



**Deutsches Zentrum
für Luft- und Raumfahrt e.V.**
in der Helmholtz-Gemeinschaft

S5P/TROPOMI

Total Column Water Vapour ATBD



sentinel-5p

document number : S5P-L2-DLR-ATBD-TCWV
authors : Ka Lok CHAN
CI identification : CI-400A-ATBD
issue : 1.2
date : 2021-12-15
status : Review

Document approval record

	digital signature	
prepared:	Ka Lok Chan	<i>Chan Ka Lok</i>
	Sander Slijkhuis	
checked:		
approved PI:		
approved PM:		
approved CM:		

Contents

1 Introduction	7
1.1 Purpose of the ATBD.....	7
1.2 Document overview.....	7
1.3 Acknowledgements.....	7
2 Applicable and reference documents	8
2.1 Applicable documents.....	8
2.2 Standard documents.....	8
2.3 Reference documents.....	8
2.3.1 Electronic references.....	8
3 Terms, definitions and abbreviated terms	9
3.1 Terms and definitions.....	9
3.2 Acronyms and abbreviations.....	9
4 Introduction to the S5P Data Products	12
4.1 Heritage.....	13
4.2 Product requirements.....	14
4.3 Overview of the retrieval method.....	14
4.3.1 Off-Line (OFFL) algorithm.....	14
4.4 General design considerations.....	14
5 This ATBD Algorithm descriptions	15
5.1 The blue band TCWV retrieval.....	15
5.1.1 Water vapour slant column retrieval.....	15
5.1.2 Air mass factor calculations.....	18
5.1.3 Box air mass factor look-up table.....	18
5.1.4 Dynamic search of a priori water vapour profile.....	19
5.1.5 Surface Albedo.....	22
5.1.6 Partially cloudy scene observations.....	22
5.1.7 Aerosol.....	23
6 Input-Output description	24
6.1 S5P TCWV product description and size.....	24
6.2 Auxiliary information needs.....	25
6.3 Level 1 information needs.....	27
7 Error estimation	29
7.1 Slant column error.....	29
7.2 Clear sky air mass factor error.....	29

7.3 Cloudy sky air mass factor error.....	30
7.4 Cloudy sky air mass factor error.....	31
7.5 Total error.....	31
7.6 Averaging kernel.....	31
7.7 QA value.....	31
8 Validation.....	32
8.1 Data sets used for validation.....	32
8.1.1 GOME-2 TCWV product.....	32
8.1.2 MODIS TCWV product.....	32
8.1.3 SSMIS TCWV product.....	33
8.2 Validation results.....	33
8.2.1 GOME-2 comparison.....	33
8.2.2 MODIS comparison.....	35
8.2.3 SSMIS comparison.....	35
9 Conclusions.....	36
10 References.....	37

Table of Figures

Figure 1: Water vapour absorption cross section. (a) The horizontal bars show the spectral band available to various satellite sensors. The wavelength range used in this study and the operational GOME-2 product are highlighted in blue and red colour, respectively. Zoom in of (a) to the blue and red bands are shown in (b) and (c), respectively. The red curves show the water vapour absorption cross section convoluted with the instrument slit function. Note that the scale of the y axis of each plot is different.....12

Figure 2: An example of the DOAS retrieval of water vapour slant column for a TROPOMI spectrum taken on 1st July 2018 over the Pacific Ocean. The spectrum was taken with solar zenith angle of 34.6 and viewing zenith angle of 61.1°. The retrieved water vapour slant column is 89.8 kg m⁻².....17

Figure 3: Statistical analysis of water vapour vertical profiles from ECMWF ERA-Interim reanalysis data over a small region of Pacific Ocean (5°S-5°N, 180°W-170°W) in July of 2008 to 2018. Water vapour profiles are sorted by their total column density into 8 ranges, which are (a) 20-25kg m⁻², (b) 25-30 kg m⁻², (c) 30-35 kg m⁻², (d) 35-40 kg m⁻², (e) 40-45 kg m⁻², (f) 45-50 kg m⁻², (g) 50-55 kg m⁻² and (h) 55-60 kg m⁻². Shaded areas indicate the 1 σ standard deviation variation of water vapour mixing ratio. (i) shows the normalized mean water vapour profile shapes of these 8 ranges.....21

Figure 4: Monthly averaged TCWV retrieved from (a) TROPOMI observations in July 2018, (b) GOME-2 observations in July 2018, (d) TROPOMI observations in January 2019 and (e) GOME-2 observations in January 2019. The scatter plots of TROPOMI and GOME-2 TCWV for July 2018 and January 2019 are shown in (c) and (f), respectively.....34

Figure 5: Monthly averaged TCWV retrieved from (a) TROPOMI observations in July 2018, (b) MODIS observations in July 2018, (d) TROPOMI observations in January 2019 and (e) MODIS observations in January 2019. The scatter plots of TROPOMI and MODIS TCWV for July 2018 and January 2019 are shown in (c) and (f), respectively.....35

Figure 6: Monthly averaged TCWV retrieved from (a) TROPOMI observations in July 2018, (b) SSMIS observations in July 2018, (d) TROPOMI observations in January 2019 and (e) SSMIS observations in January 2019. The scatter plots of TROPOMI and SSMIS TCWV for July 2018 and January 2019 are shown in (c) and (f), respectively.....35

Index of Tables

Table 1: Summary of the water vapour slant column retrieval results with different spectral fitting window for TROPOMI measurements taken on 1st July 2018 (orbit 3698 - 3711) over the tropics (30°S - 30°N).....	16
Table 2: Parameters in the box air mass factor look-up table.....	19
Table 3: List of output fields in TCWV level-2 product.....	24
Table 4: Dynamic auxiliary information needs in theTCWV retrieval algorithms.....	26
Table 5: Static auxiliary information needs in the TCWV algorithms.....	27

Listof<TBD>'s

1 Introduction

1.1 Purpose of the ATBD

Atmospheric water vapour is the most important natural greenhouse gas in the troposphere, accounting for more than 60% of the greenhouse effect [Clough and Iacono, 1995; Kiehl and Trenberth, 1997]. Despite this importance, its roles in climate and its reactions to climate change are still difficult to assess. As the atmosphere gets warmer, water vapour contents are expected to rise faster than the total precipitation amount, which is governed by the surface heat budget through evaporation [Trenberth and Stepaniak, 2003]. This results in a 'positive water vapour feedback' that further amplifies the original warming effect [Colman, 2003; Soden et al., 2005; Soden and Held, 2006]. On the other hand, clouds are known to have positive effects on cooling the earth's surface [Bellomo et al., 2014; Brown et al., 2016]. However, the net cooling or warming effect of clouds in a continuous warming atmosphere is not yet well understood [Boucher et al., 2013]. To investigate these complex interactions and evaluate climate models, continuous monitoring of the spatio-temporal variations of total column water vapour (TCWV) on a global scale is necessary [Hartmann et al., 2013].

The TROPOMI (Tropospheric Monitoring Instrument) is the payload instrument for the Sentinel 5 Precursor (S5P) Mission. The S5P platform was launched into a sun-synchronous low-earth orbit in October 13th, 2017. TROPOMI is a nadir-viewing atmospheric chemistry instrument measuring at moderate spectral resolution from the UV to the near infrared [RD7]. TROPOMI/S5P provides global measurements of a number of trace species, including ozone, nitrogen dioxide, formaldehyde, sulphur dioxide. TROPOMI/S5P also covers the blue wavelength band which can be used for water vapour retrieval.

This document describes the TROPOMI/S5P total column water vapour retrieval algorithm. The algorithm consists of two major steps. The first step is the retrieval of water vapour slant columns. The second step is the conversion of the water vapour slant columns to vertical columns. The present document is the Algorithm Theoretical Basis Document (ATBD) of the TCWV retrieval algorithm. The purpose of the ATBD is to provide detailed mathematical and physical descriptions of the algorithm, along with discussions of algorithm inputs and outputs, data products, algorithm validation and error analysis. The ATBD also includes a summary description of the TROPOMI instrument.

1.2 Document overview

Following reference sections on applicable documentation (Chapter 2) and terms of reference (Chapter 3), the instrument is summarized in Chapter 4. The main algorithm descriptions are found in Chapters 5 and 6, while Chapters 7, 8 and 9 contain notes on Feasibility, Error Analyses and Validation respectively.

1.3 Acknowledgements

The authors would like to thank the following people working on algorithm development, validation, and performing the internal review of this document:

- Pieter Valks (DLR), Sander Slijkhuis (DLR), Claas Köhler (DLR), Diego Loyola (DLR)

2 Applicable and reference documents

2.1 Applicable documents

- [AD1] Sentinel-5P Product Algorithm Laboratory – Statement of Work -;
source: ESA; **ref:** ESA-EOPG-CSCOP-SOW-28; **issue:** 1; **date:** 2019-03-14

2.2 Standard documents

There are no standard documents

2.3 Reference documents

- [RD1] 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
- [RD2] Terms, and symbols in the TROPOMI Algorithm Team;
source: KNMI; **ref:** SN-TROPOMI-KNMI-049; **date:** 2011-09-28
- [RD3] S5P/TROPOMI Clouds ATBD;
source: DLR; **ref:** S5P-L2-DLR-ATBD-400I; **issue:** 2.3; **date:** 20 21-06-25
<https://sentinel.esa.int/documents/247904/2476257/Sentinel-5P-TROPOMI-ATBD-Clouds>
- [RD4] Algorithm theoretical basis document for TROPOMI L01b data processor;
source: KNMI; **ref:** S5P-KNMI-L01B-0009-SD; **issue:** 9.0.0; **date:** 2019-07-19
<https://sentinel.esa.int/documents/247904/2476257/Sentinel-5P-TROPOMI-Level-1B-ATBD>
- [RD5] Input output data specification for TROPOMI L01b data processor;
source: KNMI; **ref:** S5P-KNMI-L01B-0012-SD; **issue:** 10.0.0; **date:** 2019-07-19
<https://sentinel.esa.int/documents/247904/3119978/Sentinel-5P-Level-01B-input-output-data-specification>
- [RD6] 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
<https://sentinel.esa.int/documents/247904/0/S5P-TROPOMI-Static-input-for-Level-2-processors.pdf/f8cc0000-4016-4a9b-b067-44d22a644387>
- [RD7] TROPOMI Instrument and Performance Overview;
source: KNMI; **ref:** S5P-KNMI-L2-0010-RP; **issue:** 0.10.0; **date:** 2014-03-15
- [RD8] Sentinel-5 precursor/TROPOMI Level 2 ProductUser Manual Total Column Water Vapour
source: DLR; **ref:**S5P-L2-DLR-PUM-400A; **issue:**2.2**date:** 2020-06-15

2.3.1 Electronic references

There are no electronic references.

3 Terms, definitions and abbreviated terms

Terms, definitions and abbreviated terms that are used in development program for the TROPOMI/S5P L0-1b data processor are described in [RD1]. Terms, definitions and abbreviated terms that are used in development program for the TROPOMI/S5P L2 data processors are described in [RD2]. Terms, definitions and abbreviated terms that are specific for this document can be found below.

3.1 Terms and definitions

Currently none.

3.2 Acronyms and abbreviations

AK	Averaging Kernel
AMF	Air Mass Factor
ATBD	Algorithm Theoretical Basis Document
AUTH	Aristotle University of Thessaloniki
BDM	Brion-Daumont-Malicet
BRDF	Bi-directional Reflectance Distribution Function
BOA	Bottom of Atmosphere
CAL	Clouds As scattering Layers
CBH	Cloud base height
CBP	Cloud base pressure
CFR	Cloud Fraction
CCI	Climate Change Initiative
COT	Cloud optical thickness
CRB	Clouds as Reflecting Boundaries
CTA	Cloud top albedo
CTH	Cloud top height
CTP	Cloud top pressure
DFS	Degrees of Freedom for Signal
DLR	German Aerospace Center (Deutsches Zentrum für Luft- und Raumfahrt)
DOAS	Differential Optical Absorption Spectroscopy
DU	Dobson Unit
ECV	Essential Climate Variable
ECMWF	European Centre for Medium-range Weather Forecasting
ENVISAT	Environmental Satellite
EO	Earth Observation

ERS-2	European Remote Sensing Satellite-2
EOS-AURA	(NASA's) Earth Observing System Aura
ESA	European Space Agency
ESRIN	European Space Research Institute
ETOP	Earth Topography
FD	Finite Difference
FM	Flight Model
GAW	Global Atmospheric Watch
GDP	GOME Data Processor
GMES	Global Monitoring for Environment and Security
GODFIT	GOme Direct FITting
GOME	Global Ozone Monitoring Experiment
GTOPO	Global TOPOgraphic (Data set)
hPa	Hectopascals
IOP	Inherent Optical Property
IPA	Independent Pixel Approximation
IUP	Institut für Umwelt-Physik
K	Kelvin
KNMI	Koninklijk Nederlands Meteorologisch Instituut
LER	Lambertian Equivalent Reflectivity
LIDORT	LInearized Discrete Ordinate Radiative Transfer
LM	Levenberg-Marquardt
LLM	Labow-Logan-McPeters
LOS	Line-Of-Sight
LUT	Look Up Table
MetOp	Meteorological Operational
ML	McPeters-Labow
MLS	Microwave Limb Sounder
MSG	Meteosat Second Generation
MRTD	Mission Requirements Traceability Document
NASA	National Aeronautics and Space Administration
NetCDF-CF	Network Common Data Format – Climate and Forecast (CF) convention
NDACC	Network for the Detection of Atmospheric Composition Change
NISE	Near-real-time global Ice and Snow Extent

NDSC	Network for the Detection of Stratospheric Change
NOAA	National Oceanic and Atmospheric Administration
NRTI	Near-real-time
OCRA	Optical Cloud Recognition Algorithm
OE	Optimal Estimation
OFFL	Off-line
OMI	Ozone Monitoring Instrument
P-S	Pseudo-Spherical
PTH	Pressure-Temperature-Height
QA	Quality Assurance
QOS	Quadrennial Ozone Symposium
RRS	Rotational Raman Scattering
RT	Radiative Transfer
ROCINN	Retrieval of Cloud Information using Neural Networks
SCIAMACHY	SCanning Imaging Absorption spectroMeter for Atmospheric Cartography
SBUV	Solar Backscatter UltraViolet
S5P	Sentinel 5 Precursor
SGP	SCIAMACHY Ground Processor
SRD	Systems Requirement Document
SZA	Solar Zenith Angle
TO3	Total Ozone
TOA	Top of Atmosphere
TOMS	Total Ozone Mapping Spectrometer
TROPOMI	TROPOspheric Monitoring Instrument
UPAS	Universal Processor for UV/VIS Atmospheric Spectrometers
UV	Ultra Violet
V8	Version 8
VCD	Vertical Column Density
VLIDORT	Vector LInearized Discrete Ordinate Radiative Transfer
WMO	World Meteorological Office
WOUDC	World Ozone and Ultraviolet Data Centre

4 Introduction to the S5P Data Products

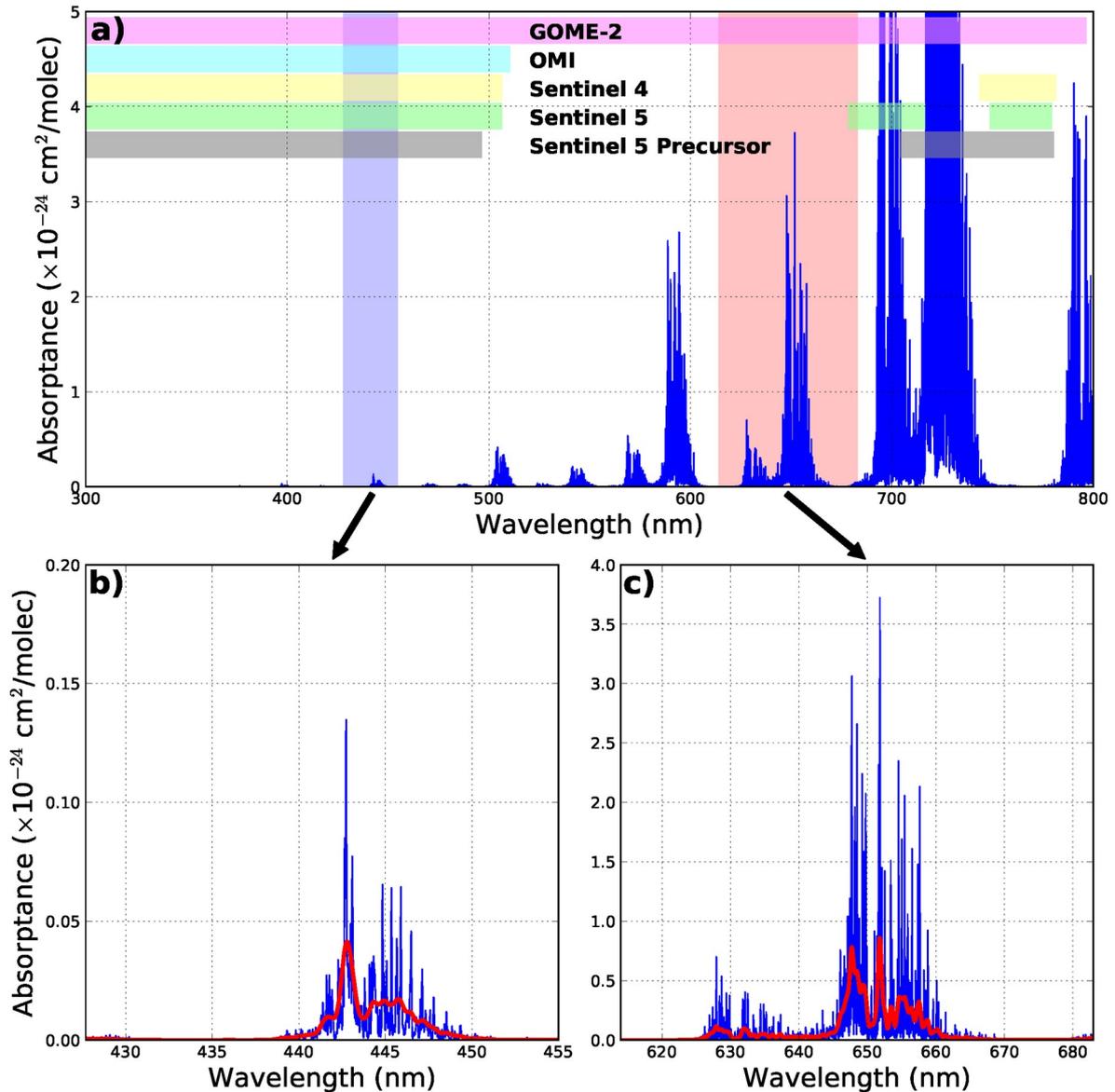


Figure 1: Water vapour absorption cross section. (a) The horizontal bars show the spectral band available to various satellite sensors. The wavelength range used in this study and the operational GOME-2 product are highlighted in blue and red colour, respectively. Zoom in of (a) to the blue and red bands are shown in (b) and (c), respectively. The red curves show the water vapour absorption cross section convoluted with the instrument slit function. Note that the scale of the y axis of each plot is different.

4.1 Heritage

Satellite remote sensing observations are an effective way to monitor the spatio-temporal variations of column amount water vapour on a global scale. High quality water vapour data can be derived from a large number of satellite sensors operating in various wavelength regions (optical, infrared, microwave) [Kaufman and Gao, 1992; Bauer and Schluessel, 1993; Noël et al., 1999, 2004; Li et al., 2006; Wagner et al., 2006; Pougatchev et al., 2009; Wang et al., 2014; Grossi et al., 2015]. Each sensor has its specific advantages and limitations, be it for spatio-temporal resolution, for truly global coverage, for sensitivity, or for the long timelines required for climate monitoring. An extensive overview of satellite measurement of water vapour can be found in Schröder et al. (2018).

Total column water vapour can be retrieved from nadir viewing spectroscopic satellite observations in the visible wavelength range. This kind of satellite observation has long been conducted since the Global Ozone Monitoring Experience (GOME) mission launched in 1995 [Burrows et al., 1999]. Together with other follow up satellite missions, for example, SCanning Imaging Absorption SpectroMeter for Atmospheric CHartography (SCIAMACHY) [Bovensmann et al., 1999], Global Ozone Monitoring Experience 2 (GOME-2) [Callies et al., 2000] and Ozone Monitoring Instrument (OMI) [Levelt et al., 2006], these observations provide a global record of earthshine radiance in the UV and Vis spectral range for more than 25 years. The recent satellite mission TROPOspheric Monitoring Instrument (TROPOMI) [Veefkind et al., 2012] on board the European Space Agency (ESA) Sentinel 5 Precursor (S5P) satellite provides daily global observations of earthshine radiance in the UV and Vis range with much finer spatial resolution (3.5km× 7km) compared to its predecessors. The TROPOMI/S5P and the upcoming Sentinel 5 (S5) missions will provide indispensable global observations of earthshine radiance in the UV and Vis ranges in the next decade. Retrieving TCWV from these observations can provide important independent data sets for climate studies and contribute towards TCWV climate data records [Beirle et al., 2018; Schröder et al., 2018].

TCWV is typically retrieved in the visible red and near infrared (NIR) spectral range [Grossi et al., 2015]. However, the TROPOMI sensor does not cover the red band, it is necessary to develop new water vapour retrieval in the available spectral bands. Most of the spectroscopic satellite borne instruments, e.g., GOME, GOME-2, OMI, TROPOMI, etc, cover the blue spectral band as it is essential for the monitoring of major atmospheric pollutants, i.e., nitrogen dioxide (NO₂) [Richter and Burrows, 2002; Valks et al., 2011; Boersma et al., 2011; Krotkov et al., 2017]. Retrieving TCWV in the blue wavelength band can provide a consistent long time series of climate record from similar type of satellite sensors. Figure 1 shows the water vapour absorption cross section in the UV and Vis bands together with the spectral band available to the current GOME-2, OMI and S5P sensors as well as the forthcoming S4 and S5 instruments. The red shadowed area indicates the spectral range used in the current GOME-2 operational water vapour retrieval. The blue shadowed area denotes the wavelength band used for TCWV retrieval from TROPOMI/S5P. Previous studies have demonstrated the feasibility of retrieving water vapour slant columns and total columns from GOME-2 and OMI satellite observations in the blue band [Wagner et al., 2013]. Based on the similar approach, Wang et al. (2014) has derived TCWV from OMI observations using a priori information from the Goddard Earth Observing System version 5 (GEOS-5) model assimilation products. The baseline retrieval algorithm for TROPOMI/S5P is designed to take into consideration of the requirements for off-line (OFFL) data sets compatible with long-term historical data records.

The DOAS technique is used to retrieve water vapour slant columns from TROPOMI/S5P spectral observations. The DOAS spectral fitting technique has long been used for the spectral retrieval of various species since the GOME/ERS-2 mission in 1995 (GDP Versions 1-4, see for example [Spurr et al., 2005] and [Van Roozendaal et al., 2006]).

4.2 Product requirements

The TROPOMI/S5P TCWV retrieval algorithm has to fulfil the following requirements. Firstly, the algorithm should be feasible to be implemented to forthcoming satellite sensors, such as S4 and S5. Secondly, the retrieval should not rely on input from chemistry transport model (CTM) to avoid propagating model error into the climatological measurement records. Lastly, the retrieval should provide a realistic error estimation as measurement uncertainty is an important parameter for data assimilation and future harmonization of satellite data. Based on the results from previous studies, we have further optimized spectral analysis settings for the TCWV retrieval and developed a statistical analysis approach to optimize the a priori water vapour profile used in the retrieval.

4.3 Overview of the retrieval method

4.3.1 Off-Line (OFFL) algorithm

The OFFL algorithm (S5P_TCWV_OFFL) comprises the DOAS spectral fitting retrieval of water vapour columns and the radiative transfer calculation of air mass factor for the conversion of slant column to vertical column. As the other water vapour retrieval bands, e.g., visible red, near infrared, are not available to TROPOMI/S5P, therefore, the retrieval wavelength band is restricted to the visible band. There are several advantages of retrieving water vapour in the blue band. Firstly, the surface albedo in the blue spectral region is similar over land and ocean and yields better sensitivity in the lowest troposphere and over ocean. Secondly, no correction for spectral saturation effects is necessary due to relatively weak water vapour absorption in the blue spectral band. As a downside, there is no oxygen absorption in the blue spectral band which can be used as reference for the conversion of slant columns to vertical columns. Thus, the conversion of slant columns to vertical columns from observations in the blue band is entirely based on radiative transfer calculations, which sets higher requirements on the accuracy of the geophysical input parameters such as surface albedo and a priori H₂O profile.

4.4 General design considerations

The science behind the prototype TCWV algorithms is well established; simulations of radiative transfer in the atmosphere are based on the widely used VLIDORT models. The radiative transfer model VILDORT has also been used on TROPOMI/S5P as well as other satellite platforms for the retrieval of various species [Theys et al., 2017; De Smedt et al., 2018].

The DOAS spectral fitting algorithm has already been employed to retrieve NO₂ and water vapour columns at the similar wavelength window [Valks et al., 2011; Krotkov et al., 2017; Liu et al., 2019; Wagner et al., 2013; Wang et al., 2014, 2016; Chan et al., 2020]. Level 1b input requirements for both algorithms are the same (calibrated geolocated radiance and irradiance measurements in the visible range of 435-455 nm).

5 This ATBD Algorithm descriptions

5.1 The blue band TCWV retrieval

In this chapter, we describe the algorithm for the retrieval of TCWV from TROPOMI/S5P visible blue observations. The TROPOMI/S5P TCWV retrieval algorithm in the blue spectral range follows the classical differential optical absorption spectroscopy (DOAS) approach, which is a standard spectroscopic method for the retrieval of weakly absorbing trace gases [Platt and Stutz, 2008]. The method consists of two major steps. The first step is the retrieval of water vapour slant columns. The second step is the conversion of the water vapour slant columns to vertical columns. Details of the retrieval algorithm and error estimation are presented in the following.

5.1.1 Water vapour slant column retrieval

Typical absorption spectroscopy describes the attenuation properties of radiation along an optical path by the Beer-Lambert-Bouguer law. For satellite measurements, the equation can be written as the following equation.

$$I(\lambda) = I_0(\lambda) e^{\left(-\varepsilon_M(\lambda) - \varepsilon_R(\lambda) - L \sum_{i=1}^n \sigma_i(\lambda) c_i\right)} R(\lambda) \quad 1$$

$I_0(\lambda)$ refers to the direct sun irradiance spectrum taken at the top of atmosphere (TOA), while $I(\lambda)$ is the earthshine radiance spectrum taken by looking down from space towards the nadir direction measuring sun light reflected by the earth's surface and atmosphere. L represents the effective optical path length from TOA to the earth's surface and reflected from the earth's surface back to the satellite. σ_i denotes the absorption cross section of gas i , and c_i is its average concentration along the effective optical path. ε_M and ε_R are the Mie and Rayleigh extinction integrated along the light path, respectively. $R(\lambda)$ represents the reflectance of the earth. The optical density $\tau(\lambda)$ can then be calculated by taking logarithm of the ratio between $I_0(\lambda)$ and $I(\lambda)$ as shown in following equation.

$$\tau(\lambda) = \left(\frac{I_0(\lambda) R(\lambda)}{I(\lambda)} \right) \quad 2$$

In practice Equation 1 cannot be directly applied for trace gas retrieval, as some of the extinction processes, i.e., Mie and Rayleigh scattering are not quantified. The DOAS method utilizes the fact that atmospheric scattering processes only show broad band spectral characteristics while trace gases exhibit narrow band absorption structures [Platt and Stutz, 2008]. Therefore, the optical density $\tau(\lambda)$ can be separated into narrow (or differential) $\tau_o(\lambda)$ and broad $\tau_b(\lambda)$ band contributions. The broad band contribution τ_b can be approximated by a low order polynomial $p(\lambda)$. The broad band structures in $R(\lambda)$ can also be accommodated by $p(\lambda)$ and narrow band features in $R(\lambda)$ can be included as pseudo cross sections in the spectral fit. Thus, the equation can be rewrite as Equation 3.

$$\tau(\lambda) = \tau'(\lambda) + \tau_b(\lambda) = \sum_{i=1}^n \sigma_i(\lambda) c_i + p(\lambda) \quad 3$$

Characteristic absorption features of different trace gases are then used to determine their concentrations c_i along the effective optical path L .

Slant column densities (SCDs) of water vapour are retrieved from TROPOMI/S5P spectra by applying the DOAS spectral fitting technique. The SCD is defined as the integrated concentration along the optical path from TOA through the atmosphere to the earth's surface and reflected back to the satellite sensor ($L \times c_i$). The DOAS spectral fit is applied to the wavelength range of 435 - 455 nm. The following absorption cross sections are employed in the DOAS fit: water vapour at 293 K from HITEMP database [Rothman et al., 2010] and scaled by Lampel et al. (2015), NO₂ at 220 K [Vandaele et al., 2002], O₃ at 228 K [Brion et al., 1998], O₄ at 293 K [Thalman and Volkamer, 2013], and liquid water at 297 K [Pope and Fry, 1997] as well as a Ring spectrum. The broad band spectral structures caused by broad band absorption of trace gases, instrumental effects, Rayleigh and Mie scattering are removed by including a 4th order polynomial in the spectral fitting. Shift and stretch parameters of radiance spectra are also fitted during the spectral fitting process to compensate for the instability due to small thermal variations of the spectrograph.

Wavelength grids for backscattered radiance and solar irradiances might be slightly different due to the difference in viewing modes. Using the solar data as a baseline, this mismatch is treated by fitting shift and squeeze parameters to the backscattered radiance wavelength grid and re-sampling the measurements. Therefore, the DOAS fit has non-linear dependence on these shift and squeeze parameters, which is usually determined iteratively. In practice, spectral retrieval is performed iteratively to optimize the shift and squeeze parameters until a suitable convergence is reached.

Table 1: Summary of the water vapour slant column retrieval results with different spectral fitting window for TROPOMI measurements taken on 1st July 2018 (orbit 3698 - 3711) over the tropics (30°S - 30°N).

Fitting Window	Length of fitting window	Median SCD (kg m ⁻²)	Median RMS	Reference
430.0-450.0nm	20.0nm	40.01	6.94×10 ⁻⁴	Wagner et al. (2013)
430.0-480.0nm	50.0nm	35.70	7.19×10 ⁻⁴	Wang et al. (2014)
427.7-465.0nm	37.3nm	38.21	6.97×10 ⁻⁴	Wang et al. (2016)
432.0-466.5nm	34.5nm	38.01	6.67×10 ⁻⁴	Wang et al. (2019)
427.7-455.0nm	27.3nm	40.45	7.05×10 ⁻⁴	Chan et al. (2020)
435.0-455.0nm	20.0nm	39.96	6.41×10 ⁻⁴	This work

The spectral fitting window for the retrieval of water vapour slant columns is selected based on sensitivity study of TROPOMI water vapour slant columns with different spectral fitting window. The DOAS fitting range and the water vapour slant column retrieval results are shown in Table 1. We vary the spectral fitting window with other retrieval settings unchanged. These spectral fitting windows for water vapour retrieval cover the featuring water vapour absorption structure at 441 - 448 nm. The median water vapour slant columns and the median root mean squares of spectral fit for measurements taken on 1st July 2018 (orbit 3698 - 3711) over the tropics (30° S - 30° N) are shown in Table 1. The spectral fitting range at 435.0 - 455.0 nm leads to the lowest root mean square and the median SCD is also close to the median value amount all settings. Therefore, the fit window of 435 - 455 nm is chosen as the standard setting in this study. An example of the DOAS retrieval of water SCD from a TROPOMI spectrum taken on 1st July 2018 over the Pacific Ocean is shown in Figure 2.

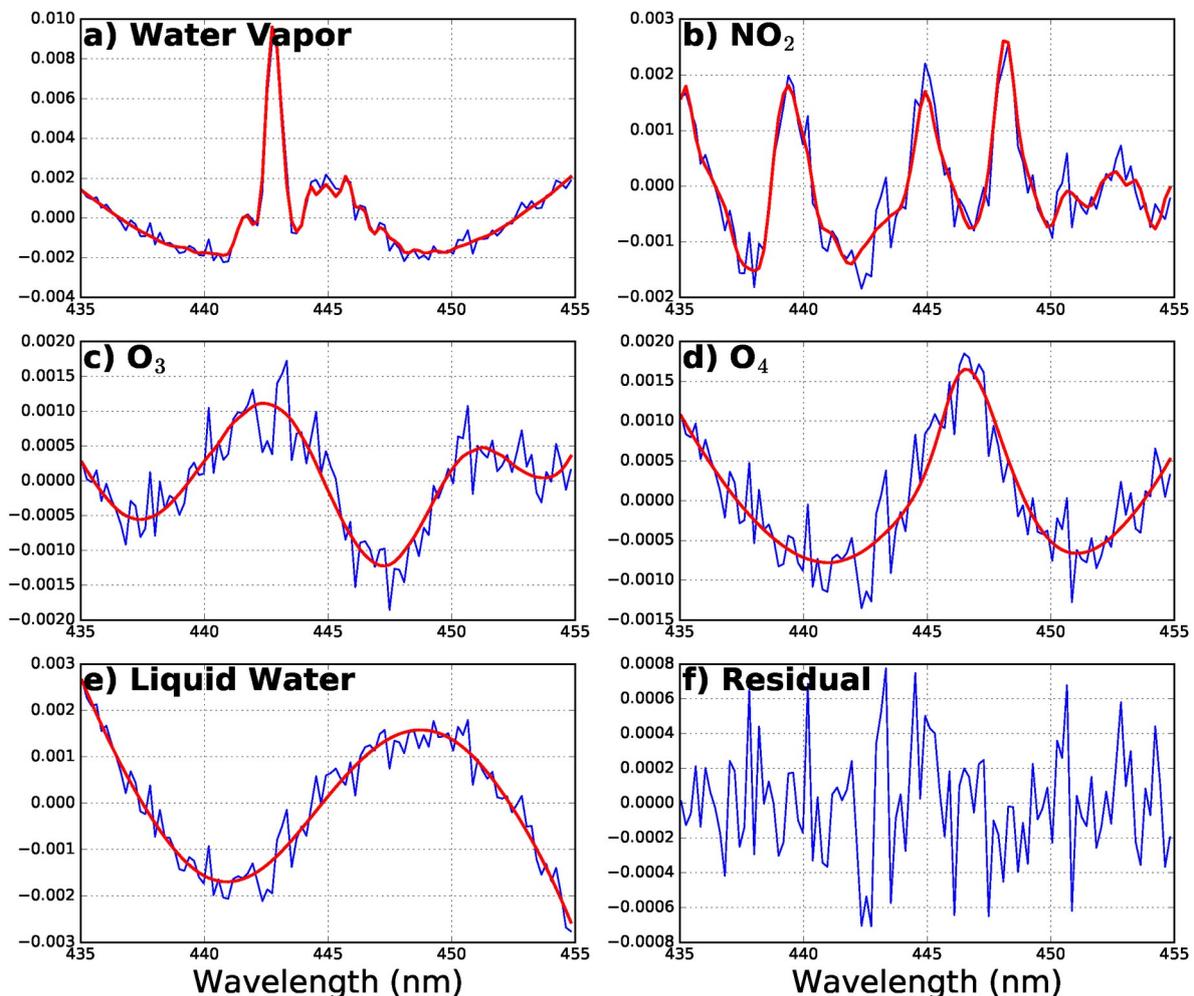


Figure 2: An example of the DOAS retrieval of water vapour slant column for a TROPOMI spectrum taken on 1st July 2018 over the Pacific Ocean. The spectrum was taken with solar zenith angle of 34.6 and viewing zenith angle of 61.1°. The retrieved water vapour slant column is 89.8 kg m⁻².

5.1.2 Air mass factor calculations

The next step of the TCWV retrieval is the conversion of water vapour SCDs to vertical column densities (VCDs). The VCD (or total column) is defined as the vertical integral of water vapour from the surface to the top of atmosphere. The SCD to VCD conversion is accomplished by using the concept of air mass factor (AMF) [Solomon et al., 1987]. As water vapour SCDs are retrieved within a relatively narrow spectral window, we can assume the wavelength dependency of the optical path is negligible. Thus, the AMFs need only be calculate at a representative wavelength. Due to the relatively strong water vapour absorption feature at 442nm, the AMFs are calculated at this wavelength. The AMF can be expressed as Equation 4.

$$AMF = \frac{SCD}{VCD} \quad 4$$

Light travelling in the atmosphere can be scattered by air molecules, aerosols, and clouds, resulting in a complex optical path. To resolve the optical path as well as the box air mass factor (ΔAMF), comprehensive multiple scattering radiative transfer calculations are required. The ΔAMF is defined as the AMF of each individual vertical layer. Typically, the height dependent air mass factor can be decoupled from the vertical distribution of optically thin absorbers [Palmer et al., 2001]. As a result, the AMF can then be calculated from the ΔAMF following Equation 5.

$$AMF = \frac{SCD}{VCD} = \frac{\sum_{l=TOA}^{l=surface} \Delta AMF_l \times \Delta z_l \times c_l}{\sum_{l=TOA}^{l=surface} \Delta z_l \times c_l} \quad 5$$

where Δz_l and c_l are the thickness and the number density of the absorber at layer l , respectively. c_l is taken from the a priori profile. The $\Delta AMFs$ are independent of the vertical distribution of the absorber, but strongly dependent on viewing geometry, solar position, surface albedo and surface altitude.

5.1.3 Box air mass factor look-up table

The ΔAMF can be calculated using a radiative transfer model. To reduce the processing time, $\Delta AMFs$ are pre-calculated with a number of representative observation and solar geometries, surface albedo and surface pressure and stored in a look-up table. In the current version of retrieval algorithm, the ΔAMF look-up table is calculated with the radiative transfer model VLIDORT version 2.7 [Spurr, 2008] at 442nm with an aerosol free US standard atmosphere [Anderson et al., 1986]. The $\Delta AMFs$ for each particular TROPOMI/S5P observation can then be derived by interpolating within the look-up table. Details of the parameterization of the ΔAMF look-up table are shown in Table 2.

Table 2: Parameters in the box air mass factor look-up table.

Parameter (unit)	Symbol	Number of grid points	Grid values
Viewing zenith angle (°)	α	10	0, 10, 20, 30, 40, 50, 60, 65, 70, 75
Solar zenith angle (°)	θ	20	0, 10, 20, 30, 40, 45, 50, 55, 60, 65, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88
Relative azimuth angle (°)	φ	7	0, 30, 60, 90, 120, 150, 180
Surface albedo	A_s	14	0, 0.01, 0.025, 0.05, 0.075, 0.1, 0.15, 0.2, 0.25, 0.3, 0.4, 0.6, 0.8, 1.0
Surface pressure (hPa)	P_s	17	1063.10, 1037.90, 1013.30, 989.28, 965.83, 920.58, 876.98, 834.99, 795.01, 701.21, 616.60, 540.48, 411.05, 308.00, 226.99, 165.79, 121.11
Pressure level (hPa)	P_l	64	1056.77, 1044.17, 1031.72, 1019.41, 1007.26, 995.25, 983.38, 971.66, 960.07, 948.62, 937.31, 926.14, 915.09, 904.18, 887.87, 866.35, 845.39, 824.87, 804.88, 785.15, 765.68, 746.70, 728.18, 710.12, 692.31, 674.73, 657.60, 640.90, 624.63, 608.58, 592.75, 577.34, 562.32, 547.70, 522.83, 488.67, 456.36, 425.80, 396.93, 369.66, 343.94, 319.68, 296.84, 275.34, 245.99, 210.49, 179.89, 153.74, 131.40, 104.80, 76.59, 55.98, 40.98, 30.08, 18.73, 8.86, 4.31, 2.18, 1.14, 0.51, 0.14, 0.03, 0.01, 0.001

In the TCWV retrieval, linear interpolation is applied to the cosine of viewing zenith angle ($\cos \alpha$), cosine of solar zenith angle ($\cos \theta$), relative azimuth angle (φ) and surface albedo (A_s) dimensions, while nearest neighbour interpolation is used for the interpolation in the surface pressure (P_s) dimension. The viewing zenith angle, solar zenith angle and relative azimuth angle are taken from the TROPOMI level 1B product. Surface albedo is derived from TROPOMI observation at the same wavelength band using the full-physics inverse learning machine (FP_ILM) algorithm (see Section 5.1.5). The resulting Δ AMF profile is then linearly interpolated to match the vertical grid of water vapour a priori profile. The interpolated Δ AMFs are used to calculate the AMF following Equation 5.

5.1.4 Dynamic search of a priori water vapour profile

The vertical distribution of water vapour is important for the conversion of slant columns water vapour to vertical columns as expressed in Equation 5. Most of the trace gas retrievals from satellite measurements in the UV and Vis spectral range use vertical profile information from chemistry transport model simulations (e.g., Wang et al., [2014], [2016]; Krotkov et al., [2017]; De Smedt et al., [2018]). Previous studies use a priori profile from GEOS-5 and MERRA-2 model products to retrieve TCWV from OMI observations (Wang et al., 2014, 2016). In this study, we used the statistical analysis of historical profiles a priori information, so that the influences from model simulation are greatly reduced and more suitable for climatological study. We developed an iterative approach to optimize the a priori water

vapour vertical profile used in the satellite retrieval to make the satellite measurements independent from model simulations, and avoid propagating model errors into the measurement. The iterative a priori profile optimization approach is based on the statistical analysis of water vapour vertical distribution over 11 years from 2008 to 2018. Figure 3 shows the statistical analysis of water vapour profiles from ECMWF ERA-Interim reanalysis data [Dee et al., 2011; Berrisford et al., 2011] over a small region of Pacific Ocean (5°S-5°N, 180°W-170°W) in July of 2008-2018. Water vapour profiles are sorted by their total column densities into 8 ranges from 20kg m⁻² up to 60kg m⁻². Colour coded lines indicate the mean profile of each range, while shadowed areas represent the 1 σ standard deviation variation of water vapour mixing ratio. The normalized mean profiles for each range are also indicated in Figure 3i. These profiles are normalized by dividing their total columns and multiplying with the mean total column calculated from all measurements. The analysis result shows that water vapour vertical profile shapes are strongly related to their own column densities. Water vapour profiles with similar total columns show very similar vertical distribution. Water vapour is typically concentrated close to the surface below 800hPa when the total column is small (i.e., less than 30kg m⁻²). It starts to extend to higher altitudes with increasing total column and changes the profile shape. A much larger portion of water vapour is located above 800hPa when the total column is larger than 40 kg m⁻². The small standard deviation of the water vapour mixing ratio profile also indicates that the water vapour profile shape only varies slightly within each range.

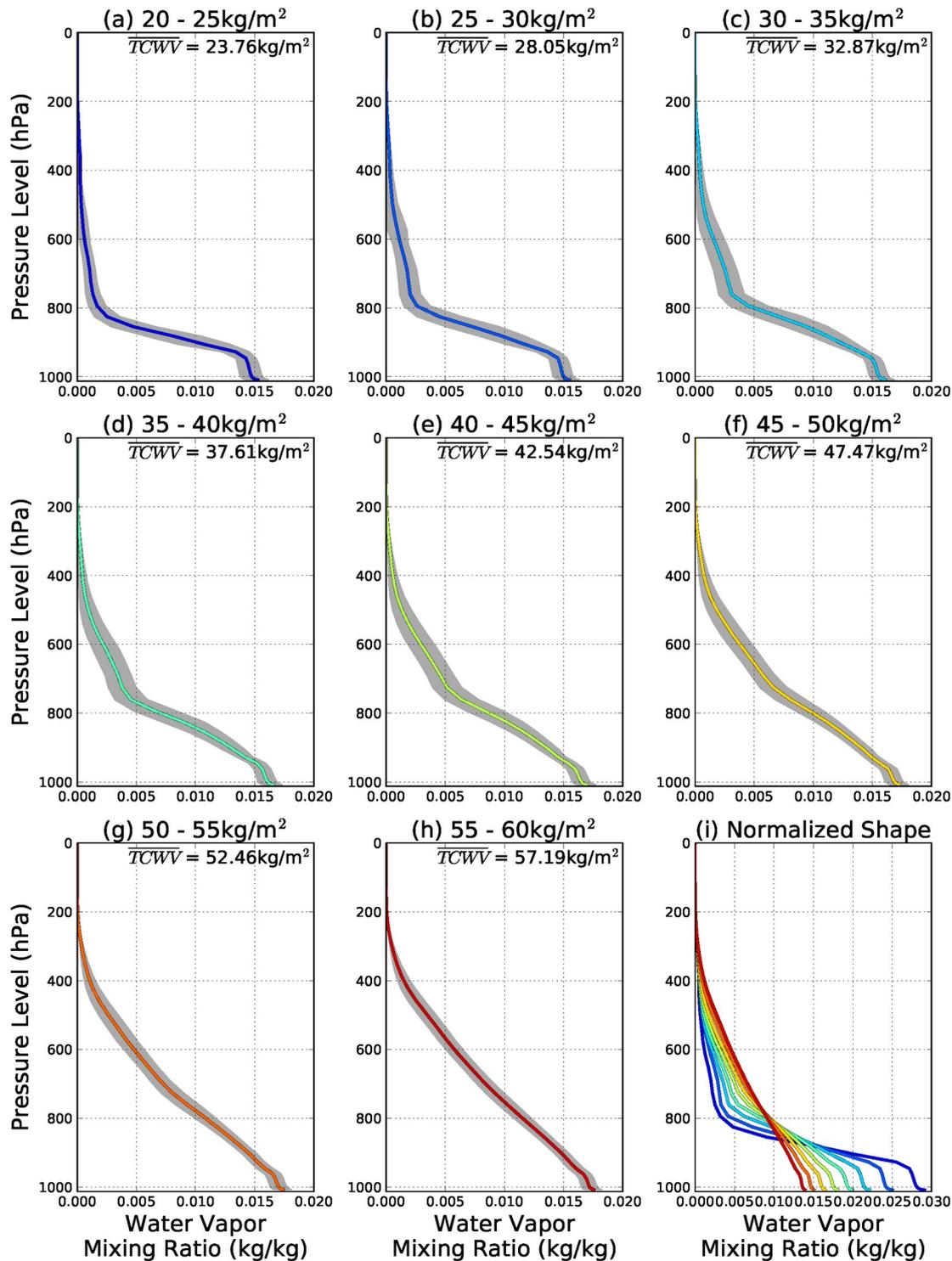


Figure 3

Figure 3: Statistical analysis of water vapour vertical profiles from ECMWF ERA-Interim reanalysis data over a small region of Pacific Ocean (5°S-5°N, 180°W-170°W) in July of 2008 to 2018. Water vapour profiles are sorted by their total column density into 8 ranges, which are (a) 20-25kg m⁻², (b) 25-30 kg m⁻², (c) 30-35 kg m⁻², (d) 35-40 kg m⁻², (e) 40-45 kg m⁻², (f) 45-50 kg m⁻², (g) 50-55 kg m⁻² and (h) 55-60 kg m⁻². Shaded areas indicate the 1 σ standard deviation variation of water vapour mixing ratio. (i) shows the normalized mean water vapour profile shapes of these 8 ranges.

By making use of the characteristic that water vapour profile shapes are strongly correlated to their total columns, we have formulated a water vapour vertical profile shape look-up table for the entire globe with a spatial resolution of 0.75° . Water vapour profiles are sorted into five ranges for each geolocation and for each month of the year. The mean profiles, the total columns, and the standard deviation of total columns for each range are stored in a look-up table. The water vapour vertical profile shape look-up table is interpolated linearly in the spatial dimension to the satellite measurement location for each range. The iterative optimization of a priori water profile begins by using the overall mean profile of the satellite measurement location of the corresponding month. This mean water vapour profile is then used together with the corresponding Δ AMFs to calculate an initial AMF following Equation 5. The water vapour slant column is divided by this initial AMF to retrieve the initial vertical column. The look-up table is then linearly interpolated in the total column dimension to the retrieved initial column to retrieve the corresponding vertical profile shape. The interpolated profile is again used to retrieve the second vertical column. This process repeats until the difference between the input and output water vapour column is less than 1% or the number of iteration reaches the limit. As the retrieval of more than 99% of TROPOMI/S5P measurements stopped within 3 iterations, the limit of maximum number of iteration in the current version of retrieval is set to 5.

5.1.5 Surface Albedo

Surface albedo is an important parameter for the calculation of the air mass factor. The sensitivity of satellite observations, especially to the lower troposphere where most water vapour is resided, is strongly related to surface albedo. Surface albedo used in this study is retrieved from TROPOMI observations using the full-physics inverse learning machine (FP_ILM) algorithm [Loyola et al., 2020]. The FP-ILM algorithm is a machine learning based approach that aims to derive the relationship (inverse function) between the parameter of interest [surface albedo in this case] and measured radiance spectrum (with other atmospheric parameters). Training of the FP_ILM algorithm is driven by a set of synthetic data generated with a radiative transfer model. Synthetic TROPOMI radiance spectra at 435 – 455 nm are simulated using the radiative transfer model VLIDORT version 2.7 [Spurr, 2008] together with the smart sampling technique [Loyola et al., 2016]. For the training, the inputs are DOAS fitted parameters, as well as the solar/satellite viewing geometry and surface pressure. The inversion result is the Geometry-dependent effective Lambertian equivalent reflectivity (GE_LER) and it is retrieved for clear sky observations (cloud fraction <0.05). Compared to the Lambertian equivalent reflectivity (LER) climatology derived from Ozone Monitoring Instrument (OMI) observations [Kleipool et al., 2008], which is being used in most of the operational TROPOMI products, albedo data derived from TROPOMI itself is expected to be better in characterizing the actual surface conditions, especially with regard to temporal variability.

5.1.6 Partially cloudy scene observations

Clouds are treated as opaque Lambertian surfaces in the retrieval algorithm. The treatment of partially cloudy pixels is based on the independent pixel approximation [Martin et al., 2002;

Boersma et al., 2004] where the pixel is separated into two independent parts, one with fully cloud cover and the other one is completely cloud free. Air mass factors are calculated separately for both clear sky and cloudy parts. Cloud information, including cloud fraction (CF), cloud albedo (A_c), and cloud top pressure (P_c), are taken from the TROPOMI/S5P operational cloud product [Loyola et al., 2018, Schuessler et al., 2014].

The cloudy AMF (AMF_{cld}) is calculated from the ΔAMF look-up table by setting the surface pressure to cloud top pressure and replacing the surface albedo with the cloud albedo. It should be noted that the same a priori water vapour profile is assumed in both the cloudy AMF and clear sky AMF (AMF_{clr}) calculations. Following Equation 5, the calculation of SCD for the cloudy scene is insensitive to water vapour below cloud and $\Delta AMFs$ below cloud are 0. On the other hand, VCD is calculated by integrating the water vapour profile from the surface to the top of atmosphere which includes the part below cloud. This ‘invisible’ column below the cloud (also known as ‘ghost column’) is taken from the a priori profile.

AMFs of partially cloudy pixels are calculated as the intensity weighted average of the AMF_{cld} and AMF_{clr} . This weighting is commonly known as intensity weighted cloud fraction (CF_{iw}) which is defined by Equation 6.

$$CF_{iw} = \frac{CFR \times I_{cld}}{CFR \times I_{cld} + (1 - CFR) \times I_{clr}} \quad 6$$

where I_{cld} and I_{clr} represents the radiance intensity for the cloudy and clear sky scenes, respectively. The radiance intensities are pre-calculated using radiative transfer model VLIDORT at 442nm for a number of representative observation and solar geometries, surface albedo, and surface pressure and stored in a look-up table. The settings of the intensity look-up table are the same as the ΔAMF look-up table but without the pressure level dimension. The AMF can then be calculated following Equation 7.

$$AMF = AMF_{cld} \times CF_{iw} + AMF_{clr} \times (1 - CF_{iw}) \quad 7$$

The resulting AMFs are used to divide the measured slant columns to convert water vapour slant columns to vertical columns. This AMF is used for the iterative optimization of a priori profile of partially cloudy pixels.

5.1.7 Aerosol

The presence of aerosols affects the radiative transfer in the atmosphere and may influence the retrieval of surface properties, cloud and atmospheric water vapour [Bhatia et al., 2015, 2018]. As the aerosol properties, e.g., extinction profile, single scattering albedo, asymmetry parameter, etc., are unknown, there is no general and easy solution to explicitly account for aerosols in the retrieval. On the other hand, it is very difficult to separate cloud and aerosol in the cloud retrieval due to their similarity in optical properties. As a result, the aerosol effect is already implicitly considered in the cloud product [Boersma et al., 2004, 2011]. Therefore, no additional treatment of aerosol is applied in the water vapour retrieval algorithm.

6 Input-Output description

Detailed info in PFS [ref].

6.1 S5P TCWV product description and size

The S5P TCCWV product will be provided in netCDF-CF. The following information will be included for each ground pixel:

- measurement time and geolocation, taken from the L1b product
- total ozone, error, averaging kernel
- climatology and other relevant parameters used in the retrieval
- fit results (RMS, etc.)

Table 3 presents detailed lists of the output fields that are required in TCWV level-2 files for S5P_TCWV_OFFL. Additional output parameters such as geolocation, input data etc are also included in the L2 product and specified in the corresponding PUM [RD8]

Table 3: List of output fields in TCWV level-2 product.

Name/Data	Symbol	Unit	Description	Data type per pixel	Dimensions
Number of measurements	N	---	Number of measurements included in the file. $N = n_{\text{Along}} \times n_{\text{Across}}$	Integer	1
Orbit number	n_o	---	Satellite orbit number	integer	1
Time	---	---	Date and time of measurement [YYMMDDHHMMSS.MS]	Character	1
Latitudes	lat	degree	Latitudes of the pixel center and corners	float	5
Longitudes	lon	degree	Longitudes of the pixel center and corners	float	5
SZA	θ_0	degree	Solar zenith angle at pixel center	float	1
RAA	α	degree	Relative azimuth angle at pixel center	float	1
SCD	SCD	kg m ⁻²	Slant column density	float	1
VCD	VCD	kg m ⁻²	Vertical column density	float	1
Clear sky AMF	AMF_{clr}	---	Clear sky air mass factor	float	1
Cloudy AMF	AMF_{cld}	---	Cloudy air mass factor	float	1

CFR	f_c	---	Cloud fraction	float	1
CTP/CTH	p_c/Z_c	Pa	Cloud top pressure	float	1
CTA/COT	A_c / τ_c	---	Cloud top albedo / optical thickness	float	1
Albedo	A_s	---	Climatological surface albedo	float	1
Surface pressure	p_s	Pa	Effective surface pressure of the satellite pixel	float	1
fitted_root_mean_square	RMS	---	residual of the DOAS fit	float	1
SCD random error	σN_{s_rand}	kg m-2	N_s random error	float	1
Pressure grid	p	Pa	Atmospheric pressure grid	float	15
Averaging kernels	AK	---	Layer averaging kernels	float	14
Water vapour profile	n_a	kg m-2	a-priori partial water vapor column profile	float	14
QA_value	qa	---	Data quality assurance value	float	1

6.2 Auxiliary information needs

All dynamic and static auxiliary data are described in Table 4 and Table 5. The needs are similar for both algorithms.

Table 4: Dynamic auxiliary information needs in theTCWV retrieval algorithms

Name/Data	Symbol	Unit	Source	Pre-process needs	Backup if not available
S5P level 1B Earth radiance	I	$W m^{-2}nm^{-1}sr^{-1}$	S5P L1b product		No retrieval
S5P level 1B sun irradiance	I_0	$W m^{-2}nm^{-1}$	S5P L1b product	Wavelength recalibrated using a high-resolution reference solar spectrum	Use previous measurement
Cloud fraction	f_c	---	S5P L2 operational cloud product	---	No retrieval
Cloud top height	p_c	Pa	S5P L2 operational cloud product	---	No retrieval
Cloud top albedo/optical thickness	A_c/τ_c	---	S5P L2 operational cloud product	---	No retrieval
Snow/ice flag	---	---	Near real-time global Ice and Snow Extent (NISE) data from NASA.	---	Use snow/ice climatology

Table 5: Static auxiliary information needs in the TCWV algorithms

Name/Data	Symbol	Unit	Source	Pre-process needs	Comments
Instrument slit function	SF	---	Slit function provided by wavelength/detector	---	---
High-resolution reference solar spectrum	E_s	$W\ m^{-2}nm^{-1}$	Chance and Kurucz [2010]	---	---
Water vapour absorption cross-section	σ_{O_3}	$cm^2\ molec.^{-1}$	Rothman et al. [2010]; Lampel et al. [2015]	Convolution at the instrumental spectral resolution using the provided slit function	---
Ring cross-section	σ_{ring}	---	Generated internally.	A high-resolution reference solar spectrum and the instrument slit function are needed to generate the data set.	---
A-priori water vapour vertical profile shapes	n_a	$kg\ m^{-2}$		---	---
Digital elevation map	Z_s	m	ECMWF same DEM for all L2 products	---	---

6.3 Level 1 information needs

In the heritage algorithms for GOME, the GOME backscattered radiance and solar irradiance were generated by the GDP Level 0-to-1b extractor [Slijkhuis et al., 2004]. In addition, Level-1 wavelength calibration was improved selectively through application of window-dependent pre-shifts to parts of the solar spectrum [Van Roozendaal et al., 2006].

The main Level 1 measurement data sets for both S5P TCWV algorithms are the Level 1b geolocated and calibrated backscatter UV Earthshine radiances and the Level 1b calibrated solar irradiances in for wavelengths in the range 435-455 nm. These measurements are provided in the TROPOMI UVN band 4 (400-495 nm) radiance and UVN irradiance products. Unless otherwise specified, Level 1b radiances are in units of [Photon/nm/sr/cm²]. Radiances and irradiances are accompanied by error quantities in the same units. Wavelengths are calibrated values in [nm]. To ensure an accurate wavelength registration, a recalibration procedure, based on a cross-correlation with a high-resolution solar spectrum (e.g. [Chance and Kurucz, 2010]), is applied to both the radiance and irradiance measurement spectra [Van Roozendaal et al., 2006].

The radiance data is assumed corrected for polarization (TROPOMI uses a scrambler to remove polarization at the instrument level). The Level 1b product also comes with full geolocation information - solar and viewing zenith and azimuth angles at the bottom of the atmosphere (BOA) in addition to those at TOA and at the spacecraft.

The cloud products from OCRA and ROCINN require additional Level 1b data. OCRA needs broad-band integrated measurement data (radiance/irradiance) from the TROPOMI, UVIS detector bands 3 and 4 (310-405 nm and 405-500 nm) and in the future NIR detector band 6 (725-775 nm) will be included. The ROCINN algorithm needs measurements in and around the oxygen A-band (TROPOMI NIR detectorband6). For more details on the cloud algorithm Level 1b requirements, refer to the accompanying ATBD [RD3].

7 Error estimation

The error of the TCWV is composed of many sources. Major sources of error can be divided into two parts: one is related to the measurement itself and the other is related to the uncertainties of assumptions in the retrieval. The uncertainty of the TCWV can be derived analytically through error propagation. As the retrieval of TCWV is separated into two major steps, slant column retrieval and AMF calculation, the error estimation also follows these two steps. The uncertainty of TCWV can be expressed as Equation 8.

$$\sigma_{vcd}^2 = V C D^2 \times \left(\left(\frac{\sigma_{scd}}{SCD} \right)^2 + \left(\frac{\sigma_{amf}}{AMF} \right)^2 \right) \quad 8$$

where σ_{vcd} , σ_{scd} and σ_{amf} are the uncertainty of TCWV, the error of water vapour slant column, and air mass factor uncertainty, respectively. Details of the estimation of the water vapour slant column uncertainty and air mass factor error are presented in the following.

7.1 Slant column error

The uncertainties of water vapour slant column are mainly attributed to the instrument noise, instrument characteristics, and the uncertainties related to the DOAS retrieval of slant column. Instrument noise is expected to cause random error and this error can be quantified by analyzing the DOAS fit residual [Stutz and Platt, 1996]. Other sources of error, related to the instrument, are the uncertainties of instrument slit function, incomplete removal of stray light, and wavelength calibration uncertainties. In addition, we have uncertainties of absorption cross sections and temperature dependency of the absorption cross sections. The contributions of systematic errors to the slant column uncertainties are estimated through sensitivity tests with absorption cross section with different effective temperature and different assumptions of instrument slit function shape. We estimated the systematic error of the slant column is about 3%. The total error of the slant column can be calculated following Equation 9.

$$\sigma_{vcd}^2 = \sigma_{scd_r}^2 + (0.03 + SCD)^2 \quad 9$$

Where $\sigma_{scd_r}^2$ is the random error estimated by analyzing the DOAS fit residual.

7.2 Clear sky air mass factor error

The uncertainty of the AMF is mainly related to the uncertainties of each input parameter used in the AMF calculation. These input parameters include the solar and viewing geometries, surface albedo, surface pressure, and water vapour vertical profile. The solar and viewing geometries are well calibrated and their errors are mainly related to the interpolation of the box AMF look-up table. These uncertainties are negligible compared to other sources of error. The contribution to the AMF uncertainty of the remaining sources of error can be estimated by the AMF sensitivity (or Jacobian) with respect to each parameter [Boersma et al., 2004]. The Jacobian is derived from the box air mass factor look-up table using the finite difference method.

In this study, Surface albedo is taken from the climatology monthly minimum Lambertian equivalent reflector (LER) product [Kleipool et al., 2008] at 442 nm derived from Ozone Monitoring Instrument (OMI) satellite observations. The uncertainty of surface albedo (A_s) is assumed to be the difference between albedo derived at 425nm and 442nm to account for the small variation of albedo within the spectral fitting window. Information of surface pressure (P_s) is taken from a digital elevation model (DEM) which is considered rather accurate and the uncertainty of surface pressure is mostly related to the variation within the TROPOMI/S5P footprint. We have analyzed this variation of surface pressure and find it is mostly (95%) below 10hPa. Therefore, we set the uncertainty of P_s to 10hPa.

The error related to a priori vertical distribution of water vapour is determined by using the a priori water vapour from the last iteration plus 1σ standard deviation which is also included in the look-up table. This new profile is then used to calculate the corresponding AMF. The difference between this AMF and the original AMF is taken as the uncertainty from the a priori profile. The uncertainty of the water vapour slant column can potentially affect the dynamic search of the a priori profile. As the slant column uncertainty can be much higher than the slant column itself over dry areas in the upper latitudes, considering this effect in the vertical profile uncertainty estimation would further amplify the uncertainty and results in unrealistic high error. Therefore, we assume this effect is well covered by the vertical profile variation and accounted in the vertical profile uncertainty estimation. The error of the clear sky AMF can be calculated following Equation 10.

$$\sigma_{amf_{clr}}^2 = \left(\frac{\partial AMF_{clr}}{\partial A_s} \sigma_{A_s} \right)^2 + \left(\frac{\partial AMF_{clr}}{\partial P_s} \sigma_{P_s} \right)^2 + \left(\frac{\partial AMF_{clr}}{\partial c_l} \sigma_{c_l} \right)^2 \quad 10$$

Where $\sigma_{amf_{clr}}$, σ_{A_s} , σ_{P_s} and σ_{c_l} are the uncertainty of the clear sky AMF, surface albedo, surface pressure, and water vapor profile, respectively. This error is in general <5% for TROPOMI/S5P measurements over the tropics (30°S-30°N).

7.3 Cloudy sky air mass factor error

The calculation of the uncertainty of the cloudy AMF is similar to the one used for the clear sky AMF, with surface albedo and surface pressure uncertainties replaced by cloud albedo and cloud top pressure errors. In this study, cloud top pressure error is assumed to be 50hPa [Theys et al., 2017; De Smedt et al., 2018]. Previous studies show that the error of cloud albedo is compensated by the corresponding error of cloud fraction and results a negligible net effect on trace gas retrieval [Van Roozendaal et al., 2006; Lutz et al., 2016]. Therefore, we assumed a cloud albedo uncertainty of 0.02 and intensity weighted cloud fraction uncertainty of 0.02. The combined effect of the assumed cloud albedo and cloud fraction uncertainties on water vapor retrieval is comparable to the assumption with just cloud fraction error of 0.05 [Theys et al., 2017; De Smedt et al., 2018]. The error of the cloudy AMF can be express as Equation 11.

$$\sigma_{amf_{cld}}^2 = \left(\frac{\partial AMF_{cld}}{\partial A_c} \sigma_{A_c} \right)^2 + \left(\frac{\partial AMF_{cld}}{\partial P_c} \sigma_{P_c} \right)^2 + \left(\frac{\partial AMF_{cld}}{\partial c_l} \sigma_{c_l} \right)^2 \quad 11$$

where $\sigma_{amf_{cld}}$, σ_{A_c} , σ_{P_c} and σ_{c_i} are the uncertainty of the cloudy AMF, cloud albedo, cloud top pressure, and water vapour profile, respectively. The error of the cloudy AMF ($\sigma_{amf_{cld}}$) in general varies from 25% (25th percentile) to 40% (75th percentile) for TROPOMI/S5P measurements over the tropics (30°S-30°N).

7.4 Cloudy sky air mass factor error

Following Equation 7, the uncertainty of the total AMF can be derived from the clear sky and cloudy AMFs through error propagation. The error of the total AMF can be calculated following Equation 12.

$$\sigma_{amf}^2 = (AMF_{cld} \times CF_{iw})^2 \times \left(\left(\frac{\sigma_{amf_{cld}}}{AMF_{cld}} \right)^2 + \left(\frac{\sigma_{cf_{iw}}}{CF_{iw}} \right)^2 \right) + (AMF_{cld} \times (1 - CF_{iw}))^2 \times \left(\frac{\sigma_{amf_{clr}}}{AMF_{clr}} \right)^2 \quad 12$$

where $\sigma_{cf_{iw}}$ is the uncertainty of intensity weighted cloud fraction which is assumed to be 0.02 in the retrieval. The uncertainty of AMF (σ_{amf}) for TROPOMI/S5P measurements over the tropics (30°S-30°N) varies in a range of 6-22% (25th and 75th percentile), while the error reduces to ~6% if the measurements are filtered for intensity weighted cloud fraction below 0.5. The uncertainty of AMF only shows a small latitudinal dependency on surface properties (albedo) and cloud patterns, observation and solar geometries. When all measurements are considered, the uncertainty of AMF varies from 8% (25th percentile) to 24% (75th percentile) with median value of 16% while the mean error remains at ~6% for measurements with intensity weighted cloud fraction below 0.5.

7.5 Total error

Combining the slant column density error with the AMF error, the error of TCWV can then be calculated following Equation 8. The error of TCWV of TROPOMI/S5P measurements over the tropics (30°S-30°N) is on average about 19% under clear sky conditions (intensity weighted cloud fraction <0.5).

7.6 Averaging kernel

The TCWV product provides the corresponding averaging kernel for each ground pixel. The averaging kernel is the altitude-dependent AMF divided by the total air-mass factor. The averaging kernel is important for users who wish to minimise the discrepancies between the profile assumed in the retrieval and their application of interest, for example for validation, data assimilation, or model comparison to a model.

7.7 QA value

Quality assurance (QA) value is provided in the TCWV product. The QA value is determined based on observation geometry, measurement noise, cloud information and air mass factor. The QA value varies between 0 (bad data) and 1 (good data). It is recommended that users not to use data with QA value below 0.5.

8 Validation

8.1 Data sets used for validation

8.1.1 GOME-2 TCWV product

The GOME-2 water vapour product [Grossi et al., 2015] is used as reference to validate the TROPOMI TCWV data set. Data from both GOME-2 instruments on board the MetOp-A and MetOp-B satellites are used. The MetOp satellites orbit at an altitude of ~820km on sun-synchronous orbits with a repeat cycle of 29 days (412 orbits) and a local equator overpass time of 09:30 LT (local time) on the descending node. The spatial resolution of the GOME-2 instrument on board the MetOp-A satellite (GOME-2A) is 40km (across-track) × 40km (alongtrack) while the spatial resolution the GOME-2 instrument on board the MetOp-B satellite (GOME-2B) is 80km (across-track) × 40km (alongtrack).

The GOME-2 water vapour product is processed with GOME Data Processor (GDP) version 4.8 at the German Aerospace Center (DLR) within the framework of EUMETSAT's Satellite Application Facility on Atmospheric Composition Monitoring (AC-SAF). Slant columns water vapour are retrieved in the visible red wavelength range of 614-683nm. The water vapour slant columns are then converted to vertical columns using air mass factors derived from the oxygen slant columns measured in the same wavelength band. The GOME-2 water vapour product has been validated intensively by radiosonde and Global Positioning System (GPS) measurements [Antón et al., 2015; Román et al., 2015; Kalakoski et al., 2016; Vaquero-Martínez et al., 2018]. The GOME-2 water vapour product has been reported to significantly underestimate TCWV over central Africa and India, while a wet bias is found over oceans in the tropics during summer of the northern hemisphere [Grossi et al., 2015]. Compared to radiosonde and GPS data, the GOME-2 product has in general a dry bias of 3-11% [Antón et al., 2015; Román et al., 2015; Kalakoski et al., 2016; Vaquero-Martínez et al., 2018].

8.1.2 MODIS TCWV product

The MODerate resolution Imaging Spectroradiometer (MODIS) instruments are passive nadir viewing imaging sensors [Salomonson et al., 1989; King et al., 1992] on board the Earth Observing System's (EOS) Terra and Aqua satellites. The Terra satellite orbits on sun-synchronous orbit with a local equator overpass time of 13:30 LT (local time) on descending node, while the local equator overpass time for the Aqua satellite is 10:30 LT on ascending node. MODIS measures earthshine radiance at 36 discrete wavelength bands from 0.4µm up to 14.4µm with various spatial resolutions, providing global observation every 1-2 days. Columnar water vapor content can be derived from MODIS observations in the near infrared (NIR) from 865-1240nm. The inversion of water vapour columns is based on the attenuation of radiation through the atmosphere. A more detailed description of the MODIS water vapour retrieval algorithm can be found in Kaufman and Gao (1992); Gao and Kaufman (2003). Due to the similar overpass time, water vapour product derived from the MODIS instruments on board the Terra satellite is used to validate the TROPOMI TCWV data set. Monthly level 3 data product with spatial resolution of 0.1°×0.1° is used. The MODIS water vapour product in general shows a dry bias of 3-13kg m⁻² (Liu et al., 2006; Prasad and Singh, 2009).

8.1.3 SSMIS TCWV product

Another data set used to validate the new TROPOMI water vapour retrieval is the atmospheric water vapour product derived from the Special Sensor Microwave Imager Sounder (SSMIS) on board the United States Air Force Defense Meteorological Satellite Program (DMSP) F16 polar orbiting satellite. The F16 satellite orbit at an altitude of ~848km on sun-synchronous orbits with local equator crossing time of 18:20 LT on the ascending and 06:20 LT on descending node. The SSMIS instrument on board the F16 satellite is in operation since 2005, providing long term climate records of wind speed, cloud liquid water, atmospheric water vapour and rainfall rate during both day and night time [Wentz, 2015]. SSMIS water vapour data are processed by Remote Sensing Systems with funding from the NASA MEaSUREs Program and the NASA Earth Science Physical Oceanography Program. The retrieval of water vapour columns is based on the radiative transfer calculation of brightness temperature over oceans [Wentz, 1997]. Monthly averaged atmospheric water vapour data from SSMIS is used to validate the TROPOMI observations of total column water vapour.

8.2 Validation results

In this section, we present a validation study of TCWV retrieved from TROPOMI observations. TROPOMI TCWV is validated by comparing to GOME-2, MODIS and SSMIS satellite observations.

8.2.1 GOME-2 comparison

Figure 4a and d show the monthly averaged TROPOMI TCWV for July 2018 and January 2019, respectively. Monthly averaged GOME-2 TCWV for July 2018 and January 2019 are shown in Figure 4b and e, respectively. Measurements in July and January are chosen as representative examples for winter and summer. Both TROPOMI and GOME-2 data sets are cloud filtered for intensity weighted cloud fraction smaller than 0.5. Considering the differences in spatial resolution and spatial coverage between TROPOMI and GOME-2, we have regridded both data sets with spatial resolution of $0.5^{\circ} \times 0.5^{\circ}$ for comparison. Missing data over the Tibet Plateau and Andes Mountains in the GOME-2 TCWV product are due to invalid retrieval over high altitude areas. Both TROPOMI and GOME-2 measurements show very similar spatial patterns of TCWV with higher columns over the tropics and lower values at upper latitudes. Compared to GOME-2, TROPOMI in general shows slightly lower columns over ocean in the tropics, especially over the West Pacific Ocean during summer. On the other hand, TROPOMI measures higher TCWV over eastern part of India, Bangladesh and Myanmar. Previous study reported that the GOME-2 TCWV product is underestimating water vapour columns over land and overestimating over oceans in the tropics [Grossi et al., 2015]. The discrepancies between the two data sets indicate that TROPOMI measurements perform better over these areas.

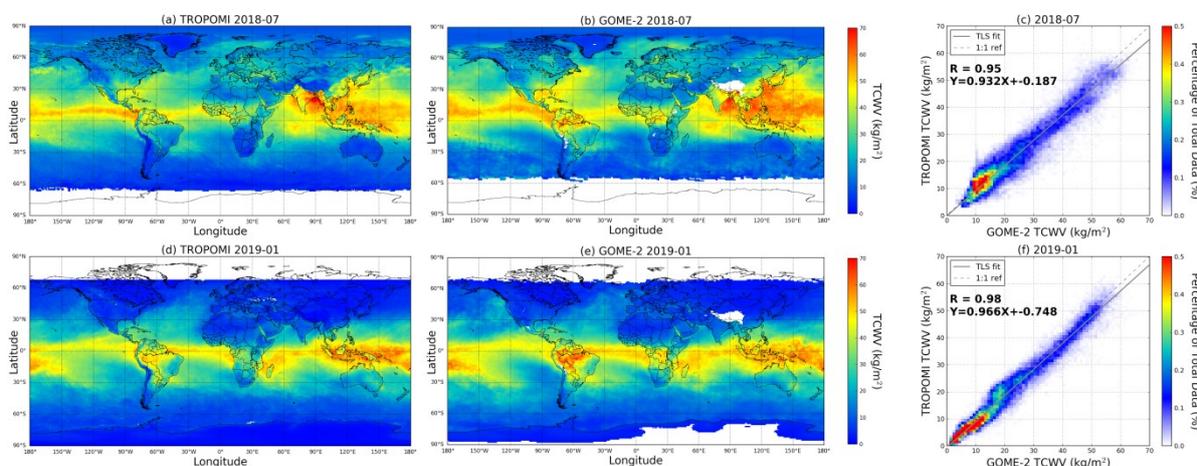


Figure 4: Monthly averaged TCWV retrieved from (a) TROPOMI observations in July 2018, (b) GOME-2 observations in July 2018, (d) TROPOMI observations in January 2019 and (e) GOME-2 observations in January 2019. The scatter plots of TROPOMI and GOME-2 TCWV for July 2018 and January 2019 are shown in (c) and (f), respectively.

The scatter plots of TROPOMI and GOME-2 TCWV for July 2018 and January 2019 are shown in Fig 4c and f, respectively. The two data sets show very good correlation with Pearson correlation coefficients (R) ranging from 0.95 up to 0.98. The correlation is slightly better during winter (January). The slope of the total least squares regression line varies from 0.91 (summer) to 0.94 (winter) with intercept ranging from -0.5 kg m^{-2} to 0.3 kg m^{-2} . Regression line with a slope of slightly less than 1 indicated a small dry bias of TROPOMI observations compared to GOME-2 data. Negative bias mainly occurs over ocean in the tropics where GOME-2 measurements were reported overestimating the TCWV [Grossi et al., 2015].

8.2.2 MODIS comparison

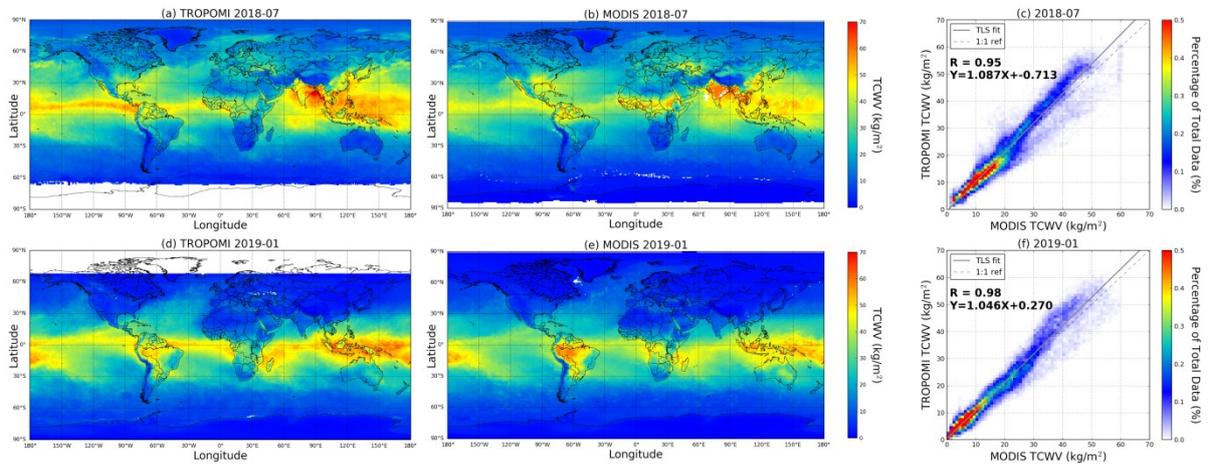


Figure 5: Monthly averaged TCWV retrieved from (a) TROPOMI observations in July 2018, (b) MODIS observations in July 2018, (d) TROPOMI observations in January 2019 and (e) MODIS observations in January 2019. The scatter plots of TROPOMI and MODIS TCWV for July 2018 and January 2019 are shown in (c) and (f), respectively.

8.2.3 SSMIS comparison

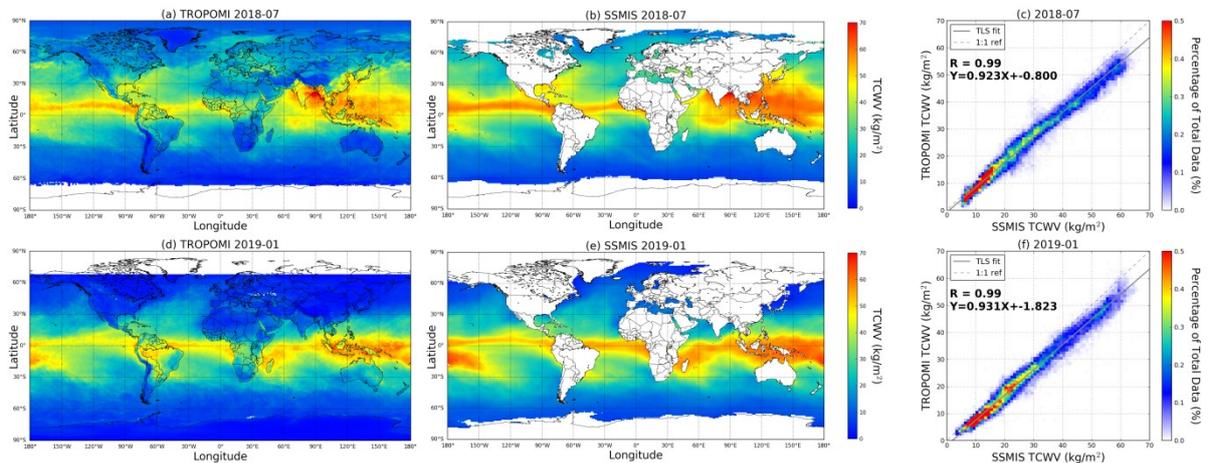


Figure 6: Monthly averaged TCWV retrieved from (a) TROPOMI observations in July 2018, (b) SSMIS observations in July 2018, (d) TROPOMI observations in January 2019 and (e) SSMIS observations in January 2019. The scatter plots of TROPOMI and SSMIS TCWV for July 2018 and January 2019 are shown in (c) and (f), respectively.

9 Conclusions

The prototype algorithms for the generation of the operational S5P total column water vapour products are presented in this ATBD. The new retrieval algorithm makes use of the water vapour absorption characteristic in the visible blue band (435-455nm) to retrieve total column water vapour. The retrieval consists of two major parts, (1) the spectral of water vapour slant column and (2) conversion of slat column to vertical column using air mass factor. For the AMF calculation part, a dynamical search approach has been developed to select the most appropriate a priori profile for the retrieval of water vapour column. The major advantage of the new approach is that it does not rely on a priori information from a chemistry transport model. This improvement makes the satellite product independent from model simulations and avoids model errors propagating to the measurement, making the data more suitable for climate studies.

TCWV data retrieved from TROPOMI/S5P is compared to GOME-2 operational production for validation. The result shows very good correlation with Pearson correlation coefficients (R) ranging from 0.95 up to 0.98. Compared to GOME-2 data, TROPOMI/S5P observations of TCWV show a small dry bias over oceans in the tropics, where GOME-2 measurements were reported overestimating the TCWV. The results indicated the TCWV retrieval is stable and the data quality fulfils the requirements.

The algorithms for the S5P total column products meet the accuracy requirements formulated for this mission. The products are provided in NetCDF and contain water vapour column, and error information including averaging kernels.

10 References

- Anderson, G. P., Clough, S. A., Kneizys, F., Chetwynd, J. H., and Shettle, E. P.: AFGL atmospheric constituent profiles (0.120 km), Tech. rep., AIR FORCE GEOPHYSICS LAB HANSCOM AFB MA, 1986.
- Antón, M., Loyola, D., Román, R., and Vömel, H.: Validation of GOME-2/MetOp-A total water vapour column using reference radiosonde data from the GRUAN network, *Atmospheric Measurement Techniques*, 8, 1135–1145, <https://doi.org/10.5194/amt-8-1135-2015>, 2015.
- Bauer, P. and Schluessel, P.: Rainfall, total water, ice water, and water vapor over sea from polarized microwave simulations and Special Sensor Microwave/Imager data, *Journal of Geophysical Research: Atmospheres*, 98, 20 737–20 759, <https://doi.org/10.1029/93JD01577>, 1993.
- Beirle, S., Lampel, J., Wang, Y., Mies, K., Dörner, S., Grossi, M., Loyola, D., Dehn, A., Danielczok, A., Schröder, M., and Wagner, T.: The ESA GOME-Evolution “Climate” water vapor product: a homogenized time series of H₂O columns from GOME, SCIAMACHY, and GOME-2, *Earth System Science Data*, 10, 449–468, <https://doi.org/10.5194/essd-10-449-2018>, 2018.
- Bellomo, K., Clement, A., Mauritsen, T., Ri, G., and Stevens, B.: Simulating the Role of Subtropical Stratocumulus Clouds in Driving Pacific Climate Variability, *Journal of Climate*, 27, 5119–5131, <https://doi.org/10.1175/JCLI-D-13-00548.1>, 2014.
- Berrisford, P., Dee, D., Poli, P., Brugge, R., Fielding, M., Fuentes, M., Kållberg, P., Kobayashi, S., Uppala, S., and Simmons, A.: The ERA-Interim archive Version 2.0, p. 23, <https://www.ecmwf.int/node/8174>, 2011.
- Bhatia, N., Tolpekin, V. A., Reusen, I., Sterckx, S., Biesemans, J., and Stein, A.: Sensitivity of Reflectance to Water Vapor and Aerosol Optical Thickness, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 8, 3199–3208, <https://doi.org/10.1109/JSTARS.2015.2425954>, 2015.
- Bhatia, N., Iordache, M.-D., Stein, A., Reusen, I., and Tolpekin, V. A.: Propagation of uncertainty in atmospheric parameters to hyperspectral unmixing, *Remote Sensing of Environment*, 204, 472 – 484, <https://doi.org/https://doi.org/10.1016/j.rse.2017.10.008>, 2018.
- Boersma, K. F., Eskes, H. J., and Brinksma, E. J.: Error analysis for tropospheric NO₂ retrieval from space, *Journal of Geophysical Research: Atmospheres*, 109, <https://doi.org/10.1029/2003JD003962>, 2004.
- 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.: An improved tropospheric NO₂ column retrieval algorithm for the Ozone Monitoring Instrument, *Atmospheric Measurement Techniques*, 4, 1905–1928, <https://doi.org/10.5194/amt-4-1905-2011>, 2011.
- Boucher, O., Randall, D., Artaxo, P., Bretherton, C., Feingold, G., Forster, P., Kerminen, V.-M., Kondo, Y., Liao, H., Lohmann, U., Rasch, P., Satheesh, S., Sherwood, S., Stevens, B., and Zhang, X.: Clouds and Aerosols, book section 7, pp. 571–658, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA,

- <https://doi.org/10.1017/CBO9781107415324.016>, 2013. www.climatechange2013.org
- Bovensmann, H., Burrows, J., Buchwitz, M., Frerick, J., Noël, S., Rozanov, V., Chance, K., and Goede, A.: SCIAMACHY: Mission objectives and measurement modes, *Journal of the Atmospheric Sciences*, 56, 127–150, [https://doi.org/10.1175/1520-0469\(1999\)056<0127:SMOAMM>2.0.CO;2](https://doi.org/10.1175/1520-0469(1999)056<0127:SMOAMM>2.0.CO;2), 1999.
- Brown, P. T., Li, W., Jiang, J. H., and Su, H.: Unforced Surface Air Temperature Variability and Its Contrasting Relationship with the Anomalous TOA Energy Flux at Local and Global Spatial Scales, *Journal of Climate*, 29, 925–940, <https://doi.org/10.1175/JCLI-D-15-0384.1>, 2016.
- Burrows, J. P., Weber, M., Buchwitz, M., Rozanov, V., Ladstätter-Weißenmayer, A., Richter, A., DeBeek, R., Hoogen, R., Bramstedt, K., Eichmann, K.-U., et al.: The global ozone monitoring experiment (GOME): Mission concept and first scientific results, *Journal of the Atmospheric Sciences*, 56, 151–175, [https://doi.org/10.1175/1520-0469\(1999\)056<0151:TGOMEG>2.0.CO;2](https://doi.org/10.1175/1520-0469(1999)056<0151:TGOMEG>2.0.CO;2), 1999.
- Callies, J., Corpaccioli, E., Eisinger, M., Hahne, A., and Lefebvre, A.: GOME-2-Metop's second-generation sensor for operational ozone monitoring, *ESA bulletin*, 102, 28–36, 2000.
- Chan, K. L., Valks, P., Slijkhuis, S., Köhler, C., and Loyola, D.: Total column water vapor retrieval for GOME-2 visible blue observations, *Atmos. Meas. Tech. Discuss.*, <https://doi.org/10.5194/amt-2020-96>, in review, 2020.
- Chance, K. and Kurucz, R.: An improved high-resolution solar reference spectrum for earth's atmosphere measurements in the ultraviolet, visible, and near infrared, *Journal of Quantitative Spectroscopy and Radiative Transfer*, 111, 1289–1295, <https://doi.org/http://dx.doi.org/10.1016/j.jqsrt.2010.01.036>, 2010.
- Clough, S. A. and Iacono, M. J.: Line-by-line calculation of atmospheric fluxes and cooling rates: 2. Application to carbon dioxide, ozone, methane, nitrous oxide and the halocarbons, *Journal of Geophysical Research: Atmospheres*, 100, 16 519–16 535, <https://doi.org/10.1029/95JD01386>, 1995.
- Colman, R.: A comparison of climate feedbacks in general circulation models, *Climate Dynamics*, 20, 865–873, <https://doi.org/10.1007/s00382-003-0310-z>, 2003.
- Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, I., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, *Quarterly Journal of the Royal Meteorological Society*, 137, 553–597, <https://doi.org/10.1002/qj.828>, 2011.
- De Smedt, I., Theys, N., Yu, H., Danckaert, T., Lerot, C., Compennolle, S., Van Roozendael, M., Richter, A., Hilboll, A., Peters, E., Pedernana, M., Loyola, D., Beirle, S., Wagner, T., Eskes, H., van Geffen, J., Boersma, K. F., and Veefkind, P.: Algorithm theoretical baseline for formaldehyde retrievals from S5P TROPOMI and from the QA4ECV

- project, *Atmospheric Measurement Techniques*, 11, 2395–2426, <https://doi.org/10.5194/amt-11-2395-2018>, 2018.
- Grossi, M., Valks, P., Loyola, D., Aberle, B., Slijkhuis, S., Wagner, T., Beirle, S., and Lang, R.: Total column water vapour measurements from GOME-2 MetOp-A and MetOp-B, *Atmospheric Measurement Techniques*, 8, 1111–1133, <https://doi.org/10.5194/amt-8-1111-2015>, 2015.
- Hartmann, D., KleinTank, A., Rusticucci, M., Alexander, L., Brönnimann, S., Charabi, Y., Dentener, F., Dlugokencky, E., Easterling, D., Kaplan, A., Soden, B., Thorne, P., Wild, M., and Zhai, P.: Observations: Atmosphere and Surface, book section 2, pp. 159–254, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, <https://doi.org/10.1017/CBO9781107415324.008>, www.climatechange2013.org, 2013.
- Kalakoski, N., Kujanpää, J., Sofieva, V., Tamminen, J., Grossi, M., and Valks, P.: Validation of GOME-2/Metop total column water vapour with ground-based and in situ measurements, *Atmospheric Measurement Techniques*, 9, 1533–1544, <https://doi.org/10.5194/amt-9-1533-2016>, 2016.
- Kaufman, Y. J. and Gao, B. : Remote sensing of water vapor in the near IR from EOS/MODIS, *IEEE Transactions on Geoscience and Remote Sensing*, 30, 871–884, <https://doi.org/10.1109/36.175321>, 1992.
- Krotkov, N. A., Lamsal, L. N., Celarier, E. A., Swartz, W. H., Marchenko, S. V., Bucsela, E. J., Chan, K. L., Wenig, M., and Zara, M.: The version 3 OMI NO₂ standard product, *Atmospheric Measurement Techniques*, 10, 3133–3149, <https://doi.org/10.5194/amt-10-3133-2017>, 2017.
- Kiehl, J. T. and Trenberth, K. E.: Earth's Annual Global Mean Energy Budget, *Bulletin of the American Meteorological Society*, 78, 197–208, [https://doi.org/10.1175/1520-0477\(1997\)078<0197:EAGMEB>2.0.CO;2](https://doi.org/10.1175/1520-0477(1997)078<0197:EAGMEB>2.0.CO;2), 1997.
- Kleipool, Q. L., Dobber, M. R., de Haan, J. F., and Levelt, P. F.: Earth surface reflectance climatology from 3 years of OMI data, *Journal of Geophysical Research: Atmospheres*, 113, <https://doi.org/10.1029/2008JD010290>, 2008.
- Lampel, J., Pöhler, D., Tschirter, J., Frieß, U., and Platt, U.: On the relative absorption strengths of water vapour in the blue wavelength range, *Atmospheric Measurement Techniques*, 8, 4329–4346, <https://doi.org/10.5194/amt-8-4329-2015>, 2015.
- Levelt, P., Van den Oord, G. H. J., Dobber, M., Malkki, A., Visser, H., de Vries, J., Stammes, P., Lundell, J., and Saari, H.: The ozone monitoring instrument, *Geoscience and Remote Sensing, IEEE Transactions on*, 44, 1093–1101, <https://doi.org/10.1109/TGRS.2006.872333>, 2006.
- Li, Z., Fielding, E. J., Cross, P., and Muller, J.-P.: Interferometric synthetic aperture radar atmospheric correction: Medium Resolution Imaging Spectrometer and Advanced Synthetic Aperture Radar integration, *Geophysical Research Letters*, 33, <https://doi.org/10.1029/2005GL025299>, 2006.
- Liu, S., Valks, P., Pinardi, G., De Smedt, I., Yu, H., Beirle, S., and Richter, A.: An improved total and tropospheric NO₂ column retrieval for GOME-2, *Atmospheric Measurement Techniques*, 12, 1029–1057, <https://doi.org/10.5194/amt-12-1029-2019>, 2019.
- Loyola, D. G., Gimeno García, S., Lutz, R., Argyrouli, A., Romahn, F., Spurr, R. J. D., Pedernana, M., Doicu, A., Molina García, V., and Schüssler, O.: The operational

- cloud retrieval algorithms from TROPOMI on board Sentinel-5 Precursor, *Atmospheric Measurement Techniques*, 11, 409–427, <https://doi.org/10.5194/amt-11-409-2018>, 2018.
- Loyola, D.G., Pedernana, M., Gimeno García, S.. Smart sampling and incremental function learning for very large high dimensional data. *Neural Networks* 2016;78:75–87. doi:10.1016/j.neunet.2015.09.001.
- Loyola, D.G., Xu, J., Heue, K.P., Zimmer, W.. Applying FP ILM to the retrieval of geometry-dependent effective lambertian equivalent reflectivity (GE LER) daily maps from UVN satellite measurements. *Atmospheric Measurement Techniques* 2020;13(2):985–999. doi:10.5194/amt-13-985-2020.
- Lutz, R., Loyola, D., Gimeno García, S., and Romahn, F.: OCRA radiometric cloud fractions for GOME-2 on MetOp-A/B, *Atmospheric Measurement Techniques*, 9, 2357–2379, <https://doi.org/10.5194/amt-9-2357-2016>, 2016.
- 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.: An improved retrieval of tropospheric nitrogen dioxide from GOME, *Journal of Geophysical Research: Atmospheres*, 107, ACH 9–1–ACH 9–21, <https://doi.org/10.1029/2001JD001027>, 2002.
- Noël, S., Buchwitz, M., Bovensmann, H., Hoogen, R., and Burrows, J. P.: Atmospheric water vapor amounts retrieved from GOME satellite data, *Geophysical Research Letters*, 26, 1841–1844, <https://doi.org/10.1029/1999GL900437>, 1999.
- Noël, S., Buchwitz, M., and Burrows, J. P.: First retrieval of global water vapour column amounts from SCIAMACHY measurements, *Atmospheric Chemistry and Physics*, 4, 111–125, <https://doi.org/10.5194/acp-4-111-2004>, 2004.
- Palmer, P. I., Jacob, D. J., Chance, K., Martin, R. V., Spurr, R. J. D., Kurosu, T. P., Bey, I., Yantosca, R., Fiore, A., and Li, Q.: Air mass factor formulation for spectroscopic measurements from satellites: Application to formaldehyde retrievals from the Global Ozone Monitoring Experiment, *Journal of Geophysical Research: Atmospheres*, 106, 14 539–14 550, <https://doi.org/10.1029/2000JD900772>, 2001.
- Platt, U. and Stutz, J.: *Differential optical absorption spectroscopy - principles and applications*, Springer, 2008.
- Pope, R. M. and Fry, E. S.: Absorption spectrum (380-700 nm) of pure water. II. Integrating cavity measurements, *Appl. Opt.*, 36, 8710–8723, <https://doi.org/10.1364/AO.36.008710>, 1997.
- Pougatchev, N., August, T., Calbet, X., Hultberg, T., Oduleye, O., Schlüssel, P., Stiller, B., Germain, K. S., and Bingham, G.: IASI temperature and water vapor retrievals - error assessment and validation, *Atmospheric Chemistry and Physics*, 9, 6453–6458, <https://doi.org/10.5194/acp-9-6453-2009>, 2009.
- Richter, A. and Burrows, J.: Tropospheric NO₂ from GOME measurements, *Advances in Space Research*, 29, 1673 – 1683, [https://doi.org/https://doi.org/10.1016/S0273-1177\(02\)00100-X](https://doi.org/https://doi.org/10.1016/S0273-1177(02)00100-X), 2002.
- Román, R., Antón, M., Cachorro, V., Loyola, D., de Galisteo, J. O., de Frutos, A., and Romero-Campos, P.: Comparison of total water vapor column from GOME-2 on MetOp-A against ground-based GPS measurements at the Iberian Peninsula,

- Science of The Total Environment, 533, 317 – 328, <https://doi.org/https://doi.org/10.1016/j.scitotenv.2015.06.124>, 2015.
- Rothman, L., Gordon, I., Barber, R., Dothe, H., Gamache, R., Goldman, A., Perevalov, V., Tashkun, S., and Tennyson, J.: HITEMP, the high-temperature molecular spectroscopic database, *Journal of Quantitative Spectroscopy and Radiative Transfer*, 111, 2139 – 2150, <https://doi.org/https://doi.org/10.1016/j.jqsrt.2010.05.001>, 2010.
- Schröder, M., Lockhoff, M., Fell, F., Forsythe, J., Trent, T., Bennartz, R., Borbas, E., Bosilovich, M. G., Castelli, E., Hersbach, H., Kachi, M., Kobayashi, S., Kursinski, E. R., Loyola, D., Mears, C., Preusker, R., Rossow, W. B., and Saha, S.: The GEWEX Water Vapor Assessment archive of water vapour products from satellite observations and reanalyses, *Earth System Science Data*, 10, 1093–1117, <https://doi.org/10.5194/essd-10-1093-2018>, 2018.
- Schuessler, O., D. Loyola, A. Doicu, and R. Spurr. Information Content in the Oxygen A-band for the Retrieval of Macrophysical Cloud Parameters, *IEEE Transactions on Geoscience and Remote Sensing*, 52, 3246-3255, 2014.
- Slijkhuis, S., B. Aberle, and D. Loyola, GOME Data Processor Extraction Software User's Manual, ER-SUM-DLR-GO-0045, 2004.
- Soden, B. J. and Held, I. M.: An Assessment of Climate Feedbacks in Coupled Ocean - Atmosphere Models, *Journal of Climate*, 19, 3354–3360, <https://doi.org/10.1175/JCLI3799.1>, 2006.
- Soden, B. J., Jackson, D. L., Ramaswamy, V., Schwarzkopf, M. D., and Huang, X.: The Radiative Signature of Upper Tropospheric Moistening, *Science*, 310, 841–844, <https://doi.org/10.1126/science.1115602>, 2005.
- Solomon, S., Schmeltekopf, A. L., and Sanders, R. W.: On the interpretation of zenith sky absorption measurements, *Journal of Geophysical Research: Atmospheres*, 92, 8311–8319, <https://doi.org/10.1029/JD092iD07p08311>, 1987.
- Spurr, R., D. Loyola, W. Thomas, W. Balzer, E. Mikusch, B. Aberle, S. Slijkhuis, T. Ruppert, M. Van Roozendael, J.-C. Lambert, and T. V. Soebijanta, GOME Level 1-to-2 Data Processor Version 3.0: A Major Upgrade of the GOME/ERS-2 Total Ozone Retrieval Algorithm, *Applied Optics*, 44, 7196-7209, 2005.
- Spurr, R., LIDORT and VLIDORT: Linearized pseudo-spherical scalar and vector discrete ordinate radiative transfer models for use in remote sensing retrieval problems. *Light Scattering Reviews*, Volume 3, ed. A. Kokhanovsky, Springer, 2008.
- Stutz, J. and Platt, U.: Numerical analysis and estimation of the statistical error of differential optical absorption spectroscopy measurements with least-squares methods, *Applied Optic*, 35, 6041–6053, <https://doi.org/10.1364/AO.35.006041>, 1996.
- Thalman, R. and Volkamer, R.: Temperature dependent absorption cross-sections of O₂-O₂ collision pairs between 340 and 630 nm and at atmospherically relevant pressure, *Phys. Chem. Chem. Phys.*, 15, 15 371–15 381, <https://doi.org/10.1039/C3CP50968K>, 2013.
- Theys, N., De Smedt, I., Yu, H., Danckaert, T., van Gent, J., Hörmann, C., Wagner, T., Hedelt, P., Bauer, H., Romahn, F., Pedergnana, M., Loyola, D., and Van Roozendael, M.: Sulfur dioxide retrievals from TROPOMI onboard Sentinel-5 Precursor: algorithm theoretical basis, *Atmospheric Measurement Techniques*, 10, 119–153, <https://doi.org/10.5194/amt-10-119-2017>, 2017.

-
- Trenberth, K. E. and Stepaniak, D. P.: Covariability of Components of Poleward Atmospheric Energy Transports on Seasonal and Interannual Timescales, *Journal of Climate*, 16, 3691–3705, [https://doi.org/10.1175/1520-0442\(2003\)016<3691:COCOPA>2.0.CO;2](https://doi.org/10.1175/1520-0442(2003)016<3691:COCOPA>2.0.CO;2), 2003.
- Valks, P., Pinardi, G., Richter, A., Lambert, J.-C., Hao, N., Loyola, D., Van Roozendael, M., and Emmadi, S.: Operational total and tropospheric NO₂ column retrieval for GOME-2, *Atmospheric Measurement Techniques*, 4, 1491–1514, <https://doi.org/10.5194/amt-4-1491-2011>, 2011.
- Van Roozendael, M., D. Loyola, R. Spurr, D. Balis, J.-C. Lambert, Y. Livschitz, P. Valks, T. Ruppert, P. Kenter, C. Fayt, and C. Zehner, Ten years of GOME/ERS2 total ozone data: the new GOME Data Processor (GDP) Version 4: I. Algorithm Description, *J. Geophys Res.*, doi: 10.1029/2005JD006375, 2006.
- Vaquero-Martínez, J., Antón, M., de Galisteo, J. P. O., Cachorro, V. E., Álvarez-Zapatero, P., Román, R., Loyola, D., ao Costa, M. J., Wang, H., Abad, G. G., and Noël, S.: Inter-comparison of integrated water vapor from satellite instruments using reference GPS data at the Iberian Peninsula, *Remote Sensing of Environment*, 204, 729 – 740, <https://doi.org/https://doi.org/10.1016/j.rse.2017.09.028>, 2018.
- Veefkind, J., Aben, I., McMullan, K., Frster, H., de Vries, J., Otter, G., Claas, J., Eskes, H., de Haan, J., 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.: 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, *Remote Sensing of Environment*, 120, 70 – 83, <https://doi.org/https://doi.org/10.1016/j.rse.2011.09.027>, 2012.
- Wagner, T., Beirle, S., Grzegorski, M., and Platt, U.: Global trends (1996-2003) of total column precipitable water observed by Global Ozone Monitoring Experiment (GOME) on ERS-2 and their relation to near-surface temperature, *Journal of Geophysical Research: Atmospheres*, 111, <https://doi.org/10.1029/2005JD006523>, 2006.
- Wagner, T., Beirle, S., Sihler, H., and Mies, K.: A feasibility study for the retrieval of the total column precipitable water vapour from satellite observations in the blue spectral range, *Atmospheric Measurement Techniques*, 6, 2593–2605, <https://doi.org/10.5194/amt-6-2593-2013>, 2013.
- Wang, H., Liu, X., Chance, K., González Abad, G., and Chan Miller, C.: Water vapor retrieval from OMI visible spectra, *Atmospheric Measurement Techniques*, 7, 1901–1913, <https://doi.org/10.5194/amt-7-1901-2014>, 2014.
- Wang, H., Gonzalez Abad, G., Liu, X., and Chance, K.: Validation and update of OMI Total Column Water Vapor product, *Atmospheric Chemistry and Physics*, 16, 11 379–11 393, <https://doi.org/10.5194/acp-16-11379-2016>, 2016.
- Wang, H., Souri, A. H., González Abad, G., Liu, X., and Chance, K.: Ozone Monitoring Instrument (OMI) Total Column Water Vapor version 4 validation and applications, *Atmospheric Measurement Techniques*, 12, 5183–5199, <https://doi.org/10.5194/amt-12-5183-2019>, 2019.