals in the ultraviolet. As a result, we find the OMPS/SNPP total ozone cloud fraction systematically underestimates that of the Raman cloud product by 0.05 to 0.12, depending on the season.

In order to determine a more accurate cloud fraction for OMPS that can be applied to both OMPS/SNPP and OMPS/NOAA-20 in a consistent manner, we calculate an updated cloud fraction for the HCHO product. We find the OMPS/SNPP Raman cloud fraction has a nearly linear dependence on the observed reflectivity at 360 nm included in the total ozone product, although the fit coefficients vary across a year. We determine the relationship for each month of the year, and use this in combination with the NMTO3 reflectivity to produce a corrected cloud fraction for the HCHO product that

agrees more closely with the Raman cloud fraction. This effectively assumes a constant global reflectivity, similarly to the NMTO3 cloud retrieval, but produces a cloud distribution much closer to that of the Raman product. The estimated cloud fractions are not valid over ice and snow.

The OMPS HCHO retrievals use the total ozone cloud pressures directly from the total ozone product. These pressures are from a climatology of the cloud optical centroid pressure determined using rotational Raman Scattering with the Ozone Monitoring Instrument's OMCLDRR cloud retrieval (Joiner, 2006). The use of climatological pressures rather than observed pressures is a source of error in HCHO measurements, and associated uncertainties are discussed in Section 3.5.2.

2.2 TROPOMI

The Sentinel-5P/TROPOMI instrument is the state-of-the-art UV/visible hyperspectral remote sensing instrument in low Earth orbit, and is in an orbit 5 minutes behind Suomi NPP. TROPOMI was launched on 13 October 2017 and uses backscattered radiation in the UV, visible, and shortwave infrared to measure a suite of aerosols and trace gases, including HCHO. We compare OMPS HCHO observations with TROPOMI HCHO in Section 4.

TROPOMI HCHO data products are available from 14 May 2018. The retrievals have a spatial resolution of 3.5×7 km^2 prior to 6 August 2019 and 3.5×5.5 km^2 afterwards. We use the offline HCHO products processed with the S5P Version 1 processor (ESA & DLR, 2019a, 2019b) up to 13 July 2020 and the Version 2 processor (ESA & DLR, 2020) after that date.

The TROPOMI HCHO retrieval uses the wavelength region 328.5 – 359 nm to fit the HCHO slant column density using differential optical absorption spectroscopy (DOAS). The vertical column is determined using a pre-computed look-up table of vertically-resolved air mass factors (AMF). The surface reflectance used in the AMF calculation is from $0.5^\circ \times 0.5^\circ$ OMI Lambertian equivalent reflectance (LER) monthly surface reflectance climatologies (Kleipool et al., 2008). A priori HCHO profiles are from daily TM5-MP model profiles at $1^\circ \times 1^\circ$ resolution (Williams et al., 2017). Cloud parameters are from a separate TROPOMI cloud retrieval (Loyola et al., 2018). The cloud correction is applied using the independent pixel approximation (Martin et al., 2002; Boersma et al., 2004)

for cloud fractions greater than 0.1. The vertical columns are reference-sector corrected using a background vertical column from the TM5 model and with a bias correction determined from the previous four days of data. Further details of the HCHO retrieval can be found in De Smedt et al. (2018) and De Smedt et al. (2021).

We filter the TROPOMI HCHO by only using observations with qa_value ≥ 0.5. In the Version 1 processor, this removes retrievals with an error flag, SZA $> 70°$, AMF < 0.1, or cloud radiative fraction at 340 nm > 0.6. We additionally filter out data flagged as snow/ice or with albedo > 0.2 (flagged by default in the Version 2 qa_value), and exclude data where effective cloud fractions > 0.4 for consistency with our OMPS analysis.

Vigouroux et al. (2020) validated TROPOMI HCHO using a network of 25 Fourier-Transform Infrared Spectrometers (FTIR) (Vigouroux et al., 2018), and found TROPOMI overestimated HCHO in clean background regions (HCHO columns $< 2.5\times10^{15}$ molecules cm^{-2}) by 26±5 % relative to the FTIRs and underestimated HCHO by 30.8±1.4 % at more polluted sites ($> 8\times10^{15}$ molecules cm^{-2}). De Smedt et al. (2021) performed a validation with 18 multi-axis DOAS (MAX-DOAS) instruments, and similarly found TROPOMI HCHO to be 25 % lower than MAX-DOAS at very polluted sites, but in good agreement in moderately polluted sites.

3 Formaldehyde retrieval algorithm

3.1 Overview

The OMPS retrieval algorithm uses a three-step approach to produce a Level 2 HCHO vertical column product. Figure 1 shows a flow chart that summarizes the algorithm components and major inputs that will be described in this section.

First, after a preliminary spectral calibration, a spectral fitting algorithm determines the HCHO slant column density for each ground pixel by fitting a modeled spectrum to the observed spectrum. This spectral fitting makes use of a spectrum determined from measurements over the Pacific ocean as a clean reference (i.e., where only minimal background HCHO is present). In the second step, a separate algorithm determines the AMF that describes the light path through the atmosphere. Third, the retrieval calculates the geometry-independent vertical column density using the retrieved slant column density, the AMF, and a reference sector correction that corrects for background HCHO

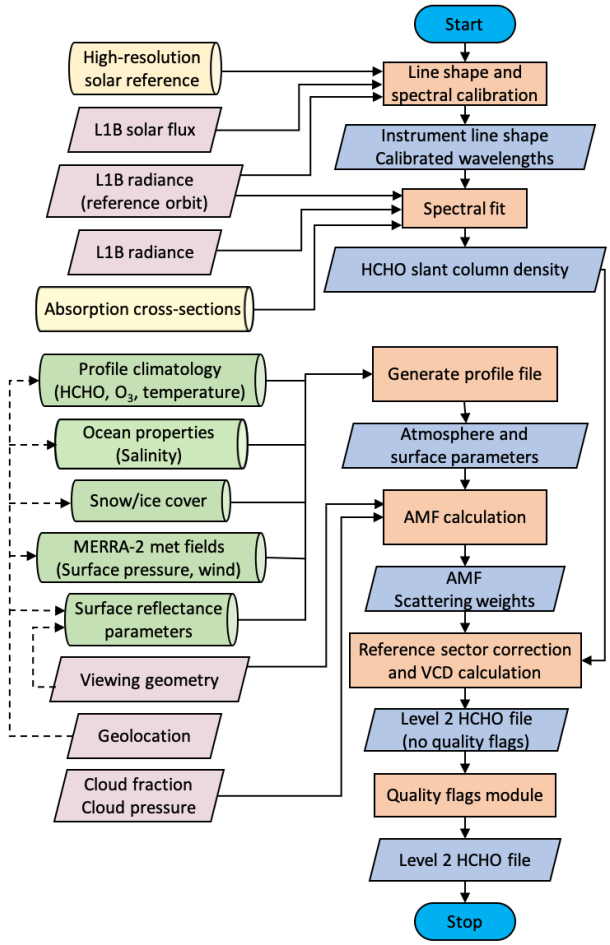

Figure 1. Flow chart of the OMPS HCHO algorithm, showing input databases for the slant column fit (yellow cylinders) and AMF calculation (green cylinders), inputs from the OMPS radiance files and cloud information (pink parallelograms), algorithm outputs/inputs (blue parallelograms) and main processes (orange rectangles).

in the clean nadir reference and any remaining background biases in the retrieval. In addition to these three major steps, we add quality flags to the final Level 2 file in a post-processing step.

In general terms, the vertical column density (VCD) of a trace gas is related to the slant column density (SCD) seen by the remote sensing instrument through an AMF that describes the mean photon path through the gas by

$$VCD = \frac{SCD}{AMF}. \tag{1}$$

The HCHO SCD is determined using a reference spectrum collected over a relatively clean region. As a result, the SCD retrieved through spectral fitting is in fact a differential SCD defined by

$$\Delta SCD = SCD - SCD_R, \tag{2}$$

where SCD is the slant column amount in the nadir observation of interest and SCD_R is the background slant column in the reference spectrum. In the case of the OMPS retrievals, SCD_R is determined using

$$SCD_R = VCD_R \cdot AMF_R, \tag{3}$$

where VCD_R is the reference vertical column density estimated from a chemical transport model and AMF_R is the AMF at the reference spectrum location.

In addition, OMPS HCHO, like many satellite retrievals of HCHO and other weak absorbers, shows latitude-dependent biases in the slant column, which are likely due to interfering absorbers and insufficiently-corrected instrument calibration issues. These biases are corrected using modeled columns of HCHO and accounted for using a slant column bias correction term SCD_B. Following Equations 1 and 2, and considering the bias correction, we determine the final vertical column density using

$$VCD = \frac{\Delta SCD + SCD_R + SCD_B}{AMF}. \tag{4}$$

Figure 2 shows the variables in Equation 4 determined for orbits on 28 July 2019, which passed over eastern Asia and regions of anthropogenic and wildfire emissions. The remainder of Section 3 describes the detailed derivation of differential slant column densities, AMFs, and reference and bias corrections used in the calculation of the final vertical column densities.

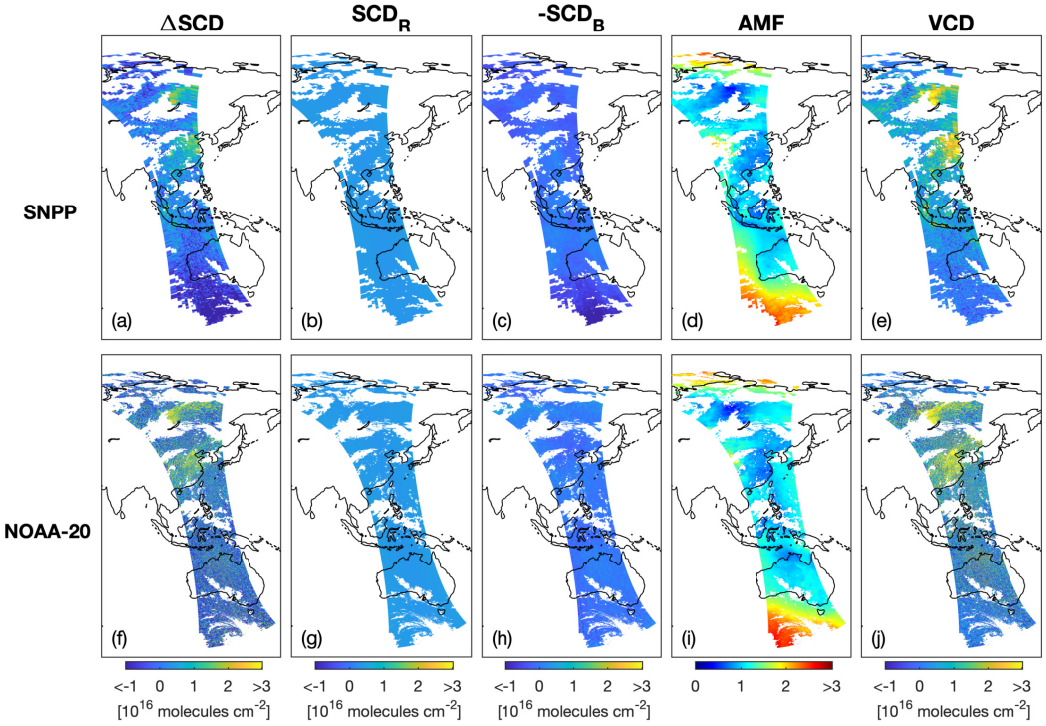

Figure 2. Formaldehyde on 28 July 2019 for SZA $< 80^{\circ}$ and cloud fractions < 0.4 retrieved from OMPS/SNPP orbit 40149 and OMPS/NOAA-20 orbit 8752, showing (a, f) differential slant column densities (ΔSCD), (b, g) reference background corrections (SCD_R), (c, h) bias corrections ($-SCD_B$), (d, i) air mass factors (AMF), and (e, j) final vertical column densities (VCD). The negative of the bias correction is shown for a more direct comparison with the ΔSCD. These orbits pass over likely anthropogenic and biogenic HCHO sources in eastern China, and a large wildfire plume in Siberia. The orbits are offset in longitude due to the 50-minute orbital separation between Suomi NPP and NOAA-20.

3.2 Spectral fitting

3.2.1 On-orbit spectral calibration

OMPS/SNPP solar measurements are provided in the L1B data files and are based on four solar measurements taken in March/April 2012, while OMPS/NOAA-20 solar measurements are similarly derived from four measurements in March/April 2018. As a first step in the retrieval, we derive the OMPS slit function s and wavelength registration as a function of across-track position using a well-established calibration approach developed for SAO trace gas retrievals (e.g., Bak et al., 2017; Sun et al., 2017).

We model the slit with three super-Gaussian terms (Beirle et al., 2017): 1) the slit's half width at $1/e$, w; 2) the shape parameter k; and 3) the asymmetry parameter a_w using

$$s(\Delta\lambda) = exp\left[-\left|\frac{\Delta\lambda}{w + \mathrm{sgn}(\Delta\lambda)a_w}\right|^k\right], \tag{5}$$

where $\Delta\lambda$ describes the wavelength distance from the center and sgn() is the sign function used to define the two sides of the slit function.

For each across-track position, the spectral fitting algorithm simultaneously determines the slit parameters and a shift in the detector-pixel-to-wavelength registration in the HCHO absorption region through a fit of a modeled solar spectrum based on a high resolution reference solar spectrum (Chance & Kurucz, 2010) to the observed solar spectrum. The slit parameters are then saved and the wavelength registration calibration is performed again using the clean radiance reference spectrum at each across-track position (described in Section 3.2.2) to fine-tune the daily wavelength calibration. The retrieved shape parameter k varies from 2.2 to 2.5 for OMPS/SNPP, depending on across-track position, and from 2.2 to 2.7 for OMPS/NOAA-20 ($k = 2$ for a standard Gaussian). The slit functions are mostly symmetric in the center of the swaths, with a_w ranging from -0.04 to 0.02 off-center, depending on across-track position.

Figure 3 shows the full width at half maxima (FWHM) retrieved from the on-orbit solar spectra as part of the HCHO retrieval, as well as those derived from on-orbit estimates by the NASA OMPS L1B calibration team. Pre-flight measurements (not shown) and on-orbit estimates of the instrument line shape indicate the FWHM in the HCHO fitting window varies only slightly as a function of wavelength (by 1–5%, with the largest deviations at the far off-nadir pixels). In our retrievals, we derive a single line shape for

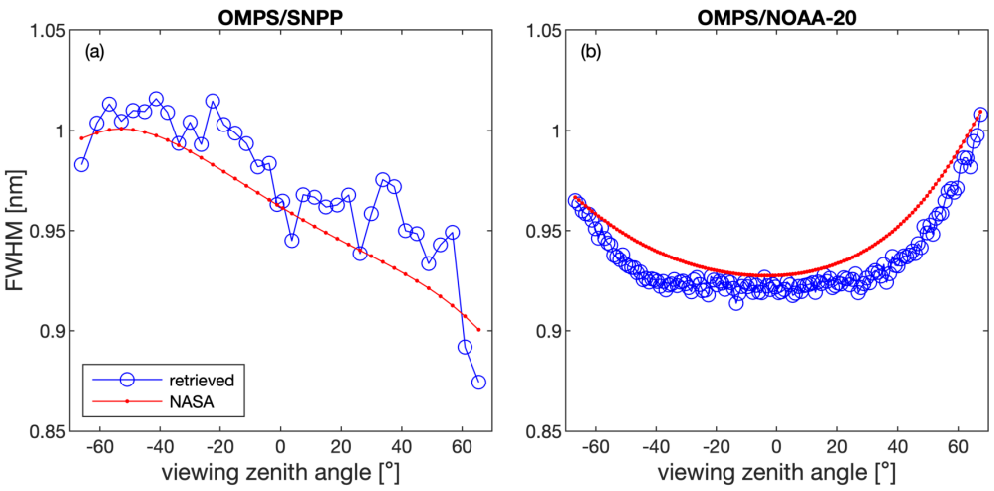

Figure 3. (a) Full width at half maximum (FWHM) of on-orbit slit functions derived for OMPS/SNPP (36 across-track positions) and (b) OMPS/NOAA-20 (140 across-track positions) using OMPS solar measurements in the HCHO fitting wavelength region, as well as slit functions derived by NASA's OMPS calibration team.

the entire HCHO wavelength window for each across-track position. Retrieved slit widths are similar to those provided in NASA calibration files, with OMPS/SNPP FWHM changing across the orbital swath from 0.9 – 1 nm, and OMPS/NOAA-20 showing a stable FWHM of 0.92 nm through the central part of the orbit and increasing slightly further off-nadir.

3.2.2 Reference spectrum

To minimize across-track striping in the HCHO retrievals, we use reference spectra derived from relatively clean observations at each across-track position in place of a direct solar irradiance measurement (e.g., González Abad et al., 2015, 2016). We determine the radiance reference spectrum at each across-track position by averaging all spectra collected at that position between latitudes 30°S and 30°N from the orbit closest in time and with an equatorial crossing closest to 160°W and within 140°W and 180°W (i.e., over the clean Pacific).

3.2.3 Spectral fitting details

We derive the differential HCHO slant column density ΔSCD for each nadir observation using a direct fit of the radiance. The direct spectral fitting approach applied in the SAO trace gas retrievals is described elsewhere in detail (e.g., Chan Miller et al., 2014; Nowlan et al., 2018). Briefly, we fit a modeled radiance to the observed radiance using non-linear least squares Levenberg-Marquart minimization by adjusting a state vector $\mathbf{x}$. The radiance is modeled at each wavelength, with pre-defined model parameters $\mathbf{b}$ as

$$F(\lambda) = [x_a I_0(\lambda) + b_u(\lambda)x_u + b_r(\lambda)x_r]e^{-\sum_i b_i(\lambda)x_i} \sum_j (\lambda - \bar{\lambda})^j x_j^{SC} + \sum_k (\lambda - \bar{\lambda})^k x_k^{BL}. \quad (6)$$

In this equation, I_0 is the reference spectrum described in Section 3.2.2, scaled by a retrieved intensity parameter x_a (which mostly describes reflectivity from the surface or clouds). The term x_u represents scaling for a wavelength-dependent correction $b_u(\lambda)$ that describes spectral undersampling (Chance et al., 2005). The term x_r represents the strength of rotational Raman (Ring) scattering described in a pre-computed Ring spectrum $b_r(\lambda)$ (Chance & Spurr, 1997).

The wavelength-dependent trace gas absorption cross sections are represented by $b_i(\lambda)$ and their differential slant column densities (ΔSCD) by x_i. Table 3 lists the trace gases modeled in the spectral fit, which include HCHO, NO_2, O_3, BrO and O_2-O_2, and their reference cross sections. The low frequency effects of aerosol and molecular scattering, wavelength-dependent surface reflectance, and instrument artifacts are considered by scaling (x^{SC}) and baseline (x^{BL}) polynomial terms of orders j and k. In addition, we simultaneously retrieve a wavelength shift that represents the difference in the nadir radiance fitting window wavelengths to those of the reference spectrum. This shift in wavelength calibration is typically due to thermal changes in the instrument over the course of an orbit and inhomogeneous scene illumination (Voors et al., 2006; Noël et al., 2012).

The wavelength region used in the fitting is 328.5–356.5 nm. This region includes major HCHO spectral features but attempts to minimize effects from strong ozone absorption at shorter wavelengths and correlations with BrO and O_2-O_2. This is the wavelength window previously used in OMI HCHO retrievals (González Abad et al., 2015).

Table 3. Parameters fit in OMPS HCHO retrieval.

Parameter	Details
HCHO	Chance and Orphal (2011), 300 K
NO_2	Vandaele et al. (1998), 220 K
O_3	Serdyuchenko et al. (2014), 223 and 243 K
BrO	Wilmouth et al. (1999), 228 K
O_2-O_2	Finkenzeller and Volkamer (2022), 293 K
Undersampling	Chance et al. (2005)
Ring spectrum	Chance and Spurr (1997)
Scaling polynomial	3rd order
Baseline polynomial	3rd order
Wavelength shift	

Figure 4 shows the differential slant column densities derived for OMPS/SNPP and OMPS/NOAA-20 measurements over eastern Asia on 28 July 2019. Figure 5 shows modeled and observed optical depth spectra for sample individual spectra collected over background HCHO, moderate (anthropogenic) and highly polluted (wildfire) from the same orbit. Both these figures show the larger noise in OMPS/NOAA-20 observations relative to those from OMPS/SNPP, resulting from the higher spatial resolution of OMPS/NOAA-20 which was achieved at a cost of decreased signal-to-noise.

Typical fitting uncertainties are on the order of $3.5 \pm 0.8 \times 10^{15}$ molecules cm^{-2} for OMPS/SNPP at SZA$< 45°$ with cloud fractions < 0.4. With its finer spatial resolution of 12×17 km^2, OMPS/NOAA-20 fitting uncertainties are on the order of $1.1 \pm 0.2 \times 10^{16}$ molecules cm^{-2}. The corresponding relative root mean square (RMS) of the fit is on average 2.9×10^{-4} for OMPS/SNPP and 8.9×10^{-4} for OMPS/NOAA-20. As shown in Figure 4, when the retrieved slant columns from OMPS/NOAA-20 Level 1B radiances are averaged at the OMPS/SNPP spatial resolution, the resulting uncertainties are similar in magnitude to those of OMPS/SNPP.

In Figure 4, fitting uncertainties increase at southern high latitudes due to lower signal-to-noise from large solar zenith angles. In addition, larger systematic fitting residuals in OMPS/SNPP fitting at low total radiance (SZA $>\sim 65°$) result in larger cal-

culated fitting uncertainties. These fitting residuals are possibly due to calibration issues in OMPS/SNPP that are not as significant for OMPS/NOAA-20 in the HCHO wavelength fitting window.

3.3 Air mass factor calculation

3.3.1 Overview

The AMF describes the mean photon path through the trace gas of interest. For the OMPS HCHO product, we calculate the AMF for each ground pixel using the formulation of Palmer et al. (2001) and Martin et al. (2002) for an assumed optically thin atmosphere. This formulation describes the AMF as a function of altitude-dependent scattering weights $W(z)$ and profile shape factors $S(z)$, and is defined as

$$AMF = \int_z W(z)S(z)\mathrm{d}z. \tag{7}$$

The scattering weights are determined using a radiative transfer model. The shape factor is the normalized profile shape, and is determined from the partial columns of the trace gas at each layer, $n(z)$, using

$$S(z) = \frac{n(z)}{\sum_z n(z)\mathrm{d}z}. \tag{8}$$

The AMF for a partly cloudy scene is determined by

$$AMF = (1-w)\cdot AMF_{clear} + w \cdot AMF_{cloudy} \tag{9}$$

where AMF_{clear} is the AMF calculated for a completely clear scene and AMF_{cloudy} is the AMF calculated for a completely cloudy scene. The cloud radiance fraction w is defined as

$$w = \frac{fI_{cloud}}{(1-f)I_{clear} + fI_{cloud}} \tag{10}$$

where I_{clear} and I_{cloud} are the radiance intensities for a completely clear and a completely cloudy scene, respectively. These are taken from the radiative transfer calculation. The scene's cloud fraction ($0 \leq f \leq 1$) is the OMPS cloud fraction discussed in Section 2.1.3.

Table 4 summarizes the major inputs to the air mass factor calculations, including radiative transfer calculation inputs and atmospheric profiles from a global chemical transport model.

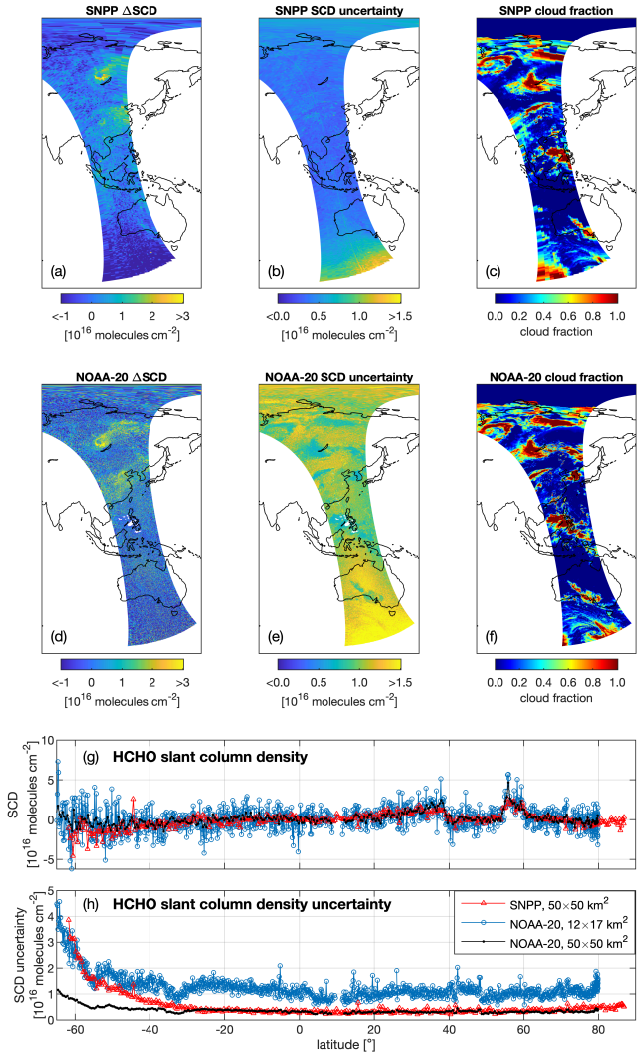

Figure 4. Formaldehyde differential slant column densities on 28 July 2019 for SZA $< 80°$ retrieved from (a) OMPS/SNPP orbit 40149 and (d) OMPS/NOAA-20 orbit 8752. The corresponding fitting uncertainties (panels b and e) and cloud fractions (c and f) are also shown. Panel (g) shows slant column densities and panel (h) shows their corresponding fitting uncertainties along one across-track position that passes through the wildfire for OMPS/NPP (across-track position 30), OMPS/NOAA-20 (position 57) and OMPS/NOAA-20 averaged to OMPS/SNPP spatial resolution (from positions 55–58). Only the ascending part of the orbit is shown. The missing retrievals in OMPS/NOAA-20 near 10°N are due to the exclusion of spectra flagged as saturated in the Level 1B data.

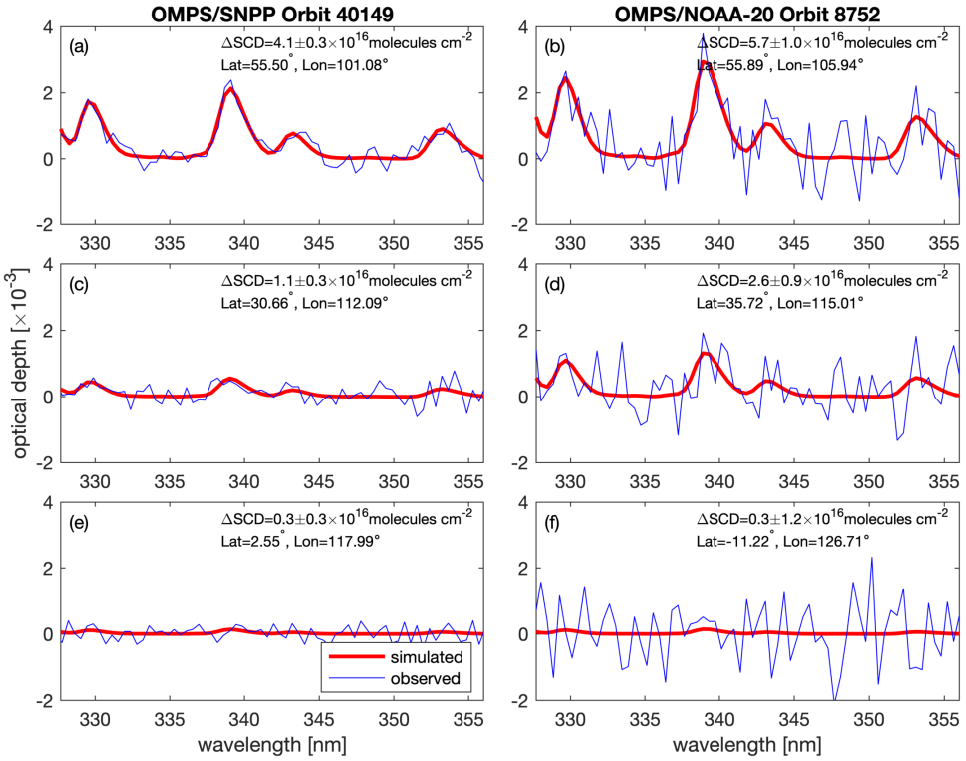

Figure 5. Modeled and observed optical depth spectra from OMPS/SNPP orbit 40149 and OMPS/NOAA-20 orbit 8752 on 28 July 2019, showing sample simulated (red) and observed (blue) HCHO optical depths from (a, b) a large Siberian wildfire, (c, d) moderate pollution over China and (e, f) a clean background. Each panel shows the differential retrieved slant column density ΔSCD between an observed spectrum and a radiance reference spectrum, and its associated fitting uncertainty.

Table 4. Baseline inputs to air mass factor calculations.

Parameter	Details
Radiative transfer model	VLIDORT V2.8 (Spurr, 2008)
Wavelength for calculation	340 nm
Trace gas profiles	GEOS-Chem 2018 monthly climatology
Temperature profile	GEOS-Chem 2018 monthly climatology
Digital elevation model	GLOBE (Hastings & Dunbar, 1999)
Surface pressure	MERRA-2 (GMAO, 2015)
Number of vertical layers in RTM	47 (reduced GEOS-5 grid)
Surface reflectance (land)	MODIS BRDF product MCD43C1 (Schaaf & Wang, 2015) extended to UV using SCIAMACHY
Surface reflectance (water)	Cox-Munk slope distribution (Cox & Munk, 1954)
Wind speed and direction	MERRA-2 (GMAO, 2015)
Ocean salinity	World Ocean Atlas 2009 (Antonov et al., 2010)
Cloud fraction	Derived from OMPS total ozone product reflectivity (Jaross, 2017b)
Cloud pressure	OMPS total ozone product (Jaross, 2017b)
Aerosols	not included explicitly

3.3.2 Radiative transfer calculation

We determine scattering weights $W(z)$ using the Vector LInearized Discrete Ordinate Radiative Transfer (VLIDORT) radiative transfer model Version 2.8 (Spurr, 2006, 2008). The scattering weights describe the sensitivity of the measurement to different altitude layers and are a function of the instrument viewing geometry, the ozone profile (which influences the penetration altitude of photons in the UV), aerosol and molecular scattering, and surface reflectance. For the HCHO retrievals, we calculate the scattering weights at 340 nm, and assume the wavelength dependency of the photon path to be constant within the narrow HCHO wavelength fitting window. Before the AMF calculation is run, we create a file for each orbit that includes trace gas profiles, surface reflectance parameters, and relevant meteorological variables such as temperature profiles, surface pressure and surface winds (see Figure 1). This file is then used as input to the radiative transfer code.

The radiative transfer calculation is performed on 47 layers from the surface to 0.01 hPa, defined by the reduced GEOS-5 vertical grid commonly used for GEOS-Chem simulations (`http://wiki.seas.harvard.edu/geos-chem/index.php/GEOS-Chem_vertical_grids#47-layer_reduced_vertical_grid`). This vertical grid maintains the GEOS-5 vertical layers in the troposphere, but reduces the stratosphere to 11 layers.

3.3.3 Surface reflectance

The surface reflectance for the AMF uses the MODIS observation-geometry dependent bidirectional reflectance distribution function (BRDF) product MCD43C1 (Schaaf & Wang, 2015). As the shortest wavelength band available from MODIS is at 469 nm, we have developed an approach to predict UV BRDFs from the first four MODIS bands (469 – 859 nm) by fitting a principal component analysis (PCA) decomposition model derived from surface spectral libraries (Zoogman et al., 2016). Since the original model only extended the BRDFs down to 400 nm, we have recently extended it to the UV by merging the original spectral databases with observations from the SCIAMACHY surface albedo database (Tilstra et al., 2017). For the OMPS HCHO retrievals, we calculate the surface reflectance at 340 nm.

We use VLIDORT to determine the surface reflectance over water using the Cox-Munk slope distribution (Cox & Munk, 1954) to represent sea surface roughness. The

surface wind speed and direction at each pixel is determined from the hourly Modern-Era Retrospective analysis for Research and Applications Version 2 (MERRA-2) product, which has a 0.5°×0.625° spatial resolution (Gelaro et al., 2017; GMAO, 2015). Ocean salinity is taken from the World Ocean Atlas 2009 (Antonov et al., 2010) at 1° resolution. The MODIS BRDF product is not available over open ocean, and its use over coastal and inland turbid waters to model surface reflectance is unreliable (Fasnacht et al., 2019). As a result, we use the Cox-Munk approach over all water bodies, recognizing there are likely large uncertainties in the surface reflectance in turbid and shallow waters.

The radiative effects of snow and ice cover are included implicitly through the use of MODIS BRDF data. However, we additionally include snow and ice cover fraction in the Level 2 data product for diagnostic reasons, even though these are not currently used in the AMF calculation. The northern hemisphere snow and sea ice fraction for each ground pixel is derived from the 4-km Interactive Multisensor Snow and Ice Mapping System (IMS) product (US National Ice Center, 2008). The southern hemisphere snow fraction is from the ancillary percent snow cover product included in the MODIS MCD43C1 product. Southern hemisphere sea ice fraction is estimated from the daily 25 × 25 km^2 Sea Ice Index product (Fetterer et al., 2017).

3.3.4 Atmospheric profiles

A priori atmospheric HCHO profiles are a key requirement of the retrieval as they are needed for determining the profile shape factor. The OMPS HCHO AMF calculation differs from that used in most previous retrievals of HCHO (i.e., González Abad et al., 2015, 2016; De Smedt et al., 2018) as it uses an online radiative transfer calculation that requires ozone and temperature profiles rather than using look-up tables built using standard profiles. We expect the influence of new profiles on the AMF to be minimal for temperature (<1 %), but somewhat larger in the case of potentially more accurate ozone profiles used in place of standard profiles (10 % in the most extreme cases).

We construct monthly climatologies of hourly HCHO, ozone and temperature profiles using output from a 2018 GEOS-Chem high-performance (GCHP) simulation (Eastham et al., 2018; Bindle et al., 2021) at 0.5°×0.5° spatial resolution on 72 vertical layers with a 1-year spin-up. GEOS-Chem is a global chemical transport model with detailed HO_x-NO_x-VOC-aerosol-halogen tropospheric chemistry (Bey et al., 2001). In this simulation,

the model is driven by meteorological fields from MERRA-2 (Gelaro et al., 2017). Global anthropogenic emissions are from the Community Emissions Data System (CEDS) (Hoesly et al., 2018), with Asian emissions replaced with the MIX inventory (M. Li et al., 2017). Biogenic emissions are determined online using the Model of Emissions of Gases and Aerosols from Nature (MEGAN) (Guenther et al., 2012). Biomass burning emissions use the fourth-generation Global Fire Emissions Database (GFED4) (Giglio et al., 2013).

We replace the monthly climatological surface pressure at each model grid box using regridded hourly surface pressures from MERRA-2 for the specific date and time of the satellite overpass. To account for differences between model resolution and satellite ground pixel resolution which may affect surface pressures in regions with changing terrain height, we additionally adjust the surface pressure of a satellite ground pixel by applying a terrain height correction using the 1-km Global Land One-kilometer Base Elevation (GLOBE) digital elevation model (Hastings & Dunbar, 1999) following the approach described by Zhou et al. (2009) and Boersma et al. (2011). Mixing ratio profiles are conserved but partial columns used in the shape factor are updated to reflect the new surface pressure on the satellite pixel footprint.

3.3.5 Clouds and aerosols

We apply the independent pixel approximation (Martin et al., 2002) to determine the effective AMF using Equation 9 with the cloud fractions described in Section 2.1.3, and cloud pressures from the OMPS total ozone product. Clouds are modeled in the radiative transfer simulation as Lambertian surfaces with albedo 0.8. As the radiative effects of aerosols are implicitly considered in the existing cloud retrievals, we do not currently model these in the AMF calculation but rather consider aerosols as a source of uncertainty in the final product (Jung et al., 2019).

3.4 Reference sector correction

The reference sector correction procedure consists of two steps. First, we add a background HCHO column which accounts for HCHO in the reference spectrum. Second, we apply a correction for the background bias that may be present in the retrievals due to unresolved instrument calibration or spectral fitting issues.

3.4.1 Correction for background HCHO in reference spectrum

We determine the background column SCD_R in Equation 4 by calculating the mean HCHO slant column density of the ground pixels used in the calculation of the reference spectrum. Each pixel's SCD_R is determined using Equation 3 with a VCD_R determined from the GEOS-Chem model climatology and the associated AMF_R from the reference orbit. The final SCD_R used in the Equation 4 correction is the mean of these individual pixel columns. As each across-track position uses a different reference spectrum, this results in a different SCD_R being applied for each across-track position in the orbit of interest. After this reference sector background column is determined for each across-track position, it is smoothed by fitting a third-order polynomial to the column as a function of across-track position. The reference sector background correction is typically on the order of 3.5–4.5$\times 10^{15}$ molecules cm^{-2}.

3.4.2 Bias correction

To account for unexplained background patterns in the HCHO retrievals which may be due to instrument or retrieval issues, we finally apply a latitude and solar zenith angle dependent bias correction.

First, we gather retrieved slant columns from the 30 Pacific reference orbits closest in time to the orbit of interest (typically covering a window of 30 days). While smaller temporal windows of ten days or less also work under clean conditions, the 30-day window helps to minimize the impact of large Arctic wildfires on the bias correction. Second, we calculate the difference between these retrieved columns and the theoretical modeled slant columns from the climatological profiles for each orbit using Equation 4. Third, we determine the mean difference (bias) between the modeled and retrieved columns for 1° latitude bins and 2° solar zenith angle bins. This binning ensures that the bias correction can be applied to occasional orbits that do not have the nominal number of OMPS observations. (Nominal observations are 36×400 for SNPP, and 104×1201 or 140×1201 for NOAA-20. However, on occasion orbits may have fewer along-track observations.) In this bias correction step, data are excluded that fall more than 3σ outside the median value in a window defined by across-track and along-track dimensions of 15×3 (SNPP), 45×9 (NOAA-20, 17×17 km^2), or 60×9 (NOAA-20, 12×17 km^2). Finally, the median value of each bin is saved as the bias. Then, for each pixel in the orbit of interest, the

bias SCD_B is determined for the pixel's latitude and solar zenith angle. The final bias correction is further smoothed using wavelet denoising.

Figure 2c and 2h show the bias correction determined for a sample orbit. We find across-track, latitude and SZA variability in the OMPS/NOAA-20 bias to be minimal for this wavelength fitting window. On the other hand, OMPS/SNPP biases are more significant at high latitudes, and largely increase as the measured radiance decreases with larger solar zenith angles. As these biases are not consistent between the two instruments, this likely points to unresolved calibration uncertainties or instrument differences rather than spectral fitting or radiative transfer issues at high latitudes.

3.4.3 Quality flags

We apply quality flags in post-processing. Pixels are flagged as 0) good, 1) suspect or 2) bad. We assign a bad flag to any observation with $|VCD| > 2 \times 10^{17}$ molecules cm^{-2}, $VCD+3\sigma_{VCD} < 0$, $AMF < 0.1$ or geometric AMF > 5. Although AMFs outside this range can be valid, they are likely highly uncertain. We flag pixels as suspect if $VCD+2\sigma_{VCD} < 0$ or geometric AMF > 4, or if they are snow or ice-covered (where the cloud retrieval is currently inaccurate). Furthermore, in general, we do not recommend the HCHO product be used in the case of high cloud fraction (> 0.4) due to large potential biases, or where SZA $> 70°$, when signal-to-noise is low and retrievals often show larger biases.

3.5 Uncertainties

Random and systematic uncertainties in the OMPS HCHO vertical column are introduced by uncertainties in the retrieved slant column density, the AMF and the reference sector corrections. Estimated uncertainties are summarized in Table 5 and discussed below. It should be noted that due to noise in the measurements, science applications of HCHO products frequently use temporal and/or spatial averaging. As a result, although we provide random and systematic uncertainty estimates, random uncertainties often become close to negligible in averaged columns, while systematic uncertainties remain.

3.5.1 Slant column density uncertainties

The random slant column density uncertainty in individual measurements is typically dominated by the random fitting uncertainty introduced by instrument noise. For OMPS/SNPP measurements with SZA $< 45°$ and cloud fractions < 0.4 these uncertainties are on the order of 3.5×10^{15} molecules cm^{-2}. OMPS/NOAA-20 SCD fitting uncertainties are on the order of 9.5×10^{15} molecules cm^{-2} (17×17 km^2) and 1.1×10^{16} molecules cm^{-2} (12×17 km^2).

Systematic errors in the slant column result from model parameter errors in the cross sections and instrument line shapes, and calibration uncertainties, as well as model errors that include the choice of polynomial fitting order and wavelength fitting window. De Smedt et al. (2018) provide a detailed error budget for HCHO slant column fitting uncertainties, and estimate a total systematic uncertainty from model parameters in HCHO background-corrected slant columns of 20%.

3.5.2 AMF uncertainties

Uncertainties in the AMF can result from uncertainties in inputs to the AMF calculation (model parameter errors), and from approximations in the calculation itself (forward model errors). Using different ancillary AMF inputs from different retrieval groups, Lorente et al. (2017) showed that structural uncertainties in the NO_2 AMF are on average 31 % in clean and 42 % in polluted regions. AMF errors are typically dominated by uncertainties in the assumed surface reflectance, aerosols, profile shape and cloud parameters. This section describes uncertainties introduced by those parameters into the OMPS retrieval.

Estimated surface reflectance uncertainties from MODIS BRDF vary by surface type, but generally have an RMSE < 0.0318 and bias within 0.0076 over land (Wang et al., 2018). Over open ocean, we estimate uncertainties of 0.018 (RMSE) and 0.015 (bias) in the surface reflectance, based on Fasnacht et al. (2019). These result in random uncertainties in the AMF on the order of 10 % (land) and 5 % (water) and AMF systematic uncertainties on the order of 3 % (land) and 5 % (water). As we also apply the open ocean approximation over coastal and turbid waters, the uncertainties in these regions are likely much higher but have not been quantified.

Aerosols are not explicitly considered in the AMF calculation, which, depending on the type and aerosol altitude, can lead to large errors when aerosol loading is high. Aerosols can either enhance or decrease sensitivity to HCHO, depending on the height of aerosols relative to HCHO and the aerosol optical properties. The cloud fraction retrieval is not able to differentiate between the effects of aerosols and clouds, and as a result implicitly considers aerosols to some extent (Boersma et al., 2004, 2011). Jung et al. (2019) used independent OMI aerosol measurements to explicitly consider the effect of aerosols on HCHO retrievals in clear-sky pixels. They found global mean biases in HCHO VCDs are largest in the presence of smoke aerosols (27±11%), with smaller differences from dust (6±6%) and sulfate (0.3±4%), while regional mean biases in VCDs ranged from -3 to 35 %. Uncertainties in individual measurements with the highest aerosol loading can be significantly larger (>100 %). As a result it is difficult to generalize random uncertainties due to aerosols. Caution is advised when using HCHO retrievals in regions of high aerosol loading.

Profile shape uncertainties due to the use of a climatology can contribute to the AMF uncertainty, although these can be mitigated by users through the re-calculation of the AMF by applying the included scattering weights to model profiles with higher spatial and temporal resolution (Laughner et al., 2019; Qin et al., 2020). To estimate the random uncertainties introduced by the use of a climatology, we examine the variability of the model's daily profiles. While there is little variability in the free troposphere, in the boundary layer the daily profile variability is on the order of ±30 % in polluted regions, relative to the climatological profiles. The resultant uncertainty in the climatological profile shape can result in AMF uncertainties of ∼30 % in those regions. Profile shape variability in clean regions is smaller, and results in AMF uncertainties closer to 10 %.

Zhu et al. (2020) found that over land, GEOS-Chem simulations show HCHO biased low near the surface as compared with in situ aircraft observations during multiple field campaigns, possibly due to inaccuracies in mixing depths and VOC emissions. When AMFs were recalculated using observed profile shapes, the seven most polluted regions saw HCHO VCD increase by 10±6 %. Remote ocean VCDs were less affected (changes of 0 and 5 % in two campaigns). Consequently, we estimate potential systematic biases of 5 % in clean regions and 10 % in polluted regions due to profile shape. Pro-

file shape uncertainties in fires are likely to be much larger, due to the inability of climatological profiles to accurately represent intermittent fire plumes.

Uncertainties in cloud fraction and cloud pressure can propagate significant uncertainties to the AMF. Through comparisons of our derived cloud fractions with those from the OMPS/SNPP Raman cloud product, we estimate potential overall systematic biases in the cloud fractions to be <0.005, which results in small systematic uncertainties on the order of ∼1 %. Random uncertainty is ∼0.02 for individual measurements, which translates to AMF random uncertainty of ∼2 % in low-HCHO observations, and ∼6 % in polluted ground pixels.

Uncertainties in the cloud pressures are expected to be one of the largest contributors to uncertainties in the OMPS HCHO products due to the use of climatological pressures. Figure 6 shows a histogram of global cloud pressure differences for a typical day in July between cloud pressures used in the HCHO AMF calculation (taken from the cloud climatologies), and those from the OMPS/SNPP Raman cloud product, as well as corresponding HCHO AMF differences. The very large pressure differences in clear and nearly-clear skies ($0 \leq f < 0.1$) are expected as here the cloud pressure retrieval struggles to retrieve the correct cloud pressure (in fact, for $0 \leq f \leq 0.05$, the Raman cloud product reports the surface scene pressure as the cloud pressure). More significantly, the cloud pressures of partly cloudy pixels ($0.1 \leq f < 0.4$) appear to be at least 150 hPa larger in the Raman cloud product. This results in large potential biases of 25 % in the AMF (Figure 6b). (Potential biases are smaller at other times of year, with January cloud pressure differences resulting in AMF biases on the order of 6 – 10 %.) Previous work has also shown OMPS Raman cloud pressures to be larger than those derived from OMI at higher cloud fractions (>0.3) at some latitudes (Vasilkov et al., 2014). As it is not clear that the OMPS Raman cloud pressures are "truth", we estimate potential biases in the climatology of 50 hPa, which is more in line with previous assessments of OMI Raman clouds (Joiner et al., 2012); this results in an AMF systematic uncertainties of 5 – 15 % in partly cloudy pixels. We estimate random uncertainties in individual OMPS HCHO AMFs due to the use of climatological pressures of ∼8 – 15 %, depending on cloud fraction amount.

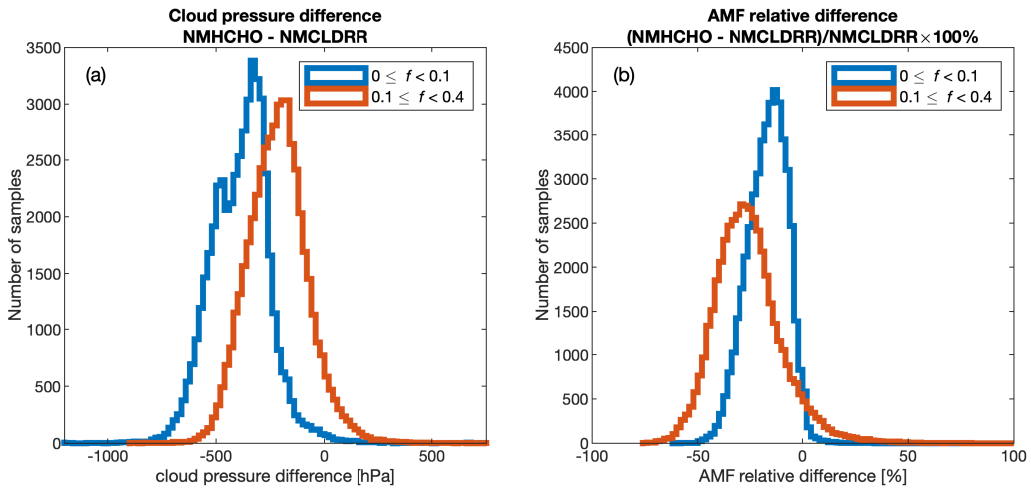

Figure 6. (a) Differences in cloud pressures and (b) air mass factors between the OMPS HCHO product (NMHCHO) and those calculated using cloud inputs from the NMCLDRR OMPS cloud product for all OMPS/SNPP orbits on 15 July 2019. The NMHCHO product uses a corrected cloud fraction based on NMCLDRR and the climatological cloud pressures from the NMTO3 product. The NMCLDRR cloud product retrieves cloud fraction and cloud pressure from rotational Raman scattering.

3.5.3 Reference sector correction uncertainties

The reference sector uncertainties result from uncertainties in the reference background correction and the bias correction. In situ airborne observations of HCHO in the remote Pacific during the Atmospheric Tomography Mission (ATom)-1 and ATom-2 campaigns show a HCHO background column of 3.0×10^{15} molecules cm^{-2} (Zhu et al., 2020) in the OMI reference region, close to the estimated modeled background VCD of $\sim 3.2\times 10^{15}$ molecules cm^{-2} used in the OMPS reference correction during those times. The random uncertainty in reference VCDs is small, and further minimized by averaging. Assuming a reference AMF uncertainty of 15 %, both the random and systematic uncertainties in the reference slant column are dominated by the reference AMF uncertainties, and are on the order of 6×10^{14} molecules cm^{-2}.

Though the bias correction is calculated using data from multiple reference orbits and smoothed, in the current implementation we do find it increases the random error on individual corrected slant columns by $\sim 1 \times 10^{15}$ molecules cm^{-2}. Systematic un-

certainty introduced by the bias correction is difficult to quantify, as its purpose is to remove unexplained systematic biases present in the differential slant columns. As a result, we estimate the overall systematic uncertainties in corrected slant column densities, including reference background and bias correction, to be on the order of 20 % after De Smedt et al. (2018).

3.6 Global products

Figure 7 shows seasonal HCHO means from the two OMPS instruments for 2019 (December is from 2018), regridded to 0.1°×0.1° using physical oversampling (Sun et al., 2018). HCHO vertical columns from the two OMPS instruments are very similar, showing nearly identical spatial distribution and columns very similar in magnitude (this will be assessed quantitatively in Section 4). The figure clearly shows regional HCHO source variability throughout the year, including increases in HCHO from isoprene emissions in the northern hemisphere summer (particularly large in the southeastern US), seasonal variation in biomass burning in South America, Africa and Southeast Asia, and regions with significant anthropogenic HCHO production (often mixed with biogenic sources), such as northern India and East China.

4 Satellite intercomparisons

In this section, we intercompare HCHO derived from the two OMPS instruments, and cross-validate with the Sentinel-5P/TROPOMI HCHO product. We do not perform comparisons with OMI HCHO in this paper as SAO OMI products are being transitioned to Collection 4 Level 1B spectra and an updated OMI HCHO product is forthcoming. Direct orbit-to-orbit comparisons of OMPS/SNPP and OMPS/NOAA-20 are complicated by the 50-minute orbital offset, which causes the instruments to view the same location at different times with different geometries, and when atmospheric conditions and cloud cover may have changed. In lieu of direct comparisons, we examine OMPS/SNPP and OMPS/NOAA-20 long-term monthly averages to explore the consistency and stability of the two instruments. In the following comparisons, we filter data using the OMPS and TROPOMI main quality flags, and exclude data with SZA $> 70°$ and cloud fractions $>$ 0.4.

Table 5. Sources of uncertainty in individual OMPS HCHO retrievals.

Source	Uncertainty	Notes
Slant column density (random)		
Measurement noise	3.5×10^{15} (SNPP)	Units are molecules cm^{-2}
	9.5×10^{15} (NOAA-20: 17×17 km^2)	SZA $< 45°$, cloud fraction < 0.4
	1.1×10^{16} (NOAA-20: 12×17 km^2)	
Bias correction	1×10^{15}	Units are molecules cm^{-2}
		SZA $< 45°$
Slant column density (systematic)		
Systematic errors	20 %	Uncertainty in corrected SCD,
		based on De Smedt et al. (2018)
Air mass factor (random)		
Surface reflectance	10 %, 5 %	Land, water
Aerosols	0 – >100 %	Depends on aerosol loading and type
Profile shape	10 %, 30 %	low HCHO, high HCHO
Cloud fraction	2 %, 6 %	low HCHO, high HCHO
Cloud pressure	8 – 15 %	
Air mass factor (systematic)		
Surface reflectance	3 %, 5 %	Land, water
Aerosols	-3 – 35 %	Jung et al. (2019) regional biases,
		depends on aerosol loading and type
Profile shape	5 %, 10 %	low HCHO, high HCHO
Cloud fraction	1 %	
Cloud pressure	5 – 15 %	

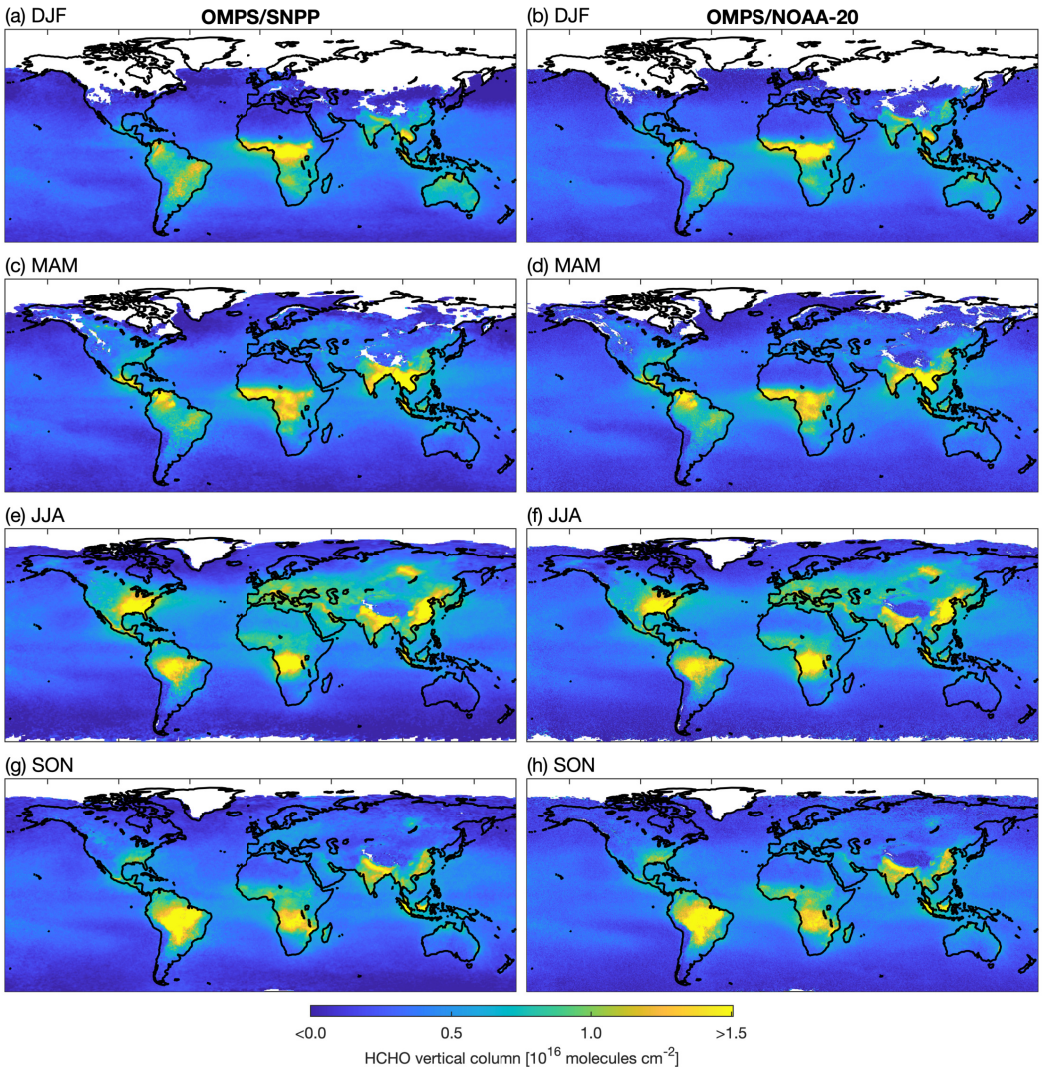

Figure 7. Seasonal 2019 mean HCHO vertical column densities at 0.1° × 0.1° resolution from OMPS/SNPP and OMPS/NOAA-20 (calculated for SZA < 70°, cloud fractions < 0.4, excluding snow and ice) for (a,b) December (2018)/January/February (DJF), (c,d) March/April/May (MAM), (e,f) June/July/August (JJA) and (g,h) September/October/November (SON).

Figure 8 shows mean August 2019 HCHO vertical columns and AMFs from both OMPS instruments and TROPOMI. For these plots, OMPS is regridded to $0.1° \times 0.1°$, with TROPOMI regridded to $0.05° \times 0.05°$. Overall, OMPS/SNPP and OMPS/NOAA-20 monthly averages are highly consistent in magnitude and spatial distribution. In some regions, OMPS HCHO during this month shows large spatial deviations with TROPOMI HCHO. These differences primarily result from differences in the AMFs, which use different ancillary inputs for surface reflectance, profile shape and clouds. Notable examples include the northern part of South America, where AMFs are much larger in TROPOMI data; northern India, where OMPS a priori profile shapes have much larger surface concentrations; and Siberian wildfires that are visible in OMPS data but missing from TROPOMI data due to those pixels being removed by the quality flag. In general, OMPS AMFs are smaller than those of TROPOMI over land, and show sharper land-water differences that result from different surface reflectance assumptions over land and water.

The black boxes in Figure 8 show the geographic regions that we use for an examination of long-term monthly averages. The boundaries of these regions are defined in Table 6. These regions are chosen based on those used in the original OMPS satellite intercomparisons by González Abad et al. (2016). We additionally add the Middle East, India and Southeast Asia regions to this comparison. Figure 9 shows a time series of monthly mean HCHO vertical columns from OMPS/SNPP, OMPS/NOAA-20 and TROPOMI in those regions. Figure 10 shows the corresponding slant columns, corrected for reference background and biases (before application of the AMF). Figure 11 presents the HCHO monthly mean vertical columns from Figure 9 as correlation plots.

As noted in Figure 11, OMPS/SNPP and OMPS/NOAA-20 HCHO monthly means are highly correlated ($r = 0.98$), and show overall negligible biases in magnitude, with a proportional bias of 2 % and an offset bias of 2×10^{14} molecules cm^{-2}. While the overall agreement between OMPS/SNPP and OMPS/NOAA-20 monthly means is excellent, there are some differences visible in the regional time series data of vertical column densities (Figure 9). The largest differences occur in the Southeast US and East China winters, where OMPS/SNPP tends to underestimate HCHO relative to OMPS/NOAA-20 by as much as 30%. However, in general, deviations between OMPS/SNPP and OMPS/NOAA-20 monthly means rarely exceed 1×10^{15} molecules cm^{-2}.

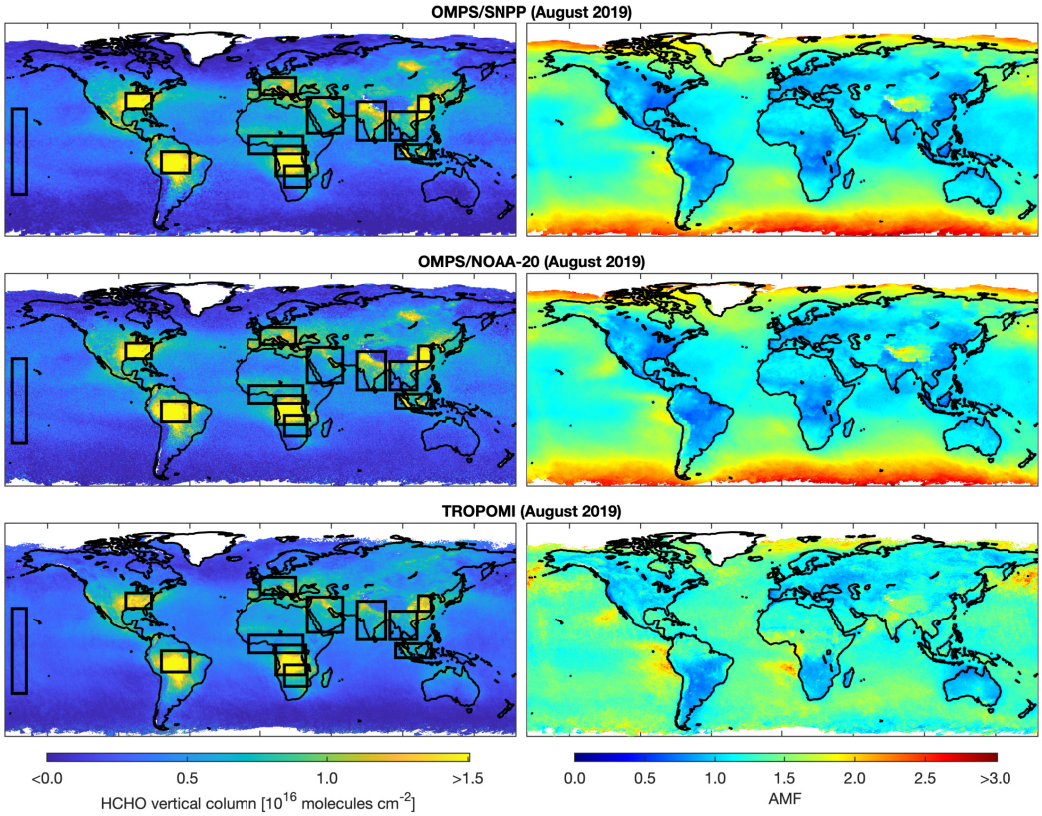

Figure 8. August 2019 mean HCHO vertical column densities and air mass factors for OMPS/SNPP, OMPS/NOAA-20 and TROPOMI. The black boxes in the left column HCHO maps show the geographic regions examined in Figures 9 and 10 described in Table 6.

OMPS/SNPP and OMPS/NOAA-20 monthly means are also highly correlated with TROPOMI ($r = 0.92$), but comparisons show a slope of ∼1.18 relative to TROPOMI. For context, these differences are well within the potential biases in HCHO retrievals discussed in Section 3.5. At higher levels of HCHO ($> 8 \times 10^{15}$ molecules cm^{-2}), OMPS HCHO vertical columns are consistently higher than HCHO derived from TROPOMI by 10±16%. Less polluted regions and times show better agreement between OMPS and TROPOMI, although there are also differences in cleaner regions such as the European and Southern African winters, where TROPOMI HCHO is much higher than OMPS HCHO.

The OMPS HCHO products are validated using ground-based FTIR measurements in a separate companion paper (Kwon et al., 2022). Similar to the TROPOMI validations (Vigouroux et al., 2020; De Smedt et al., 2021) which found TROPOMI HCHO to be larger than HCHO measured by ground-based measurements at clean sites, this study finds OMPS HCHO VCDs are larger than FTIR columns at clean sites (by 20 % and 32 % for SNPP and NOAA-20, respectively). Polluted sites ($> 4 \times 10^{15}$ molecules cm^{-2}) show closer agreement with the FTIR columns with biases of -15 % (OMPS/SNPP) and +0.5 % (OMPS/NOAA-20). As previously mentioned in Section 2.2, TROPOMI HCHO at polluted sites has previously been found to be 25 – 31 % lower than coincident HCHO measured from FTIR and MAX-DOAS networks (Vigouroux et al., 2020; De Smedt et al., 2021).

Most of the differences between OMPS and TROPOMI can be attributed to differences in the AMF. As shown in Figure 10, the corrected slant columns show minimal bias between the three instruments in most regions, implying differences in AMFs drive the VCD differences. OMPS/SNPP and OMPS/NOAA-20 monthly mean corrected slant columns show excellent correlation with each other ($r = 0.97$, slope = 0.97, intercept $= 3 \times 10^{14}$ molecules cm^{-2}), as do those of TROPOMI with OMPS/SNPP ($r = 0.96$, slope = 0.92, intercept = -1×10^{14} molecules cm^{-2}) and OMPS/NOAA-20 ($r = 0.96$, slope = 0.91, intercept = -2×10^{13} molecules cm^{-2}). Biases between TROPOMI and OMPS corrected slant columns of approximately 40 % during winter in the highest latitude regions (Europe and Southern Africa) may explain the associated biases seen in those regions in the VCD, and may be due to differences in the reference sector corrections.

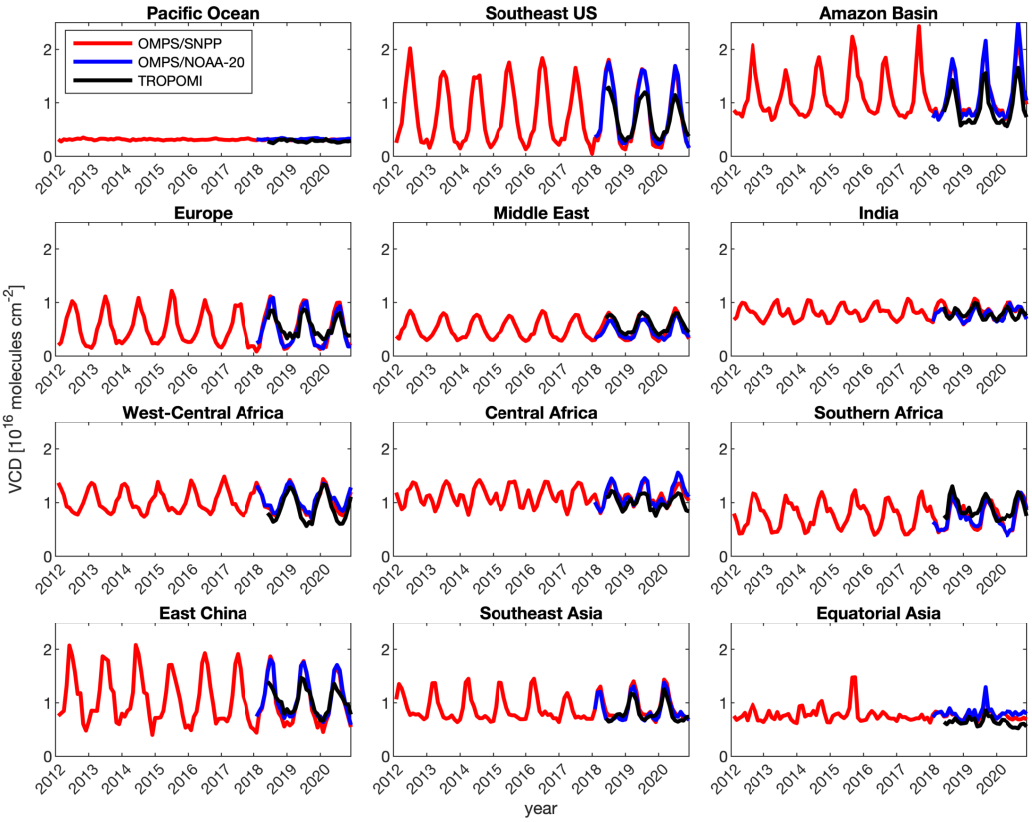

Figure 9. Time series of monthly average HCHO vertical column densities from OMPS/SNPP, OMPS/NOAA-20 and TROPOMI for the regions illustrated in Figure 8.

Table 6. Geographic limits of regions used in Figures 11 and 10 times series.

Region	Latitude Limits	Longitude Limits
Pacific Ocean	30°S – 30°N	175°W – 165°W
Southeast US	30°N – 41°N	95°W – 77°W
Amazon Basin	15°S – 0°	70°W – 50°W
Europe	40°N – 52°N	0° – 25°E
Middle East	13°N – 38°N	33°E – 58°E
India	8°N – 35°N	68°E – 88°E
West-Central Africa	1.5°S – 11°N	8°W – 30°E
Central Africa	17°S – 4°N	11°E – 32°E
Southern Africa	25°S – 10°S	17°E – 33°E
East China	28°N – 39°N	111°E – 120°E
Southeast Asia	8°N – 28°N	91°E – 110°E
Equatorial Asia	5°S – 5°N	95°E – 120°E

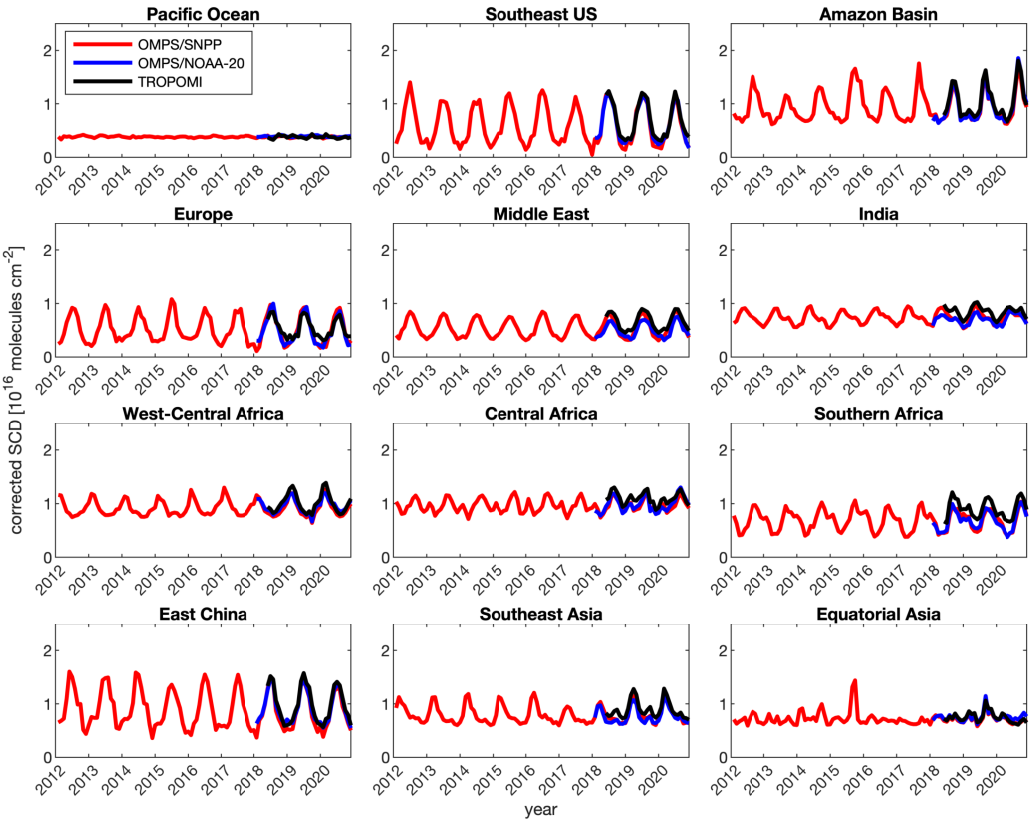

Figure 10. Same as Figure 9 but for corrected slant column densities.

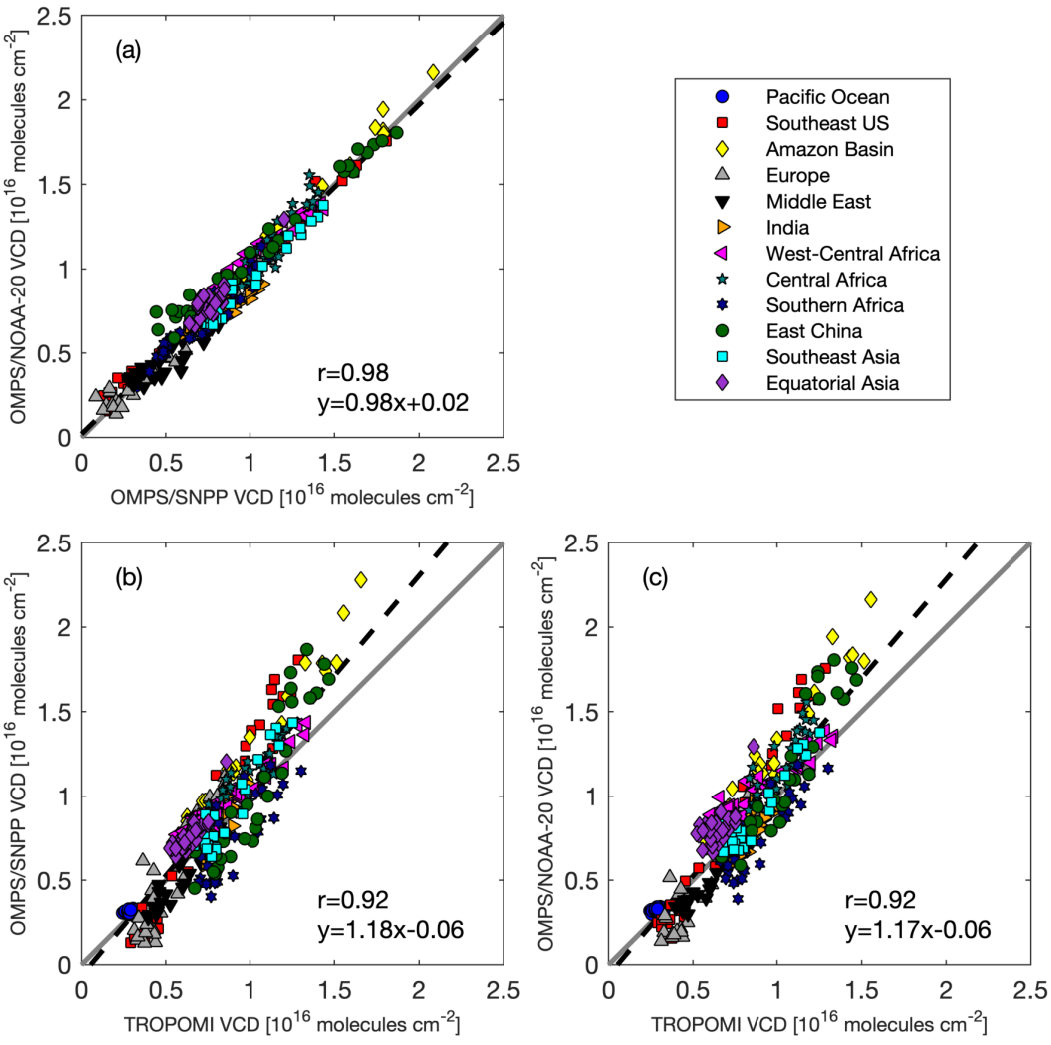

Figure 11. Correlation plot of monthly average HCHO vertical column densities for geographic regions shown in Figure 9 for (a) OMPS/SNPP vs. OMPS/NOAA-20, (b) TROPOMI vs. OMPS/SNPP and (c) TROPOMI vs. OMPS/NOAA-20.

5 Summary and future work

The OMPS instruments on Suomi NPP and NOAA-20 have been used to produce publicly-available, multi-year data records of HCHO. These data are retrieved using a three-step procedure: 1) spectral fitting of slant column density, following an on-orbit instrument line shape and wavelength calibration; 2) a scene-by-scene AMF calculation; and 3) a reference sector correction that applies a background HCHO column and a bias correction.

Monthly mean formaldehyde derived from the two OMPS instruments shows excellent agreement over 12 geographic regions. Overall, these comparisons of OMPS/SNPP versus OMPS/NOAA-20 show excellent correlation ($r = 0.98$) with a slope of 0.98 and intercept of $2{\times}10^{14}$ molecules cm^{-2}. The correlation with TROPOMI is also very good ($r = 0.92$ for both OMPS instruments), but OMPS HCHO is higher overall (slope = 1.18, intercept = -6×10^{14} molecules cm^{-2}).

Future OMPS HCHO work will include assessing OMPS HCHO with other HCHO products derived from satellite instruments as part of the long-term Making Earth System Data Records for Use in Research Environment (MEaSUREs) program at the SAO, comparisons with the forthcoming OMI Collection 4 HCHO operational product, and comparisons with the long-term data records from the European Quality Assurance for Essential Climate Variables (QA4ECV) project (De Smedt et al., 2018).

While beyond the scope of this paper, investigations that fully explore ancillary inputs to the OMPS and TROPOMI AMF calculations would help in the interpretation of HCHO measurements and their validation, and could lead to improvements in both products. In future, OMPS HCHO could also benefit from independent, validated cloud products derived from both OMPS/SNPP and OMPS/NOAA-20, and independently derived cloud fractions at 340 nm that use the same inputs as the HCHO AMF calculation. Improved cloud products could also improve the reliability of the HCHO retrievals over snow and ice.

OMPS HCHO products extend and compliment the global HCHO afternoon data records that began with OMI in 2004. The retrieval described in this paper has the potential to be applied to the OMPS instruments on future JPSS satellites, which would ensure a consistent long-term stable data record of global afternoon HCHO into the 2030's.

6 Open research

6.1 Data availability statement

The OMPS HCHO products described in this paper are available from the NASA GES DISC for OMPS/Suomi-NPP (https://doi.org/10.5067/IIM1GHT07QA8) and OMPS/NOAA-20 (https://doi.org/10.5067/CIYXT9A4I2F4).

The OMPS/Suomi-NPP datasets used to generate the HCHO products are available from the NASA GES DISC for the Level 1B radiances (https://doi.org/10.5067/DL081SQY7C89), total ozone (https://doi.org/10.5067/0WF4HAAZ0VHK), and rotational Raman cloud products (https://doi.org/10.5067/CJAALTQUCLO2). The OMPS/NOAA-20 radiances and total ozone products are available from the OMPS project website at https://ozoneaq.gsfc.nasa.gov/omps/.

MODIS BRDF data used in the AMF calculation are available from the NASA Land Processes Distributed Active Archive Center at https://doi.org/10.5067/MODIS/MCD43C1.006.

MERRA-2 data use for meteorological variables are available from the NASA GES DISC at https://doi.org/10.5067/VJAFPLI1CSIV.

Sentinel-5P/TROPOMI HCHO data are available from https://doi.org/10.5270/S5P-vg1i7t0.

Acknowledgments

This study was supported by the NOAA Climate Program Office's Atmospheric Chemistry, Carbon Cycle, and Climate program (NA18OAR4310108), NASA's The Science of Terra, Aqua and SuomiNPP (80NSSC18K0691), and NASA's Making Earth System Data Records for Use in Research Environments (80NSSC18M0091). Computations in this paper were conducted on the Smithsonian High Performance Cluster (SI/HPC), Smithsonian Institution (https://doi.org/10.25572/SIHPC). We thank Jinkyul Choi at the University of Colorado Boulder and Amir Souri at the SAO for feedback on the data, and Lawrence Flynn at NOAA for helpful discussions.

References

Ahmad, Z., Bhartia, P. K., & Krotkov, N. (2004). Spectral properties of backscattered UV radiation in cloudy atmospheres. *Journal of Geophysical Research:*

Atmospheres, *109*(D1). doi: 10.1029/2003JD003395

Antonov, J. I., Seidov, D., Boyer, T. P., Locarnini, R. A., Mishonov, A. V., Garcia, H. E., ... Johnson, D. R. (2010). *World ocean atlas 2009, volume 2: Salinity* (Vol. NOAA Atlas NESDIS 69; S. Levitus, Ed.). Washington, D.C.: U.S. Government Printing Office.

Bak, J., Liu, X., Kim, J.-H., Haffner, D. P., Chance, K., Yang, K., & Sun, K. (2017). Characterization and correction of OMPS nadir mapper measurements for ozone profile retrievals. *Atmospheric Measurement Techniques*, *10*(11), 4373–4388. doi: 10.5194/amt-10-4373-2017

Barkley, M. P., Palmer, P. I., Kuhn, U., Kesselmeier, J., Chance, K., Kurosu, T. P., ... Guenther, A. (2008). Net ecosystem fluxes of isoprene over tropical South America inferred from Global Ozone Monitoring Experiment (GOME) observations of HCHO columns. *Journal of Geophysical Research: Atmospheres*, *113*(D20). doi: 10.1029/2008JD009863

Bauwens, M., Stavrakou, T., Müller, J.-F., De Smedt, I., Van Roozendael, M., van der Werf, G. R., ... Guenther, A. (2016). Nine years of global hydrocarbon emissions based on source inversion of OMI formaldehyde observations. *Atmospheric Chemistry and Physics*, *16*(15), 10133–10158. doi: 10.5194/acp-16-10133-2016

Beirle, S., Lampel, J., Lerot, C., Sihler, H., & Wagner, T. (2017). Parameterizing the instrumental spectral response function and its changes by a super-Gaussian and its derivatives. *Atmospheric Measurement Techniques*, *10*(2), 581–598. doi: 10.5194/amt-10-581-2017

Bey, I., Jacob, D. J., Yantosca, R. M., Logan, J. A., Field, B. D., Fiore, A. M., ... Schultz, M. G. (2001). Global modeling of tropospheric chemistry with assimilated meteorology: Model description and evaluation. *Journal of Geophysical Research: Atmospheres*, *106*(D19), 23073-23095. doi: 10.1029/2001JD000807

Bhartia, P. K., & Wellemeyer, C. W. (2002). TOMS-V8 Total O_3 Algorithm. In P. K. Bhartia (Ed.), *OMI Algorithm Theoretical Basis Document, Volume II: OMI Ozone Products* (chap. 2). NASA Goddard Space Flight Center. (ATBD-OMI-02, Version 2.0)

Bindle, L., Martin, R. V., Cooper, M. J., Lundgren, E. W., Eastham, S. D., Auer, B. M., ... Jacob, D. J. (2021). Grid-stretching capability for the GEOS-Chem

13.0.0 atmospheric chemistry model. *Geoscientific Model Development*, *14*(10), 5977–5997.

Boersma, K. F., Eskes, H. J., & Brinksma, E. J. (2004). Error analysis for tropospheric NO_2 retrieval from space. *Journal of Geophysical Research: Atmospheres*, *109*(D4). doi: 10.1029/2003JD003962

Boersma, K. F., Eskes, H. J., Dirksen, R. J., van der A, R. J., Veefkind, J. P., Stammes, P., ... Brunner, D. (2011). An improved tropospheric NO_2 column retrieval algorithm for the Ozone Monitoring Instrument. *Atmospheric Measurement Techniques*, *4*(9), 1905–1928. doi: 10.5194/amt-4-1905-2011

Chance, K., Kurosu, T. P., & Sioris, C. E. (2005). Undersampling correction for array detector-based satellite spectrometers. *Applied Optics*, *44*(7), 1296–1304. doi: 10.1364/AO.44.001296

Chance, K., & Kurucz, R. L. (2010). An improved high-resolution solar reference spectrum for Earth's atmosphere measurements in the ultraviolet, visible, and near infrared. *Journal of Quantitative Spectroscopy and Radiative Transfer*, *111*(9), 1289–1295. doi: 10.1016/j.jqsrt.2010.01.036

Chance, K., & Orphal, J. (2011). Revised ultraviolet absorption cross sections of H_2CO for the HITRAN database. *Journal of Quantitative Spectroscopy and Radiative Transfer*, *112*(9), 1509–1510. doi: 10.1016/j.jqsrt.2011.02.002

Chance, K., Palmer, P. I., Spurr, R. J., Martin, R. V., Kurosu, T. P., & Jacob, D. J. (2000). Satellite observations of formaldehyde over North America from GOME. *Geophysical Research Letters*, *27*(21), 3461–3464. doi: 10.1029/2000GL011857

Chance, K., & Spurr, R. J. D. (1997). Ring effect studies: Rayleigh scattering, including molecular parameters for rotational raman scattering, and the Fraunhofer spectrum. *Applied Optics*, *36*(21), 5224–5230. doi: 10.1364/AO.36.005224

Chan Miller, C., Gonzalez Abad, G., Wang, H., Liu, X., Kurosu, T., Jacob, D. J., & Chance, K. (2014). Glyoxal retrieval from the Ozone Monitoring Instrument. *Atmospheric Measurement Techniques*, *7*(11), 3891–3907. doi: 10.5194/amt-7-3891-2014

Choi, J., Henze, D. K., Cao, H., Nowlan, C. R., González Abad, G., Kwon, H.-A., ... Weinheimer, A. J. (2022). An inversion framework for optimizing non-

methane VOC emissions using remote sensing and airborne observations in northeast Asia during the KORUS-AQ field campaign. *Journal of Geophysical Research: Atmospheres*, *127*(7), e2021JD035844. doi: 10.1029/2021JD035844

Cox, C., & Munk, W. (1954). Measurement of the roughness of the sea surface from photographs of the Sun's glitter. *Journal of the Optical Society of America*, *44*(11), 838–850. doi: 10.1364/JOSA.44.000838

De Smedt, I., Müller, J.-F., Stavrakou, T., Van Der A, R., Eskes, H., & Van Roozendael, M. (2008). Twelve years of global observations of formaldehyde in the troposphere using GOME and SCIAMACHY sensors. *Atmospheric Chemistry and Physics*, *8*(16). doi: 10.5194/acp-8-4947-2008

De Smedt, I., Pinardi, G., Vigouroux, C., Compernolle, S., Bais, A., Benavent, N., ... Van Roozendael, M. (2021). Comparative assessment of TROPOMI and OMI formaldehyde observations and validation against MAX-DOAS network column measurements. *Atmospheric Chemistry and Physics*, *21*(16), 12561–12593. doi: 10.5194/acp-21-12561-2021

De Smedt, I., Stavrakou, T., Hendrick, F., Danckaert, T., Vlemmix, T., Pinardi, G., ... Van Roozendael, M. (2015). Diurnal, seasonal and long-term variations of global formaldehyde columns inferred from combined OMI and GOME-2 observations. *Atmospheric Chemistry and Physics*, *15*(21), 12519–12545. doi: 10.5194/acp-15-12519-2015

De Smedt, I., Theys, N., Yu, H., Danckaert, T., Lerot, C., Compernolle, S., ... Veefkind, P. (2018). Algorithm theoretical baseline for formaldehyde retrievals from S5P TROPOMI and from the QA4ECV project. *Atmospheric Measurement Techniques*, *11*(4), 2395–2426. doi: 10.5194/amt-11-2395-2018

De Smedt, I., Van Roozendael, M., Stavrakou, T., Müller, J.-F., Lerot, C., Theys, N., ... van der A, R. (2012). Improved retrieval of global tropospheric formaldehyde columns from GOME-2/MetOp-A addressing noise reduction and instrumental degradation issues. *Atmospheric Measurement Techniques*, *5*(11), 2933–2949. doi: 10.5194/amt-5-2933-2012

Eastham, S. D., Long, M. S., Keller, C. A., Lundgren, E., Yantosca, R. M., Zhuang, J., ... Jacob, D. J. (2018). GEOS-Chem High Performance (GCHP v11-02c): a next-generation implementation of the GEOS-Chem chemical transport model for massively parallel applications. *Geoscientific Model Development*,

11(7), 2941–2953. doi: 10.5194/gmd-11-2941-2018

ESA, & DLR. (2019a). *Sentinel-5P TROPOMI Tropospheric Formaldehyde HCHO 1-Orbit L2 5.5km x 3.5km* [datset]. Greenbelt, MD, USA: Goddard Earth Sciences Data and Information Services Center (GES DISC). (Copernicus Sentinel data processed by ESA, German Aerospace Center (DLR)) doi: 10.5270/S5P-vg1i7t0

ESA, & DLR. (2019b). *Sentinel-5P TROPOMI Tropospheric Formaldehyde HCHO 1-Orbit L2 7km x 3.5km* [dataset]. Greenbelt, MD, USA: Goddard Earth Sciences Data and Information Services Center (GES DISC). (Copernicus Sentinel data processed by ESA, German Aerospace Center (DLR)) doi: 10.5270/S5P-vg1i7t0

ESA, & DLR. (2020). *Sentinel-5P TROPOMI Tropospheric Formaldehyde HCHO 1-Orbit L2 5.5km x 3.5km* [dataset]. Greenbelt, MD, USA: Goddard Earth Sciences Data and Information Services Center (GES DISC). (Copernicus Sentinel data processed by ESA, German Aerospace Center (DLR)) doi: 10.5270/S5P-vg1i7t0

Fasnacht, Z., Vasilkov, A., Haffner, D., Qin, W., Joiner, J., Krotkov, N., ... Spurr, R. (2019). A geometry-dependent surface Lambertian-equivalent reflectivity product for UV–Vis retrievals – part 2: Evaluation over open ocean. *Atmospheric Measurement Techniques*, *12*(12), 6749–6769. doi: 10.5194/amt-12-6749-2019

Fetterer, F., Knowles, K., Meier, W. N., Savoie, M., & Windnagel, A. K. (2017). *Sea ice index, version 3.* Boulder, Colorado USA: NSIDC: National Snow and Ice Data Center. (updated daily) doi: 10.7265/N5K072F8

Finkenzeller, H., & Volkamer, R. (2022). O_2–O_2 CIA in the gas phase: Cross-section of weak bands, and continuum absorption between 297–500 nm. *Journal of Quantitative Spectroscopy and Radiative Transfer*, *279*, 108063. doi: 10.1016/j.jqsrt.2021.108063

Flynn, L., Long, C., Wu, X., Evans, R., Beck, C. T., Petropavlovskikh, I., ... Seftor, C. (2014). Performance of the Ozone Mapping and Profiler Suite (OMPS) products. *Journal of Geophysical Research: Atmospheres*, *119*(10), 6181–6195. (2013JD020467) doi: 10.1002/2013JD020467

Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., ...

Zhao, B. (2017, jul). The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2). *Journal of Climate*, *30*(14), 5419–5454. doi: 10.1175/JCLI-D-16-0758.1

Giglio, L., Randerson, J. T., & van der Werf, G. R. (2013). Analysis of daily, monthly, and annual burned area using the fourth-generation global fire emissions database (GFED4). *Journal of Geophysical Research: Biogeosciences*, *118*(1), 317-328. doi: 10.1002/jgrg.20042

GMAO. (2015). *MERRA-2 tavg1_2d_slv_Nx: 2d,1-Hourly,Time-Averaged,Single-Level,Assimilation,Single-Level Diagnostics V5.12.4* [dataset]. Greenbelt, MD, USA: Goddard Earth Sciences Data and Information Services Center (GES DISC). doi: 10.5067/VJAFPLI1CSIV

Goldberg, M. D., Kilcoyne, H., Cikanek, H., & Mehta, A. (2013). Joint Polar Satellite System: The United States next generation civilian polar-orbiting environmental satellite system. *Journal of Geophysical Research: Atmospheres*, *118*(24), 13,463-13,475. doi: 10.1002/2013JD020389

González Abad, G. (2022a). *OMPS-N20 L2 NM Formaldehyde (HCHO) Total Column swath orbital* [dataset]. Greenbelt, MD, USA: Goddard Earth Sciences Data and Information Services Center (GES DISC). doi: 10.5067/CIYXT9A4I2F4

González Abad, G. (2022b). *OMPS-NPP L2 NM Formaldehyde (HCHO) Total Column swath orbital* [dataset]. Greenbelt, MD, USA: Goddard Earth Sciences Data and Information Services Center (GES DISC). doi: 10.5067/IIM1GHT07QA8

González Abad, G., Liu, X., Chance, K., Wang, H., Kurosu, T. P., & Suleiman, R. (2015). Updated Smithsonian Astrophysical Observatory Ozone Monitoring Instrument (SAO OMI) formaldehyde retrieval. *Atmospheric Measurement Techniques*, *8*(1), 19–32. doi: 10.5194/amt-8-19-2015

González Abad, G., Vasilkov, A., Seftor, C., Liu, X., & Chance, K. (2016). Smithsonian Astrophysical Observatory Ozone Mapping and Profiler Suite (SAO OMPS) formaldehyde retrieval. *Atmospheric Measurement Techniques*, *9*(7), 2797–2812. doi: 10.5194/amt-9-2797-2016

Guenther, A. B., Jiang, X., Heald, C. L., Sakulyanontvittaya, T., Duhl, T., Emmons, L. K., & Wang, X. (2012). The Model of Emissions of Gases and

963 Aerosols from Nature version 2.1 (MEGAN2.1): an extended and updated
964 framework for modeling biogenic emissions. *Geoscientific Model Development*,
965 *5*(6), 1471–1492. doi: 10.5194/gmd-5-1471-2012

966 Hastings, D. A., & Dunbar, P. K. (1999). *Global Land One-kilometer Base Eleva-*
967 *tion (GLOBE) Digital Elevation Model, Documentation, Volume 1.0* (Tech.
968 Rep.). Boulder, Colorado USA: National Oceanic and Atmospheric Admin-
969 istration, National Geophysical Data Center. (Key to Geophysical Records
970 Documentation (KGRD) 34)

971 Hoesly, R. M., Smith, S. J., Feng, L., Klimont, Z., Janssens-Maenhout, G., Pitka-
972 nen, T., ... Zhang, Q. (2018). Historical (1750–2014) anthropogenic emis-
973 sions of reactive gases and aerosols from the Community Emissions Data
974 System (CEDS). *Geoscientific Model Development*, *11*(1), 369–408. doi:
975 10.5194/gmd-11-369-2018

976 Ingmann, P., Veihelmann, B., Langen, J., Lamarre, D., Stark, H., & Courrèges-
977 Lacoste, G. B. (2012). Requirements for the GMES Atmosphere Service and
978 ESA's implementation concept: Sentinels-4/-5 and-5p. *Remote Sensing of*
979 *Environment*, *120*, 58–69. doi: 10.1016/j.rse.2012.01.023

980 Jaross, G. (2017a). *OMPS/NPP L1B NM Radiance EV Calibrated Geolocated Swath*
981 *Orbital V2* [dataset]. Greenbelt, MD, USA: Goddard Earth Sciences Data and
982 Information Services Center (GES DISC). doi: 10.5067/DL081SQY7C89

983 Jaross, G. (2017b). *OMPS-NPP L2 NM Ozone (O3) Total Column swath orbital V2*
984 [dataset]. Greenbelt, MD, USA: Goddard Earth Sciences Data and Information
985 Services Center (GES DISC). doi: 10.5067/0WF4HAAZ0VHK

986 Jaross, G., Bhartia, P. K., Chen, G., Kowitt, M., Haken, M., Chen, Z., ... Kelly,
987 T. (2014). OMPS Limb Profiler instrument performance assessment.
988 *Journal of Geophysical Research: Atmospheres*, *119*(7), 4399-4412. doi:
989 10.1002/2013JD020482

990 Jin, X., Fiore, A. M., Murray, L. T., Valin, L. C., Lamsal, L. N., Duncan, B., ...
991 Tonnesen, G. S. (2017). Evaluating a space-based indicator of surface
992 ozone-NO_x-VOC sensitivity over midlatitude source regions and application
993 to decadal trends. *Journal of Geophysical Research: Atmospheres*, *122*(19),
994 10,439-10,461. doi: 10.1002/2017JD026720

995 Joiner, J. (2006). *OMI/Aura Cloud Pressure and Fraction (Raman Scattering)*

1-Orbit L2 Swath 13x24 km V003 [dataset]. Greenbelt, MD, USA: Goddard Earth Sciences Data and Information Services Center (GES DISC). doi: 10.5067/Aura/OMI/DATA2010

Joiner, J. (2020). *OMPS-NPP L2 NM Cloud Pressure and Fraction swath orbital V2* [dataset]. Greenbelt, MD, USA: Goddard Earth Sciences Data and Information Services Center (GES DISC). doi: 10.5067/CJAALTQUCLO2

Joiner, J., Vasilkov, A. P., Gupta, P., Bhartia, P. K., Veefkind, P., Sneep, M., ... Spurr, R. (2012). Fast simulators for satellite cloud optical centroid pressure retrievals; evaluation of OMI cloud retrievals. *Atmospheric Measurement Techniques*, *5*(3), 529–545. doi: 10.5194/amt-5-529-2012

Jung, Y., González Abad, G., Nowlan, C. R., Chance, K., Liu, X., Torres, O., & Ahn, C. (2019). Explicit aerosol correction of OMI formaldehyde retrievals. *Earth and Space Science*, *6*(11), 2087-2105. doi: 10.1029/2019EA000702

Kaiser, J., Jacob, D. J., Zhu, L., Travis, K. R., Fisher, J. A., González Abad, G., ... Wisthaler, A. (2018). High-resolution inversion of OMI formaldehyde columns to quantify isoprene emission on ecosystem-relevant scales: application to the southeast US. *Atmospheric Chemistry and Physics*, *18*(8), 5483–5497. doi: 10.5194/acp-18-5483-2018

Kim, J., Jeong, U., Ahn, M.-H., Kim, J. H., Park, R. J., Lee, H., ... Choi, Y. (2020). New Era of Air Quality Monitoring from Space: Geostationary Environment Monitoring Spectrometer (GEMS). *Bulletin of the American Meteorological Society*, *101*(1), E1-E22. doi: 10.1175/BAMS-D-18-0013.1

Kleipool, Q. L., Dobber, M. R., de Haan, J. F., & Levelt, P. F. (2008). Earth surface reflectance climatology from 3 years of OMI data. *Journal of Geophysical Research: Atmospheres*, *113*(D18). (D18308) doi: 10.1029/2008JD010290

Kwon, H.-A., González Abad, G., Nowlan, C. R., Chong, H., Souri, A. H., Vigouroux, C., ... Smale, D. (2022). Validation of OMPS Suomi NPP and OMPS NOAA-20 formaldehyde vertical columns with Fourier-transform infrared observations.
(To be submitted to Journal of Geophysical Research: Atmospheres)

Kwon, H.-A., Park, R. J., González Abad, G., Chance, K., Kurosu, T. P., Kim, J., ... Burrows, J. (2019). Description of a formaldehyde retrieval algorithm for the Geostationary Environment Monitoring Spectrometer

(GEMS). *Atmospheric Measurement Techniques*, *12*(7), 3551–3571. doi: 10.5194/amt-12-3551-2019

Laughner, J. L., Zhu, Q., & Cohen, R. C. (2019). Evaluation of version 3.0B of the BEHR OMI NO_2 product. *Atmospheric Measurement Techniques*, *12*(1), 129–146. doi: 10.5194/amt-12-129-2019

Li, C., Joiner, J., Krotkov, N. A., & Dunlap, L. (2015). A new method for global retrievals of HCHO total columns from the Suomi National Polar-orbiting Partnership Ozone Mapping and Profiler Suite. *Geophysical Research Letters*, *42*(7), 2515-2522. doi: 10.1002/2015GL063204

Li, M., Zhang, Q., Kurokawa, J.-I., Woo, J.-H., He, K., Lu, Z., ... Zheng, B. (2017). MIX: a mosaic Asian anthropogenic emission inventory under the international collaboration framework of the MICS-Asia and HTAP. *Atmospheric Chemistry and Physics*, *17*(2), 935–963. doi: 10.5194/acp-17-935-2017

Lorente, A., Folkert Boersma, K., Yu, H., Dörner, S., Hilboll, A., Richter, A., ... Krol, M. (2017). Structural uncertainty in air mass factor calculation for NO_2 and HCHO satellite retrievals. *Atmospheric Measurement Techniques*, *10*(3), 759–782. doi: 10.5194/amt-10-759-2017

Loyola, D. G., Gimeno García, S., Lutz, R., Argyrouli, A., Romahn, F., Spurr, R. J. D., ... Schüssler, O. (2018). The operational cloud retrieval algorithms from TROPOMI on board Sentinel-5 Precursor. *Atmospheric Measurement Techniques*, *11*(1), 409–427. doi: 10.5194/amt-11-409-2018

Marais, E. A., Jacob, D. J., Jimenez, J. L., Campuzano-Jost, P., Day, D. A., Hu, W., ... McNeill, V. F. (2016). Aqueous-phase mechanism for secondary organic aerosol formation from isoprene: application to the southeast United States and co-benefit of SO_2 emission controls. *Atmospheric Chemistry and Physics*, *16*(3), 1603–1618. doi: 10.5194/acp-16-1603-2016

Marais, E. A., Jacob, D. J., Kurosu, T. P., Chance, K., Murphy, J. G., Reeves, C., ... Mao, J. (2012). Isoprene emissions in Africa inferred from OMI observations of formaldehyde columns. *Atmospheric Chemistry and Physics*, *12*(14), 6219–6235. doi: 10.5194/acp-12-6219-2012

Martin, R. V., Chance, K., Jacob, D. J., Kurosu, T. P., Spurr, R. J. D., Bucsela, E., ... Koelemeijer, R. B. A. (2002). An improved retrieval of tropospheric nitrogen dioxide from GOME. *Journal of Geophysical Research: Atmospheres*,

107(D20), ACH 9-1-ACH 9-21. doi: 10.1029/2001JD001027

Millet, D. B., Jacob, D. J., Boersma, K. F., Fu, T.-M., Kurosu, T. P., Chance, K., ... Guenther, A. (2008). Spatial distribution of isoprene emissions from North America derived from formaldehyde column measurements by the OMI satellite sensor. *Journal of Geophysical Research: Atmospheres*, *113*(D2). doi: 10.1029/2007JD008950

Noël, S., Bramstedt, K., Bovensmann, H., Gerilowski, K., Burrows, J. P., Standfuss, C., ... Veihelmann, B. (2012). Quantification and mitigation of the impact of scene inhomogeneity on Sentinel-4 UVN UV-VIS retrievals. *Atmospheric Measurement Techniques*, *5*(6), 1319–1331. doi: 10.5194/amt-5-1319-2012

Nowlan, C. R., & González Abad, G. (2022). *README Document for OMPS_NPP_NMHCHO_L2 and OMPS_N20_NMHCHO_L2.* Greenbelt, MD, USA: Goddard Earth Sciences Data and Information Services Center (GES DISC).

Nowlan, C. R., Liu, X., Janz, S. J., Kowalewski, M. G., Chance, K., Follette-Cook, M. B., ... Weinheimer, A. J. (2018). Nitrogen dioxide and formaldehyde measurements from the GEOstationary Coastal and Air Pollution Events (GEO-CAPE) Airborne Simulator over Houston, Texas. *Atmospheric Measurement Techniques*, *11*(11), 5941–5964. doi: 10.5194/amt-11-5941-2018

Palmer, P. I., Abbot, D. S., Fu, T.-M., Jacob, D. J., Chance, K., Kurosu, T. P., ... Sumner, A. L. (2006). Quantifying the seasonal and interannual variability of North American isoprene emissions using satellite observations of the formaldehyde column. *Journal of Geophysical Research: Atmospheres*, *111*(D12). doi: 10.1029/2005JD006689

Palmer, P. I., Jacob, D. J., Chance, K., Martin, R. V., Spurr, R. J., Kurosu, T. P., ... Li, Q. (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. doi: 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. doi: 10.1002/2017EA000336

Qin, K., Han, X., Li, D., Xu, J., Loyola, D., Xue, Y., ... Yuan, L. (2020). Satellite-based estimation of surface NO_2 concentrations over east-central China: A comparison of POMINO and OMNO2d data. *Atmospheric Environment*, *224*, 117322. doi: 10.1016/j.atmosenv.2020.117322

Schaaf, C., & Wang, Z. (2015). *MCD43C1 MODIS/Terra+Aqua BRDF/AlbedoModel Parameters Daily L3 Global 0.05Deg CMG V006* [dataset]. NASA EOSDIS Land Processes DAAC. doi: 10.5067/MODIS/MCD43C1.006

Seftor, C. J., Jaross, G., Kowitt, M., Haken, M., Li, J., & Flynn, L. E. (2014). Post-launch 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. doi: 10.1002/2013JD020472

Serdyuchenko, A., Gorshelev, V., Weber, M., Chehade, W., & Burrows, J. P. (2014). High spectral resolution ozone absorption cross-sections – part 2: Temperature dependence. *Atmospheric Measurement Techniques*, *7*(2), 625–636. doi: 10.5194/amt-7-625-2014

Sofieva, V. F., Kyrölä, E., Laine, M., Tamminen, J., Degenstein, D., Bourassa, A., ... Bhartia, P. K. (2017). Merged SAGE II, Ozone_cci and OMPS ozone profile dataset and evaluation of ozone trends in the stratosphere. *Atmospheric Chemistry and Physics*, *17*(20), 12533–12552. doi: 10.5194/acp-17-12533-2017

Souri, A. H., Nowlan, C. R., González Abad, G., Zhu, L., Blake, D. R., Fried, A., ... Chance, K. (2020). An inversion of NO_x and non-methane volatile organic compound NMVOC emissions using satellite observations during the KORUS-AQ campaign and implications for surface ozone over East Asia. *Atmospheric Chemistry and Physics*, *20*(16), 9837–9854. doi: 10.5194/acp-20-9837-2020

Spurr, R. (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. doi: 10.1016/j.jqsrt.2006.05.005

Spurr, R. (2008). LIDORT and VLIDORT: Linearized pseudo-spherical scalar and vector discrete ordinate radiative transfer models for use in remote sensing retrieval problems. In *Light Scattering Reviews 3: Light Scattering and Reflection* (pp. 229–275). Berlin, Heidelberg: Springer Berlin Heidelberg. doi:

10.1007/978-3-540-48546-9_7

Stavrakou, T., Müller, J.-F., De Smedt, I., Van Roozendael, M., van der Werf, G. R., Giglio, L., & Guenther, A. (2009). Global emissions of non-methane hydrocarbons deduced from SCIAMACHY formaldehyde columns through 2003–2006. *Atmospheric Chemistry and Physics*, *9*(11), 3663–3679. doi: 10.5194/acp-9-3663-2009

Su, W., Liu, C., Hu, Q., Zhang, C., Liu, H., Xia, C., ... Chen, Y. (2022). First global observation of tropospheric formaldehyde from Chinese GaoFen-5 satellite: Locating source of volatile organic compounds. *Environmental Pollution*, *297*, 118691. doi: 10.1016/j.envpol.2021.118691

Su, W., Liu, C., Hu, Q., Zhao, S., Sun, Y., Wang, W., ... Kim, J. (2019). Primary and secondary sources of ambient formaldehyde in the Yangtze River Delta based on Ozone Mapping and Profiler Suite (OMPS) observations. *Atmospheric Chemistry and Physics*, *19*(10), 6717–6736. doi: 10.5194/acp-19-6717-2019

Sun, K., Liu, X., Huang, G., González Abad, G., Cai, Z., Chance, K., & Yang, K. (2017). Deriving the slit functions from OMI solar observations and its implications for ozone-profile retrieval. *Atmospheric Measurement Techniques*, *10*(10), 3677–3695. doi: 10.5194/amt-10-3677-2017

Sun, K., Zhu, L., Cady-Pereira, K., Chan Miller, C., Chance, K., Clarisse, L., ... Zondlo, M. (2018). A physics-based approach to oversample multi-satellite, multispecies observations to a common grid. *Atmospheric Measurement Techniques*, *11*(12), 6679–6701. doi: 10.5194/amt-11-6679-2018

Thomas, W., Hegels, E., Slijkhuis, S., Spurr, R., & Chance, K. (1998). Detection of biomass burning combustion products in Southeast Asia from backscatter data taken by the GOME spectrometer. *Geophysical Research Letters*, *25*(9), 1317-1320. doi: 10.1029/98GL01087

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. doi: 10.1002/2016JD025940

US National Ice Center. (2008). *IMS Daily Northern Hemisphere Snow and Ice Analysis at 1 km, 4 km, and 24 km Resolutions* [dataset]. Boulder, Colorado

USA. (Version 1, updated daily) doi: 10.7265/N52R3PMC

Valin, L. C., Fiore, A. M., Chance, K., & González Abad, G. (2016). The role of OH production in interpreting the variability of CH_2O columns in the southeast U.S. *Journal of Geophysical Research: Atmospheres*, *121*(1), 478-493. doi: 10.1002/2015JD024012

Vandaele, A. C., Hermans, C., Simon, P. C., Carleer, M., Colin, R., Fally, S., ... Coquart, B. (1998). Measurements of the NO_2 absorption cross-section from 42,000 cm^{-1} to 10,000 cm^{-1} (238-1000 nm) at 220 K and 294 K. *Journal of Quantitative Spectroscopy and Radiative Transfer*, *59*, 171-184. doi: 10.1016/S0022-4073(97)00168-4

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. doi: 10.5194/amt-7-2897-2014

Veefkind, J. P., Boersma, K. F., Wang, J., Kurosu, T. P., Krotkov, N., Chance, K., & Levelt, P. F. (2011). Global satellite analysis of the relation between aerosols and short-lived trace gases. *Atmospheric Chemistry and Physics*, *11*(3), 1255–1267. doi: 10.5194/acp-11-1255-2011

Vigouroux, C., Bauer Aquino, C. A., Bauwens, M., Becker, C., Blumenstock, T., De Mazière, M., ... Toon, G. (2018). NDACC harmonized formaldehyde time series from 21 FTIR stations covering a wide range of column abundances. *Atmospheric Measurement Techniques*, *11*(9), 5049–5073. doi: 10.5194/amt-11-5049-2018

Vigouroux, C., Langerock, B., Bauer Aquino, C. A., Blumenstock, T., Cheng, Z., De Mazière, M., ... Winkler, H. (2020). TROPOMI–Sentinel-5 Precursor formaldehyde validation using an extensive network of ground-based Fourier-transform infrared stations. *Atmospheric Measurement Techniques*, *13*(7), 3751–3767. doi: 10.5194/amt-13-3751-2020

Voors, R., Dobber, M., Dirksen, R., & Levelt, P. (2006). Method of calibration to correct for cloud-induced wavelength shifts in the Aura satellite's Ozone Monitoring Instrument. *Applied Optics*, *45*(15), 3652–3658. doi: 10.1364/AO.45.003652

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*, 50-64. doi: 10.1016/j.rse.2018.02.001

Williams, J. E., Boersma, K. F., Le Sager, P., & Verstraeten, W. W. (2017). The high-resolution version of TM5-MP for optimized satellite retrievals: description and validation. *Geoscientific Model Development*, *10*(2), 721–750. doi: 10.5194/gmd-10-721-2017

Wilmouth, D. M., Hanisco, T. F., Donahue, N. M., & Anderson, J. G. (1999). Fourier transform ultraviolet spectroscopy of the $A^2\Pi_{3/2} \leftarrow X^2\Pi_{3/2}$ transition of BrO. *The Journal of Physical Chemistry A*, *103*(45), 8935-8945. doi: 10.1021/jp991651o

Wolfe, G. M., Nicely, J. M., St. Clair, J. M., Hanisco, T. F., Liao, J., Oman, L. D., ... Dean-Day, J. (2019). Mapping hydroxyl variability throughout the global remote troposphere via synthesis of airborne and satellite formaldehyde observations. *Proceedings of the National Academy of Sciences*, *116*(23), 11171–11180. doi: 10.1073/pnas.1821661116

Zhou, Y., Brunner, D., Boersma, K. F., Dirksen, R., & Wang, P. (2009). An improved tropospheric NO_2 retrieval for OMI observations in the vicinity of mountainous terrain. *Atmospheric Measurement Techniques*, *2*(2), 401–416. doi: 10.5194/amt-2-401-2009

Zhu, L., González Abad, G., Nowlan, C. R., Chan Miller, C., Chance, K., Apel, E. C., ... Wolfe, G. M. (2020). Validation of satellite formaldehyde (HCHO) retrievals using observations from 12 aircraft campaigns. *Atmospheric Chemistry and Physics*, *20*(20), 12329–12345. doi: 10.5194/acp-20-12329-2020

Zhu, L., Jacob, D. J., Keutsch, F. N., Mickley, L. J., Scheffe, R., Strum, M., ... Shah, V. (2017). Formaldehyde (HCHO) as a hazardous air pollutant: mapping surface air concentrations from satellite and inferring cancer risks in the United States. *Environmental Science & Technology*, *51*(10), 5650-5657. doi: 10.1021/acs.est.7b01356

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. doi:

10.1016/j.jqsrt.2016.04.003

Zoogman, P., Liu, X., Suleiman, R., Pennington, W., Flittner, D., Al-Saadi, J., ... Chance, K. (2017). Tropospheric emissions: Monitoring of pollution (TEMPO). *Journal of Quantitative Spectroscopy and Radiative Transfer*, *186*, 17-39. doi: 10.1016/j.jqsrt.2016.05.008

Earth and Space Science

Supporting Information for

Global Formaldehyde Products from the Ozone Mapping and Profiler Suite (OMPS) Nadir Mappers on Suomi NPP and NOAA-20

Caroline R. Nowlan[1], Gonzalo González Abad[1], Hyeong-Ahn Kwon, Zolal Ayazpour[2], Christopher Chan Miller[1,3], Kelly Chance[1], Heesung Chong[1], Xiong Liu[1], Ewan O'Sullivan[1], Huiqun Wang[1], Lei Zhu[4], Isabelle De Smedt[5], Glen Jaross[6], Colin Seftor[6,7], Kang Sun[2]

[1]Center for Astrophysics | Harvard & Smithsonian, Cambridge, MA, USA

[2]University at Buffalo, Buffalo, NY, USA

[3]Harvard University, Cambridge, MA, USA

[4]Southern University of Science and Technology, Shenzhen, China

[5]Royal Belgian Institute for Space Aeronomy, Brussels, Belgium

[6]NASA Goddard Space Flight Center, Greenbelt, MD, USA

[7]Science Systems and Applications, Inc., Lanham, MD, USA

Contents of this file

Introduction

This document contains supporting information for the article ``Global Formaldehyde Products from the Ozone Mapping and Profiler Suite (OMPS) Instruments on Suomi NPP and NOAA-20", which describes the formaldehyde retrieval algorithm and publicly-available data products for the OMPS nadir mapper instruments on the Suomi-NPP and NOAA-20 satellites.

The Level 2 OMPS HCHO (OMPS_NPP_NMHCHO_L2 and OMPS_N20_NMHCHO_L2) files consist of orbital swath total vertical column densities of formaldehyde from the OMPS nadir mappers. 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. Each file contains Level 2 swath data for a single orbit. There are typically 14 to 15 orbits per day.

The OMPS_NPP_NMHCHO_L2 and OMPS_N20_NMHCHO_L2 files are in netCDF (version 4) format. The file 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, 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.

Table S1 summarizes the main possible dimensions of data fields. Tables S2 – S6 list the data fields included in the OMPS Level 2 HCHO files by group. Each variables also includes attributes such as units and flag meanings. The data field `support_data/surface_pressure` includes the Eta coordinates for calculating vertical pressure grids. See the OMPS HCHO README document (Nowlan and González Abad, 2022) for further details on the file format.

Table S1. Main data field dimensions. Not all data fields have all dimensions.

Name	Description	Nominal Dimensions
`along_track`	Number of ground pixels along the satellite track	400 (SNPP) 1201 (NOAA-20)
`cross_track`	Number of ground pixels across the satellite track	36 (SNPP) 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

Table S2. Data fields in `key_science_data` group.

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 S3. Data fields in `geolocation` group.

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	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	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 S4. Data fields in `qa_statistics` group.

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 S5. Data fields in `support_data` group.

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 S6. Data fields in `uncertainty_budget` group.

Data Field Name	Description	Type	Dimensions	Units
`amf_total_uncert`	Total AMF uncertainty	32-bit floating-point	`along_track, cross_track`	%
`bias_uncertainty`	Estimated uncertainty in bias correction	32-bit floating-point	`along_track, cross_track`	molecules/cm^2
`ref_sector_uncertainty`	Estimated uncertainty in reference background correction	32-bit floating-point	`along_track, cross_track`	molecules/cm^2