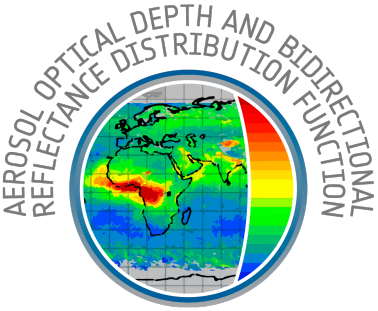
TROPOMI ATBD of the directionally dependent surface Lambertian-equivalent reflectivity



document number	:	S5P-KNMI-L3-0301-RP
authors	:	L. G. Tilstra
CI identification	:	CI-7430-ATBD
issue	:	2.1.0
date	:	2023-09-04
status	:	released

Document approval record

	digital signature
Prepared:	Gijsbert Tilstra Digitally signed by Gijsbert Tilstra Date: 2023.09.04 10:33:16 +02'00'
Checked:	
Approved PM:	
Approved PI:	

Document change record

Issue	Date	Item	Comments
0.0.1	2020-03-09	All	Initial version
0.0.2	2020-03-25	All	Version shared with team (confidential)
0.0.3	2020-06-16		Corrections and update after review:
		1.2	- further explained the handling of radiometric calibration errors
		6.2.2	- added references for the absorption cross-sections that were used
		6.2.2	- added a reference to the TROPOMI ISRF slit function database
		6.2.4	- clarified the storage of the intermediate L2 product
		8.1.2	- clarified the meaning of the “stop” (no processing) decision/action
		10.2	- added section 10.2 on preliminary validation results
		11	- updated section 11
0.1.0	2020-08-25	All	Minor changes but now with document number
1.0.0	2020-11-23		Corrections and update after review:
		5.2.1	- clarified “sufficient cloud free” requirement
		5.4	- clarified “sufficient quality” statement
		10.2	- updated section 10.2 and figure 15
		11	- updated section 11
1.1.0	2021-05-12		Update of manuscript:
		6.2.3	- updated the cloud fraction description
		6.3.3	- extended description of clear/snow-ice fields; added figures 7 and 8
		8	- updated table 8 and tables 10–15
		10.2	- updated section 10.2 and figure 17
1.2.0	2022-01-13	All	Released alongside version 1.0 of the DLER database
2.0.0	2023-08-29		Update of manuscript:
		1.2	- added update on the radiometric correction
		3.1	- added some symbols to “terms and definitions”
		3.2	- added a few acronyms to “acronyms and abbreviations”
		5	- updated tables 1 and 2
		6.2.4	- added section 6.2.4 on cloud shadow detection
		6.3.2	- added section 6.3.2 on cloud shadow filtering
		6.3.3	- added a sentence on the use of cloud shadow filtering
		6.3.4	- updated section 6.3.4 on distinction between clear/snow-ice fields
		6.3.4	- added figure 5; updated figures 6 and 7 and their captions
		8.1.2	- updated section 8.1.2 and table 6: added cloud fraction/height
		8.1.3	- updated table 7: added cloud shadow threshold
		8.2	- updated section 8.2 and table 8: added cloud shadow flag
		8.3.1	- updated table 10: added cloud shadow flag
		10.2	- updated section 10.2 and figure 18
2.1.0	2023-09-04	All	Released alongside version 2.0 of the DLER database

Contents

List of Tables

List of Figures

1 Introduction

1.1 Identification

This document is identified as S5P-KNMI-L3-0301-RP.

1.2 Purpose and objective

The purpose of this Algorithm Theoretical Baseline Document (ATBD) is to describe the theoretical basis and the implementation of the directionally dependent Lambertian-equivalent reflectivity (DLER) of the Earth's surface derived from observations by TROPOMI. The document is maintained during the development phase of the data products in the context of the S5p+I project. Several updates of the document are planned.

Note: The algorithm described in this ATBD assumes that the stability of the Earth reflectance over time is guaranteed. That is, it is expected that the Earth reflectance, derived from the radiance and solar irradiance measurements in the TROPOMI L1 product, will not show a trend caused by optical and/or electronic instrument degradation, or by any other cause of changes in detector performance.

If, however, calibration errors and/or trends in the Earth reflectances are found, and if these are not likely to be addressed by future updates of the TROPOMI L1 data processor, then dedicated reflectance correction factors may be applied to the TROPOMI L1 reflectances. Reflectance validation methods that derive these correction factors have already been been set up [RD1] by the development team responsible for this ATBD.

However, it is still expected that the need for such corrections will eventually disappear completely with the introduction of future updates of the TROPOMI L1 data processor.

Update: TROPOMI collection-3 data are sufficiently corrected for the effects of radiometric degradation. The data can be considered radiometrically stable and there is no need to apply correction factors.

1.3 Document overview

The structure of this ATBD is as follows. In section 2 applicable, standard and reference documents are listed. Section 3 introduces terms and definitions used in this ATBD, along with a list of acronyms and abbreviations that are used throughout the ATBD. Section 4 provides a reference to a general description of the TROPOMI instrument, which is common to all TROPOMI ATBDs. Section 5 presents an introduction to the algorithm setup that is defined in this ATBD. It also discusses the heritage and the most important properties of the end product, which is the TROPOMI surface DLER climatology. In section 6 the entire algorithm setup is discussed in detail. Section 7 presents results from a study in which the differences between surface DLER and surface BRDF are analysed. In section 8 the feasibility analysis is discussed. Section 9 discusses and explains the error analysis that is used in the algorithm setup. Section 10 suggests various methods for validation. Finally, section 11 provides a quick summary on the status and maturity of the ATBD and related activities.

2 Applicable and reference documents

2.1 Applicable documents

[AD1] TROPOMI Instrument and Performance Overview.
source: KNMI; **ref:** S5P-KNMI-L2-0010-RP; **issue:** 0.10.0; **date:** 2014-03-15.

2.2 Standard documents

[SD1] Space Engineering – Software.
source: ESA/ECSS; **ref:** ECSS-E-ST-40C; **issue:** 3; **date:** 2009-03-06.

2.3 Reference documents

[RD1] L. G. Tilstra, M. de Graaf, P. Wang *et al.*; In-orbit Earth reflectance validation for TROPOMI on board the Sentinel-5 Precursor satellite. *Atmos. Meas. Tech.*; **13** (2020) (8), 4479; doi:10.5194/amt-13-4479-2020. URL https://amt.copernicus.org/articles/13/4479/2020/.

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

[RD3] Terms and symbols in the TROPOMI Algorithm Team.
source: KNMI; **ref:** S5P-KNMI-L2-0049-MA; **issue:** 2.0.0; **date:** 2016-05-17.

[RD4] R. B. A. Koelemeijer, J. F. de Haan and P. Stammes; A database of spectral surface reflectivity in the range 335–772 nm derived from 5.5 years of GOME observations. *J. Geophys. Res.*; **108** (2003) (D2); doi:10.1029/2002JD002429. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2002JD002429.

[RD5] Q. L. Kleipool, M. R. Dobber, J. F. de Haan *et al.*; Earth surface reflectance climatology from 3 years of OMI data. *J. Geophys. Res.*; **113** (2008) (D18); doi:10.1029/2008JD010290. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2008JD010290.

[RD6] L. G. Tilstra, O. N. E. Tuinder, P. Wang *et al.*; Surface reflectivity climatologies from UV to NIR determined from Earth observations by GOME-2 and SCIAMACHY. *J. Geophys. Res. Atmos.*; **122** (2017) (7), 4084; doi:10.1002/2016JD025940. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2016JD025940.

[RD7] F. E. Nicodemus, J. C. Richmond, J. J. Hsia *et al.*; Geometrical Considerations and Nomenclature for Reflectance. *In Radiometry*; (pp. 94–145) (Jones and Bartlett Publishers, Inc., USA., 1992).

[RD8] L. G. Tilstra, O. N. E. Tuinder, P. Wang *et al.*; Directionally dependent Lambertian-equivalent reflectivity (DLER) of the Earth's surface measured by the GOME-2 satellite instruments. *Atmos. Meas. Tech.*; **14** (2021) (6), 4219; doi:10.5194/amt-14-4219-2021. URL https://amt.copernicus.org/articles/14/4219/2021/.

[RD9] D. F. Heath, A. J. Krueger, H. A. Roeder *et al.*; The Solar Backscatter Ultraviolet and Total Ozone Mapping Spectrometer (SBUV/TOMS) for NIMBUS G. *Opt. Eng.*; **14** (1975) (4), 323; doi:10.1117/12.7971839. URL https://doi.org/10.1117/12.7971839.

[RD10] J. R. Herman and E. A. Celarier; Earth surface reflectivity climatology at 340–380 nm from TOMS data. *J. Geophys. Res.*; **102** (1997) (D23), 28003; doi:10.1029/97JD02074. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/97JD02074.

[RD11] J. P. Burrows, M. Weber, M. Buchwitz *et al.*; The Global Ozone Monitoring Experiment (GOME): Mission concept and first scientific results. *J. Atmos. Sci.*; **56** (1999) (2), 151; doi:10.1175/1520-0469(1999)056<0151:TGOMEG>2.0.CO;2. URL https://doi.org/10.1175/1520-0469(1999)056<0151:TGOMEG>2.0.CO;2.

[RD12] P. F. Levelt, G. H. J. van den Oord, M. R. Dobber *et al.*; The ozone monitoring instrument. *IEEE Trans. Geosci. Remote Sens.*; **44** (2006) (5), 1093; doi:10.1109/TGRS.2006.872333. URL https://ieeexplore.ieee.org/document/1624590.

[RD13] R. Munro, R. Lang, D. Klaes *et al.*; The GOME-2 instrument on the Metop series of satellites: instrument design, calibration, and level 1 data processing – an overview. *Atmos. Meas. Tech.*; **9** (2016) (3), 1279; doi:10.5194/amt-9-1279-2016. URL https://www.atmos-meas-tech.net/9/1279/2016/.

[RD14] GOME-2 surface LER product – Algorithm Theoretical Basis Document.
source: KNMI; **ref:** SAF/AC/KNMI/ATBD/003; **issue:** 4.1; **date:** 2023-01-05.

[RD15] H. Bovensmann, J. P. Burrows, M. Buchwitz *et al.*; SCIAMACHY: Mission objectives and measurement modes. *J. Atmos. Sci.*; **56** (1999) (2), 127; doi:10.1175/1520-0469(1999)056<0127:SMOAMM>2.0.CO;2. URL https://doi.org/10.1175/1520-0469(1999)056<0127:SMOAMM>2.0.CO;2.

[RD16] L. G. Tilstra, M. de Graaf, I. Aben *et al.*; In-flight degradation correction of SCIAMACHY UV reflectances and Absorbing Aerosol Index. *J. Geophys. Res.*; **117** (2012) (D6); doi:10.1029/2011JD016957. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2011JD016957.

[RD17] TROPOMI ATBD of the directionally dependent surface Lambertian-equivalent reflectivity.
source: KNMI; **ref:** S5P-KNMI-L3-0301-RP; **issue:** 2.0.0; **date:** 2023-08-29.

[RD18] A. Lorente, K. F. Boersma, P. Stammes *et al.*; The importance of surface reflectance anisotropy for cloud and NO_2 retrievals from GOME-2 and OMI. *Atmos. Meas. Tech.*; **11** (2018) (7), 4509; doi:10.5194/amt-11-4509-2018. URL https://www.atmos-meas-tech.net/11/4509/2018/.

[RD19] S5p+I – Requirements Baseline Document.
source: KNMI, GRASP, Catalysts; **ref:** D1-RBD; **issue:** 1.0; **date:** 2019-10-09.

[RD20] TROPOMI ATBD of the total and tropospheric NO_2 data products.
source: KNMI; **ref:** S5P-KNMI-L2-0005-RP; **issue:** 1.4.0; **date:** 2019-02-06.

[RD21] H. Akima; A new method of interpolation and smooth curve fitting based on local procedures. *J. ACM*; **17** (1970) (4), 589.

[RD22] TROPOMI ATBD of the UV aerosol index.
source: KNMI; **ref:** S5P-KNMI-L2-0008-RP; **issue:** 1.1; **date:** 2018-06-15.

[RD23] S. Chandrasekhar; *Radiative Transfer* (Dover, Mineola, N. Y., 1960).

[RD24] J. F. de Haan, P. B. Bosma and J. W. Hovenier; The adding method for multiple scattering calculations of polarized light. *Astron. & Astrophys.*; **183** (1987), 371.

[RD25] P. Stammes; Spectral radiance modelling in the UV-visible range. *IRS 2000: Current Problems in Atmospheric Radiation*; (2001), 385.

[RD26] AFGL atmospheric constituent profiles (0–120 km).
source: Air Force Geophys. Lab., Hanscom AFB, Mass.; **ref:** Environ. Res. Pap. 954, Rep. AFGL-TR-86-0110.

[RD27] A. M. Bass and R. J. Paur; The ultraviolet cross-sections of ozone: Part I. The measurements. *In Proceedings of the Quadrennial Ozone Symposium on Atmospheric Ozone* (edited by S. Zerefos, A. Ghazi and D. Reidel); volume 1; (pp. 606–610) (Halkidiki, Greece, 1985).

[RD28] A. C. Vandaele, C. Hermans, P. C. Simon *et al.*; Measurements of the NO_2 absorption cross-section from 42 000 cm^{-1} to 10 000 cm^{-1} (238–1000 nm) at 220 K and 294 K. *J. Quant. Spectrosc. Radiat. Transfer*; **59** (1998), 171.

[RD29] L. S. Rothman, I. E. Gordon, A. Barbe *et al.*; The HITRAN 2008 molecular spectroscopic database. *J. Quant. Spectrosc. Radiat. Transfer*; **110** (2009) (9), 533; doi:10.1016/j.jqsrt.2009.02.013. URL http://www.sciencedirect.com/science/article/pii/S0022407309000727.

[RD30] S5P-NPP Cloud Processor ATBD.
source: RAL Space; **ref:** S5P-NPPC-RAL-ATBD-0001; **issue:** 1.0.0; **date:** 2016-02-12.

[RD31] V. J. H. Trees, P. Wang, P. Stammes *et al.*; DARCLOS: a cloud shadow detection algorithm for TROPOMI. *Atmos. Meas. Tech.*; **15** (2022) (10), 3121; doi:10.5194/amt-15-3121-2022. URL https://amt.copernicus.org/articles/15/3121/2022/.

[RD32] O. Torres, P. K. Bhartia, J. R. Herman *et al.*; Derivation of aerosol properties from satellite measurements of backscattered ultraviolet radiation: Theoretical basis. *J. Geophys. Res.*; **103** (1998) (D14), 17099; doi:10.1029/98JD00900. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/98JD00900.

[RD33] M. de Graaf, P. Stammes, O. Torres *et al.*; Absorbing Aerosol Index: Sensitivity analysis, application to GOME and comparison with TOMS. *J. Geophys. Res.*; **110** (2005) (D1); doi:10.1029/2004JD005178. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2004JD005178.

[RD34] MODIS BRDF/Albedo Product: Algorithm Theoretical Basis Document.
source: MODIS Science Team; **ref:** MODIS Product ID: MOD43; **issue:** 5.0; **date:** April 1999.

[RD35] C. Popp, P. Wang, D. Brunner *et al.*; MERIS albedo climatology for FRESCO+ O_2 A-band cloud retrieval. *Atmos. Meas. Tech.*; **4** (2011) (3), 463; doi:10.5194/amt-4-463-2011. URL https://www.atmos-meas-tech.net/4/463/2011/.

[RD36] F. Gao, C. B. Schaaf, A. H. Strahler *et al.*; MODIS bidirectional reflectance distribution function and albedo Climate Modeling Grid products and the variability of albedo for major global vegetation types. *J. Geophys. Res. Atmos.*; **110** (2005), D01104; doi:10.1029/2004JD005190. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2004JD005190.

2.4 Electronic references

[ER1] URL https://www.temis.nl/surface/albedo/toms_ler.php.

[ER2] URL https://www.temis.nl/surface/albedo/gome_ler.php.

[ER3] URL https://www.temis.nl/surface/albedo/omi_ler.php.

[ER4] URL https://www.temis.nl/surface/albedo/scia_ler.php.

[ER5] URL https://www.temis.nl/surface/albedo/gome2_ler.php.

[ER6] URL http://spectrolab.aeronomie.be.

[ER7] URL http://www.tropomi.eu/data-products/isrf-dataset/.

[ER8] URL https://lpdaac.usgs.gov/products/mcd43c2v061/.

3 Terms, definitions and abbreviated terms

Terms, definitions and abbreviated terms that are used in the documentation of the TROPOMI L0-1b data processor are described in [RD2]. Terms, definitions and abbreviated terms for TROPOMI Level 2 algorithms are described in [RD3]. Terms, definitions and abbreviated terms specific for this document are defined below.

3.1 Terms and definitions

θ	viewing zenith angle
θ_0	solar zenith angle
ϕ	viewing azimuth angle
ϕ_0	solar azimuth angle
μ	cosine of the viewing zenith angle
μ_0	cosine of the solar zenith angle
i	index of detector pixel in the spectral direction
s	index of detector pixel in the spatial direction
j	index of wavelength band ($j = 1 \ldots 21$)
λ_i	wavelength assigned to detector pixel i
λ_j^c	central wavelength of wavelength band j
ω_j	half of the bandwidth of wavelength band j
I	Earth radiance
E_0	solar irradiance
R	reflectance
R_λ	reflectance of wavelength band, labelled with λ
R_j	reflectance of wavelength band, indexed by j ($j = 1 \ldots 21$)
A_s	surface albedo
A_scene	scene LER
A_DLER	surface DLER
T	atmospheric transmission
$s^\star$	spherical albedo of the atmosphere for illumination from below
N_j	number of detector pixels contributing to band reflectance j
w_i^j	weighting factor specifying the contribution of detector pixel i to band reflectance j
$S_i(\lambda)$	spectral response function of detector pixel i
$G_j(\lambda)$	spectral response function of wavelength band j
$R^\mathrm{sim}(\lambda)$	simulated high-spectral-resolution reflectance spectrum
R_j^sim	simulated reflectance of wavelength band j
m	index of observation inside a grid cell
v_m	weighting factor specifying the contribution of observation m to A_LER
Γ	SCNLER–DLER contrast parameter used for the cloud shadow detection
Γ_0	contrast threshold for Γ

3.2 Acronyms and abbreviations

AAI	Absorbing Aerosol Index
ATBD	Algorithm Theoretical Baseline Document
BRDF	Bidirectional Reflectance Distribution Function
BSA	Black-Sky Albedo
CAMS	Copernicus Atmosphere Monitoring Service
CF	Climate and Forecast metadata conventions
DAK	Doubling-Adding KNMI

DARIUS	Development of Advanced Retrieval of Aerosol and Surface Properties from S5P
DLER	Directionally dependent Lambertian-Equivalent Reflectivity
DU	Dobson Units, 2.69×10^{16} molecules cm^{-2}
ECMWF	European Centre for Medium-Range Weather Forecast
ENVISAT	Environmental Satellite
EOS-Aura	Earth Observing System – Aura satellite
EPS-SG	EUMETSAT Polar System – Second Generation
ERS	European Remote Sensing Satellite
ESA	European Space Agency
EUMETSAT	European Organisation for the Exploitation of Meteorological Satellites
FOV	Field-of-View
FRESCO	Fast Retrieval Scheme for Clouds from the Oxygen A band
GMTED2010	Global Multi-resolution Terrain Elevation Data 2010
GOME	Global Ozone Monitoring Experiment
GRASP	Generalized Retrieval of Atmosphere and Surface Properties
HDF	Hierarchical Data Format
HITRAN2008	High-Resolution Transmission molecular absorption database, 2008 edition
IR	Infrared
ISRF	Instrument Spectral Response Function
KNMI	Koninklijk Nederlands Meteorologisch Instituut
LER	Lambertian-Equivalent Reflectivity
LUT	Look-Up Table
L2OP	Level-2 Operational Processor
L2PP	Level-2 Prototype Processor
MERIS	Medium Resolution Imaging Spectrometer
METOP	Meteorological Operational Satellite
MLS	Mid-Latitude Summer
NASA	National Aeronautics and Space Administration
NETCDF	Network Common Data Form, NetCDF
NIR	Near-Infrared
NISE	Near-real-time Ice and Snow Extent
NRT	Near-Real-Time
OMI	Ozone Monitoring Instrument
RAA	Relative Azimuth Angle
RMSE	Root-Mean-Square Error
RTM	Radiative Transfer Model
SAA	Solar Azimuth Angle
SCIAMACHY	Scanning Imaging Absorption Spectrometer for Atmospheric Chartography
Suomi NPP	Suomi National Polar-orbiting Partnership
SW	Software
SWIR	Short-Wavelength Infrared
SZA	Solar Zenith Angle
S5	Sentinel-5 mission
S5P	Sentinel-5 Precursor mission
TBA	To be Added
TBC	To be Confirmed
TBD	To be Defined
TOA	Top-of-Atmosphere

TOMS	Total Ozone Mapping Spectrometer
TROPOMI	Tropospheric Monitoring Instrument
UTC	Coordinated Universal Time
UV	Ultraviolet
UVNS	Ultraviolet Visible Near-infrared Shortwave spectrometer
VAA	Viewing Azimuth Angle
VIIRS	Visible Infrared Imaging Radiometer Suite
VIS	Visible
VZA	Viewing Zenith Angle

4 TROPOMI instrument description

A description of the TROPOMI instrument and performance can be found in [AD1].

5 Introduction to the TROPOMI DLER database

5.1 Background

Surface reflectivity databases are needed for cloud, aerosol and trace gas retrievals. Examples are the retrieval of trace gases such as ozone, NO_2, BrO, CH_2O, H_2O, CO_2, CO, and CH_4, and of cloud information and aerosol optical depth. The TROPOMI surface DLER product is the directionally dependent Lambertian-equivalent reflectivity (LER) of the Earth's surface observed by TROPOMI. It is the improved follow-up of earlier surface LER databases based on observations performed by GOME-1 (on ERS-2) [RD4], OMI (on the Aura satellite) [RD5], as well as SCIAMACHY (on Envisat) and GOME-2 (on the MetOp satellites) [RD6]. In this chapter, the term "surface reflectivity database" refers to global climatologies of surface albedo, available for several wavelength bands and for each month. The databases are generally extracted from several years of data and the term "surface reflectivity database" therefore always refers to a L4 product.

5.2 Heritage

The heritage from the traditional surface LER databases is described in section 5.2.1. These databases adhere to the principle of Lambertian surface reflection, which means that they completely disregard the fact that surface reflection is described by a bi-directional reflectance distribution function (BRDF) [RD7] which takes into account the dependence on the incoming and outgoing directions of the light reflected by the surface. Lambertian surface reflection is by definition non-directional. A directionally dependent LER (acronym: DLER) was first derived for the GOME-2 instrument [RD8]. This database will be discussed in section 5.2.2.

5.2.1 Lambertian-equivalent reflectivity (LER) databases

One of the first surface reflectivity databases retrieved using UV satellite remote sensing techniques is the Total Ozone Mapping Spectrometer (TOMS) [RD9] surface LER database [RD10]. The retrieved reflectivity is the so-called Lambertian-equivalent reflectivity (LER) of the surface found from scenes which are assumed to be cloud free. The retrieval method relies on the removal of the (modelled) atmospheric contribution from the (observed) top-of-atmosphere (TOA) reflectance. In this approach the surface is defined to behave as a Lambertian reflector. The TOMS surface LER database, which is provided in a spatial grid of $1.25^\circ \times 1.0^\circ$, was retrieved for 340 and 380 nm only, which severly limits its usefulness.

The GOME-1 [RD11] surface reflectivity database provides the surface LER on a $1.0^\circ \times 1.0^\circ$ grid for 11 wavelength bands between 335 and 772 nm [RD4]. Although this is already quite an improvement with respect to the TOMS surface LER database, the database is still limited in quality by the low number of measurements from which the surface LER had to be extracted and the large GOME footprint size (see Table 1). In particular, pixels over sea are often affected by residual cloud contamination. In these cases the surface LER was retrieved from scenes which were not sufficiently cloud free, evidenced by LER values at 772 nm exceeding 0.05 [RD4]. In other cases, e.g. snow surfaces, the surface LER was retrieved from a few measurements which were not representative for the entire month.

A large improvement on these points is the OMI surface reflectivity database [RD5]. First, the OMI instrument [RD12] has a much smaller footprint size (24 $\times$ 13 km^2 at nadir) combined with a larger global coverage (see Table 1). This leads to better statistics and results in a higher accuracy for the surface LER retrieval. Second, the higher number of measurements allows for inspecting the distribution of scene LERs for each grid cell, and for making a more sophisticated selection of representative (cloud-free) scenes instead of directly taking the minimum scene LER value like in the case of the TOMS and GOME-1 databases. Third, the provided OMI surface LER database has a higher spatial resolution of $0.5^\circ \times 0.5^\circ$. The limiting factor is the OMI wavelength range. The longest wavelength in the OMI surface LER database is 499 nm.

The GOME-2 series of satellite instruments [RD13] does not have the limitations of the above instruments and therefore can be used to create a better surface LER database. The GOME-2 instrument has the spectral range of GOME but a much smaller footprint (80 $\times$ 40 km^2) which is constant over the full swath width. The number of measurements that are available per longitude/latitude cell in the database grid is smaller than that of OMI, but enough to perform a statistical analysis on the distribution of retrieved scene LERs. The intrinsic spatial resolution of the GOME-2 surface LER database is $1.0^\circ \times 1.0^\circ$ [RD6], except near the coastlines and for certain region such as snow covered mountain ranges where it is $0.25^\circ \times 0.25^\circ$ [RD8, RD14].

The SCIAMACHY instrument [RD15] is comparable to the GOME-2 instrument, but covers a much larger spectral range. The derived SCIAMACHY surface LER database offers 34 wavelength bands between 328 and 2314 nm [RD6]. The main advantage of the GOME-2 and SCIAMACHY surface LER databases with respect

Surface LER database →	**TOMS**	**GOME-1**	**OMI**	**SCIAMACHY**	**GOME-2**	**Sentinel-5P**
surface reflectivity type	LER	LER	LER	LER	LER / DLER	LER / DLER
dataset time range [(1)]	1978–1993	1995–2000	2004–2007	2002–2012	2007→	2018→
selected wavelength bands	2	11	23	34	27	21
wavelength range covered [nm]	340–380	335–772	328–499	328–2314	328–772	328–2314
band width [nm]	1.0	1.0	1.0	1.0	1.0	1.0
spatial resolution [°lon × °lat]	1.25 × 1.0	1.0 × 1.0	0.5 × 0.5	0.5 × 0.5 [(2)]	0.25 × 0.25 [(2)]	0.125 × 0.125
reference	[RD10]	[RD4]	[RD5]	[RD6]	[RD6, RD8]	[RD17]
instrument	TOMS	GOME	OMI	SCIAMACHY	GOME-2	TROPOMI
satellite	Nimbus-7	ERS-2	Aura	Envisat	MetOp-A/B/C	Sentinel-5P
equator crossing time (LT)	12:00	10:30	13:45	10:00	09:30	13:30
dayside flight direction	S→N	N→S	S→N	N→S	N→S	S→N
number of days for global coverage	1	3	1	6	1.5	1
pixel size at nadir [km × km]	50 × 50	320 × 40	24 × 13	60 × 30	80 × 40	5.6 × 3.6 [(3)]
number of usable pixels per orbit	∼12000	∼1300	∼83000	∼4000	∼11000	∼1300000

Table 1: Characteristics and properties of the surface LER databases, and of the satellite instruments from which they are derived. Wavelength band information can be found in Table 2.

to the OMI surface LER database is the wider wavelength range of the GOME-2 and SCIAMACHY instrument. Additionally, the retrieval algorithm uses aerosol information, available via the Absorbing Aerosol Index (AAI) product [RD16], to filter out scenes with large amounts of aerosols, as these scenes can result in inaccurate values of the retrieved surface LER. This filtering is especially important for locations over desert areas.

The SCIAMACHY and GOME-2 surface LER databases are derived by similar retrieval codes using similar techniques. Together with the OMI surface LER database they provide all the heritage available to develop the TROPOMI surface DLER database. Table 1 summarises the properties of the surface reflectivity databases discussed in this section. References to papers and other sources of information are also provided. Table 2 provides detailed information on the wavelength bands that are contained in the surface LER databases.

5.2.2 Directional LER databases

The concept of a directional LER was first introduced to the GOME-2 surface DLER database [RD8]. The GOME-2 surface DLER database provides, next to the traditional LER database, a DLER database which is essentially a function of the viewing direction (and of month, wavelength, latitude, and longitude). For the GOME-2 surface DLER database, the difference between DLER and LER is the largest for the west viewing geometries. For example, at 772 nm the surface DLER for a vegetated scene over Amazonia can be twice as large in the west viewing direction as in the east viewing direction [RD18]. The traditional LER databases discussed in section 5.2.1 completely miss this directional dependence, and generally present a value which is very close to the minimum of the DLER over all viewing angles.

5.2.3 Links to the heritage LER databases

The heritage LER databases can be download from the following locations:

- TOMS surface LER : [ER1]
- GOME-1 surface LER : [ER2]
- OMI surface LER : [ER3]
- SCIAMACHY surface LER : [ER4]
- GOME-2 surface DLER : [ER5]

(1) The longer the time period, the larger the number of times a certain region has been observed. This increases the chances of having observed this region under clear-sky conditions. (2) The *intrinsic* resolution is in most cases 1.0°×1.0°. (3) The TROPOMI spatial sampling in the along track direction was changed from 7.2 km to the indicated 5.6 km on 6 August 2019, at the start of orbit 9388.

5.3 TROPOMI DLER database

The TROPOMI surface DLER database is in many aspects a huge step forward compared to the other databases described in Table 1. The spatial resolution of the TROPOMI DLER database will be $0.125^\circ \times 0.125^\circ$. This high spatial resolution is made possible by the small TROPOMI footprint size of 5.6 $\times$ 3.6 km^2. The smaller footprint size also results in a much larger number of cloud-free pixels that may be collected over the course of a given time period. This is beneficial for the stability and the quality of the retrieved DLER spectra. Figure 1 provides a good indication of the spatial resolution for the surface LER climatologies derived from GOME-1, GOME-2, and TROPOMI. Obviously, the higher spatial resolution of TROPOMI is an important improvement.

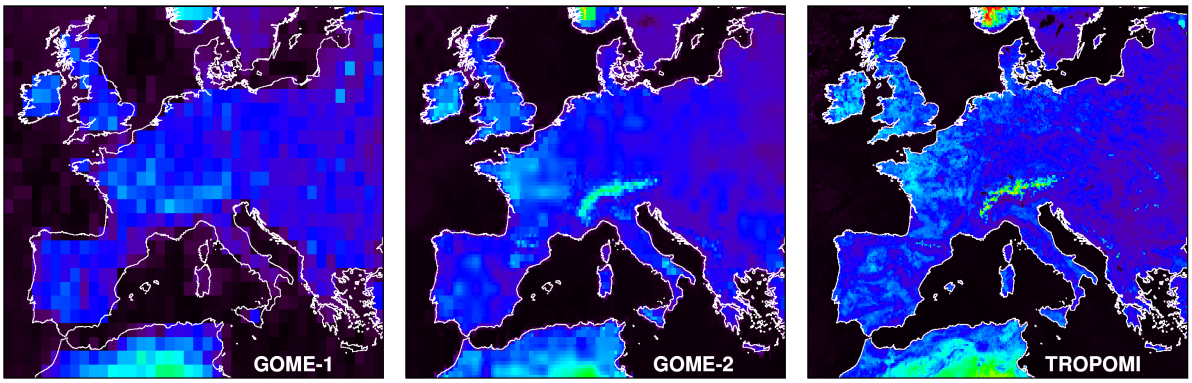


Figure 1: Surface LER in March over Europe according to the GOME-1, GOME-2, and TROPOMI databases.

Another step forward is the way in which the cloud screening is performed. For all surface LER databases listed in Table 1 the cloud screening is based on statistical methods. The advantage of this is that external cloud information is not required. The disadvantage is the risk of residual cloud contamination for areas known to be nearly always covered by persistent cloud decks. Another disadvantage is the high computational effort needed: all scenes need to be handled at once, cloud-free or not, which can be quite a task when handling a satellite instrument like TROPOMI with more than a million footprints per orbit (see Table 1).

For TROPOMI the cloud screening will be based primarily on the S5P NPP-VIIRS cloud product. The NPP-VIIRS product is based on observations performed by the VIIRS instrument onboard the Suomi NPP satellite. The S5P and NPP satellites are kept in a “loose” formation in very similar orbits, resulting in only a relatively small time difference (currently the difference is approximately 3 minutes, reduced from 5 minutes at the start of the mission). The VIIRS cloud information is provided for each individual TROPOMI footprint. More precisely, the product provides cloud information for four sets of boxes surrounding the TROPOMI measurement footprint. For our purpose, we use the nominal size of these boxes, corresponding to the smallest FOV, which matches precisely the size of the TROPOMI footprints as found in the L1 and L2 products. Separate S5P NPP-VIIRS products are provided for spectral band 3, 6, and 7. This means that there is collocated cloud information available for all used TROPOMI spectral bands (these are: bands 3/4, 5/6, and 7).

Another major improvement is in the fact that for determining the TROPOMI surface DLER, the information can be studied as a function of viewing angle without the use of viewing-angle containers. For the GOME-2 DLER, these viewing-angle containers are needed to be able to perform the statistical cloud screening [RD8]. The fact that the “container” approach is not needed for deriving the TROPOMI surface DLER is, therefore, a direct result of the decision to use the S5P NPP-VIIRS cloud product for the cloud screening.

Generally speaking, the approach that was used for the GOME-2 and SCIAMACHY surface reflectivity databases was followed closely and applied in the TROPOMI surface DLER retrieval algorithm.

5.4 Wavelength bands

In Table 2 we list the wavelength bands of the surface reflectivity databases discussed in section 5.2, and their application. As can be seen, the selection of the wavelength bands for TROPOMI was influenced largely by the already existing surface LER databases. Below 325 nm the surface contribution to the TOA reflectance is low, which prevents an accurate retrieval of the surface LER below this wavelength. The TROPOMI surface LER wavelength band at 328 nm is retrieved, but for the time being only for testing and monitoring purposes. It will not be part of the TROPOMI surface DLER database unless it is proven to be of sufficient quality. Sufficient quality means that the requirements set in the Requirements Baseline Document [RD19] need to be met.

Band	TOMS	GOME	OMI	SCIAMACHY	GOME-2	TROPOMI	Application
328			+ (1)	+ (1)	+ (1)	+ (1)	ozone, HCHO, SO_2
335		+	+	+	+	+	ozone, HCHO
340	+			+	+	+	cloud, aerosol, HCHO, BrO
342			+				cloud, aerosol, HCHO, BrO
345			+				cloud, aerosol, HCHO, BrO
354			+	+	+	+	cloud, aerosol, HCHO, BrO, OClO
367			+	+	+	+	cloud, aerosol, OClO
372			+				cloud, aerosol, OClO
376			+				cloud, aerosol, OClO
380	+	+	+	+	+	+	cloud, aerosol, OClO
388			+	+	+	+	cloud, aerosol, OClO
402						+	cloud, aerosol
406			+				cloud, aerosol
416		+	+		+	+	cloud, aerosol
418			+				cloud, aerosol
425			+	+	+	+	cloud, aerosol, NO_2
440		+	+	+	+	+	cloud, aerosol, NO_2
442			+				cloud, aerosol, NO_2
452			+				cloud, aerosol, NO_2
463		+	+	+	+	+	cloud, aerosol, NO_2, O_2-O_2
471			+				cloud, aerosol, NO_2, O_2-O_2
477			+ (2)				cloud, aerosol, NO_2, O_2-O_2
488			+				cloud, aerosol, NO_2, O_2-O_2
494		+	+	+	+	+	cloud, aerosol, NO_2, O_2-O_2
499			+				cloud, aerosol
510				+	+		cloud, aerosol
526				+	+		cloud, aerosol
546				+	+		cloud, aerosol
555		+		+	+		cloud, aerosol
564				+	+		cloud, aerosol, O_2-O_2
585					+		cloud, aerosol, O_2-O_2, H_2O
610		+			+		cloud, aerosol, H_2O
614				+			cloud, aerosol, H_2O
640				+	+		cloud, aerosol, H_2O
670		+		+	+	+	cloud, aerosol, H_2O, O_2-B
685				+	+	+	cloud, aerosol, H_2O, O_2-B
697				+	+	+ (3)	cloud, aerosol, H_2O, O_2-B
712				+	+	+ (3)	cloud, aerosol, H_2O, O_2-B
747					+	+	cloud, aerosol, H_2O
758		+		+	+	+	cloud, aerosol, O_2-A
772		+		+	+	+	cloud, aerosol, O_2-A
862				+			cloud, aerosol, H_2O
875				+			cloud, aerosol, H_2O

1030				+			cloud, aerosol, H_2O
1053				+			cloud, aerosol, H_2O
1245				+			cloud, aerosol, H_2O
1557				+			cloud, aerosol, H_2O, CH_4, CO_2
1593				+			cloud, aerosol, H_2O, CH_4, CO_2
1630				+			cloud, aerosol, H_2O, CH_4, CO_2
1670				+			cloud, aerosol, H_2O, CH_4, CO_2
2314				+		+ (4)	cloud, aerosol, H_2O, CH_4, CO

Table 2: Wavelength bands of the surface LER databases discussed in this paper, and their atmospheric applications. All wavelength bands are located outside strong gaseous absorption bands in order to avoid complicated modeling of the radiative transfer involved. The wavelength bands are 1 nm wide in most cases.

For most of the wavelength bands the exact central wavelength is not very critical, and the exact central position in the algorithm setup is exactly that as shown in Table 2. Exceptions are the 697 and 712-nm TROPOMI DLER wavelength bands. These have been given a central wavelength of 696.97 and 712.70 nm, respectively. The positions of these wavelength bands are based on very precise spectral calculations with the goal to minimise the impact of surrounding water vapour absorption bands. The width of the wavelength bands will be one nm in most cases. Exceptions are the wavelength bands at 697, 712, and 2314 nm. The 697 and 712-nm wavelength bands will have a bandwidth of 0.3 nm. The 2314-nm wavelength band will have a bandwidth of 0.5 nm. The bandwidths that are defined for TROPOMI are in line with those used in the earlier databases.

5.5 Products

The end product of the processing chain is the TROPOMI surface DLER database. As explained in section 6, intermediate products are produced in the processing chain, such as the L2 LER product. This L2 product contains the scene LER as defined in section 6.2.1. The scene LER is in general not equal to the surface LER and should only be used as such if the scene is absolutely cloud- and aerosol-free. The other intermediate product, the monthly L3 product, does contain genuine surface albedo values, but is not being distributed because this product did not receive the many post-processing steps that the DLER database did receive.

5.6 Generation frequency

The generation frequency of the L2 products is once per L1B product. The L3 products that are generated by component B are created once per month, triggered by a time signal at the end of the month + an added number of m days for safety (making sure that all L2 files were properly processed). Note that component B will provide a L3 file based on the data of the respective month of all previous years. That is, the “March” L3 product will be based on measurements from the month March of all available years. The L4 product will also be generated once per month, at the end of the month + a number of n days added for safety, where $n > m+1$. The exact values for m and n should be determined by the party responsible for the processing. The L4 product contains the surface LER fields for all twelve calendar months. Table 3 summarises the scheme.

Product	Component	Generation frequency	Time trigger	Section
L2	A	once per L1B	normal processing	6.2
L3	B	once per month	end of month + m days	6.3
L4	C	once per month	end of month + n days	6.4

Table 3: Generation frequency for the L2, L3, and L4 products.

(1) Retrieval of the surface LER below 330 nm is challenging because of the small contribution of the surface to the TOA reflectance. (2) The wavelength band at 477 nm is ill-positioned because here the retrieved (OMI) surface LER is affected by O_2-O_2 absorption. (3) The 697 and 712-nm wavelength bands are surrounded by water vapour absorption bands. Their bandwidth is 0.3 nm and their exact positions are 696.97 and 712.70 nm, respectively. (4) The 2314-nm wavelength band will have a bandwidth of 0.5 nm.

6 Algorithm description

6.1 Overview

The final objective is to create the TROPOMI surface DLER climatology, contained in one file, providing the surface DLER for a set of wavelength bands relevant for atmospheric retrieval. To achieve this in the most effective way, the algorithm is set up in three components (named A, B, and C). The complete processing scheme that will create the surface DLER climatology is summarised graphically in Figure 2.

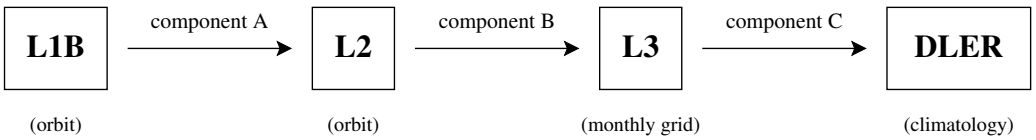


Figure 2: Processing scheme described in this section, illustrating the steps needed for the conversion of TROPOMI L1B products into the end product, the TROPOMI surface DLER climatology. The L2 product created by component A is an intermediate product containing the scene LER. Monthly grids, produced by component B, are intermediate L3 products containing the TROPOMI surface DLER. Finally, component C produces the TROPOMI surface DLER climatology after having performed several post-processing steps.

The first component (A) creates from the S5P L1B files S5P LER L2 files containing the calculated scene LER. As far as the generation of the TROPOMI surface DLER climatology data processor is concerned, this L2 scene LER product is merely an intermediate product. Component A is described in section 6.2.

Component B calculates, for each of the twelve calendar months, latitude/longitude grids containing surface LER and DLER fields. These L3 fields are based on the intermediate S5P LER L2 data from the respective month, collected over the entire mission. To explain this more clearly, if in April 2021 the situation is such that S5P LER L2 measurements are available for March 2019, March 2020, and March 2021, then all the data from these three “March” months are used to create the “March” S5P LER L3 product. The conversion from scene LER to surface LER and surface DLER, which requires cloud and aerosol screening, and the analysis and handling of the viewing geometry dependence of the surface DLER are described in section 6.3.

Finally, component C combines the S5P LER L3 grids created by component B, and applies various post-processing corrections, such as a correction for residual cloud contamination over the oceans and the filling up of empty grid cells by donor cells. Component C is described in section 6.4. The end result is the TROPOMI surface DLER climatology, contained in one file, providing the surface LER and DLER for all wavelength bands, for all twelve calendar months, as a function of the viewing geometry.

Having the processing chain consist of separate steps (A/B/C) is the most efficient approach. It allows reprocessing from any of the three stages within the processing chain, where the first stage by component A is extremely computationally demanding, where the second stage by component B is less computationally demanding but needs to repeated at least every year, and where the last stage is not computationally demanding but needs to be repeated every month for as long as TROPOMI data are being delivered.

6.2 Component A: calculation of the scene LER

6.2.1 Description

The Earth reflectance, in this ATBD, is defined in the usual way as

$$R = \frac{\pi I}{\mu_0 E_0} \ , \tag{1}$$

where I is the Earth radiance reflected by the Earth atmosphere, E_0 is the incident solar irradiance at the top of the atmosphere (TOA) and perpendicular to the solar beam, and μ_0 is the cosine of the solar zenith angle θ_0. In the algorithm, the band reflectances R_λ are determined for each wavelength band centred around wavelength λ in the following manner and for each individual footprint. First the solar irradiance E_0 is ported to the wavelength grid of the Earth radiance I. This is done using the method explained in appendix B of the S5P NO2 ATBD [RD20]. For this particular algorithm, this could also be done very accurately using Akima interpolation [RD21]. However, to stay in line with other S5P retrievals the decision was made to use the

method described in the NO2 ATBD. After this, the Earth reflectance R is calculated for all detector pixels using equation (1). Next, we identify those detector pixels that lie within the wavelength ranges $[\lambda-\omega_\lambda,\lambda+\omega_\lambda]$, where $2\omega_\lambda$ is the width of the wavelength band. In most cases, the wavelength bands are 1 nm wide, i.e., we have the situation that $2\omega_\lambda = 1.0$ nm. The reflectances R of these detector pixels are then averaged using a triangular function as a weighting function. This weighting function is defined as:

$$w_i^j = \begin{cases} 1 - \dfrac{|\lambda_i - \lambda_j^{\mathrm{c}}|}{\omega_j} & , \quad \text{for } |\lambda_i - \lambda_j^{\mathrm{c}}| \le \omega_j \\ 0 & , \quad \text{for } |\lambda_i - \lambda_j^{\mathrm{c}}| > \omega_j \end{cases} \tag{2}$$

In this equation, λ_i is the wavelength of detector pixel i, λ_j^{c} is the central wavelength of wavelength band j, and $2\omega_j$ is the width of wavelength band j. The procedure is illustrated graphically in Figure 3.

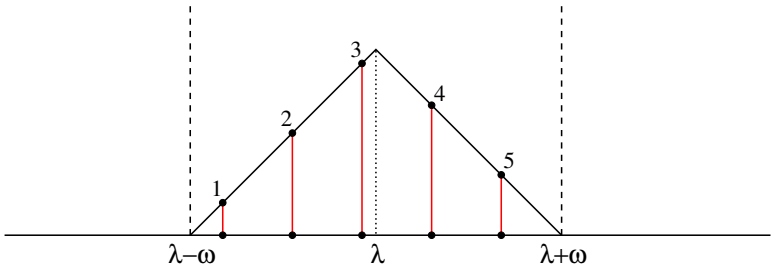


Figure 3: A simple triangular function provides the weighting factors for the detector pixels that lie within the bandwidth of the wavelength band. The procedure and need for this approach are both explained in the text.

The situation depicted in Figure 3 is that of an arbitrary (hypothetical) wavelength band centred around wavelength λ with a bandwidth of 2ω. A total of five detector pixels are found within the wavelength range, indicated with labels i ranging from 1 to 5. The height of the triangular function, indicated in red for each of the detector pixels i, provides the weighting factors w_i that are used to calculate the weighted average of the five reflectances R_i. That is, we calculate R_λ according to $R_\lambda = \sum w_i \cdot R_i / \sum w_i$. Doing so we obtain the band reflectance R_λ for all measurement footprints. This is done for each of the wavelength bands in Table 2 that is indicated with a + sign in the TROPOMI column. The triangular function stabilises the band reflectance R_λ. It makes it less susceptible to detector pixels close to the band edges that in some cases are part of the averaging and in other cases are not. This procedure, which is also used for the S5P AAI L2 product [RD22], very effectively prevents jumps along the across-track direction of the orbit swath.

Negative band reflectances R_λ are in principle not expected, but may occasionally occur. When it happens, it happens mostly for large SZA, when the Earth radiance signal I can be very low leading to small errors. If for a particular measurement the reflectance of a wavelength band is smaller than $R_{\min} = -0.05$, then the measurement is to be rejected. Suspiciously high reflectances can also occur. Reflectances can locally be larger than one, especially near the poles, but reflectances above $R_{\max} = 1.5$ are not trusted and rejected. The threshold values $R_{\min}$ and $R_{\max}$ are defined as algorithm configuration parameters (see Table 7).

Next, we calculate the values of the surface albedos that are needed to match simulated reflectances to the measured Earth reflectances. These surface albedos are in fact scene albedos, because they include the effects of surface, clouds, and aerosols. The necessary simulations assume a Rayleigh scattering atmosphere which is bounded below by a Lambertian surface. According to [RD23], the contribution of the surface to the top-of-atmosphere (TOA) reflectance may then be separated from that of the atmosphere:

$$R(\mu,\mu_0,\phi-\phi_0,A_{\mathrm{s}}) = R^0(\mu,\mu_0,\phi-\phi_0) + \frac{A_{\mathrm{s}}\,T(\mu,\mu_0)}{1-A_{\mathrm{s}}s^\star} \tag{3}$$

In equation (3), the first term is the so-called path reflectance R^0, which is the purely atmospheric contribution to the Earth reflectance. In other words, it corresponds to the TOA reflectance of the Rayleigh atmosphere when it is bounded below by a black surface that does not reflect any of the incoming radiation. The second term represents the contribution of the Lambertian surface (with surface albedo A_{s}). The parameter T is the total atmospheric transmission for the given zenith angles, $s^\star$ is the spherical albedo of the atmosphere for illumination from below, μ is the cosine of the viewing zenith angle θ, and, as before, μ_0 is the cosine of the solar zenith angle θ_0. The quantities R^0, T and $s^\star$ can be determined from radiative transfer calculations.

Using equation (3) and by demanding that the simulated Rayleigh reflectance R_λ equals the measured reflectance $R_\lambda^{\rm obs}$, we find, after some algebra, the following expression for the surface albedo $A_{\rm s}$, which is actually a scene albedo, or more specifically, the scene's Lambertian-equivalent reflectivity:

$$A_{\rm s} = \frac{R_\lambda^{\rm obs} - R_\lambda^0}{T_\lambda(\mu,\mu_0) + s_\lambda^\star (R_\lambda^{\rm obs} - R_\lambda^0)} \tag{4}$$

In this equation, R_λ^0 denotes the simulated (path) reflectance at wavelength λ, calculated for the actual atmospheric situation, but without the surface reflection contribution. The path reflectance R^0 can be expanded in a Fourier series. In our case, with a simple Rayleigh atmosphere, this expansion is exact with only three terms in the azimuth angle difference $\phi - \phi_0$. That is, the expansion of R^0 just reads

$$R^0 = a_0(\mu,\mu_0) + 2a_1(\mu,\mu_0)\cos(\phi-\phi_0) + 2a_2(\mu,\mu_0)\cos 2(\phi-\phi_0)\ . \tag{5}$$

The idea of component A is that with look-up tables (LUTs) of a_0, a_1, a_2, T, and $s^\star$ we can easily calculate R_λ^0 using equation (5) and then $A_{\rm s}$ using equation (4). The advantage of the above approach is that both the azimuthal dependence and the dependence on surface albedo are treated analytically, and are therefore not part of the LUTs. Interpolation over the remaining parameters is necessary. In this case we have to interpolate over μ and μ_0, surface elevation $z_{\rm s}$, ozone column Ω, and water vapour column Ψ. Linear interpolation is sufficient for this, because the number of nodes in the look-up tables is high enough for linear interpolation.

6.2.2 Look-up tables

The look-up tables (LUTs) are created using the radiative transfer code DAK, which is an acronym of "Doubling-Adding KNMI" [RD24, RD25]. This vector radiative transfer model (RTM) takes polarisation into account, as well as absorption by various trace gases and Lambertian surface reflection. The DAK simulations basically represent a homogeneous clear-sky Rayleigh scattering atmosphere which is bounded below by a Lambertian surface. We used version 3.2.0 of the DAK RTM. This version supports pseudo-spherical treatment of the Earth's atmosphere. Absorption by ozone, NO_2 and by the O_2–O_2 collision complex is included in the radiative transfer calculations, and for some of the wavelength bands also absorption by oxygen (O_2) or water vapour (H_2O) is taken into account. For most of the wavelength bands we use monochromatic calculations, which is possible because these bands are in parts of the spectrum where there is little or no absorption by trace gases. In Table 4, these wavelength bands are tagged "M" and the bandwidth information is not used in the radiative transfer calculations. The slit function of the TROPOMI instrument is then also not needed.

Monochromatic calculations:

The calculations for the monochromatic wavelength bands λ were done for three surface albedos $A_{\rm s} = \{0.0, 0.5, 1.0\}$, for a relative azimuth angle $\phi - \phi_0 = 0^\circ$, for 42×42 combinations of the zenith angle cosines μ and μ_0, for cloud-free conditions in a standard Mid-Latitude Summer (MLS) atmosphere [RD26], for 7 ozone column values $\Omega = \{50, 200, 300, 350, 400, 500, 650\}$ DU, and for 10 surface elevations $z_{\rm s}$ ranging from 0 to 9 km in 1 km steps. The variation of the surface elevation was achieved by removing an appropriate number of layers from the bottom of the model atmosphere. Such a removal of layers affects the ozone columns to a (very small) degree, which was compensated for by scaling the entire ozone profile in such a way that the original ozone column value was reinstated. The absorption cross-sections that are used are from [RD27] (O_3), from [RD28] (NO_2), and from [ER6] (O_2–O_2). The calculations are fast, because they are monochromatic.

The Fourier coefficients a_0, a_1, and a_2, as defined in equation (5), were delivered directly by the DAK code from the runs with zero surface albedo ($A_{\rm s} = 0.0$). The parameters T and $s^\star$ were calculated from the simulated DAK reflectances $R_\lambda(\mu,\mu_0,A_{\rm s})$, calculated for the three surface albedos $A_{\rm s}$ mentioned before, by applying equation (3) to each of the three DAK reflectances and adding or subtracting the three equations in the right manner and order. This yields, after some algebra, the following outcome:

$$s_\lambda^\star = \frac{R_\lambda(\mu,\mu_0,1.0) - 2R_\lambda(\mu,\mu_0,0.5) + R_\lambda(\mu,\mu_0,0.0)}{R_\lambda(\mu,\mu_0,1.0) - R_\lambda(\mu,\mu_0,0.5)}\ , \tag{6}$$

which is – by definition – independent on μ and μ_0, but dependent on surface elevation $z_{\rm s}$, ozone column Ω and wavelength λ, and

$$T_\lambda(\mu,\mu_0) = (1 - s_\lambda^\star)\cdot\Big(R_\lambda(\mu,\mu_0,1.0) - R_\lambda(\mu,\mu_0,0.0)\Big)\ , \tag{7}$$

wavelength band	328	335	340	354	367	380	388	402	416	425	440
instrument channel	3	3	3	3	3	3	3	4	4	4	4
central wavelength (nm)	328.0	335.0	340.0	354.0	367.0	380.0	388.0	402.0	416.0	425.0	440.0
bandwidth (nm)	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
spectral/monochromatic	S	M	M	M	M	M	M	M	M	M	M
ozone absorption	+	+	+	+	+	+	+	+	+	+	+
NO_2 absorption	+	+	+	+	+	+	+	+	+	+	+
O_2–O_2 absorption	+	+	+	+	+	+	+	+	+	+	+
oxygen absorption	–	–	–	–	–	–	–	–	–	–	–
water vapour absorption	–	–	–	–	–	–	–	–	–	–	–
wavelength band	463	494		670	685	697	712	747	758	772	2314
instrument channel	4	4		5	5	5	5	6	6	6	7
central wavelength (nm)	463.0	494.0		670.0	685.0	696.97	712.7	747.0	758.0	772.0	2314.0
bandwidth (nm)	1.0	1.0		1.0	1.0	0.3	0.3	1.0	1.0	1.0	0.5
spectral/monochromatic	M	M		M	M	S	S	M	S	S	M
ozone absorption	+	+		+	+	+	+	+	+	+	+
NO_2 absorption	+	+		+	+	+	+	+	+	+	+
O_2–O_2 absorption	+	+		+	+	+	+	+	+	+	+
oxygen absorption	–	–		–	–	+	+	–	+	+	–
water vapour absorption	–	–		–	–	+	+	–	–	–	–

Table 4: Definition of the wavelength bands. The reflectance simulations are performed spectrally or monochromatically. Absorption by oxygen and/or water vapour is added for some of the wavelength bands.

which is dependent on μ and μ_0, surface elevation z_s, ozone column Ω, and wavelength λ. The bandwidth information shown in Table 4 is not used in the calculations. The TROPOMI slit function is also not used.

Spectral calculations:

For the wavelength bands at 328, 697, 712, 758 and 772 nm high-resolution spectral calculations involving the TROPOMI slit function are to be used (see Table 4). Figure 4 explains why for some wavelength bands monochromatic calculations suffice, while for other wavelength bands spectral calculations are needed. The black curves are reflectance spectra calculated using DAK with a spectral resolution of 0.01 nm. These spectra were calculated for clear-sky conditions, for a surface albedo $A_s = 0.5$ at sea level ($z_s = 0$ km), for nadir view and local noon ($\theta = \theta_0 = 0^\circ$), for an ozone column Ω of 350 DU, and for a water vapour column Ψ of 4.0 g/cm^2. The oxygen (O_2) and water vapour (H_2O) absorption cross-sections are taken from the HITRAN2008 database [RD29]. For comparison, the horizontal green curves, when present, represent the reflectance spectra without taking absorption by oxygen and/or water vapour into account in the radiative transfer calculations.

The dotted vertical lines indicate the centre and the edges of the wavelength bands. The vertical green lines indicate the wavelengths λ_i of the detector pixels that contribute to the reflectance of the wavelength band. The blue curves represent the effective slit function of the combined detector pixels, in arbitrary units. The slit functions were calculated from the TROPOMI Instrument Spectral Response Function (ISRF) dataset [ER7]. For the wavelength bands at 670, 758 and 2314 nm the impact of absorption by the indicated trace gases is obviously very small. On the other hand, for the wavelength band at 328 nm the variation of absorption by ozone over the width of the wavelength band prevents using a monochromatic approach. For the wavelength band at 697 nm absorption by water vapour influences the band reflectance slightly but not insignificantly.

To proceed, we first define $G_j(\lambda)$, the spectral response function of wavelength band j, as a weighted superposition of the slit functions $S_i(\lambda)$ of the individual detector pixels by using the triangular weighting function defined in equation (2). That is, for the response function G_j we have

$$G_j(\lambda) = \frac{1}{w^j_{\text{tot}}} \sum_{i=1} w^j_i \cdot S_i(\lambda) \ , \tag{8}$$

where $S_i(\lambda)$ is the normalised slit function of detector pixel i from the appropriate TROPOMI spectral band and

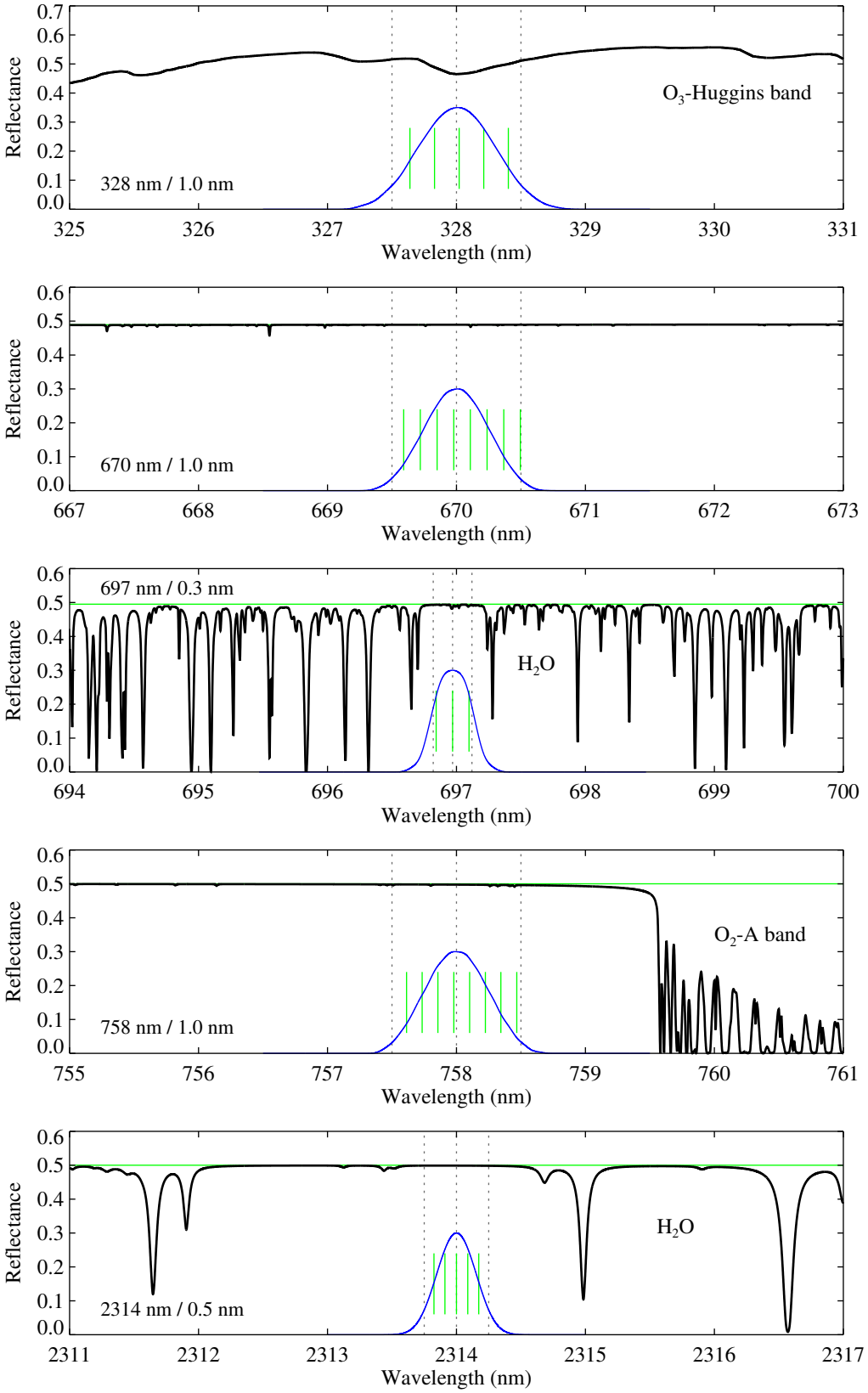


Figure 4: Reflectance spectra showing the impact of absorption by trace gases in the neighbourhood of the wavelength bands. In the case of 670 and 2314 nm the impact is negligible. For 328 and 697 nm this impact has to be taken into account in the radiative transfer calculations. Further details are provided in the main text.

$w^j_{\rm tot}$ is the sum of the weighting factors w^j_i. The resulting response functions G_j are presented in Figure 4 as the blue curves, in arbitrary units. Next, we calculate the simulated band reflectance. For this we first need to simulate the spectrum surrounding the wavelength band at a high spectral resolution. We use a spectral resolution of 0.01 nm, as was done for the spectra shown in Figure 4. The spectral sampling is then increased by a factor of 100 using Akima interpolation [RD21]. This allows for accurate numerical integration and the resulting expression for the simulated band reflectance is

$$R^{\rm sim}_j = \sum_k \Delta\lambda \cdot G_j(\lambda_k) \cdot R^{\rm sim}(\lambda_k) \ , \tag{9}$$

where the summation over k involves a summation over the wavelengths λ_k and $\Delta\lambda$ = 10^{-4} nm.

From Figure 4 one can conclude that it is not really necessary to perform spectral calculations for the 758-nm wavelength band. This is true, and the same actually applies also to the 772-nm wavelength band (not shown in Figure 4), where the impact of absorption (by oxygen) is also negligible. However, the amount of oxygen in the atmosphere is well known, requires no extra input, and any improvement is welcome so it was decided to perform spectral calculations for these two wavelength bands as well (see Table 4).

In the monochromatic case the calculations were done for a relative azimuth angle $\phi - \phi_0$ of 0°. Here the calculations are performed for three relative azimuth angles: $\phi - \phi_0 = \{0^\circ, 90^\circ, 180^\circ\}$. By applying equation (5) to the three cases one can find that the following equations are valid:

$$a_1 = \big(R_\lambda(\mu,\mu_0,0^\circ) - R_\lambda(\mu,\mu_0,180^\circ)\big)/4 \tag{10}$$
$$a_0 = \big(R_\lambda(\mu,\mu_0,0^\circ) + R_\lambda(\mu,\mu_0,90^\circ)\big)/2 - a_1 \tag{11}$$
$$a_2 = \big(a_0 - R_\lambda(\mu,\mu_0,90^\circ)\big)/2 \tag{12}$$

In these equations, all reflectances $R(\mu,\mu_0,\phi-\phi_0)$ relate to cases for which $A_{\rm s} = 0.0$. The calculation of the parameters T and $s^\star$ can be done as before using equations 6 and 7 and the R_λ calculated from equation (9) for the three surface albedo values. Here we use the DAK calculations for $\phi - \phi_0 = 90^\circ$, but this is an arbitrary decision, because T and $s^\star$ are independent on $\phi - \phi_0$.

The monochromatic and spectral results are combined into one LUT file. The LUT contains the parameters a_0, a_1, a_2, T, and $s^\star$ for each of the wavelengths. All parameters except $s^\star$ are prepared as a function of μ and μ_0, surface elevation, ozone column, and water vapour column. The parameter $s^\star$ does not depend on μ and μ_0, and is given as a function of surface elevation, ozone column, and water vapour column. The LUT file is provided in NetCDF format. Sufficient explanation about the contents is provided in the NetCDF file itself.

6.2.3 Cloud information

For information on the cloudiness of the scenes we make use of the S5P NPP-VIIRS L2 product. The S5P NPP-VIIRS product is based on observations performed by the VIIRS instrument onboard the Suomi NPP satellite. The S5P and NPP satellites are kept in a “loose” formation in very similar orbits, resulting in only a relatively small time difference (currently the difference is approximately 3 minutes, reduced from 5 minutes at the start of the mission). The VIIRS cloud information is provided for each individual TROPOMI footprint. More precisely, the product provides cloud information for four sets of boxes (0,1,2,3) surrounding the TROPOMI measurement footprint. Using the smallest surrounding box size (0) provides cloud information precisely for the TROPOMI footprint which can be used to find isolated cloud-free footprints in cloudy surroundings. Using the largest surrounding box size (3) for cloud screening assures that also the surroundings of the cloud-free TROPOMI footprint are cloud-free, which effectively removes cloud shadows at no extra cost.

The S5P NPP-VIIRS cloud product reports for each TROPOMI footprint the number of VIIRS observations which were confidently clear, probably clear, probably cloudy, and confidently cloudy [RD30]. The respective fields in the S5P NPP-VIIRS product are vcm_confidently_clear, vcm_probably_clear, vcm_confidently_cloudy, and vcm_probably_cloudy. In these fields, the numbers are reported for four different definitions of (artificial) FOV inside the cloud product [RD30]. From all possible definitions of geometrical cloud fraction that can be made, the following three typical definitions of cloud fraction were adopted and included in the L2 product:

$$c_{\rm f,strict} = \frac{N_{\rm c.cld} + N_{\rm p.cld}}{N_{\rm c.clr} + N_{\rm p.clr} + N_{\rm c.cld} + N_{\rm p.cld}} \qquad \text{(FOV: 3)} \tag{13}$$
$$c_{\rm f,nominal} = \frac{N_{\rm c.cld} + N_{\rm p.cld}}{N_{\rm c.clr} + N_{\rm p.clr} + N_{\rm c.cld} + N_{\rm p.cld}} \qquad \text{(FOV: 0)} \tag{14}$$
$$c_{\rm f,relaxed} = \frac{N_{\rm c.cld}}{N_{\rm c.clr} + N_{\rm p.clr} + N_{\rm c.cld} + N_{\rm p.cld}} \qquad \text{(FOV: 0)} \tag{15}$$

The cloud fractions found this way provide important information about the amount of cloud cover present in the observed scenes. This information is needed by component B to be able to perform cloud screening. Component B uses the "relaxed" definition of cloud fraction as given in equation (15). Separate S5P NPP-VIIRS products are available for spectral band 3, 6, and 7. This means that there is collocated cloud information available for all TROPOMI spectral bands that are needed by the algorithm (3–4–5–6–7).

6.2.4 Cloud shadow detection

Cloud shadow was not a serious problem for earlier surface albedo databases based on, for instance, SCIAMACHY or GOME-2, because of the large footprint sizes of the measurements. The fraction of the footprint area covered in shadow is relatively low for these instruments. For TROPOMI, with a relatively small footprint size of 5.6×3.6 km^2, cloud shadows can significantly reduce the reflectance and can lower the retrieved scene LER such that it can even become negative [RD31, Fig. 5]. To be able to filter out observations affected by cloud shadows we implemented the cloud shadow detection algorithm by [RD31]. This algorithm is based on a two-step approach. In the first step, a potential cloud shadow flag (PCSF) is calculated. This flag is based on the geometrical situation at hand: the position of the Sun, the viewing direction, and the height of the cloud responsible for the cloud shadow. The PCSF is designed to filter out cloud shadows, but it does it very rigorously, thereby throwing away too many observations. For that reason, we do not use it as a direct filter, but use it in combination with the second step of the cloud shadow filtering approach.

In this second step, the spectral cloud shadow flag (SCSF) is calculated. This is basically a contrast parameter which is based on the measured scene LER value and the expected surface DLER value:

$$\Gamma(\lambda) = \frac{A_{\text{scene}}(\lambda) - A_{\text{DLER}}(\lambda)}{A_{\text{DLER}}(\lambda)} \times 100\,\%\,. \tag{16}$$

According to [RD31], observations with Γ smaller than –15% are most likely affected by cloud shadows. The cloud shadow detection approach proposed in [RD31] flags observations for which (1) the PCSF was raised while at the same time (2) the contrast parameter Γ is smaller than –15%. This works well, but the complication here is that we have to use the DLER as input for the DLER calculation itself. This hurdle is by-passed by using a version of the DLER created without cloud shadow filtering. As a result, only the strongest cases of cloud shadow are removed. For future updates, the DLER input to equation (16) will remain to be the previous version of DLER, so that the input DLER for equation (16) will increase steadily in quality.

The cloud shadow flag is not determined for each of the wavelength bands. It is only determined for the longest wavelength in each of the three band duos: 494 nm for band duo 3/4, 772 nm for band duo 5/6, and 2314 nm for band duo 7/8. For the cloud information needed by the cloud shadow detection algorithm (cloud fraction and cloud height) we use the TROPOMI FRESCO cloud product for band duo 5/6. For band duo 3/4 the information is taken from the NO2 product. For band duo 7/8 the values are determined from the FRESCO product using spatial interpolation from the band 5/6 footprints to the band 7/8 footprints. The contrast threshold Γ_0 is defined in Table 7. Other algorithm settings that are used were taken from [RD31].

6.2.5 Intermediate L2 product

The output of component A is written to an intermediate L2 product, the L2 scene LER product. This is done for each TROPOMI orbit, following the frequency scheme shown in Table 3. Contents of the output product is specified in section 8.2. The scene LER found in the L2 product should in general not be used as a surface albedo, at least not before it is clear that the scene is absolutely cloud-free and – preferably – also aerosol-free. If these two conditions are met, then the scene LER can be used as a surface albedo. Moreover, in this case it most likely more suited for the situation at hand than the climatological value of the DLER database.

6.3 Component B: from scene LER to surface LER

6.3.1 Cloud and aerosol filtering

Most of the scenes for which the scene LER was calculated by component A contain clouds and/or aerosols which obstruct the view on the surface. As a result, in most cases the retrieved scene LER is not representative for the surface LER. To solve this issue, the retrieval algorithm needs to filter out scenes containing clouds and/or aerosols. Cloud screening can be achieved in two ways:

- The “statistical” approach does not require cloud information as input. It tries to identify cloud-free scenes based on statistical analyses. This has many advantages and the retrievals of surface LER from the TOMS, GOME-1, OMI, SCIAMACHY, and GOME-2 instruments have all shown that this method works well. The main advantage of the statistical approach is that cloud information from an external cloud product is not needed and as a result it does not depend on the quality, stability, and existence of this external product. The statistical approach can also be the backup approach in case the external cloud product is not available, for instance in case of loss of the respective satellite instrument.
- Cloud filtering can also be done using an external cloud product, in this case the S5P NPP-VIIRS L2 cloud product (see section 6.2.3). This “direct cloud filtering” approach works well for TROPOMI, because the S5P NPP-VIIRS information is collocated well w.r.t. the TROPOMI observations. The approach has the advantage that it is immediately clear whether a scene contains clouds or not, and this reduces the amount of processing power that is needed, as well as the amount of the storage needed for the intermediate products. It also simplifies the algorithm w.r.t. the traditional statistical approach.

For the algorithm described in this ATBD we adopt the direct cloud filtering approach, making use of the NPP-VIIRS geometrical cloud fraction $c_{f,relaxed}$ that was defined in equation (15). The threshold value $c_{f,max}$ that is used to filter out cloudy scenes is defined in Table 11.

Aerosol screening is performed using the well-known Absorbing Aerosol Index (AAI) product [RD32, RD33, RD16]. We use a relatively high threshold value r_{max} of 2.0 index points (see Table 11). This only removes scenes containing the highest concentrations of absorbing aerosol.

Next to the active cloud and aerosol screening methods described above, the algorithm also bases its results only on the lowest clear-sky reflectances (see section 6.3.3). This is done to be able to cope with unexpected outliers, and automatically screens out scenes that still contain clouds or that contain scattering aerosols. Note that this approach is essentially adopted from the statistical approach.

6.3.2 Cloud shadow filtering

Cloud shadows produced by clouds in neighbouring scenes are detected by the cloud shadow detection algorithm described in section 6.2.4. Filtering is performed for each wavelength band in a band duo whenever the cloud shadow flag was set for the longest wavelength in the band duo. Cloud shadows are removed in almost all cases, at least to the level that they can no longer be detected by eye.

6.3.3 Gridding

The Earth’s surface is described as a longitude/latitude grid containing grid cells of 0.125 by 0.125 degrees. The TROPOMI measurements are subjected to pre-screening: scenes with too high values for the geometrical cloud fraction and scenes with high amounts of aerosols are filtered out (see section 6.3.1). Scenes which were affected by a solar eclipse are also removed. The removal of such scenes is absolutely essential. Only measurements from the ascending orbit part are allowed to participate. Measurements taken in the descending orbit part are filtered out. The maximum solar zenith angle that is allowed is 85° and the maximum allowed viewing zenith angle is 70°. The many remaining TROPOMI footprints are distributed over the longitude/latitude grid, using the centre longitude and latitude to determine in which cell the measurements belong. For each grid cell the observations collected in one month are converted into the grid cell’s surface LER and into the grid cell’s surface DLER. This is described in sections 6.3.4 and 6.3.5, respectively.

6.3.4 Determining the surface LER

Two surface LER fields are constructed: one field which represents the Earth’s surface in the absence of snow/ice (“clear”) and a second field which represents the Earth’s surface in the presence of snow/ice (“snice”).

CLEAR field:

The surface LER is determined for each wavelength band and for each grid cell by taking the 10% lowest values of the (clear-sky) scene LER values that were collected. The idea behind this is that these lowest LER values correspond to the lower envelope of the scene LER distribution, which is populated only by clear-sky scenes because clouds tend to increase the Earth reflectance and, therefore, the scene LER. The approach of selecting the lower envelope of the scene LER distribution is illustrated in the left window of Figure 5.

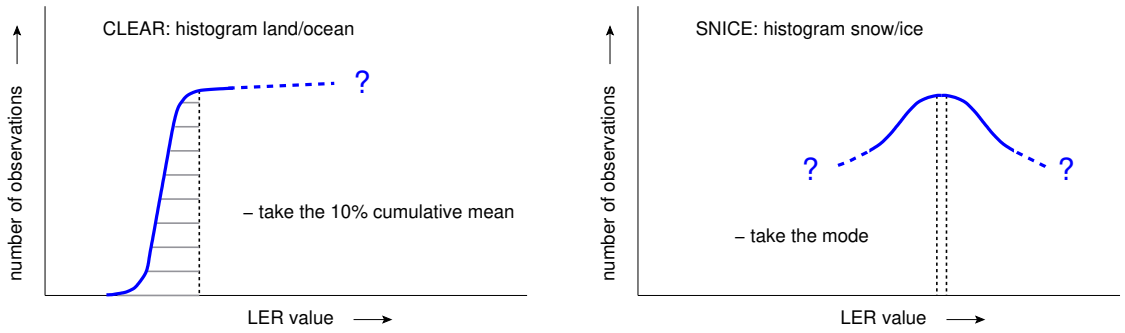


Figure 5: Approach used to calculate the surface LER for the "clear" and "snice" fields.

From the 10% lowest snow/ice-free scene LER values a weighted average is determined:

$$A_{\mathrm{LER}} = \frac{\sum_{m=1}^{M} v_m \cdot A_{\mathrm{scene},m}}{\sum_{m=1}^{M} v_m} \,, \tag{17}$$

where m is the index of the m-th measurement that is part of the M clear-sky snow/ice-free measurements that constitute the 10% lowest values of the clear-sky snow/ice-free scene LER collection. The v_m is the appropriate weighting factor of measurement m, which is determined by the precision $\delta A_{\mathrm{scene},m}$ assigned to the scene LER:

$$v_m = 1/\delta A_{\mathrm{scene},m} \tag{18}$$

The surface LER value that is calculated this way is in practice very close to the minimum scene LER value. As a result, the surface LER is by definition close to the so-called MIN-LER product as defined in, for instance, [RD6] and available in all heritage LER databases (see section 5.2.1).

SNICE field:

The surface LER is determined from (clear-sky) scenes which contain snow and/or ice. The distribution of the scene LER values in this case does not look like the one in the left window of Figure 5 but more like the one in the right window of Figure 5. That is, the scene LER distribution has a specific mode which is assumed to be the best representative value for the surface LER [RD6]. In practise, the scene LER histogram is analysed and the scenes which fall in the specific histogram bin where the mode of the distribution resides are identified. From these scenes the weighted average is determined using equation (17).

Examples:

Examples of the "clear" and "snice" fields are shown in Figures 6 and 7. These fields are raw fields in the sense that the post-processing corrections that will be discussed in section 6.4 are not applied yet. The fields were taken from the intermediate L3 surface LER product mentioned in section 6.3.6.

Figures 8 and 9 present examples of the TROPOMI surface LER database fields based on only one month of TROPOMI data (for each calendar month). In Figure 8 the surface LER is presented for the month March, for the wavelength bands at 380, 463, 758, and 2314 nm. In Figure 9 the surface LER is presented for 758 nm and for the months February, May, August, and November. Grid cells coloured in gray represent grid cells for which no surface LER could be determined. There are in practice four possible reasons for this happening:

1. persistent cloud presence for certain regions
2. grid cell in polar night or $\theta_0 > 85^\circ$
3. systematic errors in the NPP-VIIRS cloud information for certain regions
4. reflectance data systematically rejected for certain regions

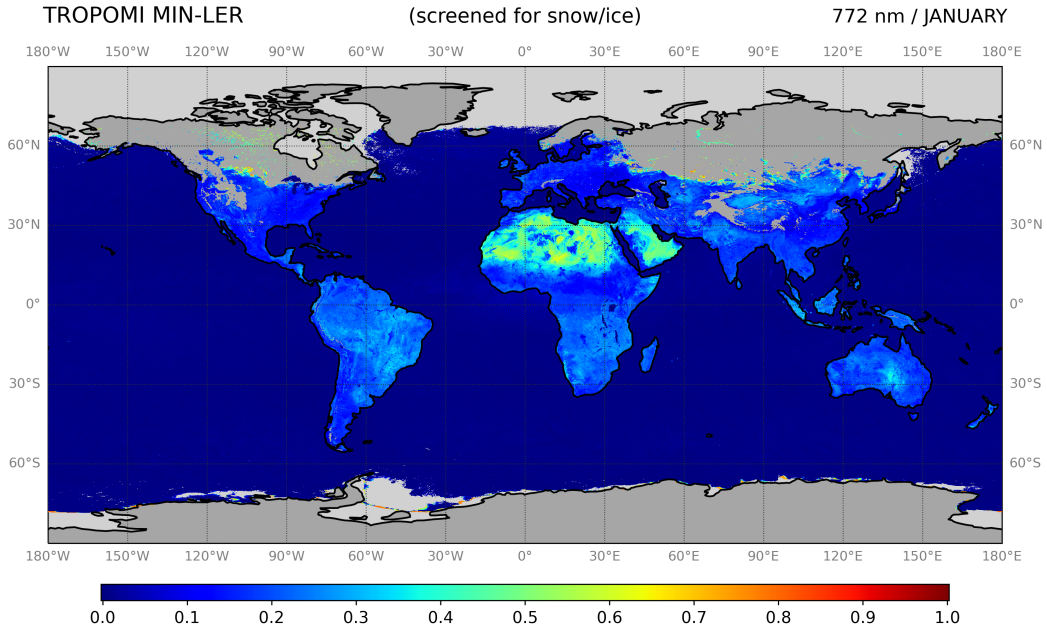


Figure 6: Global map of TROPOMI surface LER at 772 nm according to the "clear" field for the month January (from the years 2019–2023). Only snow-free and ice-free scenes were used for the retrieval. The field was taken from the intermediate L3 surface LER product, not from the final surface DLER database.

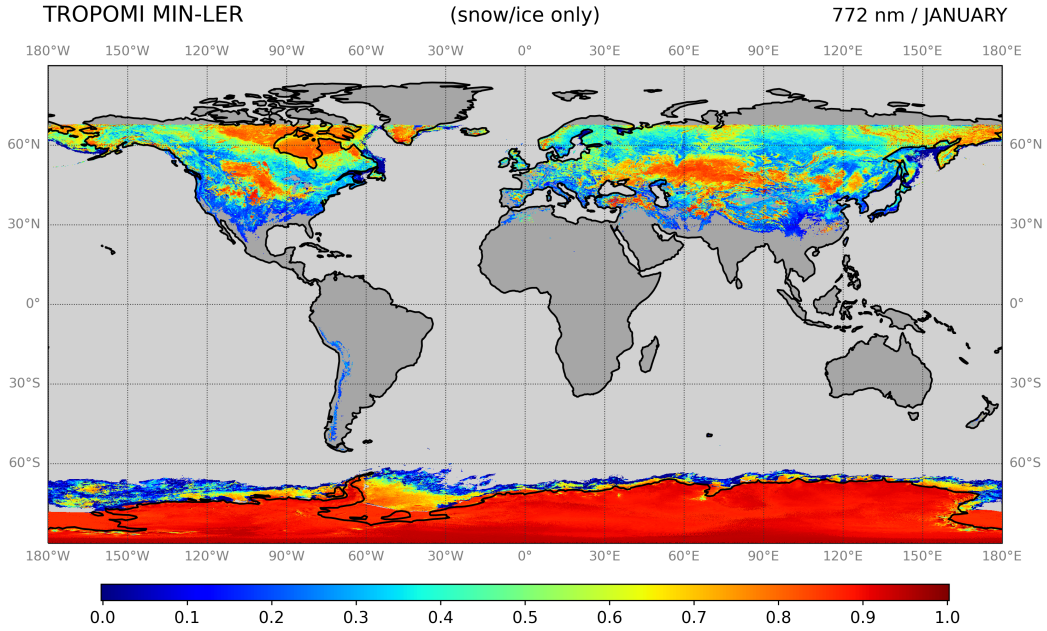


Figure 7: Global map of TROPOMI surface LER at 772 nm according to the "snice" field for the month January (from the years 2019–2023). Only scenes containing snow or ice were used for the retrieval. The field was taken from the intermediate L3 surface LER product, not from the final surface DLER database.

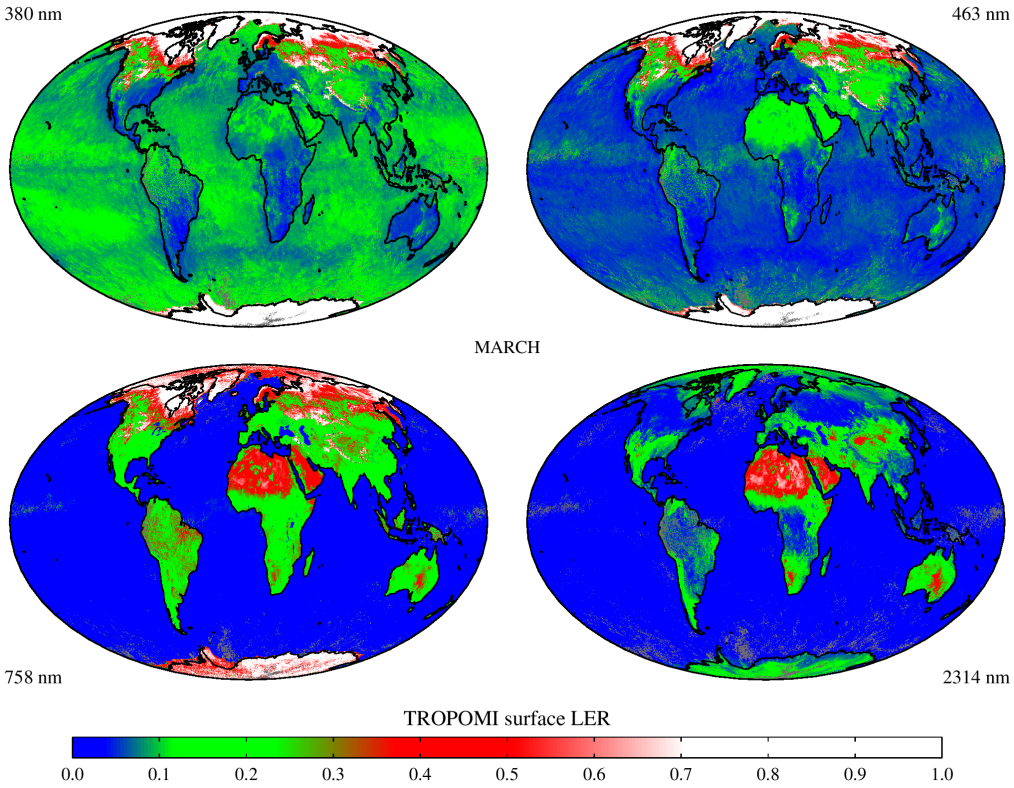


Figure 8: Global maps of TROPOMI surface LER for the month March at 380, 463, 758, and 2314 nm.

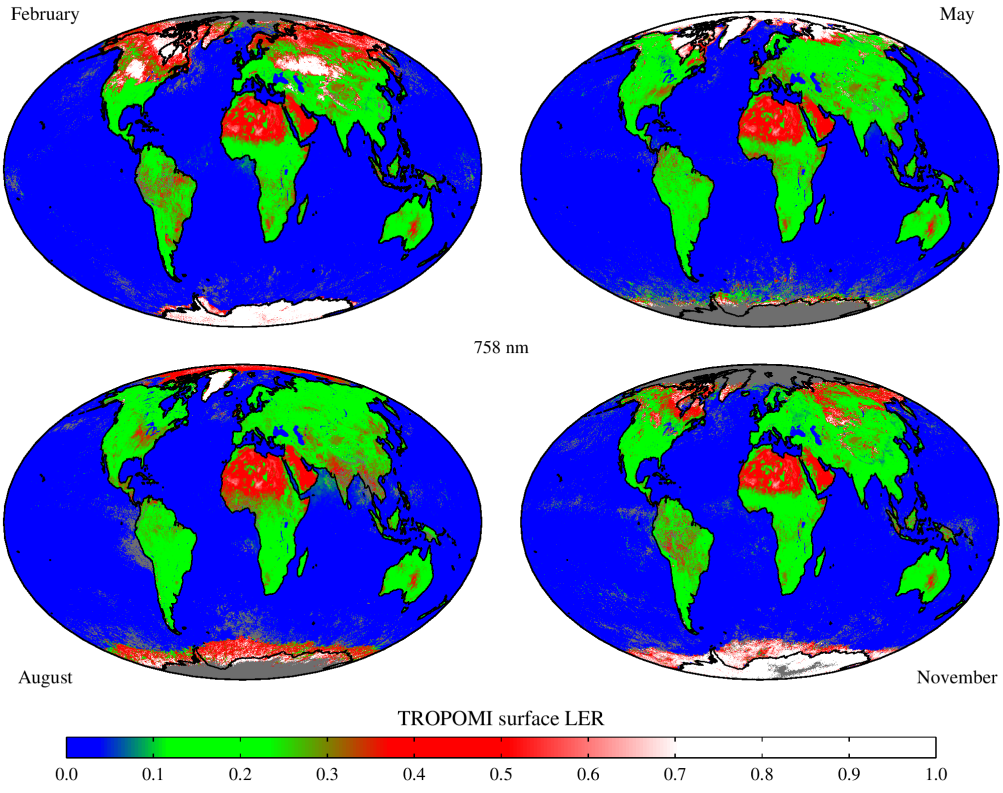


Figure 9: Global maps of TROPOMI surface LER at 758 nm for February, May, August, and November.

As for (1), persistent cloud presence is a problem only for certain areas. With only one month of data used, chances of this happening are quite large. With every year added to the TROPOMI mission, the chances of finding clear-sky scenes increase. Nevertheless, even after 5 years of data this can happen for certain areas. This issue is addressed in component C of the algorithm chain (see section 6.4.1).

The grid cell can also be contained in polar night or close it ($\theta_0 > 85°$). This (2) happens near the polar regions, for a couple of months per year (see Figure 9). A cosmetic correction is applied by component C of the algorithm chain, making use of donor grid cells from neighbouring months (see section 6.4.2).

Systematic errors in the NPP-VIIRS cloud information (3) can lead to the situation that for specific regions no clear-sky scenes are observed the entire month. Another possible reason for missing data could be systematic rejection of the reflectances for certain regions (4). Given the relaxed thresholds on R_{min} and R_{max} (see Table 7), this should not happen.

6.3.5 Determining the surface DLER

The surface LER that was defined in section 6.3.4 was based on the assumption that all surface types act as Lambertian reflectors. That is, the amount of light being reflected by the surface is assumed not to depend on the directions of the incoming and reflected light. This principle of Lambertian surface reflection is illustrated in Figure 10a. The assumption of Lambertian surface reflection is, however, not always justified. A more realistic description of the reflective properties of the surface requires a description using a bi-directional reflectance distribution function (BRDF) [RD7] which takes into account the dependence on the incoming and outgoing directions. This situation is depicted in Figure 10b. The (very) hypothetical surface BRDF shown contains a specular lobe, resulting from specular reflection by water in the forward scattering direction, and a retroreflection lobe, corresponding to reflection by vegetation in the backscattering direction.

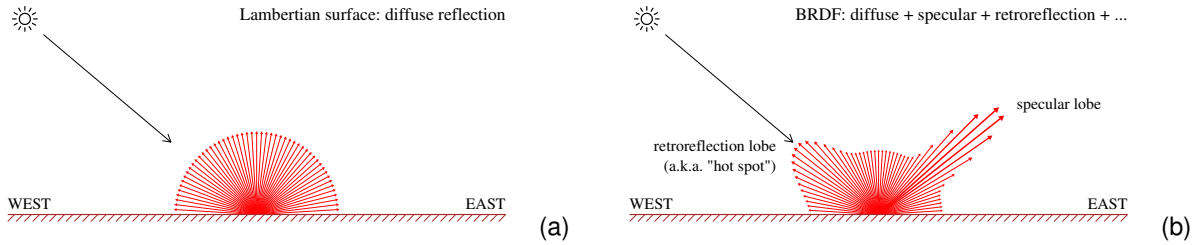


Figure 10: Lambertian surface reflection (a) versus a more realistic situation of surface reflection described by a surface BRDF (b). Unlike the Lambertian model, which does not depend on the exact scattering geometry, the BRDF depends on the directions of the incoming and reflected light.

During each calendar month, a certain geographical location is observed by the TROPOMI instrument from a variety of viewing geometries. This situation is explained graphically in Figure 11a. Each month, a certain location is observed on average 30 times, or more if the location is closer to one of the poles. The full range in viewing angles is about 134 degrees (see Figure 11b). The surface BRDF sketched in Figure 11 contains a “hot spot” of increased reflectivity near backscattering geometries, which is typical for vegetated surfaces. The reflectivity of the surface is observed by TROPOMI from viewing directions ranging from West to East. Obviously, for the situation sketched in Figure 11, in a “west” viewing direction the observed surface albedo is lower than in an “east” viewing direction (i.e., near the “hot spot”).

In the data processor, the dependence of the non-directional surface LER on the viewing angle is analysed for every grid cell. For this to lead to a reliable result, enough clear-sky scenes need to be available. Note that a certain geographical location near the equator is only observed roughly 30 times per year for a certain calendar month. Only part of these are clear-sky observations. These observations should span the entire viewing angle range. To achieve this, enough clear-sky observations need to be available, which in general is not possible with only one month of data. Multiple years of data are therefore needed so as to have clear-sky observations for each of the viewing directions.

After that, the surface DLER is parameterised as a function of the viewing angle (denoted as θ_{v}) in the following way:

$$A_{\mathrm{DLER}} = A_{\mathrm{LER}} + c_0 + c_1 \cdot \theta_{\mathrm{v}} + c_2 \cdot \theta_{\mathrm{v}}^2 + c_3 \cdot \theta_{\mathrm{v}}^3 \ , \tag{19}$$

where θ_{v} is negative on the east side of the orbit swath ($\theta_{\mathrm{v}} = -\theta$), and positive on the west side of the orbit swath ($\theta_{\mathrm{v}} = \theta$). For water bodies the coefficients c_0, c_1, c_2, and c_3 are set to zero. They are also set to zero

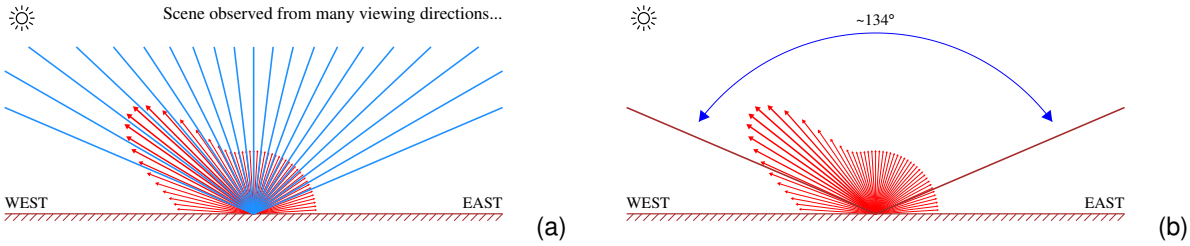


Figure 11: Left: During one month of observations, a certain location is observed many times by TROPOMI for many different viewing geometries (angles). At the equator, each location is observed roughly 30 times per month. Right: The range in viewing angles covered this way is approximately 134 degrees.

for coastal areas and for areas for which the DLER could not be retrieved. A failure to retrieve the DLER occurs when not the whole viewing angle range is covered. This happens systematically near the polar regions, for very high solar zenith angles, when part of the TROPOMI scanline has $\theta_0 < 85^\circ$ and the other part has $\theta_0 > 85^\circ$. It can also happen when a region is suffering from structural cloud presence.

6.3.6 Intermediate L3 product

The result is written to an intermediate L3 product. This is done for each calendar month, as explained in section 6.1. Component B is run every month, processing the last complete calendar month according to the scheme shown in Table 3. Contents of the output product is specified in section 8.4.

6.4 Component C: post-processing steps

Component B creates intermediate L3 files containing the retrieved surface LER fields. The surface LER fields in these intermediate files are not ready to be used. First of all, some of the grid cells over the ocean need to be corrected for residual cloud contamination. Secondly, for regions in polar night no surface information can be retrieved. This results in a polar gap in the surface LER fields for most of the calendar months. Thirdly, in the "clear" field (see Figure 6) there are sometimes gaps because of persistent snow/ice presence during the entire month. Likewise, in the "snice" field (see Figure 7) there will be gaps because of lack of snow/ice presence. Fourthly, other types of gaps (missing data) can also occur, especially at the start of the mission, when the number of TROPOMI mission years is still limited. These and other issues are addressed by component C.

6.4.1 Cloud contamination

For grid cells located over the oceans, the data processor is constantly looking for cloud-free scenes from which to determine the surface LER. Sometimes, for certain grid cells, and despite having many years of data available, a cloud-free scene is never observed. This happens for regions which are known to be suffering from persistent clouds. As an example, Figure 12a presents the TROPOMI surface LER at 758 nm for the calendar month June, based on only one month of data. There are suspiciously high values found for quite a few regions. Gaps are also present (in the colour gray). The values that are suspiciously high are a clear indication that for these cells, the retrieved surface LER is contaminated by the presence of clouds.

In the data processor, we correct for this effect by looking for donor cells which were not cloud contaminated. The process starts with identifying the contaminated cells. This is done for band duos 3/4, 5/6, and 7/8 separately. For example, for band duo 5/6 we use the retrieved 772-nm surface LER in combination with a threshold $A_{\mathrm{max},772}$ (see Table 14) to find the cloud contaminated grid cells over the ocean. The resulting mask is shown in Figure 12b, where a blue colour indicates a contaminated grid cell. After that, a correction is performed for each of the contaminated cells by finding a donor cell in the vicinity of the contaminated cell. This donor cell is searched in a box around the contaminated cell with dimensions of 10° latitude and 30° longitude. However, for contaminated cells in the region near the equator where the absolute latitude is less than 30°, the longitude range of the box is extended to 60°. The donor cell is the ocean cell in the box with the lowest retrieved 772-nm surface LER. The effect of the correction is demonstrated by Figure 12d.

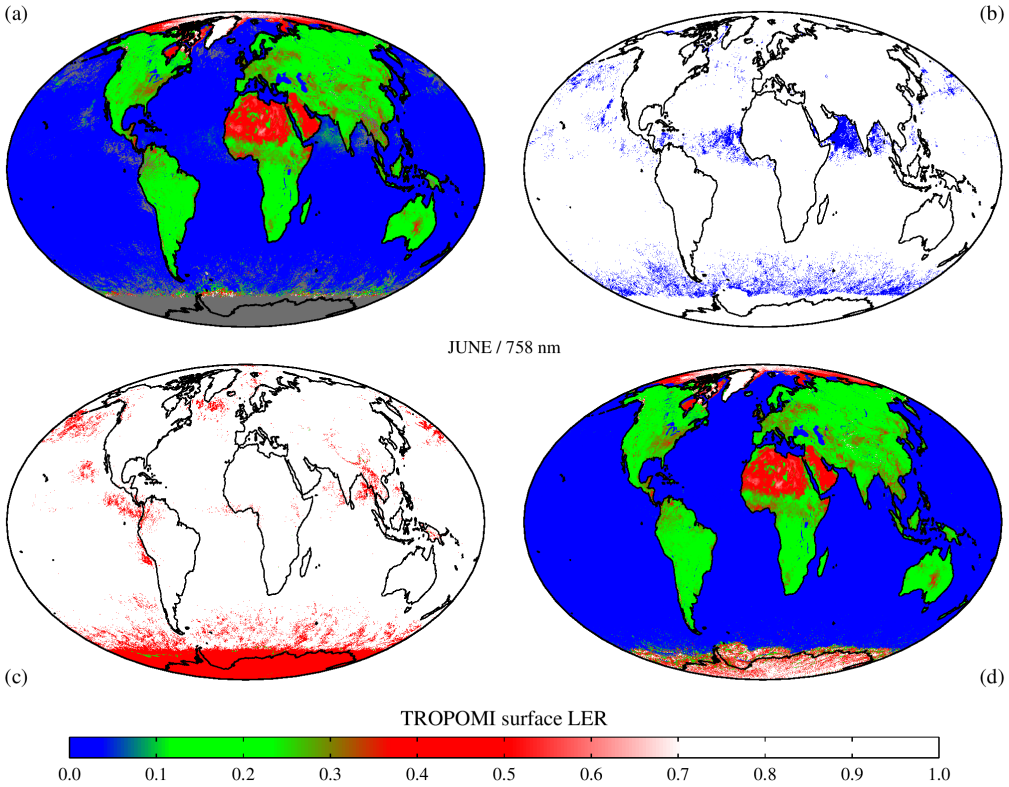


Figure 12: Some intermediate processing results for the month June, based on only one month of data (June 2018). (a) LER field at 758 nm showing gaps and residual cloud contamination. (b) Detection of residual clouds. (c) Detection of missing data and gaps. (d) Surface LER map after replacing cloud contaminated grid cells with near-by donor grid cells and filling the gaps near the polar regions with data from other months.

6.4.2 Gaps and missing data

The next processing step deals with missing data. The surface (D)LER grids inside the intermediate L3 files are only filled for the sunlit part of the globe (for which $\theta_0 < 85°$, see section 6.3.3). In certain cases no data remained after cloud screening, resulting in missing data. Then there will be the normal gaps in the “clear” and “snice” fields (see Figure 6 and 7). The occurrence of gaps and missing data is indicated in Figure 12c by red or green grid cells. To be able to offer the users of the TROPOMI surface DLER database also meaningful data for the missing parts of the grid, grid cells with no data are filled with surface (D)LER values from other months for which the grid cell did contain a trustworthy value.

In Figure 12c the red colour indicates that a suitable donor cell could be found. A green colour indicates that no suitable donor cell could be found. Near the edge of the Antarctic ice field some green grid cells are indeed present in Figure 12c. It is known from experience with the GOME-2 surface LER database that when a sufficient amount of data is used (i.e., not just one month of data as used for Figure 12) the amount of grid cells with cloud contamination or without a valid result goes down dramatically. Green grid cells as in Figure 12c are therefore not expected when multiple years of data are used, and red grid cells are normally only found in polar night. The effect of the filling in of missing data is demonstrated in Figure 12d.

For the DLER, the correction for cloud contamination is automatically done, because over the ocean the DLER and LER fields are identical (see section 6.3.5). For the cases of missing data we perform for the DLER field the same type of replacement as used for the LER field.

6.4.3 End product

A quality flag is maintained that indicates the actions that were followed. The uncertainty in the surface LER values is calculated according to the method outlined in section 9. The results for all twelve calendar months are stored in the output product. Contents of the output product is specified in section 8.6.

7 DLER versus BRDF

7.1 Introduction

This section introduces a number of straightforward model comparisons between the surface DLER and the surface BRDF. The surface BRDF offers the most complete description of the surface reflection, being dependent on the viewing zenith angle θ, the solar zenith angle θ_0, and the relative azimuth angle $\phi - \phi_0$. However, using surface BRDF as input for radiative transfer calculations requires a radiative transfer code that can model directional surface reflection. In most practical situations, this is not possible, or not feasible. In such cases Lambertian surface reflection is used in the model environment. In these cases, using the DLER is the right thing to do. Using the BRDF as input in combination with Lambertian surface reflection in the radiative transfer calculations will in general lead to errors. These errors are studied in the next sections.

7.2 Approach

The DAK radiative transfer code was described in section 6.2.2. Here we make use of version 3.5.0 of the DAK code, which is able to handle surface reflection defined by a BRDF. More specifically, the user can specify the three kernel coefficients ($f_{\rm iso}$, $f_{\rm vol}$, $f_{\rm geo}$) as defined in the MODIS ATBD [RD34]. Using this setup, the TOA reflectance is calculated at a number of wavelengths, for the VZA and SZA nodes that are also part of the LUTs described in section 6.2.2, and for 360 equidistant values of the relative azimuth angle $\phi - \phi_0$. The surface elevation is set to zero (sea level) and the ozone column to 350 DU. Cloud and aerosols are not included. The calculations are performed monochromatically at 380, 463, 772, and 2314 nm.

Next, the surface DLER is retrieved from the simulated TOA reflectances using a similar setup as the one described in section 6.2. The only difference here is that the TROPOMI measurements are replaced by the reflectances calculated by DAK 3.5.0, i.e., including surface reflection described by the BRDF kernel coefficients ($f_{\rm iso}$, $f_{\rm vol}$, $f_{\rm geo}$). The differences are analysed and interpreted in the next section.

7.3 Results

The results for 380, 463, 772, and 2314 nm are presented in Figures 13 to 16. In all cases the solar zenith angle is 32°. Figure 13 presents the results for 2314 nm. The left upper window presents the BRDF, characterised by the kernel coefficients ($f_{\rm iso}$, $f_{\rm vol}$, $f_{\rm geo}$) = (0.40, 0.26, 0.04). These kernel coefficients are representative for vegetated surfaces such as forests [RD18]. The right upper window shows the TOA reflectance calculated by the DAK RTM. Note that the reflectance is pretty much identical to the BRDF. The left bottom window presents the retrieved DLER. There is no clear difference with the BRDF. This is confirmed by the bottom right window, which presents the difference between BRDF and DLER. Differences are not found.

The results are not unexpected, because at 2314 nm the Rayleigh optical thickness is very low (3×10^{-7}), so scattering in the atmosphere is unlikely to happen. This means that only scattering at the surface occurs. This scattering happens only once, because light cannot be scattered back to the surface after the first encounter with the surface. In this situation the retrieved DLER and BRDF are numerically identical.

At 772 nm (see Figure 14), the Rayleigh optical thickness is still quite low (2×10^{-2}), so scattering via the surface is still dominating but at the same time multiple surface scattering is still rare. This is confirmed by the simulations: differences between BRDF, reflectance, and DLER are again small. Only close to the "hot spot" and for large viewing zenith angles there are clear differences between BRDF and DLER. Note that we changed the kernel coefficients compared to the 2314-nm case, to make them fit better to 772 nm.

The situation changes quite a bit at 463 nm. Figure 15 shows that at this wavelength the typical value of the BRDF is much lower than at 772 and 2314 nm. Because of increased Rayleigh scattering in the atmosphere, the TOA reflectance is very different from the surface BRDF. The retrieved DLER shows quite some differences compared to the BRDF. This is caused by an increased occurrence of scattering in the atmosphere (Rayleigh optical thickness of about 0.2) and multiple scattering via the surface. The effects may seem to be modest in an absolute sense, but relative to the typical value of the BRDF (0.03) they can be quite large, depending on the viewing and solar angles that are involved. Note that the behaviour of the BRDF for large viewing zenith angles in the forward scattering direction is suspicious, because the BRDF becomes negative for solar zenith angles close to 90°. In the DAK RTM, the surface BRDF is not allowed to become negative.

Finally, we study the situation at 380 nm. Here the Rayleigh optical thickness has increased to about 0.4. Atmospheric scattering dominates, and a Lambertian description of the surface, whether it is directional or not, inevitably breaks down. In Figure 16 the increased atmospheric scattering can be seen as an increase in the reflectance w.r.t. the reflectance at 463 nm. The retrieved DLER deviates quite a bit from the BRDF,

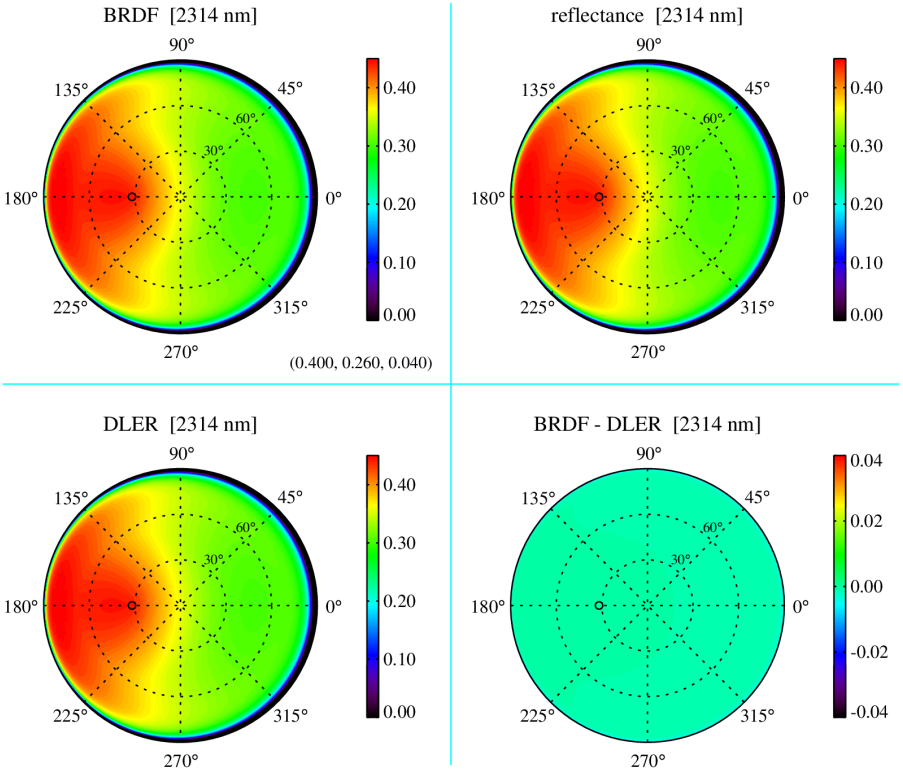


Figure 13: Top row: Surface BRDF at 2314 nm for a solar zenith angle of 32° and the resulting simulated TOA reflectance. Bottom row: Retrieved surface DLER and the difference between BRDF and DLER.

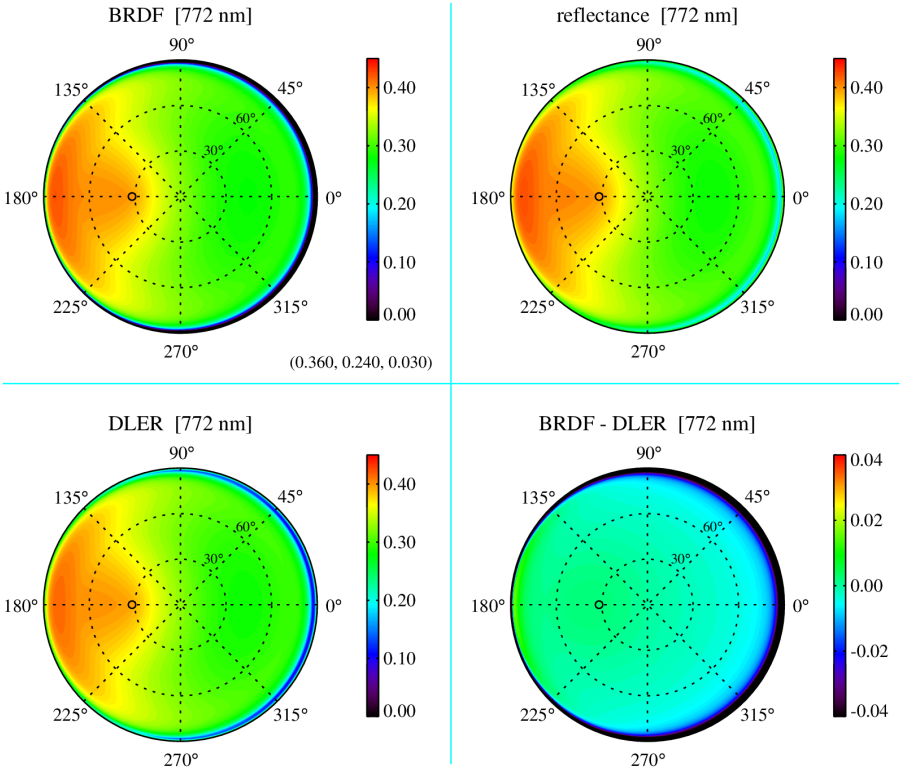


Figure 14: Similar to Figure 13 but now for 772 nm. The BRDF kernel coefficients (f_{iso}, f_{vol}, f_{geo}) that were used are indicated. They are different from the ones at 2314 nm but again representative for vegetation.

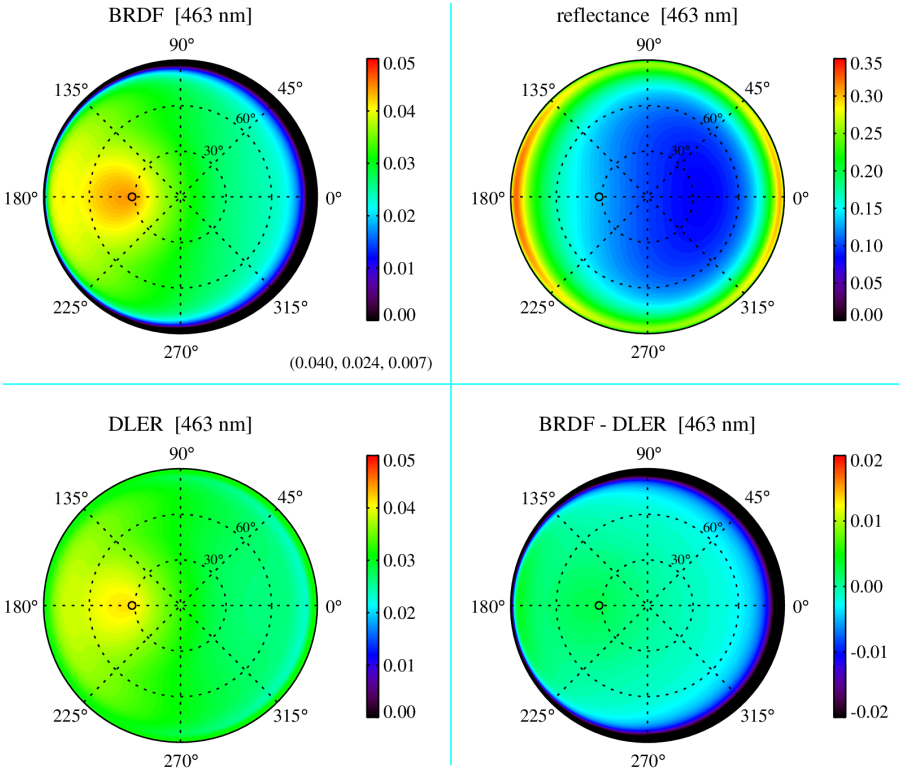


Figure 15: Similar to Figure 13 but now for 463 nm. The differences between DLER and BRDF are now larger, especially near the “hot spot” and for large viewing zenith angles.

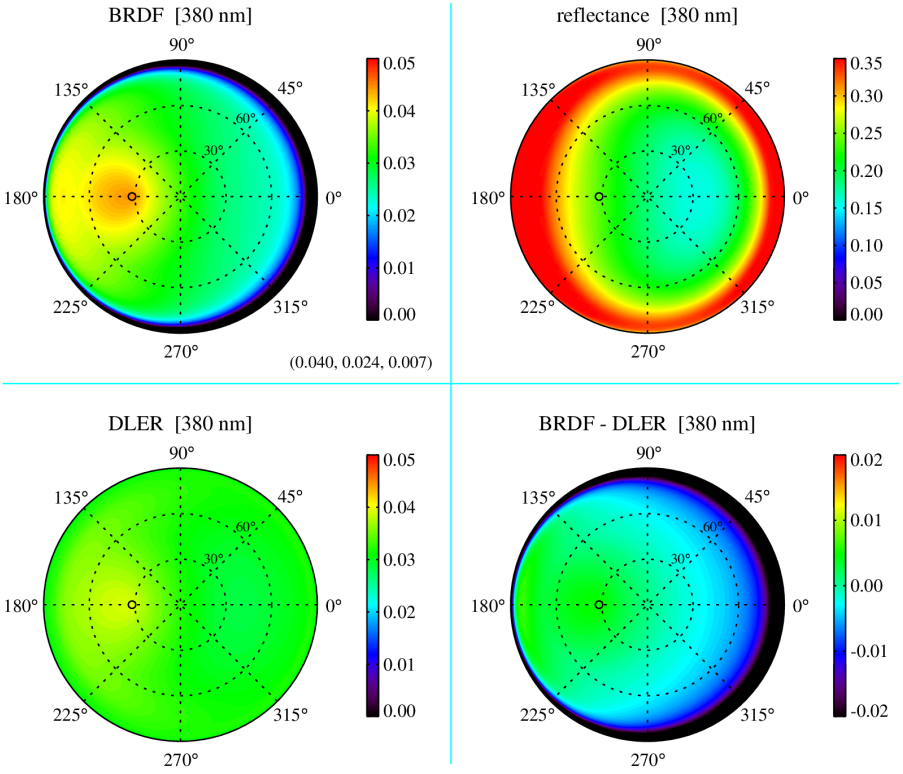


Figure 16: Similar to Figure 15 but now for 380 nm. The surface BRDF is identical to the one at 463 nm (the same kernel coefficients) but the DLER differs more from the BRDF than was the case at 463 nm.

especially in the forward scattering direction. For most moderate viewing zenith angles (i.e., $\theta < 60^{\circ}$), however, the deviations are below 0.01. This number has to be compared to the typical value of the BRDF (0.03).

7.4 Conclusion

For the NIR and SWIR wavelengths, differences between DLER and BRDF are numerically very small (example: 772 nm), or even absent (example: 2314 nm). That is, for these wavelengths the DLER can be used as a replacement for the BRDF (and vice versa). For the UV and VIS wavelengths (examples: 380 nm and 463 nm), the differences between DLER and BRDF are larger. For these wavelength bands, the DLER is not the optimal replacement for the BRDF (and the BRDF is not the optimal replacement for the DLER).

However, the DLER is not meant to be a replacement for the BRDF. The BRDF should be used in situations where the radiative transfer calculations are performed taking surface reflection according to a BRDF into account. The DLER is meant to be used in situations where the radiative transfer calculations are performed using Lambertian surface reflection. In situations where Lambertian surface reflection is used in the radiative transfer modelling, the DLER should be used. Using BRDF instead will lead to errors in the UV-VIS.

The differences that were presented in section 7.3 are not to be interpreted as shortcomings in the DLER. They represent numerical differences between two completely different parameters.

8 Feasibility

This section will list all the input and output parameters, the units that are to be used, the source of the input parameters, and what should happen when the input parameters are not available. For the output parameters also the dimensions are provided. Computation effort is also discussed, and open issues are noted.

8.1 Input – component A

8.1.1 Static input

Parameter	Symbol	Physical unit	Source	Use	If unavailable
Wavelength	λ'	[nm]	LUT	U	No processing
Surface elevation	$z_s{}'$	[km]	LUT	U	No processing
Ozone column	Ω'	[DU]	LUT	U	No processing
Water vapour column	Ψ'	[g/cm^2]	LUT	U	No processing
$\mu = \cos\theta$	μ'	–	LUT	U	No processing
$\mu_0 = \cos\theta_0$	$\mu_0{}'$	–	LUT	U	No processing
Fourier coefficient a_0	$a_0{}'$	–	LUT	U	No processing
Fourier coefficient a_1	$a_1{}'$	–	LUT	U	No processing
Fourier coefficient a_2	$a_2{}'$	–	LUT	U	No processing
Transmission	T'	–	LUT	U	No processing
Spherical albedo	$s^{\star\prime}$	–	LUT	U	No processing

Table 5: Static input parameters for component A. U=Used in computation; F=Filter; C=Copied to output.

The look-up table (LUT) is static and will normally not change during the mission life time, unless the wavelength band definition as reported in Table 2 is changed. Reprocessing of the LUT does not require much time. The origin of the LUT, the radiative transfer model that was used, and other details are described in section 6.2.2. The parameter μ is the cosine of the VZA. The array containing μ_0, which is the cosine of the SZA, is by definition identical to the array containing μ. The matrices in the LUT are all symmetrical in μ and μ_0.

8.1.2 Dynamic input

Parameter	Symbol	Physical unit	Source	Use	If unavailable
Radiance	I	[mol s^{-1} m^{-2} nm^{-1} sr^{-1}]	S5P L1B product	U	No processing
Irradiance	E_0	[mol s^{-1} m^{-2} nm^{-1}]	S5P L1B product	U	No processing
Radiance precision	δI	[mol s^{-1} m^{-2} nm^{-1} sr^{-1}]	S5P L1B product	U	No processing
Irradiance precision	δE_0	[mol s^{-1} m^{-2} nm^{-1}]	S5P L1B product	U	No processing
Radiance wavelength	λ_{rad}	[nm]	S5P L1B product	U	No processing

Parameter	Symbol	Physical unit	Source	Use	If unavailable
Irradiance wavelength	λ_{irr}	[nm]	S5P L1B product	U	No processing
Scanline	s	–	S5P L1B product	C	No processing
Ground pixel	g	–	S5P L1B product	C	No processing
Time	t	[s]	S5P L1B product	C	No processing
VZA	θ	[degrees]	S5P L1B product	U	No processing
SZA	θ_0	[degrees]	S5P L1B product	U	No processing
VAA	ϕ	[degrees]	S5P L1B product	U	No processing
SAA	ϕ_0	[degrees]	S5P L1B product	U	No processing
Centre longitude		[degrees east]	S5P L1B product	U	No processing
Centre latitude		[degrees north]	S5P L1B product	U	No processing
Corner longitudes		[degrees east]	S5P L1B product	C	No processing
Corner latitudes		[degrees north]	S5P L1B product	C	No processing
Solar eclipse flag		–	S5P L1B product	C	No processing
VCM confidently clear	$N_{\mathrm{c.clr}}$	–	S5P L2 VIIRS product	U	No processing
VCM probably clear	$N_{\mathrm{p.clr}}$	–	S5P L2 VIIRS product	U	No processing
VCM confidently cloudy	$N_{\mathrm{c.cld}}$	–	S5P L2 VIIRS product	U	No processing
VCM probably cloudy	$N_{\mathrm{p.cld}}$	–	S5P L2 VIIRS product	U	No processing
AAI (340/380 nm)	r	–	S5P L2 AAI product	C	No processing
Ozone column	Ω	[DU]	S5P L2 AAI product	U	No processing
Surface elevation (3/4)	z_s	[km]	S5P L2 AAI product	U	No processing
Surface elevation (5/6)	z_s	[km]	S5P L2 FRESCO product	U	No processing
Surface elevation (7/8)	z_s	[km]	S5P L2 CH4 product	U	No processing
Cloud fraction (3/4)	c_{eff}	–	S5P L2 NO2 product	U	No processing
Cloud fraction (5/6)	c_{eff}	–	S5P L2 FRESCO product	U	No processing
Cloud fraction (7/8)	c_{eff}	–	S5P L2 FRESCO product	U	No processing
Cloud height (3/4)	z_c	[km]	S5P L2 NO2 product	U	No processing
Cloud height (5/6)	z_c	[km]	S5P L2 FRESCO product	U	No processing
Cloud height (7/8)	z_c	[km]	S5P L2 FRESCO product	U	No processing
Water vapour column	Ψ	[g/cm^2]	AUX product (CAMS)	U	Use $\Psi = 0.1$
Land/sea mask	s_t	–	AUX product (NASA SDP Toolkit)	C	No processing

Parameter	Symbol	Physical unit	Source	Use	If unavailable
Snow/ice flag		–	AUX product (NISE)	C	No processing

Table 6: Dynamic input parameters for component A. U=Used in computation; F=Filter; C=Copied to output.

The total ozone column Ω is taken from the S5P L2 AAI product and needed for the atmospheric correction. Therefore, if the AAI product is not available, processing should be stopped. A "stop" means that the processing of the orbit is abandoned and that no output product is produced for this orbit. If, however, the value of the AAI is only out-of-bounds, processing can continue and a fill value should be provided. Surface elevation z_s is taken from the AAI product (valid for spectral bands 3/4), from the FRESCO product (valid for spectral bands 5/6), and from the CH4 product (valid for spectral bands 7/8). If the AAI and/or FRESCO product are/is not available, then processing should be stopped. If only the CH4 product is not available, then the z_s from the FRESCO product can be used instead. Note that the parameter z_s has little to no impact on the retrieved scene LER at 2.3 μm because of the low Rayleigh optical thickness of the atmosphere. Cloud fraction c_{eff} and cloud height z_c for spectral bands 3/4 are taken from the NO2 product. For spectral bands 5/6 these parameters are taken from the FRESCO product. For spectral bands 7/8 they are determined from the parameters determined for spectral bands 5/6 by using spatial interpolation. If, for some reason, the FRESCO product is not available then the parameters are determined by spatial interpolation from the NO2 product. If the S5P NPP-VIIRS cloud product is not available, or for some reason not usable, then the processing should be stopped. If the CAMS water vapour information is not available, processing should not be stopped, because only 2 of the 21 wavelength bands rely on water vapour input. Instead, a background value of 0.1 g/cm^2 should be used.

8.1.3 Algorithm configuration parameters

Parameter	Symbol	Physical unit	Meaning	Value
Band centre	λ	[nm]	Central wavelength of the wavelength bands	See Table 2
Bandwidth	$2\omega_\lambda$	[nm]	Width of the wavelength bands; see section 5.4	1.0 nm, except: 697 + 712 nm (0.3 nm) 2314 nm (0.5 nm)
Maximum SZA	$\theta_{0,\max}$	[degrees]	Maximal SZA value accepted	88.0
Minimum reflectance	$R_{\min}$	–	Minimal reflectance value accepted	–0.05
Maximum reflectance	$R_{\max}$	–	Maximal reflectance value accepted	1.5
Cloud shadow thresh.	Γ_0	–	Threshold used for Γ; see section 6.2.4	–15%

Table 7: Algorithm configuration parameters for component A.

8.2 Output – component A

Parameter	Symbol	Physical unit	Dimensions	Comments
Wavelength	λ	[nm]	[21]	

Parameter	Symbol	Physical unit	Dimensions	Comments
Scanline	s	–	[scanl. x *3*]	
Ground pixel	g	–	[pixels x *3*]	
Time	t	[s]	[scanl. x *3*]	
VZA	θ	[degrees]	[scanl. x pixels x *3*]	
SZA	θ_0	[degrees]	[scanl. x pixels x *3*]	
VAA	ϕ	[degrees]	[scanl. x pixels x *3*]	
SAA	ϕ_0	[degrees]	[scanl. x pixels x *3*]	
Centre longitude		[degrees east]	[scanl. x pixels x *3*]	Centre of footprint
Centre latitude		[degrees north]	[scanl. x pixels x *3*]	Centre of footprint
Corner longitudes		[degrees east]	[scanl. x pixels x 4 x *3*]	Corners of footprint
Corner latitudes		[degrees north]	[scanl. x pixels x 4 x *3*]	Corners of footprint
Land/sea mask		–	[scanl. x pixels x *3*]	
Cloud fraction “strict”	$c_{f,strict}$	–	[scanl. x pixels x *3*]	See equation (13)
Cloud fraction “nominal”	$c_{f,nominal}$	–	[scanl. x pixels x *3*]	See equation (14)
Cloud fraction “relaxed”	$c_{f,relaxed}$	–	[scanl. x pixels x *3*]	See equation (15)
AAI	r	–	[scanl. x pixels x *3*]	
Ozone column	Ω	[DU]	[scanl. x pixels x *3*]	Optional
Surface elevation	z_s	[km]	[scanl. x pixels x *3*]	
Water vapour column	Ψ	[g/cm^2]	[scanl. x pixels]	Only spectral band 5/6
Snow/ice flag		–	[scanl. x pixels x *3*]	
Cloud shadow flag		–	[scanl. x pixels x *3*]	See section 6.2.4
Spectral quality flag		–	[scanl. x pixels x 21]	Multi-purpose flag
Reflectance	R_λ	–	[scanl. x pixels x 21]	
Reflectance precision	δR_λ	–	[scanl. x pixels x 21]	
Scene LER	A_{scene}	–	[scanl. x pixels x 21]	
Scene LER precision	δA_{scene}	–	[scanl. x pixels x 21]	See section 9.3.1

Table 8: Output parameters for component A. Breakpoints are listed in Table 9.

The typical number of scanlines in a TROPOMI orbit is 3246 (up to and including orbit 9387 of 6 August 2019) or 4173 (for orbit 9388 and later orbits). The number of pixels in a scanline depends on the spectral band. For band duo 3/4 it is 450, for band duo 5/6 it is 448 and for band duo 7/8 it is 215. The number of wavelength bands is 21. Note that in Table 8 often the number 3 is printed in italics. This is to indicate that the parameters

of two different wavelength bands are identical as long as both wavelength bands reside in the same spectral band duo (that is, both in band duo 3/4 or both in band duo 5/6 or both in band 7/8). In practice, in the output product, these parameters do not have dimension 21 but instead are provided for each of the three band duos. This reduces the size of the files and makes sense from an instrument perspective.

Parameter	Symbol	Physical unit	Dimensions	Comments
Radiance	I	[mol s^{-1} m^{-2} nm^{-1} sr^{-1}]	[scanl. x pixels x wavel.]	Breakpoint
Irradiance	E_0	[mol s^{-1} m^{-2} nm^{-1}]	[pixels x wavel.]	Breakpoint
Radiance wavelength	λ_{rad}	[nm]	[scanl. x pixels x wavel.]	Breakpoint
Irradiance wavelength	λ_{irr}	[nm]	[pixels x wavel.]	Breakpoint
Reflectance	R	–	[scanl. x pixels x wavel.]	Breakpoint
Fourier coefficient a_0	a_0	–	[scanl. x pixels x 21]	Breakpoint
Fourier coefficient a_1	a_1	–	[scanl. x pixels x 21]	Breakpoint
Fourier coefficient a_2	a_2	–	[scanl. x pixels x 21]	Breakpoint
Transmission	T	–	[scanl. x pixels x 21]	Breakpoint
Spherical albedo	$s^\star$	–	[scanl. x pixels x 21]	Breakpoint
Path reflectance	R^0	–	[scanl. x pixels x 21]	Breakpoint

Table 9: Breakpoint output parameters for component A.

8.3 Input – component B

8.3.1 Dynamic input

Parameter	Symbol	Physical unit	Source	Use	If unavailable
Wavelength	λ	[nm]	S5P L2 LER product	U	No processing
Scanline	s	–	S5P L2 LER product	C	No processing
Ground pixel	g	–	S5P L2 LER product	C	No processing
Time	t	[s]	S5P L2 LER product	U	No processing
SZA	θ_0	[degrees]	S5P L2 LER product	U	No processing
Centre longitude		[degrees east]	S5P L2 LER product	U	No processing
Centre latitude		[degrees north]	S5P L2 LER product	U	No processing
Land/sea mask		–	S5P L2 LER product	U	No processing
Cloud fraction “relaxed”	$c_{\mathrm{f,relaxed}}$	–	S5P L2 LER product	U	No processing
AAI	r	–	S5P L2 LER product	U	Skip filtering

Parameter	Symbol	Physical unit	Source	Use	If unavailable
Snow/ice flag		–	S5P L2 LER product	U	No processing
Cloud shadow flag		–	S5P L2 LER product	U	No processing
Spectral quality flag		–	S5P L2 LER product	U	No processing
Scene LER	A_{scene}	–	S5P L2 LER product	U	No processing
Scene LER precision	$\delta A_{\mathrm{scene}}$	–	S5P L2 LER product	U	No processing

Table 10: Dynamic input parameters for component B. U=Used in computation; F=Filter; C=Copied to output.

8.3.2 Algorithm configuration parameters

Parameter	Symbol	Physical unit	Meaning	Value
Maximum cloud fraction	$c_{\mathrm{f,max}}$	–	Maximal cloud fraction value accepted	0.03
Maximum SZA	$\theta_{0,\mathrm{max}}$	[degrees]	Maximal SZA value accepted	85.0
Maximum AAI	r_{max}	–	Maximal AAI value accepted	2.0
Minimum scene LER	A_{min}	–	Minimal scene LER value accepted	0.0
Maximum scene LER	A_{max}	–	Maximal scene LER value accepted	1.5
Fraction of cloud-free observations		–	Fraction of cloud-free observations used to calculate the accumulative mean surface LER	0.10

Table 11: Algorithm configuration parameters for component B. If the scene LER for any of the wavelength bands in a band duo is outside the specified range ($A_{\mathrm{scene}} < A_{\mathrm{min}}$ or $A_{\mathrm{scene}} > A_{\mathrm{max}}$), then the scene LER values from all wavelength bands from that specific band duo are rejected and the observation is skipped.

8.4 Output – component B

Parameter	Symbol	Physical unit	Dimensions	Comments
wavelength	λ	[nm]	[21]	
longitude		[degrees east]	[2880]	Centre of cell
latitude		[degrees north]	[1440]	Centre of cell
grid_num_obs_clear		–	[2880 x 1440]	
grid_num_obs_snice		–	[2880 x 1440]	

Parameter	Symbol	Physical unit	Dimensions	Comments
grid_surface_type_clear		–	[2880 x 1440]	
grid_surface_type_snice		–	[2880 x 1440]	
grid_nise_clear		–	[2880 x 1440]	
grid_nise_snice		–	[2880 x 1440]	
ler_clear	$A_{\mathrm{LER,clr}}$	–	[2880 x 1440 x 21]	
ler_snice	$A_{\mathrm{LER,snw}}$	–	[2880 x 1440 x 21]	
ler_uncertainty_clear	$\delta A_{\mathrm{LER,clr}}$	–	[2880 x 1440 x 21]	See section 9.3.2
ler_uncertainty_snice	$\delta A_{\mathrm{LER,snw}}$	–	[2880 x 1440 x 21]	See section 9.3.2
dler_coefficients_clear	$c_{0\ldots3,\mathrm{clr}}$	–	[2880 x 1440 x 21 x 4]	See equation (19)
dler_coefficients_snice	$c_{0\ldots3,\mathrm{snw}}$	–	[2880 x 1440 x 21 x 4]	See equation (19)

Table 12: Output parameters for component B.

8.5 Input – component C

8.5.1 Dynamic input

Parameter	Symbol	Physical unit	Source	Use	If unavailable
wavelength	λ	[nm]	S5P L3 LER product	U	No processing
longitude		[degrees east]	S5P L3 LER product	U	No processing
latitude		[degrees north]	S5P L3 LER product	U	No processing
grid_num_obs_clear		–	S5P L3 LER product	U	No processing
grid_num_obs_snice		–	S5P L3 LER product	U	No processing
grid_surface_type_clear		–	S5P L3 LER product	U	No processing
grid_surface_type_snice		–	S5P L3 LER product	U	No processing
grid_nise_clear		–	S5P L3 LER product	U	No processing
grid_nise_snice		–	S5P L3 LER product	U	No processing
ler_clear	$A_{\mathrm{LER,clr}}$	–	S5P L3 LER product	U	No processing
ler_snice	$A_{\mathrm{LER,snw}}$	–	S5P L3 LER product	U	No processing
ler_uncertainty_clear	$\delta A_{\mathrm{LER,clr}}$	–	S5P L3 LER product	U	No processing
ler_uncertainty_snice	$\delta A_{\mathrm{LER,snw}}$	–	S5P L3 LER product	U	No processing
dler_coefficients_clear	$c_{0\ldots3,\mathrm{clr}}$	–	S5P L3 LER product	U	No processing
dler_coefficients_snice	$c_{0\ldots3,\mathrm{snw}}$	–	S5P L3 LER product	U	No processing

Parameter	Symbol	Physical unit	Source	Use	If unavailable

Table 13: Dynamic input parameters for component C. U=Used in computation; F=Filter; C=Copied to output.

8.5.2 Algorithm configuration parameters

Parameter	Symbol	Physical unit	Meaning	Value
Maximum ocean LER for band 3/4	$A_{max,494}$	–	Maximal ocean LER accepted at 494 nm	0.08
Maximum ocean LER for band 5/6	$A_{max,772}$	–	Maximal ocean LER accepted at 772 nm	0.04
Maximum ocean LER for band 7/8	$A_{max,2314}$	–	Maximal ocean LER accepted at 2314 nm	0.04

Table 14: Algorithm configuration parameters for component C.

8.6 Output – component C

Parameter	Symbol	Physical unit	Dimensions	Comments
wavelength	λ	[nm]	[21]	
longitude		[degrees east]	[2880]	Centre of cell
latitude		[degrees north]	[1440]	Centre of cell
flag_clear		–	[2880 x 1440 x 21 x 12]	
flag_snice		–	[2880 x 1440 x 21 x 12]	
age_clear		[months]	[2880 x 1440 x 21 x 12]	
age_snice		[months]	[2880 x 1440 x 21 x 12]	
LER_clear		–	[2880 x 1440 x 21 x 12]	
LER_snice		–	[2880 x 1440 x 21 x 12]	
LER_uncertainty_clear		–	[2880 x 1440 x 21 x 12]	
LER_uncertainty_snice		–	[2880 x 1440 x 21 x 12]	
DLER_coefficients_clear	$c_0 \ldots c_3$	–	[2880 x 1440 x 21 x 12 x 4]	See equation (19)
DLER_coefficients_snice	$c_0 \ldots c_3$	–	[2880 x 1440 x 21 x 12 x 4]	See equation (19)

Table 15: Output parameters for component C.

8.7 Open Issues

There are no open issues.

9 Error analysis

9.1 Introduction

The uncertainty on the retrieved surface LER value depends highly on the number of clear-sky scenes that were collected for a given grid cell. But there are also error sources related to the modelling of the scene and related to errors in the reflectances measured by the TROPOMI instrument. The error in the reflectance is the largest source of errors for the retrieved scene LER values. This is discussed in the next section.

9.2 Impact of errors in the Earth reflectance

The sensitivity dA/dR of the retrieved scene/surface LER A to errors in the measured Earth reflectance R can be calculated in a straightforward way from equation (4) by differentiation. This yields

$$\frac{dA}{dR} = \frac{(1 - As^\star)^2}{T} \ , \tag{20}$$

where T is the transmission matrix and $s^\star$ the spherical albedo as defined before in section 6.2.1.

Figure 17 presents a plot of the sensitivity dA/dR as a function of surface LER for a selection of wavelengths ranging from 325 nm to 772 nm. The calculations for this plot were performed using the radiative transfer model DAK, for clear-sky conditions in a standard MLS atmosphere (ozone column 334 DU), for nadir view ($\theta = 0^\circ$) and solar zenith angle $\theta_0 = 60^\circ$. In the visible wavelength range an error of 0.01 in the reflectance leads to an error of approximately 0.01 in the retrieved surface LER, as expected. In the UV, this error increases enormously due to the smaller contribution of the surface to the TOA reflectance. Below 325 nm, retrieval of the surface LER is challenging (if not impossible) and would require a near perfect radiometric calibration as well as reliable information about the shape of the ozone profile for each of the measurement footprints. This is currently not feasible, but at the same time there is also no real existing demand from the remote sensing community for surface albedo information below 325 nm.

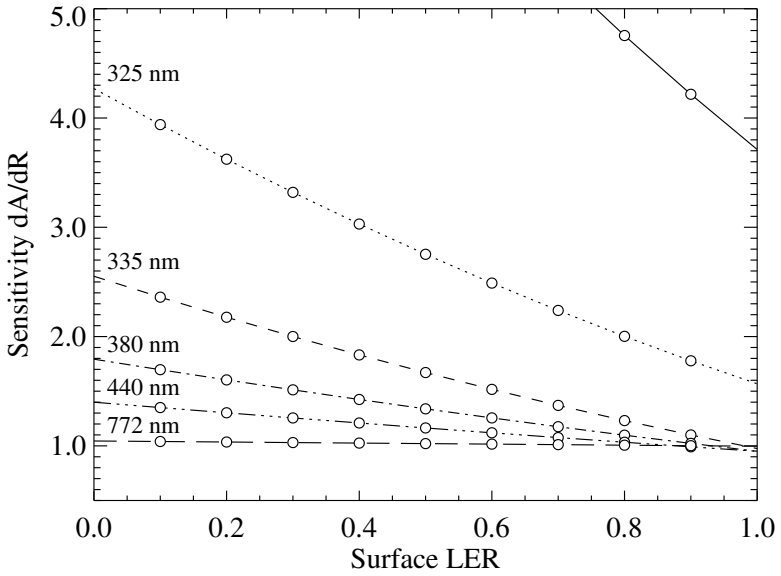


Figure 17: Sensitivity of the surface LER to changes in the Earth reflectance, as a function of surface LER and for different wavelengths. Below 330 nm the sensitivity increases dramatically.

9.3 Error calculation

9.3.1 Component A

The error in the measured Earth reflectance is calculated assuming that the errors in the radiance and solar irradiance as provided in the L1B product are completely independent. That is, we have

$$\delta R = R\sqrt{(\delta I/I)^2 + (\delta E_0/E_0)^2}\ , \tag{21}$$

where R is the reflectance, I the radiance, E_0 the solar irradiance, and where δI and δE_0 are the precisions of radiance and solar irradiance, respectively.

Given an absolute error δR in the reflectance we can use equation (20) to determine the contribution of this error to the retrieved scene LER A via $\delta A \approx (dA/dR)\cdot\delta R$. Note that the parameters T and $s^\star$ in equation (20) are already known, and that the parameter A is the scene LER itself. The error that is calculated this way is correct for all observations. But, by definition, it is the error on the scene LER, not the error on the surface LER. As such, it does not indicate whether the scene LER is representative for the true surface LER.

9.3.2 Component B

Component B will only calculate errors for the non-directional LER field. The statistical error δA_{LER} is defined as the weighted standard deviation of the scene LER values of all the selected clear-sky observations. That is,

$$\delta A_{\mathrm{LER}} = \sqrt{\frac{\sum_{m=1}^{M} v_m\cdot\left(A_{\mathrm{scene},m} - A_{\mathrm{LER}}\right)^2}{\frac{(M-1)}{M}\sum_{m=1}^{M} v_m}}\ , \tag{22}$$

where M is the number of clear-sky observations that were selected, i.e., those observations that have the 10% lowest values of the clear-sky scene LER values collected for the grid cell. The weighting factor v_m was defined already in section 6.3.4. The parameter A_{LER} is calculated using equation (17).

9.3.3 Component C

Component C simply copies the errors that were calculated by component B. In the case the information in a grid cell is replaced by that of a donor cell (as explained in section 6.4) the errors also are replaced.

10 Validation

10.1 Proposed validation plan

Validation of the retrieved TROPOMI surface LER database may be performed by comparison with the other surface LER databases that were discussed in section 5.2.1. From these databases, the SCIAMACHY surface LER database [RD6] makes most sense as a reference for the TROPOMI surface LER database, because of the largely overlapping set of surface LER wavelength bands, and because it makes use of a similar surface LER retrieval approach as the one described in this ATBD. The GOME-2 surface LER database [RD6, RD8] is also created using the same retrieval approach and is therefore also well suitable as a reference.

The OMI surface LER database [RD5] may be used as a reference for the wavelengths below 500 nm. The retrieval algorithm that was used for the OMI surface LER database is comparable to the algorithm used for the TROPOMI surface LER. Moreover, the OMI instrument has a local equator crossing time (13:45 LT) which is close to TROPOMI's (13:30 LT). This is certainly an advantage because it ensures that the solar angles for TROPOMI and OMI are similar. As a result, surface BRDF effects are also similar.

The GOME-1 surface LER database [RD4] unfortunately is fairly outdated and is suffering much more than the other databases from measurement errors, residual cloud contamination, and the effects of instrument degradation. As a result, its quality is lower than that of the other heritage databases [RD6], making it a less reliable reference. It also offers only a rather limited number of wavelength bands for comparison. As a result, comparing with the GOME-1 database does not seem worth the effort.

A comparison with non-Lambertian surface albedos, such as the MERIS black-sky albedo (BSA) [RD35], is in principle also possible. This is strictly speaking not correct, because the BSA is the integral of the bidirectional reflectance distribution function (BRDF) over the entire hemisphere whereas the LER is derived from the much smaller range of viewing angles of the satellite's observation geometry. Note that a comparison with the MERIS BSA database would only make sense over land, because the MERIS surface albedo values over sea are not retrieved from MERIS observations. They were taken directly from the GOME-1 surface LER database [RD35]. Comparisons between SCIAMACHY and MERIS surface albedo databases suggest that a comparison with the MERIS BSA is indeed not worth the effort [RD6].

In summary, for the TROPOMI surface LER product it makes most sense to perform the validation by comparison with the SCIAMACHY, GOME-2, and OMI surface LER databases. Examples of how such a comparison for surface LER products would look can be found in [RD6].

For the directional TROPOMI surface DLER the only sensible form of validation is a comparison with one of the established MODIS surface BRDF products [RD36]. For that we select the MCD43C2 product [ER8]. The spatial resolution of this product is $0.05^\circ \times 0.05^\circ$. The available MODIS wavelength bands are 469, 555, 645, 859, 1240, 1640, and 2130 nm. This means, that more or less exact comparisons are only possible for the 463-nm wavelength band and – to a lesser degree – the 670-nm wavelength band. For the other bands the wavelength differences between the TROPOMI surface DLER and MODIS surface BRDF wavelength bands are simply too large (see Table 2). But we do expect that the directional dependence can be checked in a qualitative way. The MCD43C2 product only provides BRDF coefficients over land.

10.2 Preliminary validation results

Figure 18 presents an example of preliminary validation results that have been obtained. The three global maps present the difference between the TROPOMI surface LER database and the OMI, SCIAMACHY, and GOME-2 surface LER databases, respectively. The differences are calculated for calendar month March and for the wavelength band positioned at 494 nm which exists in all four databases. In the validation report (VR) these analyses are performed for as many wavelength bands as possible, and for each of the twelve calendar months. Analyses are focused on the part of the globe between 60°S and 60°N, because the polar regions are more variable because of variations in the snow/ice fields, and therefore difficult to compare.

In Figure 18 the snow/ice regions are the surfaces which seem to have the highest (and systematic) differences compared to the heritage surface LER databases. Note that snow and ice also have the highest surface albedo of all surface types, so it is not surprising that the absolute error that is found is also the largest for these surfaces. Moreover, the snow/ice regions should not be compared because the retrieval of surface LER in these regions is not the same for the TROPOMI surface LER database as it is for the other LER databases. Instead, it is better to focus on the snow/ice-free regions. Here the agreement between TROPOMI surface LER and the heritage LER databases is very good.

The occurrence of similar features in all three maps presented in Figure 18 suggests that there are systematic features present in the comparisons. Worth mentioning is the blue feature in the north of South

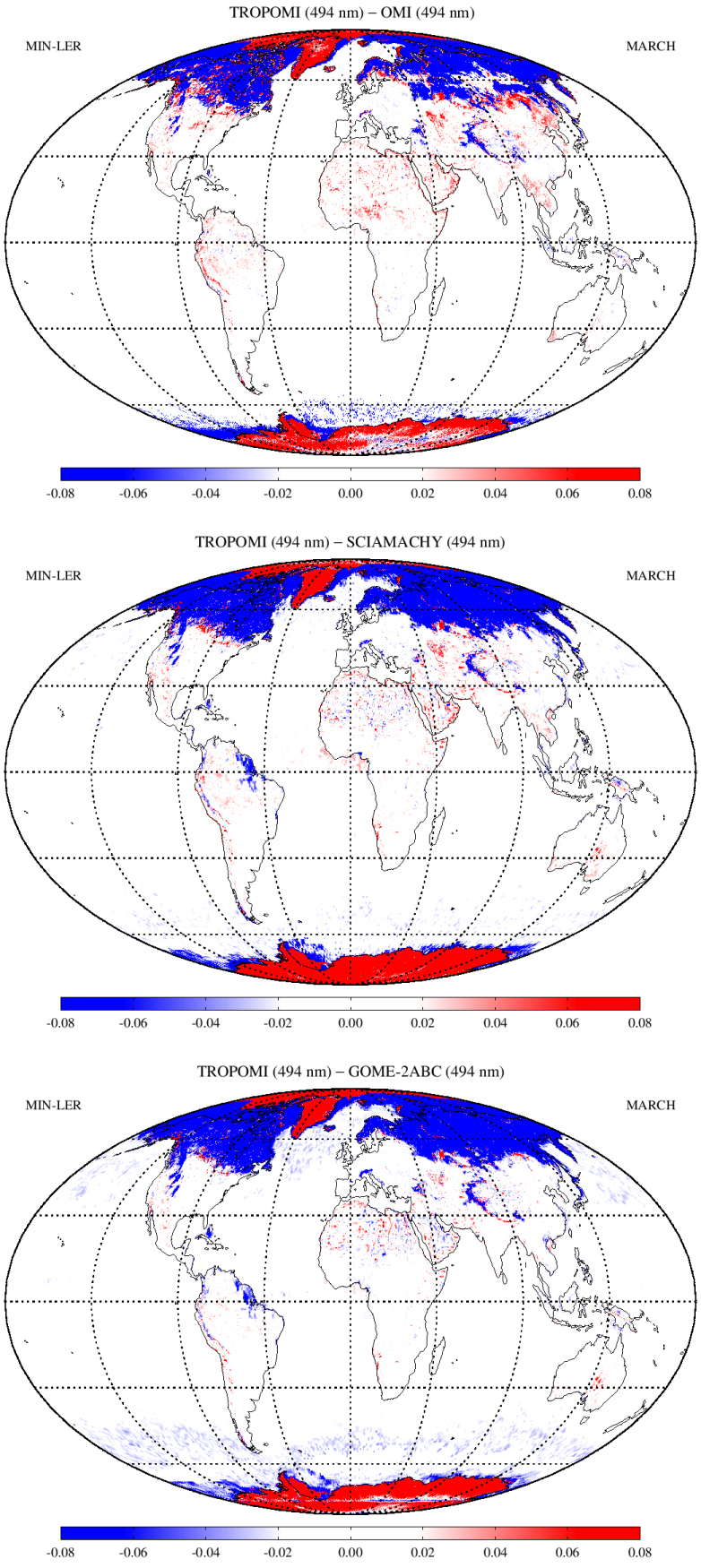


Figure 18: Example of typical validation results: Map of the difference between the 494-nm surface LER from TROPOMI and that of the OMI, SCIAMACHY, and GOME-2 surface LER databases.

America, which only appears in the comparisons with SCIAMACHY and GOME-2. The reason why it is not found in the comparison with OMI is that TROPOMI and OMI both are in afternoon orbits, while SCIAMACHY and GOME-2 are both in morning orbits. Early in the day the specific blue region is often covered by persistent cloud fields, thereby leading to too high values in the SCIAMACHY and GOME-2 surface LER. The TROPOMI and OMI surface LER databases do not show this issue for this particular region. Other features related to cloud contamination in particular areas can be seen in some of the plots in Figure 18.

Results for other wavelength bands are made available in the validation report (VR), and this is done much more in depth, for all twelve calendar months. The outcome of the validation is that, in general, the visible wavelength bands (derived from spectral bands 5/6) show minor issues, with a very low systematic error on the order of 0.001–0.002. For the UV-VIS (spectral bands 3/4) there used to be considerable systematic errors in previous versions (up to version 1.2) of the database. With the appearance of TROPOMI collection-3 data, systematic errors in the level-1 radiometric calibration [RD1] disappeared almost completely. The systematic errors in the surface LER are now on the order of 0.01–0.02. The 328-nm wavelength band can now also be used as it now also meets the requirements.

Finally, for the 2314-nm wavelength band (spectral band 7) the result is unclear. For this wavelength band we only have the SCIAMACHY surface LER database as a reference. It is at this point in time not possible to say whether TROPOMI or SCIAMACHY is responsible for the differences that are found. More details on the validation activities and results are made available in the VR.

11 Conclusion

This version of the ATBD is considered to be mature. Validation results are available and are mentioned in section 10. Next to this ATBD, a validation report (VR) and a product user manual (PUM) are also available.