Royal Netherlands
Meteorological Institute
Ministry of Infrastructure and the Environment

TROPOMI ATBD of the total and tropospheric NO_2 data products

TROPOMI

sentinel-5p

document number : S5P-KNMI-L2-0005-RP
authors : J.H.G.M. van Geffen, K.F. Boersma, H.J. Eskes, J.D. Maasakkers and J.P. Veefkind
CI identification : CI-7430-ATBD
issue : 1.0.0
date : 2016-02-05
status : released

Document approval record

	digital signature
Prepared:	Jos van Geffen Digitally signed by Jos van Geffen DN: c=NL, st=Utrecht, o=Ministry of Infrastructure and the Environment, ou=Royal Netherlands Meteorological Institute, cn=Jos van Geffen, email=geffen@knmi.nl Date: 2016.02.03 12:07:53 +01'00'
Checked:	S. Beirle, A. Hilboll, A. Richter
Approved PM:	Arnoud Apituley Digitally signed by Arnoud Apituley Date: 2016.02.03 14:40:56 +01'00'
Approved PI:	J. Pepijn Veefkind Digitally signed by J. Pepijn Veefkind DN: c=NL, st=Utrecht, o=Ministry of Infrastructure and the Environment, ou=Royal Netherlands Meteorological Institute, cn=J. Pepijn Veefkind, email=veefkind@knmi.nl Date: 2016.02.17 11:45:46 +01'00'

Document change record

issue	date	item	comments
0.0.1	2012-08-17	All	Initial draft version
0.0.2	2012-09-12	All	Major reordering, adding text and references throughout
0.0.3	2012-09-26	6.1–6.2 7.1–7.2	Several small corrections and additions
0.1.0	2012-09-26	—	First official release
0.2.0	2012-11-15	5.1–5.4 6.1–6.2	Large number of updates throughout the text after internal reviewing by the TROPOMI Level-2 Working Group
0.2.1	2013-04-10	5, 6, 7	Document number corrected, reorganisation of Sect. 6; removal of the discussion on an alternative retrieval approach; various minor corrections, updates and additions
0.2.2	2013-06-03	5, 6, 7 8, 9, 10	Major updates and further reorganisation of Sect. 5–7 first versions of Sect. 8–10
0.3.0	2013-06-04	—	Release for Level-2 Working Group review
0.3.1	2013-06-19	5.1–5.3 6.2–6.6 7.3–7.5 8.4, 9.2	Correction and additions resulting from v0.3.0 internal Level-2 Working Group review, and other minor corrections and additions
0.5.0	2013-06-20	—	Release for external review
0.5.1	2013-09-05	5.1–5.2 6.1–6.6 7.2, 7.5	First round of updates taking into account comments and suggestions of the reviewers of version 0.5.0
0.5.2	2013-11-21	5.5 6.2–6.7 7.1, 7.2 8.2, 8.3	Further updates in view of the reviewers comments, and providing more details regarding the fit procedure and the processing chain
0.9.0	2013-11-27	—	Release to ESA and external reviewers
0.9.1	2014-03-17	4.	Instrument overview section now in separate document; section numbering of this document unchanged to maintain change record
0.9.2	2014-04-07	6.2	Some small typographic corrections and updates
0.10.0	2014-04-15	—	Release to ESA
0.10.1	2014-09-19	6.1–6.6 7.1–7.4	Minor updates and corrections in text and tables; old Sects. 6.5 and 6.7 combined; Sects. 6.6 updated; references updated. Descriptions updated and tables of input and output data added. Notation of variables improved or clarified in view of the IODD [RD1].
0.10.2	2014-09-25	6.6	Table added with overview user applications and data the users need
0.11.0	2014-10-02	—	Release to ESA
0.11.1	2015-08-27	6.4.1 6.4.4.1 6.6, 7.4	Update of AMF look-up table entries Update text regarding using cloud fraction from NO_2 spectral window Minor updates in product data set tables, incl. dataset units
0.13.0	2015-09-14	—	Limited release to S5P Validation Team
0.13.1	2015-12-08	6.2, 7.5 7.1	Minor textual corrections Correction of units of input datasets
0.14.0	2015-12-11	—	Release to ESA
1.0.0	2016-02-05	—	Public release

Contents

List of Tables

List of Figures

1 Introduction to the document

1.1 Identification

This document, identified as S5P-KNMI-L2-0005-RP, is the Algorithm Theoretical Basis Document (ATBD) for the TROPOMI total and tropospheric NO_2 data products. It is part of a series of ATBDs describing the TROPOMI Level-2 data products.

1.2 Purpose and objective

The purpose of this document is to describe the theoretical basis and the implementation of the NO_2 Level-2 algorithm for TROPOMI. The document is maintained during the development phase and the lifetime of the data products. Updates and new versions will be issued in case of changes of the algorithms.

1.3 Document overview

Sections 2 and 3 list the applicable and reference documents and the terms and abbriviations specific for this document; references to peer-reviewed papers and other scientific publications are listed in Section 11. Section 4 gives a general description of the TROPOMI instrument, which is common to all ATBDs of the TROPOMI Level-2 data products. Section 5 provides an introduction to the NO_2 data products, their heritage, the set-up of their retrieval, the requirements of the products, and their availability. Section 6 gives an overview of the TROPOMI NO_2 data processing system and important aspects of the various steps in the processing. Section 7 lists some aspects regarding the feasibility of the NO_2 data products, such as the computational effort and the auxiliary information needed for the processing. Section 8 deals with an error analysis of the NO_2 data product. Section 9 gives a brief overview of validation issues and possibilities, such as campaigns and satellite intercomparions. Section 10 formulates some conclusion regarding the NO_2 data products.

1.4 Acknowledgements

The authors would like to thank the following people for useful discussions, information, reviews of earlier versions of this document and other contributions: Andreas Hilboll, Andreas Richter, Bram Sanders, Deborah Stein – Zweers, Dominique Brunner, Huan Yu, Isabelle De Smedt, Jason Williams Johan de Haan, Maarten Sneep, Mark ter Linden, Michel Van Roozendael, Piet Stammes, Pieter Valks, Ronald van der A, Steffen Beirle, Thomas Wagner.

2 Applicable and reference documents

2.1 Applicable documents

[AD1] Level 2 Processor Development – Statement of Work.
source: ESA/ESTEC; **ref:** S5P-SW-ESA-GS-053; **issue:** 1.1; **date:** 2012-05-21.

[AD2] GMES Sentinel-5 Precursor – S5p System Requirement Document (SRD).
source: ESA/ESTEC; **ref:** S5p-RS-ESA-SY-0002; **issue:** 4.1; **date:** 2011-04-xx.

[AD3] NO2 PGE Detailed Processing Model.
source: Space Sytems Finland; **ref:** TN-NO2-0200-SSF-001; **issue:** 1.2; **date:** 2010-04-21.

2.2 Standard documents

There are no standard documents

2.3 Reference documents

[RD1] Sentinel 5 precursor/TROPOMI KNMI and SRON level 2 Input Output Data Definition.
source: KNMI; **ref:** S5P-KNMI-L2-0009-SD; **issue:** 4.0.0; **date:** 2015-11-02.

[RD2] Terms, definitions and abbreviations for TROPOMI L01b data processor.
source: KNMI; **ref:** S5P-KNMI-L01B-0004-LI; **issue:** 3.0.0; **date:** 2013-11-08.

[RD3] Terms and symbols in the TROPOMI Algorithm Team.
source: KNMI; **ref:** S5P-KNMI-L2-0049-MA; **issue:** 1.0.0; **date:** 2015-07-16.

[RD4] TROPOMI Instrument and Performance Overview.
source: KNMI; **ref:** S5P-KNMI-L2-0010-RP; **issue:** 0.10.0; **date:** 2014-03-15.

[RD5] GMES Sentinels 4 and 5 Mission Requirements Document.
source: ESA/ESTEC; **ref:** EOP-SMA/1507/JL-dr; **issue:** 3; **date:** 2011-09-21.

[RD6] Science Requirements Document for TROPOMI. Volume I: Mission and Science Objectives and Observational Requirements.
source: KNMI, SRON; **ref:** RS-TROPOMI-KNMI-017; **issue:** 2.0.0; **date:** 2008-10-30.

[RD7] CAPACITY: Operational Atmospheric Chemistry Monitoring Missions – Final report and technical notes of the ESA study.
source: KNMI; **ref:** CAPACITY; **date:** Oct. 2005.

[RD8] CAMELOT: Observation Techniques and Mission Concepts for Atmospheric Chemistry – Final report of the ESA study.
source: KNMI; **ref:** RP-CAM-KNMI-050; **date:** Nov. 2009.

[RD9] TRAQ: Performance Analysis and Requirements Consolidation – Final report of the ESA study.
source: KNMI; **ref:** RP-ONTRAQ-KNMI-051; **date:** Jan. 2010.

[RD10] Algorithm theoretical basis document for the TROPOMI L01b data processor.
source: KNMI; **ref:** S5P-KNMI-L01B-0009-SD; **issue:** 6.0.0; **date:** 2015-09-22.

[RD11] S5P/TROPOMI Static input for Level 2 processors.
source: KNMI/SRON/BIRA/DLR; **ref:** S5P-KNMI-L2CO-0004-SD; **issue:** 3.0.0; **date:** 2015-02-27.

[RD12] Cloud retrieval algorithm for GOME-2: FRESCO+.
source: EUMETSAT/KNMI; **ref:** EUM/CO/09/4600000655/RM; **issue:** 1.3; **date:** 2010-10-18.

[RD13] Determine the effective cloud fraction for a specific wavelength.
source: KNMI; **ref:** S5P-KNMI-L2-0115-TN; **issue:** 1.0.0; **date:** 2013-10-15.

[RD14] S5P/TROPOMI ATBD Cloud Products.
source: DLR; **ref:** S5P-DLR-L2-ATBD-400I; **issue:** 0.11.0; **date:** 2014-09-30.

[RD15] Quality Assurance for Essential Climate Variables (QA4ECV).
source: KNMI; **ref:** EU-project 607405, SPA.2013.1.1-03; **date:** November 2012.

[RD16] Dutch OMI NO_2 (DOMINO) data product v2.0 – see http://www.temis.nl/airpollution/no2.html.
source: KNMI; **ref:** OMI_NO2_HE5_2.0_2011; **date:** 18 August 2011.

2.4 Electronic references

[ER1] TEMIS website: NO2 data product page. URL `http://www.temis.nl/airpollution/no2.html`.

[ER2] Vandaele et al. NO2 cross sections. URL `http://spectrolab.aeronomie.be/no2.htm`.

[ER3] TM5 website. URL `http://www.projects.science.uu.nl/tm5/`.

[ER4] MERIS-based albedo climatologies. URL `http://www.temis.nl/data/meris.html`.

[ER5] Q. L. Kleipool, M. R. Dobber, J. F. De Haan *et al.*; OMI Surface Reflectance Climatology (2010). URL `http://disc.sci.gsfc.nasa.gov/Aura/data-holdings/OMI/omler_v003.shtml`.

[ER6] EUMETSAT Ocean & Sea Ice Satellite Application Facility. Updated daily; URL `http://osisaf.met.no/`.

[ER7] A. Nolin, R.L. Armstrong and J. Maslanik; 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; URL `http://nsidc.org/data/docs/daac/nise1_nise.gd.html`.

[ER8] SPEC CPU2006 Results Published by SPEC. URL `http://www.spec.org/cpu2006/results/cpu2006.html`.

[ER9] MACC website. URL `http://www.gmes-atmosphere.eu/`.

3 Terms, definitions and abbreviated terms

Terms, definitions and abbreviated terms that are used in development program for the TROPOMI L0-1b data processor are described in [RD2]. Terms, definitions and abbreviated terms that are used in development program for the TROPOMI L2 data processors are described in [RD3]. Terms, definitions and abbreviated terms that are specific for this document can be found below.

3.1 Terms and definitions

The most important symbols related to the data product described in this document – some of which are not in [RD3] – are the following; see also the data product overview list in Table 10.

M	total air-mass factor
M_{cl}	cloudy air-mass factor
M_{cr}	clear-sky air-mass factor
M^{trop}	tropospheric air-mass factor
M^{strat}	stratospheric air-mass factor
N_{s}	total slant column density
$N_{\mathrm{s}}^{\mathrm{trop}}$	tropospheric slant column density
$N_{\mathrm{s}}^{\mathrm{strat}}$	stratospheric slant column density
N_{v}	total vertical column density
$N_{\mathrm{v}}^{\mathrm{trop}}$	tropospheric vertical column density
$N_{\mathrm{v}}^{\mathrm{strat}}$	stratospheric vertical column density
$N_{\mathrm{v}}^{\mathrm{sum}}$	sum of tropospheric and stratospheric vertical column density

3.2 Acronyms and abbreviations

AAI	Absorbing Aerosol Index
ACE	Atmospheric Chemistry Experiment
AMF	Air-mass factor
CTM	Chemistry Transport Model
DOAS	Differential Optical Absorption Spectroscopy
DOMINO	Dutch OMI NO_2 data products of KNMI for OMI
ECMWF	European Centre for Medium-Range Weather Forecast
ENVISAT	Environmental Satellite
EOS-Aura	Earth Observing System (Chemistry & Climate Mission)
ERBS	Earth Radiation Budget Satellite
ERS	European Remote Sensing satellite
FRESCO	Fast Retrieval Scheme for Clouds from the Oxygen A band
GOME	Global Ozone Monitoring Experiment
HALOE	Halogen Occultation Experiment
IPA	Independent pixel approximation
LUT	Look-up table
MACC	Monitoring Atmospheric Composition and Climate
MERIS	Medium Resolution Imaging Spectrometer
MetOp	Meteorological Operational Satellite
NISE	Near-real-time Ice and Snow Extent
NRT	near-real time (i.e. processing within 3 hours of measurement)
OE	Optimal Estimation
OMI	Ozone Monitoring Instrument
OMNO2A	OMI NO2 slant column data product (at NASA)

OSIRIS	Optical Sspectrograph and Infrared Imager System
OSISAF	Ocean & Sea Ice Satellite Application Facility
PDGS	Sentinel-5Precursor Payload Data Ground Segment (at DLR)
POAM	Polar Ozone and Aerosol Measurements
SAGE	Stratospheric Gas and Aerosol Experiment
SCIAMACHY	Scanning Imaging Absorption Spectrometer for Atmospheric Cartography
SDC	Satellite Data Centre (at KNMI)
SME	Solar Mesosohere Explorer
SNR	Signal-to-Noise Ratio
SPOT	Système Pour l'Observation la Terre
TM4, TM5	Data assimilation / chemistry transport model (version 4 or 5)
TM4NO2A	NO_2 data products of KNMI for GOME, SCIAMACHY and GOME-2
TROPOMI	Tropospheric Monitoring Instrument
UARS	Upper Atmosphere Research Satellite

4 TROPOMI instrument description

A description of the TROPOMI instrument and performance, referred to from all ATBDs, can be found in [RD4].

5 Introduction to the TROPOMI NO_2 data products

5.1 Nitrogen dioxide in troposphere and stratosphere

Nitrogen dioxide (NO_2) and nitrogen oxide (NO) – together usually referred to as nitrogen oxides (NO_x= NO +NO_2) – are important trace gases in the Earth's atmosphere, present in both the troposphere and the stratosphere. They enter the atmosphere as a result of anthropogenic activities (notably fossil fuel combustion and biomass burning) and natural processes (such as microbiological processes in soils, wildfires and lightning). Approxiately 95% of the NO_x emissions is in the form of NO. During daytime, i.e. in the presence of sunlight, a photochemical cycle involving ozone (O_3) converts NO into NO_2 (and vice versa) on a timescale of minutes, so that NO_2 is a robust measure for concentrations of nitrogen oxides (Solomon [1999], Jacob [1999]).

In the troposphere NO_2 plays a key role in air quality issues, as it directly affects human health [World Health Organisation, 2003]. In addition nitrogen oxides are essential precursors for the formation of ozone in the troposphere (e.g. Sillman et al. [1990]) and they influence concentrations of OH and thereby (shorten) the lifetime of methane (CH_4) (e.g. Fuglestvedt et al. [1999]). Although NO_2 is a minor greenhouse gas in itself, the indirect effects of NO_2 on global climate change are probably larger, with a presumed net cooling effect mostly driven by a growth in aerosol concentrations through nitrate formation from nitrogen oxides and enhanced levels of oxidants (e.g. Shindell et al. [2009]). Deposition of nitrogen is of great importance for eutrification [Dentener et al., 2006], the response of the ecosystem to the addition of substances such as nitrates and phosphates – negative environmental effects include the depletion of oxygen in the water, which induces reductions in fish and other animal populations.

For typical levels of OH the lifetime of NO_x in the lower troposphere is less than a day. For Riyadh, for example, Beirle et al. [2011] find a lifetime of about 4.0 ± 0.4 hours, while at higher latitudes (e.g. Moscow) the lifetime can be considerably longer, up to 8 hour, in winter, because of a slower photochemistry in that season. For Switzerland Schaub et al. [2007] report lifetimes of 3.6 ± 0.8 hours in summer and $13.1 \pm (3.8)$ hours in winter. With lifetimes in the troposphere of only a few hours, the NO_2 will remain relatively close to its source, making it well detectable from space. As an example, Fig. 1 shows distinct hotspots of NO_2 pollution over the highly industrialised and urbanised regions of London, Rotterdam and the Ruhr area in the monthly average tropospheric NO_2 for June 2005 over Europe derived from OMI data.

In the stratosphere NO_2 is involved in some photochemical reactions with ozone and thus affects the ozone

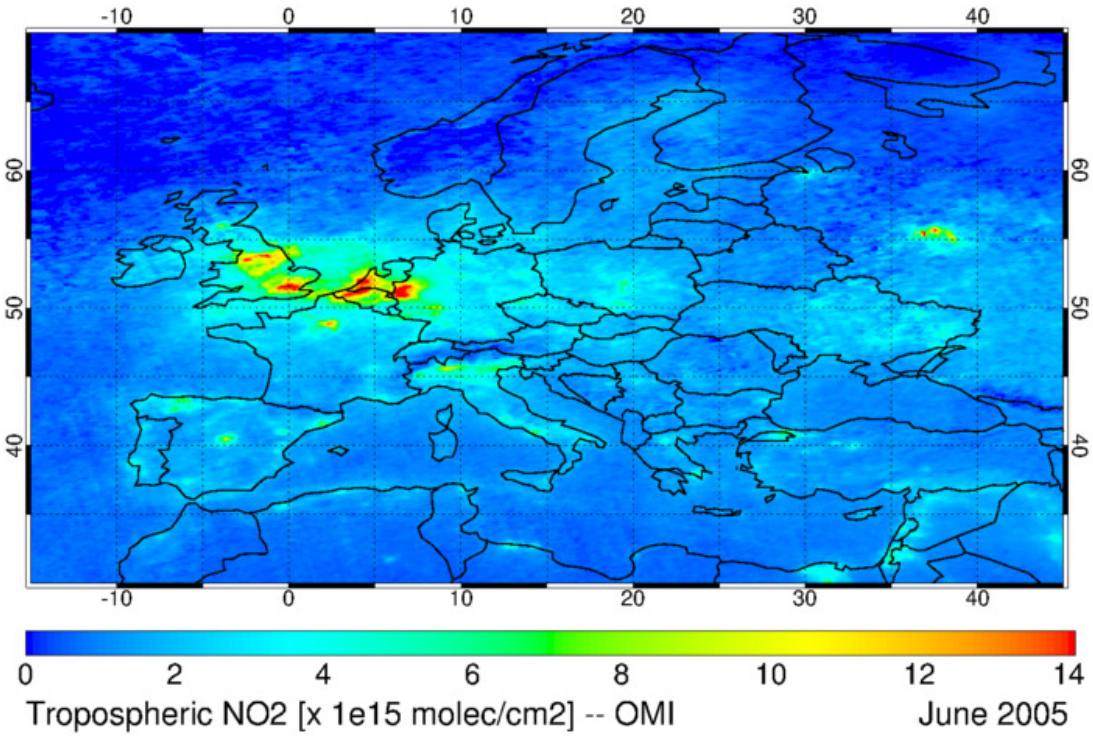

Figure 1: Monthly average distribution of tropospheric NO_2 columns for June 2005 over Europe based on OMI data, as derived by the DOMINO processing. (Data source: TEMIS website [ER1].)

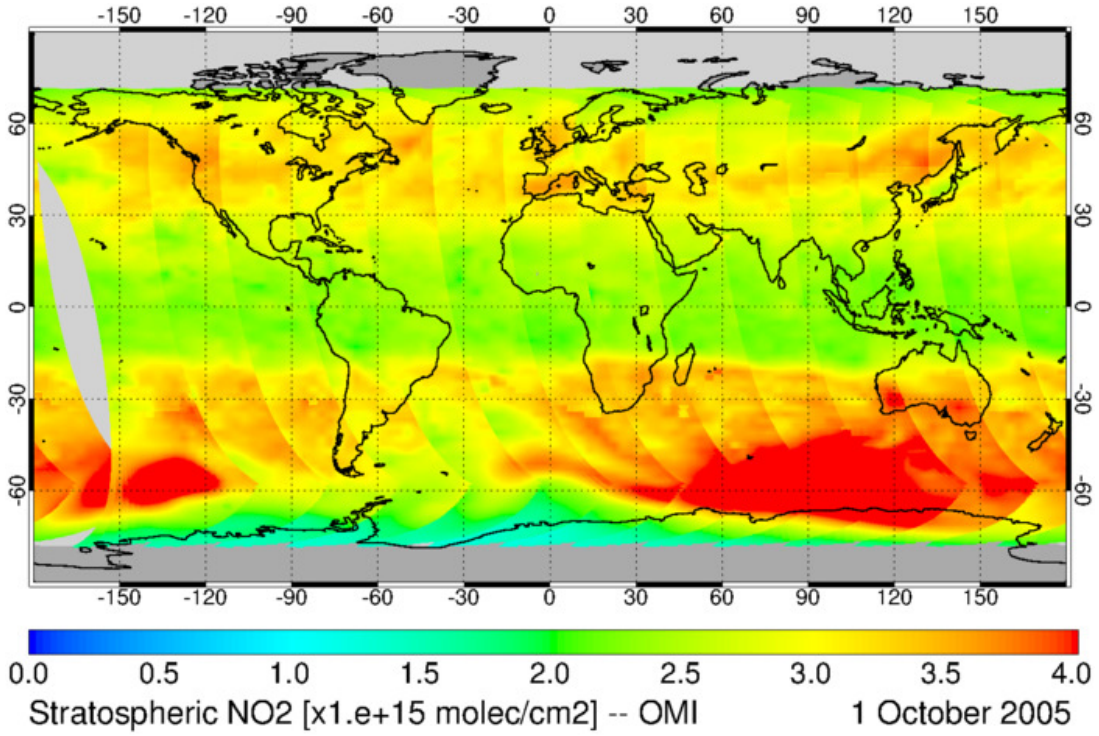

Figure 2: Distribution of stratospheric NO_2 on 1 October 2005 along the individual OMI orbits, as derived by the DOMINO processing. Note that the colour scale range is different from range in Fig. 1. (Data source: TEMIS website [ER1].)

layer (Crutzen et al. [1970]; Seinfeld and Pandis [2006]). The origin of NO_2 in the stratosphere is mainly from oxidation of N_2O in the middle stratosphere, which leads to NO_x, which in turn acts as a catalyst for ozone destruction (Crutzen et al. [1970]; Hendrick et al. [2012]). But NO_x can also suppress ozone depletion by converting reactive chlorine and hydrogen compounds into unreactive reservoir species (such as $ClONO_2$ and HNO_3; Murphy et al. [1993]).

Figure 2 shows, as an example, the stratospheric NO_2 distribution derived from OMI measurements on 1 October 2005. In a study into the record ozone loss, triggerd by enhanced NO_x levels, in the exceptionally strong Arctic polar vortex in Spring 2011, Adams et al. [2013] showed the usefulness of such data when investigating the anomalous dynamics and chemistry in the stratosphere. With its higher spatial resolution and signal-to-noise ratio, TROPOMI will clearly be well-suited to help understand the stratospheric NO_2 content and its implications for the ozone distribution.

From observed trends in N_2O emissions one would expect a trend in stratospheric NO_2 with potential implications for persistent ozone depletion well into the 21st century [Ravishankara et al., 2009]. There have been some reports of such trends in stratospheric NO_2, for instance from New Zealand [Liley et al., 2000] and northern Russia [Gruzdev and Elokhov, 2009]. On the other hand, Hendrick et al. [2012] report that changes in the NO_x partitioning in favour of NO may well conceal the effect of trends in N_2O. TROPOMI will continue the important record of stratospheric NO_2 observations that started with GOME in 1995, and improve the detectability of trends.

Over unpolluted regions most NO_2 is located in the stratosphere (typically more than 90%). For polluted regions 50–90% of the NO_2 is located in the troposphere, depending on the degree of pollution. Over polluted regions, most of the tropospheric NO_2 is found in the planetary boundary layer, as has been shown among others in campaigns using measurements made from aeroplanes, such as INTEX (e.g. Hains et al. [2010]). In areas with strong convection, enhanced NO_2 concentrations are observed at higher altitudes due to production of NO_x by lightning (e.g. Ott et al. [2010]).

The important role of NO_2 in both troposphere and stratosphere implies that it is not only important to know the total column density of NO_2, but rather the tropospheric NO_2 and stratospheric NO_2 concentrations separately. A proper separation between the two is therefore important, in particular for areas with low pollution, where the stratospheric concentration forms a significant part of the total column.

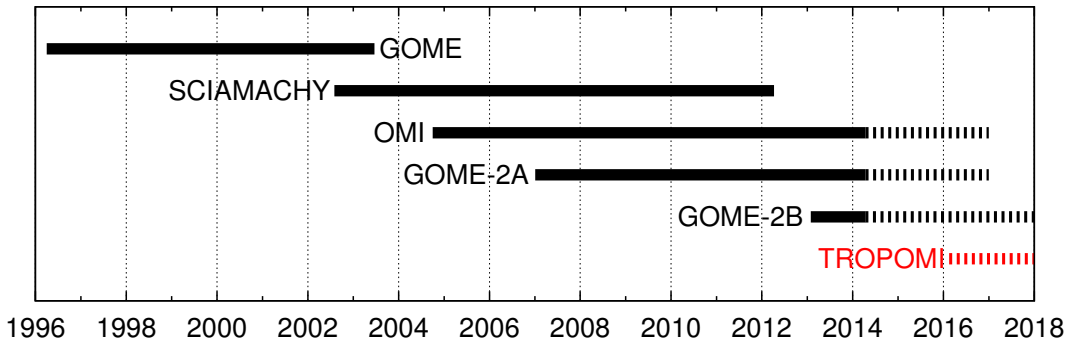

Figure 3: Overview of the UV/Vis backscatter satellite instruments capable of retrieving tropospheric and stratospheric NO_2 columns.

5.2 NO_2 data retrieval heritage

Tropospheric concentrations of NO_2 are monitored all over the world by a variety of remote sensing instruments – ground-based, in-situ (balloon, aircraft) or satellite-based – each with its own specific advantages, and to some extent still under development.

Stratospheric NO_2 has been measured by a number of satellite instruments since the 1980s, such the spectrometer aboard SME (1981-1989; Mount et al. [1984]), SAGE-II/III (ERBS/Meteor-3M, 1984-2005; Chu and McCornick [1986]), HALOE (UARS, 1991-2005; Gordley et al. [1996]), POAM (SPOT-3, 1993-1996; Randall et al. [1998]), SCIAMACHY (ENVISAT, 2002–2012; Bovensmann et al. [1999], Sierk et al. [2006]), OSIRIS (Odin, 2001–present; Llewellyn et al. [2004]), and ACE (SCISAT-1, 2003–present; Bernath et al. [2005]).

Over the past 17 years tropospheric NO_2 has been measured from UV/Vis backscatter satellite instruments such as GOME (ERS-2, 1995–2011; Burrows et al. [1999]), SCIAMACHY (ENVISAT, 2002–2012; Bovensmann et al. [1999]), OMI (EOS-Aura, 2004–present; Levelt et al. [2006]) and the GOME-2 instruments [Munro et al., 2006] aboard MetOp-A (2006–present) and MetOp-B (2012–present). TROPOMI (see [RD4]; Veefkind et al. [2012]) will extend the records of these observations. TROPOMI, OMI and the GOME-2 instruments provide (near-)global coverage in one day. Figure 3 shows the timelines of the NO_2 data records of these instruments.

For the UV/Vis backscatter instruments that observe NO_2 down into the troposphere, KNMI operates – in close collaboration with BIRA-IASB, NASA and DLR – a continuous data processing system, the results of which are freely available via the TEMIS website [ER1]. The data has been used for a variety of studies in areas like validation (see e.g. Boersma et al. [2009], Hains et al. [2010], Lamsal et al. [2010]), trends (see e.g. Van der A et al. [2008], Stavrakou et al. [2008], Dirksen et al. [2011], Castellanos and Boersma [2012], DeRuyter et al. [2012]), and NO_x emission and lifetime estimates (see e.g. Lin et al. [2010], Beirle et al. [2011], Mijling and Van der A [2012], Wang et al. [2012]).

The approach is called DOMINO (for OMI) and TM4NO2A (for GOME, SCIAMACHY and GOME-2), and is based on a DOAS retrieval, a pre-calculated air-mass factor (AMF) look-up table and a data assimilation / chemistry transport model for the separation of the stratospheric and tropospheric contributions to the NO_2 column (see Sect. 6 for details). The differences between the processing systems for the different instruments are small and related to instrument issues, such as available spectral coverage and wavelength calibration, other absorbing trace gases fitted along, and details of the cloud cover data retrieval.

5.3 Separating stratospheric and tropospheric NO_2 with a data assimilation system

The NO_2 data processing system starts with a retrieval step that determines the NO_2 slant column density, which represents the total amount of NO_2 along the line of sight, i.e. from Sun via Earth's atmosphere to satellite. To determine the tropospheric NO_2 slant column density, the stratospheric slant NO_2 column density is subtracted from the total slant column provided by a DOAS retrieval performed on a spectrum of backscattered light measured by a satellite instrument, after which both slant sub-columns are converted to the tropospheric and stratospheric vertical NO_2 column.

Several approaches to estimate the stratospheric NO_2 amount are used. The DOMINO / TM4NO2A approach uses information from a chemistry transport model by way of data assimilation to estimate the stratospheric NO_2 column [Boersma et al., 2004]. Other methods applied elsewhere include the following.

a) The wave analysis method uses subsets of satellite measurements over unpolluted areas to remove known areas of pollution, i.e. areas with potentially large amounts of tropospheric NO_2, from a 24-hour composite of the satellite measured NO_2 and expands the remainder with a planetary wave analysis across the whole stratosphere, followed where necessary by a second step to mask pollution events (e.g. Bucsela et al. [2006]). This approach has been used between 2004 and 2012 for the OMI NO_2 Standard Product (SP) of NASA/KNMI.

b) The reference sector method method uses a north-to-south region over the Pacific Ocean that is assumed to be free of tropospheric NO_2, as there are no (surface) sources of NO_2, so that all NO_2 measured is assumed to be in the stratosphere (e.g. Richter and Burrows [2002], Martin et al. [2002]). This stratospheric NO_2 is then assumed to be valid in latitudinal bands for all longitudes. In some implementaions this method is extended with a spatial filtering to include other relatively clean areas across the world (e.g. Bucsela et al. [2006], Valks et al. [2011]).

c) Image processing techniques assume that the stratospheric NO_2 shows only smooth and low-amplitude latitudinal and longitudinal variations (e.g. Leue et al. [2001], Wenig et al. [2003]). This approach will probably miss the finer details in the stratospheric NO_2 distribution (as is the case for methods *a* and *b* above). The next version of the OMI NO_2 SP will use a similar approach [Bucsela et al., 2013].

d) Independent stratospheric NO_2 data, such as collocated limb measurements (e.g. Beirle et al. [2010], Hilboll et al. [2013b]) or data taken from a chemistry transport model (e.g. Hilboll et al. [2013a]), can be substracted from the total (slant) column measurements to find the tropospheric NO_2 concentrations. Unfortunately limb collocated stratospheric measurements are not available for satellite retrievals from the GOME(-2), OMI, and TROPOMI sensors. Nevertheless this approach is potentially very useful for comparison and validation studies. Possible cross-calibration problems between the stratospheric and the total measurements would complicate the approach.

These ways of treating the stratospheric NO_2 field may not be accurate enough to capture the variability of the stratospheric NO_2 in latitudinal and longitudinal direction, as well as in time. At the same time it is not certain whether these methods do actually separate stratospheric NO_2: some of the NO_2 interpreted as "stratospheric" may be in the (higher) troposphere.

The use of a data assimilation system to provide stratospheric NO_2 concentrations has been shown to provide realistic results, as indicated by validation studies. For example, Hendrick et al. [2012] found very good agreement between satellite retrievals using data assimilation to estimate the stratospheric NO_2 column (GOME, SCIAMACHY and GOME-2) and ground-based measurements at the station of Jungfraujoch.

In a data assimilation system, meteorological fields are used to drive a chemistry transport model (CTM), while NO_2 slant column data are assimilated to regularly update the three-dimensional NO_2 distribution of the CTM. The data assimilation ensures that the model simulations of the stratospheric NO_2 column agrees closely with the satellite measurements. The advantages of the use of data assimilation are manifold:

- Data assimilation provides a realistic error estimate of the stratospheric NO_2 column [Dirksen et al., 2011].
- The height of the tropopause, obtained from the meteorological data, provides an accurate point of separation of the stratospheric from the tropospheric NO_2 column.
- The result of the data assimilation is a comprehensive understanding of 3-D NO_2 distributions that covers the whole world, taking into account the temporal variability of the NO_2 profiles.

5.4 NO_2 data product requirements

The GMES Sentinels-4, -5 and -5Precursor Mission Requirements Document [RD5] and the Science Requirements Document for TROPOMI [RD6] provide the requirements for the TROPOMI mission, aboard the Sentinel-5Precursor (S5P) mission. For the NO_2 column data products the requirements mentioned in these documents, both for the near-real time and the off-line processing, are listed in Table 1. The requirements are based on the findings of the CAPACITY [RD7], CAMELOT [RD8] and TRAQ [RD9] studies.

The uncertainties stated in Table 1 include retrieval errors as well as instrument errors. Over polluted areas retrieval errors will dominate the uncertainties; these relate to the presence of clouds and aerosols and to the surface albedo. Over rural areas, with low NO_2 concentrations, errors in tropospheric NO_2 are mostly driven by random noise related to the instrument's Signal-to-Noise Ratio (SNR), to estimates of the stratospheric NO_2 column, and to uncertainties in the NO_2 profile [RD6].

Table 1: NO_2 data product requirements for TROPOMI as given in [RD5], which are based on [RD7] (see also [RD6]). The required horizontal resolution is indicated with the targeted requirement and the threshold requirement. Note that the horizontal resolution for TROPOMI is now fixed to a nominal 7×7 km^2 at nadir.

NO_2 data product	Horizontal resolution	Uncertainty	Off-line	NRT	Tables in [RD5]
Total column	$5-20$ km	1.3×10^{15} molec/cm^2	y	y	B1, B2, B3, C1
Tropospheric column	$5-20$ km	1.3×10^{15} molec/cm^2	y	y	B1, B2, B3, C1
Stratospheric column	$50-200$ km	20%	y	y	A2, A3
PBL column	$5-20$ km	10%	y	y	B1, B2, B3

5.5 NO_2 data retrieval for TROPOMI

The TROPOMI data processing of total and tropospheric NO_2 will be based on the DOMINO system (see Sect. 6.1), thus extending the long-term record of NO_2 data, produced using a reliable, well-established and well-described processing system (see Boersma et al. [2004], Boersma et al. [2007] and Boersma et al. [2011]).

For the DOMINO system a number of improvements related to spectral fitting, stratosphere-troposphere separation, and the air-mass factor are underway [Maasakkers et al., 2013] and the TROPOMI data product is expected to benefit from these activities. In addition we plan to investigate a number of possible other improvements with the coming of TROPOMI (see Sect. 6.3 and 6.4).

The TROPOMI NO_2 processing will take place in two locations (Sect. 6.5). The first step of the processing, the DOAS retrieval, and the third step, the conversion of the slant column into the tropospheric and stratospheric NO_2 columns, will take place at the official Level-2 processing site at DLR, the Sentinel-5Precursor Payload Data Ground Segment (PDGS). The data assimilation system providing the information necessary to split the total slant column into its stratospheric and tropospheric contributions and providing NO_2 profile data, will be performed in the Satellite Data Centre (SDC) at KNMI in a processing set-up similar to the current set-up for OMI data (with NASA) and for SCIAMACHY and GOME data (with BIRA-IASB).

In order to comply with the SI unit definitions, the TROPOMI NO_2 data product files give trace gas concentrations in mol/m^2, rather than in the commonly used unit molec/cm^2. For convenience sake, the text and figures of this document will remain in the latter unit; only the tables listing the input (Sect. 7.1) and output (Sect. 7.4) dataset use the SI based units. The multiplication factor to convert mol/m^2 to molec/cm^2 is 6.02214×10^{19} (the multiplication factor to convert mol/m^2 to DU is 2241.15).

5.6 NO_2 data product availability and access

The aim is to start the NO_2 processing as soon as possible after "first light", providing data for initial checks and validations, after which the data will be made available. These steps are planned to take place in the first three months after the commissioning phase, so that the TROPOMI NO_2 data product system would be operational about nine months after launch.

Parallel to the off-line datastream, TROPOMI NO_2 data will be processed in near-real time (NRT), i.e. the data will be available within 3 hours after measurement. All data will be accessible via the official TROPOMI data website and data archive at DLR, the Sentinel-5P Core Service.

6 Algorithm description

6.1 Overview of the NO_2 retrieval algorithm

The TROPOMI NO_2 processing system will be based on the DOMINO system, with improvements related to TROPOMI and state-of-the-art scientific insights. The basis for the processing at KNMI is a retrieval-assimilation-modelling system which uses the 3-dimensional global TM5 chemistry transport model (CTM) as an essential element. The retrieval consists of a three-step procedure, performed on each measured Level-1b spectrum:

1. the retrieval of a total NO_2 slant column density (N_s) from the Level-1b radiance and irradiance spectra measured by TROPOMI using a DOAS (Differential Optical Absorption Spectroscopy) method,
2. the separation of the N_s into a stratospheric (N_s^{strat}) and a tropospheric (N_s^{trop}) part on the basis of information coming from a data assimilation system, and
3. the conversion of the tropospheric slant column density (N_s^{trop}) into a tropospheric vertical column density (N_v^{trop}) and of the stratospheric slant column density (N_s^{strat}) into a stratospheric vertical column density (N_v^{strat}), by applying an appropriate air-mass factor (AMF) based on a look-up table of altitude dependent AMFs and actual, daily information on the vertical distribution of NO_2 from the TM5 model on a $1^\circ \times 1^\circ$ grid.

The data assimilation system in the second step uses a Kalman Filtering Approach that ensures consistency between tropospheric and stratospheric NO_2 profiles in the TM5 model and the retrieved NO_2 slant column N_s. The Kalman forcing is also applied to species that are chemically closely related to NO_2 in the stratosphere, i.e. NO, NO_3, N_2O_5 and HNO_4 [Dirksen et al., 2011]. The TM5 model is driven by up-to-date ECMWF meteorological data (pressure, temperature and wind fields). The tropopause is determined based on the WMO 1985 definition, and by summing all layers above the tropopause in the assimilated field, the stratospheric column is obtained. The choice of the definition of the tropopause or the possible occurrence of a double tropopause is not critical. The pressure level of the tropopause following the WMO 1985 definition is included in the NO_2 data product. NO_x has a C-shape profile and the air around the tropopause has only a small contribution to the total column.

The TM5 model is also used to predict the vertical profile of tropospheric NO_2. This is one of the critical input parameters to a pre-calculated look-up table that provides the appropriate tropospheric AMF, M^{trop}, to convert the tropospheric slant column, N_s^{trop}, into the vertical column, N_v^{trop}. The determination of M^{trop} further requires information on the effective cloud fraction and cloud pressure, derived in a separate processing of the Level-1b spectra, and the surface albedo and terrain height. In addition information regarding geolocation and viewing geometry is needed (from the Level-1b spectrum).

The tropospheric and stratospheric NO_2 vertical column density, as well as the sum of these two, are the final data products of the processing system, along with the intermediate data used.

6.2 Spectral fitting

The baseline method to determine NO_2 total slant columns is Differential Optical Absorption Spectroscopy (DOAS; see Platt [1994], Platt and Stutz [2008]). The DOAS fitting function for TROPOMI follows the current non-linear fitting approach for OMI (Boersma et al. [2011], [AD3]).

The reflectance spectrum $R_{meas}(\lambda)$ observed by the satellite instrument is the ratio of the radiance at the top of the atmosphere, $I(\lambda)$, and the extraterrestrial solar irradiance, $E_0(\lambda)$ (where I also depends on the viewing geometry, but those arguments are left out for brevity):

$$R_{meas}(\lambda) = \frac{\pi I(\lambda)}{\mu_0 E_0(\lambda)} \tag{1}$$

where E_0 is recorded at the same detector pixel as I (once per day), and $\mu_0 = \cos(\theta_0)$ the cosine of the solar zenith angle. In space-borne DOAS, R_{meas} is related to the extinction of light by scattering and absorbing species along the average photon path between Sun and satellite instrument. The effective absorption of NO_2 along the average photon path is interpreted as the total NO_2 slant column density (N_s). The DOAS spectral fitting is performed for all satellite ground pixels with $\theta_0 < 88^\circ$, so that there is no potential danger from the division by μ_0 in Eq. (1). The FRESCO+ cloud data product (Sect. 6.4.4.1) uses this θ_0 cut-off as well. Note that in the conversion of the slant columns into vertical columns (Sect. 6.4.1) θ_0 is cut-off at 80°.

The DOAS spectral fitting attempts to find the optimal modelled reflectance spectrum $R_{\mathrm{mod}}(\lambda)$ by minimising the chi-squared merit function, i.e. the smallest possible differences between the observed and modelled reflectance spectrum:

$$\chi^2 = \sum_{i=1}^{N_\lambda} \Big(\frac{R_{\mathrm{meas}}(\lambda_i) - R_{\mathrm{mod}}(\lambda_i)}{\Delta R_{\mathrm{meas}}(\lambda_i)} \Big)^2 \tag{2}$$

with N_λ the number of wavelengths in the fit window and $\Delta R_{\mathrm{meas}}(\lambda_i)$ the precision of the measurements, which depends on the precision of the radiance and irradiance measurements (omitting the subscript i for brevity):

$$\Delta R_{\mathrm{meas}}(\lambda) = \frac{1}{E_0(\lambda)} \sqrt{\big(\Delta I(\lambda)\big)^2 + \big(\Delta E_0(\lambda)\big)^2 \cdot \big(R_{\mathrm{meas}}(\lambda)\big)^2} \tag{3}$$

The magnitude of χ^2 is a measure for how good the fit is. Another measure for the goodness of the fit is the so-called root-mean-square (RMS) error, which is defined as follows:

$$R_{\mathrm{RMS}} = \sqrt{\frac{1}{N_\lambda} \sum_{i=1}^{N_\lambda} \Big(R_{\mathrm{meas}}(\lambda_i) - R_{\mathrm{mod}}(\lambda_i) \Big)^2} \tag{4}$$

where the difference $R_{\mathrm{meas}}(\lambda) - R_{\mathrm{mod}}(\lambda)$ is usually referred to as the residual of the fit.

The baseline fitting function for TROPOMI follows the approach for OMI and reads as follows:

$$R_{\mathrm{mod}}(\lambda) = P(\lambda) \cdot \exp\Big[- \sum_{k=1}^{N_k} \sigma_k(\lambda) \cdot N_{\mathrm{s},k} \Big] \cdot \Big(1 + C_{\mathrm{ring}} \frac{I_{\mathrm{ring}}(\lambda)}{E_0(\lambda)} \Big) \tag{5}$$

with $\sigma_k(\lambda)$ the cross section and $N_{\mathrm{s},k}$ the slant column amount of molecule $k = 1, \ldots, N_k$ taken into account in the fit (NO_2, O_3, etc.), C_{ring} the Ring fitting coefficient and $I_{\mathrm{ring}}(\lambda)/E_0(\lambda)$ the sun-normalised synthetic Ring spectrum. The Ring spectrum describes the differential spectral signature arising from inelastic Raman scattering of incoming sunlight by N_2 and O_2 molecules. The last term in Eq. (5) describes both the contribution of elastic scattering to the differential absorption signatures (i.e. the 1), and the modification of these differential structures by inelastic scattering (the $+C_{\mathrm{ring}} I_{\mathrm{ring}}(\lambda)/E_0(\lambda)$ term) to the reflectance spectrum.

In the modelled spectrum of Eq. (5) a polynomial of order N_p with coefficients a_m is defined:

$$P(\lambda) = \sum_{m=0}^{N_p} a_m \lambda^m \tag{6}$$

This polynomial is introduced to account for spectrally smooth structures resulting from molecular (single and multiple) scattering and absorption, aerosol scattering and absorption, and surface albedo effects. Because of the polynomial term, only the highly structured differential absorption features contribute to the fit of the slant column densities. In order to prevent the numerical value of the polynomial components in Eq. (6) to become very large or very small (for the $405-465$ nm fit window, for example, usually $m = 5$), the wavelengths of the fit window will be scaled to the range $[-1 : +1]$ prior to the fit. All wavelengths mentioned here are assumed to be scaled.

Figure 4 shows an example of a reflectance spectrum observed by OMI on 1 July 2005, along with the modelled spectrum obtained from the DOAS fit using Eq. (5), with cross-sections for NO_2, Ozone (O_3) and water vapour (H_2O_{vap}), a Ring spectrum and a 5th order polynomial as fitting parameters. The scene is above the equator at about 140° W, with OMI looking down in almost nadir (viewing zenith angle $\theta = 3.44^\circ$) with solar zenith angle $\theta_0 = 34.23^\circ$ (i.e. $\mu_0 = 0.83$). The retrieved total NO_2 slant column $N_{\mathrm{s}} = 3.79 \times 10^{15}$ molec/cm^2. The residual (shown in the bottom panel of Fig. 4) is of the order of 10^{-4}, corresponding to an unexplained differential optial depth of that magnitude.

We can expect that the TROPOMI Level-1b solar irradiance spectra are very well wavelength calibrated ([RD10]). The earth radiance spectra, however, will only have been assigned wavelengths, and thus need to be calibrated before they are usable. Using the subscripts 'ass' and 'cal' to denote assigned and calibrated wavelengths, respectively, the calibrated radiance to be used in Eq. (1) is then given by:

$$I(\lambda_{\mathrm{cal}}) = I(\lambda_{\mathrm{ass}} + w_s + w_q \lambda_{\mathrm{ass}}) \tag{7}$$

where w_s represents a wavelength shift and w_q a wavelength stretch ($w_q > 0$) or squeeze ($w_q < 0$). Once TROPOMI Level-1b spectra are available we will investigate whether including w_q is necessary.

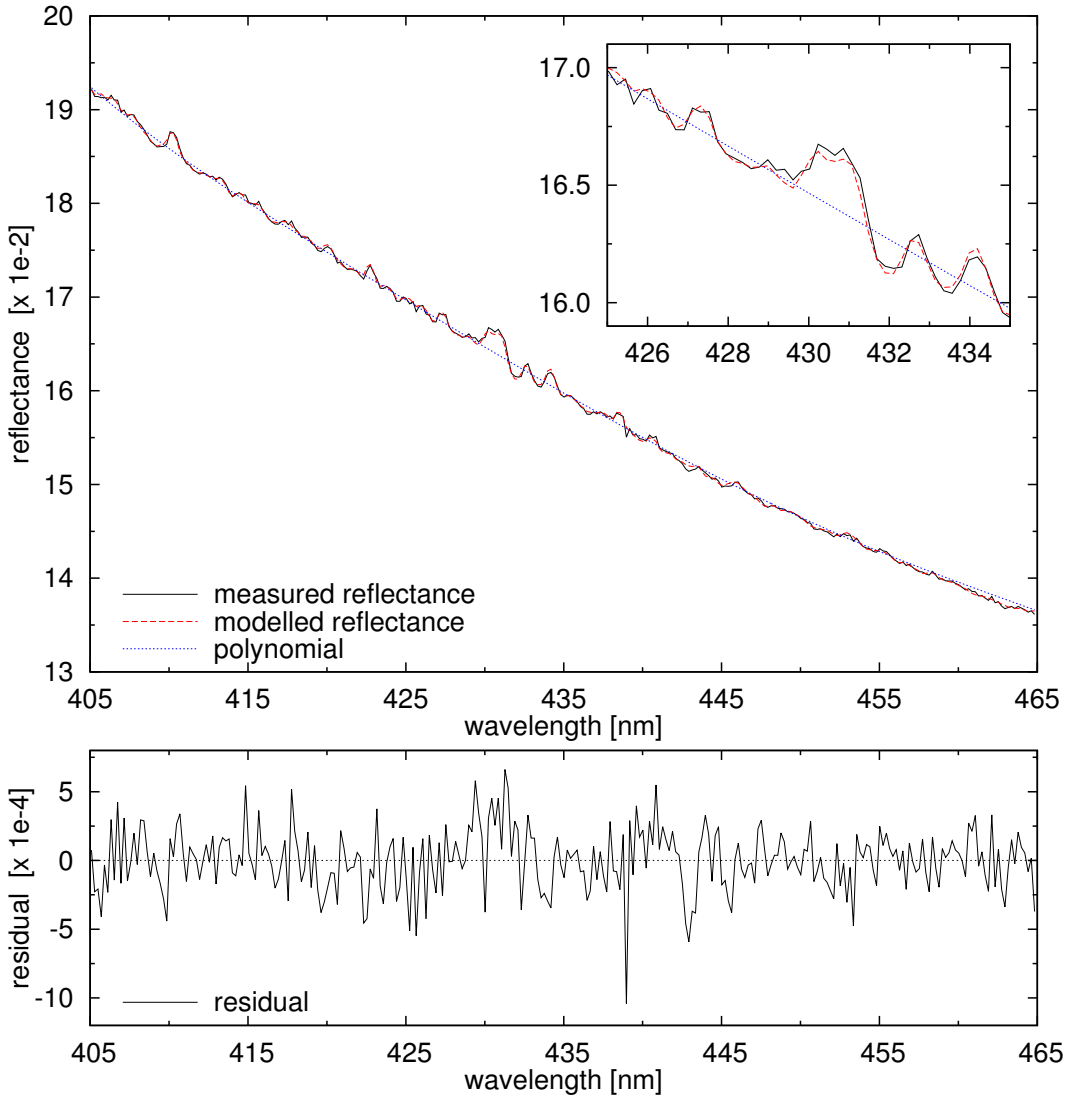

Figure 4: The *top panel* shows an example of a reflectance spectrum (black solid line) obtained by OMI on 1 July 2005, the spectrum modelled in the DOAS fit procedure (dashed red line), and the polynomial of the DOAS fit (blue dotted line). The inset shows an enlargement of a 10 nm wide part of the fit window. The *bottom panel* shows the residual of the DOAS fit, i.e. the measured minus the modelled reflectance spectrum; note that the vertical scale is a factor of 100 smaller than the scale in the top panel. The χ^2 of this particular (rather good) fit is 99.98, and the RMS is 2.23×10^{-4} sr^{-1}

Slant column densities $N_{s,k}$, the Ring coefficent C_{ring}, and the polynomial coefficients a_m are obtained through a non-linear least squares fitting that minimises the χ^2 of Eq. (2), i.e. the differences between the observed and modelled reflectances, based on routines available in the SLATEC mathematical library [Vandevender and Haskell, 1982].

Many implementations of DOAS deploy a linearised version of Eq. (5), including the Ring effect as a pseudo-absorber, giving the equation in terms of optical depth rather than in terms of reflectances: $ln[R_{\text{mod}}(\lambda)] = P^*(\lambda) - \sum\sigma_k(\lambda)\cdot N_{s,k} - \sigma_{\text{ring}}(\lambda)\cdot C^*_{\text{ring}}$, where the Ring coefficient C^*_{ring} and the polynomial $P^*(\lambda)$ are essentially different from C_{ring} and $P(\lambda)$ in Eq. (5). In this approach the Ring cross section $\sigma_{\text{ring}}(\lambda)$ is constructed from the Ring radiance spectrum $I_{\text{ring}}(\lambda)$ divided by a reference solar spectrum. This linearisation allows then for the use of a linear least squares fitting routine, which is computationally potentially faster.

We feel, however, that the Ring effect is physically described better by the non-linear approach of Eq. (5) and we will therefore use that in the NO_2 data processing for TROPOMI. Apart from dropping the physical description of the Ring effect, a disadvantage of the linearised approach is that the error propagation is no longer straightforward, because taking the logarithm of the observed spectra implies that the error no longer

Table 2: Main settings of the operational DOAS retrieval of NO_2 for TROPOMI, and for the current and previous satellite instruments in the operational processing of KNMI, which converts the NO_2 slant column data products into tropospheric and stratospheric vertical column data. For OMI the settings that will be used for the reprocessing of DOMINO version 2.0 to 3.0 are given, while for TROPOMI the same settings will be used; see Sect. 6.2.2 for a brief discussion.

	TROPOMI	OMI (DOMINO v3.0)	GOME-2 (TM4NO2A v2.3)	SCIAMACHY (TM4NO2A v2.3)
wavelength range [nm]	$405-465$	$405-465$	$425-450$	$426.5-451.5$
secondary trace gases	O_3, H_2O_{vap}, O_2–O_2, H_2O_{liq}	O_3, H_2O_{vap}, O_2–O_2, H_2O_{liq}	O_3, H_2O_{vap}, O_2–O_2	O_3, H_2O_{vap}, O_2–O_2
pseudo-absorbers	Ring	Ring	Ring	Ring
fitting method	non-linear	non-linear	linear	linear
degree of polynomial	5	5	3	2
slant column processing	DLR	NASA / KNMI	DLR / BIRA-IASB	BIRA-IASB
references	—	[Boersma et al., 2011] [Van Geffen et al., 2015]	[Valks et al., 2011]	[Van Roozendael et al., 2006]

has a Gaussian distribution.

A number of fitting diagnostics will be provided by the fitting procedure. Estimated slant column and fitting coefficient uncertainties will be obtained from the covariance matrix of the standard errors. This covariance matrix is given as a standard output of the SLATEC non-linear least squares procedure. All fitting coefficients will be provided as diagnostic data.

Table 2 provides an overview of the operational DOAS fit settings planned for TROPOMI and those used for some current and past UV/Vis backscatter satellite instruments: the fit window, the reference spectra used in the fit (see Sect. 6.2.1) and the degree of the DOAS polynomial. A few of the settings for the OMI NO_2 data product are currently under investigation (see Sect. 6.2.2) and this may impact the settings for TROMPOMI. Note that for the processing of GOME(-1) data it was necessary to include a correction for the undersampling of the spectra, i.e. the fact that the spectral sampling is of the same order as the FWHM of the instrument slit function. For the instruments listed in Table 2 this correction is not necessary: their spectral resolution, i.e. the FWHM of the slit function, is 2–3 times as large as their spectral sampling. For TROPOMI, for example, the spectral sampling is about 0.2 nm and the FWHM is about 0.55 nm [RD4].

6.2.1 Reference spectra

The selection of the reference spectra for the trace gas cross sections in Eq. (5) is driven by whether a species shows substantial absorption in the wavelength range relevant for NO_2 retrieval, and will exploit the best available sources. Experience with OMI has shown that NO_2, ozone, water vapour, and Rotational Raman Scattering (RRS), i.e. the inelastic part of the Rayleigh scattering (the so-called "Ring effect"), are most relevant in the wavelength interval relevant to NO_2. Van Geffen et al. [2015] (cf. Sect. 6.2.2) show that including also absorption in liquid water improves the fit for open ocean water ground pixels of OMI, hence it will be included for TROPOMI. High-resolution laboratory measured absorption cross sections will be convolved with the TROPOMI slit function to create the necessary cross sections. The chosen set of reference spectra (see also [RD11] and Fig. 5) is:

- trace gas cross sections $\sigma_k(\lambda)$ in Eq. (5):
  - NO_2 from Vandaele et al. [1998]; see [ER2]
  - O_3 from Brion et al. [1993], Brion et al. [1998] and from Gronshelev et al. [2014], Serdyuchenko et al. [2014]
  - Water vapour (H_2O_{vap}) based on HITRAN 2012 data (see Van Geffen et al. [2015] and Sect. 4.1 of [RD11])
  - O_2–O_2 from Thalman and Volkamer [2013]
  - Liquid water (H_2O_{liq}) from Pope and Frey [1997]
- an effective Ring spectrum $I_{ring}(\lambda)$ in Eq. (5) from Chance and Spurr [1997] (see Van Geffen et al. [2015] and Sect. 4.2 of [RD11])

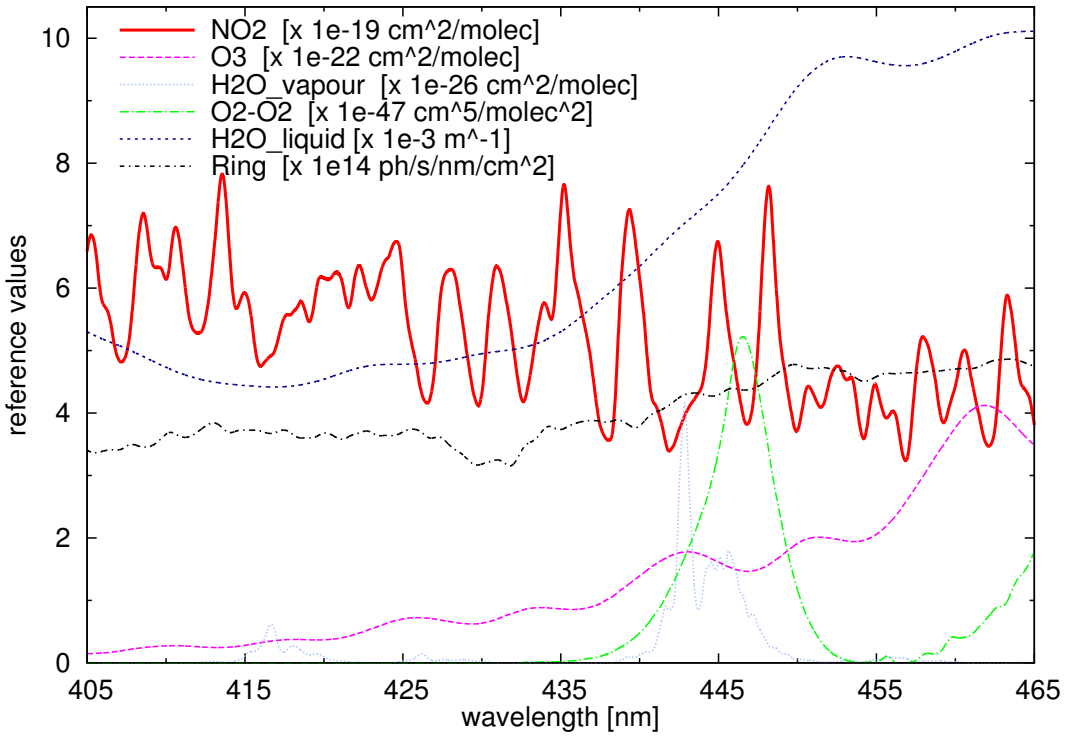

Figure 5: Absorption cross sections $\sigma_k(\lambda)$ for NO_2, O_3, water vapour, O_2–O_2 and liquid water, as well as the Ring spectrum $I_{\text{ring}}(\lambda)$, the pseudo-absorber which accounts for the Ring effect, in Eq. (5) for the $405-465$ nm wavelength range used in the OMI data processor. The reference spectra have been multiplied by the factors given in the plot legend to make the spectral signatures visible in one plot.

The inclusion of absorption by soil (as discussed by e.g. Richter et al. [2011]; Merlaud et al. [2012]) is not considered for TROPOMI, as its potential absorption signal lies well above 465 nm, the upper limit of the fit window considered for the retrieval. Also currently not being considered for inclusion in the fit is the vibrational Raman scattering in open ocean waters (e.g. Vasilkov et al. [2002], Vountas et al. [2003]), as its potential effect on the fit is currently poorly understood.

The temperature for the O_3, H_2O_{vap} and O_2–O_2 cross section spectra is fixed. Variation of these cross section temperatures has little effect on the retrieval of NO_2 slant columns. But the temperature sensitivity of the NO_2 absorption cross section is considerable, and neglecting this temperature dependency can lead to misinterpretations of the slant column on the order of 15%. This temperature dependency does not affect the quality of the spectral fit, since the shape of the differential NO_2 cross section is invariant. It is important to apply an a posteriori correction for deviations of the effective atmospheric temperature from the temperature at which the NO_2 cross section has been established. In the case of TROPOMI, the baseline is to use an NO_2 cross section that has been measured for 220 K. The resulting NO_2 slant column will be corrected for deviations from 220 K at later retrieval steps, i.e. when all the necessary information to calculate the effective atmospheric NO_2 temperature are known, as described in Sect. 6.4.2.

6.2.2 DOAS fit details for OMI and TROPOMI

Comparisons of OMI NO_2 data from the DOMINO version 2.0 processing system to independent data from other instruments have shown that OMI slant NO_2 columns are higher than columns derived from GOME-2 and SCIAMACHY (as first stated by N. Krotkov at the OMI Science Meeting in Sept. 2012), as well as columns derived from groundbased measurements. Due to the separation between stratospheric and tropospheric NO_2, which proceeds in the same way for the three satellite instruments, the high bias in the NO_2 slant columns is propagated to the stratospheric column [Belmonte Rivas et al., 2014].

Van Geffen et al. [2015] show that improving the OMI wavelength calibration of the Level-1b spectra in the OMNO2A data processing of the NO_2 slant columns used by DOMINO v2.0 reduces both the total NO_2 slant column values and the RMS of the DOAS fit. Van Geffen et al. [2015] further show that including both O_2–O_2

and H_2O_{liq} (discussed by e.g. Richter et al. [2011], Lerot et al. [2010]) in the fit improves the OMI NO_2 fit results and ensures that fit coefficients for O_3 and O_2–O_2 have realistic values. Criteria for establishing what are the best settings for the fit can be summarised as follows: (a) a low error on the NO_2 slant column, (b) a low RMS error value, (c) inclusion of secondary trace gases that clearly improve the fit, e.g. by removing specific features in the fit residual, (d) physically realistic values for the slant column values of these secondary trace gases.

The improvements described by Van Geffen et al. [2015] for OMNO2A will be used for a reprocessing of all OMI NO_2 slant column data, which will subsequently be input for DOMINO v3.0 data, and are the basis for the TROPOMI NO_2 slant column processing. The OMNO2A wavelength calibration determines a spectral shift. For the TROPOMI NO_2 processing a spectral shift (and perhaps also a spectral stretch/squeeze) will be included in the NO_2 slant column fit, as described by Eq. (7).

6.3 Separation of stratospheric and tropospheric NO_2

The baseline method for the TROPOMI NO_2 algorithm to separate the stratospheric and tropospheric contribution to the NO_2 total slant columns is by data assimilation of slant columns in a chemistry transport model (CTM). KNMI has considerable experience with this method, and in the absence of collocated independent (e.g. limb) information on stratospheric NO_2, we consider this to be the most viable method to distinguish stratospheric from tropospheric NO_2.

The central idea of the data assimilation is to regularly update a CTM simulation of the three-dimensional, coupled troposphere-stratosphere NO_2 distribution with available measurement data in such a way that the CTM simulation of the stratospheric NO_2 column achieves close agreement with the TROPOMI slant columns over areas known to have little or no tropospheric NO_2. The assimilation effectively relies on slant columns observed over regions where the model predicts the NO_2 column to be dominated by stratospheric NO_2 (e.g. over the remote oceans). For those regions and times, the modeled slant column, i.e. the inner product of the observation operator **H** and the simulated vertical distribution $\vec{x}$, is effectively forced to the observed state. For regions and times where the model predicts large tropospheric contributions, the slant column is not as good a proxy for stratospheric NO_2, and the assimilation relies more on its actual state. Because total reactive nitrogen (NO_y) is a well-conserved quantity in the stratosphere, with relatively small source and sink contributions, the information from the observations can be stored in the model over long time periods.

The assimilation scheme is based on the Kalman filter technique, with a prescribed parameterisation of the horizontal correlations between forecast errors to reduce computational effort. If NO_2 slant columns are available with a measurement time within 15 minutes of the model time, the model field is updated by the Kalman filter. In the Kalman filter update, the forecast model state is adjusted toward the observations, replacing the forecast with the analysis. This analyzed profile field $\vec{x}_a$ includes NO_2 in both troposphere and stratosphere, and is calculated from the forecast $\vec{x}_f$ and the 2-D field of so-called superobservations $\vec{y}$ by:

$$\vec{x}_a = \vec{x}_f + \mathbf{P}\mathrm{H}^T(\mathbf{H}\mathrm{P}\mathrm{H}^T + \mathbf{R})^{-1}(\vec{y} - \vec{y}_f) \tag{8}$$

with matrix **H** the observation operator, **P** the forecast error covariance matrix, and **R** the combined observation and representativeness error covariance (Eskes et al. [2003]; Dirksen et al. [2011]). The term $\mathbf{P}\mathrm{H}^T(\mathbf{H}\mathrm{P}\mathrm{H}^T + \mathbf{R})^{-1}$ determines the most likely adjustment of the model state, given the difference between observed and forecast model column ($\vec{y} - \vec{y}_f$, observation minus forecast). Note that the total slant column $\vec{y}$ includes the NO_2 in both troposphere and stratosphere. The relative size of the adjustment depends on the ratio between the uncertainties in the model forecast and observations, and the model analysis will closely follow the observations when this ratio is large.

The observation operator **H** is proportional to the averaging kernel [Eskes and Boersma, 2003], a 34-element vector that contains the sensitivity of TROPOMI to NO_2 in each model layer. The scalar product of the observation operator vector and the model NO_2 profile at the location of the individual TROPOMI observations yields the slant column that would be observed by TROPOMI given the modeled profile $\vec{x}_f$. The average of all TROPOMI observations (and model equivalents) with center coordinates inside a $1^\circ \times 1^\circ$ (longitude $\times$ latitude) model grid cell is treated as a single measurement, a so-called superobservation (and model equivalent). $\vec{y}_f$ is the model forecast of the superobservations, given by $\mathbf{H}\vec{x}_f$. In order to reduce the computational effort, the Kalman filter is applied for these superobservations, rather than for all individual observations. An additional advantage of this approach is that because of averaging the TROPOMI observations into superobservations, much of the noise in the TROPOMI observations cancels out, allowing a tight observational constraint on the model state.

We will use the TM5 CTM (Huijnen et al. [2010a]; Huijnen et al. [2010b]; [ER3]) for the assimilation of TROPOMI NO_2 slant columns. This is a major improvement over the current data assimilation systems

operated at KNMI for GOME, SCIAMACHY, OMI, and GOME-2, which use an older version of the TM CTM (TM4; e.g. Dentener et al. [2003]). The main advantage of the transition to TM5 is the better spatial resolution ($1^\circ \times 1^\circ$), updated information on (NO_x) emissions, and improved description of relevant physical (photolysis rate constants) and chemical (reaction rate constants) processes in that model. The assimilation system operates at a resolution of $1^\circ \times 1^\circ$ (longitude $\times$ latitude), with 34 sigma pressure layers up to 0.1 hPa in the vertical direction. TM5 uses forecast and analyzed 6-hourly meteorological fields, (3-hourly for boundary layer fields) from the European Centre for Medium Range Weather Forecast (ECMWF) operational model. These fields include global distributions of wind, temperature, surface pressure, humidity, (liquid and ice) water content, and precipitation.

Once the TROPOMI slant columns have been assimilated, the integral from the layer above the tropopause to the upper TM5 layer provides the stratospheric slant column that can be isolated from the total slant column, giving the tropospheric slant column (cf. Sect. 6.4):

$$N_{\mathrm{s}}^{\mathrm{trop}} = N_{\mathrm{s}} - N_{\mathrm{s}}^{\mathrm{strat}} \tag{9}$$

For the tropopause definition the WMO-1985 temperature gradient criterion is followed, but other definitions would not lead to significantly different results (e.g. Bucsela et al. [2013]).

The data assimilation system provides the following information, necessary for the subsequent processing in the calculation of the air-mass factor (AMF; see Sect. 6.4) needed for the conversion of the tropospheric slant column to the tropospheric vertical column and the final NO_2 data product (see Sect. 6.6):

- the stratospheric slant and vertical columns: $N_{\mathrm{s}}^{\mathrm{strat}}$ and $N_{\mathrm{v}}^{\mathrm{strat}}$
- an estimate of the error on the stratospheric vertical column: $\sigma(N_{\mathrm{v}}^{\mathrm{strat}})$
- the NO_2 profile shape: $n_{l,\mathrm{NO_2}}$, with l the index for the layer number $1, 2, \ldots, N_l$
- the temperature profile at the layers: T_l^{TM5}, for $l = 1, 2, \ldots, N_l$
- the pressure level coefficients: A_l^{TM5}, B_l^{TM5}, for $l = 0, 1, \ldots, N_l$
- the index of the pressure level of the tropopause: $l_{\mathrm{tp}}^{\mathrm{TM5}}$, using the WMO-1985 temperature gradient criterion
- the surface elevation and pressure: $z_{\mathrm{s}}^{\mathrm{TM5}}$ and $p_{\mathrm{s}}^{\mathrm{TM5}}$, at the $1^\circ \times 1^\circ$ model resolution

Note that the model divides the atmosphere in N_l layers. The pressure level coefficients determine the pressure at the $N_l + 1$ levels separating the layers: $p_l = A_l^{\mathrm{TM5}} + B_l^{\mathrm{TM5}} \cdot p_{\mathrm{s}}$, for $l = 0, 1, \ldots, N_l$, with p_{s} the surface pressure for the given TROPOMI ground pixel. The pressure for the layer l, for which the concentration (volume mixing ratio) $n_{l,\mathrm{NO_2}}$ and the temperature T_l^{TM5} are given, is then midway between the level pressures p_{l-1}^{TM5} and p_l^{TM5}. And the layer with index $l_{\mathrm{tp}}^{\mathrm{TM5}}$ contains the tropopause.

6.3.1 Stratospheric chemistry in the TM5 model

TM5 is primarily a tropospheric chemistry model. NO_x-O_x-HO_y chemical processes are implemented according to the Carbon Bond Mechanism 4 (CBM-4) chemistry scheme, which includes non-methane hydrocarbons to account for loss by reactions with OH [Houweling et al., 1998]. Because the chemistry version of TM5 does not simulate N_2O, the actual source of NO_x to the stratosphere, NO_x is derived from simulated HNO_3 concentrations, which follow climatological HNO_3:O_3 ratios observed by UARS (baseline option) in the 1990s, and the multi-sensor reanalysis of stratospheric O_3 columns [Van der A et al., 2010]. In this way the model partly compensates for the biases that occur due to the missing N_2O source globally, and the missing reactions involving halogens which are important in the polar vortex. Climatological HNO_3:O_3 ratios obtained from more recent ODIN satellite measurements are considered as a non-baseline option [Maasakkers et al., 2013]. Processes included in the TM5 tracer evolution are advection, convection, diffusion, photolysis and deposition. Rapid changes in stratospheric NO_2 due to e.g. sudden stratospheric warmings or changes in the vortex edge location are largely accounted for through the use of the ECMWF analysis. Solar proton events are not included in the model, but the related biases are largely removed by the assimilation. NO_x emissions are based on the RETRO-REAS emission inventories for 2006. For more details, the reader is referred to Dirksen et al. [2011].

The data assimilation provides a regular update of the TM5 simulation, with a time step of 30 minutes, of the NO_2 distribution in the atmosphere on the basis of available observations: if NO_2 slant columns are available with a measurement time within 15 minutes of the model time, the model field is updated, i.e. the forecast TM5 state is adjusted towards the observations. The observation error covariance matrix is set up in such a way that it effectively filters out measurements with increased tropospheric NO_2 load by attributing less weight to observations over known polluted areas. As a result of this, the simulated stratospheric NO_2 concentrations are forced toward the observations, with only a marginal adjustment of the simulated tropospheric NO_2 field.

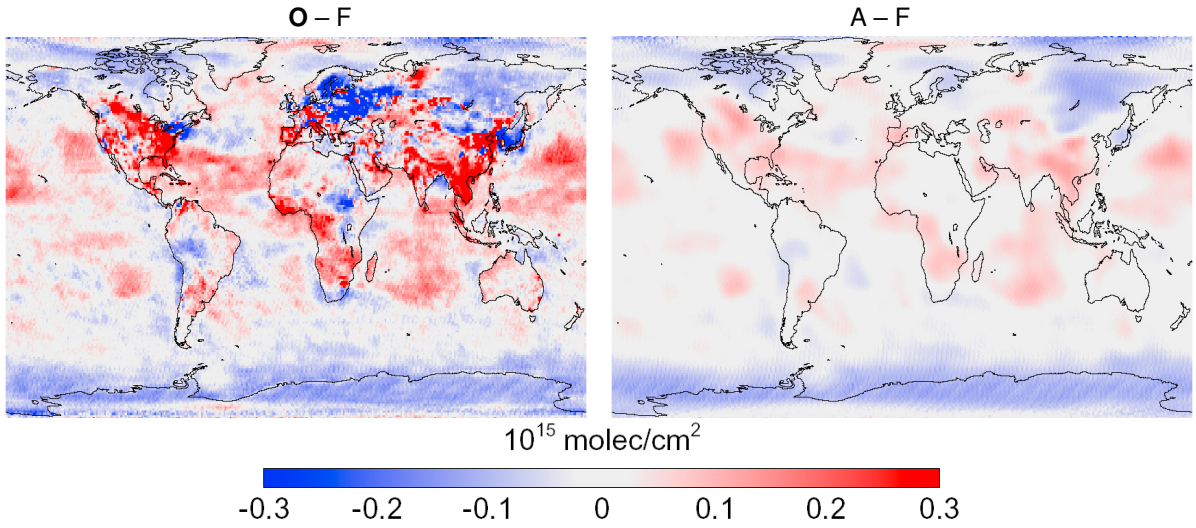

Figure 6: Monthly mean "observation minus forecast" (O–F; left panel) and "analysis minus forecast" (A–F; right panel) differences in NO_2 slant columns for March 2005, at a resolution of $1^\circ \times 1^\circ$, based on OMI data. (Image source: Dirksen et al. [2011].)

The error estimate is based on "observation minus forecast" statistics (over relatively unpolluted areas) in the assimilation. Our experience with NO_2 data assimilation using GOME, SCIAMACHY, OMI, and GOME-2 in TM has shown that the model chemistry responds smoothly to the updates forced by the satellite measurements.

Figure 6, from Dirksen et al. [2011], provides an example of the monthly mean "observation minus forecast" (O–F) and the model forcing ("analysis minus forcast", A–F) for March 2005. The difference between the two panels of Fig. 6 illustrates the effect of the assimilation: considerable O–F differences, resulting mostly from (anthropogenic) tropospheric NO_2 sources, have only a minor influence on the analysis. On the other hand, synoptic-scale structures in O–F persist in the A–F differences. That the A–F differences are much smaller (generally less than $\pm 0.15 \times 10^{15}$ molec/cm^2) than the O–F differences (up to $\pm 0.4 \times 10^{15}$ molec/cm^2) demonstrates that most tropospheric contributions are effectively discounted by the assimilation procedure.

Baseline – climatological HNO_3:O_3 ratios from UARS
Non-baseline – climatological HNO_3:O_3 ratios from ODIN or other limb-sensor (SCIAMACHY, MIPAS)

6.4 Air-mass factor and vertical column calculations

The TROPOMI NO_2 algorithm will use as default pre-calculated air-mass factor look-up tables to convert the tropospheric and stratospheric slant columns into meaningful vertical columns. The air-mass factor (AMF), denoted by the symbol M, is the ratio of the slant column density of the absorbing trace gas along the (slant) optical path from sun to satellite, and the vertical column density above the point at the surface area the satellite is viewing. The total vertical column density then follows from the retrieved total slant column density:

$$N_\mathrm{v} = N_\mathrm{s}/M \tag{10}$$

The AMF depends on the vertical profile of the trace gas and can be written as (Palmer et al. [2001]; Eskes and Boersma [2003]):

$$M = \frac{\sum_l m_l n_l c_l}{\sum_l n_l}, \qquad m_l \equiv \delta N_\mathrm{s}/\delta n_l \tag{11}$$

with m_l the altitude dependent AMFs (see Sect. 6.4.1) that describe the vertically resolved sensitivity to NO_2, n_l the column density, and c_l the temperature correction term discussed below (see Sect. 6.4.2) for layer $l = 1, 2, \ldots, N_l$ [Boersma et al., 2004]. The altitude-dependent AMFs depend on retrieval (forward model) parameters, including the satellite viewing geometry, as well as surface albedo and surface pressure, cloud fraction, and cloud pressure.

The data assimilation provides an estimate for the stratospheric vertical profile. Summation over the layers above the tropopause level ($l > l_\mathrm{tp}^\mathrm{TM5}$) to top-of-atmosphere ($l = N_l$) provides the stratospheric AMF, from which

the stratospheric slant column can then be calculated:

$$N_{\mathrm{s}}^{\mathrm{strat}} = N_{\mathrm{v}}^{\mathrm{strat}} * M^{\mathrm{strat}} = \sum_{l=l_{\mathrm{tp}}^{\mathrm{TM5}}+1}^{N_l} m_l \, n_l \, c_l \tag{12}$$

Subtracting this from the total slant column and using the tropospheric AMF, determined by integration over the layers from the surface ($l = 1$) up to and including the tropopause level ($l = l_{\mathrm{tp}}^{\mathrm{TM5}}$), then gives the tropospheric vertical column:

$$N_{\mathrm{s}}^{\mathrm{trop}} = N_{\mathrm{s}} - N_{\mathrm{s}}^{\mathrm{strat}} \qquad \Rightarrow \qquad N_{\mathrm{v}}^{\mathrm{trop}} = N_{\mathrm{s}}^{\mathrm{trop}} / M^{\mathrm{trop}} \tag{13}$$

Note that the total vertial column N_{v} in Eq. (10) is *not* the same as sum of the partial vertical columns:

$$N_{\mathrm{v}}^{\mathrm{sum}} \equiv N_{\mathrm{v}}^{\mathrm{trop}} + N_{\mathrm{v}}^{\mathrm{strat}} \neq N_{\mathrm{v}} \tag{14}$$

Our best physical estimate of the NO_2 vertical column at any given place is the sum $N_{\mathrm{v}}^{\mathrm{sum}}$. Users who, for example, wish to assimilate NO_2 total columns should, however, use the total column N_{v} for this.

6.4.1 Altitude dependent AMFs

The altitude-dependent AMFs, or vertical sensitivities, will be calculated with a radiative transfer model by adding a finite amount of NO_2 to the model atmosphere layer l and subsequently ratioing the excess NO_2 slant column (simulated with a radiative transfer model) to the vertical column added to that layar ($m_l = \delta N_{\mathrm{s}} / \delta n_l$). The model atmosphere does not include aerosols and describes the Earth's surface as a Lambertian reflector.

As radiative transfer model we will use the Doubling Adding KNMI (DAK) radiative transfer model (De Haan et al. [1987]; Stammes et al. [2001]), version 3.2, which has the possibility to include a pseudo-sphericity correction. The radiative transfer calculations will take the sphericity of the atmosphere into account, with Rayleigh scattering (including multiple scattering effects) and polarisation correction included (see Boersma et al. [2011] and references therein). The DAK model atmosphere consists of a Lambertian surface albedo, and an adjustable number of atmospheric layers. Atmospheric data are from the standard AFGL midlatitude summer profile. We calculate the AMF at 439 nm, in the middle of the spectral fitting window, for the corresponding TROPOMI NO_2 slant column retrievals.

The altitude-dependent AMFs are stored in a look-up table (LUT) as a function of solar zenith angle (θ_0), viewing zenith angle (θ), relative azimuth angle ($\phi - \phi_0$), Lambertian surface albedo (A_{s}), surface pressure (p_{s}), and (midlevel) atmospheric pressure (p_l). This 6-dimensional LUT is to be extended with more reference points compared to earlier versions in order to respect the anticipated increase in variability of TROPOMI retrieval parameters (coarser OMI pixels have less variability in spatially smeared surface albedo and surface pressure values than anticipated for TROPOMI) and to minimise interpolation errors when looking up the appropriate altitude-dependent AMF. Pixel-specific altitude-dependent AMFs are obtained by using the best estimates for forward model parameters and a 6-D linear interpolation scheme. The dimensions for the LUT are chosen to balance sufficiently accurate interpolation with computational efficiency and resource economy. Table 3 gives an overview of the references points for the quantities that make up the 6 dimensions. In the current OMI NO_2 data product only ground pixels with $\theta_0 < 80^\circ$ ($\cos(\theta_0) = 0.174$) are used in the conversion to vertical columns. For TROPOMI we will investigate whether it is possible to extend this range to $\theta_0 < 88^\circ$ ($\cos(\theta_0) = 0.035$), hence the lower limit of $\cos(\theta_0)$ of 0.03 in Table 3.

6.4.2 Temperature correction

For the TROPOMI NO_2 retrieval, a temperature correction will be applied in the air-mass factor step (see Eq. 11). The NO_2 cross-sections used in the DOAS retrieval, taken from Vandaele et al. [1998] [ER2], are valid for NO_2 at a temperature of 220 K. The temperature at which the NO_2 cross-section is evaluated does significantly influence the fit: amplitudes of the differential NO_2 absorption features decrease with increasing temperature, while the overall shape of the differential cross-section is independent of temperature.

To account for the temperature sensitivity, a correction factor has been determined for the difference between the effective temperature of the NO_2 (which is derived from the ECMWF temperature profile and the modelled profiles in the data assimilation system) and the temperature of the cross-section, where the temperature dependence is assumed to be linear. For layer l of the NO_2 profile the correction factor c_l is:

$$c_l = 1 - 0.00316(T_l - T_\sigma) + 3.39 \times 10^{-6} (T_l - T_\sigma)^2 \tag{15}$$

Table 3: Quantities and their reference points in the AMF look-up table to be used in the TROPOMI NO_2 data processing to convert the tropospheric slant column into the tropospheric vertical column. The lower limit of $\cos(\theta)$ in the list is related to the maximum value of θ for TROPOMI, which is 72° (as for OMI).

Quantity	Number of reference points	Values at reference points
Solar zenith angle $\cos(\theta_0)$	17	1.00, 0.95, 0.90, 0.80, 0.70, 0.60, 0.50, 0.45, 0.40, 0.35, 0.30, 0.25, 0.20, 0.15, 0.10, 0.05, 0.03
Viewing zenith angle $\cos(\theta)$	11	1.00, 0.95, 0.90, 0.80, 0.70, 0.60, 0.50, 0.45, 0.40, 0.35, 0.30
Relative azimuth angle $180^\circ - \lvert\phi - \phi_0\rvert$	10	0°, 20°, 40°, 60°, 80°, 100°, 120°, 140°, 160°, 180°
Surface albedo A_s	26	0.00, 0.01, 0.02, 0.03, 0.04, 0.05, 0.06, 0.07, 0.08, 0.09, 0.10, 0.12, 0.14, 0.16, 0.18, 0.20, 0.25, 0.30, 0.35, 0.40, 0.50, 0.60, 0.70, 0.80, 0.90, 1.00
Surface pressure p_s [hPa]	14	1048, 1036, 1024, 1013, 978, 923, 840, 754, 667, 554, 455, 372, 281, 130
Atmospheric pressure p_l [hPa]	174	1054.995, 1042.82, 1030.78, 1018.89, 1007.13, 995.51, 984.0309, 972.67, 961.45, 950.35, 939.39, 928.55, 917.84, 907.24, 896.71, 886.24, 875.88, 865.65, 855.54, 845.54, 835.67, 825.90, 816.26, 806.72, 797.12, 787.47, 777.93, 768.51, 759.21, 750.01, 740.93, 731.96, 723.09, 714.33, 705.65, 697.04, 688.54, 680.14, 671.85, 663.65, 655.56, 647.56, 639.66, 631.86, 624.07, 616.30, 608.62, 601.03, 593.54, 586.15, 578.85, 571.63, 564.51, 557.48, 550.44, 543.39, 536.43, 529.56, 522.77, 516.08, 509.47, 502.9492, 496.50, 490.14, 483.75, 477.32, 470.97, 464.71, 458.53, 452.44, 446.42, 440.49, 434.63, 428.86, 423.12, 417.42, 411.80, 406.26, 400.79, 395.39, 390.07, 384.82, 379.64, 374.52, 369.43, 364.37, 359.37, 354.44, 349.57, 344.78, 340.05, 335.38, 330.78, 326.24, 321.70, 317.15, 312.66, 308.24, 303.89, 299.59, 295.35, 291.18, 287.06, 283.00, 261.31, 225.35, 193.41, 165.49, 141.03, 120.12, 102.68, 87.82, 75.12, 64.30, 55.08, 47.20, 40.535, 34.79, 29.86, 25.70, 22.14, 19.08, 16.46, 14.20, 12.30, 10.69, 9.29, 8.06, 6.70, 6.11, 5.37, 4.70, 4.10, 3.57, 3.12, 2.74, 2.41, 2.12, 1.87, 1.65, 1.46, 1.29, 1.141, 1.01, 0.89, 0.79, 0.69, 0.61, 0.54, 0.48, 0.42, 0.37, 0.33, 0.29, 0.23, 0.18, 0.13, 0.10, 0.07, 0.05, 0.04, 0.030, 0.020, 0.014, 0.0099, 0.0066, 0.004471, 0.002997, 0.002005, 0.001352, 0.0009193, 0.0006300, 0.0004387, 0.000307

with T_l and T_σ the temperature of the profile layer and cross-section, respectively. The function in Eq. (15) is an update[1] w.r.t. the correction used for the OMI NO_2 data in DOMINO v2 (Boersma et al. [2002], Boersma et al. [2004], Bucsela et al. [2013]). Note that the temperature sensitivity given in the above equation is determined for the default wavelength window 405–465 nm used for the fit; depending on the fit window and on TROPOMI's spectral resolution details, the function may need to be adapted.

[1] M. Zara, Jan. 2016, KNMI Technical Report, in preparation.

6.4.3 Cloud correction

The AMF formulation accounts for cloud-contaminated pixels. Following Martin et al. [2002] and Boersma et al. [2002] the independent pixel approximation (IPA) is used to express the AMF as a linear combination of a cloudy AMF (M_{cl}) and a clear-sky AMF (M_{cr}):

$$M = w M_{\mathrm{cl}} + (1-w) M_{\mathrm{cr}} \tag{16}$$

with w the radiance weighted cloud fraction, which depends on the effective cloud fraction (f_{eff}):

$$w = \frac{f_{\mathrm{eff}} I_{\mathrm{cl}}}{R} = \frac{f_{\mathrm{eff}} I_{\mathrm{cl}}}{f_{\mathrm{eff}} I_{\mathrm{cl}} + (1-f_{\mathrm{eff}}) I_{\mathrm{cr}}} \tag{17}$$

where I_{cl} is the radiance from the cloudy part of the pixel, I_{cr} the radiance from the clear part of the pixel, and R the total scene radiance. Both I_{cl} and I_{cr} depend on the viewing geometry and the assumed (cloud) albedo, and these are calculated following Vermote and Tanré [1992].

The AMF calculations for TROPOMI will use the effective cloud fraction (f_{eff}) and cloud pressure (p_{c}) from the O_2 A-band cloud retrieval, using the FRESCO+ method (see Sect. 6.4.4.1). The TROPOMI cloud and NO_2 retrievals will be consistent in the sense that both use the independent pixel approximation, which represents clouds as opaque Lambertian surfaces of albedo 0.8.

6.4.4 Retrieval parameters

6.4.4.1 Cloud cover data

For TROPOMI NO_2 data, the FRESCO+ algorithm will be the baseline for the retrieval of cloud parameters. The FRESCO+ algorithm (Wang et al. [2008]; [RD12]) retrieves cloud information from the O_2 A-band around 758 nm: the cloud fraction and the cloud pressure, for all satellite ground pixels with solar zenith angle $\theta_0 < 88^\circ$. The surface albedo database that will be used by the FRESCO+ algorithm is a combination of MERIS data over land and GOME data over oceans at 754 and 775 nm (see Popp et al. [2011] and references therein, and [ER4]), which is described further in Sect. 6.4.4.2.

FRESCO+ does not provide the geometric cloud fraction but rather a radiometric equivalent cloud fraction: an effective cloud fraction, f_{eff}, that results in the same top-of-atmosphere radiance as the real cloud, based on an optically thick Lambertian cloud with a fixed albedo of $A_{\mathrm{c}} = 0.8$ at the cloud pressure level, p_{c}. This approach has proven to be useful for trace gas retrieval, and the errors introduced to the trace gas retrievals are usually small (and minimal for a fixed cloud albedo of $A_{\mathrm{c}} = 0.8$; see Wang et al. [2008], who evaluated this for ozone and NO_2) when compared to scattering cloud models (Koelemeijer et al. [2001]; Stammes et al. [2008]).

Because of the large difference in wavelength between the O_2 A-band and the NO_2 retrieval window, the cloud fraction retrieved by FRESCO+ in the O_2 A-band may not be exactly representative for the cloud fraction in the NO_2 window, although Van Diedenhoven et al. [2007] found that cloud parameters retrieved from UV and O_2 A-band measurements showed good consistency for cloud fractions > 0.2; for mostly clear skies, FRESCO+ provides somewhat higher cloud fractions than UV-based retrievals.

In addition, a small misalignment between ground pixel field-of-view of the VIS and NIR bands, containing the NO_2 retrieval window and the O_2 A-band, respectively, is expected for the TROPOMI measurements.

For these reasons, the baseline option for the TROPOMI NO_2 retrieval is to retrieve the cloud fraction from the NO_2 spectral window itself, $f_{\mathrm{eff,NO_2}}$. This can be done by fitting the observed continuum reflectance to a simulated reflectance constructed with the independent pixel approximation and radiative transfer calculations for the clear-sky and cloudy-sky part of the pixel, using the appropriate surface albedo in that spectral window as forward model parameter. The approach is very similar to FRESCO+ but now around the 450 nm domain, and explicitly accounts for Rayleigh scattering; see [RD13]. The work involves the calculation of LUTs with the TOA reflectance as a function of viewing geometry, surface/cloud albedo, and surface pressure. Cloud pressure would still be obtained from the O_2 A-band.

Apart from the support product FRESCO+, TROPOMI cloud parameters will be provided by an algorithm currently under development at DLR [RD14]. Once the validity and reliability of this cloud data product is established, its cloud parameters will be tested in the NO_2 processor and the results will be compared against the results found with FRESCO+ cloud data.

Baseline – use the cloud fraction retrieved in the NO_2 fit window and the FRESCO+ cloud pressure

from the O_2 A-band

Non-baseline – cloud parameters from other sources of information (e.g. the TROPOMI data product provided by DLR)

6.4.4.2 Surface albedo

The baseline surface albedo climatology for TROPOMI NO_2 retrievals is the OMI database, aggregated to a grid of $0.5^\circ \times 0.5^\circ$; see Kleipool et al. [2008], which describes a climatology made from 3 years of OMI data. Meanwhile the climatology has been improved by using 5 years of data, based on the the same method [ER5]. This 5 years based climatology (version 3) is currently in use for the OMI NO_2 retrieval, and it will also be the baseline for the TROPOMI NO_2 retrievals.

The OMI albedo climatology is considered to be the best currently available source of information for the surface albedo, because of its spectral coverage in the NO_2 fit region, its relatively high spatial resolution, and the seamless transition between land and sea. An additional advantage is that the Kleipool-climatology [Kleipool et al., 2008] has been derived from observations taken at similar local times and under similar viewing conditions as the TROPOMI observations will be taken.

The Kleipool surface albedo climatology is based on OMI data, which does not cover the near-infrared wavelengths in use by the FRESCO+ algorithm to derive cloud properties (Sect. 6.4.4.1). Instead the black-sky albedo of the channels at 754 nm and 775 nm of the MERIS albedo climatology, aggregated to a grid of $0.25^\circ \times 0.25^\circ$ (see Popp et al. [2011] and references therein, and [ER4]), will be used for FRESCO+. Unfortunately, however, the MERIS data does not provide albedo values over oceans. For oceans the GOME monthly surface albedo climatology [Koelemeijer et al., 2003] is used, which has albedo values for both the O_2 A-band and the NO_2 window at a resolution of $1^\circ \times 1^\circ$ (see Popp et al. [2011] and [ER4]). Because of the relatively coarse spatial resolution of the GOME albedo data and possible discontinuities between the GOME and MERIS data at land/ocean boundaries, the MERIS/GOME data is considered less suitable as albedo climatology for the NO_2 retrieval, motivating our choice for the Kleipool surface albedo climatology.

The use of the MERIS/GOME albedo climatology for FRESCO+ may imply an inconsistency between the cloud parameters and the NO_2 retrieval. We will investigate the consistency between the MERIS/GOME and the OMI based albedo values to determine the potential effect of any discrepencies on the NO_2 retrieval. The MERIS/GOME albedo climatology is available as fallback option for the NO_2 retrieval, in case the discrepencies are found to be too large. Another issue to be addressed once TROPOMI is operational is the non-perfect co-registration of the visible and near-infrared channels of the TROPOMI data, a consequence of the way the TROPOMI detector system is constructed [RD4].

Accounting for the anisotropic properties of surface reflectance is currently not foreseen as a baseline for the TROPOMI NO_2 retrieval algorithm. A recent study showed that accounting for BRDF in OMI NO_2 retrievals has a generally small effect (<5%) with substantial effects only occurring at extreme viewing angles at high solar zenith angles [Zhou et al., 2010]. The specific choice of albedo data set is considered to be more important than accounting for surface BRDF effects. In the near future we will consider actual developments in generating improved surface albedo data, for instance from the ADAM (see [RD11], Sect. 6.1) and QA4ECV [RD15] projects.

Baseline – use of Kleipool surface albedo climatology

Non-baseline – use of MERIS/GOME surface albedo climatology

– account for BRDF effects on the surface albedo

6.4.4.3 Snow and ice cover

Because the surface albedo for a satellite ground pixel will be somewhat different from the value given in the surface albedo climatology, small errors may be introduced in the retrieval. Substantial errors are introduced if the real albedo differs considerably from what is expected, for example in the case of the sudden snowfall or ice cover. Correcting the surface albedo from the climatology using knowledge of actual snow/ice cover will therefore improve the final data product, in terms of the retrieval itself and for flagging such cases. Suitable candidates for snow/ice cover information are the daily updated OSISAF [ER6] and NISE [ER7] data sets or the snow/ice data provided by the ECMWF (see also [RD1]).

The FRESCO+ algorithm (Sect. 6.4.4.1) provides two sets of data (Wang et al. [2008]; citeRDATBD-FRESCO): (i) the effective cloud fraction f_{eff} and cloud pressure p_{c} using a cloud albedo $A_{\text{c}} = 0.8$, and (ii)

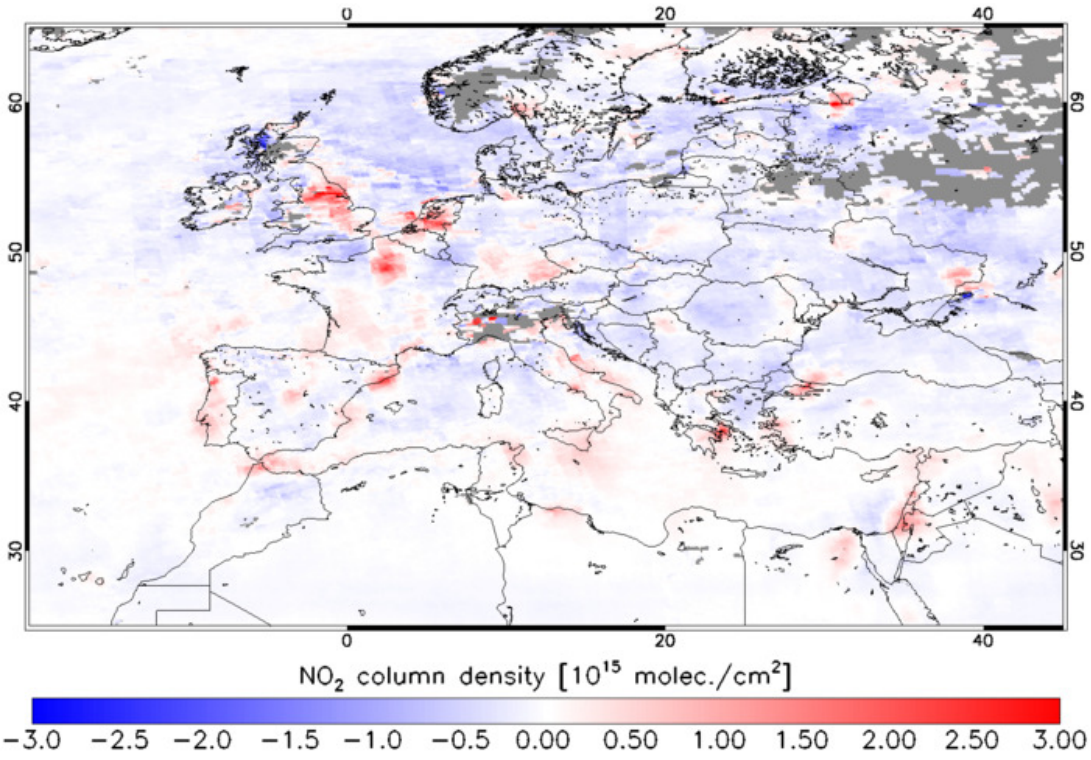

Figure 7: Tropospheric NO_2 from OMI retrieved with TM5 at a resolution of $1^\circ \times 1^\circ$ minus retrieved with TM5 at a resolution of $3^\circ \times 2^\circ$ for 20–30 October 2004 over Europe.

the scene albedo A_{sc} and the scene pressure p_{sc} assuming a cloud fraction $f_{\text{eff}} = 0.0$ – the user of the data can then select which is the appropriate set to use. With the snow/ice flag, which will be incorporated in the TROPOMI support FRESCO+ cloud product, the NO_2 processing can select which of these two sets to use. The snow/ice flag will be transferred to the NO_2 data product.

6.4.4.4 Surface pressure

The (altitude dependent) AMFs in Eq. (11) depend on the surface pressure, p_{s}. For the TROPOMI NO_2 retrieval this information will be obtained from the TM5 model ($1^\circ \times 1^\circ$) driven by ECMWF meteorological data. Because the TM5 information is representative for spatially coarse pressures, the TM5 results will be corrected based on the method described in Zhou et al. [2009] and Boersma et al. [2011]. This correction computes a new surface pressure based on the difference between the corresponding spatially coarse terrain height and the actual, pixel-averaged terrain height based on a 3-km resolution digital elevation map [Maasakkers et al., 2013].

6.4.4.5 A priori vertical NO_2 profiles

A chemistry transport model (CTM) is considered to be the best source of information for a priori NO_2 vertical profiles. The baseline for the TROPOMI NO_2 retrieval algorithm is to use TM5 vertical NO_2 profiles simulated at a $1^\circ \times 1^\circ$ (longitude $\times$ latitude) spatial resolution for 34 layers. The a priori profiles are interpolated to the centre of the TROPOMI ground pixel based on four nearest neighbour TM5 cell centres. Using TM5 instead of TM4 constitutes a significant improvement in itself: TM5 v3 is a fully benchmarked model version (Huijnen et al. [2010a]; Huijnen et al. [2010b]; [ER3]), with more up-to-date NO_x emissions (from the RETRO-REAS inventory), chemistry, and ongoing improvements of ship, soil and lightning NO_x emission descriptions.

Using TM5 with a global $1^\circ \times 1^\circ$ resolution is an important improvement over previous global satellite NO_2 retrievals that used vertical profiles computed at spatial resolutions of $2^\circ \times 2.5^\circ$ or $3^\circ \times 2^\circ$ (e.g. Lamsal et al. [2010], Boersma et al. [2011]). Obviously, there are still spatial gradients in NO_2 concentrations over scales smaller than a degree, but a resolution of $1^\circ \times 1^\circ$ should capture the most relevant gradients much better than a resolution of $3^\circ \times 2^\circ$. Using higher resolution models in combination with the TROPOMI averaging kernels

will in effect further improve the spatial resolution in the a priori NO_2 fields for advanced users interested in regionally focused investigations (e.g. Huijnen et al. [2010b]).

The effect of the improved spatial resolution is illustrated by Figure 7, which shows the difference between averaged tropospheric NO_2 columns from the OMI sensor from 20–30 October 2004 retrieved with TM5 at $3^\circ \times 2^\circ$ and at $1^\circ \times 1^\circ$. The retrieval with the higher resolution profile shapes clearly captures the pollution hotspots in Europe (e.g. Madrid, Paris) much better, leading to more pronounced contrasts between the sources of pollution and background (ventilated) pollution. To better capture the sources of air pollution is an important target of the TROPOMI mission.

6.4.5 Averaging kernels

For each ground pixel, the TROPOMI data product will provide the averaging kernel. The averaging kernel for DOAS retrievals is defined as the altitude-dependent AMF ratioed (decoupled from the NO_2 vertical distribution) by the total air-mass factor [Eskes and Boersma, 2003]. The tropospheric averaging kernel can be obtained by scaling the kernel by M/M^{trop} (see [RD16]) and setting all elements of the kernel to zero above the tropopause layer, i.e. for $l > l_{\text{tp}}^{\text{TM5}}$. Similarly, the stratospheric averaging kernel can be obtained by scaling the kernel by M/M^{strat} and setting all elements of the kernel to zero up to and including the tropopause layer, i.e. for $l \leq l_{\text{tp}}^{\text{TM5}}$. Using the averaging kernel is important for data users who wish to minimise the discrepancies between the assumptions in the TROPOMI retrieval and their application of interest, for example for validation, data assimilation, or comparison to a model (e.g. Silver et al. [2013]).

6.4.6 De-striping the NO_2 data product

The OMI measurements show across-track stripes resulting from viewing zenith angle dependent calibration errors in the OMI backscatter reflectances. For the DOMINO NO_2 data product Boersma et al. [2011] developed an empirical post-hoc de-striping correction based on the daily mean across-track dependency of the NO_2 slant columns. This correction is applied in the final step of the NO_2 processing, i.e. after the conversion to vertical columns.

Given that TROPOMI will be measuring with a CCD detector similar to the one used by OMI, the possibility of stripes occuring in the TROPOMI data cannot be ruled out. For this reason an option will be included in the Level-2 processor that allows for a de-striping correction on the NO_2 slant column data, prior to the ingestion in the data assimilation scheme. Once TROPOMI Level-1b spectra are available we will investigate whether the de-striping correction option needs to be turned on.

6.5 Processing chain elements

6.5.1 Off-line (re)processing

The off-line (re)processing of the TROPOMI NO_2 retrieval algorithm, schematically displayed in Fig. 8, will take place at two locations (for more details, see [RD1]):

(1) The first step of the NO_2 processing system, illustrated in the top left part of Fig. 8, the DOAS retrieval, ingests the Level-1b spectra and will be running in the TROPOMI processing system, i.e. the PDGS at DLR. Also performed in the PDGS, in a separate processing chain (not shown), is the retrieval of cloud information, needed by several Level-2 data products. The processor uses the slant column and cloud cover data to assemble a "first guess" NO_2 vertical column product, based on the NO_2 vertical profile forecast made for the near-real time (NRT) processing at the observation date (cf. Sect. 6.5.2). The result of this serves as a backup NO_2 data product.

(2) The NO_2 backup data product is then transferred to the SDC at KNMI, where once a day the data of all orbits is ingested in the data assimilation / chemistry transport model TM5, as illustrated in the right part of Fig. 8, to provide the off-line NO_2 profiles and the "final" nominal NO_2 data product.

(3) The nominal NO_2 data product is then transferred back to the PDGS (bottom left part of Fig. 8), where it is made available for the users of the TROPOMI NO_2 data products via the TROPOMI website, the Sentinel-5P Core Service.

The motivation for this set-up is to take full advantage of the available processing elements at DLR and KNMI, and at the same time keep the number of data transfers limited. DLR will operate in the PDGS a suite of processors geared to handling large amounts of TROPOMI spectra, including the processing of NO_2 column

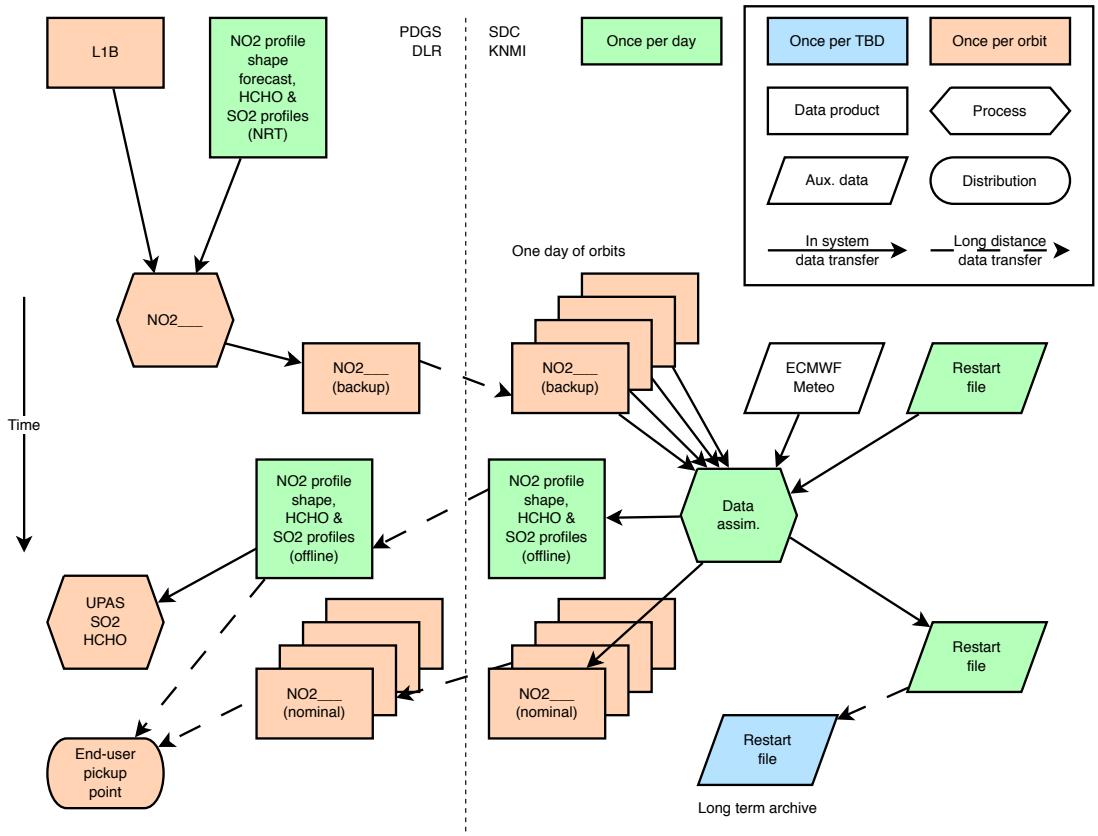

Figure 8: Schematic representation of the TROPOMI processing of tropospheric NO_2 data from a Level-1b spectrum received in the PDGS in the off-line mode. The dotted line marks the division of the processing locations: the Payload Data Ground Segment (PDGS) at DLR on the left and the Satellite Data Centre (SDC) at KNMI on the right. (Source of the figure: [RD1].)

data from TROPOMI spectra. The SDC at KNMI hosts a complete data assimilation system based on the Dutch TM5 model, and has considerable experience in both the off-line and on-line retrieval of NO_2 from the GOME, SCIAMACHY, OMI, and GOME-2 instruments. The essential inputs for the processing of TROPOMI NO_2 data are (1) the Level-1b spectra measured by TROPOMI at the PDGS, and (2) the ECMWF meteo data at the SDC.

As illustrated by Fig. 8, the data assimilation system not only provides vertical profiles for the processing of NO_2 data, but also for some other TROPOMI data products: formaldehyde (HCHO) and sulphur dioxide (SO_2). Unlike NO_2, HCHO and SO_2 are not assimilated in the TM5 model: their profiles are output of the TM5 model, based on the chemistry involving these species.

6.5.2 Near-real time processing

The near-real time (NRT) processing of TROPOMI NO_2 is based on the same principles as the off-line processing, described in Sect. 6.5.1 (Fig. 8). The main difference between the NRT processing, depicted in Fig. 9, and the off-line processing is the timing of the data assimilation step and the use of forecast ECMWF rather than analyzed ECMWF meteorological fields. For the NRT processing of TROPOMI data, the TM5 model is run once per day in the SDC at KNMI, and ingests the NO_2 slant columns from the orbits that have been observed thus far. Based on the assimilated "state" of day i, the TM5 model provides a forecast of the NO_2 vertical distribution for days $i+1$ until $i+5$. This information is then transferred to the PDGS, as illustrated in Fig. 9, for the NO_2 NRT data product.

This procedure ensures that as soon as new TROPOMI measurements are available in NRT, all necessary information from the TM5 model is ready to be processed in the PDGS to provide an NO_2 vertical column data product, without the need for a (time consuming) model run first. With NO_2 profile data available in a 5-day forecast, an interruption of the data stream from the SCD is not an immediate problem for the NRT processing system. In case the interruption lasts longer than 5 days, the PDGS processing system will use the latest

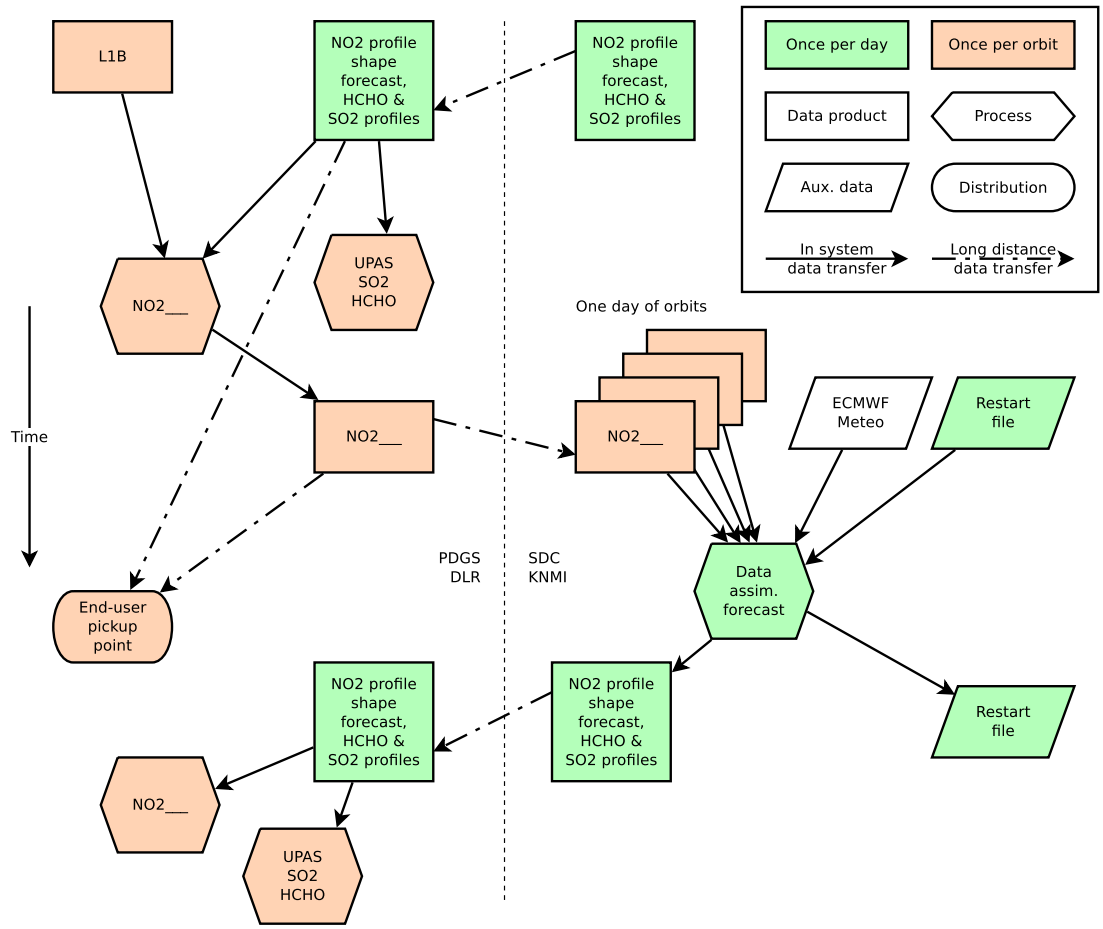

Figure 9: Schematic representation of the TROPOMI processing of tropospheric NO_2 data from a Level-1b spectrum received in the PDGS in the near-real time (NRT) mode. The dotted line marks the division of the processing locations: the Payload Data Ground Segment (PDGS) at DLR on the left and the Satellite Data Centre (SDC) at KNMI on the right. (Source of the figure: [RD1].)

available NO_2 profile as a fall-back to be able to continue provinding NO_2 data in NRT. As fall-back the latest available NO_2 profile is used, rather than NO_2 profile data from a climatology, because a switch to climatology data would constitute an evident discontinuity in the NO_2 data.

In the NRT processing, the TM5 data assimilation run is started just after midnight, as soon as the ECMWF meteo data has arrived. In that run, the system incorporates all the NO_2 slant column data that has been processed since the previous data assimilation run (from 24 hours before). Since the (ECMWF) forecast is provided up to 5 days ahead, the NRT processing is capable of providing tropospheric NO_2 data, even after a period of missing data or unforeseen slowdowns of the data assimilation itself. Previous analysis has shown [Boersma et al., 2007] that the forecast is accurate enough to provide reliable NO_2 tropospheric columns for a few days ahead.

6.6 The NO_2 data product

The final NO_2 vertical column data product for publication at the TROPOMI website – the Sentinel-5P Core Service – as official data product, shall have the data sets listed in Table 4 for each ground pixel. It consists of the data sets from the DOAS NO_2 retrieval, the data assimilation and the AMF calculation. Table 5 provides a list of six main classes of possible TROPOMI NO_2 data users and the data sets that these users will need for their usage.

In order to comply with the SI unit definitions, the TROPOMI NO_2 data product file gives trace gas concentrations in mol/m^2, rather than in the commonly used unit molec/cm^2. The multiplication factor to convert mol/m^2 to molec/cm^2 is 6.02214×10^{19} (the multiplication factor to convert mol/m^2 to DU is 2241.15). The O_2–O_2 concentration is given in mol^2/m^5; the multiplication factor to convert this to the commenly used unit

Table 4: Overview of data sets for each ground pixel in the final NO_2 data product assembled for dissemination via the TROPOMI website, the Sentinel-5P Core Service. Where relevant, the precision of a data set is provided as well. Data sets marked with $*$ are not part of the official Level-2 data product, but will be provided in a separate support data file. A more detailed overview can be found in Tables 10 and 11.

origin of data set	*for each ground pixel*	*symbols*
Level-1b spectrum	measurement time	t
	ground pixel centre and corner coordinates	ϑ_{geo}, δ_{geo}
	viewing geometry data	θ_0, θ, ϕ_0, ϕ
Databases	surface albedo in the NO_2 window	$A_{\text{s,NO}_2}$
	surface albedo used for the cloud retrieval	A_{s}
	surface elevation and pressure	z_{s}, p_{s}
Cloud retrieval	cloud fraction and cloud pressure FRESCO	f_{eff}, p_{c}
	scene pressure and scene albedo FRESCO	p_{sc}, A_{sc}
	cloud fraction in the NO_2 window	$f_{\text{eff,NO}_2}$
	cloud radiance fraction in the NO_2 window	w_{NO_2}
DOAS retrieval	NO_2 slant column	$N_{\text{s,NO}_2}$
	slant columns of secondary trace gases	$N_{\text{s,O}_3}$, $N_{\text{s,H}_2\text{O}_{\text{vap}}}$, $N_{\text{s,}\ldots}$
	Ring effect coefficient	C_{ring}
	polynomial coefficients	$a_m \ \ [m = 0, 1, ..., N_p]$
	wavelength calibration coefficients	w_s, w_q
	RMS and χ^2 of the fit	R_{RMS}, χ^2
Data assimilation & AMF calculation	NO_2 tropospheric vertical column	$N_{\text{v}}^{\text{trop}}$
	NO_2 stratospheric vertical column	$N_{\text{v}}^{\text{strat}}$
	NO_2 total vertical columns	$N_{\text{v}} \equiv N_{\text{s}}/M$, $N_{\text{v}}^{\text{sum}} \equiv N_{\text{v}}^{\text{trop}} + N_{\text{v}}^{\text{strat}}$
	tropospheric, stratospheric and total AMF	M^{trop}, M^{strat}, M
	averaging kernel	$\mathbf{A}$
	TM5 tropopause layer index	$l_{\text{tp}}^{\text{TM5}}$
	TM5 pressure level coefficients	A_l^{TM5}, B_l^{TM5}
	$*$ NO_2 profile for stratosphere and troposphere	n_{l,NO_2}
	$*$ TM5 temperature profile	T_l^{TM5}
	$*$ TM5 surface elevation and pressure	$z_{\text{s}}^{\text{TM5}}$, $p_{\text{s}}^{\text{TM5}}$
Flags	quality assurance value	—
	processing quality flags	—
	absorbing aerosol index	—
	snow/ice flag and land/water classification	—

molec^2/cm^5 is 3.62662×10^{37}.

The output for each ground pixel will be accompanied by two flags indicating the status of the results of the processing. The "quality assurance value" is a continuous variable, ranging from 0 (no output) to 1 (all is well). Warnings that occur during processing or results of the processing can be reasons to decrease the flag value. The "processing quality flags" contains the individual event that led to processing failure, or a precise record of the warnings that occurred during processing. The definitions and usage of these flags will be harmonised between the Level-2 data products of TROPOMI and will be documented elsewhere. The NO_2 data product has the Absorbing Aerosol Index (AAI) as flag, as additional information for the NO_2 data users, both in the off-line and the NRT processing mode.

The data product will consist of two files: one with the main retrieval results and a separate file with vertical information on atmospheric temperature and the NO_2 profile at the $1^\circ \times 1^\circ$ grid of TM5 on a half-hourly basis. The additional data sets will probably not be used by most NO_2 data users, but have shown to be useful to advanced users interested in improving particular aspects of the retrieval for their detailed (regional) purposes.

Table 5: Overview of different user applications of NO_2 data and the data sets from the TROPOMI NO_2 data product the users will need. In addition all users may need pixel related data, such as measurement time, geolocation, viewing geometry, quality flags, etc.

	user application	*data sets needed*
# 1	Tropospheric chemistry / air quality model evaluation and data assimilation Validation with tropospheric NO_2 profile measurements (aircraft, balloon, MAX-DOAS)	$N_\mathrm{v}^\mathrm{trop}$, $\sigma\left(N_\mathrm{v}^\mathrm{trop}\right)$ M^trop, M, $\mathbf{A}$ † A_l^TM5, B_l^TM5, $l_\mathrm{tp}^\mathrm{TM5}$, p_s
# 2	Tropospheric column comparisons, e.g. with other NO_2 column retrievals	$N_\mathrm{v}^\mathrm{trop}$, $\sigma\left(N_\mathrm{v}^\mathrm{trop}\right)$
# 3	Stratospheric chemistry model evaluation and data assimilation Validation with stratospheric NO_2 profile measurements (limb/occultation satellite observations)	$N_\mathrm{v}^\mathrm{strat}$, $\sigma\left(N_\mathrm{v}^\mathrm{strat}\right)$ M^strat, M, $\mathbf{A}$ ‡ A_l^TM5, B_l^TM5, $l_\mathrm{tp}^\mathrm{TM5}$, p_s
# 4	Stratospheric column comparisons, e.g. with ground-based remote sensors	$N_\mathrm{v}^\mathrm{strat}$, $\sigma\left(N_\mathrm{v}^\mathrm{strat}\right)$
# 5	Whole atmosphere (troposphere + stratosphere) data assimilation systems	N_v, $\sigma\left(N_\mathrm{v}\right)$ § $\mathbf{A}$ A_l^TM5, B_l^TM5, $l_\mathrm{tp}^\mathrm{TM5}$, p_s
# 6	Visualisation of the NO_2 product	$N_\mathrm{v}^\mathrm{trop}$, $N_\mathrm{v}^\mathrm{strat}$, $N_\mathrm{v}^\mathrm{sum}$ §

† The tropospheric kernel $\mathbf{A}^\mathrm{trop}$ is derived from the total kernel $\mathbf{A}$ and the air-mass factors M and M^trop.
‡ The stratospheric kernel $\mathbf{A}^\mathrm{strat}$ is derived from the total kernel $\mathbf{A}$ and the air-mass factors M and M^strat.
§ Note that the total NO_2 vertical column $N_\mathrm{v} \equiv N_\mathrm{s}/M$ is *not* the same as the sum $N_\mathrm{v}^\mathrm{sum} \equiv N_\mathrm{v}^\mathrm{trop} + N_\mathrm{v}^\mathrm{strat}$

7 Feasibility

The baseline approach for the TROPOMI NO_2 data product is to provide a full retrieval for all S5P ground pixels in an orbit. The only criteria that will be used to start the slant column retrieval is whether the Level-1b spectra for NO_2 and cloud retrieval are appropriate, i.e. flagged as valid. Subsequently, the set of NO_2 slant columns and cloud parameters is input to further processing to determine the tropospheric and stratospheric vertial column data. At this stage pixels only need to pass general validity and solar zenith angle (SZA $< 88^\circ$) criteria.

7.1 Required input

The processing of TROPOMI NO_2 data poses different demands for different retrieval steps. As described in Sect. 6.5 and illustrated in Figs. 8 and 9, the main processing will take place at the PDGS at DLR, which ingests information of a data assimilation system running in the SDC at KNMI, in order to convert the NO_2 slant column data into the respective tropospheric and stratospheric vertial column data.

The dynamic and static input data needed in the PDGS for the off-line and NRT processing of the NO_2 data product are listed below and summarised in Tables 6 and 7, respectively. Table 6 also mentions what the fall-back is in the processing of a given ground pixel, in case the dynamical data is not available. The FRESCO/KNMI cloud product is a TROPOMI Level-2 support product, provided by KNMI software running in the PDGS. The S5P/DLR cloud product is optional; the actual use of this product will be investigated post-launch. For the snow/ice cover data, NISE [ER7] and ECMWF assimilated data are requested, at least one is required, with daily updates near the polar region, less frequent updates closer to the equator. (See also the general TROPOMI documents [RD11] and [RD1].)

7.1.1 Spectral fitting inputs

In the PDGS at DLR, the following input is required, making a distinction between: (a) static (constant) input data and dynamic input data, which changes every orbit, and (b) data needed for the spectral fitting, and information needed in the subsequent processing step. After the DOAS NO_2 retrieval, the PDGS assembles the NO_2 vertical column data product, using information from the data assimilation system from the SDC at KNMI for further processing, as llustrated in Fig. 8.

Spectral fitting input data for the DOAS fit

- Dynamic input:
  - Level-1b Earthshine and Solar spectra
- Static input:
  - Reference spectra (convolved with the TROPOMI slit function) for NO_2, O_3, H_2O_{vap}, O_2–O_2, H_2O_{liq}, Ring effect

NO_2 data product input data

- Dynamic input:
  - NO_2 slant column density & errors from the DOAS fit
  - NO_2 profile shape from the data assimilation system
  - Geolocation data (incl. pixel corner coordinates)
  - Viewing geometry
  - Effective cloud fraction and cloud pressure
  - Snow and ice cover data
- Static input:
  - Pixel-average representative (interpolated) surface albedo at:
    - 439 nm (representative for the NO_2 window)
  - Pixel-average representative (interpolated) terrain height from a digital elevation map, including a land/water classification
  - Altitude-dependent AMF look-up table
  - Cloud fraction look-up table

Table 6: Overview of the dynamic input data needed for both the off-line and the NRT NO_2 data processing in the PDGS. The table does not list the input needed by the data assimilation system in the SDC. See Sect. 7.1 for further remarks.

name/data	*symbol*	*unit*	*source*	*pre-process needs*	*backup if not available*	*comments*
S5P Level-1b Earth radiance VIS band	$I(\lambda)$	mol/s/m^2/nm/sr	S5P Level-1b product	per pixel	no retrieval	—
S5P Level-1b Solar irradiance VIS band	$E_0(\lambda)$	mol/s/m^2/nm	S5P Level-1b product	per pixel	use previous	—
NO_2 profile shape	$n_{l,\mathrm{NO_2}}$	mol/m^2	TM5 model	per pixel	latest available [†] N/A	NRT off-line
FRESCO/KNMI cloud product	f_{eff} p_{c}	1 Pa	S5P Level-2 support product	—	no VCD product	—
S5P/DLR cloud product	f_{eff} p_{c}	1 Pa	S5P Level-2 cloud product	—	no VCD product	optional
snow/ice cover	—	—	NISE / ECMWF	per pixel	latest available [†] climatology	NRT off-line
aerosol absorbing index	AAI	1	S5P Level-2 AAI product	—	set AAI fill value set AAI fill value	NRT off-line

[†] Latest available value for that day.

Table 7: Overview of the static input data needed for both the off-line and the NRT NO_2 data processing in the PDGS. The table does not list the input needed by the data assimilation system in the SDC. See Sect. 7.1 for further remarks.

name/data	*symbol*	*unit*	*source*	*pre-process needs*	*comments*
absorption cross sections					
NO_2	$\sigma_{\mathrm{NO_2}}(\lambda)$	m^2/mol	Vandaele et al. [1998]	convolution	—
O_3	$\sigma_{\mathrm{O_3}}(\lambda)$	m^2/mol	Gronshelev et al. [2014] & Serdyuchenko et al. [2014]	convolution	—
O_2–O_2	$\sigma_{\mathrm{O_2–O_2}}(\lambda)$	m^5/mol^2	Thalman and Volkamer [2013]	convolution	—
H_2O_{vap}	$\sigma_{\mathrm{H_2O_{vap}}}(\lambda)$	m^2/mol	HITRAN 2012 data	convolution	[†]
H_2O_{liq}	$\sigma_{\mathrm{H_2O_{liq}}}(\lambda)$	1/m	Pope and Frey [1997]	convolution	—
Ring reference spectrum	$I_{\mathrm{ring}}(\lambda)$	mol/s/m^2/nm/sr	Chance and Spurr [1997]	convolution	[†]
retrieval input settings	—	—	KNMI	—	—
air-mass factor lookup table	—	—	KNMI	—	—
cloud fraction lookup table	—	—	KNMI	—	[‡]
digital elevation map, with land/water classification	—	—	—	per pixel	—
surface albedo database	$A_{\mathrm{s,NO_2}}$	1	Kleipool et al. [2008]; [ER5]	per pixel	—

[†] Created e.g. as in Van Geffen et al. [2015]; see also [RD11].
[‡] For the cloud fraction retrieval in the NO_2 fit window and for the cloud radiance fraction.

7.1.2 Data assimilation and air-mass factor inputs

In the SDC at KNMI, the NO_2 slant column data received from the PDGS is used in the data assimilation system to determine the NO_2 profile shape needed for the conversion of the NO_2 slant columns from the DOAS fit into the stratospheric and tropospheric NO_2 columns at the PDGS. For this step the following input is required, where making a distinction between the input needed for the data assimilation system.

Data assimilation input data

- Dynamic input:
  - NO_2 slant column density + errors
  - ECMWF meteorological fields (pressure, temperature, wind, ...)
  - NO_2 vertical profile shape
  - Geolocation data (incl. pixel corner coordinates)
  - Viewing geometry

Table 8: Estimate of the computational effort for the off-line TROPOMI NO_2 processing. Any delays introduced by the different processing steps having to wait for data to be available are not included. The numbers given for the data assimilation and AMF conversion are based on tests with the current setup of OMI data; it is probable that for TROPOMI these codes will be run multi-threaded.

	Time needed for processing one TROPOMI orbit	*Time needed for processing one day of TROPOMI data*
Spectral fitting	10 min (10 processors)	2 hours
Data transfer DLR $\rightarrow$ KNMI	$<$ 1 min	$<$ 15 min
Data assimilation with TM5	20–30 min (one core)	5–6 hours
Data transfer KNMI $\rightarrow$ DLR	$<$ 1 min	$<$ 15 min
AMF conversion	10–15 min (one core)	3–4 hours
Total processing time	40–55 min	10–12 hours

  - Effective cloud fraction and cloud pressure
  - Snow and ice cover data
- Static input:
  - NO_x (and other) emission inventories
  - Pixel-average representative (interpolated) surface albedo at:
    - 439 nm (representative for the NO_2 window)
    - 758 nm (representative for the O_2 A-band)

7.2 Estimated computational effort

The algorithm for the spectral fitting of slant columns, described in Sect. 6.2, has been tested extensively on actual OMI Level-1b spectra under all possible conditions. For OMI, the DOAS retrieval of NO_2 slant column densities takes about 12 minutes per orbit. The DOAS retrieval step is performed at NASA on a machine with a Intel(R) Xeon(R) CPU E5420 2.50GHz processor, with a SPEC cpu 2006 of about 20 [ER8], and the resulting data file, the OMNO2A data product, is then transferred to KNMI for further processing. A little more than 50% of the OMI Level-2 data, including the NO_2 slant column, is available half an hour after acquisition, i.e. after the last measurement of a given orbit. And about 99% of the OMI Level-2 data is available one hour after acquisition.

With TROPOMI having about 8.6 times as many ground pixels as OMI does (about 860,000 versus about 100,000 per orbit), the production of an orbit of TROPOMI NO_2 slant column densities is expected to take about 100 minutes on a single processing core. The baseline for TROPOMI spectral fitting is that 10 processors will be used in parallel at the PDGS to limit the computational time for producing the TROPOMI slant columns to 10 minutes. The code will be developed at KNMI in C++ and transferred and tested at DLR.

While not truly part of the computational effort, the transfer of slant columns from the PDGS to the SDC may take a considerable amount of time. For the OMI NO_2 data processing, the OMNO2A slant column data product is transferred from NASA to KNMI at a rate of approximately 1 Mb/sec. Since the OMNO2A files are about 13 Mb per orbit, the data transfer takes less than 1 min. The TROPOMI NO_2 slant column data files will be larger than the OMNO2A file (see Sect. 7.4), but the data connection between DLR and KNMI is faster and the transfor of a single TROPOMI slant column file can be expected to take less than 1 min.

The off-line processing of TROPOMI NO_2 data relies largely on the speed of running the TM5 model with full chemistry at a resolution of $1^\circ \times 1^\circ$. One full day worth of data assimilation at $1^\circ \times 1^\circ$ of OMI data currently takes approximately 6 hours on a workstation using one processor. When running TM5 in parallel (in the KNMI SDC), a speed-up from 6 to 2 hours for one full day can be foreseen. The assimilation is not expected to slow down substantially because of the larger amount of TROPOMI ground pixels, as the TROPOMI pixels will be binned to so-called superobservations at $1^\circ \times 1^\circ$ (Sect. 6.3). The number of superobservations to be assimilated will be thinned out (via a checkerboard approach) to reduce the computational burden.

The NO_2 profile shape determined by the TM5 model is transferred to the PDGS. In the NRT processing this transfer takes place once a day (cf. Fig. 9); in the off-line mode the transfer rate remains to be determined (cf. Fig. 8). The subsequent AMF conversion of NO_2 slant columns in the PDGS will be somewhat slower than in the OMI case, because of the larger amount of pixels. The conversion from the DOAS retrieved slant column to the tropospheric NO_2 vertical column is based on a look-up table. For one OMI orbit, this conversion takes

Table 9: Estimate of the computational effort for the near-real time TROPOMI NO_2 processing. Any delays introduced by the different processing steps having to wait for data to be available are not included.

	Time needed for processing one TROPOMI orbit in NRT
Spectral fitting	10 min (10 processors)
Data transfer DLR $\rightarrow$ KNMI	$<$ 1 min
Data assimilation with TM5	N/A
Data transfer KNMI $\rightarrow$ DLR	$<$ 15 min (once a day)
AMF conversion	10–15 min (one core)
Total processing time	20–30 min

about 1.5 minutes, so for TROPOMI it is estimated to take between 10 and 15 minutes.

Table 8 provides a conservative estimate of the computational effort for the off-line processing of all TROPOMI pixels. This estimate is based on our current TROPOMI test set-up that is being applied on OMI data. We foresee that implementing the TM5 data assimilation system on the KNMI SDC will reduce the computational time by 4 hours, which would bring the total processing time within 12 hours for one full day.

7.3 Near-real time timeliness

For the NRT Level-2 data to be available within the required 3 hours after measurement, it is required that the processing of Level-2 data does not take more than about 30 minutes per oribt.

The data assimilation run is done at KNMI once a day (just after midnight) to provide a forecast of the NO_2 profile shapes for the coming 5 days, based on assimilation of TROPOMI slant columns observed over the previous day. While such a forecast run will probably take a fair amount of processing time ($\approx$20 hrs), the NRT chain does not need to wait for this, as mentioned in Sect. 6.5.2. This constitutes an important speed-up compared to the off-line processing and effectively eliminates the time used for the TM5 run effort in the NRT processing.

Based on these considerations, Table 9 provides a conservative estimate of the computational effort for the NRT processing of all TROPOMI pixels of a given orbit. The times in this table do not include any delays introduced by time scheduling of the different process steps.

In the current OMI NO_2 NRT data processing steps at KNMI, the processes check once every 10 minutes whether new data is available: first whether new OMNO2A slant column files have arrived at KNMI, then whether new data is available for the AMF conversion, and finally whether there is new data for publication on the website. In the worst case situation, this therefore may introduce a delay of 30 min. For the TROPOMI NO_2 processing, it is clear that such inter-process delays will have to be reduced to a minimum.

7.4 NO_2 product description and size

The TROPOMI NO_2 data output product consists of the retrieved tropospheric and stratospheric NO_2 columns, along with error estimates and the (total) averaging kernel. A general overview of the data product contents is given in Sect. 6.6 and Table 4. Table 10 provides a more detailed overview of the data sets, their unit, type, etc. in the main output data product.

In order to comply with the SI unit definitions, the TROPOMI NO_2 data product file gives trace gas concentrations in mol/m^2, rather than in the commonly used unit molec/cm^2. The multiplication factor to convert mol/m^2 to molec/cm^2 is 6.02214×10^{19} (the multiplication factor to convert mol/m^2 to DU is 2241.15). The O_2–O_2 concentration is given in mol^2/m^5; the multiplication factor to convert this to the commenly used unit molec^2/cm^5 is 3.62662×10^{37}.

Given the number of data per ground pixel listed in Table 10, the main output file is estimated to be 450MB per TROPOMI orbit.

The averaging kernel describes how the retrieved NO_2 columns relate to the true NO_2 profile [Eskes and Boersma, 2003]. The averaging kernel should be used in validation exercises, model evaluations, and assimilation or inverse modelling attempts with TROPOMI NO_2 data. The output product will also contain the necessary information (surface pressure and TM5 sigma coordinates) to construct the pressure grid to which the averaging kernel values correspond.

Table 10: Overview of the data set units, types and sizes in the main data output product file, listed alphabetically; cf. Table 4. All quantities followed by a $*$ in the "symbol" column consist of the value and the associated precision (for these the number of data per pixel is doubled in the 6th column). The data sets in the support data file are listed in Table 11.

name/data	*symbol*	*unit*	*description*	*type*	*data per pixel*	*comments*
aerosol absorbing index	—	1	from S5P Level-2 AAI product	float	1	added as flag
air-mass factor	M^{trop}	1	tropospheric AMF	float	1	—
	M^{strat}	1	stratospheric AMF	float	1	—
	M	1	total AMF	float	1	—
averaging kernel	$\mathbf{A}$	1	—	float	N_l	†
chi-squared	χ^2	1	χ^2 of the NO_2 DOAS fit	float	1	cf. Eq. (2)
cloud albedo	A_{c}	1	used in the cloud retrieval	float	1	fixed at 0.8
cloud pressure	p_{c}	Pa	from the cloud retrieval	float	1	—
cloud radiance fraction	w_{NO_2}	1	for the NO_2 VCD	float	1	in NO_2 fit window
DOAS fit results	$N_{\text{s,NO}_2}$ *	mol/m^2	total NO_2 SCD	float	1 ×2	—
	$N_{\text{s,H}_2\text{O}_{\text{liq}}}$ *	m	H_2O_{liq} coeff. in NO_2 window	float	1 ×2	—
	$N_{\text{s,H}_2\text{O}_{\text{vap}}}$ *	mol/m^2	H_2O_{vap} SCD in NO_2 window	float	1 ×2	—
	$N_{\text{s,O}_2\text{–O}_2}$ *	mol^2/m^5	O_2–O_2 SCD in NO_2 window	float	1 ×2	—
	$N_{\text{s,O}_3}$ *	mol/m^2	O_3 SCD in NO_2 window	float	1 ×2	—
	C_{ring} *	1	Ring coeff. in NO_2 window	float	1 ×2	—
effective cloud fraction	$f_{\text{eff,NO}_2}$	1	for the NO_2 VCD	float	1	in NO_2 fit window
	f_{eff}	1	from the cloud retrieval	float	1	—
ghost column	$N_{\text{v}}^{\text{ghost}}$	mol/m^2	NO_2 column below the clouds	float	1	‡
ground pixel coordinates	δ_{geo}	°	VIS pixel – latitude	float	5	centre, 4 corners
	ϑ_{geo}	°	VIS pixel – latitude	float	5	centre, 4 corners
ground pixel index	—	1	across-track pixel index	int	1	—
land/water classification	—	1	surface classification	int	1	—
measurement time	t	s	VIS pixel	float	2	—
number of wavelengths	N_λ	1	in the NO_2 fit window	int	1	#
number of iterations	N_i	1	from the DOAS fit	int	1	—
polynomial coefficients	a_m *	1	in the NO_2 DOAS fit	float	(N_p+1) ×2	cf. Eq. (6) §
processing quality flags	—	1	—	int	1	cf. Sect. 6.6
quality assuarance value	—	1	—	float	1	cf. Sect. 6.6
root-mean-square	R_{RMS}	1	RMS of the NO_2 DOAS fit	float	1	cf. Eq. (4)
satellite coordinates	z_{sat}	m	altitude of the satellite	float	1	—
	δ_{sat}	°	latitude sub satellite point	float	1	—
	ϑ_{sat}	°	longitude sub satellite point	float	1	—
	φ_{sat}	1	relative offset in orbit	float	1	—
scanline index	—	1	along-track pixel index	int	1	—
scene albedo	A_{sc}	1	from the cloud retrieval	float	1	—
scene pressure	p_{sc}	Pa	from the cloud retrieval	float	1	—
snow-ice flag	—	1	snow/ice case flagging	int	1	—
surface albedo	$A_{\text{s,NO}_2}$	1	for cloud fraction NO_2 window	float	1	—
	A_{s}	1	for the cloud retrieval	float	1	—
surface elevation	z_{s} *	m	VIS pixel	float	1 ×2	—
surface pressure	p_{s}	Pa	VIS pixel	float	1	—
TM5 pressure level	A_l^{TM5}	Pa	—	float	0	¶
coefficients	B_l^{TM5}	1	—	float	0	¶
TM5 tropopause layer index	$l_{\text{tp}}^{\text{TM5}}$	1	—	int	1	—
vertical column density	$N_{\text{v,NO}_2}^{\text{trop}}$ *	mol/m^2	tropospheric NO_2 VCD	float	1 ×2	—
	$N_{\text{v,NO}_2}^{\text{strat}}$ *	mol/m^2	stratospheric NO_2 VCD	float	1 ×2	—
	$N_{\text{v,NO}_2}$ *	mol/m^2	total NO_2 VCD	float	1 ×2	$\equiv N_{\text{s}}/M$
	$N_{\text{v,NO}_2}^{\text{sum}}$ *	mol/m^2	total NO_2 VCD	float	1 ×2	$\equiv N_{\text{v}}^{\text{trop}} + N_{\text{v}}^{\text{strat}}$

Table continues on next page

Table 10: — *continued.*

name/data	*symbol*	*unit*	*description*	*type*	*data per pixel*	*comments*
viewing geometry data	θ_0	°	solar zenith angle	float	1	—
	ϕ_0	°	solar azimuth angle	float	1	—
	θ	°	viewing zenith angle	float	1	—
	ϕ	°	viewing azimuth angle	float	1	—
wavelength calibration	w_s *	nm	wavelength shift	float	1 ×2	cf. Eq. (7)
	w_q *	1	wavelength stretch	float	1 ×2	cf. Eq. (7)
	χ^2_w	1	χ^2 of the calibration	float	1	—

† The number of TM5 layers is $N_l = 34$ for the DOMINO v2.0 processing; this may change when the layer distribution is optimised.
‡ The NO_2 ghost column is the NO_2 profile shape from TM5 integrated from the surface to the cloud pressure level.
# The actual number of wavelengths N_λ used in the fit (cf. Eq. (2)), i.e. after removel of, for example, bad pixels within the fit window.
§ The degree of the DOAS polynomial is $N_p = 5$ in the current OMNO2A; there will be room in the data product to change this.
¶ One set of $N_l + 1$ (see note †) TM5 pressure level coefficients per data granule.

Table 11: Overview of the data set units, types and sizes in the support output product file; this file is also used to store the profiles of HCHO and SO_2, delivered along with the NO_2 profile by the TM5 model. The data is provided on the TM5 grid resolution of $1^\circ \times 1^\circ$ on a half-hourly basis, rather than on TROPOMI pixel basis. The data sets in the main data file are listed in Table 10.

name/data	*symbol*	*unit*	*description*	*type*	*data per grid cell*	*comments*
HCHO profile	$n_{l,\mathrm{HCHO}}$	1	volume mixing ratio	float	N_l	†
NO_2 profile	$n_{l,\mathrm{NO_2}}$	1	volume mixing ratio	float	N_l	†
SO_2 profile	$n_{l,\mathrm{SO_2}}$	1	volume mixing ratio	float	N_l	†
TM5 temperature profile	T_l^{TM5}	K	—	float	N_l	†
TM5 pressure level coefficients	A_l^{TM5}	Pa	—	float	0	¶
	B_l^{TM5}	1	—	float	0	¶
TM5 surface elevation	$z_{\mathrm{s}}^{\mathrm{TM5}}$	m	—	float	1	‡
TM5 surface pressure	$p_{\mathrm{s}}^{\mathrm{TM5}}$	Pa	—	float	1	—
TM5 tropopause layer index	$l_{\mathrm{tp}}^{\mathrm{TM5}}$	1	—	int	1	—
date & time	—	1	year, month, day, hour, min, sec	int	0	§
time	d	days	no. of days since 1 Jan. 1950	float	0	§

† The number of TM5 layers is $N_l = 34$ for the DOMINO v2.0 processing; this may change when the layer distribution is optimised.
¶ One set of $N_l + 1$ (see note †) TM5 pressure level coefficients per data file.
‡ This data set is provided via a separate static TROPOMI digital elevation map file.
§ One set per data file.

For advanced users, a separate support file will be made available that contains the temperature and NO_2 vertical profile. This data is given at the TM5 grid resolution of $1^\circ \times 1^\circ$ on a half-hourly basis, rather than on TROPOMI pixel basis, to reduce the size of the file that needs to be transferred; conversion of the gridded profile data to the TROPOMI ground pixel is done in the PDGS. The temperature and NO_2 profiles are not included in the standard Level-2 product, because most users will not need these and we wish to keep the size of the main TROPOMI Level-2 files reasonable. Nevertheless, the temperature and NO_2 profiles 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]. The support data file will also contain the vertical profiles for SO_2 and HCHO, in support of the respective TROPOMI data products, as mentioned in Sect. 6.6. Table 11 provides an overview of the data sets in the support output data product.

7.5 Selection of the data assimilation system to be used

For the separation of the total slant column into a tropospheric and stratospheric part, as outlined in Sect. 6, the data assimilation and AMF calculation system currently operational in the SDC at KNMI for data of past and current satellite instruments is planned to be used also for the TROPOMI NO_2 processing system.

Within the GMES Atmospheric Services, the MACC project (Monitoring Atmospheric Composition and Climate; see [ER9]) operates a large data assimilation system that also assimilates NO_2 data. The MACC forecast NO_2 profiles could, in principle, be used as alternative for the TM5 NO_2 a priori profiles, by applying

the averaging kernel to be provided in the TROPOMI NO_2 data product. The use of the MACC data assimilation system, however, is undesirable for the data assimilation of TROPOMI NO_2 slant columns for the following reasons.

1. The MACC data assimilation system uses NO_2 data from several retrievals, including NO_2 retrievals from KNMI, as well as data of other gases, in the same assimilation. This leads to potential, hard to quantify and hard to pinpoint feedbacks to the TROPOMI NO_2 retrieval from the MACC model.
2. Using the MACC data assimilation system for TROPOMI NO_2 data would introduce a discontinuity in the long-term NO_2 data record at KNMI, since the GOME-1, SCIAMACHY, OMI and GOME-2 NO_2 data are all derived using the same data assimilation system (currently the data assimilation is driven by the TM4 model, but this will be replaced by TM5 in 2016 and the NO_2 data records will be processed).
3. When using the the MACC data assimilation system, the TROPOMI NO_2 retrieval would be dependent on the implementation of the NO_2 chemistry and up-to-date NO_x emission inventories, which lies beyond the control of the TROPOMI NO_2 retrieval team.

8 Error analysis

The TROPOMI NO_2 retrieval algorithm generates stratospheric and tropospheric vertical column densities for all pixels. Since assumptions differ considerably for stratospheric and tropospheric retrievals, the error budget for each case will be treated separately below.

The overall error for the retrieved tropospheric columns is determined through propagation of the three main error sources: (a) measurement noise and spectral fitting affecting the slant columns, (b) errors related to the separation of stratospheric and tropospheric NO_2, and (c) systematic errors due to uncertainties in model parameters such as clouds, surface albedo, and a priori profile shape, affecting the tropospheric air-mass factor. For the stratospheric NO_2 column, the errors are driven by slant column errors, errors in the estimate of the stratospheric contribution to the slant column, and stratospheric AMF (observation operator) errors.

For NO_2, the overall error budget thus consists of several different error source terms. Errors in the slant columns will be driven in part by instrumental noise (random errors), and in part by necessary choices on the physical model and reference spectra used (systematic errors). Errors in the AMF will be mostly systematic (e.g. assumptions on albedo) but will also have random contributions (e.g. from observed cloud parameters). It is thus not possible to make a clear distinction between these error types in the total error reported in the TROPOMI NO_2 data product. This implies that by averaging TROPOMI pixels over time or over a larger area, the random part of the overall error can be largely eliminated, but systematic effects may still persist in averaged retrievals.

Experience with errors in OMI NO_2 over polluted regions, largely stemming from theoretical error analysis and practical validation studies, indicates that overall errors on the order of 25% for individual tropospheric NO_2 column retrievals may be expected. Validation studies show that the systematic part of this error is on the order of 10-15% (e.g. Hains et al. [2010]; Irie et al. [2012]; Ma et al. [2013]). For stratospheric NO_2 columns, the errors are considerably smaller and depend mostly on the absolute accuracy of the slant columns, and on the separation of the stratospheric and tropospheric contributions. The stratospheric NO_2 column error is expected to have errors on the order of 5-10% (e.g. Hendrick et al. [2012]).

8.1 Slant column errors

Instrument noise is the main source of errors in the spectral fitting of TROPOMI Level-1b spectra. The anticipated radiometric signal-to-noise ratio (SNR) of TROPOMI in the $400-500$ nm range is $800-1000$ for an individual Level-1b spectrum [RD4]. Experience with OMI spectral fitting in the $405-465$ nm spectral domain showed that the uncertainty in OMI NO_2 slant column densities retrieved with very similar signal-to-noise ratios as for TROPOMI is on the order of 0.7×10^{15} molec/cm^2 ($< 10-20$%) [Boersma et al., 2007]. The quoted OMI uncertainty contains contributions from striping effects that may not occur for the TROPOMI sensor, so 0.7×10^{15} molec/cm^2 is adopted as a conservative estimate for the TROPOMI slant column error. Other, potentially systematic, errors include inaccuracies in the NO_2 cross-section spectrum (Vandaele et al. [1998]; [ER2]), and in the temperature dependence of the NO_2 cross section, but these have been shown to be of little concern for the slant column errors [Boersma et al., 2002].

Figure 10 shows as function of the signal-to-noise ratio (SNR) an estimate of the uncertainty of the retrieved slant column density determined by a DOAS fit in the wavelength window $405-465$ nm with polynomial degree 5. Spectra were simulated with a radiative transfer code using an atmosphere with two NO_2 profiles, taken from the CAMELOT study [RD8], with the same profile shape in the stratosphere:

(*a*) European background profile, simulated with a total vertical column $N_v = 2.5\times10^{15}$ molec/cm^2

(*b*) European polluted profile, simulated with a total vertical column $N_v = 7.5\times10^{15}$ molec/cm^2

The simulations are performed with surface albedo $A_s = 0.05$, no clouds, solar zenith angle $\theta_0 = 50^\circ$, and looking down in nadir. The legend of Fig. 10 gives the total slant column N_s in 10^{15} molec/cm^2. The retrieved N_s varies very little with the SNR: about 3×10^{12} molec/cm^2 between SNR$=700$ and 1100. For profile (*a*) the retrieved N_s is within 5% of the inital N_s and for profile (*b*) it is within 3%. Given this a good accuracy of the DOAS fits can be expected, with uncertainties in the range of $10-15\%$ for background NO_2 cases and $5-10\%$ for polluted cases.

8.2 Errors in the stratospheric (slant) columns

Data assimilation of TROPOMI NO_2 slant columns in TM5 provides the estimate of the stratospheric contribution to the NO_2 slant columns. The accuracy of these estimates is largely determined by the accuracy of the slant columns, as the TM5 stratospheric NO_2 distributions are scaled to become consistent with the retrieved slant

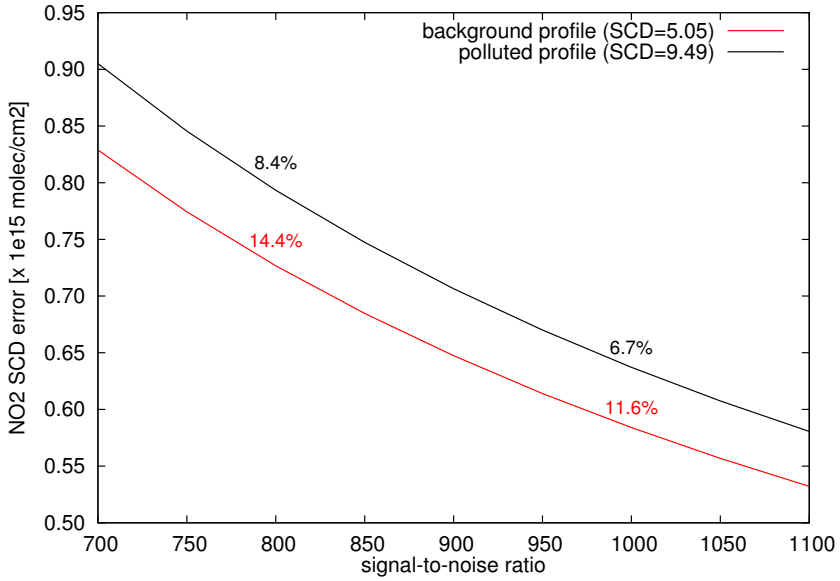

Figure 10: DOAS retrieval slant column uncertainty estimate [in 10^{15} molec/cm^2] as function of the SNR for two NO_2 profiles. The plot legend gives the retrieved slant column in 10^{15} molec/cm^2. At SNR equal 800 and 1000 the relative slant column uncertainty is plotted. For further details see the text.

columns. Random error estimates are derived from the assimilation approach: a considerable advantage of the assimilation scheme is that it provides a statistical estimate of the uncertainties in the stratospheric (slant) columns through the standard deviation of the differences between the TM5 model analysis and forecast stratospheric NO_2 ("A–F"). Generally, the uncertainty for the stratospheric NO_2 columns is of the order of $0.1-0.2\times10^{15}$ molec/cm^2 [Dirksen et al., 2011], much smaller than the slant column uncertainty. This is partly the result of the using superobservations to strongly reduce random errors in the stratospheric slant column estimates. Figure 11 shows the average A–F difference for 5–15 November 2004 in the data assimilation system based on TM5, which will be the default for the TROPOMI NO_2 processing. The A–F differences are generally within 0.5×10^{15} molec/cm^2, suggesting similar uncertainties for stratospheric NO_2 columns.

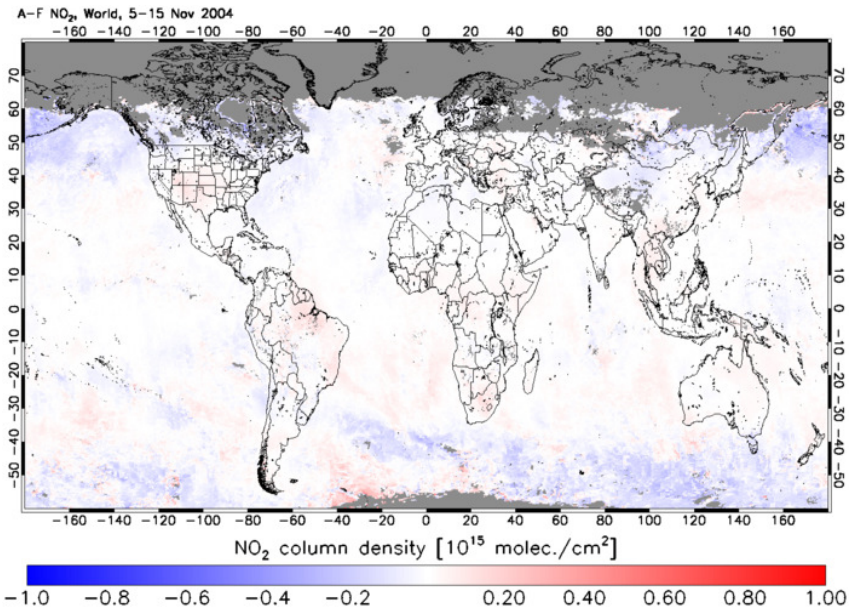

Figure 11: The difference between the model forecast before and after the analysis ("A–F") in the data assimilation of the NO_2 slant column divided by the geometric air-mass factor (right panel) averaged for 5–15 Nov. 2004, based on OMI data, assimilated with TM5 at $3^\circ\times2^\circ$ spatial resolution.

Table 12: Estimate of the error in the AMF due to several error sources ('BL' stands for Boundary Layer.) The estimated AMF errors are considered to be representative of 'typical' retrieval scenarios over regions of interest, i.e. with substantial NO_2 pollution for mostly clear-sky situations, and non-extreme boundary conditions for surface albedo and pressure.

Error type	*Estimated error*	*Corresponding AMF error*
Cloud fraction	± 0.02	$\pm 10\%$
Cloud pressure	± 50 hPa	$\pm[0-10]\%$
Surface albedo	± 0.015	$\pm 10\%$
Surface pressure	± 20 hPa	$\pm[0-5]\%$
A priori NO_2 profile shape	BL height & mixing schemes	$\pm 10\%$
A priori NO_x emissions	$\pm[0-25]\%$	$\pm[0-10]\%$
Aerosol-related errors		$\pm[0-10]\%$
Overall error		$\pm[15-25]\%$

Forward (radiative transfer) model calculations are important for, but contribute little to errors in the assimilation procedure. The observation operator **H** (see Eq. (8)) is proportional to the averaging kernel [Eskes and Boersma, 2003], the vector that contains the vertical sensitivity of TROPOMI to NO_2 in each layer. The scalar product of the observation operator vector and the TM5 NO_2 profile at the location of the individual TROPOMI observations yields the slant column that would be observed by TROPOMI given the modeled profile. Stratospheric radiative transfer calculations around 439 nm are relatively straightforward compared to those for the troposphere, where multiple scattering occurs, and the effects of clouds and aerosols interact with the vertical distribution of NO_2. The main forward model parameter influencing errors in the stratospheric estimate is the a priori stratospheric NO_2 profile shape (and associated temperature correction), but sensitivity tests suggest that uncertainties in the exact shape of this profile are of little influence to the overall error of the stratospheric NO_2 column.

One potential source of error is the sphericity correction in the radiative transfer model. These errors are negligible for most viewing geometries, but may need to be considered for far off-nadir viewing angles and high solar zenith angles. Spurr [2002] evaluated the difference between a regular and an enhanced pseudo-spherical correction on the modelled intensities around 325 nm. These differences provide an estimate for errors in the sphericity correction and they decrease with wavelength and increase with altitude. For the far off-nadir viewing angles of OMI and TROPOMI ($\pm 55^\circ$ from nadir) at SZA$= 85^\circ$ the differences are around 3% for 335.5 nm [Spurr, 2002]. For stratospheric NO_2, at higher wavelengths and located comparable altitude as ozone, the error due to errors in the sphericity correction is estimated to be smaller than 2%.

8.3 Errors in the tropospheric air-mass factors

The tropospheric air-mass factor (AMF) is calculated with a forward model (here version 3.2 of the DAK radiative transfer model) and depends on the a priori assumed profile shape and forward model parameters (cloud fraction, cloud pressure, surface albedo, surface pressure and aerosol properties). The AMF also depends on the solar zenith, viewing zenith and relative azimuth angles, but the measurement geometry is known with high accuracy and therefore does not contribute significantly to the AMF errors. The forward model itself is assumed to represent the physics of the measurement accurately, so that forward model errors can be characterised in terms of model parameters only.

The most important AMF errors are cloud fraction, surface albedo, and a priori profile shape. Cloud parameters are obtained from TROPOMI observations, and these will have random as well as systematic components. Surface albedo and NO_2 profile shape are obtained from a priori assumptions (i.e. a pre-calculated climatology and CTM simulations), and much depends on the accuracy of these assumptions that will be different for different retrieval situations (e.g. season, surface type etc.). Because the retrieved cloud parameters depend on similar (if not the same) surface albedo assumptions as the NO_2 air-mass factors, errors will be dampened to some extent [Boersma et al., 2004].

In Table 12 the most probable uncertainties of the forward model parameters to provide a cautious error prediction for TROPOMI NO_2 AMFs are listed. For this the theoretical error propagation framework used in Boersma et al. [2004] is followed. This approach takes into account the sensitivity of the AMF to uncertainties around the actual value of a particular forward model parameter (e.g. the AMF is much more sensitive to albedo

Table 13: Tropospheric NO_2 vertical column values and per pixel uncertenty estimates based on OMI data for selected regions averaged over 5–15 November 2004 using cloud-free ground pixels only,

Region	*Average tropospheric column*	*Average pixel uncertainty*	*box size ranges*	
			longitude	*latitude*
World	0.52×10^{15} molec/cm^2	0.58×10^{15} molec/cm^2 [111%]	$-180:180$	$-90:90$
Pacific	0.24×10^{15} molec/cm^2	0.39×10^{15} molec/cm^2 [163%]	$-180:-140$	$-50:25$
China	3.13×10^{15} molec/cm^2	1.51×10^{15} molec/cm^2 [48%]	$85:125$	$30:40$
USA East	4.45×10^{15} molec/cm^2	2.14×10^{15} molec/cm^2 [48%]	$-90:-68$	$35:45$
Europe	1.99×10^{15} molec/cm^2	1.39×10^{15} molec/cm^2 [70%]	$-20:45$	$35:65$
Netherlands	6.58×10^{15} molec/cm^2	3.49×10^{15} molec/cm^2 [53%]	$4:7$	$51:53.5$
Barcelona	6.30×10^{15} molec/cm^2	1.82×10^{15} molec/cm^2 [29%]	$1.5:2.5$	$41:42$

errors for dark surfaces than for brighter surfaces).

Aerosol-related errors are intimately coupled to cloud parameter errors. The O_2 A-band cloud algorithm currently does not correct for the presence of aerosols, so that an effective cloud fraction and cloud pressure will be retrieved. It is a matter of ongoing research whether or not the disentanglement of aerosol and cloud effects will improve the quality of the AMFs (Leitão et al. [2010]; Boersma et al. [2011]; Lin et al. [2014]).

The results in Table 12 provide a general estimate of air-mass factor errors that may be expected for TROPOMI NO_2 data under polluted conditions. But error analysis for individual retrievals show considerable variability on these estimates [Boersma et al., 2004]. For instance, regions with a low surface albedo are very sensitive to albedo uncertainties, and this can be reflected in AMF errors of more than 50%. For TROPOMI NO_2 a full error propagation that takes these sensitivities into account will be provided, and as well as a unique error estimate for every pixel.

8.4 Total errors in the tropospheric NO_2 columns

The overall error in the TROPOMI tropospheric NO_2 columns is driven by error propagation of the error terms discussed before, i.e. (1) slant column errors, (2) errors associated with the separation of the stratospheric and tropospheric contributions to the slant column, and (3) tropospheric air-mass factor errors.

The overall error variance for each pixel is written as in Boersma et al. [2004]:

$$\left\langle \varepsilon^2 \right\rangle = \left(\frac{\sigma(N_\mathrm{s})}{M^\mathrm{trop}} \right)^2 + \left(\frac{\sigma(N_\mathrm{s}^\mathrm{strat})}{M^\mathrm{trop}} \right)^2 + \left(\frac{(N_\mathrm{s} - N_\mathrm{s}^\mathrm{strat}) \cdot \sigma(M^\mathrm{trop})}{(M^\mathrm{trop})^2} \right)^2 \tag{18}$$

with $\sigma(N_\mathrm{s})$ the slant column error (0.7×10^{15} molec/cm^2), $\sigma(N_\mathrm{s}^\mathrm{strat})$ the stratospheric slant column error (0.2×10^{15} molec/cm^2) and $\sigma(M^\mathrm{trop})$ the estimated error in the tropospheric air-mass factor ($\pm25\%$). We see immediately that the total error depends on details in the retrieval and therefore differs from one pixel to the next. For small tropospheric excess slant columns, the overall retrieval uncertainty is dominated by the random errors in spectral fitting, whereas for large tropospheric slant columns, the retrieval uncertainty is dominated by air-mass factor uncertainties (the last term in Eq. (18)).

Figure 12 shows the absolute and relative error in the tropospheric NO_2 column retrieved for clear-sky scenes from OMI data on 10 November 2004. We see that over the oceans and the remote continental regions, the overall tropospheric retrieval uncertainty is dominated by errors in the spectral fitting and the stratospheric column estimate and is more than 100% (indicated by purple colours in the bottom panel of Fig. 12). For larger columns over continental areas, the relative uncertainty in the retrieved column reduces to $20-50\%$, and is dominated by the uncertainty in the tropospheric air-mass factor. Retrieval results are generally best for regions with strong NO_2 sources and/or high surface albedos.

Based on the instrumental performance expectations for TROPOMI, and our experience with OMI tropospheric NO_2 retrievals (see Fig. 12 and Table 13), the overall error budget for individual TROPOMI tropospheric NO_2 retrievals can tentatively be approximated as $\varepsilon = 0.5\times10^{15}$ molec/cm$^2 + [0.25-0.50]\cdot N_\mathrm{v}^\mathrm{trop}$. This is a more complete and realistic error statement than the requirements from [RD5] ($\varepsilon = 1.3\times10^{15}$ molec/cm$^2 + 0.1\cdot N_\mathrm{v}^\mathrm{trop}$ for a horizontal resolution of $5-20$ km; cf. Table 1). Note that in Table 13 the entries for The Netherlands and Barcelona are the most realistic ones as they represent small "hotspot" sized areas: the errors given for these two (53% and 29%, resp.) are in agreement with the above mentioned error estimate. Further note that for the numbers in Table 13 only an 11-day period has been used: the errors can be quite different for other

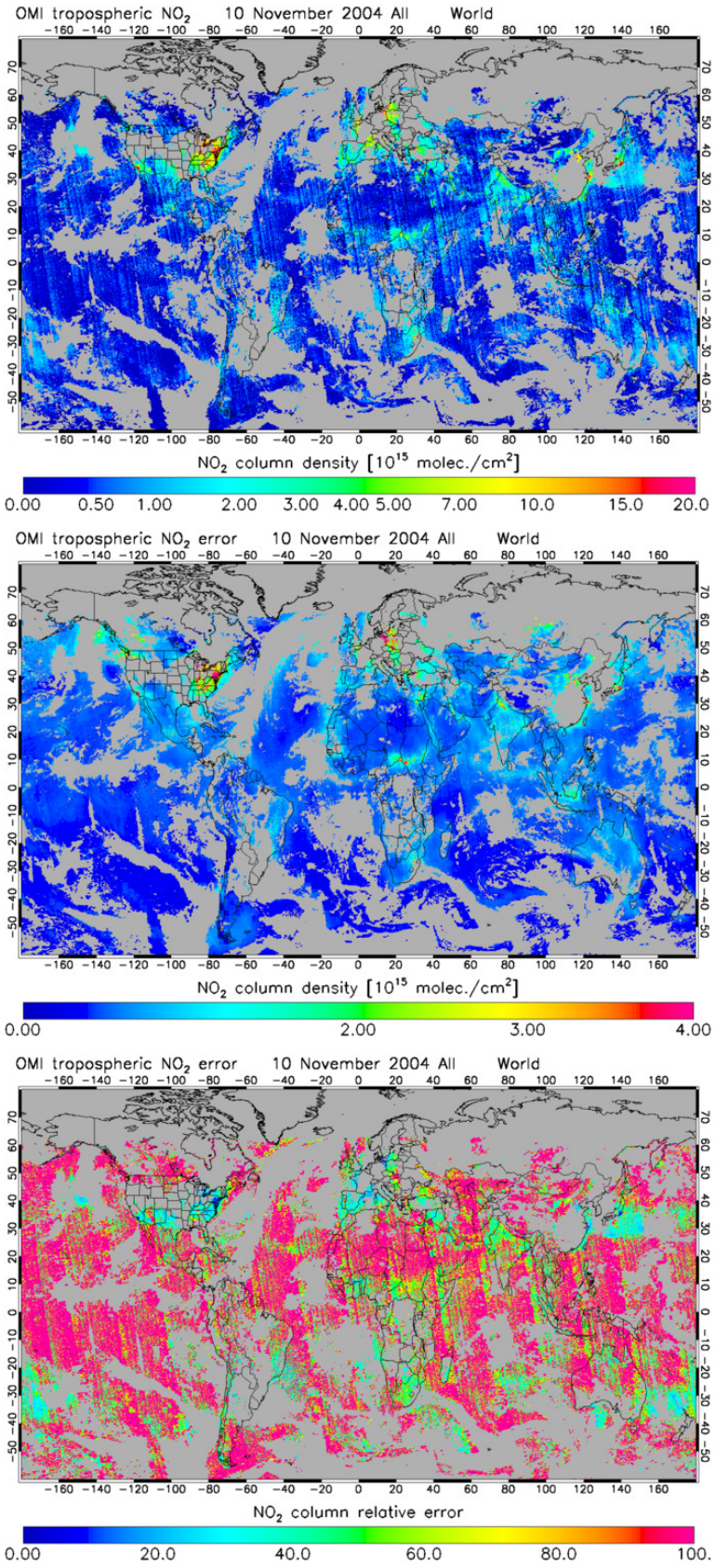

Figure 12: Tropospheric NO_2 vertical column values (top panel; in 10^{15} molec/cm^2), the corresponding absolute error estimate (middle panel; in 10^{15} molec/cm^2), and the relative error (bottom panel; in %) for 10 November 2004, based on OMI data. Large relative errors are seen mostly over areas with small NO_2 column values: oceans and remote continental regions.

averaging period, especially for less cloudy periods. In the very optimistic case that all errors in Eq. (18) are truly random in nature (which they are probably not), the accuracy required in [RD5] could be achieved by co-adding 4 TROPOMI retrievals.

9 Validation

9.1 Introduction

The most important validation need for TROPOMI NO_2 is for tropospheric NO_2 under polluted and clean conditions. Under polluted NO_2 conditions, column and profile information in the lower troposphere is essential for column density validation. The NO_2 data from OMI has been validated in several studies over the past years, based on satellite inter-comparisons (e.g. GOME-2, SCIAMACHY) as well as comparisons against ground-based (e.g. MAX-DOAS, Lidar) and in-situ (e.g. aircraft, balloon) measurements. And these validation activities have covered both tropospheric and stratospheric NO_2 data.

Despite these validation activities large uncertainties remain. These uncertainties are partly related to the quality of the independent NO_2 data used for the validation. And they are partly related to the issue of representativity of the often point-size ground-based and in-situ measurements w.r.t. the finite-sized satellite ground pixels.

KNMI is involved in a number of NO_2 validation activities, ranging from satellite intercomparisons to MAX-DOAS measurements [Piters et al., 2012] and the development of a special balloon-borne NO_2 sonde [Sluis et al., 2010]. Some of the validation activities are performed within the framework of dedicated international validation campaigns. Furthermore, KNMI has a close collaboration with other institutes involved in NO_2 validation activities.

For the validation of the TROPOMI NO_2 data we wish to extend the validation activities, concentration on a better characterisation of the quality of the independent NO_2 data and any difference between these data sets and the TROPOMI NO_2 data. These validation activities should make a distinction between stratospheric and tropospheric NO_2.

There is also a need for correlative surface albedo data to investigate the accuracy of the OMI-based surface albedo climatology. Retrieval of TROPOMI NO_2 depends on cloud information, thus cloud properties, from the FRESCO+ retrieval, must be validated carefully with correlative measurements.

Below plans for pre-launch (algorithm testing) and post-launch activities related to validation of the TROPOMI NO_2 data product are listed. Post-launch activities include comparisons of geophysical data comparisons between TROPOMI and correlative NO_2 data from a variety of sources.

9.2 Algorithm testing and verification

This important activity provides confidence in the retrieval algorithms, including forward and inverse models, based on simulations, and comparisons between different techniques and software programs. Much of TROPOMI's verification phase will address this issue in a thorough way. This activity also includes reviews and updates of the TROPOMI NO_2 ATBD.

In the (pre-launch) verification phase, the intention is to test the TROPOMI retrieval algorithm on data from existing satellite instruments, especially OMI. This has actually already been done for this ATBD but this effort will be extended. Furthermore, we anticipate that a extensive comparison of the TROPOMI NO_2 algorithm with alternative scientific retrievals from other groups (U. Bremen, MPI Mainz, DLR, BIRA-IASB) will provide useful validation and verification of the TROPOMI NO_2 algorithm proposed here. This activity would ideally be completed at launch and lead to operational data with reasonably high confidence already from the beginning of TROPOMI operations. However, experience with remote sensing has shown that redesign and fine-tuning of retrieval algorithms take place for years after launch, and this can also be expected for the TROPOMI NO_2 algorithm.

9.3 Stratospheric NO_2 validation

For stratospheric NO_2 columns, correlative (column and profile) measurements are needed in regions that are representative for a complete zonal band, and hence need to be relatively unpolluted. The currently operational NDACC network covers this need in principle, although there are concerns about the accuracy of the standard SAOZ and FTIR measurements techniques for some stations (e.g. Dirksen et al. [2011]). Nevertheless, the measurements of stratospheric NO_2 concentrations taken from high-altitude ground stations such as the Jungfraujoch station in Switzerland, are particularly valuable for validation (e.g. Hendrick et al. [2012]). Other useful sources of stratospheric NO_2 data are satellite instruments that measure in limb view, SCIAMACHY (Beirle et al. [2010], Hilboll et al. [2013b]), HIRDLS and MLS. These measurements can provide vertical profiles of NO_2 in the stratosphere, but there are difficulties in using them for direct validation as they are often only sparsely validated themselves.

In our view the main priority of the validation efforts should lie on a better characterisation of the vertical profile of stratospheric NO_2, as these profiles are an essential input to the data assimilation system in use for the separation between tropospheric and stratospheric NO_2 columns from the TROPOMI measurements. To this end the validation measurements should be well characterised, in terms of the quality of the data and of estimates of the error on the data. Improving the knowlegde of spatial and seasonal variations in the stratospheric NO_2 profiles is also important.

Stratospheric NO_2 measurements near the Arctic vortex in late winter and early spring would be useful to better test the capability of the data assimilation scheme (and other stratosphere-troposphere separation schemes) in capturing the influence of stratospheric air masses low in NO_x on stratospheric NO_2 at lower latitudes. Such excursions are known to occur and may lead to systematic errors in the separation scheme (e.g. Dirksen et al. [2011]; Bucsela et al. [2013]). Independent measurements may provide important information on how to improve these issues.

9.4 Tropospheric NO_2 validation

For validation of tropospheric NO_2 data, correlative (column and profile) measurements are needed in the highly populated polluted regions at mid-latitudes, and also in regions with natural sources of nitrogen oxides, e.g. from biomass burning, microbial soil activity and lightning. NDACC-instruments unfortunately do not meet this need, as they are often located in relatively remote and clear areas.

Information on tropospheric NO_2 concentrations – with the NO_2 in the planetary boundary layer and/or in the free troposphere – comes from in-situ instruments (at the ground, in masts, or on low-flying balloons) and from remote-sensing instruments at the ground, on balloons or aircraft. The emerging suite of MAX-DOAS instruments are particularly valuable for validation of TROPOMI measurements, but homogenisation and cross-calibration of these measurements should remain a priority for successful validation.

An important issue when comparing independent NO_2 measurements with data derived from satellite-based instruments is the question of representativity. We recommend a careful investigation into the spatial representativity of any independent NO_2 measurement, in order to facilitate a meaningful comparison with the $7\ \text{km} \times 7\ \text{km}$ TROPOMI pixel.

Important for the validation as well as for the data assimilation system in use for the separation between tropospheric and stratospheric NO_2 columns from the TROPOMI measurements is a good understanding of the vertical profile of the tropospheric NO_2. The best source of information on vertical profiles of NO_2 is still from incidental aircraft campaigns. Alternatively, experimental NO_2 profiles from (tethered) balloon sondes, and measurement towers, will provide valuable information on the vertical distribution of NO_2.

Since tropospheric retrievals depend on the concept of the air-mass factor, which has to rely on a priori information, it is important to also validate the inputs and assumptions that go into the air-mass factor calculation. This mostly concerns cloud parameters – cloud fraction and cloud pressure – that should be well characterised. Another critical issue, about which very little is known as yet, is the effect of the presence of aerosols on the NO_2 retrieval. Collocated information on the aerosol profile – e.g. coming from the TROPOMI Aerosol Layer Height data product – could be useful for this.

10 Conclusion

We have presented the baseline approach for the retrieval of the operational tropospheric and stratospheric NO_2 column products from the TROPOMI sensor. The NO_2 data will be deliverd both as an off-line product for the NO_2 data record and as a near-real time product, with the NO_2 data delivered within 3 hours after observation. The TROPOMI NO_2 data products pose an improvement over previous NO_2 data sets, particularly in their unprecedented spatial resolution (7×7 km^2), but also in the separation of the stratospheric and tropospheric contributions of the retrieved slant columns, and in the calculation of the air-mass factors used to convert slant to total columns.

The backbone of the retrieval system is the TM5 chemistry transport model, that will be operated at a global resolution of $1^\circ \times 1^\circ$. The assimilation of NO_2 slant columns in TM5 will ensure that the modelled stratospheric state becomes consistent with the TROPOMI slant columns over regions with small tropospheric NO_2 amounts. The information from the data assimilation system is used to separate the slant column into its stratospheric and tropospheric components and to provide the a priori NO_2 vertical profile required by the air-mass factor calculation.

For each TROPOMI pixel an air-mass factor is calculated, using altitude-dependent air-mass factors from a look-up table calculated with the DAK radiative transfer model, in combination with the vertical distribution of NO_2 provided by the TM5 chemistry transport model (in assimilation mode) at a spatial resolution of $1^\circ \times 1^\circ$. The AMF calculation uses local surface albedos from the OMI surface reflectance (439 nm) climatology that is based on 5 years of OMI measurements. It accounts for cloud scattering using information on effective cloud fraction and cloud pressure retrieved for every TROPOMI pixel from the FRESCO+ retrieval algorithm.

Several additional algorithm improvements are anticipated. We are currently investigating the inclusion of additional reference spectra in the DOAS spectral fit to improve the accuracy of the retrieved NO_2 slant columns. Residuals resulting from tests with the TROPOMI prototype fitting algorithm on OMI spectra suggest a need to include absorption by liquid water, in any case over cloud-free ocean scenes without substantial oceanic chlorophyll. Revisiting the OMI spectra also re-emphasised the importance of an appropriate spectral calibration that is representative for the complete fitting window. Using TM5 at a spatial resolution of $1^\circ \times 1^\circ$ (instead of a lower spatial resolution) has been shown to provide more accurate estimates of the NO_2 profiles. The conversion of the slant to vertical columns will be improved by using an air-mass factor look-up table with more nodes, in order to reduces interpolation errors.

The TROPOMI NO_2 processing chain enables us to provide a realistic error budget. The retrieval error is dominated by the spectral fitting error over oceans and regions with low tropospheric NO_2 amounts. Over the polluted regions, air-mass factor errors contribute substantially to the overall error, which can be generally approximated as 1×10^{15} molec/cm^2 $+ 25\%$ for an individual pixel. Assuming that a substantial fraction of the air-mass factor error is considered to be random (i.e. changing in time and space), the requirements for better than $\approx 10\%$ accurate retrievals can be met by averaging over 4 TROPOMI pixels.

Besides a complete error analysis, the TROPOMI data product will also provide the averaging kernel, which describes the sensitivity of TROPOMI to NO_2 in each model layer, for every pixel. The averaging kernel is especially relevant for data users who wish to minimise the discrepancies between the assumptions in the TROPOMI retrieval and their own application of interest, e.g. for data assimilation, validation, or comparison studies.

TROPOMI's high spatial resolution, with 7×7 km^2 ground pixels at nadir, will enable monitoring NO_2 concentrations with an unprecedented accuracy, both in the troposphere and the stratosphere. From these measurements we will learn more about the distribution of NO_2, its sources and sinks, its transport through the atmosphere, its role in stratospheric and tropospheric chemistry, as well as in climate issues, notably through the important role that nitrogen oxides play in the formation of secondary pollutants ozone and aerosol. The early-afternoon NO_2 data record, which started with OMI, will be extended by TROPOMI, alongside the mid-morning measurements of the GOME-2 instruments, thus providing essential information on the diurnal cycle of NO_2. Over the past 17 years various UV/Vis backscatter instruments have been used to monitor NO_2 on a global scale. The operational TROPOMI NO_2 data processing is consistent with the NO_2 retrieval record generated at KNMI, and will continue and improve that record.

11 References

[Adams et al., 2013] Adams, C., Strong, K., Zhao, X., Bourassa, A. E., Daffer, W. H., Degenstein, D., Drummond, J. R., Farahani, E. E., Fraser, A., Lloyd, N. D., Manney, G. L., McLinden, C. A., Rex, M., Roth, C., Strahan, S. E., Walker, K. A., and Wohltmann, I. (2013). The spring 2011 final stratospheric warming above Eureka: anomalous dynamics and chemistry. *Atmos. Chem. Phys.*, 13:611–624.

[Beirle et al., 2011] Beirle, S., Boersma, K. F., Platt, U., Lawrence, M. G., and T., W. (2011). Megacity emissions and lifetimes of nitrogen oxides probed from space. *Science*, 333:1737–1739.

[Beirle et al., 2010] Beirle, S., Kühl, S., Puķīte, J., and T., W. (2010). Retrieval of tropospheric column densities of NO_2 from combined SCIAMACHY nadir/limb measurements. *Atmos. Meas. Tech.*, 3:283–299.

[Belmonte-Rivas et al., 2014] Belmonte-Rivas, M., Veefkind, P., Boersma, F., Levelt, P., Eskes, H., and Gille, J. (2014). Intercomparison of daytime stratospheric NO_2 satellite retrievals and model simulations. *Atmos. Meas. Tech.*, 7:no2–atbd–ref.bib.

[Bernath et al., 2005] Bernath, P. F., McElroy, C. T., Abrams, M. C., Boone, C. D., and M. Butler, e. a. (2005). Atmospheric Chemistry Experiment (ACE): Mission overview. *Geophys. Res. Lett.*, 32(L15S01):5 pp.

[Boersma et al., 2009] Boersma, K., Jacob, D., Trainic, M., Rudich, Y., DeSmedt, I., Dirksen, R., and Eskes, H. (2009). Validation of urban NO_2 concentrations and their diurnal and seasonal variations observed from the SCIAMACHY and OMI sensors using in situ surface measurements in israeli cities. *Atmos. Chem. Phys.*, 9:3867–3879.

[Boersma et al., 2002] Boersma, K. F., Bucsela, E., Brinksma, E., and Gleason, J. F. (2002). NO_2. In *OMI Algorithm Theoretical Basis Document – Vol. 4: OMI Trace Gas Algorithms, ATBD-OMI-02 Vers. 2.0*, pages 13–28. NASA Goddard Space Flight Cent., Greenbelt, Md.

[Boersma et al., 2004] Boersma, K. F., Eskes, H. J., and Brinksma, E. J. (2004). Error analysis for tropospheric NO_2 retrieval from space. *J. Geophys. Res.*, 109(D04311):20 pp.

[Boersma et al., 2011] Boersma, K. F., Eskes, H. J., Dirksen, R. J., Van der A, R. J., Veefkind, J. P., Stammes, P., Huijnen, V., Kleipool, Q. L., Sneep, M., Claas, J., Leitão, J., Richter, A., Zhou, Y., and Brunner, D. (2011). An improved retrieval of tropospheric NO_2 columns from the Ozone Monitoring Instrument. *Atmos. Meas. Tech.*, 4:1905–1928.

[Boersma et al., 2007] Boersma, K. F., Eskes, H. J., Veefkind, J. P., Brinksma, E. J., Van der A, R. J., Sneep, M., Van den Oord, G. H. J., Levelt, P. F., Stammes, P., F., G. J., and Bucsela, E. J. (2007). Near-real time retrieval of tropospheric NO_2 from OMI. *Atmos. Chem. Phys.*, 7:2013–2128.

[Bovensmann et al., 1999] Bovensmann, H., Burrows, J. P., Buchwitz, M., Frerick, J., Noel, S., Rozanov, V. V., Chance, K. V., and Goede, A. P. H. (1999). SCIAMACHY: mission objectives and measurement modes. *J. Atmos. Sci.*, 56:127–150.

[Brion et al., 1998] Brion, J., Chakir, A., Charbonnier, J., Daumont, D., Parisse, C., and Malicet, J. (1998). Absorption spectra measurements for the ozone molecule in the 350-830 nm region. *J. Atmos. Chem.*, 30:291–299.

[Brion et al., 1993] Brion, J., Chakir, A., Daumont, D., Malicet, J., and Parisse, C. (1993). High-resolution laboratory absorption cross section of O_3. Temperature effect. *Chem. Phys. Lett.*, 213:610–612.

[Bucsela et al., 2006] Bucsela, E. J., Celarier, E. A., Wenig, M. O., Gleason, J. F., Veefkind, J. P., Boersma, K. F., and Brinksma, E. J. (2006). Algorithm for NO_2 vertical column retrieval from the ozone monitoring instrument. *IEEE Trans. Geosci. Rem. Sens.*, 44:1245–1258.

[Bucsela et al., 2013] Bucsela, E. J., Krotkov, N. A., Celarier, E. A., Lamsal, L. N., Swartz, W. H., Bhartia, P. K., Boersma, K. F., Veefkind, J. P., Gleason, J. F., and Pickering, K. E. (2013). A new stratospheric and tropospheric NO_2 retrieval algorithm for nadir-viewing satellite instruments: applications to OMI. *Atmos. Meas. Tech.*, 6:2607–2626.

[Burrows et al., 1999] Burrows, J. P., Weber, M., Buchwitz, M., Rozanov, V., Ladstätter-Weißenmayer, A., Richter, A., Debeek, R., Hoogen, R., Bramstedt, K., Eichmann, K.-U., Eisinger, M., and Perner, D. (1999). The Global Ozone Monitoring Experiment (GOME): Mission concept and first results. *J. Atmos. Sci.*, 56:151–175.

[Castellanos and Boersma, 2012] Castellanos, P. and Boersma, K. F. (2012). Reductions in nitrogen oxides over Europe driven by environmental policy and economic recession. *Scientific Reports*, 2:7 pp.

[Chance and Spurr, 1997] Chance, K. V. and Spurr, R. J. D. (1997). Ring effect studies: Rayleigh scattering, including molecular parameters for rotational Raman scattering and the Fraunhofer spectrum. *Appl. Opt.*, 36(21):5224–5230.

[Chu and McCormick, 1986] Chu, W. P. and McCormick, M. P. (1986). Sage observations of strastospheric nitrogen dioxide. *J. Geophys. Res.*, 91(D5):5465–5476.

[Crutzen, 1970] Crutzen, P. J. (1970). The influence of nitrogen oxides on the atmospheric ozone content. *Quart. J. R. Meteorol. Soc.*, 96:320–325.

[De Haan et al., 1987] De Haan, J. F., Bosma, P. B., and Hovenier, J. W. (1987). The adding method for multiple scattering in a non-homogeneous Rayleigh atmosphere. *Astron. & Astroph.*, 183:371–391.

[De Ruyter de Wildt et al., 2012] De Ruyter de Wildt, M., Eskes, H., and Boersma, K. F. (2012). The global economic cycle and satellite-derived NO_2 trends over shipping lanes. *Geophys. Res. Lett.*, 39(L01802):6 pp.

[Dentener et al., 2006] Dentener, F., Drevet, J., Lamarque, J. F., Bey, I., Eickhout, B., Fiore, A. M., Hauglustaine, D., Horowitz, L. W., Krol, M., Kulshrestha, U. C., Lawrence, M., Galy-Lacaux, C., Rast, S., Shindell, D., Stevenson, D., Noije, T. V., Atherton, C., Bell, N., Bergman, D., Butler, T., Cofala, J., Collins, B., Doherty, R., Ellingsen, K., Galloway, J., Gauss, M., Montanaro, V., Müller, J. F., Pitari, G., Rodriguez, J., Sanderson, M., Solmon, F., Strahan, S., Schultz, M., Sudo, K., Szopa, S., and Wild, O. (2006). Nitrogen and sulfur deposition on regional and global scales: A multimodel evaluation. *Global Biogeochem. Cycles*, 20(GB4003):21 pp.

[Dentener et al., 2003] Dentener, F., Van Weele, M., Krol, M., Houweling, S., and van Velthoven, P. (2003). Trends and inter-annual variability of methane emissions derived from 1979-1993 global CTM simulations. *Atmos. Chem. Phys.*, 3:73–88.

[Dirksen et al., 2011] Dirksen, R. J., Boersma, K. F., Eskes, H. J., Ionov, D. V., Bucsela, E. J., Levelt, P. F., and Kelder, H. M. (2011). Evaluation of stratospheric NO_2 retrieved from the Ozone Monitoring Instrument: Intercomparison, diurnal cycle, and trending. *J. Geophys. Res.*, 116(D08305):22 pp.

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

[Eskes et al., 2003] Eskes, H. J., Van Velthoven, P. F. J., Valks, P., and Kelder, H. M. (2003). Assimilation of GOME total ozone satellite observations in a three-dimensional tracer transport model. *Q. J. R. Meteorol. Soc.*, 129(590):1663–1681.

[Fuglestvedt et al., 1999] Fuglestvedt, J. S., Berntsen, T., Isaksen, I. S. A., Mao, H., Liang, X.-Z., and Wang, W.-C. (1999). Climatic forcing of nitrogen oxides through changes in tropospheric ozone and methane. *Atmos. Environ.*, 33(3):961–977.

[Gordley et al., 1996] Gordley, L. L., III, J. M. R., Mickley, L. J., Frederick, J. E., Park, J. H., Stone, K. A., Beaver, G. M., McInerney, J. M., Deaver, L. E., Toon, G. C., Murcray, F. J., Blatherwick, R. D., Gunson, M. R., Abbatt, J. P. D., III, R. L. M., Mount, G. H., Sen, B., and Blavier, J.-F. (1996). Validation of nitric oxide and nitrogen dioxide measurements made by the Halogen Occultation Experiment for UARS platform. *J. Geophys. Res.*, 101(D6):10241–10266.

[Gorshelev et al., 2014] Gorshelev, V., Serdyuchenko, A., Weber, M., and Burrows, J. P. (2014). High spectral resolution ozone absorption cross-sections: Part I. Measurements, data analysis and comparison around 293K. *Atmos. Meas. Tech.*, 7:609–624.

[Gruzdev and Elokhov, 2009] Gruzdev, A. N. and Elokhov, A. S. (2009). Validating NO_2 measurements in the vertical atmospheric column with the OMI instrument aboard the EOS Aura satellite against ground-based measurements at the Zvenigorod Scientific Station. *Izv. Atmos. Oceanic Phys.*, 45(4):444–455.

[Hains et al., 2010] Hains, J., Boersma, K., Kroon, M., Dirksen, R., Cohen, R., Perring, A., Bucsela, E., Volten, H., Swart, D., Richter, A. andWittrock, F., Schoenhardt, A., Wagner, T., Ibrahim, O., Roozendael, V., M., Pinardi, G., Gleason, J., Veefkind, J., and Levelt, P. (2010). Testing and improving OMI DOMINO tropospheric NO_2 using observations from the DANDELIONS and INTEX-B validation campaigns. *J. Geophys. Res.*, 115(D05301):20 pp.

[Hendrick et al., 2012] Hendrick, F., Mahieu, E., Bodeker, G., Boersma, K., Chipperfield, M., De Mazière, M., De Smedt, I., Demoulin, P., Fayt, C., Hermans, C., Kreher, K., Lejeune, B., Pinardi, G., Servais, C., Stübi, R., Van der A, R., Vernier, J.-P., and Van Roozendael, M. (2012). Analysis of stratospheric NO_2 trends above Jungfraujoch using ground-based UV-visible, FTIR, and satellite nadir observations. *Atmos. Chem. Phys.*, 12:8851–8864.

[Hilboll et al., 2013a] Hilboll, A., Richter, A., and Burrows, J. (2013a). Long-term changes of tropospheric NO_2 over megacities derived from multiple satellite instruments. *Atmos. Chem. Phys.*, 13:4145–4169.

[Hilboll et al., 2013b] Hilboll, A., Richter, A., Rozanov, A., Hodnebrog, O., Heckel, A., Solberg, S., Stordal, F., and Burrows, J. P. (2013b). Improvements to the retrieval of tropospheric NO_2 from satellite – stratospheric correction using SCIAMACHY limb/nadir matching and comparison to Oslo CTM2 simulations. *Atmos. Meas. Tech.*, 6:565–584.

[Houweling et al., 1998] Houweling, S., Dentener, F., and Lelieveld, J. (1998). The impact of nonmethane hydrocarbon compounds on tropospheric photochemistry. *J. Geophys. Res.*, 103(D9):10,673–10,696.

[Huijnen et al., 2010a] Huijnen, V., Eskes, H. J., Poupkou, A., Elbern, H., Boersma, K., Foret, G., Sofiev, M., Valdebenito, A., Flemming, J., Stein, O., Gross, A., Robertson, L., D'Isidoro, M., Kioutsioukis, I., Friese, E., Amstrup, B., Bergstrom, R., Strunk, A., Vira, J., Zyryanov, D., Maurizi, A., Melas, D., Peuch, V.-H., and Zerefos, C. (2010a). Comparison of OMI NO_2 tropospheric columns with an ensemble of global and European regional air quality models. *Atmos. Chem. Phys.*, 10:3273–3296.

[Huijnen et al., 2010b] Huijnen, V., Williams, J., Van Weele, M., Van Noije, T., Krol, M., Dentener, F., Segers, A., Houweling, S., Peters, W., De Laat, J., Boersma, F., Bergamaschi, P., Van Velthoven, P., Le Sager, P., Eskes, H., Alkemade, F., Scheele, R., Nédélec, P., and Pätz, H.-W. (2010b). The global chemistry transport model tm5: description and evaluation of the tropospheric chemistry version 3.0. *Geosc. Model Dev.*, 3(2):445–473.

[Irie et al., 2012] Irie, H., Boersma, K. F., Kanaya, Y., Takashima, H., Pan, X., and Wang, Z. F. (2012). First quantitative bias estimates for tropospheric NO_2 columns retrieved from SCIAMACHY, OMI, and GOME-2 using a common standard. *Atmos. Meas. Tech.*, 5:2403–2411.

[Jacob, 1999] Jacob, D. J. (1999). *Introduction to Atmospheric Chemistry*. Princeton University Press.

[Kleipool et al., 2008] Kleipool, Q. L., Dobber, M. R., De Haan, J. F., and Levelt, P. F. (2008). Earth surface reflectance climatology from 3 years of OMI data. *J. Geophys. Res.*, 113(D18308):22 pp.

[Koelemeijer et al., 2003] Koelemeijer, R. B. A., De Haan, J. F., and Stammes, P. (2003). A database of spectral surface reflectivity in the range 335-772 nm derived from 5.5 years of GOME observations. *J. Geophys. Res.*, 108(D2):13 pp.

[Koelemeijer et al., 2001] Koelemeijer, R. B. A., Stammes, P., W., H. J., and De Haan, J. F. (2001). A fast method for retrieval of cloud parameters using oxygen A band measurements from the Global Ozone Monitoring Experiment. *J. Geophys. Res.*, 106(D4):3475–3490.

[Lamsal et al., 2010] Lamsal, L., Martin, R., Van Donkelaar, A., Celarier, E., Bucsela, E., Boersma, K., Dirksen, R., Luo, C., and Wang, Y. (2010). 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(D05302):15 pp.

[Leitão et al., 2010] Leitão, J., Richter, A., Vrekoussis, M., Kokhanovsky, A., Zhang, Q., Beekmann, M., and Burrows, J. P. (2010). On the improvement of NO_2 satellite retrievals – aerosol impact on the airmass factors. *Atmos. Meas. Tech.*, 3:475–493.

[Lerot et al., 2010] Lerot, C., Stavrakou, T., De Smedt, I., Müller, J.-F., and Van Roozendael, M. (2010). Glyoxal vertical columns from GOME-2 backscattered light measurements and comparisons with a global model. *Atmos. Chem. Phys.*, 10:12059–12072.

[Leue et al., 2001] Leue, C., Wenig, M., Wagner, T., Klimm, O., Platt, U., and Jähne, B. (2001). Quantitative analysis of NO_x emissions from Global Ozone Monitoring Experiment satellite image sequences. *J. Geophys. Res.*, 106(D6):5493–5505.

[Levelt et al., 2006] Levelt, P. F., van den Oord, G. H. J., Dobber, M. R., Mälkki, A., Visser, H., de Vries, J., Stammes, P., Lundell, J. O. V., and Saari, H. (2006). The Ozone Monitoring Instrument. *IEEE Trans. Geosci. Rem. Sens.*, 44:1093–1101.

[Liley et al., 2000] Liley, J. B., Johnston, P. V., McKenzie, R. L., Thomas, A. J., and Boyd, I. S. (2000). Stratospheric NO_2 variations from a long time series at Lauder, New Zealand. *J. Geophys. Res.*, 105(D9):11,633–11,640.

[Lin et al., 2014] Lin, J., Martin, R. V., Boersma, K. F., Sneep, M., Stammes, P., Spurr, R., Wang, P., Roozendael, M. V., Clémer, K., and Irie, H. (2014). Retrieving tropospheric nitrogen dioxide over China from the Ozone Monitoring Instrument: Effects of aerosols, surface reflectance anisotropy and vertical profile of nitrogen dioxide. *Atmos. Chem. Phys.*, 14:1441–1461.

[Lin et al., 2010] Lin, J.-T., McElroy, M. B., and Boersma, K. F. (2010). Constraint of anthropogenic NO_x emissions in china from different sectors: a new methodology using multiple satellite retrievals. *Atmos. Chem. Phys.*, 10:63–78.

[Llewellyn et al., 2004] Llewellyn, E. J., Lloyd, N. D., Degenstein, D. A., Gattinger, R. L., and S. V. Petelina, e. a. (2004). The OSIRIS instrument on the Odin spacecraft. *Canadian Journal of Physics*, 82(6):411–422.

[Ma et al., 2013] Ma, J. Z., Beirle, S., Jin, J. L., Shaiganfar, R., Yan, P., and Wagner, T. (2013). Tropospheric NO_2 vertical column densities over Beijing: results of the first three years of ground-based MAX-DOAS measurements (2008-2011) and satellite validation. *Atmos. Chem. Phys.*, 13:1547–1567.

[Maasakkers et al., 2013] Maasakkers, J. D., Boersma, K. F., Williams, J. E., Van Geffen, J., Vinken, G. C. M., Sneep, M., Hendrick, F., Van Roozendael, M., and Veefkind, J. P. (2013). Vital improvements to the retrieval of tropospheric NO_2 columns from the Ozone Monitoring Instrument. *Geophys. Res. Abstracts*, 15(EGU2013-714):1. EGU General Assembly 2013.

[Martin et al., 2002] Martin, R. V., Chance, K., Jacob, D. J., Kurosu, T. P., Spurr, R. J. D., Bucsela, E., Gleason, J. F., Palmer, P. I., Bey, I., Fiore, A. M., Li, Q., Yantosca, R. M., and Koelemeijer, R. B. A. (2002). An improved retrieval of tropospheric nitrogen dioxide from gome. *J. Geophys. Res.*, 107(D20):4437–4457.

[Merlaud et al., 2012] Merlaud, A., Van Roozendael, M., Van Gent, J., Fayt, C., Maes, J., Toledo-Fuentes, X., Ronveaux, O., and De Maziére, M. (2012). DOAS measurements of NO_2 from an ultralight aircraft during the Earth Challenge expedition. *Atmos. Meas. Tech.*, 5:2057–2068.

[Mijling and Van der A, 2012] Mijling, B. and Van der A, R. J. (2012). Using daily satellite observations to estimate emissions of short-lived air pollutants on a mesoscopic scale. *J. Geophys. Res.*, 117(D17302):20 pp.

[Mount et al., 1984] Mount, G. H., Rusch, D. W., Noxon, J. F., Zawodny, J. M., and Barth, C. A. (1984). Measurements of stratospheric NO_2 from the Solar Mesosphere Explorer Satellite: 1. An overview of the results. *J. Geophys. Res.*, 89(D1):1327–1340.

[Munro et al., 2006] Munro, R., Eisinger, M., Anderson, C., Callies, J., Corpaccioli, E., Lang, R., Lefebvre, A., Livschitz, Y., and Albinana, A. P. (2006). GOME-2 on MetOp. In *Proceedings of the Atmospheric Science Conference 2006*, SP 628. ESA, ESA, Paris.

[Murphy et al., 1993] Murphy, D. M., Fahey, D. W., Proffitt, M. H., Liu, S. C., Chan, K. R., Eubank, C. S., Kawa, S. R., and Kelly, K. K. (1993). Reactive nitrogen and its correlation with ozone in the lower stratosphere and upper troposphere. *J. Geophys. Res.*, 98(D5):8751–8773.

[Ott et al., 2010] Ott, L. E., Pickering, K. E., Stenchikov, G. L., Allen, D. J., DeCaria, A. J., Ridley, B., Lin, R.-F., Lang, S., and Tao, W.-K. (2010). Production of lightning NO and its vertical distribution calculated from three-dimensional cloud-scale chemical transport model simulations. *J. Geophys. Res.*, 115(D04301):19 pp.

[Palmer et al., 2001] Palmer, P. I., Jacob, D., Chance, K., Martin, R. V., Spurr, R. J. D., Kurosu, T. P., Bey, I., Yantosca, R., Fiore, A., and Li, Q. (2001). Air-mass factor formulation for spectroscopic measurements frorm satellite: applications to fomaldehyde retrievals from the Global Ozone Monitoring Experiment. *J. Geophys. Res.*, 106:14539–14550.

[Piters et al., 2012] Piters, A., Boersma, K., Kroon, M., Hains, J. C., and Van Roozendael M., e. a. (2012). The Cabauw Intercomparison campaign for Nitrogen Dioxide measuring Instruments (CINDI): design, execution, and early results. *Atmos. Meas. Tech.*, 5:457–485.

[Platt, 1994] Platt, U. (1994). Differential Optical Absorption Spectroscopy (DOAS). *Air monitoring by spectroscopic techniques, Chem. Anal.*, 127:27–76. edited by M.W. Sigrist.

[Platt and Stutz, 2008] Platt, U. and Stutz, Z. (2008). *Differential Optical Absorption Spectroscopy, Principles and Applications*. Springer, Heidelberg, Germany.

[Pope and Fry, 1997] Pope, R. M. and Fry, E. S. (1997). Absorption spectrum (380-700 nm) of pure water. II. Integrating cavity measurements. *Appl. Opt.*, 36(33):8710–8723.

[Popp et al., 2011] Popp, C., Wang, P., Brunner, D., Stammes, P., Zhou, Y., and Grzegorski, M. (2011). MERIS albedo climatology for FRESCO+ O_2 A-band cloud retrieval. *Atmos. Meas. Tech.*, 4:463–483.

[Randall et al., 1998] Randall, C. E., Rusch, D. W., Bevilacqua, R. M., Hoppel, K. W., and Lumpe, J. D. (1998). Polar Ozone and Aerosol Measurement (POAM) II stratospheric NO_2 1993-1996. *J. Geophys. Res.*, 103(D21):28361–28371.

[Ravishankara et al., 2009] Ravishankara, A. R., Daniel, J. S., and Portmann, R. W. (2009). Nitrous Oxide (N2O): The Dominant Ozone-Depleting Substance Emitted in the 21st Century. *Science*, 326(5949):123–125.

[Richter et al., 2011] Richter, A., Begoin, M., Hilboll, A., and Burrows, J. P. (2011). An improved NO_2 retrieval for the GOME-2 satellite instrument. *Atmos. Meas. Tech.*, 4:1147–1159.

[Richter and Burrows, 2002] Richter, A. and Burrows, J. P. (2002). Tropospheric NO_2 from GOME measurements. *Adv. Space Res.*, 29(11):1673–1683.

[Schaub et al., 2007] Schaub, D., Brunner, D., Boersma, K. F., Keller, J., Folini, D., Buchmann, B., Berresheim, H., and Staehelin, J. (2007). SCIAMACHY tropospheric NO_2 over Switzerland: estimates of NO_x lifetimes and impact of the complex Alpine topography on the retrieval. *Atmos. Chem. Phys.*, 7:5971–5987.

[Seinfeld and Pandis, 2006] Seinfeld, J. H. and Pandis, S. N. (2006). *Atmospheric Chemistry and Physics - From Air Pollution to Climate Change (2nd Edition)*. John Wiley & Sons.

[Serdyuchenko et al., 2014] Serdyuchenko, A., Gorshelev, V., Weber, M., Chehade, W., and Burrows, J. P. (2014). High spectral resolution ozone absorption cross-sections: Part II. Temperature dependence. *Atmos. Meas. Tech.*, 7:625–636.

[Shindell et al., 2009] Shindell, D. T., Faluvegi, G., Koch, D. M., Schmidt, G. A., Unger, N., and Bauer, S. E. (2009). Improved attribution of climate forcing to emissions. *Science*, 326(5953):716–718.

[Sierk et al., 2006] Sierk, B., Richter, A., Rozanov, A., Von Savigny, C., Schmoltner, A. M., Buchwitz, M., Bovensmann, H., and Burrows, J. P. (2006). Retrieval and monitoring of atmospheric trace gas concentrations in nadir and limb geometry using the space-borne SCIAMACHY instrument. *Environmental Monitoring and Assessment*, 120:65–73.

[Sillman et al., 1990] Sillman, S., Logan, J. A., and Wofsy, S. C. (1990). The sensitivity of ozone to nitrogen oxides and hydrocarbons in regional ozone episodes. *J. Geophys. Res.*, 95(D2):1837–1851.

[Silver et al., 2013] Silver, J., Brandt, J., Hvidberg, M., Frydendall, J., and Christensen, J. (2013). Assimilation of OMI NO_2 retrievals into the limited-area chemistry-transport model DEHM (V2009.0) with a 3-D OI algorithm. *Geosc. Model Dev.*, 6:1–16.

[Sluis et al., 2010] Sluis, W. W., Allaart, M. A. F., M., P. A. J., and Gast, L. F. L. (2010). The development of a nitrogen dioxide sonde. *Atmos. Meas. Tech.*, 3:1753–1762.

[Solomon, 1999] Solomon, S. (1999). Stratospheric ozone depletion: A review of concepts and history. *Rev. Geophys.*, 37(3):275–316.

[Spurr, 2002] Spurr, R. J. D. (2002). Simultaneous derivation of intensities and weighting functions in a general pseudo-spherical radiative transfer. *J. Quant. Spectrosc. & Radiat. Transfer*, 75:129–175.

[Stammes, 2001] Stammes, P. (2001). Spectral radiance modeling in the UV-visible range. In Smith, W. and Timofeyev, Y., editors, *IRS 2000: Current Problems in Atmospheric Radiation*, pages 385–388. A. Deepak, Hampton, Va.

[Stammes et al., 2008] Stammes, P., Sneep, M., De Haan, J. F., Veefkind, J. P., Wang, P., and Levelt, P. F. (2008). Effective cloud fractions from the Ozone Monitoring Instrument: Theoretical framework and validation. *J. Geophys. Res.*, 113(D16S38):12 pp.

[Stavrakou et al., 2008] Stavrakou, T., Müller, J.-F., Boersma, K. F., De Smedt, I., and Van der A, R. J. (2008). Assessing the distribution and growth rates of NO_2 emission sources by inverting a 10-year record of NO_2 satellite columns. *Geophys. Res. Lett.*, 35(L10801):5 pp.

[Thalman and Volkamer, 2013] Thalman, R. and Volkamer, R. (2013). Temperature dependant absorption cross-sections of O2-O2 collision pairs between 340 and 630 nm at atmospherically relevant pressure. *Phys. Chem. Chem. Phys.*, 15:15371–15381.

[Valks et al., 2011] Valks, P., Pinardi, G., Richter, A., Lambert, J.-C., Hao, N., Loyola, D., Van Roozendael, M., and Emmadi, S. (2011). Operational total and tropospheric NO_2 column retrieval for GOME-2. *Atmos. Meas. Tech.*, 4:1491–1514.

[Van der A et al., 2010] Van der A, R., Allaart, M., and Eskes, H. (2010). Multi sensor reanalysis of total ozone. *Atmos. Chem. Phys.*, 10:11277–11294.

[Van der A et al., 2008] Van der A, R. J., Eskes, H. J., Boersma, K. F., Van Noije, T. P. C., Van Roozendael, M., De Smedt, I., Peters, D. H. M. U., Kuenen, J. J. P., and Meijer, E. W. (2008). rends, seasonal variability and dominant NO_x source derived from a ten year record of NO_2 measurd from space. *J. Geophys. Res.*, 113(D04302):12 pp.

[Van Diedenhoven et al., 2007] Van Diedenhoven, B., Hasekamp, O. P., and Landgraf, J. (2007). Retrieval of cloud parameters from satellite-based reflectance measurements in the ultraviolet and the oxygen A-band. *J. Geophys. Res.*, 112(D15208):15 pp.

[Van Geffen et al., 2015] Van Geffen, J. H. G. M., Boersma, K. F., Van Roozendael, M., Hendrick, F., Mahieu, E., De Smedt, I., M., S., and Veefkind, J. P. (2015). Improved spectral fitting of nitrogen dioxide from OMI in the 405 – 465 nm window. *Atmos. Meas. Tech.*, 8:1685–1699.

[Van Roozendael et al., 2006] Van Roozendael, M., Loyola, D., Spurr, R., Balis, D., Lambert, J.-C., Livschitz, Y., Valks, P., Ruppert, T., Kenter, P., Fayt, C., and Zehner, C. (2006). Ten years of GOME/ERS-2 total ozone data – The new GOME data processor (GDP) version 4: 1. Algorithm description. *J. Geophys. Res.*, 111(D14311):21 pp.

[Vandaele et al., 1998] Vandaele, A. C., Hermans, C., Simon, P. C., Carleer, M., Colin, R., Fally, S., Mérienne, M. F., Jenouvrier, A., and Coquart, B. (1998). Measurements of the NO_2 absorption cross-section from 42000 cm^{-1} to 10000 cm^{-1} (238-1000 nm) at 220 K and 294 K. *J. Quant. Spectrosc. & Radiat. Transfer*, 59:171–184.

[Vandevender and Haskell, 1982] Vandevender, W. H. and Haskell, K. H. (1982). The SLATEC mathematical subroutine library. *ACM SIGNUM Newsletter*, 17.3:16–21.

[Vasilkov et al., 2002] Vasilkov, A. P., Joiner, J., Gleason, J., and Bhartia, P. K. (2002). Ocean Raman scattering in satellite backscatter UV measurements. *Geophys. Res. L*, 29(17):1837–1840.

[Veefkind et al., 2012] Veefkind, J. P., Aben, I., McMullan, K., Förster, H., De Vries, J., Otter, G., Claas, J., Eskes, H. J., De Haan, J. F., Kleipool, Q., Van Weele, M., Hasekamp, O., Hoogeveen, R., Landgraf, J., Snel, R., Tol, P., Ingmann, P., Voors, R., Kruizinga, B., Vink, R., Visser, H., and Levelt, P. F. (2012). TROPOMI on the ESA Sentinel-5 Precursor: A GMES mission for global observations of the atmospheric composition for climate, air quality and ozone layer applications. *Rem. Sens. Environment*, 120:70–83.

[Vermote and Tanré, 1992] Vermote, E. and Tanré, D. (1992). Analytic expressions for radiative properties of planar Rayleigh scattering media, including polarization contributions. *J. Quant. Spectrosc. & Radiat. Transfer*, 47:305–314.

[Vountas et al., 2003] Vountas, M., Richter, A., Wittrock, F., and Burrows, J. P. (2003). Inelastic scattering in ocean water and its impact on trace gas retrievals from satellite data. *Atmos. Chem. Phys.*, 3:1365–1375.

[Wang et al., 2008] Wang, P., Stammes, P., Van der A, R., Pinardi, G., and Van Roozendael, M. (2008). FRESCO+: an improved O_2 A-band cloud retrieval algorithm for tropospheric trace gas retrievals. *Atmos. Chem. Phys.*, 8:6565–6576.

[Wang et al., 2012] Wang, S., Zhang, Q., Streets, D., He, K., Martin, R., Lamsal, L., Chen, D., Lei, Y., and Lu, Z. (2012). Growth in NO_x emissions from power plants in China: bottom-up estimates and satellite observations. *Atmos. Chem. Phys.*, 12:4429–4447.

[Wenig et al., 2003] Wenig, M., Spichtinger, N., Stohl, A., Held, G., Beirle, S., Wagner, R., Jähne, T., and Platt, U. (2003). Intercontinental transport of nitrogen oxide pollution plumes. *Atmos. Chem. Phys.*, 3:387–393.

[World Health Organisation, 2003] World Health Organisation (2003). *Health Aspects of Air Pollution with Particulate Matter, Ozone and Nitrogen Dioxide*. World Health Organisation, Bonn.

[Zhou et al., 2009] Zhou, Y., Brunner, D., Boersma, K., Dirksen, R., and Wang, P. (2009). An improved tropospheric NO_2 retrieval for omi observations in the vicinity of mountainous terrain. *Atmos. Meas. Tech.*, 2:401–416.

[Zhou et al., 2010] Zhou, Y., Brunner, D., Spurr, R. J. D., Boersma, K., Sneep, M., Popp, C., and Buchmann, B. (2010). Accounting for surface reflectance anisotropy in satellite retrievals of tropospheric NO_2. *Atmos. Meas. Tech.*, 3:1185–1203.