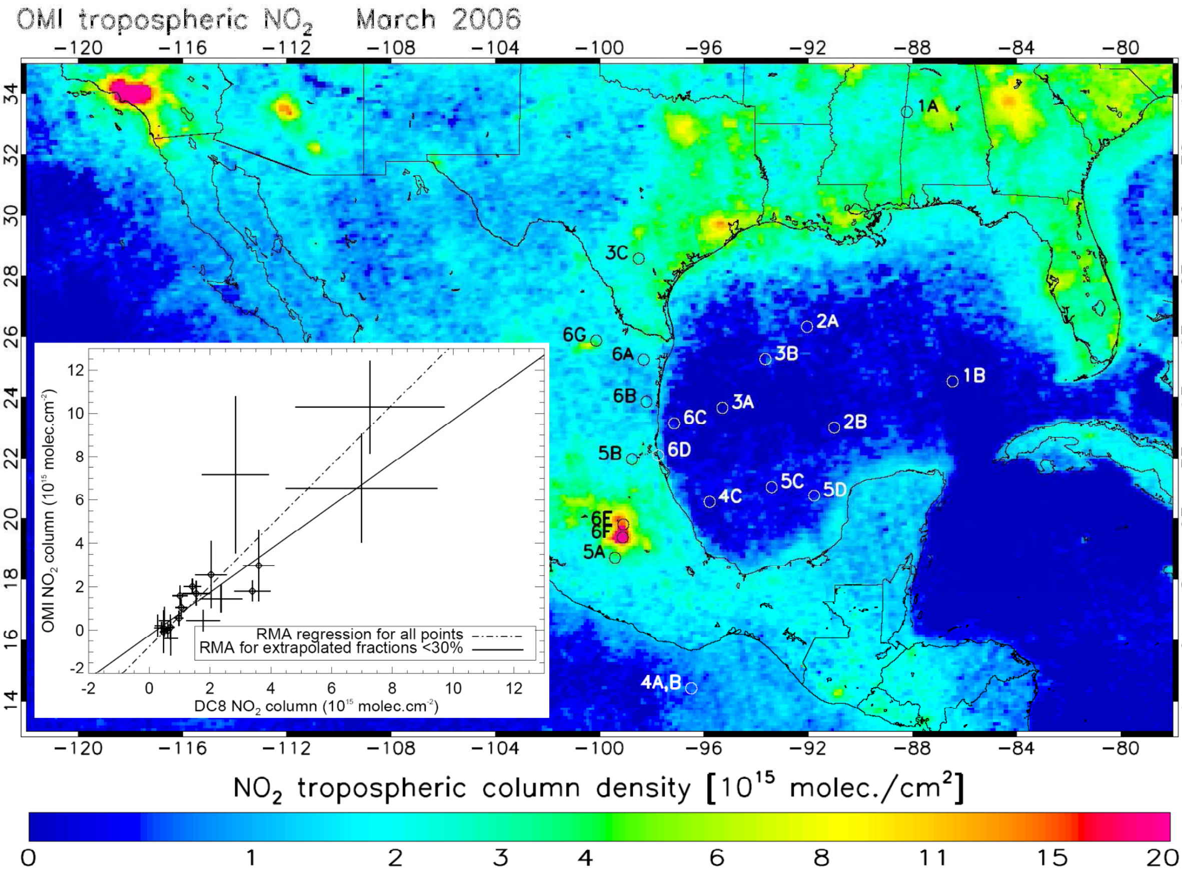
Dutch OMI NO_2 (DOMINO) data product

HE5 data file user manual

K. F. Boersma, R. J. Dirksen, J. P. Veefkind, H. J. Eskes, and R. J. van der A



29 April 2008

Contents

Version: 29 april 2008

What's new in this version of the Product Specification Document?

- Known issue with stratospheric averaging kernel values (sections 2.2 and 3.4).
- Updated information on the surface albedo used (section 3.4)
- Added Acknowledgments

Front cover figure: monthly mean tropospheric NO_2 column in March 2006 from OMI for cloud-free situations (cloud radiance fraction <50%). The circles indicate the location of aircraft profiles over sea (white) and land (black). Comparison of coincident OMI tropospheric NO_2 columns with those determined from in situ measurements from the DC-8 aircraft. Coincidences are labeled as in the left panel. Vertical bars are shown for illustration of the estimated OMI retrieval uncertainty, and horizontal bars for illustration of the estimated uncertainty of the DC-8 measurements determined from in situ measurement uncertainty and extrapolation assumptions. The dashed line represents the reduced major-axis regression through all data with a slope of 1.40 ± 0.21 (r^2 = 0.79). The solid line represents the reduced major axis (RMA) regression for situations when less than 30% of the DC-8 column has been extrapolated (diamonds) with a slope of 0.99 ± 0.17 (r^2=0.67). From Boersma et al. [2008].

1 Introduction

1.1 Purpose and data product

This document specifies the DOMINO (Dutch OMI NO_2) data product, version 1.0.1.

The DOMINO algorithm at KNMI has produced a 3+ years (October 2004 – today) set of OMI NO_2 data based on improved level-1b (ir)radiances. The product is available as data and images through www.temis.nl. For details on the DOMINO retrieval algorithm, please read Boersma et al. [2007].

The Dutch OMI NO_2 product is a post-processing data set, based on the most complete set of OMI orbits, improved level-1b (ir)radiance data (collection 3, Dobber et al. [2008]), analysed meteorological fields, and actual spacecraft data. The better data coverage, the improved calibration of level-1b data, and the use of analysed rather than forecast data make the Dutch OMI NO_2 product superior to the near-real time NO_2 data also available through www.temis.nl. The DOMINO product is the recommended product for scientific use.

1.2 Relation to GOME(-2) and SCIAMACHY NO_2 data formats

The GOME, GOME-2, and SCIAMACHY data are available as daily HDF4-files through www.temis.nl. In contrast, Dutch OMI NO_2 data, version 1.0.1 are now available in the orbital HDF-EOS5 (or HE5) format. The main reason for the transition from HDF4 to HDF-EOS5 is to bring the DOMINO product in line with all other OMI data products that are provided in the HDF-EOS5 format, at the expense of consistency with the GOME and SCIAMACHY heritage. Table 1 summarizes the differences between the HDF4 and HDFEOS5 formats.

Table 1. Overview of differences between KNMI satellite NO_2 products in the HDF4 and HDF-EOS 5 data formats.

HDF4	HDF-EOS 5
GOME(-2), SCIAMACHY	OMI v1.0.0 and OMI v1.0.1
Daily files	Orbital file[1]
1-dimensional structure (time-ordered, follows satellite ground track)	2-dimensional swath structure (time-ordered and identical to satellite ground track)

[1] For consistency with other TEMIS data products, orbital files are provided in a daily tar-file on www.temis.nl.

2 Product overview

2.1 DOMINO = Level 2 product

The DOMINO data contains geolocated column integrated NO_2 concentrations, or NO_2 columns (in units of molecules/cm^2). DOMINO data constitute a pure Level 2 product, i.e. it provides geophysical information for each and every ground pixel observed by the instrument, without the additional binning, averaging or gridding typically applied for Level 3 data. In addition to vertical NO_2 columns, the product contains intermediate results, such as the result of the spectral fit, fitting diagnostics, assimilated stratospheric NO_2 columns, the averaging kernel, cloud information, etc.

For advanced users, a second 'profile' file is made available that contains geolocated temperature and a priori NO_2 profiles at the exact pixel locations. Temperature and NO_2 profiles for each and every pixel are not included in the DOMINO product because most users will not need it and we want to keep the size of the DOMINO files reasonable. Nevertheless, the temperature and NO_2 profiles (from the TM4 chemistry-transport model) complete the a priori information used in the retrieval algorithm to compute the stratospheric NO_2 columns, the air mass factors, and the temperature correction [Boersma et al., 2007]. This product will be discussed in a separate document.

2.2 Version notes

This document applies to the Dutch OMI NO2 data product, version 1.0.1 The most important feature of version 1.0.1 is that it is based on collection 3 Level 1B data [Dobber et al., 2008]. Collection 3 Level 1B data are based on much improved instrument calibration parameters that lead to much less across-track variability, or stripes, in the OMI data products. It has therefore been decided to switch off the a posteriori stripe correction in the DOMINO retrieval algorithm. This stripe correction has been in use in pre-version 1.0 OMI near-real time products [Boersma et al., 2007].

Version 1.0.1 files suffer from one format-related issue. The averaging kernel values have been written as 16-bit integers, and need to be multiplied with a scale factor 0.0001, see Table 4 in section 3.4. Therefore the kernel elements in the file are limited to values less than +32676. Unfortunately, kernel values > 3.2676 occur on a regular basis, especially in the stratosphere and for large geometrical air mass factors. In these cases, the kernel elements have been written out with erroneous, negative values. For the time being, a simple fix can solve this issue. For negative kernel values, adding (2^{16}) 65356 will solve the problem as long as the original kernel values were smaller than 6.5356. Below find some pseudo-code that will temporarily solve the problem (for scenes with geometrical air mass factors less than approximately 6):

```
if (kernel_element < 0 and amfgeo < 5.5)
   kernel_element = kernel_element + 65356
endif
```

Note that this is a data format-related issue. The averaging kernels used in the v1.0.1 code are correct but have --for the cases mentioned above-- not been correctly written to file. We are currently addressing this issue in the soon-to-come v1.0.2.

2.3 Version 1.0.1 vs. 1.0.0

Version 1.0.0 data has been available on www.temis.nl between January 2008 and April 2008. The contents of versions 1.0.0 and 1.0.1 are identical, i.e. both are based on the same collection 3-based slant columns and have identical retrieval algorithms. The versions differ only in their data format. Version 1.0.1 furthermore corrects a minor problem[2] in version 1.0.0, but the most important change is that in Version 1.0.1 the averaging kernel has been included, and that some Data Fields have been written as integers instead of floats to save memory space. Table 2 summarizes the most important differences between Versions 1.0.0, and 1.0.1.

Table 2. Main differences between DOMINO Version 1.0.0 and 1.0.1 data files.

	Version 1.0.0	**Version 1.0.1**
Averaging kernel	Not included	Included
Surface albedo	int32	int16 + scaling by 0.0001
Cloud radiance fraction	int32	int16 + scaling by 0.01
TM4TerrainHeight	int32	int16
TM4TropoPauseLevel	int32	uint8

2.4 Product Identifier and file names

Although the OMI Science Support Team has not provided a dedicated identifier for the DOMINO product, we follow their convention as much as possible and use "OMDOMINO" for the global product. Similarly, we follow the file name convention specified in the HDF-EOS Aura File Format Guidelines [2003]. DOMINO file names will have 4 sections within the basis of the file name. Each section will be delimited by an *underscore*. The suffix will follow the basis and be delimited by a period. The four sections in the basis are Instrument ID, Data Type, Data ID and Version. Thus, the filename is constructed in the following way:

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

In Table 3 details the contents of the four sections and the suffix are given. The following is an example of a file name for the first orbit on 1 October 2004:

OMI-Aura_L2-OMDOMINO_2004m1001t0003-o01132_v003-2008m0324t184703.he5

This filename means measurement started on 1 October 2004, 00:03 UTC, orbit 1132, processed on 24 March 2008 at 18:47:03 UTC.

Table 3. Description of the different sections and the suffix of the file name.

Section	**Format**	**Description**
InstrumentID	"OMI-Aura"	ID for instrument and spacecraft
DataType	"DOMINO"	Product indicator
DataID	<start date and time>-o<orbit>	Date and orbit indicators: Date-time format: <yyyy>m<mmdd>t<hhmm> Orbit format: o<nnnnn>
Version	v<version>-<production date and time>	Version indicators: Version format <nnn> Date-time format: <yyyy>m<mmdd>t<hhmmss>
Suffix	"he5"	Suffix for product file

[2] Version 1.0.0 is not an entirely consistent set. From 26 September 2006 onwards, the Quality Flag has erroneously been set to -32767. The data can still be used, but the Quality Flag that indicates whether the Cloud Radiance Fraction exceeds 50% can not be trusted after the aforementioned date.

3 The Data File

3.1 Description and format

The OMI-Aura_L2-OMDOMINO_<yyyy>m<mmdd>t<hhmm>-o<nnnnn>_v003-<yyyy>m<mmdd>t<hhmmss>.he5 files contain data on NO_2 retrieved during one orbit. The format of the data file is HDF-EOS 5. To ease the use of EOS Aura data sets, the Aura teams have agreed to make their files match as closely as possible. To this end, the Aura teams have agreed on a set of guidelines for their file formats, as described in HDF-EOS Aura File Format Guidelines [2003].

The data file uses the HDF-EOS Swath format. Figure 1 shows an example of the structure of a DOMINO data file, when viewed using hdfview.

Figure 1. Structure of a DOMINO data file, when opened with hdfview (publicly available through http://hdf.ncsa.uiuc.edu/hdf-java-html/hdfview/).

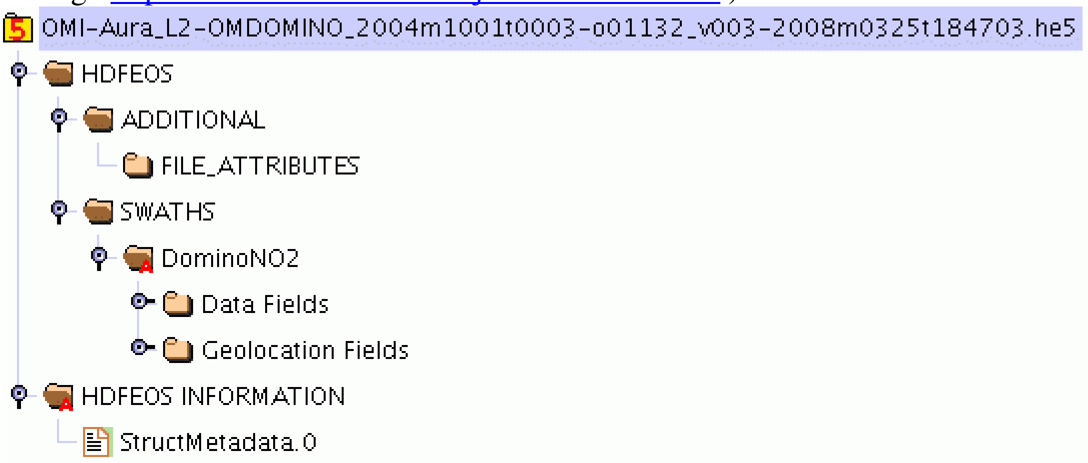


Figure 1 shows that the file contains a single swath structure named "DominoNO2". This is where all relevant retrieval data are stored. The swath structure consists of Data Fields and Geolocation Fields, but we start with StructMetadata.0, since this holds information on the size of the Data Fields and Geolocation Fields, that is being read in before the Data Fields and Geolocation Fields are read in.

3.2 StructMetadata.0

The most important information stored in StructMetadata.0 are the DIMENSIONS. For a DOMINO data file, there are four relevant DIMENSIONS. These pertain to the number of pixels across track (nXtrack), the number of measurements along track (nTimes), the number of corner points that specify the spatial extent of a pixel (nCornerpoints), and the number of pressure levels used in the air mass factor calculation (nPressureLevels). The contents of StructMeadata.0 are illustrated in Figure 2.

With the exception of nCornerpoints (always 4), the dimensions may differ between different files. For instance, nXtrack = 60 for nominal-model OMI measurements, but nXtrack = 30 for zoom mode measurements. nTimes is practically always 1644 (corresponding to 1644 2-s measurements along track). nPressureLevels = 35 for the period 1 October 2004 – 31 January 2006, and nPressureLevels = 34 from 1 February 2006 onwards. This change in the number of layers originates from a transition(from 60 to 91 layers) in the operational model ECMWF meteorological fields as of 1 February 2006. The 91 ECMWF layers are merged into 34 rather than 35 TM4 layers because this minimizes the need for interpolation.

Figure 2. Illustration of the main contents of the Structured Metadata (StructMetadata.0).

3.3 Attributes of the DominoNO2 Swath

An example of the Attributes of the DominoNO2 Swath is given in Figure 3. The Attributes are an important part of the file since they contain essential information on the versions of the retrieval algorithm and retrieval input data.

NumTimes refers to the number of OMI measurements taken along the track (same as nTimes in StructMetadata.0). Data_version refers to the version of the retrieval algorithm used to produce the DOMINO product. The subsequent date, 25 March 2008, is the date on which the file has been produced. Processing_mode refers to the retrieval mode and is always “Analysis” (i.e. not “Near-real time”) for the DOMINO v1.0.1 data. NO2_L2_file refers to the input file with NO_2 slant column and cloud information processed at the OMI SIPS in Greenbelt, Md, United States. METEO_DATA_X refers to the ECMWF meteorological input files used in the TM4 assimilation step that generates a stratospheric NO_2 slant column and a priori NO_2 profiles and a temperature correction.

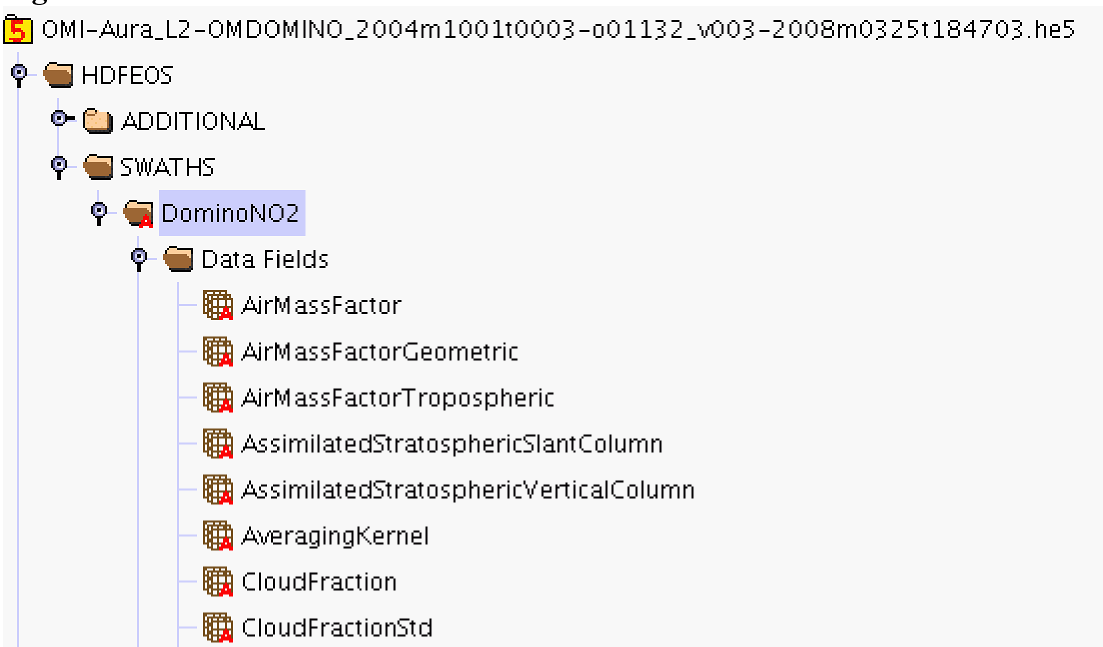
Figure 3. Attributes of the DominoNO2 swath.

Name	Value	Type	Array Size
NumTimes	1644	16-bit integer	1
Data_version	1.0.1 , 25 March 2008	String, length ...	1
PGE_version	1.0.0.10 , 25 March 2008	String, length ...	1
PGE_name	tm4no2a_omi	String, length ...	1
Authors	K.F. Boersma, R. Dirksen, H.J. Eskes, J.P. Veefkind	String, length ...	1
Email	dirksen@knmi.nl, boersma@knmi.nl, eskes@knmi.nl	String, length ...	1
Affiliation	KNMI (Royal Netherlands Meteorological Institute)	String, length ...	1
Processing_mode	Analysis	String, length ...	1
NO2_L2_file	OMI-Aura_L2A-OMNO2_2004m1001t0003-o01132...	String, length ...	1
AMF_LUT	no2_amf_lut.hdf Version 1.1	String, length ...	1
METEO_DATA_1	uvsp_20041001_21p06.hdf Version: 11	String, length ...	1
METEO_DATA_2	t_20041001_21p06.hdf Version: 10	String, length ...	1
METEO_DATA_3	q_20041001_21p06.hdf Version: 10	String, length ...	1
METEO_DATA_4	cld_20041001_21p06.hdf Version: 11	String, length ...	1
METEO_DATA_5	sub_20041001_21p06.hdf Version: 10	String, length ...	1
METEO_DATA_6	surf_20041001_21p03.hdf Version: 11	String, length ...	1

3.4 DominoNO2 Data Fields

The actual retrieved data in the DominoNO2 swath are found in Data Fields. Figure 4 shows the first couple of Data Fields to give an idea of the structure. In fact the swath holds 31 Data Fields, ordered in an alphabetical fashion.

Figure 4. Illustration of the first 8 Data Fields of the DominoNO2 swath.



All 31 Data Fields are summarized in Table 4. The information on a Data Field can also be found in the Attribute of the Data Field. These Attributes are important since they provide information on scale factors (needed to convert the values in meaningful numbers), the physical units of the field, the source of the information, and they provide interpretation for the missing data values, flags, etc.

Table 4. The Data Fields

Name	Type	Dimensions	Unit, scale factor	Description
AirMassFactor	32-bit floating point (HE5T_NATIVE_FLOAT)	nTimes × nXtrack	NoUnits	Total air mass factor used to compute the VCD (=SCD/AMF)
AirMassFactorGeometric	32-bit floating point (HE5T_NATIVE_FLOAT)	nTimes × nXtrack	NoUnits	Geometrical air mass factor (eq. (3) in Boersma et al. [2004])
AirMassFactorTropospheric	32-bit floating point (HE5T_NATIVE_FLOAT)	nTimes × nXtrack	NoUnits	Tropospheric air mass factor used to compute vcdtrop = [scd-scdstr]/amftrop)
AssimilatedStratosphericSlantColumn[3]	32-bit floating point (HE5T_NATIVE_FLOAT)	nTimes × nXtrack	molec.cm-2, 1e15	Assimilated stratospheric slant column as described in Boersma et al. [2007]
AssimilatedtratosphericVerticalColumn	32-bit floating point (HE5T_NATIVE_FLOAT)	nTimes × nXtrack	molec.cm-2, 1e15	Assimilated stratospheric vertical column as described in Boersma et al. [2007]

[3] The error on the AssimilatedStratosphericSlantColumn is estimated to be 0.25×10^{15} molec.cm^{-2} in all cases based on observation-forecast statistics, as discussed in Boersma et al. [2004, 2007].

AveragingKernel	16-bit integer (HE5T_NATIVE_INT16)	nLayer × nTimes × nXtrack	NoUnits, 0.0001	Averaging kernel as described in Eskes and Boersma [2003]
Cloud Fraction	16-bit integer (HE5T_NATIVE_INT16)	nTimes × nXtrack	NoUnits, 0.001	Effective cloud fraction as described in Acarreta et al. [2004]
CloudFractionStd	16-bit integer (HE5T_NATIVE_INT16)	nTimes × nXtrack	NoUnits, 0.001	Effective cloud fraction precision as described in Acarreta et al. [2004]
CloudPressure	16-bit integer (HE5T_NATIVE_INT16)	nTimes × nXtrack	hPa	Effective cloud pressure as described in Acarreta et al. [2004]
CloudPressureStd	16-bit integer (HE5T_NATIVE_INT16)	nTimes × nXtrack	hPa	Effective cloud pressure precision as described in Acarreta et al. [2004]
CloudRadianceFraction	16-bit integer (HE5T_NATIVE_INT16)	nTimes × nXtrack	NoUnits (%), 0.01	Cloud radiance fraction, see Eq. (14) in Boersma et al. [2004]
GhostColumn	32-bit floating point (HE5T_NATIVE_FLOAT)	nTimes × nXtrack	molec.cm-2, 1e15	TM4 vertical NO2 column between surface and effective cloud pressure, following the definition in Burrows et al. [1999]
InstrumentConfigurationId	8-bit unsigned character (HE5T_NATIVE_UINT8)	nTimes	NoUnit	Unique ID for instrument settings in current swath.
MeasurementQualityFlags	8-bit unsigned character (HE5T_NATIVE_UINT8)	nTimes	NoUnit	Bit level quality flags at measurement level. See Table AX.
SlantColumnAmountNO2	32-bit floating point (HE5T_NATIVE_FLOAT)	nTimes × nXtrack	molec.cm-2, 1e15	NO2 slant column from DOAS fit
SlantColumnAmountNO2Std	32-bit floating point (HE5T_NATIVE_FLOAT)	nTimes × nXtrack	molec.cm-2, 1e15	Precision of NO2 slant column from DOAS fit
SurfaceAlbedo	16-bit integer (HE5T_NATIVE_INT16)	nTimes × nXtrack	NoUnits, 0.0001	Surface albedo, described in Koelemeijer et al. [2003], as used in AMF calculation following Boersma et al. [2004, 2007].
TM4PressurelevelA	32-bit floating point (HE5T_NATIVE_FLOAT)	nLayer	Pa	Input for TM4 pressure levels, calculated as p = a + p_surf·b
TM4PressurelevelB	32-bit floating point (HE5T_NATIVE_FLOAT)	nLayer	NoUnit	Input for TM4 pressure levels, calculated as p = a + p_surf·b
TM4SurfacePressure	32-bit floating point (HE5T_NATIVE_FLOAT)	nTimes × nXtrack	hPa	TM4 surface pressure at the ground pixel center, as used in AMF calculation
TM4TerrainHeight	16-bit integer (HE5T_NATIVE_INT16)	nTimes × nXtrack	m	Surface elevation at ground pixel center, corresponding to the TM4 surface pressure
TM4TropoPauseLevel	8-bit unsigned character (HE5T_NATIVE_UINT8)	nTimes × nXtrack	NoUnit	TM4 level where tropopause occurs
TerrainHeight	16-bit integer (HE5T_NATIVE_INT16)	nTimes × nXtrack	m	Terrain height at ground pixel center from high-resolution database
TotalVerticalColumn	32-bit floating point (HE5T_NATIVE_FLOAT)	nTimes × nXtrack	molec.cm-2, 1e15	NO2 total vertical column (SCD/AMF)
TotalVerticalColumnError	32-bit floating point (HE5T_NATIVE_FLOAT)	nTimes × nXtrack	molec.cm-2, 1e15	Error in the NO2 total vertical column (SCD/AMF), following Boersma et al. [2004]

TroposphericColumnFlag	8-bit character (HE5_NATIVE_SCHAR)	nTimes × nXtrack	NoUnits	Flag to indicate when the retrieved tropospheric column is dominated by model information. See Appendix.
TroposphericVerticalColumn	32-bit floating point (HE5T_NATIVE_FLOAT)	nTimes × nXtrack	molec.cm-2, 1e15	NO2 tropospheric vertical column (SCD-SCDstrat)/AMFtrop
TroposphericVerticalColumnError	32-bit floating point (HE5T_NATIVE_FLOAT)	nTimes × nXtrack	molec.cm-2, 1e15	Error in the NO2 tropospheric vertical column (SCD/AMF), following Boersma et al. [2004]
TroposphericVerticalColumnModel	32-bit floating point (HE5T_NATIVE_FLOAT)	nTimes × nXtrack	molec.cm-2, 1e15	NO2 tropospheric vertical column according to TM4
VCDErrorUsingAvKernel	32-bit floating point (HE5T_NATIVE_FLOAT)	nTimes × nXtrack	molec.cm-2, 1e15	Error in NO2 total vertical column w/o profile error contribution.
VCDTropErrorUsingAvKernel	32-bit floating point (HE5T_NATIVE_FLOAT)	nTimes × nXtrack	molec.cm-2, 1e15	Error in NO2 tropospheric vertical column w/o profile error contribution.

Averaging kernel issue

Version 1.0.1 suffers from a data format issue for the averaging kernel field. For geometrical air mass factors larger than ~3.0, kernel elements in the data files may have negative values (see discussion in 2.2). These negatives values may be repaired by adding (2^{16}) 65356. This format issue, mainly occurring for stratospheric kernel elements with large geometrical amfs, will be resolved in v1.0.2.

Pressure grid

Every pixel has a unique 35-layer pressure grid that holds the 35 pressure levels that have been used to compute the averaging kernel. The equation to convert the TM4PressurelevelA, TM4PressurelevelB and TM4SurfacePressure into pressure levels (in Pascal) representative for the layering of the averaging kernel is:

$$p = a + p_{surf} \cdot b \qquad (1)$$

There is a change in TM4PressurelevelA, and TM4PressurelevelB per 1 February 2006, related to the transition in the number of TM4 layers, as discussed in Section 3.2. For detailed information on the actual meaning of the various flags, we refer to the appendix.

Surface albedo

The surface albedo reported in the data file is determined as follows. If an OMI pixel is situated within a grid cell of the Koelemeijer et al. [2003] database, it gets the corresponding value attributed (i.e. there is no spatial interpolation of the Koelemeijer et al. [2003] data base). Subsequently, there is interpolation in time: the center day of the month gets attributed the value given in the albedo database, but all other days are linear interpolations in time between the two nearest months. For instance the albedo on 17 January is determined as follows:

$$a_{sf} = w_1 \times a_{K01} + w_2 \times a_{K02} \qquad (2)$$

with w_1 = (29/29.5), and w_2 = (1/29.5), a_{K01}, and a_{K02} the values in the Koelemeijer et al. [2003] database and 29.5 being the number of days between January 16 and February 14.5, the centers of these months.

Whenever an OMI viewing scene contains snow or ice, this is detected based on the NISE [Nolin et al., 2005], and the albedo values from Koelemeijer et al. [2003] are being overwritten with specific values, i.e. 0.6 for snow over land. In such situations, cloud retrieval is attempted, but the retrieved cloud fraction and cloud pressure are dubious at best. Although the TroposphericColumnFlag is not raised in these situations, we do not yet have much confidence in the cloud values. We therefore recommend to handle scenes with a GroundPixelQualityFlag value that indicates the presence of snow or ice with utmost care. Most likely such scenes will be attributed a TroposphericColumnFlag value -1 in release v1.0.2.

3.5 Geolocation Fields

The geolocation fields are stored in the Geolocation Fields group of the DominoNO2 Swath. Table 6 gives a description of the Geolocation Fields.

Table 5. The geolocation fields

Name	Type	Dimensions	Unit, scale factor	Description
GroundPixelQualityFlag	16-bit integer (HE5T_NATIVE_INT16)	nTimes × nXtrack (1644 × 60)	NoUnits	See Table A3
Latitude	32-bit floating point (HE5T_NATIVE_FLOAT)	nTimes × nXtrack (1644 × 60)	Degrees	Latitude of groundpixel center
LatitudeCornerPoints	32-bit floating point (HE5T_NATIVE_FLOAT)	nCorner × nTimes × nXtrack (4 × 1644 × 60)	Degrees	Latitudes of the four corners of the ground pixel
Longitude	32-bit floating point (HE5T_NATIVE_FLOAT)	nTimes × nXtrack (1644 × 60)	Degrees	Longitude of groundpixel center
LongitudeCornerPoints	32-bit floating point (HE5T_NATIVE_FLOAT)	nCorner × nTimes × nXtrack (4 × 1644 × 60)	Degrees	Longitudes of the four corners of the ground pixel
SolarAzimuthAngle	32-bit floating point (HE5T_NATIVE_FLOAT)	nTimes × nXtrack (1644 × 60)	Degrees	Solar Azimuth Angle at WGS84 ellipsoid for center ground pixel, defined East-of-North
SolarZenithAngle	32-bit floating point (HE5T_NATIVE_FLOAT)	nTimes × nXtrack (1644 × 60)	Degrees	Solar Zenith Angle at WGS84 ellipsoid for center ground pixel.
Time	64-bit floating point (HE5T_NATIVE_DOUBLE)	nTimes (1644)	s	Time at start of scan (in TAI-93 format).
ViewingAzimuthAngle	32-bit floating point (HE5T_NATIVE_FLOAT)	nTimes × nXtrack (1644 × 60)	Degrees	Viewning azimuth angle at WGS84 ellipsoid for center ground pixel, defined East-of-North
ViewingZenithangle	32-bit floating point (HE5T_NATIVE_FLOAT)	nTimes × nXtrack (1644 × 60)	Degrees	Viewing Zenith Angle at WGS84 ellipsoid for center ground pixel.

Time is given in the TAI-93 format, i.e. the number of seconds passed since 01-01-1993, 00:00 UTC. WGS84 refers to the World Geodetic System 84, the commonly used reference frame for the earth dating from 1984.

4. Remarks on total vs. tropospheric NO_2 columns

The tropospheric NO_2 column is the principal DOMINO product. For historical reasons, an additional total NO_2 column is retrieved and stored in the Swath Data Fields. This total NO_2 column (TotalVerticalColumn) has been somewhat unfortunately defined as the ratio of the total slant column and the total air mass factor. For users interested in the actual total atmospheric column (integrated from the surface to the top-of-atmosphere), we strongly discourage the scientific use of TotalVerticalColumn. The reason for this is that a total air mass factor is too crude a metric to resolve the intricacies of tropospheric radiative transfer. As a matter of fact, the subtleties involved in accurate radiative transfer for species such as NO_2, concentrated in the boundary layer, are the very motivation for retrieval groups to explicitly separate the stratospheric background signal from the slant column before applying a pure tropospheric air mass factor.

Therefore, for users interested in the total NO_2 column, this quantity should be computed as the sum of the tropospheric and stratospheric vertical columns:

$$N_v = N_{v,tr} + N_{v.st} \tag{3}$$

i.e., by taking the sum of TroposphericVerticalColumn and the AssimilatedStratosphericVerticalColumn.

5. The use of the averaging kernel

Two distinct user groups can be distinguished: users that take our product ‘face value’, and more advanced users working on extensive scientific projects doing model-to-measurement comparisons and/or satellite validation studies.

(1) Basic users will be mainly interested in the tropospheric column TroposphericVerticalColumn and its error TroposphericVerticalColumnError, and/or the total vertical NO_2 column (as defined in Eq. (2)). These users may for instance want to qualitatively check preliminary results of some field experiment with the retrieved NO_2 columns.

(2) Advanced users may be interested in the relation between the (modelled or measured) 'true' vertical distribution of NO_2 and the retrieved quantity. These users will want to use the averaging kernel that provides the link between (modelled) reality and retrieval (for more details on the averaging kernel, read Eskes and Boersma [2003]. For example, those who are interested in a model – OMI comparison may want to map the modelled NO_2 profiles via the averaging kernel to what OMI would retrieve (y is the 'retrieved' quantity) as follows:

$$y = \mathbf{A} \cdot \mathbf{x} \tag{4}$$

with **A** the averaging kernel, a vector specified at nLayer pressure levels (sections 3.2, 3.4) and **x** the vertical distribution of NO_2 (in partial subcolumns) from a chemistry-transport model (or from collocated validation measurements) at the same nLayer pressure levels. The user thus needs to either convert his or her vertical (subcolumn) NO_2 profile to the pressure grid of the averaging kernel in order to construct a vertical column y as would be retrieved by OMI.

In principle, a user may also interpolate the averaging kernel vector to the grid of his or her **x**. However, since the averaging kernel is so sensitive to changes on small spatial scales, for instance due to rapid cloud changes, interpolation of the averaging kernel vector is discouraged.

Users will often be interested in the tropospheric NO_2 load. For tropospheric retrievals (with y now the tropospheric column), equation (3) reduces to:

$$y_{trop} = \mathbf{A}_{\mathbf{trop}} \cdot \mathbf{x}_{\mathbf{trop}} \tag{5}$$

with $\mathbf{A}_{\mathbf{trop}}$ the averaging kernel for tropospheric retrievals, defined as:

$$\mathbf{A}_{\mathbf{trop}} = \mathbf{A} \cdot \frac{AMF}{AMF_{trop}} \tag{6}$$

and $\mathbf{x}_{\mathbf{trop}}$ the profile shape for tropospheric levels (levels up to level number TM4TropoPauseLevel as specified in the DataField). The pressure at level TM4TropoPauseLevel does not necessarily correspond to the tropopause pressure but rather gives the pressure of the layer in which the tropopause occurs according to the WMO 1985 tropopause criterium.

For (tropospheric) applications using the averaging kernel, the error in y will reduce to VCDTropErrorUsingAvKernel since uncertainties on the a priori vertical NO_2 profile no longer contribute. A user should be aware that he or she should then no longer use VCDTropError, because this error includes the profile error term that can now be discarded.

Appendix

The TropColumnFlag is the most important error flag for users interested in the tropospheric NO_2 column. For users interested in other data products than the actually retrieved tropospheric column, the MeasurementQualityFlag (that applies to the slant column fitting), may be of interest. The GroundPixelQualityFlag does not represent an error flag, but merely provides interesting additional information on the viewing scene.

Table A1. Definition of the TropColumnFlag.

Value	Description
0	Tropospheric column for more than 50% determined by observed information.
-1	Tropospheric Column for more than 50% determined by forward model parameter assumptions.
-127	Missing data

One of the DominoNO2 Swath Data Fields is the MeasurementQualityFlag that relates to the slant column fitting. Table A2 summarizes the possible entries and their description.

Table A2. Definition of the MeasurementQualityFlags

Bit	Name	Description
0	Measurement Missing Flag	Set if all Ground Pixels give Earth Radiance Missing Flag.
1	Measurement Error Flag	Set if any of the L1B MeasurementQualityFlags bit 0, 1or 3 are set for the Radiance or for the used Solar product.
2	Measurement Warning Flag	Set if any of the L1B MeasurementQualityFlags bit 0, 1, or 3 are set for the Radiance or for the used Solar product.
3	Rebinned Measurement Flag	Set if L1B radiance MeasurmentQualityFlags bit 7 is set to 1.
4	4 SAA Flag	Set if L1B MeasurmentQualityFlags bit 10 is set to 1, for the Radiance or for the used Solar product
5	Spacecraft Maneuver Flag	Set if L1B MeasurmentQualityFlags bit 11 is set to 1, for the Radiance or for the used Solar product
6	Instrument Settings Error Flag	The Earth and Solar InstrumentConfigurationIDs are not compatible.
7	Cloud Data Not Synchronized Flag	Set if radiance anc cloud data are not synchronized

The GroundPixelQualityFlag provides information on the viewing scene. This additional information is stored as a 16-bit integer, whose meaning can be retrieved with dedicated software that will be provided on www.temis.nl. Below are two examples of how the GroundPixelQualityFlag should be interpreted:

65535 = fill value/missing data (all bits have been set)
25857 = Greenland (0110 | 0101 | 0000 | 0001)

Here, bits 0-3 are 0001 representing a numerical value of 1 (2^0 is set, 2^1, 2^2, 2^3 are not set), i.e. Land. Bits 8-14 are 0110 0101, representing a numerical value of 101 ($2^0+2^2+2^5+2^6=101$) i.e. Permanent Ice .

Table A3. Definition of the GroundPixelQualityFlag.

Bit	Description
0-3	Land/Water flags 0=Shallow Ocean 1=Land 2=Shallow Inland Water 3=Ocean coastline/Lake shoreline 4=Ephemeral (intermittent) water 5=Deep Inland Water 6=Continental Shelf Ocean 7=Deep Ocean 15=Error flag for Land/Water

4	Sun Glint possibility flag
5	Solar Eclipse possibility flag
6	Geolocation Error flag
8-14	Snow/Ice flags [based on NISE] 0=Snow-free land 1-100=Sea ice percentage (%) 101=Permanent ice (Greenland, Antarctica) 103=Dry snow 104=Ocean 124=Mixed pixels at coastline 125=Suspect ice value 126=Corners (undefined) 127=Error
15	NISE nearest neighbour filling flag 0=Not set 1=Set

InstrumentConfigurationID. Table A4 summarizes common Instrument Configurations. Instrument Configurations 0-49 are 'regular operations', encountered since the start of the OMI Nominal Operations Baseline on 8 October 2004. Before that date, OMI was in the Launch and Early Orbit operations phase, where special calibration measurements have been carried out. These special measurements with InstrumentConfigurationID values between 118-140, are not encountered during the OMI Nominal Operations Baseline.

Table A4. Overview of possible values for InstrumentConfigurationID and their description.

Instrument Configuration	**Description**
0	Global Tropical
1	Global Midlatitude
2	Global Arctic
7	Global Ozone Hole
42	Spatial Tropical
43	Spatial Midlatitude
44	Spatial Arctic
49	Spatial Ozone Hole
118	Central Tropical
120	Central Midlatitude
122	Central Arctic
124	Central Ozone Hole
126	Left Tropical
128	Right Tropical
130	Left Midlatitude
132	Right Midlatitude
134	Left Arctic
136	Right Arctic
138	Left Ozone Hole
140	Right Ozone Hole

Acknowledgments

We are thankful for very useful user feedback provided by Yipin Zhou (EMPA, Switzerland) and Achim Strunk (University of Cologne, Germany).

References

Acarreta, J. R., J. F. De Haan, and P. Stammes (2004), Cloud pressure retrieval using the O_2-O_2 absorption band at 477 nm, *J. Geophys. Res.*, *109*, D05204, doi:10.1029/2003JD003915.

Boersma, K. F., H. J. Eskes, and E. J. Brinksma (2004), Error analysis for tropospheric NO2 retrieval from space, *J. Geophys. Res., 109*, D04311, doi:10.1029/2003JD003962.

Boersma, K. F., H. J. Eskes, J. P. Veefkind, E. J. Brinksma, R. J. van der A, M. Sneep, G. H. J. van den Oord, P. F. Levelt, P. Stammes, J. F. Gleason, and E. J. Bucsela (2007), Near-real time retrieval of tropospheric NO_2 from OMI, *Atmos. Chem. Phys., 7*, 2103-2118.

Boersma, K. F., D. J. Jacob, E. J. Bucsela, A. E. Perring, R. Dirksen, R. J. van der A, R. M. Yantosca, R. J. Park, M. O. Wenig, T. H. Bertram, and R. C. Cohen (2008), Validation of OMI tropospheric NO_2 observations during INTEX-B and application to constrain NO_x emissions over the eastern United States and Mexico, *Atmos. Environm.*, doi:10.1026/j.atmosenv.2008.02.004, in press.

Burrows, J. P., et al. (1999), The Global Ozone Monitoring Experiment (GOME): Mission concept and first scientific results*, J. Atmos. Sci., 56*, 151-175.

Dobber, M., Q. Kleipool, R. Dirksen, P. Levelt, G. Jaross, S. Taylor, T. Kelly, L. Flynn, G. Leppelmeier, and N. Rozemeijer (2008), Validation of Ozone Monitoring Instrument level-1b data products, *J. Geophys. Res.*, doi:10.1029/2007JD008665, in press.

Eskes, H. J., and K. F. Boersma (2003), Averaging kernels for DOAS total-column satellite retrievals, *Atmos. Chem. Phys., 3*, 1285-1291.

HDF-EOS Aura File Format Guidelines, NCAR SW-NCA-079, Version 1.3, 27 August 2003.

Koelemeijer, R. B. A., J. F. de Haan, and P. Stammes (2003), A database of spectral surface reflectivity in the range 335-772 nm derived from 5.5 years of GOME measurements, J. Geophys. Res., 108(D2), 4070, doi:10.1029/2002JD002429.

Nolin, A., R. Armstrong, and J. Maslanik (2005), Near-real time SSM/I EASE grid daily global ice concentration and snow extent, Boulder, CO, USA: National Snow and Ice Data Center. Digital Media, 2005, updated daily.