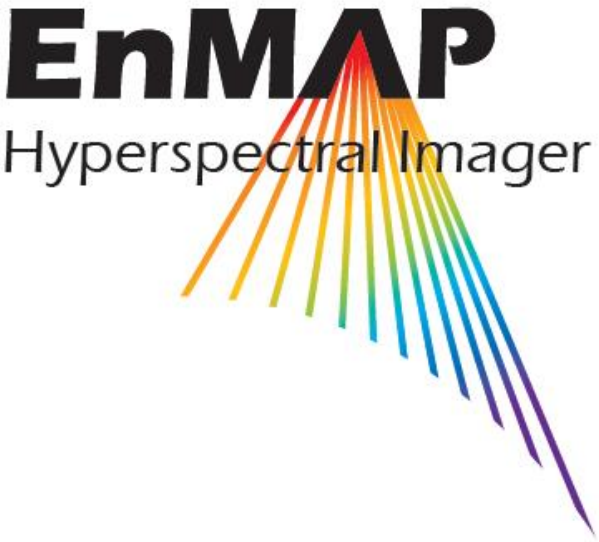
Deutsches Zentrum
für Luft- und Raumfahrt e.V.
in der Helmholtz-Gemeinschaft

EnMAP Ground Segment

Level 2A Processor (Atmospheric Correction over Land) ATBD

Doc. ID	EN-PCV-TN-6007
Issue	2.4
Date	07.09.2023

Configuration Controlled: Yes



German Remote Sensing Data Center (DFD)
Remote Sensing Technology Institute (IMF)
German Space Operation Center (GSOC)

– blank page –

TABLE OF SIGNATURES

Prepared

Date	Raquel de los Reyes Subsystem Engineer DEV
Date	Maximilian Langheinrich Subsystem Engineer DEV
Date	Martin Bachmann Subsystem Manager CGC

Reviewed

Date	Martin Bachmann Subsystem Manager CGC
Date	Rudolf Richter (DLR MF-PBA)

Approved

Date	Miguel Pato System Manager PCV

Released

Date	Emiliano Carmona GS Project Manager

– blank page –

DISTRIBUTION LIST

This EnMAP Ground Segment document is stored in the Design Definition File (DDF) of the EnMAP Ground Segment Teamsite.

This document is accessible by all participants of the EnMAP Ground Segment as well as by the EnMAP Ground Segment Customer. This document may be further distributed by the EnMAP Project Management as necessary.

After commissioning phase, the document is provided to the EnMAP user community on www.enmap.org.

CHANGE RECORD

Version	Date	Chapter	Comment
1.0	07.06.2010	all	
1.1	22.10.2010	4.1	Answer to RIDKAU-H-0211 included
1.2	17.05.2016	4, 8	System change request SCR-00018-User Product Definition and Naming
	17.05.2016	5	System change request SCR-00015-Ozone parameters for atmospheric correct
	17.05.2016	5	System change request SCR-00027-AC2020 instead of ATCOR
1.3	23.12.2016	5.1	Δ-CDR RID FIS-S-0052
		5.1	Δ-CDR RID FIS-S-0064
		5.6	Δ-CDR RID GUA-L-0384
		5.8	Δ-CDR RID HOL-A-0427
		5.8	Δ-CDR RID HOL-A-0431
		4.1	Δ-CDR RID SEG-K-0413
		8	Δ-CDR RID GUA-L-0380
		8	Δ-CDR RID GUA-L-0385
		5.3	Δ-CDR RID HOL-A-0428
1.4	08.11.2017	8.1, 8.3, added 8.4 and 8.5	Updated the information for the creation and usage of the sensor response functions and the atmospheric LUTs
	09.01.2018	8.	Corrected typos and answered comments.
	09.01.2018	All	Updated references AR to a proper list that can be automatically reference using Word tools
1.5	21.03.2018	5.3, 8	Included Ozone strategy: default values for automatic selection if the time difference between MODIS and scene is larger than 30 days (SCR-00103)
1.6	19.06.2018	6	Added documentation about the automatic ozone and season retrieval (SCR-00103)
1.7	03.12.2018	5.12	Added paragraph concerning shift of smile corrected pixels to nominal center wavelengths. (SCR-00111)
1.8	03.12.2018	all	Edition changes. Included TBD (To Be Done) section.
1.9	17.04.2019	all	Edition changes.
	14.11.2019	14.	Editorial changes related to the Data Quality (Combined and Land_Mode)
	03.04.2020	All	All the changes accepted.
	09.07.2020	2.2, 4. and 5.	General introduction of the L2A processor for land and water bodies. Editing changes: Cross reference format for the Informative references
	13.07.2020	5.1, 5.3, 5.4 and all	(SCR-00103) New processing change schema including external input dependencies. Editing: separation in different sections the pre-classification and the final EnMAP masks. Adding more detailed reference to L2A processor and cross references through the document. Updated MODTRAN version number. Editing changes: updated cross reference and IR references.
	31.07.2020	4.2 and 5.3	Added section about pixel interpolation (SCR-00208) Re-edition of the masking thresholds and formulas (SRC-00102)
	18.08.2020	5.7	Edition of chapter 5.7 to add the interpolation of HTM to hyperspectral (NCR-00152) Removed de-shadowing section since it is not applied in the L2A processor.
	21.09.2020	All, 5.2	Editing: Removed section with PACO quality layers (not EnMAP products) Adding cross-references and automatic equations numbers across the document Added more detailed information about L1Bint (SCR-00102)

	24.09.2020	6 and 7	Add information about requirements, preconditions and validation procedures
	15.10.2020	4.2, 5.1.3, 9 10 and Appendix	Added description of the interface between land and water and Metadata fields (SCR-00123)
2.0	08.06.2021	4.2.3	Added description of deactivation of strong water vapor absorption bands in the L2A user product (SCR-00214)
	06.09.2021	4.2.3	Added a description of the band interpolation default in the L2A job order
	14.09.2021	4.2.	Added general description of selectable L2A Water output products.
	11.01.2022	All	Corrected typos and grammar (from authors and reviewers)
2.1	20.01.2022		Version for GS-ORR
	06.09.2022	5.1.1.4	Corrected typo
	06.09.2022	5.1.1.4, 5.1.1.6	Updated cloud threshold for low illumination conditions and shadows threshold (TC-930-001-L2A-1)
	06.09.2022	14.1	Updated L2A quality criteria thresholds (SCR-00331, TC-930-002-L2A-2)
	20.09.2022	14.1, 14.2	Added remark reg. exclusion of L2A atm. quality rating in L0 overall quality rating, see SCR-00331
	21.09.2022	All	Typos correction
	29.09.2022	2.2, 5.3	Editorial correction
2.2	29.09.2022		Version for the FQR + CAR
	09.11.2022	5.1.2	Editorial typos
	29.11.2022	8	Addition of description of water verification
	24.01.203	8	Changes resulting from closing NCR 00341
	24.02.2023	4.2.3	Redefinition of 1400 and 1900 absorption bands ranges (as requested by GFZ)
2.3	24.02.2023	all	Typos revision and new release version
	01.03.2023		Changed signatures responsibles in managements
	10.07.2023	4.2.3	Documentation of the changes in wavelength ranges for the delivery of new bands (ICR-356-4)
	24.07.2023	5.10, 9.	Update documentation for the change of the ozone and season auxiliary information from MODIS to Copernicus Climate Change service database (ICR-00354-02)
2.4	07.09.2023	Editorial changes	Release version to solve ICR-00354-02

Custodian of this document is Raquel de los Reyes.

EXTERNAL DOCUMENT REVIEW

Dr. Barbara Bulgarelli
Joint Research Centre (JRC) of the European Commission
TP 723, Via Fermi 1, 21020 Ispra (VA), Italy
phone: ++39-0332-785778 / fax: ++39-0332-785469
e-mail: barbara.bulgarelli@jrc.it

Reviewed on 04.06.2010
Comments from external review processed and included

– blank page –

CONTENTS

1.1 Purpose 11
1.2 Scope 11
2.1 Applicable References 12
2.2 Informative References 12
4.1 EnMAP Processing chain and generated products 15
4.2 L2A atmospheric correction processor for land and water surfaces 16
4.2.1 Interpolation of spectral pixel values 17
4.2.2 Interface between land and water processors 17
4.2.3 Setting strong absorption bands in the user product output to default NoData value 18
5.1 Masks for land, water, haze, cloud, shadow 20
5.1.1 Pre-classification masks 20
5.1.2 EnMAP quality layers 24
5.1.3 EnMAP L2A Metadata 24
5.2 Radiative transfer calculations and atmospheric database 25
5.2.1 Ozone 25
5.2.2 Atmospheric database and radiative transfer LUT 27
5.3 Extraterrestrial solar irradiance spectrum 29
5.4 Retrieval of aerosol optical thickness 30
5.4.1 Notice concerning visibility iterations: 31
5.5 Retrieval of columnar water vapor 32
5.5.1 Water vapor retrieval for smile sensor 33
5.6 Haze detection and removal 33
5.6.1 Haze thickness map 34
5.6.2 Haze thickness per band 35
5.6.3 Haze removal 35
5.6.4 DEM data in dehazing method 36
5.7 Cirrus detection and removal 36
5.8 Retrieval of surface reflectance 37
5.8.1 Flat-terrain 37
5.8.2 Mountainous terrain 39
5.9 Correction of spectral smile 43
5.10 Ozone column and season automatic retrieval 44
5.10.1 Ozone column 44
5.10.2 Season 44
6.1 Tile processing 45
7.1 Land Requirements: AOT, WV and surface reflectance 46
8.1 Verification of land requirements 47
8.2 Verification of water requirements 47
9.1 Input: sensor description 48
9.2 Input: atmospheric LUTs 49
9.3 Input: ozone LUTs 50
9.4 Input: land surface temperature DB 50
9.5 Output files 50
14.1 Quality land processor 54
14.2 Quality water processor 54
14.3 Quality processor in Combined mode 55

LIST OF FIGURES

Figure 4-1: EnMAP processing chain. 16
Figure 5-1: L2A land processor workflow. From L1C products (in green) and external input databases (in blue), the land processor extract a set of internal (in grey) and final products (in red). Those products masked in red and incorporated into the final EnMAP products. The information of AOT and WV is included in the metadata file. 19
Figure 5-2 Influence of ozone on surface reflectance retrieval. 26
Figure 5-3 Retrievals for the synthetic winter scene (black: winter, grey: summer LUTs). 28
Figure 5-4: Top: Solar irradiance spectra of Fontenla 2011, Thuillier 2003 and TSIS 2021 convolved to 10 nm. Bottom: Relative difference of the above spectra (in %) between Thuillier 2003 and TSIS 2021 with respect to EnMAP solar irradiance model (Fontenla 2011). [IR-45] 29
Figure 5-5: Schematic sketch of APDA method with three channels. 32
Figure 5-6: Treatment of the adjacency effect. 39
Figure 5-7: Radiation components in mountainous terrain. 40
Figure 5-8: SPOT-5 scene from Switzerland with topographic correction 43
Figure 6-1: L2A tile land processing. To estimate the L2A products for tile i, tiles i-1, i, and i+1 are combined and processed. The L2A products corresponding to the tile i are cut and delivered. 45

LIST OF TABLES

Table 1 Enhanced classes (land, water, haze, cloud, etc.) internally generated in the L2A land processor 20
Table 2: EnMAP quality layers products (L0 and LX product specifications) with the correspondence between their label and the pre-classification mask (between brackets) in Table 1. 24
Table 3: Grid points schema for the monochromatic atmospheric database. (VIS= visibility, SZA = solar zenith angle, RAA = relative azimuth angle between sun and observer line-of-sight, ELE = ground elevation above sea level, TILT=instrument tilt angle, WVC = water vapor column) 27
Table 4 Altitude profile of the midlatitude summer atmosphere. 52
Table 5 Altitude profile of the midlatitude winter atmosphere. 52

1. Introduction

1.1 Purpose

This document defines the project standard to which all EnMAP Ground Segment project documentation shall be produced. The defined documentation structure and layout in this standard shall support the required information for use, development and correctness evidence of the EnMAP Ground Segment on a defined document quality level.

1.2 Scope

This EnMAP Ground Segment ATBD (Algorithm Theoretical Basis Document) is structured as follows:

- Chapters 1 to 3 are the usual ECSS prompted introductory chapters.
- Chapter 4 contains a brief survey on the two modules for atmospheric correction (land / water).
- Chapter 5 is the main chapter of this ATBD.
- Chapter 6 describes the requirements.
- Chapter 7 and 8 concerns verification and validation.
- Chapter 9 documents the input and output files.
- Chapter 10 contains a summary.
- The remaining chapters are appendices with more details on certain aspects.

2. References

2.1 Applicable References

The following documents are applicable to the extent specified herein.

Document ID	Document Title	Issue
AR-1	EN-KT-RS-001 Mission Requirements Document	1.0
AR-2	EN-KT-DD-001 EnMAP Design Document – Instrument, Volume III	1.0
AR-3	EN-PCV-ICD-2009-1 HSI Product Description Level 0	1.6
AR-4	EN-PCV-RSP-2013_PCV System Requirements Specifications	1.3
AR-5	EN-PCV-TN-5006: Level 1C Processor (Geometric Correction) ATBD	1.6
AR-6	EN-PCV-TN-6008: Level 2A Processor (Atmospheric Correction Water) ATBD	3.1
AR-7	EN-GS-TN-1009: EnMAP Glossary and Abbreviations	3.0
AR-8	EN-PCV-TN-1004 Definition of Product Levels	1.2
AR-9	EN-PCV-DD-2004 Processor Design Document	1.6
AR-10	EN-PCV-TN-8005 HSI Geometric Calibration and Quality Control Concept	1.3
AR-11	EN-PCV-DD-8004 EnMAP Ground Segment Quality Control Design Document	1.2
AR-12	EN-PCV-ICD-2009-2 HSI Product Description Level 1 / Level 2	1.8
AR-13	EN-PCV-ICD-7008_HSI_Calibration_Product_Specification	1.6
AR-14	EN-PCV-TN-6040_RfD-00002	1.0
AR-15	EN-PCV-TN-6020 Smile Correction in EnMAP Products	2.2
AR-16	EN-PCV-TN-4006 Level 1B Processor Systematic and Radiometric Correction	1.7
AR-17	EN-PCV-PLN-2020_DEV_Operations_Plan	1.2
AR-18	EN-PCV-TN-8006 Handling of low quality data	1.1

2.2 Informative References

The following documents, though not formally part of this document, amplify or clarify its content.

Document ID	Document Title	Issue
IR-1	Ackerman, S. A., Strabala, K. I., Menzel, W. P., Frey, R. A., Moeller, C. C., and Gumley, L. E.," Discriminating clear sky from clouds with MODIS", *J. Geophys*. *Res.*, Vol. 103, D24, 32,141-32,157 (1998).	1998
IR-2	Berk, A., et al., "MODTRAN cloud and multiple scattering upgrades with application to AVIRIS", *Remote Sensing of Environment,* Vol. 65, 367-375 (1998).	1998
IR-3	Berk, A., et al., "MODTRAN4 Version 3 Revision 1 User 's Manual", Air Force Research Laboratory, Hanscom MA (2003).	2003
IR-4	Berk, A., Anderson, G.P., Acharya, P.K., and Shettle, E.P., "MODTRAN5.2.0.0 User's Manual", Spectral Sciences Inc., Burlington, MA; Air Force Research Laboratory, Hanscom, MA (2008).	2008
IR-5	Bouvet, M. et al, 2019, A Radiometric Calibration Network for Earth Observing Imagers Operating in the Visible to Shortwave Infrared Spectral Range, Remote Sens., 11, doi:10.3390/rs11202401.	2019
IR-6	Chylek, P., Borel, C. C., Clodius, W., Pope, P. A., and Rodger, A. P., "Satellite-based columnar water vapor retrieval with the multi-spectral Thermal Imager (MTI)", *IEEE TGRS*, Vol. 41, 2767-2770 (2003).	2003
IR-7	Crist, E. P., and Cicone, R. C., "A physically-based transformation of Thematic Mapper data - the Tasseled Cap", *IEEE TGRS,* Vol. GE-22, 256-263 (1984).	1984
IR-8	De los Reyes, R. et al, 2020, "PACO: Python-based Atmospheric Correction", MDPI sensors, 20, 1428; doi:10.3390/S20051428	2020
IR-9	Dozier, J., Bruno, J., and Downey, R. C., "A faster solution to the horizon problem", *Computers & Geosciences*, Vol. 7, 145-151 (1981).	1981
IR-10	Fontenla, J.M.; Harder, J.; Livingston, W.; Snow, M.; Woods, T.; "High-resolution solar spectral irradiance from extreme ultraviolet to far infrared", J. Geophys. Res. Atmos., 2011, 116, doi:10.1029/2011JD016032.	2011
IR-11	Gao, B.-C., and Goetz, A.F.H., "Determination of total column water vapor in the atmosphere at high	1990

	spatial resolution AVIRIS data using spectral curve fitting and band rationing techniques", Proc. SPIE, Vol. 1298, 138-150, 1990.	
IR-12	Gao, B.-C., Kaufman, Y.J., Han, W., and Wiscombe, W. J., "Correction of thin cirrus path radiances in the 0.4 – 1.9 µm spectral region using the sensitive 1.375 µm cirrus detecting channel", *J. Geophys. Res.,* Vol 103, D24, 32,169-32,176 (1998).	1998
IR-13	Gao, B.-C., Yang, P., Han, W., Li, R.-R., and Wiscombe, W. J., "An algorithm using visible and 1.38 µm channels to retrieve cirrus cloud reflectances from aircraft and satellite data", *IEEE TGRS*, Vol. 40, 1659-1668 (2002).	2002
IR-14	Gao, B.-C., Meyer, K., and Yang, P., "A new concept on remote sensing of cirrus optical depth and effective ice particle size using strong water vapor absorption channels near 1.38 and 1.88 µm", *IEEE TGRS*, Vol. 42, 1891-1899 (2004).	2004
IR-15	Goetz, A. F. H, Kindel, B. C., Ferri, M., and Qu, Z., "HATCH: results from simulated radiances, AVIRIS, and Hyperion", IEEE TGRS, Vol. 41, 1215-1221 (2003).	2003
IR-16	Gomez-Chova, L., Camps-Valls, G., Calpe-Maravilla, J., Guanter, L., and Moreno, J., "Cloud-screening algorithm for ENVISAT/MERIS multispectral images", *IEEE TGRS* Vol. 45, 4105-4118.	
IR-17	Guanter, L., Richter, R., and Kaufmann, H., "On the application of the MODTRAN4 atmospheric radiative transfer code to optical remote sensing ", *Int. J. Remote Sensing*, Vol. 30, 1407-1424 (2009).	2009
IR-18	Hay, J. E., and McKay, D. C., "Estimating solar irradiance on inclined surfaces: a review and assessment of methodologies", *Int. J. Solar Energy*, Vol. 3, 203-240 (1985).	1985
IR-19	Holben, B.; et al, 1998, AERONET - A Federated Instrument Network and Data Archive for Aerosol Characterization, Remote Sens. Environ., 66, 1-16. Doi:10.1016/S0034-4257(98)00031-5	1998
IR-20	Irish, R.R., Barker, J. L., Goward, S. N., and Arvidson, T.," Characterization of the Landsat-7 ETM+ automated cloud-cover assessment (ACCA) algorithm" *Photogr*. *Engin*. *Remote Sens*. Vol. 72, 1179-1188 (2006).	2006
IR-21	Isaacs, R. G., Wang, W. C., Worsham, R. D., and Goldberg, S., "Multiple scattering LOWTRAN and FASCODE models", Applied Optics, Vol. 26, 1272-1281 (1987).	1987
IR-22	Kaufman, Y. J., et al. "The MODIS 2.1 µm channel – correlation with visible reflectance for use in remote sensing of aerosol", *IEEE TGRS,* Vol. 35, 1286-1298 (1997).	1997
IR-23	Liang, S., Fallah-Adl, H., Kalluri, S., Jaja, J., Kaufman, Y. J., and Townshend, J. R. G.," An operational atmospheric correction algorithm for Landsat Thematic Mapper imagery over the land", *J. Geophys. Res.*, Vol 102, D14, 17,173-17,186 (1997).	1997
IR-24	Marion, R., Remi, M., and Faye, C., "Measuring trace gases in plumes from hyperspectral remotely sensed data*", IEEE TGRS* Vol. 42, 854-864 (2004).	2004
IR-25	Marion, R., Michel, M., and Faye, C., "Atmospheric correction of hyperspectral data over dark surfaces via simulated annealing", *IEEE TGRS,* Vol. 44, 1566-1574	
IR-26	Mouroulis, P., Green, R. O., and Chrien, T. G., "Desing of pushbroom imaging spectrometers for optimum recovery of spectroscopic and spatial information", Applied Optics, Vol. 39, 2210-2220 (2000).	2000
IR-27	Nicodemus, F. E., "Reflectance nomenclature and directional reflectance and emissivity", *Applied Optics*, Vol. 9, 1474-1475 (1970).	1970
IR-28	Platnick, S.E.A. "MODIS Atmosphere L3 Eight-Day Product"; NASA MODIS Adaptative Processing System, Goddard Space Flight Center, USA, 2017; doi:10.5067/MODIS/MOD08-E3.061	2017
IR-29	Richter, R., "A fast atmospheric correction algorithm applied to Landsat TM images", *Int. J. Remote Sensing*, Vol. 11, 159-166 (1990).	1990
IR-30	Richter, R., "Atmospheric correction of satellite data with haze removal including a haze/clear transition region", *Computers & Geosciences*, Vol. 22, 675-681 (1996).	1996
IR-31	Richter, R., "Correction of satellite imagery over mountainous terrain", *Applied Optics*, Vol. 37, 4004-4015 (1998)	1998
IR-32	Richter, R., and Müller, A., "De-shadowing of satellite / airborne imagery ", *Int. J. Remote Sensing*, Vol. 26, 3137-3148 (2005).	2005
IR-33	Richter, R., and Schläpfer, D., "Considerations on water vapor and surface reflectance retrievals for a spaceborne imaging spectrometer ", *IEEE TGRS,* Vol. 46, 1958-1966 (2008).	2008
IR-34	Richter, R., Kellenberger, T., and Kaufmann, H., "Comparison of topographic correction methods ", *Remote Sensing*, Vol. 1, 184-196 (2009).	2009
IR-35	Richter, R., Heege, T., Kiselev, V., and Schläpfer, D., "Correction of ozone influence on TOA radiance ", Int. J. Remote Sensing, Vol. 35, 8044-8056 (2014).	2014
IR-36	Schläpfer, D., Borel, C. C., Keller, J., and Itten, K. I., "Atmospheric precorrected differential absorption technique to retrieve columnar water vapor", *Remote Sensing of Environment*, Vol. 65, 353-366 (1998).	1998
IR-37	Simpson, J. J., and Stitt, J. R., "A procedure of the detection and removal of cloud shadow from AVHRR data over land", *IEEE TGRS*, Vol. 36, 880-897 (1998).	1998
IR-38	Simpson, J. J., Jin, Z., and Stitt, J. R., "Cloud shadow detection under arbitrary viewing and	2000

	illumination conditions", *IEEE TGRS*, Vol. 38, 972-976 (2000).	
IR-39	Sirguey, P., "Simple correction of multiple reflection effects in rugged terrain", *Int. J. Remote Sensing*, Vol. 30, 1075-1081 (2009).	2009
IR-40	Stamnes, K., Tsay, S. C., Wiscombe, W. J., and Jayaweera, K., "Numerically stable algorithm for discrete-ordinate method radiative transfer in multiple scattering and emitting layered media", Applied Optics, Vol. 27, 2502-2509 (1988).	1988
IR-41	Wan, Z.; Hook, S.H.G., "MYD11A2 MODIS/Aqua Land Surface Temperature/Emissivity 8-Day L3 Global 1km SIN Grid V006"; doi:10.5067/MODIS/MYD11A2.006	
IR-42	P. Yang, K. N. Liou, K. Wyser, and D. Mitchell, "Parametrization of the scattering and absorption properties of individual ice crystals", *J. Geophys. Research*, Vol. 105, No. D4, 4699-4718 (2000).	2000
IR-43	Zhang, Y., Guindon, B., and Cihlar, J., "An image transform to characterize and compensate for spatial variations in thin cloud contamination of Landsat images", *Remote Sensing of Environment*, Vol. 82, 173-187 (2002).	2002
IR-44	Coddington, O.M.; Richard, E.C.; Harber, D.; Pilewskie, P.; Woods, T.N.; Chance, K.; Liu, X.; Sun, K. The TSIS-1 Hybrid Solar Refer- 569 ence Spectrum. Geophysical Research Letters 2021, 48, e2020GL09170	2021
IR-45	De Los Reyes, R. ; Richter, R.; Bachmann, M.; Alonso, K.; Pflug, B.; Lafrance, B.;Reinartz, P., Influence oft he Solar Spectra Models on PACO Atmopsheric Correcion, Remote Sensing, 14, 4237 (2022)	2022
IR-46	https://oceancolor.gsfc.nasa.gov/docs/rsr/f0.txt	2022
IR-47	Zibordi et al. (2009: AERONET-OC: A Network for the Validation of Ocean Color Primary Products. *J. Atmos. Oceanic Technol.*, *26*, 1634–1651, DOI: 10.1175/2009JTECHO654.1.	2009

3. Terms, Definitions and Abbreviations

Terms, definitions and abbreviations for the EnMAP Ground Segment are collected together with those for the EnMAP Space Segment in a database which is publicly accessible via Internet on the EnMAP Information Portal:

http://www.enmap.org/

(menu item: Glossary & Abbr.)

4. Outline

This ATBD describes an algorithm for the atmospheric correction of EnMAP data over land surfaces using the L2A land processor based on the PACO (Python-based Atmospheric Correction) library. A separate module (MIP = Modular Inversion Program) provided by EOMAP is available for water bodies [AR-6]. The method for land can be applied over flat or mountainous terrain and converts top-of-atmosphere radiance imagery (L1C) into surface reflectance data (L2A). The processor is also used as L1B_int processor to perform the atmospheric correction in sensor geometry. The flat terrain case can be processed with L1B data prior to geocoding or after geocoding (L1C). The two image cubes from the VNIR and SWIR spectrometer are always co-registered to the VNIR instrument. A Digital Elevation Model (DEM) is always in the L1B and L1C internal products to include a correction of topographic effects and, together with the sensor geometry (L1B) and orthorectified (L1C) product, are input to the L1B_int and L2A processor, respectively. A DEM is available for the scene and it is applied in the atmospheric correction when at least 1% of the scene pixels have slope values greater than 6 degrees.

The L2A processor is also run in the L0 chain to generate the pre-classification masks (or quality layers) as well as the quality indicators.

The document presents the physical basis of the algorithm and its assumptions and limitations.

4.1 EnMAP Processing chain and generated products

The algorithms described in this document are applied to the automatic EnMAP HSI processing chain, described in detail in AR-9 and schematically represented in Figure 4-1.
The L2A algorithms described in this document are applied in two parts of the processing chain:

- **L1Bint**: **simplified atmospheric correction** in the VNIR and SWIR sensors separately to perform the smile correction (if the sensor's final characterization includes the correction of this effect) and the pixels interpolation in pixel surface reflectance. An atmospheric inversion will return the L1B products to radiances.
- **L2A**: complete **atmospheric correction** of the orthorectified EnMAP L1C product, described in [AR-9]. The L1C processor orthorectifies image tiles from the VNIR and SWIR instrument independently. After the orthorectification the two image tiles are co-registered (requirement better than 0.2-pixel size) and form a geometric consistent product over the whole wavelength range. The result of the atmosphere correction is reflectance [in % units] values and pixel classification (quality layers).

This document will describe the algorithms used for the complete atmospheric correction and include the simplification assumption assumed for L1Bint within the description of each of the corresponding algorithms, if it applies.

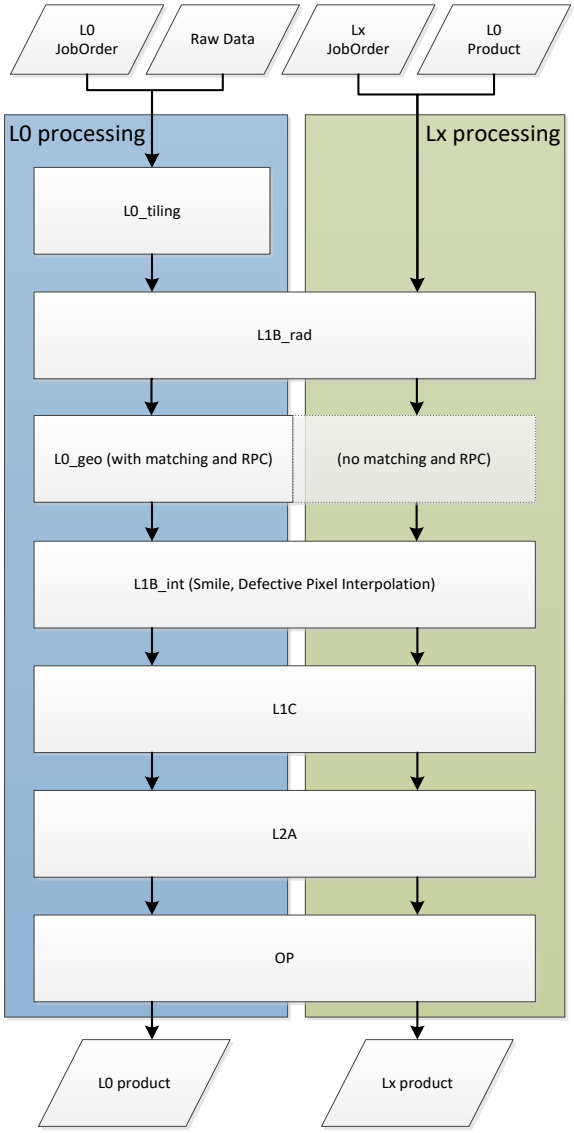


Figure 4-1: EnMAP processing chain.

4.2 L2A atmospheric correction processor for land and water surfaces

As mentioned before, the L2A atmospheric correction is performed through different algorithms depending if the pixel is in a land or water surface.

Section 5 will describe the algorithms used in the atmospheric correction to retrieve the surface reflectance (in % units). Internally, the land processor works in radiance units (mW cm^{-2} sr^{-1} μm^{-1}).
The surface reflectance L2A product is obtained in *Land_Mode* (both for land and water pixels) and only for land surfaces in *Combined* mode.

For water bodies in *Combined* and *Water_Mode*, the L2A processor will retrieve either subsurface irradiance reflectances (R_E) or normalized water-leaving remote sensing reflectance (R_rs), both expressed in % units. The given output type can be specified by the user and passed to the processor though the job order file, with R_rs being the default option. The algorithms applied are described in detail into a different document [AR-6].

The quality layers (or pixel classification) are the same for the three processing modes.
They are based on the pixel classification described in section 5.1, which will be also input for the water algorithms. But the final quality layers will contain the results from the water processor that will correct the land classification of some pixels to water, if their analysis considers it so.
Such quality layers are consistent among the rest of the product levels (L1B, L1C and L2A).

4.2.1 Interpolation of spectral pixel values

As mentioned in 4.1 the atmospheric correction process is not only applied in level L2A but in a simplified form during the L1B processing steps (in particular L1B_int). The motivation is to be able to perform any spectral interpolation steps within the processors on BOA reflectances as these reflectance spectra are generally smoother than the expected TOA radiance spectra and therefore are considered to promote better interpolation accuracy. Two different interpolation steps are applied during the EnMAP processing chain after retrieving surface reflectances in L1Bint.:

- In the case of a smile affected sensor and activated smile correction, an interpolation of the column wise retrieved surface reflectances to the nominal wavelength of each particular band is applied. This is necessary as the surface reflectances in the case of occurring smile up to this point in the chain represent the reflectances related to the shifted wavelengths per column per band.
- An interpolation of dead and defective pixels is always conducted. The information on which pixels to interpolate is provided by a binary mask flagging dead pixels and defects generated during L0 processing (AR-16). By design the interpolation is applied to surface reflectances and in spectral dimension to exploit the high spectral resolution of the EnMAP sensors and as these spectra are considered smoother in comparison to top-of-atmosphere radiances. In cases where missing spectral information may lead to large interpolation errors i.e. in damaged datatakes missing a large number of consecutive / neighboring bands or in spectral border regions of the sensors, interpolation is adaptively switched to a weighted spatial neighborhood approach.

After the interpolation steps were applied to the bottom-of-atmosphere surface reflectances, an inverted atmospheric correction model is applied to restore top-of-atmosphere at-sensor radiances for further processing.

At the wavelength range where the atmospheric water vapor absorbs a large portion of the solar radiation (depending on the content of water molecules), the radiance measured by the sensor is rather low, especially over water bodies. The uncertainty of the atmospheric correction could be of the order of magnitude of the calculated reflectance at these wavelengths, producing large fluctuations (“spikes”) in the final spectra. This would result, for example, in a large bias in the mean signal of water pixels above 700 nm.

Therefore, the surface reflectance and the water reflectance in the water vapor absorption bands retrieved during L2A processing will be set to the defined no data / background value and all the pixels of those bands will be flagged in the dead pixel mask.

4.2.2 Interface between land and water processors

As described in Section 5, the land processor generates additional information required for the land atmospheric correction. This information is also transmitted to the water processor, so in any of the processing modes the land processor must run and generate the corresponding information.

The information passed to the water processor is:

- Ozone column: corresponding ozone column value corresponding to the tile. This value is extracted from Copernicus database or is a default value if the database is not up-to-date (see Section 5.10 for details). Ozone column information is used in the L2A water processor in form of a static value mask in the size of the sensor image extents.

- Pixel classification mask: land internal pre-classification mask (Section 5.1.1) is delivered to the water processor which needs this preliminary information as an input dataset for its own land-water mask generation. During this process pixel classes of the values 2, 3, 4, 13, 15, 16 and 17 (see Table 1) are considered a trusted classification by the water processor and used similarly during internal processing. Pixels flagged with a different class than those mentioned above are double checked by the water processor and redefined to water where necessary.

This possible reclassification of spectral pixels during the water processing chain makes it necessary to update the metadata with regards to the overall number of water pixels and land-/water-related statistical properties. This update step is conducted at the end of L2A processing after L2A land and L2A water processing steps have run successfully. Further the land-water-classification mask as initially produced by the land processor is updated and given as part of the user product.

4.2.3 Setting strong absorption bands in the user product output to default NoData value

Bands within the wavelength range of strong water vapor absorption (1331.0 – 1460.0 nm and 1796.0 – 1938.0 nm) are always disabled by default in the final user product image cubes of the L2A processing chain. The deactivation is applied by setting the according bands to the default NoData / background value of -1e6.
From all processing starting 05.07.2023 onwards, the wavelength range for the 1400 nm water absorption region was changed to 1331.0 – 1448.0 nm to adjust the substitution bands interval to the new SWIR bands delivered.

The determination of bands falling in the two wavelength ranges is conducted automatically during L2A Land processing and the information is passed to the L2A Water processor to apply deactivation accordingly. For the nominal case (no missing bands, no smile) this refers to the bands 130-135. For the 1900 wavelength range, all the bands correspond to bands not transmitted

In addition, also the bands within the absorption regions around 940nm, 1130 nm, 725 nm, 760 nm and 820 nm can be set to background value, or can be provided within the user product.
This can be set by the user, and is handled through the “Band Interpolation” option in the job order, where “Yes” relates to the substitution of these bands with the background value. Note that the bandwidth of each absorption region is configurable in the L2A Land processor.

5. Algorithm Description

Input to the L2A processor (simplified or complete) are radiances (ortho-rectified or in sensor geometry) with their corresponding co-registered Digital Elevation Model (DEM) in the same geometry as the input radiances cubes. The L2A processor (see workflow in Figure 5-1) creates a *classification map* that includes land, water, haze, clouds, snow, cirrus and cloud shadow pixels (Section 5.1).

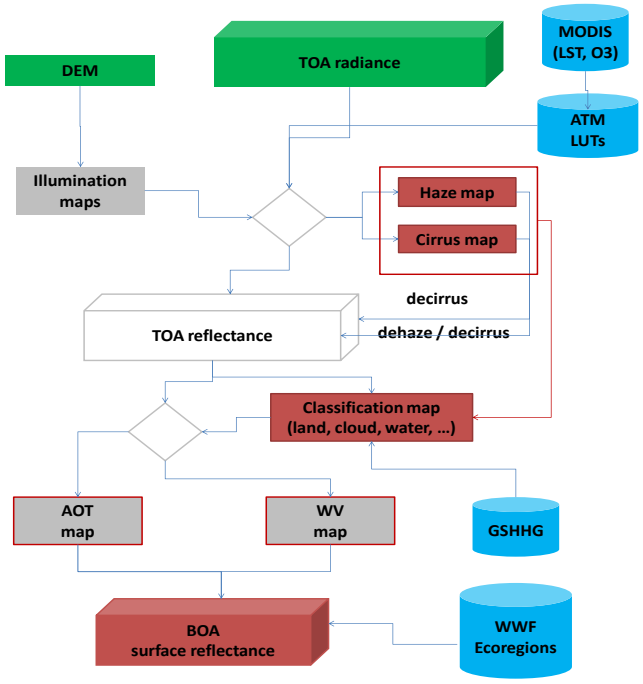


Figure 5-1: L2A land processor workflow. From L1C products (in green) and external input databases (in blue), the land processor extract a set of internal (in grey) and final products (in red). Those products masked in red and incorporated into the final EnMAP products. The information of AOT and WV is included in the metadata file.

For known atmospheric parameters, solar and view geometry, and surface reflectance, the at-sensor or top-of-atmosphere (TOA) radiance can be calculated with a radiative transfer code. With remote sensing data from a satellite instrument we have to solve the inverse problem, namely, to derive the atmospheric parameters and surface reflectance from the TOA radiance.

A popular approach, also taken for EnMAP, is the calculation of a database of look-up tables (LUTs). We used the MODTRAN5.4 radiative transfer code (IR-2, IR-3, IR-4, Berk et al. 1998, 2003, 2008) to generate a large database with the required atmospheric functions (path radiance, ground-to-sensor direct and diffuse transmittance, direct and diffuse solar flux, spherical albedo) and a multidimensional linear interpolation is performed to obtain the atmospheric parameters. The estimation of the aerosol optical thickness (AOT at 550 nm) (or its equivalent visibility, in km) and the columnar water vapour per pixel will be described in Section 5.4 and Section 5.5, respectively.

The final surface reflectance cube is computed using simulated radiative transfer (RT) functions (Section 5.2) in provided Look-Up-Tables (LUTs), which include bins in simulated atmospheric aerosol and water vapour content.

The radiative transfer equation for a homogeneous surface under clear sky conditions can be formulated as:

$$L = L_p + \frac{\tau \left(E_{dir} \cos\theta_s + E_{dif}\right) \rho / \pi}{1 - \rho s}$$

Equation 4.2-1

where L, L_p, τ, E_{dir}, E_{dif}, θ_s, ρ and s are at-sensor radiance, path radiance, ground-to-sensor transmittance, direct and diffuse solar flux on the ground, solar zenith angle, surface reflectance, and spherical albedo of the atmosphere, respectively. The total transmittance is the sum of the direct and diffuse transmittances τ

= τ_{dir} + τ_{dif}. For brevity, in Equation 4.2-1, the dependence on wavelength, solar and viewing geometry, and atmospheric parameters has been omitted. A Lambertian or isotropic reflectance law (IR-27, Nicodemus 1970) is assumed here, as no knowledge of the bidirectional reflectance distribution function (BRDF) is available prior to the atmospheric correction.

5.1 Masks for land, water, haze, cloud, shadow

The same land / water/ cloud pre-classification mask (see Section 5.1.1) is used as input for the atmospheric correction for both L2A processors (land and water) before executing an atmospheric correction. The final EnMAP masks will merge the results from this internal pre-classification map with the mask information extracted from the L2A water processor.
The final EnMAP masks are described in Section 5.1.2.

5.1.1 Pre-classification masks

The L2A processor calculates a pixel map for cirrus, haze, cloud, snow and land/water. It is an intermediate product used and produced by the L2A land processor. The following Table 1 contains the class labels of this map. It shows one class for cloud (meaning water cloud as opposed to cirrus cloud). The low optical thickness cloud is put into the haze classes which can often be successfully eliminated.

Label	Class	Label	Class
0	geocoded background	10	Thick cirrus over land
1	Shadow over land	11	Haze over land
2	Thin cirrus over water	13	Haze over water
3	Medium cirrus over water	15	Cloud over land
4	Thick cirrus over water	16	Cloud over water
5	Land	17	Water
6	Saturated (blue/green band)	18	Cirrus cloud
7	Snow/ice	19	Thick cirrus cloud
8	Thin cirrus over land	20	Bright pixels
9	Medium cirrus over land	21	Topographic shadows

Table 1 Enhanced classes (land, water, haze, cloud, etc.) internally generated in the L2A land processor

Most of the classes will be defined with spectral criteria, some might need additional information (e.g. “cloud over water” is difficult or impossible to distinguish from “cloud over land” with purely spectral information) which may not be available.

The apparent reflectance used for the masking is calculated on a per-pixel and per-wavelengths basis. This approach is efficiently applied for all other hyperspectral sensors without map quality degradation. The cirrus band apparent TOA reflectance is averaged over bands in the 1.36-1.40 µm region to reduce the noise level.

The algorithm initially assigns the label ‘land’ to all pixels. A pixel belongs to ‘geocoded background’ if the digital number of the red band (around 650 nm) is zero. The red band is preferred to the first (or blue band) since the blue band of hyperspectral imagery is often noisy and contains many “zero” pixels. The proposed spectral thresholds are empirical and similar to those used in the automated cloud-cover assessment (ACCA) of Landsat ETM+ (IR-20, Irish et al. 2006) and MODIS (IR-1, Ackerman et al. 1998), but experience from the processing of other multispectral sensor imagery (e.g., Ikonos, ALOS-AVNIR2, SPOT-5 etc.) is also being used in defining the thresholds.

For the description of the different thresholds apply for the different class masks, the symbols $\wedge$, $\vee$ denote a logical AND, OR, respectively. The asterisk (*) marks the apparent reflectance:

$$\rho^* = \frac{\pi\ L\ d^2}{E_s\ \cos\theta_s}$$

Equation 5.1-1

where *L* is the recorded radiance signal, E_s is the extra-terrestrial solar irradiance for the selected band, d is the Earth-Sun distance (in astronomical units), and θ_s is the solar zenith angle.

The spectral bands chosen to establish thresholds in apparent reflectance are those closer to the following values: blue = 0.44 µm, green = 0.549 µm, SWIR1 = 1.656 µm, SWIR2 = 2.133 µm, red = 0.659 µm, NIR = 0.84 µm, cirrus = 1.382 µm.

For some masks, the terrain profile is also considered: the cosine of the incidence angle between the solar ray and the surface normal for each pixel (cbeta / 255) and the slope.

In case all bands inside some of the wavelength range values (e.g. blue, red, NIR, ...) are missing, an alternative set of bands and thresholds is implemented.

In the simplified atmospheric correction (L1Bint), these masks are calculated for the VNIR cube, since at least one of the bands in the visible wavelength range is used in the mask criteria.

5.1.1.1 Land class

The pre-classification map is initialized by assigning all pixels to the land class. The consecutive masking algorithms will change the classification of the different according to the masking criteria.

Therefore, at the end the land mask will contain the land pixels plus those pixels that do not enter in any of the following criteria.

5.1.1.2 Water class

a. If the surface elevation of a pixel is lower than 1.2 km above sea level, then the water criterion is

$$\rho^*(blue) \leq 0.2\ and\ \rho^*(blue) > \rho^*(green) - 0.03\ and\ \rho^*(NIR) < \rho^*(green)\ and\ RVI < 1.2\ and\ \rho^*(SWIR1) < T_{water,SWIR1} + 0.015\ and\ cbeta > 26$$

Equation 5.1-2

where $T_{water,SWIR1}$ is the water threshold reflectance in the short-wave infrared SWIR1 band (around 1.6µm) as defined in the preference parameter file (default $T_{water,SWIR1}$ = 0.03).

RVI is the ratio between the apparent reflectance in the NIR and red bands.

b. If the pixel elevation is higher than 1.2 km, the criterion of a negative gradient for the apparent reflectance does not properly work (as the path radiance becomes small) and the following rules with the surface reflectance instead of apparent reflectance are used

$$\rho^*(NIR) < T_{water,NIR} + 3\ and\ RVI < 1.2\ and\ \rho^*(SWIR1) < T_{water,SWIR1}\ and\ cbeta > 26$$

Equation 5.1-3

If no SWIR channels are available, the condition in the ρ* (SWIR1) is excluded.

Note on the water mask:

Water masks are obtained with the criterion that the *TOA reflectance* $\rho*$ spectrum must have a negative gradient for bands in the visible to NIR:

$$\frac{d\rho^*(\lambda)}{d\lambda} < 0 \quad for \quad 0.4 < \lambda < 0.85\ \mu m$$

Equation 5.1-4

However, for a scene with a high average ground elevation (we use > 1.2 km above sea level) the *TOA reflectance* criterion is again replaced with the NIR *surface reflectance* criterion, because of the distinctly smaller path radiance. The water reflectance threshold for a 1600 nm band is always included as the second criterion (if such a band exists).

Sometimes the gradient criterion with Equation 5.1-4 is not adequate, and the NIR / SWIR1 reflectance thresholds yield a better water mask. This may happen in urban areas containing shadow pixels cast by buildings. Then the NIR / SWIR1 thresholds have to be defined as negative reflectance values to overwrite the gradient criterion.

In addition, all the pixels with slope > 7 degrees are excluded from the water classification. This avoids false water/shadow classification in mountains due to cloud shadows.

5.1.1.3 Saturated pixel (blue and green bands)

For EnMAP snow is considered in the linear dynamic range of the instrument and the radiometric encoding with 14 bits / pixel. Other effects such as specular reflection from water or buildings can lead to saturation in any band.

$$DN\ (blue) \geq\ T_{saturation}\ or\ DN\ (green) \geq\ T_{saturation}$$

Equation 5.1-5

where *DN(blue)* is the digital number in a blue band (around 470 nm) and the threshold $T_{saturation}$ is defined in the preference parameter file. $T_{saturation}$ = *b*encoding* (default b=0.9), encoding = 2^n where n is the number of bits. Therefore $T_{saturation}$ = 14745. The default value b=0.9 is chosen instead of the obvious b=1.0. Factor b is a saturation factor applied to the maximum radiometric encoding, e.g. for 8-bit data and b=1 all pixels with DN=255 will be marked as saturated. Setting b=0.9 implies pixels with *D*N > 230 will be considered as saturated or in the non-linear radiometric range. This factor is only used for 8 and 16-bit (signed or unsigned) data, not for float or 32-bit integer data.

5.1.1.4 Cloud over land (non-cirrus)

Pixels must satisfy the conditions:

$$\rho^*\ (blue) >\ T_c\ \ and\ \rho^*\ (red) > 0.15\ \ and\ \ \frac{\rho^*(NIR)}{\rho^*\ (red)} < 2.\ and\ \rho^*\ (NIR) > 0.8 * \rho^*\ (red)\ \ and\ \ \frac{\rho^*(NIR)}{\rho^*\ (SWIR1)} > 1\ \ and\ NDSI < 0.7\ \ \ and\ \ \ DN\ (blue) >\ T_{saturation}$$

Equation 5.1-6

Here, T_c is the cloud threshold, default T_c = 0.25, and NDSI is the normalized difference snow index.
For scenes with SZA > 45°, T_c is reduced to 0.2 to detect a larger fraction of optically thick clouds at low illumination conditions.

$$NDSI = \frac{\rho^*(green) - \rho^*(SWIR1)}{\rho^*(green) + \rho^*(SWIR1)}$$

Equation 5.1-7

Note that saturated pixels in visible bands are automatically counted as cloud (if NDSI < 0.7) although they might be something else (e.g., specular reflection from a surface).

If now SWIR channels are present, the conditions on SWIR1 channel (and therefore on NDSI) are excluded.

5.1.1.5 Cirrus cloud

Cirrus labeling is based on a physical model and uses the apparent reflectance in the 1.38 µm band. For the cirrus mask layers, the following fixed thresholds for thin, medium, and thick cirrus are used:

- Thin cirrus: 0.01 ≤ ρ* (1.38µm) < 0.015
- Medium cirrus: 0.015 ≤ ρ* (1.38µm) < 0.025
- Thick cirrus: 0.025 ≤ ρ* (1.38µm) < 0.04
- Cirrus cloud: 0.04 ≤ ρ* (1.38µm) < 0.05
- Cirrus cloud thick: ρ* (1.38µm) > 0.05

The first three cirrus classes usually can be successfully treated during the cirrus removal. The two last classes are defined for cirrus clouds, where cirrus removal might not be successful. Further details will be described in Section 5.7.
Those pixels with ρ* (1.38µm) > 0.1 will be excluded from the snow mask.

5.1.1.6 Cloud shadow

In order to be masked as cloud shadow, pixels must satisfy the spectral conditions:

$$\rho^*(blue) \leq 0.15 \; and \; 0.04 < \rho^*(NIR) < 0.12 \; and \; \rho^*(SWIR1) < 0.2 \; and \; and \; NDWI \leq 0.06$$

Equation 5.1-8

where NDWI is the normalized difference water index using green and NIR bands.

To minimize the confusion with semi-topographic shadows all pixels with cosβ < 0.78 are also excluded.

5.1.1.7 Snow

For the snow mask, pixels must satisfy the conditions:

$$(\rho^*(green) \geq 0.22 \; and \; NDSI \geq 0.4) \; or \; \left(\rho^*(green) \geq 0.22 \; and \; NDSI > 0.25 \; and \; \rho^* \frac{(SWIR2)}{\rho^*(green)} < 0.5\right)$$

Equation 5.1-9

As a last step, all the successfully classified snow pixels are removed from the cloud mask.

5.1.1.8 Cloud over water

Spectrally is not possible to determine the land/water nature of the pixels below clouds. Therefore, an external mask for land / water is provided as input of EnMAP processor (blue container in Figure 5-1).

EnMAP uses the land/water mask from the CIA world data bank II, http://www.evl.uic.edu/pape/data/WDB/ (GSHHG database in Figure 5-1) to resolve the land/water pixels under clouds.

5.1.1.9 Haze over land and haze over water

Haze class is delineated by a statistical analysis of the bands in the visible range. A more detailed description can be found in Section 5.6.

5.1.2 EnMAP quality layers

Table 2 shows the final EnMAP quality layers obtained in all the processing levels, from L0 to L2A and their correspondence to the internal pre-classification labelling described before.

In the Classes quality layer, the land pre-classification mask is corrected by the water processor for a more accurate water mask classification in EnMAP final product (QL_QUALITY). A more detailed description of the EnMAP labelling inside each class is described in the L0 and LX EnMAP products specifications (AR-3 and AR-12, respectively).

Quality Layer	EnMAP **label** in quality layers (corresponding pre-classification label)
Classes	**1** (5, 20, 1), **2** (17), **3** (0)
Cloud	**1** (15, 18, 19)
Cloud shadow	**1** (1)
Haze	**1** (11, 13)
Cirrus	**1** (2, 8), **2** (3, 9), **3** (4, 10)
Snow	**1** (7)

Table 2: EnMAP quality layers products (L0 and LX product specifications) with the correspondence between their label and the pre-classification mask (between brackets) in Table 1.

5.1.3 EnMAP L2A Metadata

Within the Metadata of every EnMAP Product a range of parameters describing the atmospheric conditions and scene content are generated by the L2A processor, incl.

- scene average AOT
- scene average WV
- scene average SZA
- cloud coverage
- cirrus coverage
- haze coverage
- snow coverage
- water coverage
- cloud shadow coverage
- non-cloud shadow coverage
- quality rating of the atmospheric correction

For more details, see AR-3 and AR-12.

Also, at the end of the processing chain, the WWF (World Wildlife Fund) ecoregions data base is queried for the EnMAP main tile location to assign a global ecoregion (biome) to the metadata file. These ecoregions are predefined location patterns across the Earth surface.

The scene average AOT metadata value is extracted from land/water pixels if the processor run in Land or Water Mode, respectively. In “Combine mode”, both AOT values are considered, but each value is weighted with the number of classified pixels as land and water.

The scene average WV is calculated over the clear land pixels for both Land Mode and Combined. For Water_Mode the value is -999 (undefined).

The band statistics in L2A are in units of irradiance reflectance (independent of solid angle) for both Land/Water and Combined.
For Combined and Water Mode, the spectral image cube and the band statistics in AR-12 include the water products selected in the job order, water leaving or subsurface irradiance reflectance.

5.2 Radiative transfer calculations and atmospheric database

In principle, MODTRAN offers several algorithms for multiple scattering and absorption in the earth’s atmosphere: DISORT (discrete-ordinate radiative transfer, IR-40) with and without the azimuth-dependence of the multiple scatter, Isaac’s multiple scatter (IR-21), and correlated k (IR-2) where k indicates the absorption coefficient. The correlated k (CK) approach reformulates the transmittance for a given atmospheric layer as a weighted sum of exponential terms with monochromatic absorption coefficients and probability-depending weighting factors to interface the MODTRAN band model technique with multiple scattering algorithms. After extensive testing, we decided to compile the atmospheric database with the azimuth-dependent scaled DISORT employing 8 streams for the atmospheric window regions (IR-17). The scaled DISORT performs DISORT calculations at 11 fixed wavelengths in window regions and scales Isaac’s results to obtain a very fast algorithm. For absorption regions we use the accurate, but time-consuming CK algorithm (IR-17). A straightforward method would use the CK algorithm over the whole spectral range at the highest spectral resolution (1 cm^{-1}). However, this approach is not efficient in terms of computer resources and the referenced paper (IR-17) demonstrates that a sufficient accuracy is obtained when employing the CK method to the absorption regions 680 - 840 nm, 890 - 990 nm, 1080 - 1240 nm, 1950 - 2100 nm, and 2300-2500nm.
Calculations were performed for temperature/humidity profiles of the so called MODTRAN mid-latitude summer and mid-latitude winter atmosphere model (see Appendix A) with a CO_2 content of 400 ppmv and an ozone column of 330 DU (sea level geometry). The selected CO2 concentration is accurate enough for the spectral range of EnMAP. The variation was tested for the 360-440 ppmv with no influence on the EnMAP spectral range.

The different atmospheric profiles are simulated yielding two different sets of radiative transfer LUTs. The selection of each of the LUTs (so called only “summer” or “winter”) will be done through the parameter “season” (see Chapter 5.10.2) and explained in more detail in Chapter 5.2.2.

5.2.1 Ozone

Ozone will not be retrieved on a per-pixel basis due to its low spatial and temporal variation, instead a constant value of 330 DU (mid-latitude summer) and 377 DU (mid-latitude winter) is taken for sea level, decreasing with elevation, when no external information is available (see Section 5.10).

On a global scale, extreme values of 200 – 500 DU can occur, but a more typical range is from 250 - 420 DU. Figure 5-2 shows that relative surface reflectance retrieval errors of up to about 5 - 7 % can be expected in the 550 – 600 nm spectrum with the simplification of a constant 330 DU concentration if compared to concentrations of 250 DU and 410 DU and two reflectance levels (ρ=0.05 and ρ=0.15). All other molecular absorbers in the atmosphere (e.g., O_2, CO, NO_x, CH_4) are assumed to have constant mixing ratios except for water vapor which varies significantly in space and time.

If external ozone information is provided, the corresponding ozone column (in Dobson Units) is specified in the input parameter file. The effect is considered in a separate ozone LUT covering the spectral range 450 – 800 nm, for solar zenith angles in the range 0° – 70° (increment 10°), and view zenith angles 0°-40° (increment 10°), see [IR-35].

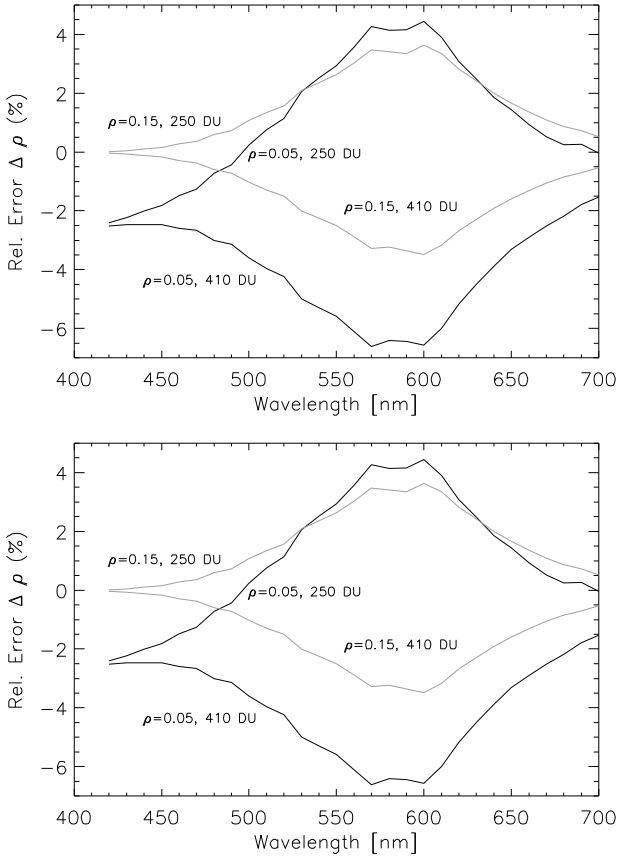


Figure 5-2 Influence of ozone on surface reflectance retrieval.

The influence of ozone is also to be included. This information is taken from an external source (Copernicus database). The default ozone column in the LUTs is 330 DU (Dobson Units) for sea level for summer (377 for winter), and since ozone can be decoupled from the other atmospheric parameters, a separate ozone LUT is compiled. In the L2A land processor it covers the ozone concentrations 200 - 500 DU, to avoid extrapolations. User selected values outside these limits will be reset by the processor to the above-mentioned range, decreasing the quality of the product.

The LUT is calculated for the wavelength range 450 - 800 nm, on a 10 nm grid, and a 10 nm spectral averaging, which is sufficient because the ozone spectral transmittance does not vary rapidly. The supported range of satellite view zenith angles is 0°-40°, increment 10°, and the supported solar zenith angle range is 0°-70°, increment 10°. The ozone LUTs are different depending on the season (summer or winter) and both are provided (ozone_luts.bin and ozone_luts_winter.bin).

Due to the small spatial variation of ozone, the ozone value is considered as constant per scene. The ozone value is specified in the input parameter or extracted from external information (if available) by the processor.

Since the ozone content peaks in the 20 - 40 km altitude region, the variation of ground elevation can be neglected for elevations 0 - 4 km (the influence is smaller than about 3%).

If the user requires the option “automatic” for the ozone retrieval and Copernicus TCO3 ozone values, not older than 30 days, will be loaded and the ozone value will be used for the atmospheric correction.

In the simplified atmospheric correction, a default value used in the generation of the monochromatic LUTs (330 DU) is considered, so no ozone correction will be performed.

5.2.2 Atmospheric database and radiative transfer LUT

The current atmospheric database is recompiled with the MODTRAN5.4.0 with 0.4 nm grid. This is the final high spectral-resolution (so called “monochromatic”) database suitable for instruments with bandwidths ≥ 3 nm.

This database is needed for the resampling of the seven atmospheric functions (path radiance, direct and diffuse transmittance ground-to-sensor, direct and diffuse solar flux, spherical albedo, and transmittance for the sun-to-earth radiance) with the channel filter curves. If the spectral calibration changes during the lifetime of a mission, the resampling of the atmospheric LUTs has to be updated.

The monochromatic database is also needed for “smile” instruments because the change of the spectral calibration across the detector array requires a spectral resampling in column direction (see Section 5.9). Note: “smile” refers to an optical aberration of push broom instruments that causes the spectrometer entrance slit, representing the across-track swath, to be projected as a curve on the rectilinear detector array (AR-15).

Table 3 contains the grid points of the atmospheric database in the employed 6-dimensional parameter space. The simulation grid is the same for all parameters for both atmospheric models, except for the water vapor column. The default aerosol type (rural) will be fixed for the operational processing to avoid possible brightness steps in surface reflectance at scene borders that might be caused by a change of aerosol type.

Parameter / Bin	#1	#2	#3	#4	#5	#6	#7	#8
VIS (km)	5	7	10	15	23	40	80	120
SZA (deg)	0	10	20	30	40	50	60	70
RAA (deg)	0	30	60	90	120	150	180	
ELE (km)	0	0.7	1.5	2.5	4.0			
TILT (deg)	0	10	20	30				
WVC (cm)	0.4	1.0	2.0	2.9	4.0	5.0		summer
WVC (cm)	0.2	0.4	0.8	1.1				winter

Table 3: Grid points schema for the monochromatic atmospheric database. (VIS= visibility, SZA = solar zenith angle, RAA = relative azimuth angle between sun and observer line-of-sight, ELE = ground elevation above sea level, TILT=instrument tilt angle, WVC = water vapor column)

The LUT elevation range is 0 – 4 km above sea level. Higher elevations up to 5.5 km are calculated with linear extrapolation of the radiative transfer terms, but elevations > 5.5 km are treated as a 5.5 km elevation.

Table 3 contains the water vapor grid points for the baseline atmosphere (mid-latitude summer) covering the range of 0.4 – 5 cm, values greater than 5 cm will be extrapolated up to 5.5 cm. Water vapor grid points for mid-latitude winter conditions contain a smaller number of grid points because of the smaller dynamic range of 0.2 – 1.1 cm. Values greater than 1.1 cm are extrapolated up to 1.6 cm. For values over 5.5 cm the reduced accuracy flag is raised (see Section 14).

The current baseline for atmospheric correction of EnMAP data is the mid-latitude summer or mid-latitude winter atmosphere with the large dynamic range of water vapor concentrations. Hourly average global air temperature (2m above surface) maps from Copernicus ERA5 database are used to enable a decision for

the each of the atmospheric models LUTs. The mid-latitude winter LUTs are selected if the scene mean LST of the scene is lower than 8 degrees Celsius (see Section 5.10.2 for details).

Figure 5-3 shows some examples of surface reflectance retrievals from a synthetic EnMAP mid-latitude winter scene with a water vapor column of 0.8 cm, where retrievals are calculated twice, with the corresponding winter atmosphere (black curves) and with the standard mid-latitude summer atmosphere (grey curves). Relative RMS water vapor retrieval errors are 2.2% (winter LUTs) and 8.1% (summer LUTs). The larger error for the summer LUTs causes larger surface reflectance retrieval errors in spectral regions affected by water vapor. This is especially visible for the soils and the constant reflectance targets. The difference in the 0.4 – 0.7 μm region is caused by the different ozone contents (377 DU for the winter scene, 330 DU for the summer LUTs).
A moderate deviation of the vertical temperature / humidity profile (e.g. tropical atmosphere versus mid-latitude summer or US standard versus mid-latitude summer) causes only small relative water vapor retrieval errors (about 2%, IR-33) which is within the accuracy of the APDA retrieval method itself.

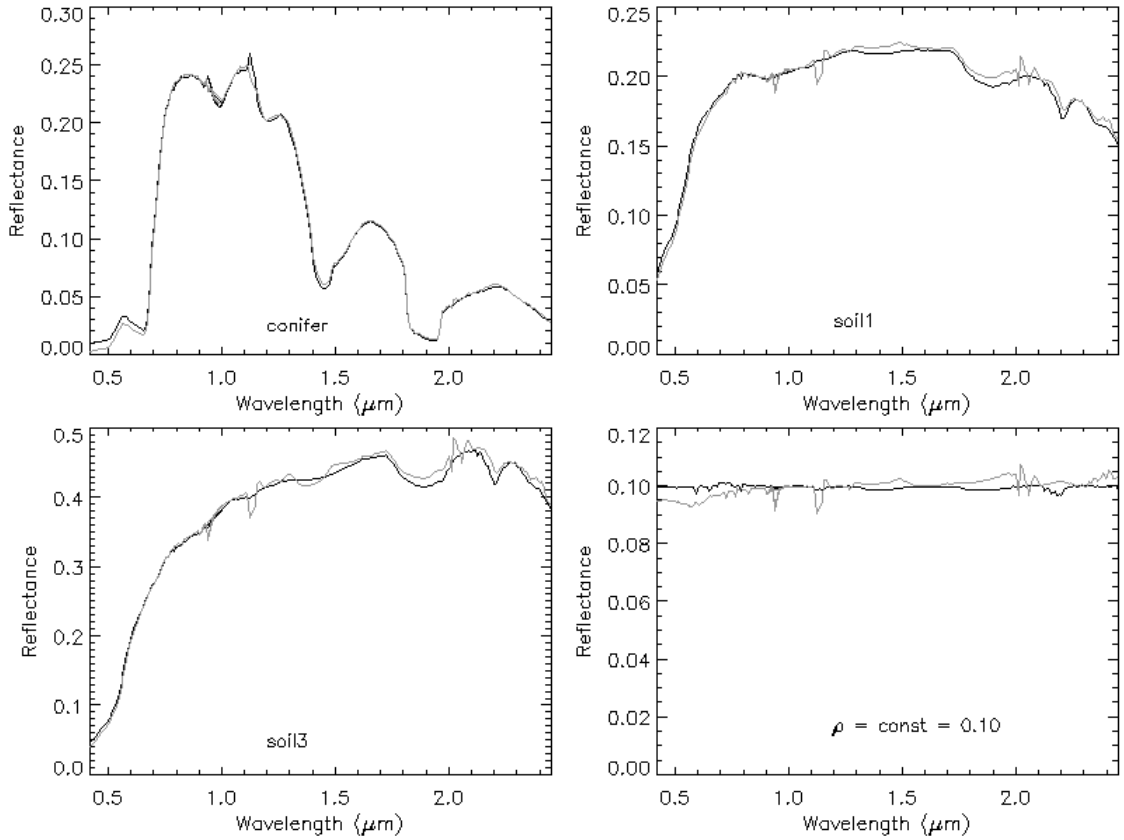


Figure 5-3 Retrievals for the synthetic winter scene (black: winter, grey: summer LUTs).

5.2.2.1 Size of high spectral resolution database

The file size of the high spectral resolution (0.4 nm) database needs some consideration because it is very large. The file size per water vapor grid point is more than 200 MB, so with the 6 water vapor grid points for the summer atmosphere we have to store 1.2 GB. If we include the winter conditions we need more than 2 GB disk space for a fixed ozone concentration. Season dependent ozone LUTs are also provided, with a size of 12 kB. In addition, about 600 MB of disc space are required to store the global set of monthly averaged land surface temperature data from Copernicus C3S service.

These high resolution LUTs will be needed for the processing of each scene if the spectral smile effect has to be accounted for. Otherwise, the resampled LUTs to the EnMAP channel spectral response

functions will be used, not only in the complete atmospheric correction (L2A) but also in the simplified one (L1Bint). The resampled LUTs will change each time there is an update of the calibration tables and will be part of the internal EnMAP products.

A detailed flow chart is given in Section 11 and the LUT file names are documented in Section 12.

5.3 Extraterrestrial solar irradiance spectrum

Although the solar constant (wavelength-integrated extraterrestrial solar irradiance) is known with an accuracy of about 1% there exist large differences between published solar irradiance spectra. Figure 5-4 shows the relative differences between the solar irradiance model used in EnMAP (Fontenla 2011) and two other solar spectra: Thuillier 2003 and TSIS 2021 [IR-44]. Thullier 2003 is the solar irradiance model used in Sentinel-2 missions from the Copernicus program while TSIS-2021 is the latest solar irradiance spectrum recommended by CEOS.
The difference between the most recent solar irradiance measurements (TSIS-2021) and EnMAP implemented model are below 3% for the full wavelength range of EnMAP, except for some wavelengths in the blue range (~6%). But these differences are smaller than differences with other spectra like Thuillier 2003.

However, a recent study [IR-45] reveals that there is no influence in the L2A results if the same solar model irradiance is used both in the calibration coefficients and in the L2A radiative transfer simulations. This is the scenario implemented in the EnMAP mission.

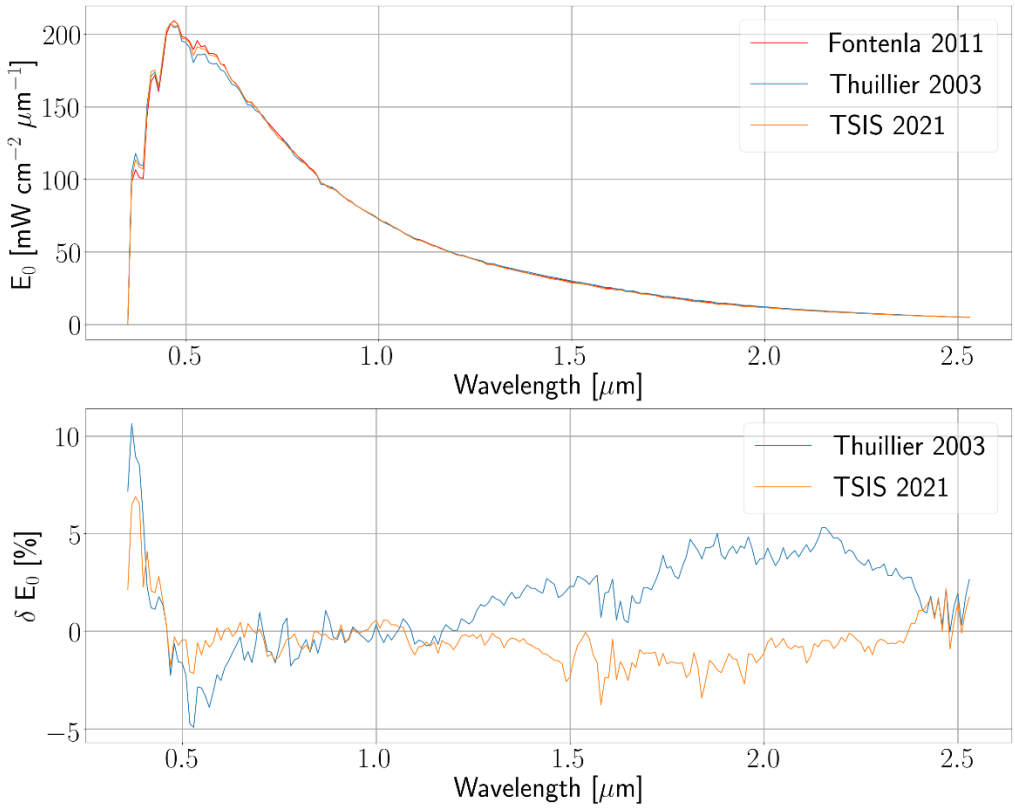


Figure 5-4: Top: Solar irradiance spectra of Fontenla 2011, Thuillier 2003 and TSIS 2021 convolved to 10 nm. Bottom: Relative difference of the above spectra (in %) between Thuillier 2003 and TSIS 2021 with respect to EnMAP solar irradiance model (Fontenla 2011). [IR-45]

The atmospheric database of the land processor compiled with MODTRAN5.4.0 used the Fontenla 2011 [IR-10] (MODTRAN option LSUNFL=5) extraterrestrial solar irradiance spectrum coming with the MODTRAN version distribution.

Both L2A processors (land and water) have their own atmospheric database LUTs but both use the same solar irradiance spectrum. Note that the E0 spectrum used will also be made available on the EnMAP web page, thereby fulfilling the CEOS recommendations and allowing for an easier validation by 3rd parties.

5.4 Retrieval of aerosol optical thickness

An automatic aerosol retrieval over land can be performed if the scene contains dark reference areas (dense dark vegetation, DDV, IR-22, Kaufman et al. 1997). Then the method can determine the aerosol type and optical thickness.

If there are not enough reference pixels found, the scene is processed with a constant visibility of 23 km.

The SWIR2 2.1 µm channel is employed to mask dark pixels in the scene. An atmospheric correction is performed for this channel using the rural aerosol and a visibility of 23 km. With these assumptions, the obtained surface reflectance map $\rho_{2.1}(x, y)$ is slightly better than with the usual assumption of $\rho_{2.1} = \rho^*_{2.1}$ employing the apparent reflectance instead of the surface reflectance. Pixels with $0.01 < \rho_{2.1} < 0.05$ are labeled as dark pixels suitable for the derivation of the visibility or aerosol optical thickness (AOT). Very dark pixels (water) have to be excluded with $0.01 < \rho_{2.1}$. If more than 2% of the scene pixels are assigned as dark, then the spectral correlation

$$\rho(0.66\mu m) = 0.5\,\rho(2.1\mu m) \quad and \quad \rho(0.47\mu m) = 0.5\,\rho(0.66\mu m) + 0.005$$

Equation 5.4-1

is used to calculate the reflectance in a red and blue channel. The spectral correlation coefficient (default =0.5) is a parameter in the processor design. A fixed correlation coefficient is used globally because no seasonal global maps of this coefficient are available for the 30 m spatial resolution of EnMAP.

If the percentage of dark pixels is less than 2% the SWIR reflectance threshold is increased to $0.01 < \rho_{2.1} < 0.10$ to include medium brightness pixels, if there are still less than 2% of reference pixels the threshold is finally increased to $0.01 < \rho_{2.1} < 0.12$. However, the accuracy of the spectral correlation decreases with increasing surface reflectance of $\rho_{2.1.}$

If the reflected signal for the blue and red channels is subtracted from the respective total radiance, the blue/red channel path radiances are obtained. A comparison of L (path, scene, blue) / L (path, scene, red) with the corresponding ratio of MODTRAN aerosols (rural, urban, maritime, desert) can determine the closest MODTRAN aerosol type. However, for the operational processing we will use the rural (continental) aerosol type and only derive the visibility or aerosol optical thickness (AOT) at 550 nm to avoid possible brightness steps in surface reflectance at tile borders caused by a change in the aerosol type.

Once the surface reflectance ρ(red), ρ(blue) is known for each dark pixel, the at-sensor radiance L(VIS) for the current viewing and solar geometry can be calculated as a function of the visibility (or AOT at 550nm) using the precalculated LUTs. The intersection of the measured radiance L(measured) with the calculated curve L(VIS) determines the per-pixel visibility which is subsequently smoothed over a 3 km box to reduce noise and decrease the influence of small-scale fluctuations in the spectral correlations (Equation 5.4-1).

Additionally, the surface reflectance of water bodies is checked in a red (660 nm) and NIR (850 nm) band if the scene mean visibility is below 110 km. If the selected visibility is too low and causes negative reflectance pixels in the red or NIR bands, it will be iteratively increased up to VIS=100 km. This check of negative reflectance pixels is done also if not enough reference pixels are found, so the default visibility of

23 km assumed in this case might be increased only if water bodies or dark pixels show more than 1% of negative reflectance pixels in the red and NIR.

It is assumed that the main contribution to the AOT is from aerosols in the lower troposphere (0-2 km boundary layer), small contributions from the upper troposphere (2 – 10 km) and stratosphere are defaulted as per MODTRAN model (e.g. mid-latitude summer).

Details of the aerosol retrieval: after the masking of the reference pixels their surface reflectance in the SWIR2 2.1µm channel is calculated with the assumption of a visibility of 23 km. Then Equation 5.4-1 is employed to calculate the surface reflectance in the red band, and subsequently the visibility for each reference pixel.

After the visibility of the reference pixels and water bodies is determined, in order to remove the influence of instrument noise and variations in the spectral correlation coefficient, a spatial averaging filter (moving window box of 3 km x 3 km) is applied, assigning the average visibility to all gap (i.e., non-reference and no water/dark) pixels. For a faster array addressing the float visibility values (between 5 km and 190 km) are mapped to a discrete integer grid named visibility index (visindex). The visindex=0 corresponds to visibility 190 km, and the visindex increment of 1 corresponds to an AOT increment of 0.006 (at sea level). A second reason for working with the visibility (and corresponding visindex) is that the visibility has only a slight dependence on surface elevation, while AOT varies strongly near steep mountain ridges. Thus, a spatial averaging of visibility/visindex in mountainous regions is not critical. Therefore, the surface reflectance retrieval is performed with the visibility / visindex parameter, and the AOT map is calculated from the visibility.

With the MODTRAN code the AOT (at 550nm) can be calculated for a given visibility as

$$AOT\ (550\ nm) = e^{a(z)+b(z)\ln(VIS)}$$

Equation 5.4-2

where z is the surface elevation (elevation simulated bins), and a(z), b(z) are coefficients obtained from a linear regression of $\ln(\mathrm{AOT})$ versus $\ln(\mathrm{VIS})$.

For the simplified atmospheric correction, a default value of 23 km is used for the complete tile and for each sensor (VNIR and SWIR).

5.4.1 Notice concerning visibility iterations:

The land processor for EnMAP always computes the visibility map based on dark reference pixels in the scene and overwrites the initial value of the *visi*b parameter.

The processor will automatically iterate the initial visibility (parameter *visib,* or the calculated visibility from reference pixels) if the number of negative reflectance pixels is larger than 1% of the scene for the red band (around 650 nm, vegetation is checked here) or the NIR band (around 850 nm, water is checked here). The lowest visibility index (highest visibility) is always kept and included in the calculation of the tile visibility.

Since the visibility is based in the DDV and water/dark pixels present in the scene (EnMAP tile), consecutive tiles might contain different information and therefore different scene visibility. This can create a step in the visibility estimation, visible at the tile borders. In order to avoid the visibility step between consecutive tiles borders (with distances below 60 meters), the adjacent tiles will be merged before the iteration to check the pixels with negative reflectance starts. In addition to this the merged tiles smooth includes the information of the neighbor tiles. Finally, a low pass filter with a 3 km x 3 km box is applied. The procedure for the tile merging is described in chapter 6.1.

5.5 Retrieval of columnar water vapor

The per-pixel columnar water vapor map will be calculated with the APDA (atmospheric pre-corrected differential absorption) algorithm (IR-36, Schläpfer et al. 1998) using measurement channels around 0.94 / 1.13 µm and appropriate reference or window channels around 0.88, 1.02 µm, and 1.08, 1.24 µm, respectively. Since this algorithm is not valid over dark surfaces (water), the water vapor columns over water surfaces will be set to the average value obtained over land. Other methods for water vapor retrieval (i.e. band fitting algorithms) were not selected since of much higher requirements on runtime and a difficulty to run on an operational basis.

The basic method uses three channels, one in the atmospheric water vapor absorption region around 940 or 1130 nm (the "measurement" or “absorption” channel), the others in the neighboring window regions ("reference" or “window” channels). The depth of the absorption feature is a measure of the water vapor column content, see Figure 5-5.

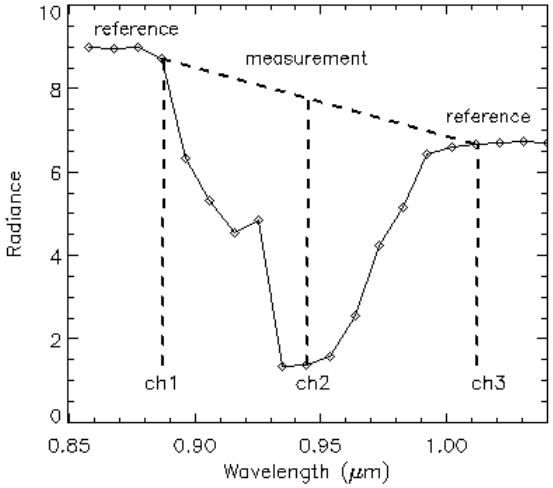


Figure 5-5: Schematic sketch of APDA method with three channels.

For three channels the water vapor dependent APDA ratio is calculated as

$$R_{APDA}(\rho,u) = \frac{L_2(\rho_2) - L_{2,p}}{w_1\left(L_1(\rho_1) - L_{1,p}\right) + w_3\left(L_3(\rho_3) - L_{3,p}\right)}$$

Equation 5.5-1

where the index 1 and 3 indicates window channels (e.g., in the 850 – 890 nm region and 1010 – 1050 nm region), respectively. Index 2 denotes a channel in the absorption region (e.g., 920 - 970 nm). *L* and L_p are the total at-sensor radiance and path radiance, respectively. The symbol *u* indicates the water vapor column. The weight factors are determined from

$$w_1 = (\lambda_3 - \lambda_2)/(\lambda_3 - \lambda_1) \quad \text{and} \quad w_3 = (\lambda_2 - \lambda_1)/(\lambda_3 - \lambda_1)$$

The problem is the estimation of the surface reflectance ρ_2 in the absorption band. Usually a linear behavior of surface reflectance is assumed across the absorption region which might influence the water vapor retrieval accuracy if this assumption is violated. The fast and robust APDA method is often used for operational processing (Chylek et al. 2003). The R_{APDA} expression can be approximated with an exponential fit function using 4 – 6 water vapor grid points for the least square’s regression (depending on the water vapor range, see chap. 5.3)

$$R_{APDA}(u) = a_0 \, exp(-a_1 \, u^{a_2})$$

Equation 5.5-2

which can be solved for the water vapor column *u:*

$$u = \left(\frac{ln \, R_{APDA} / a_0}{-a_1} \right)^{1/a_2}$$

Equation 5.5-3

The fit coefficients are calculated iteratively starting with a0=1, a1=0.01, a2=0.5. More than three channels can be used with the linear regression ratio (LIRR) method (IR-36, Schlaepfer et al. 1998).

Concerning the selection of channels for water vapor retrieval: it is preferable to employ measurement channels in the 900 – 930 nm region to avoid an overlap with the spectral absorption feature of liquid water in plants. Measurement channels in the 1130 nm region could also be employed. The final decision will be made based on the noise performance in the respective bands, presumably using more than 3 bands to apply a linear regression ratio.

Note on the non-linear influence of vegetation in water vapor calculations:

This option applies to the APDA-regression algorithm and only if the 940 nm region is selected for the water vapor retrieval. The retrieval is based on a linear interpolation across the absorption region, and errors can occur due to the non-linear behavior of the reflectance of vegetated surfaces in this region. A simple empirical correction to the water vapor map W is applied using the NDVI calculated with the apparent reflectance in the red/NIR channels

$$W(new) = W(old) - 0.1(NDVI / 0.7) \quad (cm)$$

The correction is only performed for pixels with *NDVI* ≥ 0.25 and values *NDVI* > 0.7 are reset to 0.7.

To minimize the pixel-to-pixel noise, the final water vapor estimation for the scene is smoothed over 100m. In order to avoid water vapor steps in neighbor pixels at the border of adjacent tile, the tile smooth approach is also applied here (see Chapter 6.1).

5.5.1 Water vapor retrieval for smile sensor

If spectral “smile” has been taken into account, the water vapor retrieval has to be performed on a per-column basis. With the updated EnMAP design, the spectral “smile” effect is small and can probably be neglected for the operational processing. However, the input sensor definition file contains a flag to switch the smile correction (see Section 5.9) on or off. If smile correction is enabled, the water vapor retrieval is performed on a per-column basis as part of the simplified atmospheric correction.

5.6 Haze detection and removal

The haze removal employs a simplified method where the signal in haze affected regions consists of the haze-free scene radiance and a haze contribution H:

$$L^{sensor} = L_0 + H$$

where L_0 is the sum of the (haze-free) path radiance and the surface reflected radiance. So, the influence of the haze attenuation on the surface reflected radiation is not accounted for. Assuming a linear relationship of the digital number (DN) to radiance conversion, L^{sensor} can be expressed in DN as

$$DN^{sensor} = DN_i(x,y) + HTM_i(x,y)$$

where (x,y) are the pixel coordinates in the scene, *i* is the band number, and DN_i is the *DN* in band *i*, and HTM_i(x,y) is the haze thickness map in terms of *DN*. So the haze-free image will be obtained by subtracting HTM_i(x,y) from the recorded DN_i.

5.6.1 Haze thickness map

Haze thickness is usually varying over a scene and the HTM is calculated by searching dark pixels all over the image. A small to medium ground sampling distance (GSD = 2 - 30 m) allows the recording of shadows from objects. The corresponding pixels have a very low DN value (small ground reflectance + path radiance) and if haze is present the recorded DN value is also nonzero and higher than the corresponding haze-free contribution. These nonzero DN values are employed to estimate the HTM of the image. This HTM includes the thickness of haze and of the clear-scene aerosol thickness. The latter part is restored in the final step, and will be considered during the atmospheric correction.

The search for dark pixels is conducted with a moving local non-overlapping window w (i.e., area w x w pixels). The small window is used to increase the chance to locate dark pixels in shaded regions. The other pixels in the window are assumed to have the same haze thickness as the thickness estimated for the pixel with the minimal value. A window with a smaller size (w x w) allows a better estimation of the haze spatial distribution (haze spatial thickness over the scene) and a window size of w = 3*GSD is used as default based on the analysis of different Landsat-8, RapidEye, and Worldview-2 scenes. The selected dark pixels (with window size w) are stored in a matrix of size w smaller than the original scene. This matrix is median filtered (3 x 3 pixels) and then resized with cubic interpolation to the original image size.

The presence of large bright objects (bright agricultural fields, sand, rocks, roofs, snow, etc.) causes a wrong estimate of the HTM. Therefore, these areas have to be detected and labeled as not appropriate for an HTM evaluation. The results were obtained with the following apparent reflectance criteria:

$$\rho^*_{NIR} \geq 0.1 \; and \; (\rho^*_{blue} \geq T_{blue} \; or \; \rho^*_{red} \geq T_{red})$$

Equation 5.6-1

where $\rho^*(blue)$ is the apparent blue band reflectance and $\rho^*(red)$ is the red band reflectance and:

$$T_{blue} = \bar{\rho}^*(blue) + 2 * STD(\rho^*(blue)) \qquad T_{red} = \bar{\rho}^*(red) + 2 * STD(\rho^*(red))$$

Areas labeled as 'bright object' are interpolated in the HTM. In multispectral data, the search of dark pixels should be performed in a band with a minimal ground reflectance and a maximal haze signal component. A spectral band in the blue region (0.37 - 0.49 µm) is most suitable for this purpose. Spectral bands in the red and near infrared exhibit higher ground reflectance for land and are not suitable. However, the direct usage of a blue band often leads to over-dehazing of this band. To overcome this problem, a new synthetic band is created by a linear extrapolation of the two shortest wavelength bands (haze_bands) :

$$Band_{ext}(x,y) = B1(x,y) + (B1(x,y) - 0.95 * B2(x,y))$$

Equation 5.6-2

The surface reflectance in the extrapolated (lower wavelength) band is less, haze thickness is higher, and the HTM can be estimated more precisely. A 3x3 pixel median filter can be applied to remove noise in the extrapolated band.

The quality of the HTM calculation is improved with the use of the extrapolated blue band instead of the sensor-specific blue band. Therefore, the extrapolated blue band is employed for the dark pixel search and the non-overlapping window size w = 3*GSD:

$$HTM(x,y) = DARK\_PIXEL\_SEARCH(Band_{ext}(x,y), w)$$

It is necessary to label hazy and haze-free regions. A haze mask is created to find haze-free regions. This haze mask is based on the generation of an additional HTM(w2) with a moderately large window (GSD * 20, w2 = 21) and a thresholding of HTM(w2). This large window HTM(w2) is only used for the hazy / haze-free labeling.

$$haze\_free = where(HTM(x,y,w2) < mean(HTM(x,y,w2)))$$

The haze mask does not necessarily have to be precise, because it is only used to calculate a relative haze thickness coefficient, which is robust to mislabeled outliers.

5.6.2 Haze thickness per band

The optical haze thickness is wavelength-dependent, usually decreasing with wavelength, and channels in the short-wave infrared (SWIR2: 2 - 2.5 µm) are only marginally influenced by haze.
Therefore, a channel-dependent HTM_i is computed relative to the *HTM* using the image $Band_{ext}(x; y)$. First, a temporary HTM_i' is calculated for each band and the regression coefficient (k_i) of HTM_i' versus *HTM* is stored in an array (*K*). The regression coefficient is computed using the pixels labeled as haze in the haze mask. The k_i sequence should decrease with wavelength, i.e., $k_i+1 < k_i$. A linear scaling of *K* into the range [1, s] is performed with s= [0.1, 0.15, 0.73] if the last band is [SWIR2, SWIR1, NIR], respectively (SWIR1: 1.65 µm).

$$HTM_i'(x,y) = DARK\_PIXEL\_SEARCH(Band_i(x,y), w3)$$

The no-overlapping window size is w3 = 7*GSD.

$$k_i = SLOPE(HTM(x,y), HTM_i'(x,y))$$

The final band-specific HTM is calculated for multispectral bands (based in Sentinel-2 central wavelengths) and interpolated to the rest of hyperspectral bands

$$HTM_i(x,y) = HTM_i'(x,y) * k_i$$

5.6.3 Haze removal

A subtraction of $HTM(x,y)$ from the $DN_i'(x,y)$ recovers the dehazed band:

$$DN_i(x,y) = DN_i^{sensor}(x,y) - HTM_i(x,y)$$

This subtraction removes the calculated haze influence via HTM from the scene. However, the HTM includes the effect of haze plus clear-scene aerosol, and we want to keep the clear-scene aerosol component for the subsequent atmospheric correction. Therefore, we evaluate the mean DN value of the clear areas in the original scene ($M_i^{original}$) and the corresponding mean value in the dehazed scene ($M_i^{dehazed}$) and add the difference:

$$DN_i^{final}(x,y) = DN_i(x,y) + abs(M_i^{original} - M_i^{dehazed})$$

5.6.4 DEM data in dehazing method

The previous considerations are valid for a flat terrain. In case of mountainous regions there are additional problems. For example, high mountain regions (e.g. elevations > 2000 m) are seldom affected by haze. Nevertheless, the proposed dehazing algorithm may erroneously label some of these areas as hazy, because dark pixels can occur due to topographic effects and these will cause misclassifications. With a corresponding DEM we exclude these high mountain areas and their surroundings from the haze map as a precaution.

5.7 Cirrus detection and removal

Thin cirrus clouds are difficult to detect with typical broad-band multispectral satellite sensors in the atmospheric window regions, especially over land, because land scenes are spatially inhomogeneous and this type of cloud is partially transparent. Cirrus clouds typically occur in the upper troposphere and lower stratosphere, i.e. in altitudes of 8 – 16 km. On the other hand, water vapor dominates in the lower troposphere and usually 90% or more of the atmospheric water vapor column is located in the 0 – 5 km altitude layer. Therefore, if a narrow spectral band is selected in a spectral region of strong water vapor, e.g. around 1.38 µm or 1.88 µm, the ground reflected signal will be totally absorbed, but the scattered cirrus signal will be received at a satellite sensor.
So, a channel at 1.38 µm is able to detect cirrus clouds, and if a correlation of the cirrus signal at this wavelength and other wavelengths in the VNIR and SWIR region can be found, then the cirrus contribution can be removed from the radiance signal to obtain a cirrus-corrected scene. This idea is exploited in the state-of-the-art algorithm.

The original cirrus algorithm (IR-12, IR-13, IR-14: Gao et al. 1998, 2002, 2004) differs for land and water pixels. For land pixels, the correlation of the 1.38 µm channel and a red channel (around 0.66 µm) is used to calculate the scatter plot. To obtain a high sensitivity, only vegetation pixels are taken because they have a low reflectance in the red spectral region, so the cirrus contribution is easily traced. The scatter plot is computed in terms of the apparent (or TOA) reflectance of $\rho^*_{1.38}$ versus $\rho^*_{0.66}$ where the apparent reflectance is defined as Equation 5.1-1

The empirical approach is described by Equation 5.7-1 to Equation 5.7-4:

$$\rho^*(\lambda) = \rho_c(\lambda) + \frac{T_c(\lambda)\,\rho_v(\lambda)}{1 - s_c(\lambda)\,\rho_v(\lambda)}$$

Equation 5.7-1

Here, ρ_c is the reflectance of the cirrus cloud, T_c the two-way transmittance (direct plus diffuse) through the cloud, ρ_v the reflectance of the "virtual" surface (land or ocean surface including all effects of molecular and aerosol scattering below the cirrus), and s_c is the cloud base reflectance of upward radiation. Equation Equation 5.7-1 can be simplified, because of $s_c \rho << 1$, yielding

$$\rho^*(\lambda) = \rho_c(\lambda) + T_c(\lambda)\,\rho_v(\lambda)$$

Equation 5.7-2

Gao et al. (IR-13) assume that the cirrus reflectance $\rho_c(\lambda)$ is linearly related to the cirrus reflectance at 1.38 µm

$$\rho_c(\lambda) = \rho^*(1.38\mu m) / \gamma \quad for \quad 0.4 < \lambda < 1.0\mu m$$

Equation 5.7-3

where γ is an empirical parameter derived from the scene. It depends on the scene content, cirrus cloud height, and solar and view angles. Substituting Equation 5.7-3 intoEquation 5.7-2 yields

$$T_c(\lambda)\,\rho_V(\lambda) = \rho^*(\lambda) - \rho^*(1.38\mu m)/\gamma \quad for \quad 0.4 < \lambda < 1.0\mu m$$

Neglecting the cirrus transmittance T_c (i.e., setting T_c = 1) we obtain the "cirrus path radiance corrected" image:

$$\rho_V(\lambda) = \rho^*(\lambda) - \rho^*(1.38\mu m)/\gamma \quad for \quad 0.4 < \lambda < 1.0\mu m$$

Equation 5.7-4

As the cirrus is almost on top of the atmosphere we have ρ_c(1.38µm) = ρ*(1.38µm) and the apparent cirrus reflectance can be calculated with Equation 5.1-1. So, the cirrus removal is performed prior to the "normal" atmospheric correction which is then applied after converting ρ_v to its equivalent TOA radiance.

The γ parameter is calculated as follows:

1. Create $\gamma_{vect}[j]$ as [0.6, 0.65, 0.7, 0.75, 0.8, 0.85, 0.9, 0.95, 1.0], j=[0..8]
2. For each γ in $\gamma_{vect}[j]$ and for each channel k (wavelengths up to 0.58µm):
   - If j == 0: calculate the mean TOA reflectance for the non-cirrus pixels, save into $s_{no}[k]$.
   - For cirrus pixels in band k calculate the MEAN(ρ_k – $\rho_{1.38\mu m}$ /γ), save into $s_{ci}[j, k]$.
3. Calculate $cr[j] = \left| \sum_k s_{no}[k] - s_{ci}[j,k] \right|$
4. Find the index *ind* for the minimum value in the *cr*[*j*].
5. Calculate the γ value: $\gamma = \gamma_{vect}[ind]$.
6. For the range 1.5-2.5 µm the γ value is multiplied by 2.

Notes concerning cirrus removal:
- The calculation of the water vapor column is not accurate in cirrus regions, because cirrus modifies the ratio of TOA radiances in the window and measurement channels used for the retrieval.
- The cirrus removal option is switched off if the average water vapor column W of the scene is less than a threshold W = 0.4 cm. The reason is that bright surfaces can cause a similar signal level in the 1.38 µm channel as cirrus if the water vapor content is very low.

5.8 Retrieval of surface reflectance

The steps of cirrus and haze removal are executed before the calculation of the visibility/AOT and water vapor maps. After the aerosol optical thickness and water vapor maps have been calculated, the surface reflectance (also understood as reflectance irradiance) retrieval is performed.

5.8.1 Flat-terrain

In flat terrain, the radiative transfer equation for a homogeneous surface under clear sky conditions can be formulated as:

$$L = L_p + \frac{\tau \left(E_{dir}\cos\theta_s + E_{dif}\right)\rho / \pi}{1 - \rho s}$$

Equation 5.8-1

where L, L_p, τ, E_{dir}, E_{dif}, θ_s, ρ and s are at-sensor radiance, path radiance, ground-to-sensor transmittance, direct and diffuse solar flux on the ground, solar zenith angle, surface reflectance, and spherical albedo of the atmosphere, respectively. The total transmittance is the sum of the direct and diffuse transmittances $\tau = \tau_{dir} + \tau_{dif}$. For brevity, the dependence on wavelength, solar and viewing geometry, and atmospheric parameters has been omitted.

Equation 5.8-1 is solved for the surface reflectance ρ. The initial reflectance image is iterated to take the atmospheric blurring into account ("adjacency effect"). This is caused by photons from the surrounding area that are scattered into the sensor's line-of sight. A simple formulation was proposed by (IR-29: Richter 1990) where the strength of the adjacency effect is weighted with the ratio q of the diffuse to direct ground-to-sensor transmittance

$$\rho^{(2)}(x,y) = \rho^{(1)}(x,y) + q\left(\rho^{(1)}(x,y) - \overline{\rho}(x,y)\right)$$

Equation 5.8-2

and where $\rho^{(1)}(x,y)$ is the initial surface reflectance image, and $\overline{\rho}$ the $\rho^{(1)}$ reflectance averaged over a square box of size 2 R, where R=1 km is the assumed range of the adjacency effect. The averaging can also be performed with a distance weighting factor (IR-31: Richter 1998):

$$\rho^{(2)}(x,y) = \rho^{(1)}(x,y) + q\left(\rho^{(1)}(x,y) - \sum_{i=1}^{n} \overline{\rho}_i(x,y)\, w_i\right)$$

Equation 5.8-3

where n is the number of circular rings and w_i the exponential distance weighting factor. However, the difference between both approaches is usually small over land areas as the adjacency correction is a second order effect and the spatial pattern of fields tends to repeat itself.

At the scene border the adjacency correction needs information on areas outside the image. Therefore, an estimate is needed, see Figure 5-6. For the original (raw) geometry, a possibility is to mirror the adjacency strip. Another possibility is: if the neighborhood around a pixel includes a pixel outside the scene the nearest edge pixel is used to compute the box average. The second option is implemented in the land processor as it is much faster. To reduce the influence of the estimate of the missing area, the weighting factor q in Equation 5.8-2 is replaced with the factor 0.5 q. In case of a geocoded scene (bottom of Figure 5-6) no adjacency correction is performed in the border strip. Earlier versions worked with the assumption of using the average scene reflectance for the geocoded part, but if this estimate is not realistic, inconsistent reflectance values are clearly visible in the border region. For example, this problem will occur if the border region contains dark areas, but the scene-average is much brighter. So the least artifacts are obtained without adjacency correction for the border regions for geocoded data.

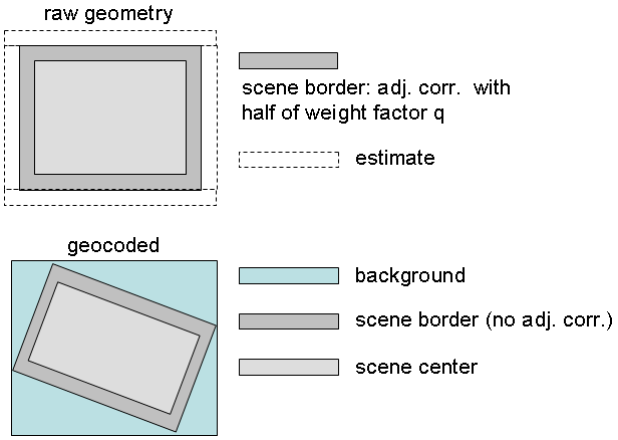


Figure 5-6: Treatment of the adjacency effect.

In the vicinity of clouds, a different approach to the adjacency correction has to be taken, because clouds are high in the atmosphere and should not be treated as bright ground surfaces. For the adjacency correction in Equation 5.8-2 all cloud pixels are assigned the scene average (non-cloud) reflectance $\bar{\rho}$ per channel.

5.8.2 Mountainous terrain

For each EnMAP scene a digital elevation model (DEM) is to be used. The DEM is to be matched to the scene, and the derived topographic maps of slope and aspect are calculated. Based on the statistics of the slope map a decision will be made to include a topographic correction as part of the atmospheric correction over land or to process the image assuming a flat terrain with the average elevation according to the DEM. The decision is:

- If more than 1% of the scene contains pixels with slopes > 6 degrees, then the topographic correction will be included, otherwise not (“quasi-flat” areas) because artifacts are likely to occur.

In mountainous terrain, the combined atmospheric / topographic correction obviously leads to more complex equations than for the flat terrain case, and some simplifications have to be done. Basically, there are four radiation components contributing to the TOA radiance: path radiance, pixel reflected radiance, adjacency radiation, and reflected radiation from the surrounding topography, see Figure 5-7.

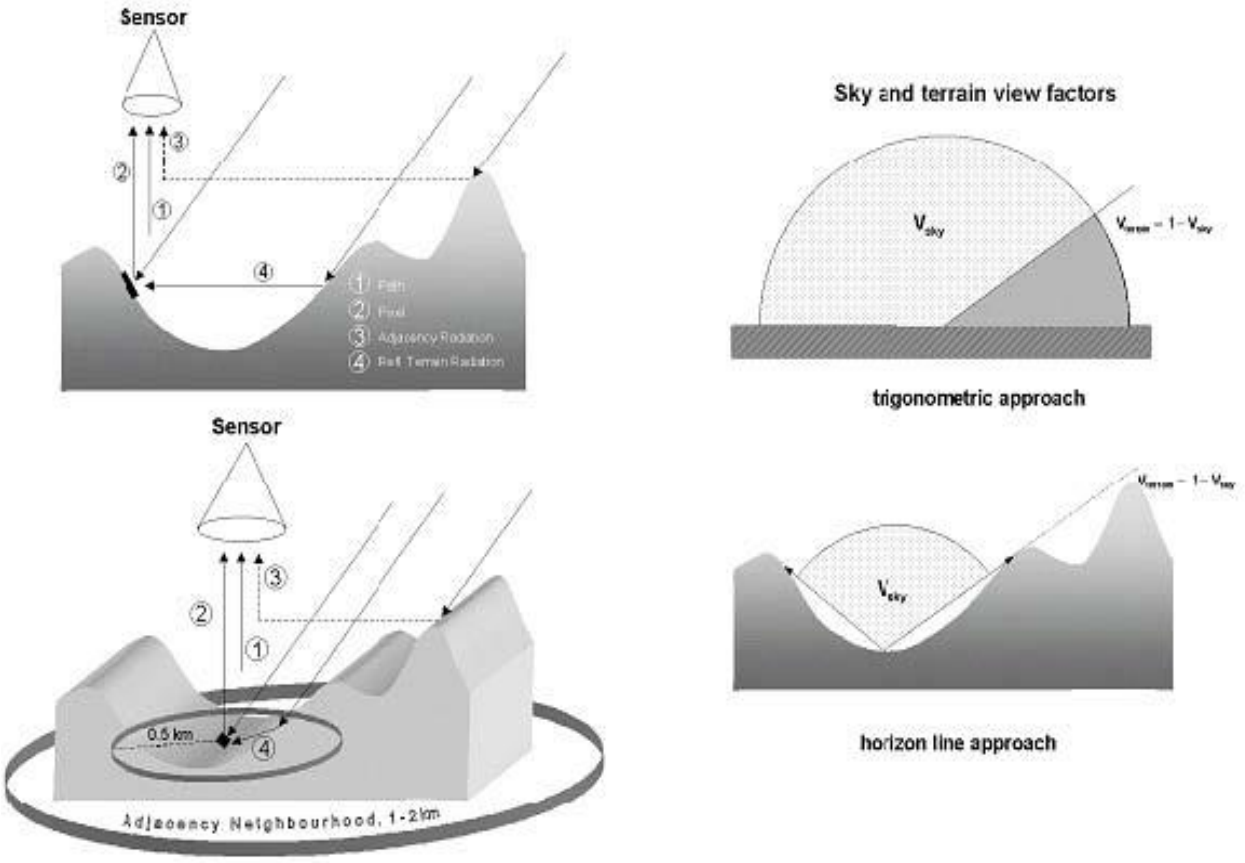


Figure 5-7: Radiation components in mountainous terrain.

The following brief summary follows the paper IR-31 (Richter 1998). If $\theta_s, \theta_n, \phi_s, \phi_n$ denote solar zenith angle, terrain slope, solar azimuth and topographic azimuth, respectively, the illumination angle β can be obtained from the DEM slope and aspect angles and the solar geometry:

$$\cos\beta(x,y) = \cos\theta_s \cos\theta_n(x,y) + \sin\theta_s \sin\theta_n(x,y) \cos(\phi_s - \phi_n(x,y))$$

Equation 5.8-4

The illumination image cos β(x,,y) is calculated within the land processor and stored as separate map. The diffuse solar flux on an inclined plane is calculated with Hay's model (Hay and McKay 1985):

$$E_d^*(x,y,z) = E_d(z)\left[b\,\tau_s(z)\cos\beta(x,y)/\cos\theta_s + \{1 - b\,\tau_s(z)\}\, V_{sky}(x,y) \right]$$

Equation 5.8-5

In rugged terrain the radiation reflected from mountains within the line-of-sight of a pixel has to be included as an additional radiance component. The fraction of pixels in the line-of-sight is called terrain view factor (V_t), and it is calculated as

$$V_t = 1 - V_{sky}$$

where V_{sky} is computed with a ray tracing program (IR-9: Dozier et al. 1981) or as a simple approximation with the local DEM slope value θ_n:

$$V_{sky} = cos^2(\theta_n / 2)$$

The use of the ray tracing program is recommended for very steep terrain (module *skyview* of the land processor creates the corresponding map), in moderately steep terrain the simple approximation is sufficient.

The surface reflectance is calculated iteratively. The first step neglects the adjacency effect and starts with a fixed terrain reflectance of $\overline{\rho}_t^{(0)}$ = 0.1

$$\rho_{(i)}^{(1)}(x,y) = \frac{\pi \left\lfloor d^2 \{ c_0 + c_1 DN(x,y) \} - L_p(z, \theta_v, \phi) \right\rfloor}{\tau_v(z, \theta_v) \left[b(x,y) E_S \tau_S(z) \cos\beta(x,y) + E^*_{dif}(x,y,z) + E_t(z) \overline{\rho}_t^{(i)} V_t(x,y) \right]}$$

Equation 5.8-6

The terms are defined as:

x, y	horizontal coordinates, corresponding to the georeferenced pixel positions
z	vertical coordinate, containing the elevation information from the DEM
DN(x,y)	digital number of georeferenced pixel ;
$L_p(z,\theta_v,\phi)$	path radiance, dependent on elevation and viewing geometry
$\tau_v(z,\theta_v)$	ground-to-sensor view angle transmittance, direct plus diffuse components
$\tau_S(z)$	Sun-to-ground beam (direct) transmittance
$\beta(x,y)$	angle between the solar ray and the surface normal (illumination, Equation 5.8-4);
b	binary factor: b=1 if pixel receives direct solar beam, otherwise b=0
E_S	extraterrestrial solar irradiance (earth-sun distance d=1 astronomical unit);
$E^*_d(x,y,z)$	diffuse solar flux on an inclined plane (see Equation 5.8-5))
$E_g(z)$	global flux (direct plus diffuse flux on a horizontal surf. at elevation z);
$E_t(z)$	terrain radiation reflected from adjacent slopes;
$\overline{\rho}_t^{(0)}$ = 0.1,	initial value of average terrain reflectance
$\overline{\rho}_t^{(i)}(x,y)$	locally varying average terrain reflectance, calculated iteratively;
$V_t(x,y)$	terrain view factor (range 0-1).

The next step iterates Equation 5.8-6 averaging the reflected radiation over a square box of 0.5 km x 0.5 km. If Equation 5.8-6 is used with $E_t = E_g$ then three iterations are usually sufficient to be independent of the start value of the terrain reflectance (IR-31: Richter 1998). However, for highly reflective surfaces (e.g., snow) and high terrain view factors, more than three iterations are necessary and a faster

convergence of $\overline{\rho}_t^{(i)}$ can be achieved with a geometric series for the terrain reflected radiation E_t as proposed in IR-39 (Sirguey 2009):

$$E_t^{(i)} = E_g \frac{\overline{\rho}_t^{(i-1)}}{1-\overline{\rho}_t^{(i-1)}\overline{V}_t}$$

The factor π in the numerator of Equation 5.8-6 assumes an isotropic (Lambertian) surface reflectance law. The next step is the adjacency correction (Equation 5.8-3).

5.8.2.1 Empirical BRDF correction

For a steep terrain the local solar zenith angle can reach values of 70° to 90° and the Lambertian reflectance assumption causes overcorrected (bright) values. To avoid this behavior the following empirical approach to topographic correction is taken (IR-34: Richter et al. 2009): if the local solar illumination angle β exceeds a threshold β_T (i.e., in areas of low illumination where the Lambertian reflectance ρ_L might become high), the updated surface reflectance is reduced according to:

$$\rho_{MM} = \rho_L \left(\frac{\cos\beta}{\cos\beta_T} \right)^b$$

Equation 5.8-7

The index *'MM'* stands for modified Minnaert. If the local illumination angle does not exceed the threshold β_T, then Equation 5.8-7 is not applied and the Lambertian reflectance is retained ($\rho_{MM} = \rho_L$). The following set of empirical rules is used:

- The threshold angle is set as $\beta_T = \theta_s + 20°$, i.e., *20°* above the solar zenith angle, if $\theta_s < 45°$.
- If $45° \leq \theta_s \leq 55°$ then $\beta_T = \theta_s + 15°$.
- If $\theta_s > 55°$ then $\beta_T = \theta_s + 10°$, i.e., a higher solar zenith angle requires a threshold closer to the solar zenith, because the correction must take place earlier.
- Exponent $b = 1/2$ for non-vegetation
- Exponent $b = 3/4$ for vegetation in the visible spectrum ($\lambda < 720\ nm$)
- Exponent $b = 1/3$ for vegetation if $\lambda \geq 720\ nm$.

In addition, if the correction factor $(\cos\beta / \cos\beta_T)^b$ is smaller than 0.25 (i.e., local solar zenith angle much higher than the threshold angle) it will be reset to 0.25 to prevent a too strong reduction. The reference (IR-34: Richter et al. 2009) demonstrates that this empirical topographic correction approach is usually superior to other commonly employed techniques, but there is currently no method which performs best in all situations.

Figure 5-8 presents an example of a SPOT-5 scene from Switzerland (IR-34: Richter et al. 2009) with a combined atmospheric and topographic correction including the empirical BRDF correction of Equation 5.8-7. The left part shows the colour coded scene (RGB = 1650, 840, 660 nm), the central part contains the illumination map, and the right part the surface reflectance product. Most of the topographic features are successfully eliminated.

Figure 5-8: SPOT-5 scene from Switzerland with topographic correction

5.9 Correction of spectral smile

Spectral “smile” is a consequence of optical aberrations that cause the spectrometer entrance slit, representing the across-track swath, to be projected as a curve on the rectilinear detector array (IR-26, Mouroulis et al. 2000; IR-15, Goetz et al. 2003). Spectral smile for EnMAP is specified to be smaller than 0.2 pixels for each channel. In case the smile values turn out to be small, a “normal” atmospheric correction without considering smile will be the standard processing. However, for scientific investigations a higher accuracy is desirable, therefore a smile correction is supported and configurable as a processor parameter. Further, any deviation from the nominal center wavelength related to a spatial pixel of a band will be given in the EnMAP user product metadata in form of fifth-order polynomial coefficients (see AR-3 and AR-12). This also applies if no smile correction was applied and / or the detected smile is below 0.2 pixels for each channel.

For each spectral channel *j* the channel center wavelength $\lambda_c(j)$ depends on the image column or pixel position x across-track. The absolute value of $\lambda_c(j)$ is specified in the wavelength file used to generate the spectral channel response functions, and it is also included in the sensor-specific solar irradiance file. If *n* is the number of image columns (raw geometry), the change $\Delta(x,j)$ of the center wavelength $\lambda_c(j)$ with the pixel position x can be accurately approximated as a 4th order polynomial (using the nm unit):

$$\Delta(x,y) = a_0(j) + a_1(j)\,x + a_2(j)\,x^2 + a_3(j)\,x^3 + a_4(j)\,x^4$$

Equation 5.9-1

$$\lambda_c(x,j) = \lambda_c(j) + \Delta(x,j)$$

Equation 5.9-2

The first left-hand image pixel is x=1, the last right-hand pixel is x=n. The coefficients $a_i(j)$ (i=0…,4) have to be stored in an ASCII file, starting with the wavelength in column 1, followed by the polynomial coefficients (beginning with a_0 and ending with a_4), one line per channel. Since the smile effect is only relevant in atmospheric absorption regions, the 5 coefficients can be set to zero in atmospheric window regions, but they must be provided for each channel. If all 5 coefficients are set to zero for a certain channel, this channel is processed in the “non-smile” mode which will expedite the processing.

Smile processing must be performed in raw image geometry (not orthorectified) in order to preserve the original image columns. During the water vapor and surface reflectance retrievals the algorithms work on

a per-column basis, i.e., to each image column its appropriate center wavelength is associated. For the nadir pixel (x=n/2) the channel center and bandwidth are defined in the sensor's wavelength file. The pre-calculated atmospheric LUTs apply to this geometry.
For off-nadir pixels the center wavelength is shifted according to the Equation 5.9-1 and Equation 5.9-2 and the monochromatic atmospheric database (see chapter 2.3) is accessed and resampled with the respective shifted channel filter function. The spectral bandwidth of a channel is assumed to be constant across the detector array, but a corresponding modification is readily included if it were necessary to consider an across-track variation of the bandwidth. In this case the same equations (Equation 5.9-1 and Equation 5.9-2 will be taken for the bandwidths using another 4^{th} order polynomial for the bandwidth variation across the detector array.

If the smile correction is enabled, the water vapor and surface reflectance retrievals are performed on a per-column basis with an increase of execution time of about a factor 6.

In order to maintain consistency with the calibration of the imaging sensors, the obtained reflectance values (taking smile into account) are shifted to the nominal center wavelength of the respective band by linear interpolation. For better discriminability the process of atmospheric correction on basis of adjusted lookup tables will be addressed as smile aware atmospheric correction, the process of shifting bands to the nominal wavelength will be addressed as smile correction in further documentation.

A detailed analysis of the performance of the EnMAP processor on data affected with smile can be found in AR-15.

5.10 Ozone column and season automatic retrieval

When the "automatic" option is specified by the user, a software module will query the Copernicus database and retrieve the historical value of the ozone column and corresponding season for the current scene location and time of data acquisition.

The Copernicus database will be updated regularly (see AR-17)

Both parameters are retrieved from Copernicus Climate Change Service (C3S) database ERA5 (ECMWF ReAnalysis v5) (DOI: 10.24381/cds.adbb2d47). ERA5 is the fifth generation ECMWF reanalysis for the global climate and weather for the past 8 decades (from 1940 until the present). The principle of the reanalysis is the data assimilation and it combines model data with observations from across the world into a globally complete and consistent dataset. This ERA5 dataset provides hourly estimates for a large number of atmospheric and land quantities, like for example the total ozone column and the temperature (at 2 meters), It is updated with a latency of 5 days, with a possible 2-3 months delay in case a serious flaw is detected. The dataset is re-gridded to a regular lat-lon grid of 0.25 x 0.25 deg.

5.10.1 Ozone column

The total ozone column is retrieved from the retrieved ERA5.

The data is extracted from the most recent data (by default not older than 30 days) in the retrieved C3S database available. If no up-to-date data is available then the default values are retrieved (ozone column of 330 DU for the summer season and 370 DU for the winter season).

The scene is projected into the global average map of ozone column (TCO3) and average the value over the scene boundary coordinates.

5.10.2 Season

The selection of the radiative LUTs for the two different MODTRAN atmospheric profiles (mid-latitude summer and mid-latitude winter) is done through a threshold in the local temperature (at 2 meters). This

parameter is the temperature of air at 2 meters above the surface of land, sea or inland waters. 2m temperature is calculated by interpolating the lowest model level and the Earth's surface, considering the atmospheric conditions
This temperature is preferred to the land surface temperature since the former one, extracted from IR measurements) depends on the albedo, the vegetation cover, the soil moisture and the duration of the solar heating.

A yearly temperature variation is assumed, so the data is extracted from the most recent data in the retrieved Copernicus database available corresponding to the hour and month of the scene data take. If there is available ERA5 data with a time difference with respect to the scene smaller than 30 days, it will be used even if it is a different month.

The scene is projected into the global average map of air temperature (T2M) and average the value over the scene boundary coordinates.

Temperature cut = 8 degrees Celsius, below which is 'winter' (or MODTRAN mid-latitude winter atmospheric profile). This cut is configurable in the EnMAP L2A processor, but its re-definition will require a different version of the software.

6. L2A-land processor

6.1 Tile processing

All EnMAP datatakes are processed in tiles of 1024 lines.
Due to the different information (e.g. DDV pixels and water/dark pixels) successive tiles may have different visibilities at tile borders. Also, the smoothed water vapor map might create visible differences at the border of two adjacent tiles.
These differences might cause different reflectance on ground pixels belonging to a similar ground material. To avoid or at least reduce this effect the following strategy is used:

- For a tile i, the tiles (i-1), i, and (i+1) are combined into one cube, processing the threes tile sequentially.
- For the determination of atmospheric parameters, smooth processes involving information over the three tiles, is performed. The visibility and water vapor determination is done in a full cube resulting of the merging of the three tiles.
- The atmospheric cubes are afterwards cropped into the original three tiles, and each tile is processed separately.
- The tile i (1000 x 1024) L2A products are delivered to the user.

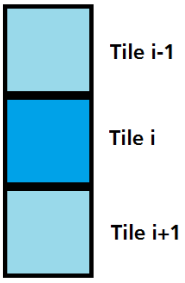


Figure 6-1: L2A tile land processing. To estimate the L2A products for tile i, tiles i-1, i, and i+1 are combined and processed. The L2A products corresponding to the tile i are cut and delivered.

The land processor model supports four aerosol types: rural / continental, maritime, urban, and desert. However, as a change in aerosol type between successive tiles might cause surface reflectance brightness steps at the tile borders, the baseline for the operational processor is to use a fixed aerosol type (rural / continental).

This strategy will produce smooth transition regions without brightness steps in most cases if tiles are later merged into a larger strip.
What it would not be avoidable is the step into parallel stripes, even if the atmospheric conditions would be the same.

7. Accuracy requirements for retrievals

EnMAP requirements of L2A processor include the uncertainty in the retrieval of atmospheric parameters (AOT and WV) and BOA reflectance of land and water (AR-14).

As already mentioned in Sections 5.1.3 and 5.8, the BOA reflectance retrieved in independent of the solid angle (irradiance) both for Land/Water and Combined products.

7.1 Land Requirements: AOT, WV and surface reflectance

The requirements in terms of uncertainty (RMSE) described here will only be applied to land-masked pixels (in Land_Mode and Combined) for areas in the scene fulfilling certain quality criteria included in each requirement (AR-14).

The uncertainty requirement in AOT is:

$$RMSE_{AOT} \leq 0.31 * AOT_{ref} + 0.05$$

within the AOT range between 0.1 and 0.8 and reflectance in the SWIR band (2.1 µm) < 0.12. These AOT values correspond of aerosol type “rural/continental” (see Section 5.4). An additional statistical pre-condition requires that at least 5% of dark pixels are found to evaluate the requirement per scene (~ 1/sqrt(1e3) < 0.05).

The requirement in WV is:

$$RMSE_{WV} \leq 0.04 * WV_{ref}\ (cm) + 0.18\ (cm)$$

only evaluated in clear land pixels (e.g. cloud pixels, water and shadow areas excluded) and for scenes with a “nominal” quality (see Section 14) and water vapor values above 1 cm.

Only the final Bottom-Of-Atmosphere (BOA) reflectance (in %, unitless) corresponding to land pixels are also described here and they refer to ground surface reflectance. The requirement will be met.

$$RMSE_{\rho} < \begin{cases} 0.01 & ,\rho < 0.1 \\ 0.02 & ,0.1 \leq \rho < 0.3 \\ 0.05 & ,0.3 \leq \rho < 0.6 \end{cases}$$

The uncertainty of the surface reflectance retrieval (in units of [0.01 * %]) is defined for a set of prerequisites: flat terrain ground surface, nadir view geometry, and a valid spectral band correlation factor based on DDV (dense dark vegetation) reference areas in a scene. In addition, the scenes should have “nominal” quality (see Section 14), that includes constrains like solar zenith angle.

As it is explained in Section 8, the surface reflectance requirements do not include the uncertainties in additional processing algorithms, selected by the user, like de-hazing and de-cirrus. Therefore, such requirements might not be fulfilled for dehazed or decirrus processed data.

8. Algorithm Verification and Validation

8.1 Verification of land requirements

A preliminary verification and validation of the atmospheric correction over land is documented in (IR-8, De los Reyes, et al, 2020).

The retrieved AOT (550 nm) atmospheric parameter from EnMAP data is validated against AERONET (IR-5) for different locations on Earth. The wavelength interpolation between the AERONET measurements and the EnMAP AOT at 550 nm, is performed via a logarithm relation.
Additionally, water vapor (WV) retrieval was performed with EnMAP scenes and compared also with AERONET results.
The validation was performed per scene within an area of < 9 km (diameter) around the AERONET station.
Only AERONET values around 20 min. of the scene acquisition are used for validation.
For all scenes where less than 5% of DDV reference pixels are contained inside the 9 km area, the validation is discarded for statistical purposes.
To avoid additional uncertainties (e.g. wind direction, speed) in the verification process, information from reference pixels outside the validation area is excluded.

Note: due to the previously described validation procedure and the size of each EnMAP tile (~30km by ~30km), it is not meaningful to use the mean AOT (worse for WV) per scene in EnMAP metadata to verify the previously mentioned requirements.

The preliminary analysis with Sentinel-2 and Landsat-8 is suitable for AOT and WV retrieval and similar results are to be expected for EnMAP, although the Landsat-8 and Sentinel-2 channels have lower spectral resolution that the ones of EnMAP.

The surface reflectance is also verified using the ground measurements data at RadCalNet sites (IR-19) as reference.

8.2 Verification of water requirements

The EnMAP L2A water products (see Section 4.2) are verified to fulfill the requirements described in (AR-14).
These requirements are only applied to water-masked pixels (in Water_Mode and Combined) for areas in the scene fulfilling certain quality criteria described in the requirements.

The verification of the AOT and water leaving reflectance product (R_rs), is performed comparing with AERONET-OC in-situ measurements [IR-47].

For the AOT over water sites, the AOT (also at 550 nm) is calculated using the same logarithm interpolation as for the AOT of the land processor.

The in-situ measurements used are the normalized water leaving radiance (so called $L_{wN,f/Q}$ inside AERONET-OC products) corrected from bi-directional effects, and it is related to the EnMAP water leaving reflectance by Equation:

$$R_{rs}[\%/100] = \pi \frac{L_{wN,f/Q}}{F_0}$$

where F_0 is the extraterrestrial solar irradiance provide by AERONET-OC [IR-46].

The subsurface water irradiance reflectance (abbreviated to R_E in Section 4.2) is verified converting the previous in-situ measurement to this unit, using EOMAP converting tools.

9. Input / Output Files and Formats

The input product data are with the standard defined by the AR-12.

The ozone concentration (in Dobson units) is a user input (Ozone value). If no value given by the user, a historic ozone data is to be used. The data are based on Copernicus database (ERA5) and shouldn't have a time difference with the scene larger than 30 days. These data are worldwide hourly ozone values (TCO3)
If the time difference is larger than 30 days, a default value will be used: 330 DU for summer and 377 DU for winter.

9.1 Input: sensor description

The sensor files are stored in a directory structure that registers the validity start time and possible different versions for the sensor descriptions (due to updates, bug corrections, etc…).
The directory structure is:
<sensor_parent_directory>/<sensor_name>/<validity_start_time>/<sensor_version>
where <sensor_parent_directory> is the directory where all the sensor files are stored. This directory must be linked as "sensor" at the same level as land processor library directory "ac2020" is located.
<validity_start_time> has the format YYYY.MM.DD (Y=year, M=month, D=day) and is the start day from which this sensor description should be applied.
<sensor_version> has the format <XX> (from 01 until 99). When running the land processor, the program always selects the last version available.

Inside the <sensor_parent_directory>/<validity_start_time>/<sensor_version> there must be the following files:

- sensor_<sensor_name>.dat: description file.The description is an ASCII file with up to 7 lines in the directory "../ac2020/sensor/enmap/". Example name: "sensor_enmap.dat":

  Line 1: across-track field-of view (degrees), pixels per line (raw geometry)
  Line 2: first and last reflective band (< 2.55 µm)
  Line 3: first and last mid-IR band (2.6 – 7.1 µm)
  Line 4: first and last thermal band (7.1 – 14 µm)
  Line 5: flag for sensor tilt capability (0=no, 1=yes)
  Line 6 0 (dummy for gain settings)
  Line 7 flag for smile (0=no, 1=yes), ID for spectral channel filter shape (ID=5: Gaussian)
  Line 7 can be omitted if the sensor smile effect is negligible.

- In case of spectral smile, the same directory also contains the file:

- "*smile_poly_ord4.dat"*, an ASCII file with 6 columns, namely the center wavelength of each channel and the 5 polynomial coefficients for the description of the fourth-order polynomial.
  E.g.

  0.420000 0.00000 0.00000 0.00000 0.00000 0.00000
  ……..
  0.681089 1.95781 -0.00782344 7.81562e-06 1.76601e-15 -7.97516e-19
  ……

  Zero coefficients correspond to a smile-free sensor.

- <sensor_name>.wvl, ASCII file with three columns: <band_number> <channel center wavelength> <channel_fwhm>
  where the wavelengths are in micro meter.
- band<bandnumber>.rsp: a list of x ASCII files (where x is the number of bands of the sensor). <bandnumber> goes from 1 to x, as three digits number, from 001 to a maximum number of bands are 999). Each ASCII file contains the channel response function for this channel. The spectral channel response is specified with one ASCII file per containing two columns: wavelength (in micrometer) versus normalized spectral response plus a header (first line) containing the number of following lines and the name of the file.

  E.g.:

  59 band001.rsp
  0.40918 0.000000
  0.40966 0.000032
  0.41013 0.000067
  ……
  0.42303 1.000000
  ……
  0.43688 0.0000000
- e0_solar_<sensor_name>.spc, containing three columns: center wavelength (in nm), fwhm (in nm) and solar irradiance for this channel (after re-sampling) in (mW/cm2 micron)

The lines 3 and 4 in the first sensor file (sensor_<sensor_name>.dat) are not needed to configure EnMAP, since EnMAP does not contain thermal bands, but are kept for file format consistency.

All the sensor files (expect the smile_poly_ord4.dat) will be created through a program provided within the L2A land processor library.

The input of the program will be the calibration tables that contain the center and fwhm for each channel (VNIR and SWIR) (see AR-13). The program will also call a function (provided by L1B) to extract this information from the calibration tables.

9.2 Input: atmospheric LUTs

The atmospheric LUTs files are stored in a directory structure that registers the sensor validity start time and version plus the version of the monochromatic LUTs they were re-sampled from.
The directory structure is:
<atm_lib_parent_directory>/<sensor_name>/<LUTs_version>/<validity_start_time>/<sensor_version>
and will contain a set of binary files that contains the atmospheric LUTs for the specific sensor.
This directory must be linked as "atm_lib" at the same level as land processor directory "ac2020" is located.

These are binary files (.atm) for the selected aerosol type and several water vapor contents, containing path radiance, direct and diffuse transmittance, direct and diffuse solar flux, spherical albedo, and sun-to-

ground transmittance, resampled for each channel of EnMAP using the monochromatic atmospheric database.

In addition to these binary files two ASCII files are created:

- Irrad_source.txt: name of the solar irradiance tables used for the LUTs (default: fmed2_2011)
- version.txt: version of the software used to produce the LUTs.

The atmospheric LUTs must be re-processed if:

- new sensor description and response functions are delivered for a certain (already created) validity time: new <sensor_version>.
- New valid sensor description is available: new <validity_start_time>/01 directory and corresponding files created.
- New monochromatic LUTs are available: then a new directory will be created <LUTs_version> and under it, all the previously available <validity_start_time> should be re-produce. Only for the latest version number, that in this new directory structure will become <sensor_version> = “01/”.

The monochromatic database is the input of the L2A land processor program that creates the sensor files It should have the following directory structure: <atm_database_parent_directory>/<LUTs_version>/<noff-nadir_max_binning>.

The latest version of <LUTs_version> (currently 1.0) will be provided as part of the L2A processor release when available for the corresponding <noff-nadir_max_bining> of EnMAP.

9.3 Input: ozone LUTs

For the ozone correction a binary file containing ozone LUTs should be provided as part of the monochromatic database. This directory must be linked as “atm_database” at the same level as land processor directory “ac2020” is located.
The file ozone_luts.bin (for summer season) and ozone_luts_winter.bin (for winter season) should exist in: <atm_database>/<LUTs_version>/ directory.

E.g.:

```
| atm_database/01.00/nadir/ozone_luts.bin
                        ozone_luts_winter.bin
                  offnadir10/……
                        ……..
                  offnadir30/…..
                        …….
```

9.4 Input: land surface temperature DB

For the season selection a Copernicus database should be provided in the previously described “atm_database” directory with the following structure:
<atm_database_parent_directory>/ECMWF/C3S/ERA5/T2M
Inside each first day of the month a hierarchical data format file (.nc) table must exists with the name of the database (ERA5), variable (T2M) and YYYYMM_YYY_DOY format in the name (e.g. of the file name ERA5_T2M_20230101_2023001.nc for January 2023)

9.5 Output files

The output product data are with the standard defined by the AR-12.

10. Summary

An automatic atmospheric correction algorithm for EnMAP imagery has been described.
The algorithms are implemented as part of:

- the smile correction and interpolation processor in sensor geometry
- atmospheric correction processor to retrieve BOA products. This processor consists of two different processors that process land and water pixels separately to retrieve land surface reflectance and water irradiance reflectance.
- Additional masks of clouds, haze, cirrus, … pixels are also delivered.

11. Appendix A: Altitude profile of summer and winter atmospheres

The following two tables contain the altitude profiles of the MODTRAN[AR-3] atmospheres mid-latitude summer and mid-latitude winter for the 0 – 5 km height range.

altitude (km)	pressure (mbar)	temperature (°C)	rel. humidity (%)	abs. humidity (g/m^3)
0	1013	21.0	76	13.9
1	902	16.5	66	9.3
2	802	12.0	55	5.9
3	710	6.0	45	3.9
4	628	0.0	39	1.9
5	554	-6.0	31	1.0

Table 4 Altitude profile of the midlatitude summer atmosphere.
Total (ground-to-space) water vapor content = 2.92 (cm or g cm^{-2}).

altitude (km)	pressure (mbar)	temperature (°C)	rel. humidity (%)	abs. humidity (g/m^3)
0	1017	-1.0	77	3.5
1	897	-4.5	70	2.5
2	789	-8.0	65	1.8
3	694	-11.5	57	1.2
4	608	-17.5	50	0.7
5	531	-23.5	47	0.4

Table 5 Altitude profile of the midlatitude winter atmosphere.
Total (ground-to-space) water vapor content = 0.85 (cm or g cm^{-2}),

12. Appendix B: Selection of atmospheric LUTs

The selection of the appropriate LUTs is done as a function of the atmospheric temperature / humidity profile (summer or winter) and the ozone content. The summer / winter decision will be based on the global monthly averaged Copernicus C3S air temperature (T2M) product available for a 0.25° For each EnMAP scene (30 km x 30 km) the corresponding month of the Copernicus product has to be accessed to average the appropriate database pixels to calculate a single temperature value. If the temperature is lower than a certain threshold (currently 8°C) then the winter LUTs are taken, otherwise the summer LUTs.

The ozone content is the input (external source) for the selected scene.

The high spectral resolution ("monochromatic" *.bp7) LUTs have to be resampled with the EnMAP channel filter functions to obtain the EnMAP specific LUTs (*.atm). The monochromatic database is calculated off-line, and the resampling has to be conducted only once if the spectral filter functions are stable. The resampled LUTs will be used during atmospheric correction. However, if the spectral smile has to be accounted for then the monochromatic LUTs are also needed and they have to be resampled with the channel filter curves on a per-column basis of the image.

13. Appendix C: LUT names

The LUT names include the aerosol type (e.g. rural), water vapor grid (e.g. wv=04 meaning wv=0.4 cm for the sea level), as well as a summer /winter identifier.

The same names are applied for the monochromatic and resampled LUTs, only the file extension is different (.bp7 for monochromatic LUTs, .atm for resampled LUTs).

Summer LUTs

File name	ozone content (DU)
h99000_wvxx_rura	330

where xx is a two-letter water vapor grid identifier: xx = 04, 10, 20, 29, 40, 50
(xx=20 means water vapor column 2.0 cm from sea level to space).

Winter LUTs

File name	ozone content (DU)
w99000_wvxx_rura	377

where xx is a two-letter water vapor grid identifier: xx = 02, 04, 08, 11
(xx=02 means water vapor column 0.2 cm from sea level to space).

14. Appendix D: Data Quality Control for L2A data

As described in AR-10 and AR-11, data quality control is also conducted for atmospherically-corrected data (L2A). The quality of the L2A will depend on the processing mode: Land, Water or combined. For the Land and Water, the corresponding data quality criteria for each processor will be applied and the criteria is described in the next sections. In addition to the automated QC described below, also an interactive (off-line) analysis is carried out, see Quality Control Operations Plan (EN-PCV-PLN-8020) and the independent product Validation Plan (PVP) by GFZ.

14.1 Quality land processor

The quality of atmospheric correction (metadata quality flag **QualityAtmosphere**) is to be reduced when certain conditions occur. Parameters for such conditions include:

- Average scene water vapor content [WV > 4cm reduced quality, WV > 5cm low quality.]
- Average scene visibility / aerosol optical thickness [VIS < 10km reduced quality, VIS <7km low quality.]
- Low sun angle [SZA >55: reduced quality, SZA >65: low quality.]
- Percentage of haze over land (> 20%: low quality)
- Percentage of cirrus (thin, medium and thick) (> 20%: reduce quality)
- Percentage of clouds (> 20%: low quality)
- Percentage of cloud shadow (>10%: low quality)
- Percentage of topographic shadow (> 10%: low quality)

where different numerical values are given depending on the quality: “nominal” =0, “reduced” =1 and “low” = 2.
These conditions will be evaluated for Land_Mode and Combined mode. In addition, Combined mode will also include the data quality criteria described in AR-6.

The performance of the L2A land processor (metadata quality flag **SceneAtmParam**) algorithm also depends on certain scene properties. Parameters affecting the data quality include:

(1) DDV & water pixels: percentage of pixels found in the scene in relation to the dark reflectance threshold
(2) Remaining percentage of negative reflectance pixels (per channel)
(3) Other warning and error messages of the L2A processor
(4) Parameters (1) & (2)
(5) Parameters (1) & (3)
(6) Parameters (2) & (3)
(7) Parameters (1) & (2) & (3)

It is important to point out that the mentioned atmospheric quality parameters do not affect the overall quality rating (overallQuality) of the L0 products (see SCR-00331 resulting from the Commissioning Phase analysis; see also IR-44, Ch. 4.2.4 and 8.5).

14.2 Quality water processor

A detailed data quality criterion for the water processor is described in AR-6.

The quality of atmospheric correction (metadata quality flag **QualityAtmosphere**) is likely to be reduced when certain conditions occur.

- Percentage of Q factor (> 25%: low quality, > 50%: reduced quality)

  The Q factor is the overall quality indicator as given by the water processor and considering occurrence of sun glint, high AOT and sun angle on a per pixel basis. For a detailed description of the Q factor quality metric and the underlying sub quality classes refer to L2A water ATBD (Section 2.8 in AR-6)

- Low sun angle [SZA >55: reduced quality, SZA >65: low quality.]

  Elevation thresholds for sun angles that lead to reduced and low water AC quality.

- Percentage of haze over water (> 20%: low quality, > 10%: reduced quality)

  Percentage thresholds for haze depended quality assessment as taken from the updated land-water-mask.

- Percentage of clouds (> 20%: low quality, > 10%: reduced quality)

Percentage thresholds for cloud depended quality assessment as taken from the updated land-water-mask.

- Percentage of sun glint (>10%: low quality, >5%: reduced quality)

  Percentage thresholds for sun glint depended quality assessment as taken from the quality mask provided by the L2A water processor.

These conditions can be tracked automatically by analyzing the corresponding data products (see Chapter 5.2). The results will be included in the metadata as well as in the Quality Quick-look.

Also, the performance of the proposed algorithm depends on certain scene properties. Parameters affecting the data quality include:

- Remaining percentage of negative reflectance pixels (per channel)
- Other warning and error messages of the L2A processor

Again, these parameters are derived automatically by parsing the log file and analyzing the corresponding data products, and will be included in the metadata.

It is important to point out that the mentioned atmospheric quality parameters do not affect the overall quality rating (overallQuality) of the L0 products (see SCR-00331 resulting from the Commissioning Phase analysis; see also IR-44, Ch. 4.2.4 and 8.5).

14.3 Quality processor in Combined mode

The data quality conditions of both processors are evaluated separately in a consecutive manner while any occurrence of a reduced- and low-quality indication overwrites previous nominal results. Therefore, it is assured that in a nominal quality scene all quality criteria explained 14.1 and 14.2 are fulfilled for both processors.

15. Appendix E: List of TBC

TBC To Be Confirmed

Number	Topic	Status	Chapter	Done
1	LUTs for atmospheric height profile of temperature / humidity: only summer, or summer & winter	Resolved. Ozone value and season are the input to the L2A chain.		08.11.2017
		Resolved. A method for ozone correction is introduced.		1.7.2016
3	Data Quality Control – Thresholds for flagging a reduced data quality (in regard of atm. conditions and illumination geometry)	Resolved.	13	1.7.2016
4	Data Quality Control – Consistency with processing using alternative atm.	Resolved.	13	1.7.2016

	correction software (proposed: FLAASH)			
5	MODIS database will be updated yearly.	Resolved. The frequency of the update is specified in the Operations Plan (AR-17)	6	12.11.2020

16. Appendix F: List of TBD

TBD To Be Defined

Number	Topic	Status	Chapter	Done
1	Update sections to current processor status	Resolved	all	17.04.2020