OMNO2 README Document
Data Product Version 4.0

The OMI Nitrogen Dioxide Algorithm Team *

December 2019
Document Version 9.0

Abstract

This README file describes the version 4.0 release of the OMI NO_2 Standard Product, OMNO2, the version 4.0 release of the OMI NO_2 gridded Level-2 (OMNO2G), and the version 4.0 release of the gridded OMNO2d product produced from it.

Species	Nitrogen Dioxide (NO_2)
Data Version	Standard Product, v4.0
Version Release Date	December 2019
Retrieved Quantities	Total slant column density Total vertical column density Stratospheric column density Tropospheric column density
Spatial Resolution	13 km x 24 km (at nadir)
Global Coverage	Approximately daily
Date Range	2004/10/01–Present
Data Screening	See data quality flags in L2 data files
Data Location	https://disc.gsfc.nasa.gov/datasets/OMNO2_V003/summary
Point of Contact	Nickolay A. Krotkov NASA/Goddard Space Flight Center nickolay.a.krotkov@nasa.gov 301-614-5553

---

[*] The OMI Nitrogen Dioxide Algorithm Team members are: Dr. Nickolay A. Krotkov, Dr. Lok N. Lamsal, Dr. Sergey V. Marchenko, Dr. William H. Swartz

Contents

1 Executive Summary 5

1.1 Improvements 5

2 Introduction 9

2.1 OMI Observations 9
2.2 Spatial coverage of OMI 9
2.3 Temporal coverage of OMI 11
2.4 Row anomaly 13
2.5 Zoom mode 13

3 NO_2 algorithm description 15

3.1 Solar spectral irradiance 17
3.2 Slant column retrieval 18
3.3 AMF calculation 18
3.4 Destriping 20
3.5 Stratosphere-troposphere separation 20
3.6 NO_2 data quality 21

4 Level-2 Data Product 22

4.1 File name 22
4.2 File organization 22
4.3 Data description 23
4.4 Limitations 25

5 The Level-2 gridded NO_2 product, OMNO2G 27

5.1 File name 27
5.2 File structure 27
5.3 Data fields 27
5.4 Limitations 28

6 The Level-3 gridded NO_2 product, OMNO2d 30

6.1 File name 30
6.2 File structure 30
6.3 Data fields 31
6.4 Limitations 32

7 Software versions 33

8 Data quality 34

9 Product availability 35

10 Reporting problems and requesting information 37

A Time calculations 40

List of Tables

2 Table of abbreviations, acronyms, and initializations (AAI) used in this document. 7

4 Criteria used to screen OMNO2 data for use in generating OMNO2d data product. 31

5 Version numbers of data products and the applications (Apps) that create them. 33

6 Fields available in the three OMI NO_2 data sets. D indicates fields in the *Data Fields* group; G indicates fields in the *Geolocation Fields* group. 38

1 Executive Summary

The OMI NO_2 algorithm retrieves estimated columns (total, tropospheric, and stratospheric) of nitrogen dioxide from OMI Level-1B calibrated radiance and irradiance data. The current version, v4.0, improves on the retrievals in the prior versions in a number of significant ways. The OMNO2 algorithm is designed to infer as much information as possible about atmospheric NO_2 from the OMI measurements, with the minimum possible dependence on model simulations.

Since its first release in 2006 [Bucsela et al., 2006, Celarier et al., 2008], research at NASA and other institutions has led to significant conceptual and technical improvements in the retrieval of NO_2 from space-based measurements, which have guided the development of the current version. Version 2.1 OMNO2 developed a new scheme for seperating stratospheric and tropospheric components [Bucsela et al., 2013, Lamsal et al., 2014], version 3.0 represented a significant advance in NO_2 slant column density (SCD) retrievals [Marchenko et al., 2015, Krotkov et al., 2017, Choi et al., 2019], and the current version includes several changes for improved NO_2 air mass factor (AMF) and vertical column density (VCD) calculations [Lamsal et al., in preparation].

1.1 Improvements

The principal improvements in v4.0 include:

1. Use of a new daily and OMI field of view (FOV) specific geometry-dependent surface Lambertian Equivalent Reflectivity (GLER) product in both NO_2 and cloud retrievals. The GLER data are derived by coupling the atmosphere with the MODIS surface bidirectional reflectance distribution function (BRDF) data for land [Vasilkov et al., 2017, Qin et al., 2019], and an observationally-constrained (VLIDORT) model of reflection and water-leaving radiance for water surfaces [Vasilkov et al., 2017, Fasnacht et al., 2019];

2. Use of improved cloud parameters (effective cloud fraction (CF) and optical centroid pressure (OCP)) from a new cloud algorithm (OMCDO2N). These parameters are retrieved consistently as NO_2 using a new algorithm for O_2-O_2 SCD data, and GLER product for terrain reflectivity [Vasilkov et al., 2018];

3. Use of a more accurate terrain pressure calculated using OMI ground pixel-averaged terrain height and monthly mean GMI terrain pressure.

The terrain height information comes from the Digital elevation model (DEM) data at 2 arcmin resolution;

4. Improved treatment over snow/ice surfaces by using the concept of scene LER and scene pressure.

The following improvements made in previous versions remain unchanged.

5. An improved DOAS algorithm for retrieving slant column densities [Marchenko et al., 2015] for NO_2. The key features of the algorithm include independent, very accurate registration of wavelength scales between radiance and irradiance spectra, iterative subtraction of rotational Raman scattering effect signal, sequential retrieval of SCD of NO_2 and interferring species (H_2O and CHOCHO), and use of stable monthly average solar irradiances derived from OMI measurements;
6. Use of improved higher resolution 1 deg (latitude) x 1.25 deg (longitude) GMI model-based monthly *a priori* NO_2 profile shapes [Douglass et al., 2004] with year-specific emissions;
7. Improved estimation of NO_2 SCD uncertainties using the curvature of the chi-squared surface around the retrieved point;
8. Reduced striping in NO_2 SCDs due to the use of alternatively processed solar data (OML1BIR2) instead of OML1BIRR;
9. An observation-based stratosphere-troposphere separation scheme to estimate stratospheric NO_2 field by spatial interpolation using retrieved SCD data over unpolluted or cloudy areas;
10. Improved de-striping approach to remove cross-track artifacts using data from unpolluted areas in tropics.

The listed items 5–8 were part of v3.0 and the items 9-10 were part of v2.1.

Table 2: Table of abbreviations, acronyms, and initializations (AAI) used in this document.

Abbr.	Meaning
AAI	Abbreviations, acronyms, and initializations
AMF	Air mass factor
AMSR-E	Advanced Microwave Scanning Radiometer–Earth Observing System
APP	Application (production software unit)
AVDC	Aura Validation Data Center
BIRA-IASB	The Royal Belgian Institute for Space Aeronomy
BRDF	Bidirectional reflectance distribution function
CRF	Cloud Radiance Fraction
CTM	Chemistry and Transport Model
DEM	Digital elevation model
DOAS	Differential Optical Absorption Spectroscopy
ECF	Effective cloud fraction
EOS	Earth Observing System
FoV	Field of view
FTIR	Fourier Transform Infrared
GES-DISC	Goddard Earth Sciences Data and Information Services Center
GEOS-5	Goddard Earth Observing System, Version 5
GLER	Geometry-dependent surface Lambertian Equivalent Reflectivity
GMI	Global Modeling Initiative
GSFC	Goddard Space Flight Center
HDF-EOS	HDF EOS data file format
IUP	Institute of Environmental Physics, University of Bremen
KNMI	Koningklijk Nederlands Meteorologisch Instituut
LECT	Local equator crossing time
L-1B	Level-1B (data product with calibrated radiances or irradiances)
L-2	Level-2 (data product with retrieved geophysical values)
L-2G	Gridded Level-2 (data product in grid format)
L-3	Level-3 gridded data product

LUT	Look-up table
MODIS	Moderate Resolution Imaging Spectroradiometer
NASA	National Aeronautics and Space Administration
NISE	Near-real-time Ice and Snow Extent
OMI	Ozone Monitoring Instrument
OML1BIRR	OMI Level-1B Solar Irradiances
OML1BIR2	Alternatively processed OMI Level-1B Solar Irradiances
OCP	Optical centroid pressure
RA	Row anomaly
SCD	Slant column density
SIPS	Science Investigator Processing System
SSMIS	Special Microwave Imager–Sounder
SW	Scattering weight
CCD	Charge-coupled device
VCD	Vertical column denisty
VIS	OMI visible-wavelength detector
VLIDORT	Vector Linearized Discrete Ordinate Radiative Transfer model

2 Introduction

Nitrogen oxides (NO_x = NO + NO_2) are species that play key roles in tropospheric and stratospheric ozone chemistry. Further, high surface level NO_2 is itself recognized to be deleterious to human health. Major sources of tropospheric NO_x include combustion, soil emissions, and lightning. Nitric oxide (NO) and nitrogen dioxide (NO_2) are in quasi-steady-state in the atmosphere, and their relative concentrations depend on emissions, solar illumination, and the concentrations of other chemical species.

2.1 OMI Observations

OMI was launched on July 15, 2004, on the NASA EOS Aura satellite, which is in a sun-synchronous ascending polar orbit with a local equator crossing time (LECT) of 13:45±0:15. Science-quality data operations began on October 1, 2004.

OMI makes simultaneous measurements in a swath of width 2600 km, divided into 60 fields of view (FoVs).[1]

The FoVs vary in size from $\sim$ 13km $\times$ 24km near nadir to $\sim$ 24km $\times$ 160km at the outermost FoVs. Figure 1 shows the relation between the OMI instrument and its viewing geometry.

2.2 Spatial coverage of OMI

Figure 2 shows the positions and sizes of these FoVs relative to the flight direction. One swath is measured every two seconds. Due to the optical characteristics of the instrument, adjacent swaths overlap considerably in their ground coverage. The width of a swath ensures that swaths from adjacent consecutive orbits are nearly contiguous at the equator and have some overlap at mid- to high-latitudes. In a single orbit, OMI measures approximately 1650 swaths from terminator to terminator. With an orbital

---

[1] There is a diversity of terminologies that are used within the EOS and OMI communities to describe an instrument's viewing area. In this document, we will use the following convention: Each measurement is made within a FoV. A set of FoVs that are measured at the same time (transverse to the orbital track–see Figure 2) is a *swath.* The set of all measurements made during a single day-side passage of the OMI instrument is a *granule.*

Some naming conventions, sometimes conflicting with ours and with each other, is, unavoidably, inherited into the NO_2 data products from various other data products and file structure specifications. Synonyms for swath include: *exposure, scan, scanline, and iTime.* Synonyms for FoV include: *pixel, groundpixel, scene, scan position, cross-track position, iXtrack, and row.* The term *orbit* refers to a granule. *Swath*, in the HDF-EOS5 convention, refers to a granule.

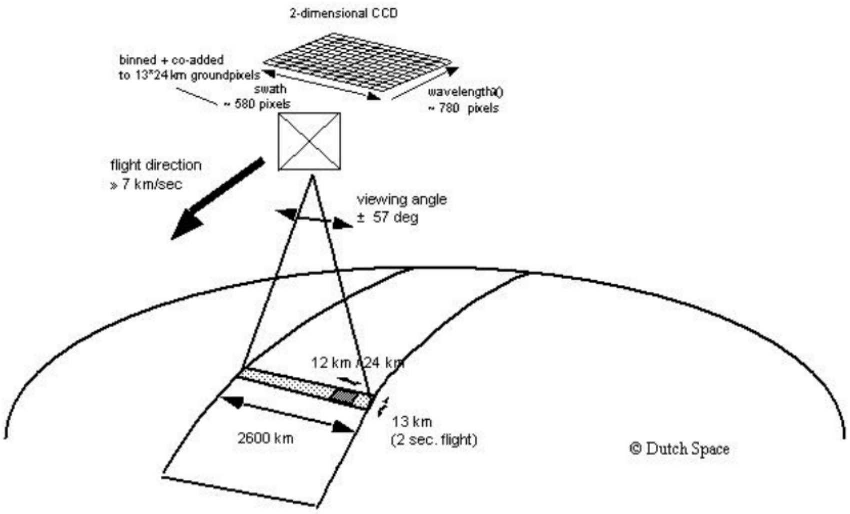


Figure 1: The OMI instrument and its viewing geometry.

period of 99 minutes, OMI views the entire sunlit portion of the Earth in 14–15 orbits.

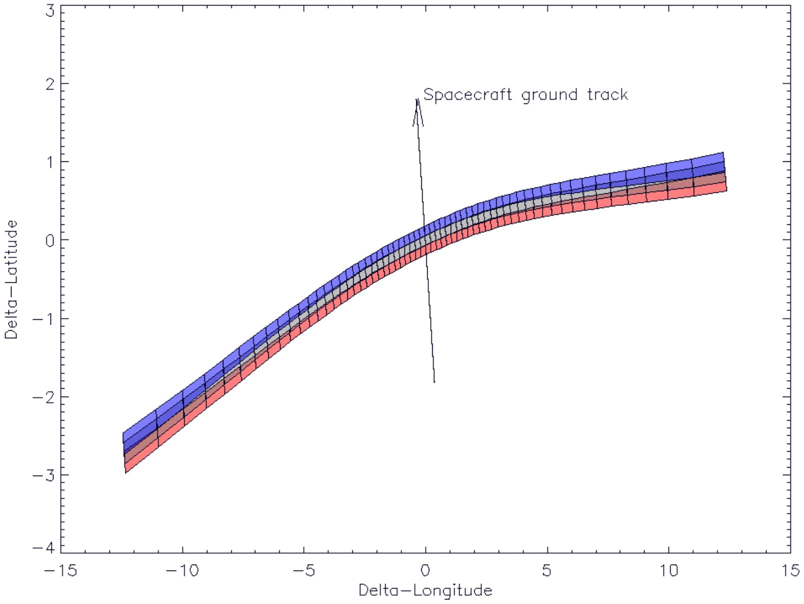


Figure 2: Geometry of the OMI FoVs near the equator. The vertical scale (Delta_latitude, degree) is exaggerated by a factor of ~ 3, relative to the horizontal scale (Delta_longitude, degree). The colored tiles show the geometries of three successive swaths (exposures) of OMI FoVs (red, gray, and blue) near the equator. In the middle of the swath, there is very little overlap between consecutive FoVs. There is greater overlap for FoVs closer to the edge of the swath. In the outermost positions, the areal overlap between consecutive FoVs is nearly 50%.

2.3 Temporal coverage of OMI

For any position on the Earth, the OMI measurement time is generally not equal to the LECT. For near-nadir FoVs, the local overpass time is generally earlier than the LECT in the Northern Hemisphere, and later in the Southern Hemisphere. Around latitudes 50 degrees South and North, the local time of observation for near-nadir FoVs is about 1 hour later or earlier, respectively. The difference is larger for off-nadir FoVs. In a swath the observational time is earlier for western FoVs and later for eastern FoVs (Fig. 3). Appendix A describes how to calculate local times for OMI observations.

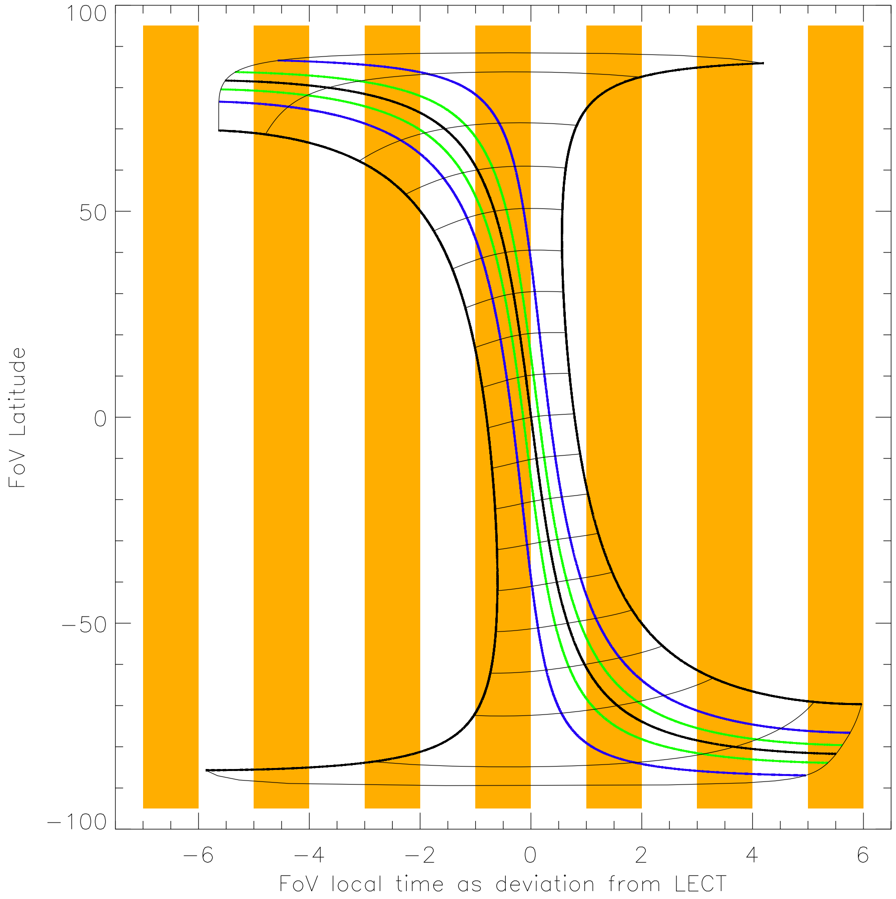


Figure 3: Deviation of local solar time (LST, "sundial time") from local equator crossing time (LECT) along the ascending leg of the OMI orbit. The local time at the subsatellite point (near scan positions 29 and 30) is shown as the thick black curve in the middle. Curves are also shown for for scan positions 0 and 59 (black), 10 and 49 (blue), and 20 and 39 (green). The nearly horizontal curves connect the FOV centers that are measured simultaneously, in a single swath. These are chosen as the times when the subsatellite point crosses latitudes $-80, -70, \cdots, +70, +80$. The slight asymmetry of the figure, between the southern and northern hemispheres is mainly due to the fact that OMI's FOVs are not arrayed symmetrically about the flight vector.

2.4 Row anomaly

Starting June 25, 2007 (presumably, even earlier [Schenkeveld et al., 2017]) , an anomaly began to appear in the L-1B radiances, attenuating the measured radiances in certain FoVs (53 and 54; numbered from 0). Then, on May 11, 2008 cross-track positions 37–42 started to be affected toward the northern end of the OMI orbit. The anomaly has developed and changed over the period since. This phenomenon has been named the "row anomaly" (RA) referring to affected FoVs (rows) of the CCD detector.

Four distinct effects on the OMI radiance specta have been identified: (1) Blockage–a decrease in radiance level. It is currently assumed that this is caused by a partial blocking of the OMI viewing port; (2) An increase in solar stray-light, scattered into the Earth-viewing port. It is assumed to be caused by the reflection of sunlight into the viewing port by the material blocking the port; (3) Wavelength shift–The blocking object causes an inhomogeneous illumination of the instrument's spectral slit, which induces a shift in the spectral dispersion; and (4) Radiance received from outside the nominal FoV. These effects are discussed in the web document "Background information about the Row Anomaly in OMI" (`http://projects.knmi.nl/omi/research/product/rowanomaly-background.php`).

The data quality for RA-affected FoVs is sufficiently poor as to prevent the retrieval of NO_2 column amounts. In version 4.0, we have abandoned the retrieval calculations of the VCDs (`ColumnAmountNO2`, *etc.*), and inserted fill-values into those fields where the RA has been identified. We found that the RA-detection algorithm implemented in the L-1B APP sometime fails to flag clearly-affected FoVs adjacent to flagged FoVs. We have implemented additional flagging in the OMNO2 and OMNO2G data products in those cases.

Row anomaly information is available in the `XTrackAnomalyFlags` data field. In the OMNO2 data product, the content of this field reflects our additional flagging, so it is not identical to the so-named field in other data products.

2.5 Zoom mode

The OMI instrument has a number of operating configurations, including *global* measurements (the ordinary measurement), *spatial-zoom*, *spectral-zoom*, and a variety of calibration modes. The mode of operation for any measurement is indicated by the `InstrumentConfigurationID` field. The retrieval code has not been optimized for use with either of the zoom-measurement modes. For that reason, the calculation of the vertical column densities (`ColumnAmountNO2`, for example) has been abandoned when the instrument is not in global measurement mode, and a fill-value will be found in each of these data fields.

There are some instances when a set of 14 or more consecutive orbits have made measurements in zoom mode. Users will find that there are no valid data in the files for these orbits.

3 NO_2 algorithm description

The OMI Level 1B data product contains calibrated earthshine radiance spectra $\mathbf{I}$ for each FOV. Earthshine radiances are divided by the solar irradiance spectrum $\mathbf{F}$ to give a normalized spectrum $\mathbf{R} = \mathbf{I}/\mathbf{F}$. We use the normalized spectrum in the visible (VIS) wavelengths (402–465 nm) to retrieve the trace gas (NO_2, H_2O, CHOCHO) slant column amounts. The slant column represents the integrated abundance of a trace gas along the average photon path from the Sun, through the atmosphere, to the satellite. The slant NO_2 column amounts are then combined with stratospheric and tropospheric air mass factors to obtain vertical column densities (VCDs). The Level 2 (L-2) NO_2 product (OMNO2) includes stratospheric, tropospheric, and total vertical column densities.

Figure 4 shows schematically the data flow through the algorithm. The individual steps are described in more detail in the following subsections.

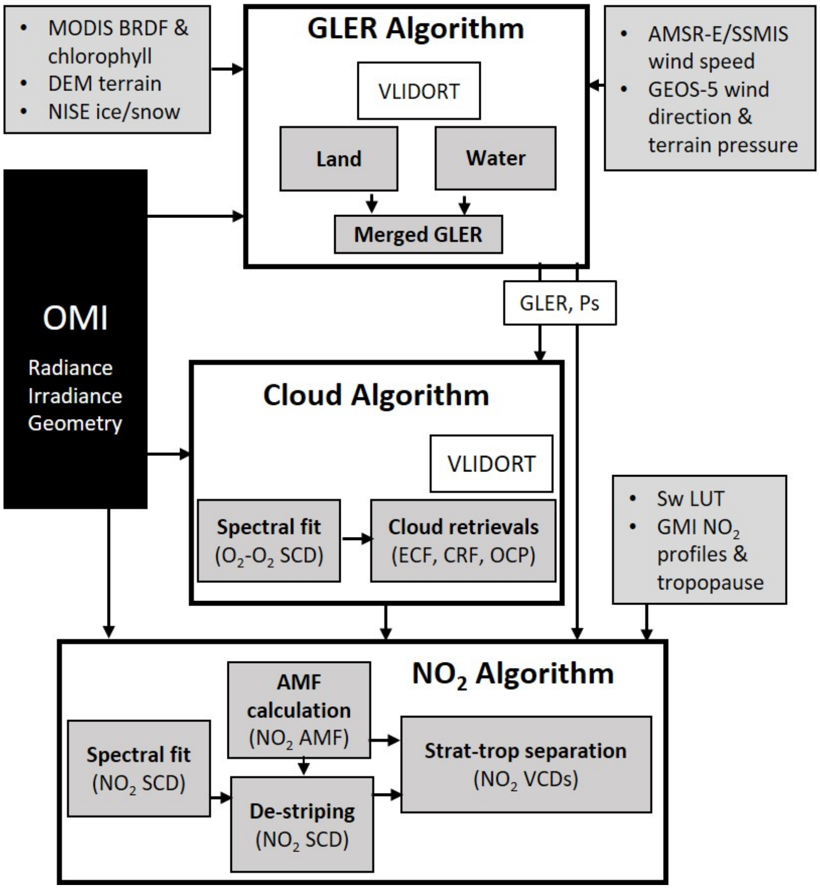


Figure 4: Schematic description of the v4.0 OMI NO_2 processing algorithm. The multi-step procedure consists of 1) a spectral fitting algorithm to calculate NO_2 slant column density SCD ; 2) determination of air mass factor (AMF) to convert SCD to vertical column density (VCD); 3) a new O_2-O_2 cloud algorithm to estimate cloud radiance fraction (CRF) and cloud optical centroid pressure (OCP), both required for the AMF calculation; 4) a new geometry-dependent surface LER (GLER) algorithm needed for both NO_2 and cloud retrievals; 5) GMI-derived inputs (*e.g.,* NO_2 vertical profile shapes) for the AMF calculation; and 6) a stratosphere-troposphere separation scheme to derive tropospheric and stratospheric NO_2 columns. Scattering weight (Sw) is taken from pre-computed look-up-table (LUT) using the radiative transfer program TOMRAD. Inputs to GLER calculation include chlorophyll concentrations from MODIS, the wind speed data from the Advanced Microwave Scanning Radiometer–Earth Observing System (AMSR-E) and the Special Microwave Imager–Sounder (SSMIS) instruments, the wind direction data from the NASA GEOS-5 model.

3.1 Solar spectral irradiance

OMI makes daily solar spectral irradiance measurements that are processed and archived. The original intent was that, in processing an orbital granule into any L-2 product, the most recent solar irradiance spectra, acquired within the preceding 24 hours, would be used. Soon after operations began, pronounced irregular patterns across a swath appeared in retrieved quantities in all OMI L-2 products; these patterns were relatively consistent along-track, and so were dubbed "striping." The stripes changed from orbit to orbit. The use of a static solar reference spectrum ameliorated much of the calibration-induced striping. However, as we have learned more about the sensitivity of OMI to solar variations, and indeed the scale of those variations, we have found it necessary, starting with version 3.0, to use monthly irradiances that are computed as a separate data product. These irradiances are carefully determined, statistically, from a month's irradiance measurements. Lately, the number of missing data in irradiance measurements has grown, affecting the quality of monthly irradiance data. Therefore, since version 3.1, we use monthly irradiance data created from alternative OML1BIR2 solar data. Our analysis revealed that the use of alternative solar data was beneficial to reduce the stripes. Nevertheless, a destriping step is still necessary in the OMNO2 processing.

The OMI NO_2 slant column retrieval algorithm starts from radiances, normalized to monthly-averaged solar irradiances (Section 3.2). We choose the monthly averages as a compromise between the required high S/N of the solar data, and the need to account for long-term instrumental drifts, as well as for the solar variability, which is quite noticeable in the $S/N \gtrsim 1000$ solar spectra [Marcheko and DeLand, 2014].

To produce the solar monthly averages, we correct the daily OMI solar irradiances for the time-varying Sun-Earth distance. At each wavelength of the 402–465 nm domain used in the NO_2 retrievals we assess stability of the wavelength registration, linearly interpolating the spectra with deviations exceeding 0.002 nm to a common wavelength grid. At each wavelength the daily irradiances are searched for $\pm 1.5\sigma$ outliers. These are substituted with a reference spectrum comprised of an unweighted 3-month (January–March 2005) average of measured solar spectra, adjusted for instrumental drifts and solar variability. The wavelength-corrected, cleaned daily spectra are then directly (no weights) averaged into monthly solar means.

3.2 Slant column retrieval

Slant column densities are computed using new algorithms for NO_2 [Marchenko et al., 2015] and O_2-O_2 [Vasilkov et al., 2018] that improve retrievals by increasing the accuracy of the wavelength registration between radiance and irradiance spectra. The wavelength registration procedure relies on the structure of the Ring effect spectrum (rotational Raman scattering), whose amplitude in the spectral region of interest is much greater than spectral features of the trace gases. Wavelength calibrations are performed independently in each of seven spectral 'micro-windows' inside the broad 402–465 nm NO_2 retrieval range. For O_2-O_2, we use five spectral 'micro-windows' over the 463–488 nm range. With this calibration, the dominant Ring feature is subtracted from $\mathbf{R}$, and interfering trace gases (*e.g.,* NO_2, H_2O, and CHOCHO) are successively estimated by fitting Ring-free $\mathbf{R}$ with laboratory-measured cross-sections of trace gases by the Differential Optical Absorption Spectroscopy (DOAS) method. A further step corrects for undersampling by the OMI instrument [Chance et al., 2005], as well as for outlying $\mathbf{R}$ values (mostly caused by cosmic ray events). With the undersampling correction in hand, the entire process of wavelength calibration followed by successive trace gas retrieval is repeated. The result of the spectral fit is a slant column density `SlantColumnAmountNO2` for NO_2.

3.3 AMF calculation

The air mass factor (AMF) is defined as the ratio of the measured slant column density S to the vertical column density V. AMFs depend upon a number of parameters including optical geometry (solar and viewing azimuth and zenith angles), surface reflectivity, cloud pressure and fraction, and the shape of the NO_2 vertical profile. The AMFs are computed from the scattering weights (Section 4.3) and a monthly mean climatology of NO_2 profile shapes constructed from the Global Modeling Initiative (GMI) Chemistry and Transport Model (CTM) simulation, with a horizontal resolution of 1.25° longitude ×1.0° latitude. The simulation is based on yearly-varying emissions, as discussed in Strode et al. [2015]. Model profiles are generated in 15 minute time steps. The profiles used for the AMFs are the averages of the profiles at timesteps from 13:00 to 14:00, local time. Use of monthly NO_2 profile shapes captures the seasonal variation in NO_2 profiles [Lamsal et al., 2010]. The present version (v4.0) uses annual monthly profiles from 2004 to 2015. For dates starting in 2016, the 2015 monthly profiles are used. Using local climatological GEOS-5 monthly temperature profiles, the scattering weights are corrected for the atmospheric temperature profile, which compensates for the fact that the SCD retrieval is done assuming a constant NO_2 temperature (220 K). Stratospheric and tropospheric AMFs are calculated, (A_{strat} and A_{trop}) separated at the climatological GEOS-5 monthly tropopause pressure.

The method of AMF calculation is similar to that described by Palmer et al. [2001]. For each FOV, AMFs are computed for clear ($\text{AMF}_{\text{clear}}$) and cloudy ($\text{AMF}_{\text{cloud}}$) conditions. The AMF of a partially clouded scene is calculated by:

$$\text{AMF} = (1 - f_r) \cdot \text{AMF}_{\text{clear}} + f_r \cdot \text{AMF}_{\text{cloud}} \tag{1}$$

where f_r is the cloud radiance fraction (CRF), *i.e.* the fraction of the measured radiation that comes from clouds and scattering aerosols. The CRF at 440 nm is computed from the effective cloud fraction (ECF) f_c, using tables constructed from VLIDORT model calculations. $\text{AMF}_{\text{clear}}$ is calculated using the new geometry-dependent surface Lambertian Equivalent Reflectivity product (GLER) calculated at 440 nm R_s at pressure P_s. $\text{AMF}_{\text{cloud}}$ is calculated assuming a Lambertian surface of reflectivity 0.8 at cloud optical centroid pressure OCP.

The v4.0 includes major updates on various input parameters used in the AMF calculation. Information on R_s is taken from the Geometry-dependent surface LER (GLER) product (OMGLER) [Qin et al., 2019, Fasnacht et al., 2019]. GLER is derived by inversion of the top-of-atmosphere radiances for a Rayleigh atmosphere simulated by VLIDORT, coupled with anisotropic surface models to account for the solar and viewing geometry dependence of LER for each OMI pixel. A MODIS RossThick Li-Sparse kernel BRDF model is used for land surface, while for water, the Cox-Munk slope distribution is used for surface reflection, and a Case 1 water model simulates light backscattered by water column. Pixel average P_s is computed for each OMI FoV using average terrain height from a digital elevation model (DEM) data at 2 arcmin resolution and GMI monthly surface pressure. P_c and f_c are obtained from a new O_2-O_2 OMCDO2N product with consistent retrieval approach and inputs (*e.g.,* GLER) used in OMNO2 [Vasilkov et al., 2018].

Special attention has been paid to retrievals over snow and ice surfaces. Over ice and snow surfaces, identified by the Near-real-time Ice and Snow Extent (NISE) flags in the OMI Level 1b data, the following treatments are made for surface reflectivity. In case of permanent ice and snow surfaces, the MODIS product provides BRDF parameters, allowing us to calculate GLER. In case of either data gaps or too small GLER values for snow-covered OMI pixels, we use OMI-derived LER but capped by a constant snow albedo of 0.6 following Boersma et al. [2011]. The concept of scene pressure is used for cloud retrievals over seasonal snow covered areas. If the NISE flags are set as true, we use the following algorithm to assign values to our CRF, OCP, and NO_2 retrievals. To avoid a possible NISE misclassification for low reflectivity scenes with scene LER < 0.2, we consider such scenes as being snow/ice free and calculate CRF, OCP, and NO_2 AMF using the standard procedure with a given GLER for those scenes.

3.4 Destriping

The measured NO_2 SCDs are corrected for an instrumental artifact that varies across the orbital track and results in the appearance of "stripes" along the track. The destriping algorithm computes the mean cross-track biases using measurements obtained at latitudes between 30S and 5N and from orbits within 2 orbits of target orbit. These are essentially a set of 60 correction constants, one for each cross-track position, that are subtracted from the measured SCDs to calculate the destriped SCD field, `SlantColumnAmountNO2Destriped` (see Section 4.3). Although the uncorrected SCDs (`SlantColumnAmountNO2`) are also stored in the Level 2 files, we do not use them to calculate the VCDs.

3.5 Stratosphere-troposphere separation

The stratospheric and tropospheric column amounts are retrieved separately under the assumption that the two are largely independent. The stratospheric field is computed first, beginning with creation of a gridded global field $V_{\mathrm{init}} = S/\mathrm{AMF}_{\mathrm{strat}}$ values, assembled from data taken within ± 7 orbits of the target orbit. An *a priori* estimate of the tropospheric contribution to this field, based on a monthly GMI model climatology and cloud measurements, is subtracted, and grid cells where this contribution exceeds 0.3×10^{15} molecules cm^{-2} are masked. Masking ensures that the model contribution to the retrieval is minimal. Note that not all highly polluted areas will be masked in this procedure, since clouds may already hide the tropospheric NO_2 from OMI in those regions. A three-step (interpolation, filtering, and smoothing) algorithm is then applied to fill in the masked regions and data gaps, and to remove residual tropospheric contamination. The resulting stratospheric vertical column field V_{strat} is converted to a slant column field using $\mathrm{AMF}_{\mathrm{strat}}$, and subtracted from S to give the tropospheric slant column. Dividing this by the tropospheric air mass factor $\mathrm{AMF}_{\mathrm{trop}}$ gives the tropospheric vertical column V_{trop}. For details see Bucsela et al. [2013].

3.6 NO_2 data quality

The algorithms that produce L-2 data products are complex (see Figure 4), and the incoming data from the satellite can be noisy—any individual measurement may be anomalous. We have gone to great lengths to automatically recognize anomalous measurements. Many anomalies may be sufficiently severe that the calculation is abandoned, and the derived quantities are assigned a fill-value. Less severe anomalies may not demand abandoning the calculation.

We provide a summary quality flag–a single bit that may be interrogated to select data. This is the least-significant bit of the field `VcdQualityFlags` (See Section 4.3). We *strongly* recommend that users select *only* data for which the least-significant bit of `VcdQualityFlags` is zero, indicating good data. This may be done using a bitwise logical *and* operation on `VcdQualityFlags` and the integer '1'. For example, in FORTRAN,

```
goodData = ( IAND( VcdQualityFlags , 1 ) .EQ. 0 )
```

4 Level-2 Data Product

4.1 File name

OMNO2 L-2 files are written in HDF-EOS version 5 (HDF-EOS5) format and have the following naming convention [Craig et al., 2006, Claas, 2011]:

`<InstrumentID>_<DataType>_<DataID>_<Version>.<Suffix>,`

where

`<DataID> = <ObservationDateTime>-o<Orbit#>`

and

`<Version> = v<Collection#>-<ProductionDateTime>`

Below is an example of an OMNO2 L2 file name:

`OMI-Aura_L2-OMNO2_2011m1010t2318-o38499_v003-2019m0816t193742.he5`

where:

`<InstrumentID>` = `OMI-Aura`
`<DataType>` = `L2-OMNO2`
`<ObservationDateTime>` = `2011m1010t2318`
`<Orbit#>` = `38499`
`<Collection#>` = `003`
`<ProductionDateTime>` = `2019m0816t193742`
`<Suffix>` = `he5`

The observation time is stated to the minute (4 digits); the processing time is stated to the second (6 digits).

4.2 File organization

The HDF-EOS5 file structure is shown in Figure 5.

As HDF-EOS5 files, OMNO2 L-2 files contain a single swath, called `ColumnAmountNO2`, composed of a `Geolocation Fields` group and a `Data Fields` group. This section briefly describes the more commonly-used data fields. A complete list of the fields and metadata information contained in the OMNO2 files can be found in Celarier et al. [2016].

Figure 5: OMNO2 HDF-EOS5 file structure.

- OMI-Aura_L2-OMNO2_2015m0626t0005-o58217_v003-2016m0805t180406.he5
  - HDFEOS
    - ADDITIONAL
      - FILE_ATTRIBUTES
    - SWATHS
      - ColumnAmountNO2
        - Data Fields
        - Geolocation Fields
  - HDFEOS INFORMATION
    - ArchivedMetadata
    - StructMetadata.0
    - coremetadata

4.3 Data description

Two kinds of data fields are found in the file: product data, and flags. While most product data fields are of a floating-point ("real") type, some, such as `CloudFraction` have been stored as integers to make the file a bit smaller. The fields have field-level metadata which characterize the values contained in the data fields. These include the fill-values that are used when no meaningful data are available, and a scale factor and offset. These are usually 1 and 0, respectively, indicating that the values have not been modified. However, `CloudFraction` is one exception, since the integer values stored are 1000 times the actual value: Its scale factor is 0.001 The field `TerrainReflectivity` is similarly scaled.

Flag fields may have 8, 16, 32, or 64 bits per word, stored as unsigned integer values (one word per FOV or per swath, as appropriate), containing a collection of bits that each indicate processing conditions that should be taken as warnings or errors, or may indicate which path was taken through one of the algorithms, or may indicate why some data field(s) have been assigned fill values. The meanings of single bits, and groups of bits, for each flag field are found in Celarier et al. [2016].

The following paragraphs describe briefly the fields that are of the greatest interest to most end-users of the data product. A complete list of data fields is in Table 6.

SlantColumnAmountNO2Destriped and **SlantColumnAmountNO2Std**: Retrieved slant column density (SCD) S and its uncertainty. S is the retrieved total areal density of NO_2 molecules along the effective optical path from the sun into the atmosphere, and then toward the satellite. This is calculated from the measured Earthshine radiance and solar irradiance using a variant of the DOAS algorithm (see section 3), with an NO_2 cross section measured at 220 K. Variations that are due to calibration differences between the detector cells have been removed using the destriping procedure described in Section 3.4. The units are molecules cm^{-2}.

ColumnAmountNO2Strat and **ColumnAmountNO2StratStd**: Estimates of the stratospheric vertical column density (VCD), V_{strat}, derived from S, and its uncertainty. The units are molecules cm^{-2}.

ColumnAmountNO2Trop and **ColumnAmountNO2TropStd**: Estimates of the tropospheric vertical column density, V_{trop}, derived from S, and its uncertainty. The units are molecules cm^{-2}.

ColumnAmountNO2 and **ColumnAmountNO2Std**: Estimates of the total (i.e., $V = V_{\mathrm{strat}} + V_{\mathrm{trop}}$) vertical column density and its uncertainty. The units are molecules cm^{-2}.

ScatteringWeight: Vector $\mathbf{A}$ [no units] that describes the relationship between slant column density, S_i, and the vertical column density, V_i, for each atmospheric layer i:

$$S = \sum_i S_i \approx \sum A_i \cdot V_i \tag{2}$$

$\mathbf{A}$ is relatively insensitive (within $\sim 20\%$) to the wavelength within the spectral region used in OMNO2, so only a single value, representative of the entire spectral fitting window, is provided. $\mathbf{A}$ is a function of the optical geometry (solar and viewing azimuth and zenith angles), surface reflectivity, and cloud pressure and cloud fraction, and contains a correction for the temperature dependence of the NO_2 cross section. The scattering weights are stored as a 3-dimensional array with dimensions (pressure levels, across track, along track; *e.g.,* [35,60,1644]). The grid of pressure levels is available as the data field `ScatteringWtPressure`.

Partial slant column (*e.g.,* tropospheric) densities may be computed from Eq. (2) using ranges of i falling within the partial column, and V_i values derived from models. The partial column Air-Mass Factor (AMF) (*e.g.,* AMFtrop, Section 3.3) can be obtained by dividing Eq. (2) by the corresponding partial vertical column (*e.g.,* V_{trop}). Methods for comparing OMI columns with external datasets may be found in Bucsela et al. [2008], Lamsal et al. [2014], and references therein.

XtrackQualityFlags: The cross-track quality flags indicate specific likely problems with the radiance measurements, due to the row anomaly (Section 2.4). As a general rule, for files with measurements after June 2007, one should not use data where the `XtrackQualityFlags` field is nonzero. For these FoVs, the current data product version, the column amount fields are set to their fill values. However, before this time, the `XtrackQualityFlags` words are set to a fill value. Thus, the user should only use Level 2 data where `XtrackQualityFlags` is equal to zero, *or* equal to the fill value. The fill value can be found in each field's metadata.

vcdQualityFlag: This variable contains quality assurance information for the tropospheric vertical column. The least significant bit is the **summary quality flag**. We recommend that users only use data for which this bit is zero (*i.e.*, `vcdQualityFlag` is an even integer).

4.4 Limitations

As with all remote sensing data sets, there are subtleties in the OMNO2 data that are due to geophysics, instrumental measurements, and the retrieval algorithm. Users of the data are encouraged to communicate directly with members of the OMI NO_2 algorithm team. We also encourage those using the data to read Bucsela et al. [2013], Lamsal et al. [2014], Marchenko et al. [2015], Krotkov et al. [2017], Choi et al. [2019], and Lamsal et al. [in preparation], which describe the algorithm in detail.

Particular attention should be paid to the various data quality flags. For most users, the Summary Quality Flag (least significant bit of the `vcdQualityFlags` data field) should suffice. In row-anomaly-affected FoVs, the column amount fields have been set to their respective fill values, so `XTrackQualityFlags` does not need to be explicitly checked. In certain periods of time, this row-anomaly problem will result in up to 50% field-of-view rejection rate.

While features inherent to the stratospheric NO_2 field are relatively large, compared to the geographical extent of OMI's larger (far-off-nadir) FoVs, many local features in tropospheric fields are smaller than OMI FoVs. This may lead to a negative bias in the column amounts when there is a local maximum within the FoV. The retrieval algorithm permits the values of any of the columns to be negative. In particular, small-magnitude negative values are not uncommon in areas that are generally unpolluted (*e.g.,* over open oceans). When computing statistics (*e.g.,* monthly or annual average) from multiple measurements, it is important to include all valid measurements, regardless of their sign, in order to avoid biases.

Caution required in the use of the data (*e.g.,* trend analyses) when some of the data are heavily flagged or filled that could lead to differences in contribution and weight from each scan position over the period of interest. For example, in a time-series analysis of data over a time-span that includes periods of RA-affected and periods of non-RA-affected measurements, the non-RA periods could be sampled to include only those cross-track positions that are unflagged in the RA period.

OMI has a number of measurement modes. Besides the normal "global" mode, it has two "zoom" modes. Zoom mode measurements are indicated by the field `InstrumentConfigurationID` having a value > 7. We have not sufficiently evaluated the performance of the retrieval algorithm for zoom mode obervations. Therefore, when the instrument is operating in a zoom mode, no NO_2 vertical amounts are calculated. This occasionally results in entire orbital granules having nothing but fill-values for the NO_2 column fields.

5 The Level-2 gridded NO_2 product, OMNO2G

The Level-2 HDF-EOS5 files, described in Section 4.3, are used to create Level-2-gridded, daily data products, called OMNO2G. These are also HDF-EOS5 files, but are Grid type, rather than Swath type files. Each $0.25° \times 0.25°$ geographical grid cell can be thought of as containing a "stack" of up to 15 L-2 FOVs' data collocated with the grid cell. In practice, only up to ~ 6 are populated. The OMNO2G data product can be useful for considering L-2 data within a geographic area of interest. It was originally conceived as a "global overpass" data set. It does have the advantage of containing a geographically sorted list of L-2 FOVs, which may be more convenient for users interested in regional NO_2 fields.

Since only as many as 15 L-2 FOVs are identified with a grid cell, there may be some selection: greater priority can be given to FOVs having the shortest optical path length (defined as $\sec\theta + \sec\theta_\circ$, where θ is the viewing zenith angle, and $\theta_\circ$ is the solar zenith angle). The user should be aware that the identification of a L-2 FOV with a grid cell is based entirely on the location of the FOV center.

Since the grid array has a spatial resolution of 0.25°, and many of the OMI fields-of-view are considerably larger than that, it is a good idea to examine data that are identified with a larger region than the actual region of interest.

5.1 File name

The file name for the OMNO2G files is of the form:

`OMI-Aura_L2G-OMNO2G_<ObservationDate>_v003-<ProductionDate>.he5`

An example is:

`OMI-Aura_L2G-OMNO2G_2019m1106_v003-2019m1107t190903.he5`

5.2 File structure

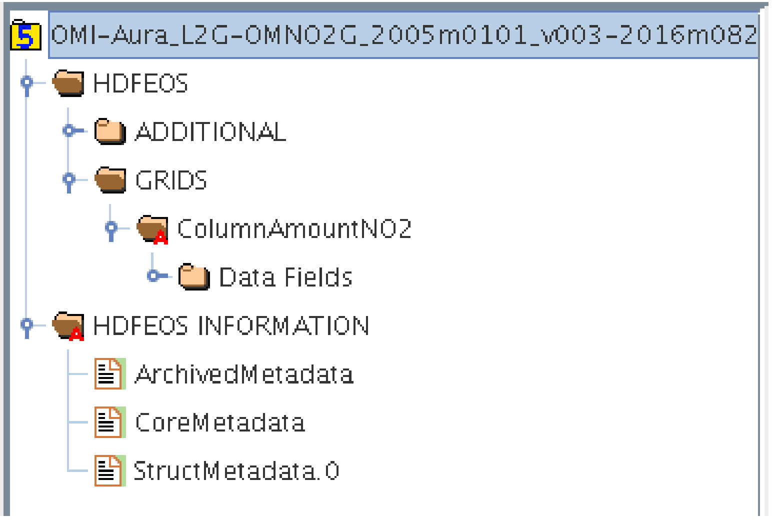
The structure of the HDF-EOS5 data file is shown in Figure 6. There are 38 data sets within the `Data Fields` group, which are selected from the L-2 data. Table 6 lists all the fields.

Each data field has dimensions [1440 , 720 , 15] (number of cells in the longitude direction, number of cells in the latitude direction, and maximum L-2 FOVs identified with each cell, respectively). Unpopulated elements in the data field are assigned a fill value.

5.3 Data fields

The data fields are listed in Table 6. While most of the fields' names and data values are inherited from the OMNO2 data product, there are several that are unique to the OMNO2G. We describe those here.

Figure 6: OMNO2G HDF-EOS5 file structure.



- OMI-Aura_L2G-OMNO2G_2005m0101_v003-2016m082
  - HDFEOS
    - ADDITIONAL
    - GRIDS
      - ColumnAmountNO2
        - Data Fields
  - HDFEOS INFORMATION
    - ArchivedMetadata
    - CoreMetadata
    - StructMetadata.0

OrbitNumber Since the OMNO2G compiles data from multiple orbits into a single daily file, the data in the stack are may be associated with different orbits. This field identifies the orbit.

LineNumber The swath (exposure) number in the along-track direction (0-based).

SceneNumber The cross-track position of the FoV(0-based).

NumberOfCandidateScenes The number of populated members in the stack. The maximum is 15. Typical values range from 0 to 6.

PathLength The length of the optical path; it is also the geometric air mass factor. It is equal to $\sec\theta_o + \sec\theta$, where θ_o is the solar zenith angle and θ is the viewing zenith angle.

5.4 Limitations

Since the L-2 data are copied directly into the OMNO2G data product, the general quality of the data is the same. (See Section 4.4.) For some purposes, in some geographical regions (*e.g.,* in polar regions), more than 15 L-2 FoVs may have their centers land in a particular cell, and some L-2 data, whose optical path lengths are longer than the others, may be excluded. This should happen rarely, but may lead to slight shifts in statistical measures.

Since the identification of a grid cell with a L-2 FoV is based solely on the location of its center, some FoVs identified with nearby grid cells may be relevant to a particular grid cell.

In the current version of OMNO2G, the FoV corners (from the product OMPIXCOR) are not copied from the L-2 files. Any work that involved detailed knowledge of the L-2 FoV geometries will have to be done using either the L-2 data product (OMNO2) or the OMPIXCOR data product.

6 The Level-3 gridded NO_2 product, OMNO2d

The L-2 HDF-EOS5 files, described in Section 4.3 are used to create L-3 daily data products, called OMNO2d. These are also HDF-EOS5 files, but are Grid type, rather than Swath type files.

In the archived data product, a day's worth of L-2 data (usually 14 or 15 orbits) are mapped into a single $0.25° \times 0.25°$ latitude-longitude grid. The parameters specifying the grid cell locations are available in the metadata included in each file. Each file contains five (5) grid fields: `ColumnAmountNO2`, `ColumnAmountNO2CloudScreened`, `ColumnAmountNO2Trop`, `ColumnAmountNO2TropCloudScreened`, and `Weight`. (See Table 6.)

In each of the first four of these fields, the value given in any grid cell is an area-weighted average of the values of the corresponding field (`ColumnAmountNO2` or `ColumnAmountNO2Trop`) in all the L-2 FOVs that have any overlap at all with that grid cell. See the OMPIXCOR Readme file for more information. The weighting scheme is described below. The L-3 data product is available from the Goddard Earth Sciences Data and Information Services Center (GES-DISC). See Section 9 for details.

6.1 File name

The names of the OMNO2d files are of the form:

```
OMI-Aura_L3-OMNO2d_<ObservationDate>_v003-<ProductionDate>.he5
```

An example is:

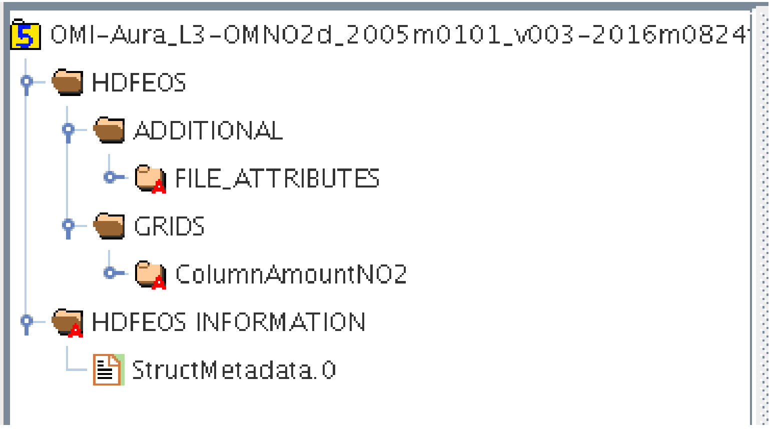
```
OMI-Aura_L3-OMNO2d_2019m0201_v003-2019m0202t141455.he5.
```

6.2 File structure

The structure of the HDF-EOS5 data file is shown in Figure 7. Metadata are found in four places:

1. Structural metadata are in `/HDFEOS INFORMATION/StructMetadata.0`;
2. Metadata concerning the source data are found in `/HDFEOS/ADDITIONAL/FILE_ATTRIBUTES`;
3. The grid metadata are in group attributes of the group `/HDFEOS/GRIDS/ColumnAmountNO2`;
4. Metadata concerning individual fields is attached to the grid fields themselves.

Figure 7: OMNO2d HDF-EOS5 file structure.



OMI-Aura_L3-OMNO2d_2005m0101_v003-2016m0824
- HDFEOS
  - ADDITIONAL
    - FILE_ATTRIBUTES
  - GRIDS
    - ColumnAmountNO2
- HDFEOS INFORMATION
  - StructMetadata.0

6.3 Data fields

The data fields contain the gridded data. The first grid cell (the one with the smallest indices) has edges at 180° West longitude and 90° South latitude. Grid cells that did not have any overlapping L-2 FoVs among the input files are assigned a fill value $(-2.^{100} \simeq -1.26765{\times}10^{30})$. The data are in units of molecules cm^{-2}.All the NO_2 data fields are produced by first screening the L-2 data, and then calculating the weighted average of the remaining data. The screening criteria are listed in Table 4.

Table 4: Criteria used to screen OMNO2 data for use in generating OMNO2d data product.

Field	"Pass" criterion
SolarZenithAngle	$< 85°$
CloudFraction	$< 30\%$ (For cloud-screened fields)
CloudFraction	No filter (For other fields)
XTrackQualityFlags	0 or 255
vcdQualityFlags	Asceding orbit; Summary flags not set

The weighted averages are computed as follows: For each grid cell (j) and each L-2 FoV (i), the area of overlap (Q_{ij} = area of overlap ÷ area of grid cell) is computed, and the area of the FoV A_i is known. The weight is linear with A_i:

$$w_{A_i} = 1 - (A_i - A_{\min})/A_{\max} \tag{3}$$

(larger area, smaller weight) , and is proportional to the area of overlap (larger overlap, larger weight.) The weight for FoV i and cell j is

$$w_{ij} = w_{A_i} \cdot Q_{ij}. \tag{4}$$

The total of all weights for cell j, w_j , is stored in the data field `Weight`. This can be used to combine gridded data from multiple L-3 files and geographical regions in order to rapidly compute spatial or temporal averages. Indexing the relevant data sets by k, compute V_j:

$$V_j = \frac{\sum_k w_{kj} V_{kj}}{\sum_k w_{kj}} \tag{5}$$

6.4 Limitations

While the L-3 data product can be used to assess the daily NO_2 column densities (or, when combined as described above, for longer time periods), it is important to remember that the values in the grid cells are weighted averages of a number of OMI measurements, and the value in a cell may not correspond to any one actual measurement.

Because the 8–10 OMI FoVs farthest from nadir are quite large (see Figure 2), their contribution to the weighted average in a grid cell may be affected by actual NO_2 columns some distance away from the cell. This is particularly important when looking at daily L-3 data, as, especially in the tropics, some grid cells may have contributions from only the OMI swath edge FoVs, while others have contributions from only the smaller, near-nadir FoVs. The natural spatial resolution of the former is coarser than the grid cells, while the spatial resolution of the latter is comparable to the grid cell size.

To compare different small areas, one should consider the Weight field values for each. The weights of better-characterized grid cells will tend to be larger than those of less-well-characterized grid cells. This is also a consideration when constructing time-series for a set of grid cells: Because of Aura's precession relative to the fixed geographical (latitude-longitude) grid, a chosen grid cell will be under large OMI FoVs on some days, and under small ones on other days. One should especially look at the weights if one finds an apparent spatial or temporal periodicity in the NO_2 columns.

The product development team has chosen a cloud screening criterion of the effective cloud fraction $f_c < 0.30$ (see Table 4) for the cloud-screened variables, which reflects a compromise between data quality and quantity.

7 Software versions

This document applies to the public release of the OMI L-2 NO_2 data, product version 4.0, archived as collection 3 and released in December 2019. The L-2 algorithm is divided into four processes, each performed by a separate application. The end result is the creation of the OMNO2 L-2 data product from the OMI Level L-1B product. The L-0 to L-1B processing version is designated "Collection 3". This is not to be confused with OMNO2 version 4.0.

The software versions used to produce product version 4.0 are listed in Table 5.

Table 5: Version numbers of data products and the applications (Apps) that create them.

Data product	Product version	App version
OMNO2	4.0	2.1.6.00
OMNO2G	4.0	1.3.0
OMNO2d	4.0	1.0.5.03
OMNO2SCD		0.2.0.06
OMNO2B		1.6.7.0
OMGLER	1.0	1.0.2.4
OMCDO2N		3.0.0.1
OMO4SCD		0.0.10

8 Data quality

The quality of the data in this release has been established by consistency checks with previous versions, which were extensively evaluated [Celarier et al., 2008, Lamsal et al., 2014, Krotkov et al., 2017, Choi et al., 2019]. Our validation effort using other independent measurements and campaign data from ground-, aircraft-, and satellite-based instruments is ongoing.

Our evaluation over unpolluted areas shows that the OMNO2products agree well with independent satellite- and ground-based Fourier Transform Infrared (FTIR) measurements [Krotkov et al., 2017]. Intercomparisons of NO_2 SCD [Zara et al., 2018] and AMF [Lorente et al., 2017] from different research groups (*e.g.,* NASA, BIRA-IASB, IUP, and KNMI) suggests that the version 4.0 retrievals are in close agreement with other independent retrievals. The fitting error in the NO_2 slant column is estimated to be $0.3 - 1 \times 10^{15}$ molecules cm^{-2}. A document detailing algorithm updates and validation with other independent observations is in preparation.

The stratosphere-troposphere separation algorithm uses modeled atmospheric profiles. One consequence of this is that, in relatively clean regions (*e.g.,* over the open ocean), the tropospheric NO_2 column is essentially model-driven, as there is no separable tropospheric column information in the slant column denisity.

9 Product availability

The OMNO2 product is archived and distributed from the NASA Goddard Earth Sciences Data & Information Services center (GES-DISC). The files can be directly downloaded from the GES-DISC Mirador site which provides parameters and spatial subset capabilities. OMI products are written in HDF-EOS5 format. GES-DISC also provides a list of tools that read HDF-EOS5 data files.

In order to download GES-DISC data via browser window, or from the command line, or via desktop applications, you must (1) register with Earthdata Login and (2) authorize NASA GES-DISC Data Access. See the first two items below.

The following is a list of data sources and resources related to the OMNO2 data.

To register with Earthdata Login (step 1):
https://urs.earthdata.nasa.gov/users/new

To authorize NASA GES-DISC Data Access (step 2):
https://urs.earthdata.nasa.gov/

OMNO2 data:
https://disc.gsfc.nasa.gov/datasets/OMNO2_V003/summary

OMNO2G data:
https://disc.gsfc.nasa.gov/datasets/OMNO2G_003/summary

OMNO2d data:
https://disc.gsfc.nasa.gov/datasets/OMNO2d_003/summary

Station overpass data:
https://avdc.gsfc.nasa.gov/pub/data/satellite/Aura/OMI/V03/L2OVP/OMNO2/

OMNO2d images:
https://avdc.gsfc.nasa.gov/

OMNO2 Readme (also covers OMNO2G and OMNO2d):
https://aura.gesdisc.eosdis.nasa.gov/data/Aura_OMI_Level2/OMNO2.003/doc/README.OMNO2.pdf

OMNO2 File description:
https://docserver.gesdisc.eosdis.nasa.gov/repository/Mission/OMI/3.3_ScienceDataProductDocumentation/3.3.2_ProductRequirements_Designs/OMNO2_v3.0_data_product_specification_20160913.pdf

OMI Data User's Guide:
https://docserver.gesdisc.eosdis.nasa.gov/repository/Mission/OMI/3.3_ScienceDataProductDocumentation/3.3.2_ProductRequirements_Designs/README.OMI_DUG.pdf

OMI FoV corners:
https://disc.gsfc.nasa.gov/datasets/OMPIXCOR_V003/summary

GES-DISC-hosted tools for reading HDF-EOS5 files:
https://disc.gsfc.nasa.gov/information/howto?keywords=aura

GES-DISC-hosted other recipes:
https://disc.gsfc.nasa.gov/information/howto

10 Reporting problems and requesting information

To report problems, or pose questions and comments related to the OMNO2 algorithm, data quality, and file structure, please sends electronic mail to nickolay.a.krotkov@nasa.gov or lok.lamsal@nasa.gov.

Table 6: Fields available in the three OMI NO_2 data sets. D indicates fields in the *Data Fields* group; G indicates fields in the *Geolocation Fields* group.

Field	OMNO2	OMNO2G	OMNO2d
AMFQualityFlags	D		
AlgorithmFlags	D		
AmfStrat	D		
AmfStratClear	D		
AmfStratCloudy	D		
AmfStratStd	D		
AmfTrop	D		
AmfTropClear	D		
AmfTropCloudy	D		
AmfTropStd	D		
CloudFraction	D	D	
CloudFractionStd	D	D	
CloudPressure	D	D	
CloudPressureStd	D	D	
CloudRadianceFraction	D	D	
ColumnAmountNO2	D	D	D
ColumnAmountNO2CloudScreened			D
ColumnAmountNO2Std	D	D	
ColumnAmountNO2Strat	D	D	
ColumnAmountNO2StratStd	D	D	
ColumnAmountNO2Trop	D	D	D
ColumnAmountNO2TropCloudScreened			D
ColumnAmountNO2TropStd	D	D	
FoV75Area	G		
FoV75CornerLatitude	G		
FoV75CornerLongitude	G		
GroundPixelQualityFlags	G		
InstrumentConfigurationId	D	D	
Latitude	G	D	
LineNumber		D	
Longitude	G	D	
MeasurementQualityFlags	D	D	
NumberOfCandidateScenes		D	
OMNO2SCD_algoFlags	D		
OMNO2SCD_algoxFlags	D		
OMNO2SCD_procFlags	D		
OMNO2SCD_radFlags	D		

Continued. . .

Table 6, continued.

Field	OMNO2	OMNO2G	OMNO2d
OMNO2SCD_scdFlags	D		
OMNO2SCD_wvlnFlags	D		
OrbitNumber		D	
OrbitPhase	G		
PathLength		D	
ScatteringWeight	D		
ScatteringWtPressure	D		
ScdApStrat	D		
ScdApTrop	D		
SceneLER	D		
SceneNumber		D	
ScenePressure	D		
SlantColumnAmountCHOCHO	D		
SlantColumnAmountCHOCHOStd	D		
SlantColumnAmountH2O	D		
SlantColumnAmountH2OStd	D		
SlantColumnAmountNO2	D	D	
SlantColumnAmountNO2Destriped	D	D	
SlantColumnAmountNO2Std	D	D	
SmallPixelRadiance	D		
SmallPixelRadiancePointer	D		
SolarAzimuthAngle	G	D	
SolarZenithAngle	G	D	
SpacecraftAltitude	G	D	
SpacecraftLatitude	G	D	
SpacecraftLongitude	G	D	
TerrainHeight	D		
TerrainPressure	D	D	
TerrainReflectivity	D	D	
Time	G	D	
TropopausePressure	D	D	
VcdApBelowCloud	D		
VcdApStrat	D		
VcdApTrop	D		
VcdQualityFlags	D	D	
ViewingAzimuthAngle	G	D	
ViewingZenithAngle	G	D	
WavelengthRegistrationCheck	D		
WavelengthRegistrationCheckStd	D		
Weight			D
XTrackQualityFlags	D	D	

A Time calculations

The local mean, civil, or apparent time at the center of any OMI FOV can be obtained from the geolocation data, using the variable `Time` for the swath and the variable `Longitude` for the FOVÅpparent time requires, additionally, calculation of the Equation of Time. The `Time` variable is given in decimal TAI-93 format, so should be converted (for sub- minute precision) to UTC. The local solar times—mean and apparent—are of importance when the photochemical lifetimes of NO_2 are important. The relevant equations are:

$$\mathrm{UTC} = \mathrm{TAI} - 32 - \mathrm{LS} \tag{6}$$

$$\mathrm{LCT} = \mathrm{UTC} + \mathrm{TZ} \tag{7}$$

$$\mathrm{LMST} = \mathrm{UTC} + \lambda/15 \tag{8}$$

$$\mathrm{LAST} = \mathrm{LMST} + \mathrm{E} \tag{9}$$

Where

UTC = Coordinated Universal Time
LCT = Local civil time
LMST = Local mean solar time
LAST = Local apparent solar time
LS = Number of leap seconds added since July 1, 2004 One-second additions occurred at midnight after Dec. 31, 2005, Dec 31, 2008, Jun 30, 2012, Jun 30, 2015, and Dec 31, 2016)
TZ = Time zone value (*e.g.,* −4 hours for U.S. Eastern Daylight Time)
λ = Longitude, in degrees (East positive, West negative)
E = Equation of Time

The Equation of Time, in minutes, can be approximated with a precision of < 6 s by the formula

$$\mathrm{E} = 9.87\sin(2B) - 7.53\cos(B) - 1.5\sin(B) \tag{10}$$

where

$$B = 360(\mathrm{DOY} - 81)/365 \tag{11}$$

$$\mathrm{DOY} = \text{Day of Year} \tag{12}$$

Formulae for higher-precision calculations of E can be found in various reference sources [Seidelmann, 2005].

References

K. F. Boersma, H. J. Eskes, R. J. Dirksen, R. J. van der A, J. P. Veefkind, P. Stammes, V. Huijnen, Q. L. Kleipool, M. Sneep, J. Claas, J. Leitão, A. Richter, Y. Zhou, and D. Brunner. An improved tropospheric no$_2$ column retrieval algorithm for the ozone monitoring instrument trends and variability in surface ozone over the united states. *Atmos. Meas. Tech.*, 2011. doi: 10.5194/amt-4-1905-2011.

E. Bucsela, E. Celarier, M. Wenig, J. Gleason, J. Veefkind, K. Boersma, and E. Brinksma. Algorithm for NO/sub 2/ vertical column retrieval from the ozone monitoring instrument. *IEEE Trans. Geosci. Remote Sens.*, 44:1245–1258, 2006. ISSN 0196-2892. doi: 10.1109/TGRS.2005.863715.

E. J. Bucsela, a. E. Perring, R. C. Cohen, K. F. Boersma, E. A. Celarier, J. F. Gleason, M. O. Wenig, T. H. Bertram, P. J. Wooldridge, R. Dirksen, and J. P. Veefkind. Comparison of tropospheric NO 2 from in situ aircraft measurements with near-real-time and standard product data from OMI. *J. Geophys. Res.*, 113 (D16):D16S31, may 2008. ISSN 0148-0227. doi: 10.1029/2007JD008838. URL http://doi.wiley.com/10.1029/2007JD008838.

E. J. Bucsela, N. A. Krotkov, E. A. Celarier, L. N. Lamsal, W. H. Swartz, P. K. Bhartia, K. F. Boersma, J. P. Veefkind, J. F. Gleason, and K. E. Pickering. A new stratospheric and tropospheric NO2 retrieval algorithm for nadir-viewing satellite instruments: applications to OMI. *Atmos. Meas. Tech.*, 6(10):2607–2626, 2013. ISSN 1867-1381. doi: 10.5194/amt-6-2607-2013.

E. A. Celarier, E. J. Brinksma, J. F. Gleason, J. P. Veerkind, A. Cede, J. R. Herman, D. Ionov, F. Goutail, J. P. P. Pommereau, J. C. C. Lambert, M. Van Roozendael, G. Pinardi, F. Wittrock, A. Schönhardt, A. Richter, O. W. Ibrahim, T. Wagner, B. Bojkov, G. Mount, E. Spinei, C. M. Chen, T. J. Pongetti, S. P. Sander, E. J. Bucsela, M. O. Wenig, D. P. J. Swart, H. Volten, M. Kroon, and P. F. Levelt. Validation of ozone monitoring instrument nitrogen dioxide columns. *J. Geophys. Res. Atmos.*, 113(6):3357–3365, mar 2008. ISSN 0021-9606. doi: 10.1029/2007JD008908.

E. A. Celarier, L. N. Lamsal, E. J. Bucsela, S. V. Marchenko, and N. A. Krotkov. OMNO2 data product specification, 2016. URL http://disc.sci.gsfc.nasa.gov/Aura/data-holdings/OMI/documents/v003/OMNO2_data_product_specification.pdf.

K. Chance, T. P. Kurosu, and C. E. Sioris. Undersampling correction for array detector-based satellite spectrometers. *Appl. Opt.*, 44(7):1296, 2005. ISSN 0003-6935. doi: 10.1364/AO.44.001296. URL http://www.opticsinfobase.org/abstract.cfm?URI=AO-44-7-1296.

S. Choi, L. N. Lamsal, M. Follette-Cook, J. Joiner, N. A. Krotkov, W. H. Swartz, K. E. Pickering, C. P. Laughner, W. Appel, G. Pfister, P. E. Saide, R. C. Cohen,

A. J. Weinheimer, and J. R. Herman. Assessment of NO_2 observations during DISCOVER-AQ and KORUS-AQ field campaigns. *Atmos. Meas. Tech. Disc.*, 2019. doi: 10.5194/amt-10-333-2017.

J. J. Claas. Discussion on L2 files, November 2011.

C. Craig, D. Cuddy, P. Veefkind, P. Leonard, P. Wagner, C. Vuu, and D. Shepard. Hdf-eos aura file format guidelines, October 2006. URL `http://disc.sci.gsfc.nasa.gov/Aura/additional/documentation/HDFEOS_Aura_File_Format_Guidelines.pdf`.

A. R. Douglass, R. S. Stolarski, S. E. Strahan, and P. S. Connell. Radicals and reservoirs in the GMI chemistry and transport model: Comparison to measurements. *J. Geophys. Res.*, 109:D16302, 2004. doi: 10.1029/2004JD004632.

Z. Fasnacht, A. Vasilkov, D. Haffner, W. Qin, J. Joiner, N. Krotkov, A. M. Sayer, and R. Spurr. A geometry-dependent surface lambertian-equivalent reflectivity product for UV/Vis retrievals: Part II. evaluation over open ocean. *Atmos. Meas. Tech. Discuss.*, 2019. doi: 10.5194/amt-2019-260.

N. A. Krotkov, L. N. Lamsal, E. A. Celarier, W. H. Swartz, S. Marchenko, E. J. Bucsela, K. L. Chan, M. Wenig, and M. Zara. The version 3 OMI NO2 standard product. *Atmos. Meas. Tech.*, 10(9):3133–3149, September 2017. ISSN 1867-1381. doi: https://doi.org/10.5194/amt-10-3133-2017.

L. N. Lamsal, R. V. Martin, A. van Donkelaar, E. A. Celarier, E. J. Bucsela, K. F. Boersma, R. Dirksen, C. Luo, Y. Wang, M. Steinbacher, E. A. Celarier, E. J. Bucsela, E. J. Dunlea, J. P. Pinto, K. F. Boersma, R. Dirksen, C. Luo, Y. Wang, N. A. Krotkov, E. A. Celarier, W. H. Swartz, K. E. Pickering, E. J. Bucsela, J. F. Gleason, R. V. Martin, S. Philip, H. Irie, A. Cede, J. Herman, A. Weinheimer, J. J. Szykman, and T. N. Knepp. Indirect validation of tropospheric nitrogen dioxide retrieved from the OMI satellite instrument: Insight into the seasonal variation of nitrogen oxides at northern midlatitudes. *J. Geophys. Res.*, 115 (D16):11587–11609, mar 2010. ISSN 2169-897X. doi: 10.1029/2009JD013351.

L. N. Lamsal, N. A. Krotkov, E. A. Celarier, W. H. Swartz, K. E. Pickering, E. J. Bucsela, J. F. Gleason, R. V. Martin, S. Philip, H. Irie, A. Cede, J. Herman, A. Weinheimer, J. J. Szykman, and T. N. Knepp. Evaluation of OMI operational standard NO2 column retrievals using in situ and surface-based NO2 observations. *Atmos. Chem. Phys.*, 14(21):11587–11609, 2014. ISSN 1680-7316. doi: 10.5194/acp-14-11587-2014.

L. N. Lamsal, N. A. Krotkov, A. Vasilkov, S. Marchenko, J. Joiner, E. S. Yang, W. Qin, Z. Fasnacht, D. Haffner, S. Choi, W. H. Swartz, and B. Fisher. Improved standard nitrogen dioxide product from Aura/OMI. *TBD*, in preparation. doi: TBD. in preparation.

A. Lorente, K. F. Boersma, H. Yu, S. Dörner, A. Hilboll, A. Richter, M. Liu, L. N. Lamsal, M. Barkley, I. De Smedt, M. Van Roozendael, Y. Wang, T. Wagner, S. Beirle, J. T. Lin, N. Krotkov, P. Stammes, P. Wang, H. J. Eskes, and M. Krol. Structural uncertainty in air mass factor calculation for NO_2 and HCHO satellite retrievals. *Atmos. Meas. Tech.*, 10(10):759–782, March 2017. ISSN 1867-1381. doi: https://doi.org/10.5194/amt-10-759-2017.

S. Marchenko, N. A. Krotkov, L. N. Lamsal, E. A. Celarier, W. H. Swartz, and E. J. Bucsela. Revising the slant column density retrieval of nitrogen dioxide observed by the Ozone Monitoring Instrument. *J. Geophys. Res.*, 120(11):5670–5692, jun 2015. ISSN 2169-897X. doi: 10.1002/2014JD022913.

S. Marcheko and M. DeLand. Solar spectral irradiance changes during cycle 24. *Astrophys. J.*, 289:117, 2014. doi: 10.1088/0004-637X/789/2/117.

P. I. Palmer, D. J. Jacob, K. Chance, R. V. Martin, R. J. D. Spurr, T. P. Kurosu, I. Bey, R. Yantosca, A. Fiore, and Q. Li. Air mass factor formulation for spectroscopic measurements from satellites' application to formaldehyde retrievals from the global ozone monitoring experiment. *J. Geophys. Res.*, 106(D13):14539–14550, 2001. doi: 10.1029/2000JD900772.

W. Qin, Z. Fasnacht, D. Haffner, A. Vasilkov, J. Joiner, N. Krotkov, B. Fisher, and R. Spurr. A geometry-dependent surface lambertian-equivalent reflectivity product for UV/Vis retrievals part I: Evaluation over land surfaces using measurements from OMI at 466nm. *Atmos. Meas. Tech.*, 2019. doi: 10.5194/amt-12-3997-2019.

V. M. E. Schenkeveld, G. Jaross, S. Marchenko, D. Haffner, Q. L. Kleipool, N. C. Rozemeijer, J. P. Veefkind, and P. F. Levelt. In-flight performance of the Ozone Monitoring Instrument. *Atmos. Meas. Tech.*, 2017. doi: 10.5194/amt-10-1957-2017.

P. K. Seidelmann. *Explanatory Supplement to the Astronomical Almanac.* University Science Books, 2005. ISBN 978-1891389450.

S. A. Strode, J. M. Rodriguez, J. A. Logan, O. R. Cooper, J. C. Witte, L. N. Lamsal, M. Damon, B. V. Aartsen, S. D. Steenrod, and S. E. Strahan. Trends and variability in surface ozone over the United States. *J. Geophys. Res.*, 2015. doi: 10.1002/2014JD022784.

A. Vasilkov, W. Qin, N. A. Krotkov, L. N. Lamsal, R. Spurr, D. Haffner, J. Joiner, E. S. Yang, and S. Marchenko. Accounting for the effects of surface BRDF on satellite cloud and trace-gas retrievals: A new approach based on geometry-dependent lambertian-equivalent reflectivity applied to OMI algorithms. *Atmos. Meas. Tech.*, 2017. doi: 10.5194/amt-10-333-2017.

A. Vasilkov, E. S. Yang, S. Marchenko, W. Qin, L. N. Lamsal, J. Joiner, N. A. Krotkov, D. Haffner, P. K. Bhatia, and R. Spurr. A cloud algorithm based on the

O_2-O_2 477 nm absorption band featuring an advanced spectral fitting method and the use of surface geometry-dependent lambertian-equivalent reflectivity. *Atmos. Meas. Tech.*, 2018. doi: 10.5194/amt-11-4093-2018.

M. Zara, K. F. Boersma, I. De Smedt, A. Richter, E. Peters, J. H. G. M. Van Geffen, S. Beirle, T. Wagner, M. Van Roozendael, S. Marchenko, L. N. Lamsal, and H. J. Eskes. Improved slant column density retrieval of nitrogen dioxide and formaldehyde for OMI and GOME-2A from QA4ECV: intercomparison, uncertainty characterization, and trends. *Atmos. Meas. Tech. Discuss.*, NA(9):1–47, January 2018. ISSN 1867-1381. doi: https://doi.org/10.5194/amt-2017-453.