Optimal Cloud Analysis (OCA) Release 1 Product Users Guide

DOI: 10.15770/EUM_SEC_CLM_0049 (0DEG)

Doc.No.	:	EUM/OPS/REP/21/1209275
Issue	:	v1K e-signed
Date	:	19 December 2022
WBS/DBS	:	

EUMETSAT
Eumetsat-Allee 1, D-64295 Darmstadt, Germany
Tel: +49 6151 807-7
Fax: +49 6151 807 555
http://www.eumetsat.int

Document Signature Table

	Name	**Function**	**Signature**	**Date**
Book captain:	Alessio Lattanzio	Climate Product Expert		
Reviewed by:	Alessio Bozzo	RSS Optical Cloud Imaging		
Reviewed by:	Marie Doutriaux-Boucher	Climate Product Expert		
Reviewed by:	Rob Roebeling	Climate Product Expert		
Reviewed by:	Andrea Di Giovanni	OPS QA		
Approved by:	Jörg Schulz	Climate Service and Product Manager		

Co-authors/Contributors

Name	***Organisation***
John Jackson	Innoflair UG

Document Change Record

Issue / Revision	***Date***	***DCN. No***	***Changed Pages / Paragraphs***
1H	16/02/2022		Internally approved
1J	18/05/2022		comments from external reviewers accepted
1K	19/12/2022		Correction of typos: product filename and MET10 operational period

Table of Contents

Table of Figures

Table of Tables

1 INTRODUCTION

1.1 Purpose and Scope

This document describes the Meteosat Second Generation (MSG) Optimal Cloud Analysis (OCA) Climate Data Record (CDR) Release 1 and provides information needed for downloading, decoding, and using the data. In addition, it provides a summary of the CDR validation, including information on the limitations of the CDR, if there are any.

The data record contains Meteosat level-2 cloud properties: cloud phase, cloud optical thickness, cloud top pressure, cloud effective particle radius (for one and two layers) and covers the period 21 February 2004 – 31 August 2019.

1.2 Reference Documents

Benas, N., Finkensieper, S., Stengel, M., van Zadelhoff, G.-J., Hanschmann, T., Hollmann, R., and Meirink, J. F.: The MSG-SEVIRI-based cloud property data record CLAAS-2, Earth Syst. Sci. Data, 9, 415–434, https://doi.org/10.5194/essd-9-415-2017, 2017.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., Berg, L. van de, Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., Rosnay, P. de, Tavolato, C., Thépaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, 137, 553–597, https://doi.org/10.1002/qj.828, 2011.

EUMETSAT: Clear Sky Reflectance Map: Product Guide, 2015.

EUMETSAT: MTG-FCI: ATBD for Optimal Cloud Analysis Product, 2016.

EUMETSAT: NetCDF Creation Guidelines; Best Practises, Conventions and Applicable Standards - EUM/OPS/STD/11/3120, 2017.

EUMETSAT: Optimal Cloud Analysis (OCA) Release 1 Validation Report, 2021.

Hocking, J., Rayer, P., Rundle, D., and Saunders, R.: RTTOV v11 Users Guide, 2015.

Meirink, J. F., Roebeling, R. A., and Stammes, P.: Inter-calibration of polar imager solar channels using SEVIRI, Atmos. Meas. Tech., 6, 2495–2508, https://doi.org/10.5194/amt-6-2495-2013, 2013.

Stöckli, R., Duguay–Tetzlaff, A., Jedrzej, B., Hollmann, R., Fuchs, P., and Werscheck, M.: CM SAF ClOud Fractional Cover dataset from METeosat First and Second Generation - Edition 1 (COMET Ed. 1), https://doi.org/10.5676/EUM_SAF_CM/CFC_METEOSAT/V001, 2017.

Watts, P. D., Bennartz, R., and Fell, F.: Retrieval of two-layer cloud properties from multispectral observations using optimal estimation, J. Geophys. Res., 116, D16203, https://doi.org/10.1029/2011JD015883, 2011.

1.3 Acronyms and abbreviations

The below table contains the acronyms and abbreviations used in this document:

Acronyms	Meaning
AMV	Atmospheric Motion Vector
ATBD	Algorithm Theoretical Basis Document
CDR	Climate Data Record
CFC	Cloud Fractional Cover
CLM	CLoud Mask
COT	Cloud Optical Thickness
CPH	Cloud PHase
CRMN	Clear sky Reflectance Map
CRE	Cloud Effective Radius
CTP	Cloud Top Pressure
CTH	Cloud top height
ECMWF	European Centre for Medium-range Weather Forecasts
ERA	European ReAnalysis
FCDR	Fundamental Climate Data Record
IR	InfraRed
LUT	LookUp Table
ML	Machine-learning
MPEF	Meteorological Product Extraction Facility
MSG	Meteosat Second Generation
NRT	Near Real Time
NWP	Numerical Weather Prediction
OCA	Optimal Cloud Analysis
RTM	Radiative Transfer Model
RTTOVS	RadiaTive Transfer TOVS
SAF	Satellite Application Facility
SEVIRI	Spinning Enhanced Visible and Infrared Imager
SSR	Sensor Spectral Response
SSP	Sub-Satellite Point
TOVS	Tiros Operational Vertical Sounders
TSK	Skin Temperature
UTC	Universal Time Coordinate
VIS	Visible
WV	Water Vapour

2 BACKGROUND

The Optimal Cloud Analysis (OCA) algorithm was first developed in a research study awarded to the Rutherford Appleton Laboratory (RAL) in 1997 and was coded as a prototype system in 2001. The algorithm was further developed by EUMETSAT, with the aim of providing a Day-2 product from the Meteosat Second Generation (MSG) Spinning Enhanced Visible and Infrared Imager (SEVIRI) instrument. The latest version of the operational algorithm allows for identification of multi-layer cloud situations and retrieves cloud properties for two-layer scenes (Watts et al., 2011). OCA also provides an estimation of the uncertainty derived by the optimal estimation method.

Since June 2013, the OCA products have been operationally generated at full repeat cycle (15 min) frequency as a demonstrational product. Cloud properties retrieved by OCA are cloud top pressure, cloud optical thickness, and cloud effective radius. The OCA algorithm has been slightly adapted for climate data record processing. The adaptation mainly consists in the usage of different inputs, because the one used for Near Real Time (NRT) were not available for the reprocessing (cloud mask, clear sky reflectance map) and also not homogenous (reanalysis) over the complete time period. The difference between NRT and CDR products are provided in the validation report (EUMETSAT, 2021).

The OCA Release 1 Climate Data Record (CDR) covers the MSG observation period from 2004 up to 2019, providing a homogenous cloud properties time series it extends the NRT product more than 9 years back in time. The OCA Release 1 is planned to be used as input for the generation of a new CDR of Atmospheric Motion Vectors (AMV) with the potential to include an uncertainty estimates of the wind vector height.

3 DATA RECORD OVERVIEW

General	Data record name	Optimal Cloud Analysis Release 1
	Data record digital identifier	DOI: 10.15770/EUM_SEC_CLM_0049 (0-degree)
	Data record short description	Reprocessed level-2 geostationary cloud properties from Meteosat second generation
	Record type	Thematic Climate Data Record
	Content	Meteosat level-2 cloud properties
Coverage	Spatial	• Meteosat disk centered at 0° • each pixel (IFOV) ground resolution of 3.0 km for SEVIRI at sub-satellite point.
	Time period	21 February 2004 – 31 August 2019
	Temporal frequency	• MSG/SEVIRI 15 minutes
Instrument	Instruments names	Spinning Enhanced Visible and Infrared Imager (SEVIRI)
	Instruments descriptions	• **SEVIRI** is a compact telescope and scan assembly, allowing the implementation of a large passive cooler which improves IR detector performances by lowering their operating temperature. The imaging SEVIRI radiometer is equipped with a patented three-mirror (3M) telescope of compact design (focal length of 5367 mm) which allows the insertion of a small black body for full-pupil calibration. It has 12 spectral channels, 11 provide measurements with a resolution of 3 km at the sub-satellite point. The twelfth, so-called HRV (High Resolution Visible) channel of SEVIRI (Spinning Enhanced Visible and Infrared Imager), provides measurements with a resolution of 1 km.
Instrument Data	Input data	• Reflectances from the SEVIRI Level 1.5 image data for the VIS0.6 µm, the VIS0.8 µm and the VIS1.6 µm channels; • Radiances from the SEVIRI Level 1.5 image data for these channels: IR3.9 µm, IR6.2 µm, IR7.3 µm, IR8.7 µm, IR9.6 µm, IR10.8 µm, IR12.0 µm and IR13.4 µm. Note: The IR3.9 µm and the IR9.6 µm channels are not actively used in the product generation. • ERA-Interim forecast • Cloud mask at pixel resolution • Clear Sky Reflectance Map (daily) • Emissivity Map (from MODIS)
	Output data	Geostationary cloud properties retrieved using the EUMETSAT algorithm: cloud phase, cloud optical thickness, cloud top pressure, cloud effective particle radius (one layer and two layers). Ancillary output: cloud mask
	Format	NetCDF4 format
Access	EUMETSAT Data Centre	The data set is available from EUMETSAT Data Centre (https://eoportal.eumetsat.int/)
	Delivery	• ftp push (after registration to the data centre)

4 DATA RECORD DEFINITION

4.1 Product retrieval

The OCA product comprises information on cloud phase (CPH), cloud top pressure (CTP), cloud optical thickness (COT) and particle effective radius (CRE) for every pixel in a SEVIRI image. An input cloud mask (CLM, section 5.2.4) is used to determine cloud presence in the field of view. Internal diagnostics of the retrieval process are used to identify the presence of significant layering in the cloud structures. In case of layered cloud structures, it is assumed two layers are present for which CTP and COT are derived, while CRE is only made available for the upper layer. A complete description of the OCA algorithm is given in the Algorithm Theoretical Baseline Document (ATBD) (EUMETSAT, 2016)

Each retrieved cloud property comes with an error estimate that is calculated by propagation of the modelled errors in the input parameters. In addition, there is, per pixel, an overall quality indicator based on the Chi-squared fit of the retrieval. While the error estimates are derived for each cloud property separately, the overall quality indicator is derived from all retrieved cloud properties for the pixel concerned. Compared to the NRT product, the OCA Release 1 climate data record provides additional information, such as several overall disk averages to allow the users for a quick check on the quality. It also includes the input cloud mask.

An example overview of the product as full disk images is given in Figure 1 where the highest (lowest pressure) cloud CTP, CRE and total cloud optical thickness are shown.

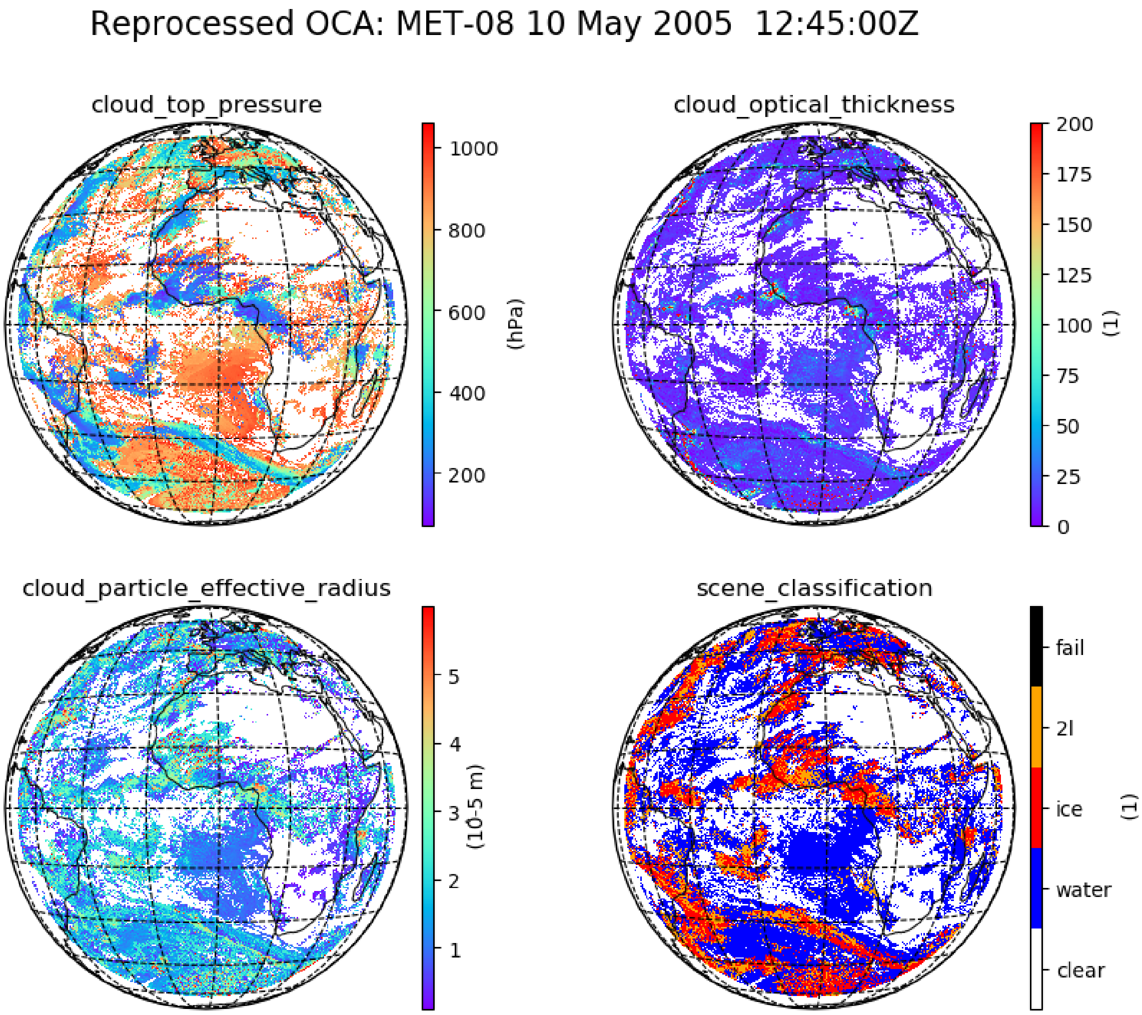

Figure 1: Example of OCA Release 1 products for 10th of May 2005 at 12:45 UTC. Upper left: cloud top pressure (hPa) of the highest layer, upper right: total cloud optical thickness, lower left: cloud top particle effective radius (m) and lower right cloud phase.

The OCA product, with its combination of phases and layers, is an unusual product compared to cloud products from other providers, such as the single layer cloud products from the CM SAF (Benas et al., 2017). To illustrate this, Figure 2 presents a pseudo-vertical cross-section of the OCA product for a sample of about 700 pixels in the West-East direction across Europe intersecting various cloud types. Where CTPs are shown as cloud top heights (CTHs) (height in km) and the COTs as geometric vertical extent. The figure shows the presence of single layer ice and liquid clouds, as well as two layers clouds. The figure excludes information on CRE, error estimates, and the overall quality indicator.

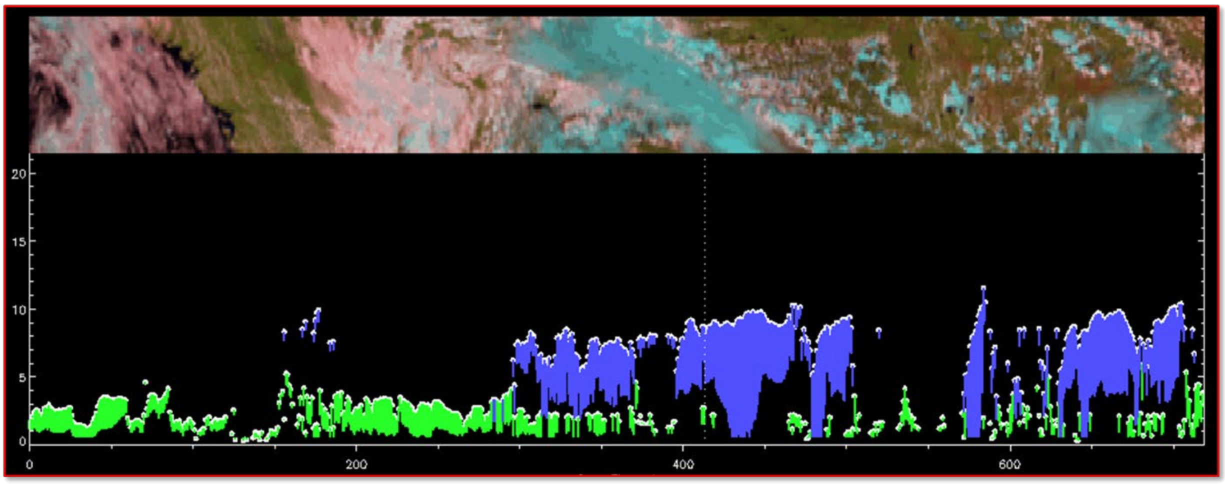

Figure 2: Illustrative vertical 'section' view of the OCA product. Upper panel is a section of the SEVIRI image in 'Natural colour' RGB and the pixels constituting the lower edge of this image are shown in the lower panel as the OCA retrieval: ice phase clouds are shown blue, liquid phase clouds in green. For purely visualisation purposes simple conversion factors were used to derive CTH (height in km) from the CTP retrievals and cloud geometrical thicknesses from the COT retrievals. Note that CRE, error estimates, and quality flags are not illustrated in this figure.

Product temporal coverage

Cloud properties are derived using images acquired from the four SEVIRI instruments on-board the series of MSG satellites. The OCA Release 1 provides quarter-hourly products derived over the full disk, with maximum extents of 65°N − 65°S and 65°W − 65°E around the nominal Sub-Satellite Point (SSP) (see Table 1) for the prime satellite (see Appendix C for details). The retrieval is performed at full SEVIRI temporal resolution of 15 minutes.

Table 1: **Satellite, instrument, mission, nominal orbit position and main services for the period 2004-2019. The period includes and Meteosat 8-11 (MSG).**

Satellite	Instrument	Mission (SSP)	Start Date	End Date
Meteosat-8	SEVIRI	0 DEG (0°)	2004-01-19	2007-05-31
Meteosat-9	SEVIRI	0 DEG (0°)	2007-05-01	2013-04-30
Meteosat-10	SEVIRI	0 DEG (0°)	2013-01-21	2018-02-20
Meteosat-11	SEVIRI	0 DEG (0°)	2018-02-20	2019-08-31

The data record stops 31 of August 2019 because this is the end of the ERA-Interim period.

4.3 Product spatial coverage

The OCA Release 1 only includes products from the MSG 0-degree missions. The spatial coverage is given in Figure 3. The retrieval is done at full spatial resolution, with a nominal pixel size at sub satellite point of 3km. Due to normal operation changes (satellite maintenance) of the prime satellite, few products are generated using the current backup platform. For this reason some products are generated at SSP of 3.4 degrees west and others at 9.5 degrees East (see Figure 3). The information about the current SSP is not included in the product in this release, but the correct area is considered in latitude and longitude. The products cover a region of about 65 degrees radium around the SSP.

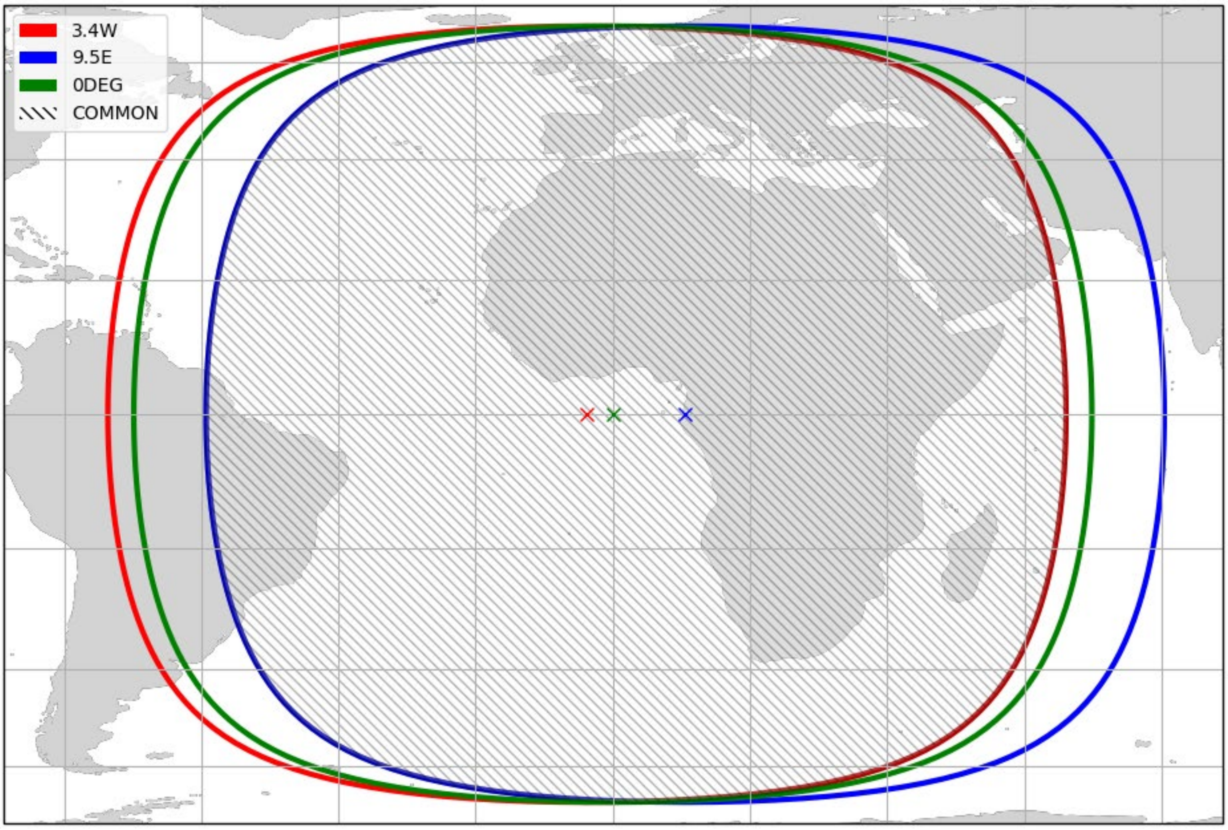

Figure 3: ***Spatial coverage of Meteosat SSP at 0° exploited for OCA up to 65 degrees. The area covered also by the backup platforms at 3.4W and at 9.5E are shown. The common retrieved area along the whole data record time coverage is also shown.***

4.4 Product Contents

The OCA Release 1 is made available in the EUMETSAT NetCDF format. For each 15-minute observation slot a separate NetCDF file is made available. Thus, OCA Release 1 consists of a maximum of 96 products per day. An estimation of expected OCA product aggregation size is given in Table **2**.

Table 2: NetCDF file representative size

Granularity	Size
Single	66.0 MB
Day	4.5 GB
Month	235.0 GB
Year	2.7 TB

4.4.1 NetCDF product description

The NetCDF files of OCA Release 1 are files in the NetCDF4 format. There are 15 global attributes as listed in Table 4. There are 41 variables, including 26 OCA full disk statistics as listed in Table 5 and 15 OCA variables as listed in Table 6. The names of the OCA variables follow, when possible, the CF-convention[1]. The OCA variables used in the validation report and of main interest for the users are described in Table 3.

Table 3: Mapping NetCDF to report variable name.

Variable in the report	NetCDF mapping variable
CTP	cloud_top_pressure in Pa (Cloud top pressure of single layer or, in two layer cases, the highest layer)
COT	10(cloud_optical_depth_log) + 10(cloud_optical_depth_log_lower_layer)
CRE	cloud_particle_effective_radius in m
Phase	scene_classification (0=Clear, 1=SL liquid water, 2=SL ice water, 3=2L, 10=failed retrieval) (see bottom right panel of Figure 1)

Table 4: Example for the list of global attributes for an OCA file generated 19:00 UTC on the 19 May 2015.

Variable	Content
title	Cloud microphysical and bulk cloud properties derived from visible and infrared imager
institution	EUMETSAT
source	Observational satellite observation: SEVIRI as input
history	1.0
references	DOI:10.1029/2011JD015883, DOI:10.5194/amt-5-1889-2012, DOI:10.15770/EUM_SEC_CLM_0049
software_version	v1.0
comment	
instrument	SEVIRI
platform	MET-10
albedo_flag	1
albedo_filename	MSG3-SEVI-MSGCRMN-0100-0100-20150519131500.000000000Z
product_date	20150519
product_time	190000Z
generation_date	20200520
generation_time	003822.729

[1] Link valid 22/09/2021 http://cfconventions.org/Data/cf-standard-names/75/build/cf-standard-name-table.html

Table 5: Content of one NetCDF4 file – OCA statistics over the full disk.

Number	Name	Long name	Unit	Type
1	iterations_disk_average	Average number of iterations within processing area	1	float32
2	measurement_cost_function_average	Average solution measurement cost of values upto 500 within processing area	1	float32
3	apriori_cost_function_disk_average	Average apriori cost within processing area	1	float32
4	measurement_cost_function_disk_average_single_layer	Average single-layer solution measurement cost of values upto 500 within processing area	1	float32
5	measurement_cost_function_disk_average_multi_layer	Average multi-layer solution measurement cost of values upto 500 within processing area	1	float32
6	apriori_cost_function_disk_average_single_layer	Average single-layer apriori cost within processing area	1	float32
7	apriori_cost_function_disk_average_multi_layer	Average multi-layer apriori cost within processing area	1	float32
8	cloud_top_pressure_disk_average_ice_water_single_layer	Average single-layer cloud top pressure (ice water) within processing area	Pa	float32
9	cloud_top_pressure_disk_average_liquid_water_single_layer	Average single-layer cloud top pressure (liquid water) within processing area	Pa	float32
10	cloud_top_pressure_disk_average_upper_multi_layer	Average multi-layer cloud top pressure (upper layer) within processing area	Pa	float32
11	cloud_top_pressure_disk_average_lower_multi_layer	Average multi-layer cloud top pressure (lower layer) within processing area	Pa	float32
12	cloud_optical_depth_disk_average_ice_water_single_layer	Average single-layer cloud optical depth (ice water) within processing area	1	float32
13	cloud_optical_depth_disk_average_liquid_water_single_layer	Average single-layer cloud optical depth (liquid water) within processing area	1	float32
14	cloud_optical_depth_disk_average_upper_multi_layer	Average multi-layer cloud optical depth (upper layer) within processing area	1	float32
15	cloud_optical_depth_disk_average_lower_multi_layer	Average multi-layer cloud optical depth (lower layer) within processing area	1	float32
16	cloud_particle_effective_radius_disk_average_ice_water_single_layer	Average single-layer cloud particle_effective_radius (ice water) within processing area	m	float32
17	cloud_particle_effective_radius_disk_average_liquid_water_single_layer	Average single-layer cloud particle effective radius (liquid water) within processing area	m	float32
18	cloud_particle_effective_radius_disk_average_upper_multi_layer	Average multi-layer cloud particle effective radius (upper layer) within processing area	m	float32
19	toa_brightness_temperature_11microns_disk_average	Average modelled TOA brightness temperature (10.8um channel) within processing area	K	float32

20	toa_brightness_temperature_6microns_disk_average	Average modelled TOA brightness temperat ure (6.2um channel) within processing ar ea	K	float32
21	cloudy_pixel_disk_fraction	Fraction of cloudy pixels within process ing area	1	float32
22	measurement_cost_function_gt500_disk_fraction	Fraction of cloudy pixels with solution measurement cost greater than 500 within processing area	1	float32
23	unprocessed_cloudy_pixels_fraction	Fraction of unprocessed cloudy pixels wi thin processing area	1	float32
24	forward_model_sw_residuals_disk_average	Average residuals for short wave channel s within processing area	1	float32
25	forward_model_lw_residuals_disk_average	Average residuals for long wave channels within processing area	K	float32
26	surface_reflectivity_clear_sky_disk_average	Average clearsky reflectance in 0.6um, 0 .8um, and 1.6um channels over land withi n processing area	1	float32

Table 6: Content of one NetCDF4 file - OCA variables

Number	**Name**	**Long name**	**Unit**	**Actual Range**	**Valid Range**	**Scale**	**Offset**	**Missing Value**	**Type**
27	latitude	Latitude	degrees	-90., 90.	-32767, 32767	0.002746666	0.0	-32768	integer16
28	longitude	Longitude	degrees	-180.,180	-32767, 32767	0.005493332	0.0	-32768	integer16
29	scene_classification	Final cloud model ass umption (0=Clear, 1=S L liquid water, 2=SL ice water, 3=2L, 10=f ailed retrieval)	1	0, 10	0, 10	1	0	-128	byte8
30	measurement_cost_funct ion_log	Solution measurement cost (in Log10)	1	-1.0, 3.0	-127, 127	0.01574803	1.0	-128	byte8
31	cloud_optical_depth_lo g	Cloud optical thickne ss of single layer or , in two layer cases, the highest layer (in Log10)	1	-1.3, 2.532	-32767, 32767	5.847347e-0 5	0.6160 001	-32768	integer16
32	cloud_optical_depth_er ror_log	Error in cloud optica l thickness of single layer or, in two laye r cases, the highest layer (in Log10)	1	-1.0, 2.15	-127, 127	0.01240158	0.575	-128	byte8
33	cloud_particle_effecti ve_radius	Cloud particle effect ive radius of single layer or, in two laye r cases, the highest layer	m	0.0, 0.00010 2	-32767, 32767	1.556444e-0 9	5.1e-0 5	-32768	integer16
34	cloud_particle_effecti ve_radius_error_log	Error in cloud partic le effective radius o	m	-7.0, -3.85	-127, 127	0.01240158	-5.425	-128	byte8

		f single layer or, in two layer cases, the highest layer (in Log 10)							
35	cloud_top_pressure	Cloud top pressure of single layer or, in two layer cases, the highest layer	Pa	5000.0, 107300.0	-32767, 32767	1.561022	56150.0	-32768	integer16
36	cloud_top_pressure_error_log	Error in cloud top pressure of single layer or, in two layer cases, the highest layer (in Log10)	Pa	1.0, 5.15	-127, 127	0.01633858	3.075	-128	byte8
37	cloud_optical_depth_log_lower_layer	Cloud optical thickness of lower layer 2 (in Log10)	1	-1.30, 2.532	-32767, 32767	5.847347e-05	0.6160001	-32768	integer16
38	cloud_optical_depth_error_log_lower_layer	Error in cloud optical thickness of lower layer 2 (in Log10)	1	-1.0, 2.15	-127, 127	0.01240158	0.575	-128	byte8
39	cloud_top_pressure_lower_layer	Cloud top pressure of lower layer	Pa	5000.0, 107300.0	-32767, 32767	1.561022	56150.0	-32768	integer16
40	cloud_top_pressure_error_log_lower_layer	Error in cloud top pressure of lower layer (in Log10)	Pa	1.0, 5.15	-127, 127	0.01633858	3.075	-128	byte8
41	cloud_probability	Ancillary information: cloud probability in seven steps 0,25,40,50,60,75,100	%	0, 100	-127, 127	-	-	-127	byte8

4.4.2 Naming convention

The names of the NetCDF files (see Table 7) of OCA Release 1 follow WMO filename conventions that were adopted for NetCDF files (EUMETSAT, 2017).

Table 7: Description of the NetCDF filename for the CDR.

Filename convention for reprocessed netCDF Meteosat geostationary OCA	
W_XX-EUMETSAT-Darmstadt,OCA+<SATID>+<INSTRUMENT>_C_<ORIGINATOR>_<START_DATE>_1_OR_FES_<SSP>_<MAJOR_RELEASE_VERSION><MINOR_RELEASE_VERSION>.nc	
Example of a filename for OCA retrieved using images from SEVIRI on-board the MSG Meteosat 10 satellite	
W_XX-EUMETSAT-Darmstadt,OCA+MSG3+SEVIRI_C_EUMG_20150519190000_1_OR_FES_E0000_0100.nc	
Description	
<SATID>	Satellite used: MSG1 to MSG4 means MET08 to MET11
<INSTRUMENT>	MVIRI or SEVIRI
<ORIGINATOR>	EUMG or EUMS
<START_DATE>	Date of the image used (yyyyMMddhhmmss)
<SSP>	Sub satellite point: 0°
< MAJOR_RELEASE_VERSION >	Release major version
< MINOR_RELEASE_VERSION >	Release minor version

5 DATA RECORD GENERATION

5.1 Major steps of the retrieval

This processor is a standalone version of the operational software, OCA version 1.1, used for SEVIRI in the EUMETSAT Meteorological Products Extraction Facility (MPEF release 2.9) ground segment.

5.2 Input data

5.2.1 Meteosat imagery

The OCA Release 1 products are derived using imagery acquired by SEVIRI instruments on-board MSG satellites. All data are taken from the EUMETSAT NRT Archive system[2] with the calibrations therein applied. For the thermal infrared channels (3.9 to 13.3 µm) no further adjustment has been made. The three shortwave solar channels at 0.6, 0.8 and 1.6 µm are corrected for a known significant calibration bias according to the analysis of SEVIRI-MODIS matchups (Meirink et al., 2013). Although the biases are known to drift somewhat and be slightly different between the SEVIRI instruments, a single value per channel has been used in the generation of this version of the CDR as shown in Table 8.

Table 8: Calibration factors applied to the SEVIRI solar channel reflectance measurements

Channel	Calibration adjustment
0.6	1.08
0.8	1.06
1.6	0.96

[2] EUMETSAT product navigator OCA landing page: https://navigator.eumetsat.int/product/EO:EUM:DAT:MSG:OCA

5.2.2 Reanalysis model data

Interpretation of satellite imagery, especially for the thermal IR channels, requires a good estimate of the atmospheric temperature and humidity pertaining at the individual pixel's location and time. To supply these, reanalysis outputs from ECMWF ERA-Interim (Dee et al., 2011) forecast data are used as input for the reprocessing. We used the 6 and 12 hour reanalysis forecasts from the 0 and 12 UTC base times, leading to 4 forecast files per day (0, 6, 12 and 18 UTC). The forecast grid (spatial resolution) is 1°x1°. The parameters used for the processing are listed in Table 9. More details on the ERA-Interim re-analysis data can be found on the ECMWF web site[3].

Table 9: ERA-Interim data fields used for OCA processing.

Model Layers (1 -60) parameter	**Surface parameter**
Temperature	Surface skin temperature
Specific humidity	Surface skin pressure
Ozone	2m pressure

An example showing temperature and specific humidity at a height of around 200 hPa is given in Figure 4.

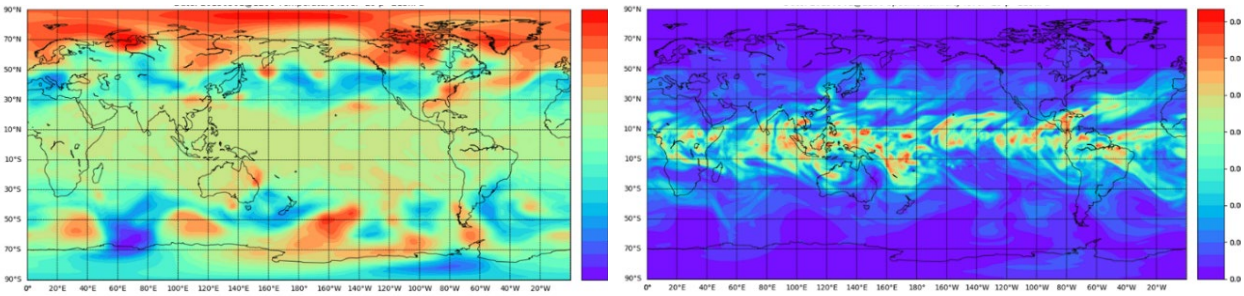

Figure 4: ERA-Interim forecast data: example of temperature (left panel) in Kelvin and specific humidity (right panel) in $kg.kg^{-1}$ @213hPa for the 1^{st} of January 2015.

5.2.3 RTTOV simulation

Re-analysis output data (see section 5.2.2) are used to feed the Radiative Transfer for TOVS (RTTOV version 11.2) (Hocking et al., 2015) radiative transfer model, in order to generate channel clear sky transmittances and radiances required to run the OCA algorithm. The RTTOV requires temperature, specific humidity, and ozone profiles at 54 specific pressure levels requiring, therefore, that the 60 ERA-Interim model levels are interpolated onto these 54 levels as described in Figure 5.

[3] Link valid 16/11/2021 https://www.ecmwf.int/en/forecasts/datasets/archive-datasets/reanalysis-datasets/era-interim

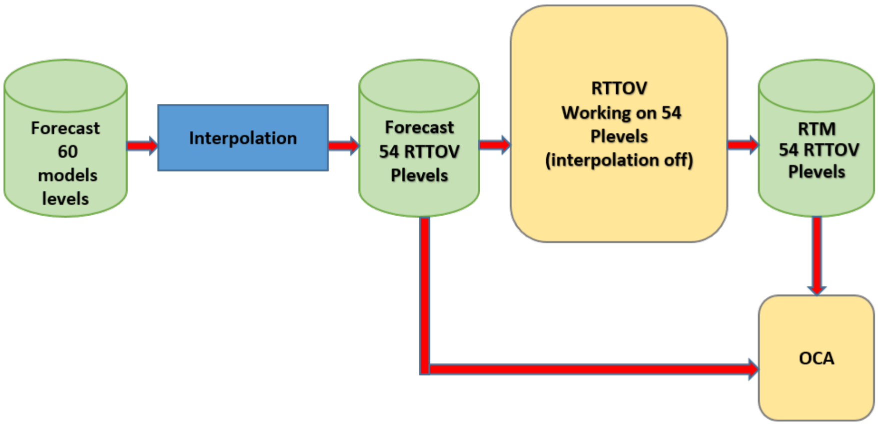

Figure 5: Preparation of input data for OCA retrieval module using forecast data from ECMWF reanalysis.

5.2.4 Cloud mask

The OCA algorithm itself does not determine whether a cloud is present or not and so a CLM is obtained from SEVIRI measurements using an algorithm developed at MeteoSwiss within the Climate Monitoring Satellite Application Facility (CM SAF) (Stöckli et al., 2017). The exploitation of a new CLM for the reprocessing of OCA was necessary because the operational NRT cloud mask is currently not archived at full resolution. The implemented scheme of the MeteoSwiss cloud mask is based on a Bayesian approach applied to a set of scores calculated exploiting only two solar (VIS) channels around 0.6 and 0.8 µm and one thermal infrared (IR) channel around 10.8 µm from SEVIRI. The algorithm also builds up a daily background reflectance map to assess potential clouds with higher reliability. The mask is provided as a cloud probability for 7 fixed steps: 0, 25, 40, 50, 60, 75, 100%. For this release, all pixels with a cloud probability higher than 40% are considered cloudy and passed to OCA for retrieval. The cloud probability has been included in the OCA products to allow users to further discriminate potential dubious solutions.

5.2.5 Surface reflectance

Interpretation of the solar reflectance measurements, even in cloudy situations, requires knowledge of the reflectivity of the underlying surface. An advantage of geostationary sensors is that they continually view the same ground locations so that, over a relatively short time, cloud-free daytime reflectance maps can be generated for the full disk. This is done operationally for the SEVIRI instruments and the resulting ‘Clear Sky Reflectance Maps’ (CRMN product) (EUMETSAT, 2015) can be obtained from the EUMETSAT archive. The NRT CRMN stores the mean cloud-free TOA reflectances for the three channels and they are updated, on a rolling basis, every day at 13:15 UT. The archive of CRMN is to a degree incomplete as it has been affected by operational outages and differing archiving procedures over time. As surface reflectance is a slowly changing phenomenon, this presented no significant resulting deficiencies. As backup, when the CRMN of a specific day was missing, the algorithm selected the first available for the same day in one of the following years. The principle behind this choice is that the surface reflectance is depending strongly on the illumination conditions, i.e. on the season (assuming of course no major changes in the surface coverage).

5.3 Cloud parameters estimation

The information on cloud properties contained in passive visible-infrared imagery is inherent in the differing relation of radiation at the various measured wavelengths with the cloud/atmosphere/surface. These relations are illustrated in Figure 6 for the most complex case handled by the OCA algorithm, that of two layer cloud. There are, broadly speaking, direct relations between: cloud top temperature[4] and thermal infra-red channel measurements, optical thickness and non-absorbing solar (visible) channel measurements and micro-physics (effective radius and phase) and absorbing solar (near-infrared) channel measurements. In the simplest possible cloud system – a very thick, plane parallel, and dense cloud over a non-reflecting (e.g. ocean) surface, these assumed relations effectively hold true. Should the cloud be optically thin though, there is transmission of thermal radiation through the cloud and measuring the top temperature becomes dependent on a knowledge of (or accounting for) the optical thickness. Similarly, the microphysics needs optical thickness as a reference to interpret the near infrared measurements.

The real world is complicated by a thermally emitting (infrared) surface and atmosphere and a reflecting (shortwave) surface which affect the measurements to an increasing degree as the cloud becomes thinner. It is not possible with the limited spectral information in imagers to simultaneously extract information on these parameters. Fortunately, the thermal state of the surface and atmosphere is accurately enough (generally) available from global numerical Weather Prediction (NWP) models (ECMWF ERA is used here) and the physical state of the surface (emissivity and albedo) are static enough that climatologies or recently accumulated estimates suffice.

The final complication arises when clouds deviate from their idealised representation in classical retrieval schemes, which is a plane parallel geometrically thin entity. The greatest deviation, and the only one tackled in the OCA used here, is the case of clouds of more than one distinct layer. The effect of layering is greatest on the cloud top temperature retrieval and then only in cases where the upper layer is thin (permitting the lower to be ‘visible’) and the lower sufficiently low compared to the top to provide a thermal contrast[5]. Retrieving the properties of the layers (two is definitely the maximum achievable) is stretching the imager capabilities to the limit and requires the thermal channels with significant atmospheric absorption, the 13 µm CO_2 channel and the two water vapour channels, to be fully utilised.

The sometimes complex dependencies of measurements on the ‘state parameters’ of the cloud retrieval problem outlined above is the reason that an optimal estimation framework based on Fast Radiative Transfer (FASTRTM) was chosen (and explains the ‘optimal’ in the algorithm name). It permits by nature the simultaneous use of the measurements available and is extendable to new measurements (e.g. with the next generation imagers) without fundamental changes (most likely only modification to the cloud model used to accommodate the new information).

[4] Cloud top *pressure* is ultimately estimated but note that the measurements are hardly sensitive to pressure directly and it is essentially inferred from the accompanying forecast temperature profile.
[5] It is worth noting here that daytime OCA retrievals are normally significantly more (and differently) sensitive to layered clouds than schemes that only use thermal infrared channels in the retrieval: Simultaneous use of the solar channels forces the retrieval to note the entire optical depth of the cloud. This may, and at first, did appear as a disadvantage of the method, but at the same time it permits the scheme to disentangle the layer properties at least in the cases where the sensitivity is high. For more details, see the references.

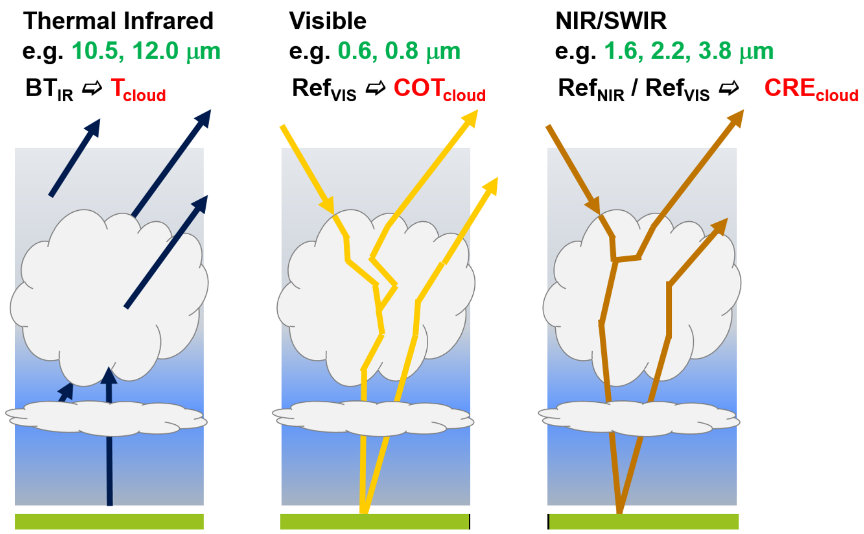

Figure 6: Relation between surface/atmosphere radiometric phenomenon and the SEVIRI channels used in OCA.

The FASTRTM is based on pre-calculated set of lookup-tables (LUTs) of cloud radiative properties per SEVIRI channel. It combines the cloud-free atmosphere channel radiances/transmittances and temperatures (section 5.2.3) with cloud emissivities, transmissivities and reflectivies from the LUTs. The LUTs have dependencies on COT, CRE and one or three dimensional geometry, the latter being applicable only to the bi-directional reflection function for solar affected channels. A set of LUTs is required for each of the two phases. The FASTRTM is completed with surface reflectivity and emissivity values for the solar and thermal components respectively.

The algorithm adjusts the cloud state to minimise, under certain constraints, the difference between the measurements and their equivalents calculated with FASTRM. This difference represented by the system 'cost' or 'penalty' function. The minimisation can branch and restart if checks on the cloud phase indicates it should change from the initial value. Similarly, but only at the termination of the minimisation, it can be re-started if diagnostics indicate that multi-layer cloud is present.
In addition to the core cloud properties of the estimation process there is a further parameter representing the skin temperature (TSK). This is included because of its strong effect on cloudy IR radiances (and optically thin cloud) and because it can be poorly estimated by the forecast model (e.g. over desert surfaces).

At the convergence of the minimisation a linear error analysis is performed to compute the expected cloud parameter errors based on the input assumptions on radiometric and FASTRTM model errors.

The penalty function normally has a single 'global' minimum but there are specific cases where distinct multiple minima occur. The most common cause is the presence of a boundary layer temperature inversion (in the NWP profile) which, through the dependency of the infrared channel measurement on temperature, can lead to minima with CTP above and below the inversion (although not necessarily with the same minimum penalty value). The minimisation algorithm cannot generally 'cross' this boundary and so it is important that the initial state ('first guess') is set correctly above or below. The product error analysis does not account for convergence to a 'false' minimum; only a resulting high solution cost would indicate a problem and this is not always the case.

Details of the handling of boundary layer inversions can be found in section 6.4.3 of the ATBD (EUMETSAT, 2016).

The retrieval consists of the following steps:

1) Initial Guess based on presumption of single layer model
2) Cost Minimization: Levenberg-Marquardt algorithm, with phase switching
3) Multi-Layer detection:
    a) Initial Guess based on presumption of two-layer model
    b) Cost Minimisation: Levenberg-Marquardt
    c) Post processing of retrieved parameters to two-layer state
4) Linear error analysis

Figure 7 shows the flow logic of these steps (apart from step 1).

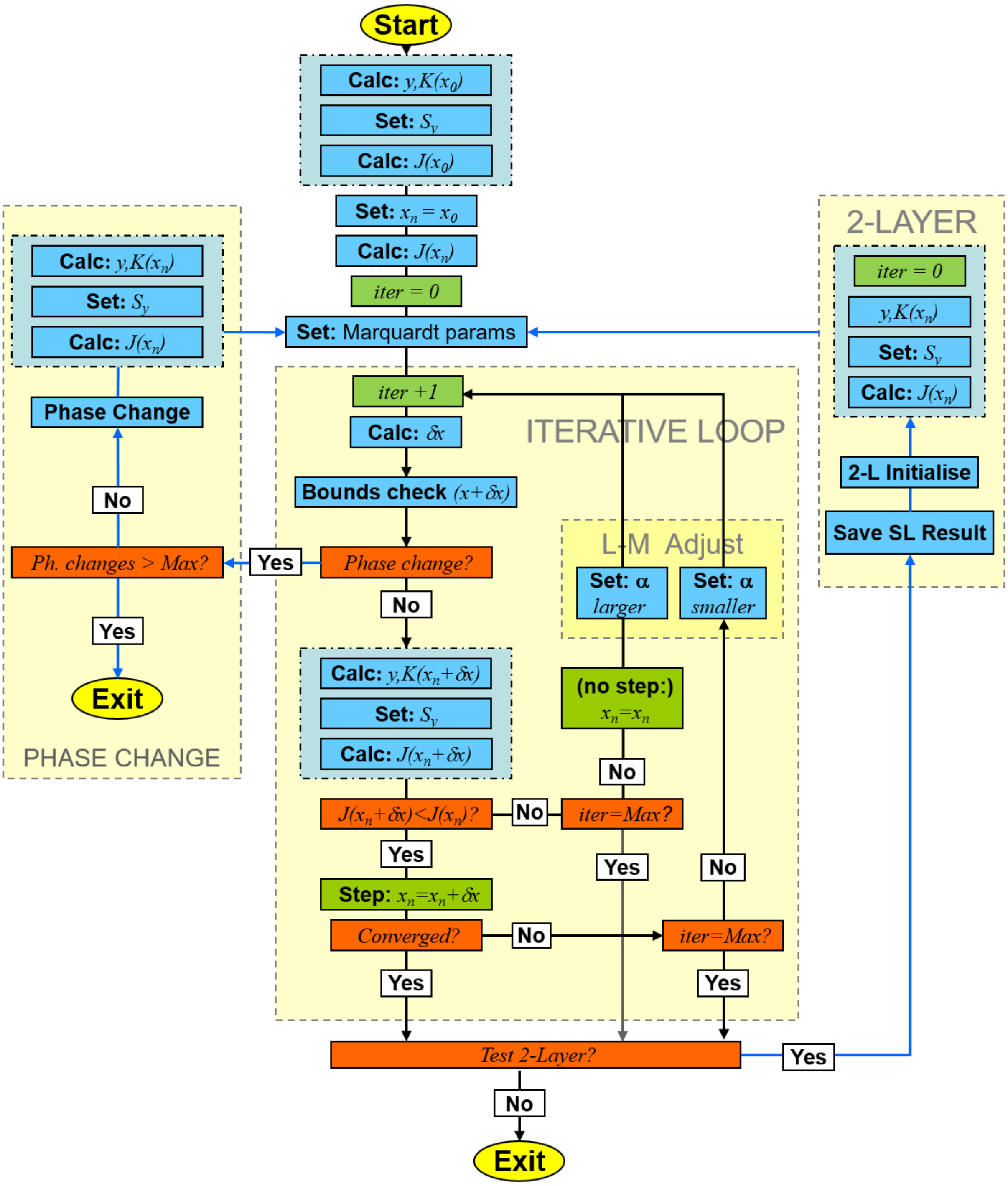

Figure 7: Cost minimization and multi-layer retrieval logical processing steps

A complete description of the OCA retrieval algorithm can be found in the ATBD (EUMETSAT, 2016)

6 VALIDATION SUMMARY

An extensive validation of OCA can be find in the validation report (EUMETSAT, 2021). The main summary of the validation are presented in this section.

The OCA CDR R1 cloud properties (cloud top pressure, cloud optical thickness, cloud particles effective radius and cloud phase) retrieved by a slightly adapted version of the NRT OCA algorithm using as inputs the 15 minutes SEVIRI measurements (from 4 Meteosat satellites) during the period 2004-2019. The OCA algorithm has been used in near real time at EUMETSAT as demonstrational SEVIRI L2 cloud product since 2013. The main feature that sets OCA apart from existing similar radiometer-based retrieval algorithms is its capability to detect and retrieve cloud properties for two overlapping cloud layers with full error estimates. The CDR extends the OCA retrieval back to 2004 by producing for the first time OCA using all of the Meteosat-9 and Meteosat-8 measurements.

A significant application of the OCA CDR is its usage as the main source for cloud height assignment and associated uncertainty for assigning the height of the SEVIRI Atmospheric Motion Vectors (AMV). The validation of the cloud top pressure is especially important. Nevertheless, the quality of other values retrieved by OCA has been included in the product validation.

The main goal of the validation was to confirm the homogeneity and consistency of the CDR and the accuracy of the reprocessing products, especially cloud top pressure/height, within the known limits of the algorithm throughout the entire period.

For this purpose, OCA CDR was compared against the operational OCA NRT archive product for the year 2018, collocated measurements and retrieval products from other independent satellite measurements such as polar orbiting radar and lidar from CloudSat and CALIPSO, and reference L3 products from MODIS and CLAAS 2.1 CDR covering the same period.

The main conclusion from the validation are as follows:

- The CDR is stable and homogeneous over the entire period.
- The CDR is consistent when compared against measurements and retrieval products from CloudSat and CALIPSO collocated with SEVIRI pixels over the period 2007-2015. The biases in the OCA products known already from previous studies, have been confirmed. In particular, when the reference cloud top height is defined by lidar measurements, radiometer-based cloud top height retrievals like OCA tend to be biased low, especially for high thin ice clouds (~1.5 km below the real top). The cloud top height retrieval for overlapping cloud layers shows larger biases than for single layer clouds. The retrieval of the cloud optical thickness and particle effective radius is most reliable for daytime measurements while the cloud top height retrieval is of comparable quality for daytime and nighttime SEVIRI slots.
- CRE is significantly more accurate when solar channels are available. Nighttime retrievals produce particle sizes that generally are too large, especially for single layer ice clouds. Similarly, the current two-layer algorithm is restricted to use of infrared channels and these retrievals also suffer from the lack of solar channel input. This last limitation can be mitigated by the use of a more complete radiative transfer model that includes solar channels when available. This is currently investigated for the next version of the algorithm.
- The CDR and archived near real time OCA products are consistent and in close agreement. However, some differences appear due to the usage of different input data for the retrieval. The differences in cloud top pressure between the CDR and NRT products are due to the usage of a

different cloud mask and different model data as inputs. The cloud mask used in the CDR has a larger amount of thin cirrus for which the reliability of OCA retrieval is low. The cloud mask is included in the data files so it can be used by users to filter out products based on different cloud probability thresholds. The CDR uses the ERA-Interim reanalyses meteorological profiles while the NRT uses operational model forecasts data. Those model data are used in the CDR at a higher vertical resolution than for the generation of the NRT products and this has an impact especially on the definition of cloud top height close to temperature inversions.

- The comparison of the OCA Release 1 CDR against other two independent satellite reference CDRs (MODIS L3 and CLAAS 2.1) shows that the products of the three different algorithms are compatible and consistent over the whole period 2004-2019. This is particularly true for the cloud top pressure and for the cloud optical thickness. However, for the cloud ice particle size where the results differ significantly. OCA shows a significant downward trend of the cloud optical thickness time series, not observed in the other two CDR. A possible cause of this is that it was not possible to incorporate the latest calibration information on the short-wave channels in the creation of this CDR.

7 LIMITATIONS AND POSSIBLE IMPROVEMENTS

The validation of the OCA Release 1 CDR highlighted the following limitations that should be considered by the users. Potential improvements to the algorithm or its inputs are given when possible and if already implemented for tests.

1. The daytime and nighttime products from OCA have different systematic and random uncertainties in COT and CRE. This limitation does not apply to the CTP. Since the ratio between daytime and nighttime COT and CRE retrievals varies with location and season, one should restrict their usage to only daytime or only nighttime observations in climate research, such as studying diurnal cycles of seasonal variations.
2. Unrealistic COT retrieval during nighttime is made worse by false or less sensitive cloud screening.
3. Anomalously high CTP values over bright surfaces. This is due to the inherently challenging cloud detection and properties retrieval over such areas. However, for the current CDR, this could come from a false or less sensitive cloud screening from the cloud mask adopted in the CDR added a negative impact. As mitigation, the input cloud mask is included in the product so that any user can decide for further pixel screening depending on the application. **Improvement**: A state of the art cloud mask or the current mask with a threshold dependent on the underlying surface conditions will be used.
4. Like other retrievals based on passive imagers, OCA CTPs are generally higher (lower altitude) than observed CTPs (e.g. as sensed by satellite or airborne lidars). Maximum differences (~1.5 km) are found for high thin ice clouds. **Improvement**: More realistic modelling of cloud extinction in the algorithm forward radiative transfer.
5. Clear sky reflectance maps were not archived operationally at EUMETSAT for the complete period (2004-2019). Maps for a different day or even for a different year (same day) were used in case of a gap. This may have introduced some level of dynamic but systematic error in the daytime retrieval of COT and CRE for thinner clouds. **Improvement**: Regeneration of the missing clear sky reflectance maps.

8 PRODUCT ORDERING

The data are accessible via the EUMETSAT Data Centre. To access the data from EUMETSAT, you need to register with the EUMETSAT Data Centre. When registered, you can order the data through a written request send to EUMETSAT's helpdesk.

8.1 Register with EUMETSAT Data Centre

Do this to register with the EUMETSAT Data Centre:

1. Register in the EUMETSAT EO-Portal (https://eoportal.eumetsat.int/) by clicking on the New User – Create New Account tab;
2. After finalisation of the registration process, an e-mail is sent to the e-mail address entered in the registration. Click the confirmation link in the e-mail to activate your account;
3. Login and subscribe to the Data Centre Service by going to the Service Subscription Tab and selecting Data Centre Service. Follow instructions issued from the web page to add needed information.

Order data

The data record described in this product user guide can also be ordered via the EUMETSAT User Service Helpdesk in Darmstadt, Germany. Please send a written request to this helpdesk, email ops@eumetsat.int, indicating the data record that you want to order including its Digital Object Identifier (DOI) number.

9 PRODUCT SUPPORT AND FEEDBACK

For enquiries about the CDR described in this product user guide, please contact the EUMETSAT User Service Helpdesk by email: *ops@eumetsat.int*.

10 PRODUCT REFERENCING

The data record described in this product user guide has a unique DOI that should be used for referencing. The product's filename provides a unique identifier for each product.

11 ACKNOWLEDGEMENTS

None

APPENDIX B DAY/NIGHT PIXEL

During the OCA processing each pixel is classified as daily or nightly according to a threshold for the solar zenith angle (SZA). The value of the threshold is 80 degrees. Pixel with a SZA<80 are classifies as daily, while pixels with a SZA≥80 are classified as nightly pixels. This information is not currently stored in the OCA product. To overcome this current limitation, the formula to calculate the SZA together with an example in python code are provided.

The SZA θ_0 can be calculated as follows[6]:

$$\cos\theta_0 = \sin\varphi\,\sin\delta + \cos\varphi\,\cos\delta\,\cosh$$

Where:

φ *is the pixel latitude*

$h = 15^\circ \times (LST - 12)$ *is the hour angle*[7]

$\left(LST = hour + \frac{minutes}{60}\right.$ *is the local solar time in 24 hours format*$\left.\right)$

The solar declination angle can be aproximanted[8]:

$$\delta = -23.45\cos\left(\frac{360}{365}(jday + 10)\right)$$

jday is the day in the year

The SZA can be calculated in python with these few lines of code.
NOTE: EUMETSAT does not support such code, provided on best effort basis only:

```
from pyorbital import astronomy
from datetime import datetime
import numpy as np

# pixel date and time
year   = 2010
month  = 5
day    = 26
hour   = 12
minute = 23

# pixel location (decimal degree)
lon = 12
lat = 23

b = datetime(int(year),int(month),int(day),int(hour),int(minute),0,0)
print('processing date:',b)
print('location: latitude=%.1f, longitude=%.1f' % (lat,lon) )

sol_zen = np.arccos(astronomy.cos_zen(b,lon,lat))*180./np.pi
print('SZA=%.1f degrees' % sol_zen)
```

output:

```
processing date: 2010-05-26 12:23:00
location: latitude=23.0, longitude=12.0
SZA=17.2 degrees
```

[6] From: https://solarsena.com/solar-elevation-angle-altitude/ (Link valid 18/05/2022)
[7] From: https://solarsena.com/solar-hour-angle-calculator-formula/ (Link valid 18/05/2022)
[8] From: https://solarsena.com/solar-declination-angle-calculator/ (Link valid 18/05/2022)

APPENDIX C MSG 0DEG PRIME SATELLITE

Due to maintenance, the prime satellite is not always the same during a period. There are several platform swaps in the period. The Table 10 indicates which Meteosat was the prime satellite during the period 2004 – 2019 used for the creation of the OCA CDR.

Table 10: Detailed list of prime MSG satellite at 0-degrees

Satellite	Start date	End date
MET08	19/01/2004	23/09/2006
MET09	25/09/2006	10/10/2006
MET08	10/10/2006	11/04/2007
MET09	11/04/2007	03/12/2007
MET08	03/12/2007	12/12/2007
MET09	12/12/2007	05/05/2008
MET09	05/05/2008	13/05/2008
MET08	13/05/2008	19/05/2008
MET09	19/05/2008	17/04/2009
MET08	17/04/2009	23/04/2009
MET09	23/04/2009	15/08/2009
MET08	15/08/2009	21/08/2009
MET09	21/08/2009	21/01/2013
MET10	21/01/2013	01/07/2013
MET08	01/07/2013	09/07/2013
MET10	09/07/2013	14/01/2014
MET09	14/01/2014	21/01/2014
MET10	21/01/2014	20/02/2018
MET11	20/02/2018	21/01/2019
MET09	21/01/2019	28/01/2019
MET11	28/01/2019	13/01/2020
MET09	13/01/2020	27/01/2020
MET11	27/01/2020	25/02/2021

APPENDIX D BLACKLISTING

A post processing analysis was made on the available[9] CDR products. Some files may need to be blacklisted by the user depending on its usage and/or needs.
A machine learning approach has been used. Two different models have been exploited to look for potential anomalous files: an Autoencoder machine-learning (ML) model[10] and a classifier based on the Mahalanobis distance[11]. Both methods run on five distinct parameters: the file size, the averaged CTP, the cost, the percentage of cloudy pixels, and the percentage of unprocessed pixels. It has been decided to flag as potential outliers only the files found in both methods to increase the robustness of the post processing blacklisting. The number of common outliers (see Figure 8) is 286 (0.053% of the total).

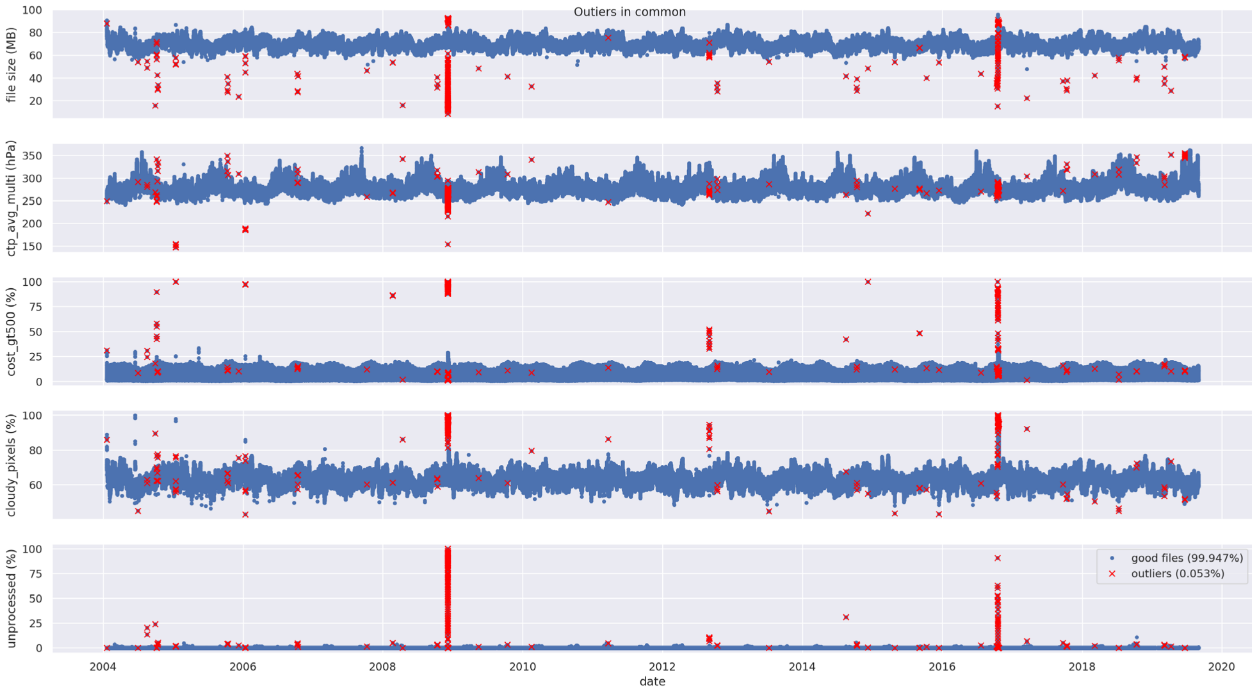

Figure 8: Common outliers retrieved by two different machine-learning methods.

The list of the potential outliers is available for the interested users.

[9] A product is generated if all needed input data (satellite image, cloud mask) are available.
[10] E.g. https://en.wikipedia.org/wiki/Autoencoder (Link valid 24/05/2022)
[11] E.g. https://en.wikipedia.org/wiki/Mahalanobis_distance (Link valid 24/05/2022)