Product User Manual

Tropospheric NO_2 columns and surface NO_2 concentrations from Sentinel-5P/TROPOMI

Signatures

Action: Name	***Affiliation***	***Function***	***Date***	***Sign.***
prepared by: P. Valks K.L. Chan S. Liu	 DLR-MF DLR-MF DLR-MF	 S-VELD Project Manager S-VELD Project Scientist S-VELD Project Scientist	01.09.2019	

Document Change Log

Issue	***Rev.***	***Date***	***Section***	***Description of Change***
1/A	1	01.09.2019	All	Completely new

TABLE OF CONTENTS

SENTINEL-BASIERTE ATMOSPHÄRENPRODUKTE ZUR BEWERTUNG DES EINFLUSSES VON VERKEHRSEMISSIONEN AUF DIE LUFTQUALITÄT IN DEUTSCHLAND (S-VELD)

Problemstellung

Luftqualität spielt eine wichtige Rolle in der aktuellen Umweltdiskussion. Millionen Menschen leben heute in Regionen mit stark verschmutzter Luft. Hohe Schadstoffbelastungen sind für Atemwegs- und Herz-Kreislauf-Erkrankungen verantwortlich und führen zu erhöhten Sterblichkeitsraten. Zum Schutz der menschlichen Gesundheit gelten deshalb europaweit unter anderem Grenzwerte für Stickstoffdioxid(NO_2) und Feinstaub. Überschreitungen der Grenzwerte beider Komponenten finden sich derzeit vor allem an durch Verkehr hoch belasteten Standorten.

Projektziel

Dieses Projekt hat zum Ziel, auf Grundlage neuester Sentinel-Satellitenmessungen von NO_2 und Feinstaub sowie Verkehrsdaten des BMVI, die verkehrsbedingten Schadstoffemissionen und deren Anteil an derUmweltbelastung in Deutschland besser zu quantifizieren.
Mit der Bereitstellung hochaufgelöster Karten zur Schadstoffbelastung und verbesserten Emissions-Daten soll das Potential der Copernicus Sentinel-Flotte für das Umwelt-Monitoring als Entscheidungshilfe fürBehörden demonstriert und als wichtige zukünftige Datenquelle etabliert werden.

Durchführung

Aus Messungen der Sentinel-Satelliten werden bodennahe NO_2- und Feinstaubkonzentrationen für dasBundesgebiet und die angrenzenden Länder bestimmt und über einen Web Mapping-Datendienst zeitnahfür die Endnutzer bereitgestellt. Diese Fernerkundungsdaten werden mit Emissions- und Ausbreitungs-Modellen kombiniert. Anhand räumlich und zeitlich detaillierter Daten zur Verkehrsbelastung der BAStunter Verwendung technologieabhängiger Emissionsfaktoren können dann wichtige Erkenntnisse überEmissionen des Kfz-Verkehrs gewonnen werden.

Verbundkoordinator

Deutsches Zentrum für Luft- und Raumfahrt e.V., Oberpfaffenhofen

Projektpartner

- Freie Universität Berlin, Institut für Meteorologie, Berlin
- IVU Umwelt GmbH, Freiburg
- TNO, Climate Air and Sustainability Unit, Utrecht, die Niederlande

Web-site: www.s-veld.de

Ansprechpartner: Pieter.Valks@dlr.de

1 INTRODUCTION

1.1 Purpose and scope

The purpose of this document is to present the user manual of the tropospheric NO_2 column and surface NO_2 concentration products derived from Sentinel-5P/TROPOMI as provided in the framework of the mFUND S-VELD project. These products contain daily and monthly mean tropospheric NO_2 column densities and surface NO_2 concentrations for Germany and surrounding areas for the period from January 2018 to December 2020.

1.2 Product description

1.2.1 Tropospheric NO_2 column densities

The algorithm used for the generation of the daily/orbit and monthly gridded products takes the Level 2 tropospheric NO_2 columns generated at DLR within the framework of S-VELD as inputs. Gridded tropospheric NO_2 columns are available over the time period from January 2018 to December 2020.

The gridded NO_2 products are organized into user-friendly and self-describing NetCDF-4 (Network Common Data Form) files, based on the temporal period of collection (daily/orbit and monthly data set).

For daily/orbit data, the logical file name convention is:
TROPOMI_NO2_ORBIT_YYYYMMDDThhmmss_ooooo.nc
The meaning of the different subset of character string is given below:

- YYYY (4 digits) is the year of the processed products
- MM (2 digits, 01-12) is the month of the year the processed products
- DD (2 digits, 01-31) is the day of the month the processed products
- hh (2 digits, 00-23) is the hour of the processed products
- mm (2 digits, 00-59) is the minute of the processed products
- ss (2 digits, 00-59) is the second of the processed products
- ooooo (5 digits) is the orbit number of the corresponding products

The attributes for the daily/orbit files are listed in the following table:

Table 1.1: List of attributes of daily/orbit data with typical values and a short description.

Global Attributes	typical values	Description
Description	Level 3 NO_2 data	short description of the file content
Conventions	CF-1.7	Climate and Forecast convention for variable names, units and dates
Filename	TROPOMI_NO2_ORBIT_YYYYMMDDThhmmss_ooooo.nc	The file name starts with the sensor name, followed by the content (NO2), the coverage, the measurement time as YYYYMMDDThhmmss and 5

		digit orbit number.
composite_type	orbit	Gridded dataset for orbit covering Germany
institution	DLR Deutsches Zentrum für Luft und Raumfahrt	Name of the institution responsible for the data

Group Attributes	typicale values	Description
creator_name	Ka Lok Chan	Name of the person responsible for the data
creator_email	Ka.Chan@dlr.de	contact address
process_time	YYYY-MM-DD HH:mm:ss(UTC)	date and time (UTC) when this file was created
baseProduct	Level 3 SVELD	Basis of this data product
baseProductVersion	v 1.0	version of the underlying data
productAlgorithmVersion	v 1.0	Version of this data product
productFormatType	netCDF	File type
productFormatVersion	4	type version
project	SVELD	
geospatial_lat_min	45.7428	minimum latitude (degree)
geospatial_lat_max	56.3780	maximum latitude (degree)
geospatial_y_resolution	2000 m	Resolution in y dimension (m)
geospatial_y_units	meter north	
geospatial_long_min	0.2427	minimum longitude (degree)
geospatial_long_max	17.9025	maximum longitude (degree)
geospatial_x_resolution	2000 m	Resolution in x dimension (m)
geospatial_x_units	meter east	
sensor	TROPOMI	sensor name
platform	Sentinel 5 Precusor	Satellite
measurement_start_time	YYYY-MM-DD HH:mm:ss(UTC)	Beginning of the measurement
measurement_end_time	YYYY-MM-DD HH:mm:ss(UTC)	End of the measurement

For monthly data, the logical file name convention is:
TROPOMI_NO2_MONTH_YYYYMM.nc
The meaning of the different subset of character string is given below:

- YYYY (4 digits) is the year of the processed products
- MM (2 digits, 01-12) is the month of the year the processed products

The attributes for the monthly files are listed in the following table:

Table 1.2: List of attributes of monthly data with typical values and a short description.

Global Attributes	typical values	Description
Description	Level 3 NO_2 data	short description of the file content
Conventions	CF-1.7	Climate and Forecast convention for variable names, units and dates
Filename	TROPOMI_NO2_MONTH_YYYYMMDD.nc	The file name starts with the sensor name, followed by the content (NO2), the coverage, the measurement month as YYYYMM.
composite_type	month	Gridded dataset for orbit covering Germany

institution	DLR Deutsches Zentrum für Luft und Raumfahrt	Name of the institution responsible for the data

Group Attributes	typicale values	Description
creator_name	Ka Lok Chan	Name of the person responsible for the data
creator_email	Ka.Chan@dlr.de	contact address
process_time	YYYY-MM-DD HH:mm:ss(UTC)	date and time (UTC) when this file was created
baseProduct	Level 3 SVELD	Basis of this data product
baseProductVersion	v 1.0	version of the underlying data
productAlgorithmVersion	v 1.0	Version of this data product
productFormatType	netCDF	File type
productFormatVersion	4	type version
project	SVELD	
geospatial_lat_min	45.7428	minimum latitude (degree)
geospatial_lat_max	56.3780	maximum latitude (degree)
geospatial_y_resolution	2000 m	Resolution in y dimension (m)
geospatial_y_units	meter north	
geospatial_long_min	0.2427	minimum longitude (degree)
geospatial_long_max	17.9025	maximum longitude (degree)
geospatial_x_resolution	2000 m	Resolution in x dimension (m)
geospatial_x_units	meter east	
sensor	TROPOMI	sensor name
platform	Sentinel 5 Precusor	Satellite

Each L3 product is produced at a spatial resolution of 2km×2km from the complete set of L2 orbit files. The Y/X grid defines the dimensions of the data set (see Table 1.3).

Table 1.3: Overview of the dimension in the data file.

Dimension			
Dimension name	**Unit**	**Size**	**Description**
Y	meter north	580	2000 m resolution in y direction
X	meter north	550	2000 m resolution in x direction

1.2.2 Surface NO_2 concentrations

The algorithm used for the generation of the daily/orbit and monthly products takes the Level 3 tropospheric NO_2 columns, meteorological data and geographic information as inputs to derive surface NO_2 concentrations. Surface NO_2 concentration data are available over the time period from January 2018 to December 2020.

The surface NO_2 concentration products are organized into user-friendly and self-describing NetCDF-4 (Network Common Data Form) files, based on the temporal period of collection (daily/orbit and monthly data set).

For daily/orbit data, the logical file name convention is:
TROPOMI_NO2_CONC_ORBIT_YYYYMMDDThhmmss_ooooo.nc
The meaning of the different subset of character string is given below:

- YYYY (4 digits) is the year of the processed products
- MM (2 digits, 01-12) is the month of the year the processed products
- DD (2 digits, 01-31) is the day of the month the processed products
- hh (2 digits, 00-23) is the hour of the processed products
- mm (2 digits, 00-59) is the minute of the processed products
- ss (2 digits, 00-59) is the second of the processed products
- ooooo (5 digits) is the orbit number of the corresponding products

The attributes for the daily/orbit files are listed in the following table:

Table 1.4: List of attributes of daily/orbit data with typical values and a short description.

Global Attributes	typical values	Description
Description	Daily/Orbit surface NO2 concentration	short description of the file content
Conventions	CF-1.7	Climate and Forecast convention for variable names, units and dates
Filename	TROPOMI_NO2_CONC_ORBIT_ YYYYMMDDThhmmss_ooooo.nc	The file name starts with the sensor name, followed by the content (NO2), the coverage, the measurement time as YYYYMMDDThhmmss and 5 digit orbit number.
composite_type	orbit	Gridded dataset for orbit covering Germany
institution	DLR Deutsches Zentrum für Luft und Raumfahrt	Name of the institution responsible for the data

Group Attributes	typicale values	Description
creator_name	Ka Lok Chan	Name of the person responsible for the data
creator_email	Ka.Chan@dlr.de	contact address
process_time	YYYY-MM-DD HH:mm:ss(UTC)	date and time (UTC) when this file was created
baseProduct	Level 3 SVELD	Basis of this data product
baseProductVersion	v 1.0	version of the underlying data
productAlgorithmVersion	v 1.0	Version of this data product
productFormatType	netCDF	File type
productFormatVersion	4	type version
project	SVELD	
geospatial_lat_min	45.7416	minimum latitude (degree)
geospatial_lat_max	56.3915	maximum latitude (degree)
geospatial_y_resolution	500 m	Resolution in y dimension (m)
geospatial_y_units	meter north	
geospatial_long_min	0.2396	minimum longitude (degree)
geospatial_long_max	17.9294	maximum longitude (degree)

Group Attributes	typicale values	Description
geospatial_x_resolution	500 m	Resolution in x dimension (m)
geospatial_x_units	meter east	
sensor	TROPOMI	sensor name
platform	Sentinel 5 Precusor	Satellite
measurement_start_time	YYYY-MM-DD HH:mm:ss(UTC)	Beginning of the measurement
measurement_end_time	YYYY-MM-DD HH:mm:ss(UTC)	End of the measurement

For monthly data, the logical file name convention is:
TROPOMI_NO2_MONTH_YYYYMM.nc
The meaning of the different subset of character string is given below:

- YYYY (4 digits) is the year of the processed products
- MM (2 digits, 01-12) is the month of the year the processed products

The attributes for the monthly files are listed in the following table:

Table 1.5: List of attributes of monthly data with typical values and a short description.

Global Attributes	typical values	Description
Description	Monthly surface NO2 concentration	short description of the file content
Conventions	CF-1.7	Climate and Forecast convention for variable names, units and dates
Filename	TROPOMI_NO2_CONC_MONTH _YYYYMMDD.nc	The file name starts with the sensor name, followed by the content (NO2), the coverage, the measurement month as YYYYMM.
composite_type	month	Gridded dataset for orbit covering Germany
institution	DLR Deutsches Zentrum für Luft und Raumfahrt	Name of the institution responsible for the data

Group Attributes	typicale values	Description
creator_name	Ka Lok Chan	Name of the person responsible for the data
creator_email	Ka.Chan@dlr.de	contact address
process_time	YYYY-MM-DD HH:mm:ss(UTC)	date and time (UTC) when this file was created
baseProduct	Level 3 SVELD	Basis of this data product
baseProductVersion	v 1.0	version of the underlying data
productAlgorithmVersion	v 1.0	Version of this data product
productFormatType	netCDF	File type
productFormatVersion	4	type version
project	SVELD	
geospatial_lat_min	45.7416	minimum latitude (degree)
geospatial_lat_max	56.3915	maximum latitude (degree)
geospatial_y_resolution	500 m	Resolution in y dimension (m)
geospatial_y_units	meter north	

Group Attributes	typicale values	Description
geospatial_long_min	0.2396	minimum longitude (degree)
geospatial_long_max	17.9294	maximum longitude (degree)
geospatial_x_resolution	500 m	Resolution in x dimension (m)
geospatial_x_units	meter east	
sensor	TROPOMI	sensor name
platform	Sentinel 5 Precusor	Satellite

Each surface NO2 concentration product is produced at a spatial resolution of 0.5km×0.5km from the complete set of L2 orbit files. The Y/X grid defines the dimensions of the data set (see Table 1.3).

Table 1.6: Overview of the dimension in the data file.

Dimension			
Dimension name	**Unit**	**Size**	**Description**
Y	meter north	2320	500 m resolution in y direction
X	meter north	2200	500 m resolution in x direction

2 TROPOMI INSTRUMENT AND TROPOSPHERIC NO_2 RETRIEVAL

2.1 Sentinel-5P/TROPOMI instrument

TROPOMI is a push broomimaging spectrometer coveringwavelength bands between the ultraviolet and the shortwave infrared. Launched inOctober 2017, the 7-year-lifetimesensor TROPOMI is the single payload aboard the Sentinel-5 Precursor satellite in a Sun-synchronous orbit with an overpass time of 13:30 local solar time (Veefkind et al., 2012).With a spatial resolution of 5.5 km×3.5 km (along×across track), TROPOMI is a major step forward compared to itspredecessors Ozone Monitoring Instrument(OMI), Scanning Imaging Absorption Spectrometerfor Atmospheric Chartography (SCIAMACHY), and Global Ozone Monitoring Experiment-2 (GOME-2). The spatial resolutionis combined with a swath width of ~2600 kminthe direction across the track of the satelliteto allow for daily global coverage.With its global coverage and open data policy, the mission will support global efforts to monitor pollution and to improve our understanding of chemical and physical atmospheric processes.

2.2 Tropospheric NO2 retrieval algorithm

The Differential Optical Absorption Spectroscopy (DOAS) method is used to determine NO_2 slant column densities from calibrated TROPOMI (ir)radiance data in the wavelength interval 405–465 nm. NO_2absorption features are prominent in this wavelength range, and interference effects are manageable. A single NO_2 cross-section reference spectrum at 220 K (Vandaele et al. 2002)is used, and the interfering species ozone, O_2-O_2, water vapor, and liquid water are included in the DOAS fit, as well as an additive Fraunhofer Ring spectrum. A linear intensity offset correction is fitted as another effective cross-section to correct for the stray light in the spectrometer, the inelastic scattering in the ocean, and remaining calibration issue. Shift and stretch parameters are applied to cross-section wavelength grids to improve the wavelength registration and compensate for inaccuracy in the wavelength calibration.

The initial total vertical column density is computed under the assumption that the troposphere is not polluted. Therefore, the air mass factor is based on stratospheric NO_2 profiles only, whereas the tropospheric NO_2 amount is assumed to be negligible. This approach is valid over large parts of the Earth, but in areas with significant tropospheric NO_2, the total column densities are underestimated and need to be corrected. The air mass factors are calculated with the VLIDORT radiative transfer model for 437.5 nm, since NO_2 is an optically thin absorber in this wavelength region.

Tropospheric NO_2columns are obtained from the initial total columns by estimating the stratospheric content and removing it from the total amount. For the stratosphere estimation, the STRatospheric Estimation Algorithm from Mainz (STREAM) method (Beirle et al. 2016) is applied. Belonging to the reference sector method, STREAM relies on the totalcolumn measurements over clean, remote regions as wellas over clouded scenes where the tropospheric column iseffectively shielded. The contribution of individual satellitemeasurements to the stratospheric estimate is controlled byvarious weighting factors.STREAM is a flexible and robustalgorithm and does not require input from chemical transportmodels.

After the stratosphere-troposphere separation, the tropospheric NO_2 columns can be determined using a tropospheric air mass factor. For the tropospheric air mass factor computation, the surface albedo is described by a climatology based on three years of OMI Lambertian-equivalent reflectivity measurements at 440 nm (Kleipool et al., 2008). Daily TM5-MP vertical NO_2profiles (Williams et al., 2017) simulated at a global 1°×1° resolution are used as a priori NO_2vertical profiles (determined for the TROPOMI overpass time). TROPOMI-derived cloud properties, determined with the OCRA and ROCINN algorithms are used to calculate the air mass factors for scenarios in the presence of clouds (see also Sect. 2.3). The calculation of the tropospheric NO_2 columns is complicated in case of (partly)

cloudy conditions. For many measurements over cloudy scenes, the cloud-top is well above the NO_2pollution in the boundary layer, and when the clouds are optical thick, the enhanced tropospheric NO_2 concentrations cannot be detected by TROPOMI. Therefore, the tropospheric NO_2 columns calculated for observations with a cloud fraction > 20% are flagged in the TROPOMI L2 product.

2.3 Cloud parameters

The presence of clouds significantly affects the retrieval of tropospheric trace gases in the visible spectral range, and it is very important to derive information on cloud properties from the TROPOMIobservations to study the indirect effects of residual cloud contamination. In the NO_2 retrieval, the computation of the NO_2 columns assumes the independent pixel approximation (IPA) for cloud treatment.

The OCRA and ROCINN algorithms (Lutz et al., 2016; Loyola et al., 2018) are used for obtaining TROPOMI cloud information. Clouds are regarded as reflecting Lambertian surfaces, and cloud information is reduced to the specification of three parameters: cloud fraction, cloud-top albedo and cloud-top pressure. The radiometric cloud fraction is determined with OCRA by separating a spectral scene into cloudy contribution and cloud-free background, and the cloud pressureand cloud albedo are derived using the ROCINN algorithm by comparing simulatedand measured radiance in and adjacent to the oxygen A-band around 760 nm.

2.4 Error estimates and validation

An estimation of the error budget for the TROPOMI tropospheric NO_2 column is provided in Table 2.1. This includes typical errors on NO_2 slant columns, errors related to the separation of stratospheric and tropospheric NO_2, and systematic errors due to uncertainties in model parameters such as clouds, surface albedo, and a priori profile shape, affecting the tropospheric airmass factor.

Table 2.1: Contributions to the overall uncertainty in the TROPOMItropospheric NO_2 column for pollutedconditions.

Error source	Uncertainty
Slant column	0.5×10^{15}molec/cm^2
Stratospheric column	0.2×10^{15}molec/cm^2
Tropospheric air mass factor	10-35%
Tropospheric column	**15-50%**

The tropospheric NO_2 productsare regularly validated by comparison with correlativeobservations from ground-based multiple-axis DOAS (MAXDOAS) spectrometers (Pinardi et al., 2015). MAXDOAS measurements at the Xianghe BIRA-IASB site, located in the highly populated polluted regions in China, are valuable sources of validation data for TROPOMI.Good correlations between TROPOMI and the ground-based MAXDOAS data are obtained, both in terms of correlation coefficients (value of 0.96) and slopes of the regression analysis (value of 1.12). Time series of TROPOMI measurements above Xianghe and the comparisons to MAXDOAS tropospheric NO_2 data show that the pollution episodes are captured well by TROPOMI. The comparisons of monthly averaged columns show consistent seasonal variations with high NO_2 in winter and low NO_2 in summer.

3 DAILY AND MONTHLY MEAN TROPOSPHERIC NO_2 DATA

3.1 Theoretical description: sampling and gridding

The level 2 TROPOMI NO_2 columns are gridded onto a high resolution horizontal grid with a spatial resolution of 2km × 2km. The gridded data is based on all valid measurementscovering Germany. Valid measurements are defined with corresponding solar zenith angle smaller than 85°, intensity weighted cloud fraction smaller than 0.5, root mean square of spectral fit residual less than 0.002, and AMF larger than 0.1.The vertical column of each valid pixel is stored in all grid points lying within the satellite ground pixel boundaries. These pixel boundaries are taken from the level 1B data. For overlapping pixels, a weighted average is calculated following the equation below.

$$VCD_g = \frac{\sum VCD_i \times w_i}{\sum w_i}$$

where the weighting (w_i) is defined by

$$w_i = \frac{1}{(1 + 3 \times CF_{iw_i})^2}$$

VCD_g is the gridded NO_2 column while VCD_i represents each individual measurement. The weighting is denoted as w which is dependent on the intensity weighted cloud fraction (CF_{iw}). As clear sky data are more reliable, the gridding scheme gives higher weights to clear sky pixels. Details of the gridding algorithm can be found in Wenig et al., 2008 andChan et al., 2012.

3.2 Parameters for daily/orbit tropospheric NO_2 product

The daily/orbit tropospheric NO_2 column for each grid cell is given in the netcdf file. The temporal resolution of the product is per orbit (2 to 3 overpasses per day). Grid maps covering Germany and its surrounding areas are in UTM coordinate with resolution of 2km × 2km. The errors associated to the L2 retrieval are computed. A weighted tessellation technique is used to compute the NO_2 columns and the associated errors, as described in Sect. 3.1.

Table 3-1: List of Level 3 mapped NO_2 products.

Variable name	Unit	Size	Description
y	Meter north	580	Y coordinate of the centre of the gridded data
x	Meter east	550	X coordinate of the centre of the gridded data Longitudinal resolution
y_corner	Meter north	581	Y coordinate of the corners of the gridded data
x_corner	Meter east	551	X coordinate of the corners of the gridded data Longitudinal resolution
latitude	Degree north	580×550	Latitude of the centre of the gridded data
longitude	Degree east	580×550	Longitude of the centre of the gridded data
latitude_corner	Degree north	581×551	Latitude of the corners of the gridded data
longitude_corner	Degree east	581×551	Longitude of the corners of the gridded data
no2_tropospheric_column	mol m^{-2}	580×550	averaged tropospheric NO_2 column

Variable name	Unit	Size	Description
no2_tropospheric_column _error	mol m^{-2}	580×550	averaged error associated to the tropospheric NO_2 column
Number_ofmeasurements	#	580×550	number of measurement used in the calculation of average value
cloud_fraction	unitless	580×550	radiometric cloud fraction
intensity_weighted_cloud_fraction	unitless	580×550	intensity weighted cloud fraction at 435nm

As an example, the daily/orbit gridded tropospheric NO_2 column (parameter no2_tropospheric_column) derived from TROPOMI/S5P measurements using weighted tessellation is shown in Fig. 3.1 for orbit 2484 taken on 6th April 2018. High tropospheric NO_2 columns can be observed over major cities, i.e., Hamburg, Mannheim and Düsseldorf.

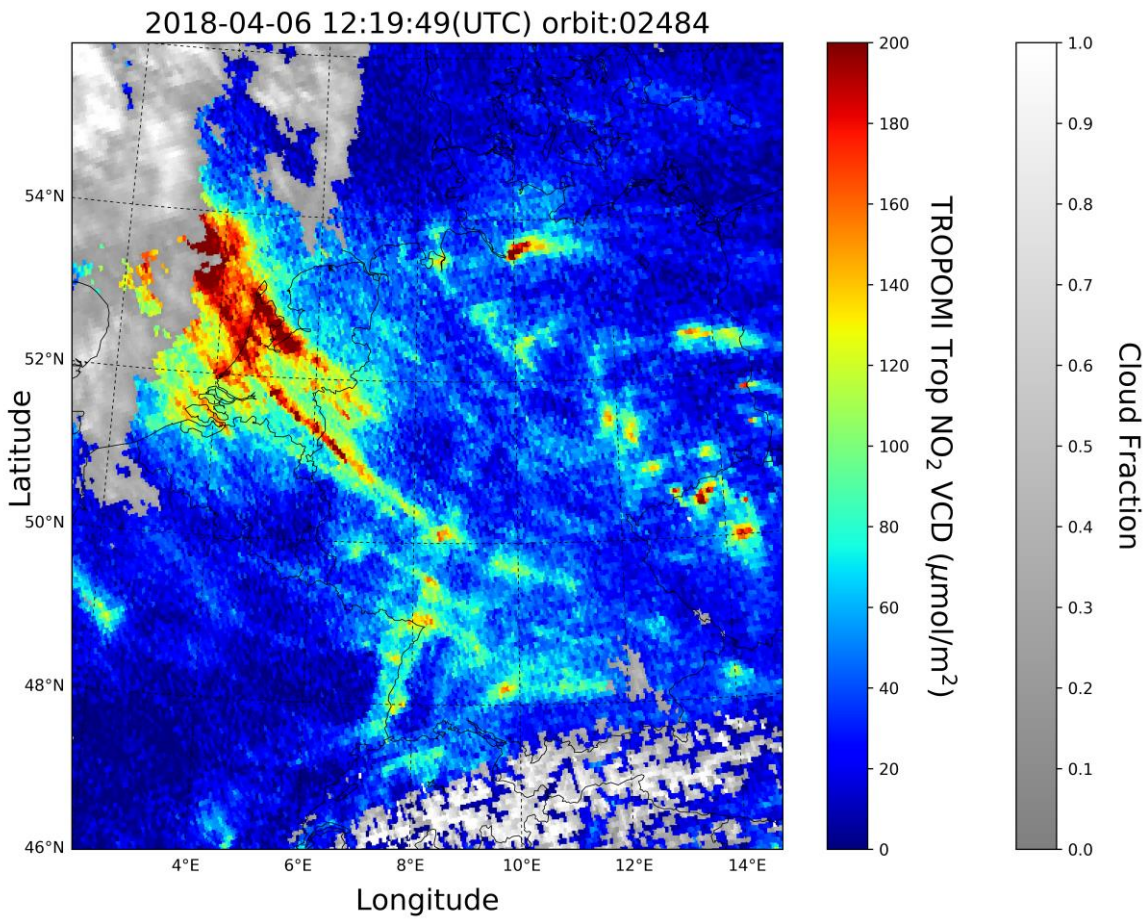

Figure 3.1: Spatial distribution of the averaged tropospheric NO_2 column derived from TROPOMI/S5P measurements on6th April 2018.

3.3 Parameters for monthly mean tropospheric NO_2 product

The monthly tropospheric NO_2 column for each grid cell is given in the netcdf file. The temporal resolution of the product is monthly (calendar). Grid maps covering Germany and its surrounding areas are in UTM coordinate with resolution of 2km × 2km. The errors associated to the L2 retrieval are computed. A weighted tessellation technique is used to compute the NO_2 columns, as described in Sect. 3.1.

Table 3-2: List of Level 3 mapped NO_2 products.

Variable name	Unit	Size	Description
y	Meter north	580	Y coordinate of the centre of the gridded data
x	Meter east	550	X coordinate of the centre of the gridded data Longitudinal resolution
y_corner	Meter north	581	Y coordinate of the corners of the gridded data
x_corner	Meter east	551	X coordinate of the corners of the gridded data Longitudinal resolution
latitude	Degree north	580×550	Latitude of the centre of the gridded data
longitude	Degree east	580×550	Longitude of the centre of the gridded data
latitude_corner	Degree north	581×551	Latitude of the corners of the gridded data
longitude_corner	Degree east	581×551	Longitude of the corners of the gridded data
no2_tropospheric_column	mol m^{-2}	580×550	averaged tropospheric NO_2 column
no2_tropospheric_column _error	mol m^{-2}	580×550	averaged error associated to the tropospheric NO_2 column
Number_ofmeasurements	#	580×550	number of measurement used in the calculation of average value

As an example, the monthly gridded tropospheric NO_2 column (parameter no2_tropospheric_column) derived from TROPOMI/S5P measurements using weighted tessellation is shown in Fig. 3.2 for April 2018. Elevated tropospheric NO_2 columns can be observed over major cities and industrial regions, i.e., Hamburg, Mannheim and lower Rhine regions.

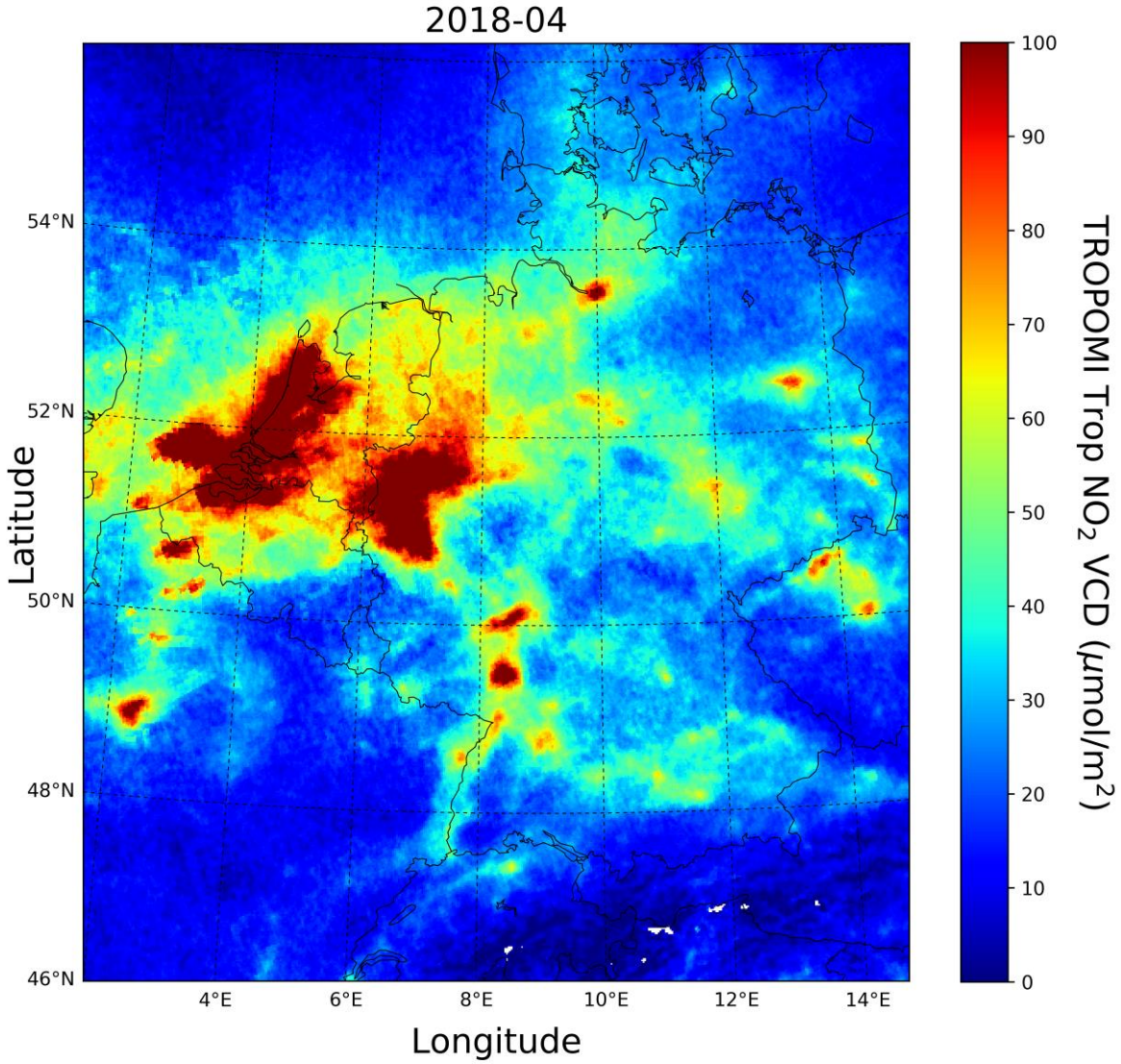

Figure 3.2: Spatial distribution of the averaged tropospheric NO_2 column derived from TROPOMI/S5P measurements in April 2018. The tropospheric NO_2 column measurements have been aggregated by weighted tessellation.

4 SURFACE NO_2 CONCENTRATION ESTIMATION

Surface NO_2 concentrations are estimated using a machine learning method. TROPOMI satellite observations of tropospheric NO_2 column, in-situ measurements of NO_2 concentration, meteorological data and geographical information are used as input to train the dedicated neural network model for the estimation of surface NO_2 concentration over Germany. The results from the neural network model have been validated again ground based in-situ measurements and chemistry transport model simulations (Chan et al., 2021). The retrieval methodology are briefly described below. More detailed description and validation results can be found in Chan et al., 2021.

4.1 Input data

4.1.1 Satellite data

Tropospheric NO_2 columns measured by TROPOM is used as input for the retrieval of surface NO_2 concentrations (see section 2). In order to pair up the satellite NO_2 data with other data sets, TROPOMI tropospheric NO_2 columns are first regridded onto regular grid with spatial resolution of 0.5km×0.5km. Satellite observations are insensitive to NO_2 below clouds, and uncertainties of satellite observations under cloudy condition are much larger than that of clear sky measurements. Therefore, proper cloud filtering is necessary to avoid cloud contaminated data affecting following data analysis. In order to obtain a balance between having sufficient measurements while minimizing the influences of cloud contaminated data, only satellite data with cloud radiance fraction (CRF) of less than 50% are considered in the analysis. Measurement uncertainties related to cloud contamination are typically within 15% for observations with CRF smaller than 0.5 over polluted regions. Therefore, the threshold of CRF of 0.5 is selected in this study. In addition to cloud filtering, we have also excluded data with solar zenith angle larger than 85° and root mean square of spectral fit residual larger than 0.001 in the following analysis.

4.1.2 Ambient Air Quality Monitoring Station Data

Ambient NO_2 concentrations data in Germany are acquired from air quality monitoring network operated by the Environment Agency of each Federal State (Landesamt für Umwelt, LfU) and summarized by the Federal Environment Agency (Umwelt Bundesamt, UBA) of Germany. The air quality monitoring network comprises 455 in-situ monitoring stations, in which 268 of them are ambient stations, 30 of them are industrial stations and 157 of them are road side stations. In this study, we only used data measured by ambient monitoring stations to avoid strong influences from very local sources, i.e., traffic and industrial. These monitoring stations cover both urban, suburban and rural areas of Germany. These air quality monitoring stations use in-situ chemiluminescence NO_2 analyzers to measure ambient NO_2 concentrations. Details of the air quality monitoring network as well as air quality monitoring data can be found on the website of the European Environment Agency (https://www.eea.europa.eu/).

4.1.3 Meteorological Data

A number of meteorological parameters are also used to improve the estimation of surface NO$_2$ concentrations over Germany. In this study, meteorological parameters, such as, boundary layer height, surface air temperature, wind speed, relative humidity, precipitation and downward ultraviolet radiation, are taken from the European Center for Medium-Range Weather Forecasts (ECMWF) ERA5 reanalysis product. The ERA5 reanalysis data covers a long time period, since 1979, and provides consistent meteorological data on a global scale. The reanalysis data are produced with a data assimilation scheme, which combined various measurements as prior information from model forecasts. The original ERA5 data set is in a spatial resolution of ~31km (T255 Spectral). The data are then transformed to the latitude-longitude coordinate system, with a horizontal resolution of 0.25°×0.25°through the Copernicus Climate Change Service. The ERA5 reanalysis data is available on the Copernicus Climate Data Store (https://cds.climate.copernicus.eu).

4.1.4 Surface Elevation Data

The territory of Germany stretches from the North Sea and the Baltic Sea across the Northern Lowland to the Alps in the South. The surface elevation of Germany ranges from sea level up to 2963m above sea level (a.s.l.) with the highest point at Zugspitze. As the surface elevation of Germany varies in such a wide range, it is also an important factor for the estimation of surface level NO_2 concentrations

from satellite observations of vertical columns. In this study, surface elevation data from the Digital Elevation Model over Europe (EU-DEM) is used. EU-DEM is a digital surface model provided by the Copernicus Land Monitoring Service. It is a hybrid product produced from the Shuttle Radar Topography Mission (SRTM) and Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) Global Digital Elevation Model (GDEM) data using a weighted averaging approach. The spatial resolution of the EU-DEM data set is about 30m.

4.2 Surface NO_2 Concentration Retrieval

Machine learning approximates a function that represents the relationships between inputs and outputs data by both linear and non-linear regressions. We use the artificial neural network (refers as neural network (NN) from hereafter) to learn the non-linear relationships between inputs and outputs. The neural network model formulates surface NO_2 concentrations as a function of all input parameters. All input and output parameters regridded or spatio-temporal interpolated to a resolution of daily 0.5km×0.5km for the training and validation of the neural network model.

The neural network model used in this study consists of 4 hidden layers. The number of neurons of the first to the last layer is 40, 20, 10 and 5, respectively. The model is trained using the stochastic gradient descent (SGD) algorithm with mean squared error (MSE) as loss function. We have also tested other settings with more layers and number of neurons. However, the performance of the model does not show any significant improvement. Therefore, we use these settings to train the neural network model.

In total, there are 76,338 pairs of data available for training and validation. Relatively more data is available in spring and summer compared to winter and autumn. It is mainly due cloudiness in the cold months. The available data pairs are then randomly splitted into two groups, 90% of the data (68,704) are used for the training of the neural network model, and the rest (7,634) are used to validate the model.

5 DAILY AND MONTHLY MEAN SURFACE NO_2 DATA

5.1 Parameters for daily surface NO_2 product

The daily surface NO_2 concentration for each grid cell is given in the netcdf file. The temporal resolution of the product is per orbit (2 to 3 overpasses per day). Grid maps only cover Germany which are in UTM coordinate with resolution of 0.5km × 0.5km.

Table 5-1: List of mapped daily surface NO_2 products.

Variable name	Unit	Size	Description
y	Meter north	2320	Y coordinate of the centre of the gridded data
x	Meter east	2200	X coordinate of the centre of the gridded data Longitudinal resolution
latitude	Degree north	2320×2200	Latitude of the centre of the gridded data
longitude	Degree east	2320×2200	Longitude of the centre of the gridded data
no2_concentration	$\mu g\ m^{-3}$	2320×2200	Surface NO_2 concentration
cloud_fraction	unitless	2320×2200	radiometric cloud fraction

As an example, the daily surface NO_2 concentration (parameter no2_concentration) derived from TROPOMI/S5P measurements using machine learning approach is shown in Fig. 5.1 for orbit 2484 taken on 6th April 2018.

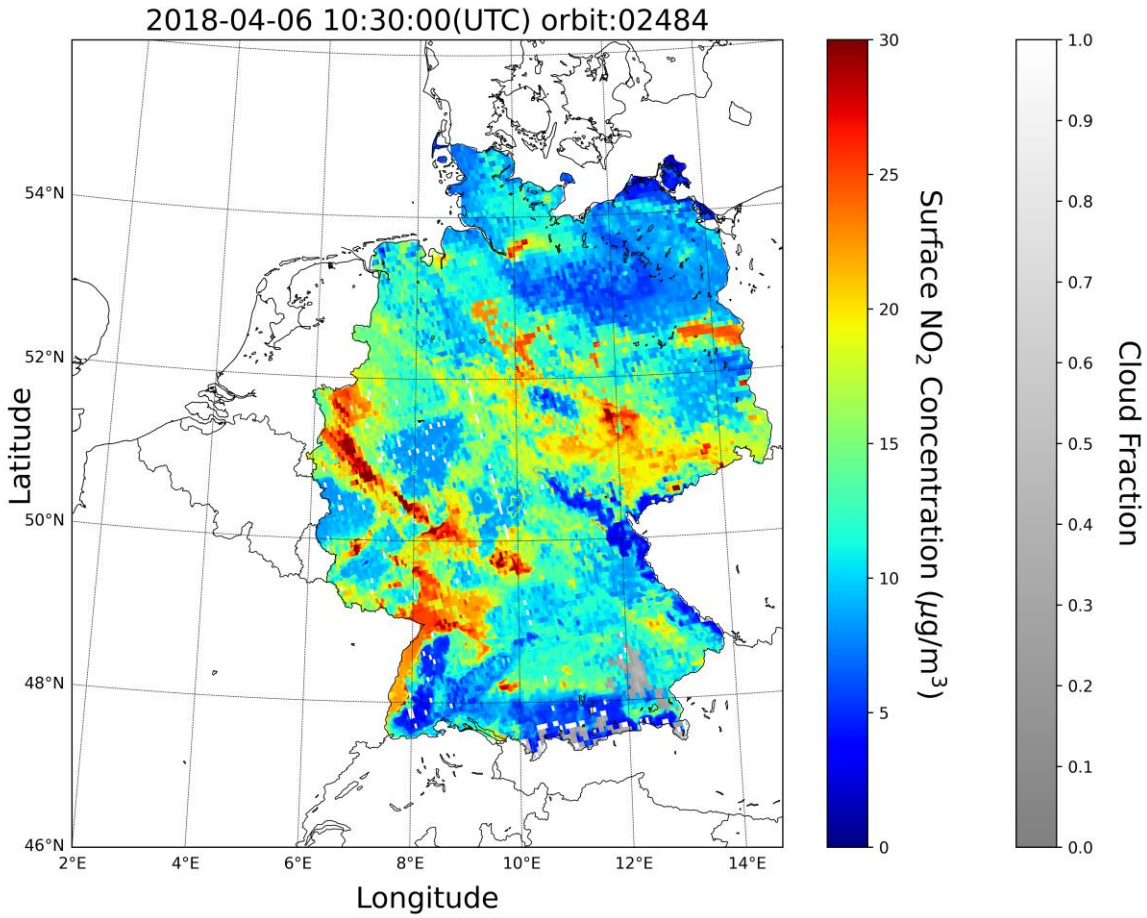

Figure 5.1: Spatial distribution of surface NO_2 concentration derived from TROPOMI/S5P measurements using the machine learning approach on 6th April 2018.

5.2 Parameters for monthly surface NO_2 product

The monthy surface NO_2 concentration for each grid cell is given in the netcdf file. The temporal resolution of the product is monthly (calendar). Grid maps only cover Germany which are in UTM coordinate with resolution of 0.5km × 0.5km.

Table 5-2: List of mapped monthly surface NO_2 products.

Variable name	Unit	Size	Description
y	Meter north	2320	Y coordinate of the centre of the gridded data
x	Meter east	2200	X coordinate of the centre of the gridded data Longitudinal resolution
latitude	Degree north	2320×2200	Latitude of the centre of the gridded data
longitude	Degree east	2320×2200	Longitude of the centre of the gridded

Variable name	Unit	Size	Description
			data
no2_concentration	µg m^{-3}	2320×2200	averaged surface NO_2 concentration
Number_ofmeasurements	#	2320×2200	number of measurement used in the calculation of average value

As an example, the monthly surface NO_2 concentration (parameter no2_concentration) derived from TROPOMI/S5P measurements using machine learning approach is shown in Fig. 5.2 for April 2018.

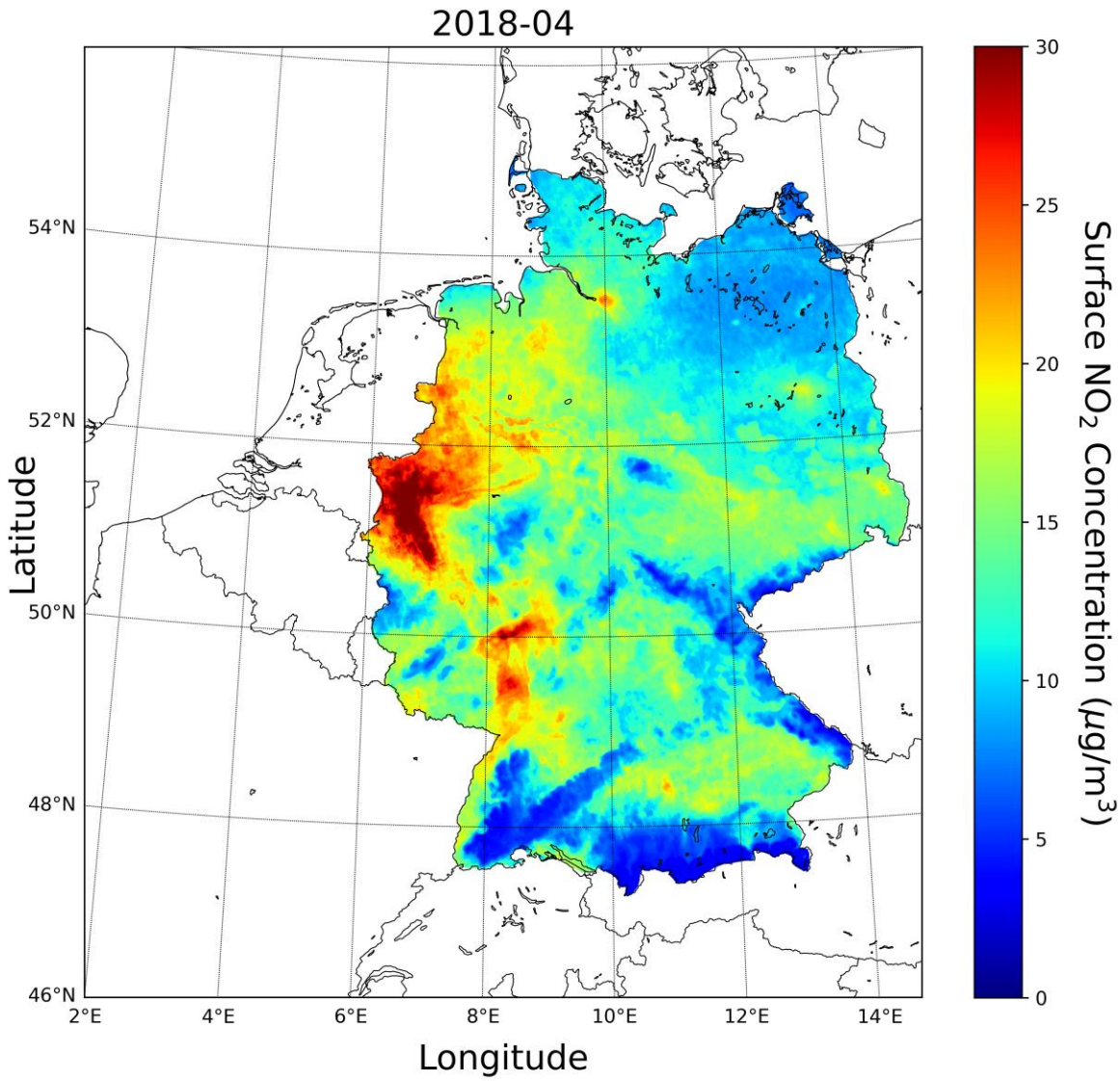

Figure 5.2: Spatial distribution of averaged surface NO_2 concentration derived from TROPOMI/S5P measurements using machine learning approach in April 2018.

REFERENCES

Beirle, S., Hörmann, C., Jöckel, P., Liu, S., Penning de Vries, M., Pozzer, A., Sihler, H., Valks, P., and Wagner, T.: The STRatospheric Estimation Algorithm from Mainz (STREAM): estimating stratospheric NO_2 from nadir-viewing satellites by weighted convolution, Atmos. Meas. Tech., 9, 2753–2779, https://doi.org/10.5194/amt-9-2753-2016, 2016.

Chan, K.L., Pöhler, D., Kuhlmann, G., Hartl, A., Platt, U., Wenig, M.O.: NO_2 measurements in Hong Kong using LED based long path differential optical absorption spectroscopy, Atmos. Meas. Tech., 5, 901–912, https://doi.org/10.5194/amt-5-901-2012, 2012.

Chan, K.L., Khorsandi, E., Liu, S., Baier, F., Valks, P.: Estimation of Surface NO_2 Concentrations over Germany from TROPOMI Satellite Observations Using a Machine Learning Method. *Remote Sens.* **2021**, *13*, 969. https://doi.org/10.3390/rs13050969

Kleipool, Q., Ludewig, A., Babić, L., Bartstra, R., Braak, R., Dierssen, W., Dewitte, P.-J., Kenter, P., Landzaat, R., Leloux, J., Loots, E., Meijering, P., van der Plas, E., Rozemeijer, N., Schepers, D., Schiavini, D., Smeets, J., Vacanti, G., Vonk, F., and Veefkind, P.: Pre-launch calibration results of the TROPOMI payload on-board the Sentinel-5 Precursor satellite, Atmos. Meas.Tech., 6439-6479, 2018.

Liu, S., Valks, P., Pinardi, G., Xu, J., Chan, K. L., Argyrouli, A., Lutz, R., Beirle, S., Khorsandi, E., Baier, F., Huijnen, V., Bais, A., Donner, S., Dörner, S., Gratsea, M., Hendrick, F., Karagkiozidis, D., Lange, K., Piters, A. J. M., Remmers, J., Richter, A., Van Roozendael, M., Wagner, T., Wenig, M., and Loyola, D. G.: An improved tropospheric NO_2 column retrieval algorithm for TROPOMI over Europe, Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2021-39, in review, 2021.

Loyola, D., Garcıá, S. G., Lutz, R., Argyrouli, A., Romahn, F., Spurr, R. J., Pedergnana, M., Doicu, A., Garcıá, V. M., and Schlüssler, O.: The operational cloud retrieval algorithms from TROPOMI on board Sentinel-5 Precursor, Atmos. Meas. Tech., 11, 409, 2018.

Lutz, R., Loyola, D., GimenoGarcıá, S., and Romahn, F.: OCRA radiometric cloud fractions for GOME-2 on MetOp-A/B, Atmos. Meas. Tech., 9, 2357-2379, 2016.

Pinardi, G., Lambert, J.-C., Y. Huan, I. De Smedt, J. Granville, M. Van Roozendael and P. Valks, O3M SAF Validation report of GOME-2 GDP 4.8 NO_2 column data, SAF/O3M/IASB/VR/NO2, Issue 1/0, Oct. 2015.

Vandaele, A. C., Hermans, C., Fally, S., Carleer, M., Colin, R., Merienne, M.-F., Jenouvrier, A., and Coquart, B.: High-resolution Fourier transform measurement of the NO_2 visible and nearinfrared absorption cross sections: Temperature and pressure effects, J. Geophys. Res. Atmos., 107, 2002.

Veefkind, J.P., Aben, I., McMullan, K., Förster, H., De Vries, J., Otter, G., Claas, J., Eskes, H.,DeHaan, J., Kleipool, Q., van Weele, M., Hasekamp, O., Hoogeveen, R., Landgraf, J., Snel, R.,Tol, P., Ingmann, P., Voors, R., Kruizinga, B., Vink, pR., Visser, H., and Levelt, P. F.: TROPOMIon the ESA Sentinel-5 Precursor: A GMES mission for global observations of the atmosphericcomposition for climate, air quality and ozone layer applications, Remote Sens. Environ., 120,70-83, 2012.

Wenig, M.O., Cede, A.M., Bucsela, E.J., Celarier, E.A.,Boersma, K.F., Veefkind, J.P., Brinksma, E.J., Gleason, J.F., Herman, J.R.. Validation of OMI tropospheric NO_2 column densities using direct-sun mode brewer measurements at NASA goddard space flight center.Journal of Geophysical Research: Atmospheres 2008;113(D16).

Williams, J. E., Boersma, K. F., Le Sager, P., and Verstraeten, W. W.: The high-resolutionversion of TM5-MP for optimized satellite retrievals: description and validation, Geosci. ModelDev., 10, 721,

2017.