THEIA

DIRECTION DU NUMERIQUE, DE L'EXPLOITATION ET DES OPERATIONS

OBSERVATION DE LA TERRE

DEVELOPPEMENT SEGMENTS SOL IMAGERIE ET SONDAGE

THEIA-ST-411-0477-CNES

Change	: 01	Date	: 01/03/2018
Issue	: 03	Date	: 12/06/2019

Ref. : intentionally empty

Distribution Code : E

TECHNICAL SPECIFICATION

FORMAT SPECIFICATION OF OBS2CO WATERCOLOR PRODUCTS

Written by : KETTIG Peter DNO/OT /IS	Date :	
Approved by : BAILLARIN Simon DNO/OT /IS DONADIEU Joëlle DNO/OT /IS HAGOLLE Olivier DSO/SI / /CB	Date :	
For application : SELLE Arnaud DNO/OT /TA PACHOLCZYK Philippe DNO/OT /TA	Date :	

INDEX SHEET

CONFIDENTIALITY : DLP 12/06/2029	KEYWORDS : Muscate, L2B, Water, OBS2CO, WaterColor, Format

TITLE :

Technical Specification
FORMAT SPECIFICATION OF OBS2CO WATERCOLOR PRODUCTS

AUTHOR(S) :

KETTIG Peter	DNO/OT /IS

SUMMARY : Document to describe the format of Level-2B-WATER products in MUSCATE

RELATED DOCUMENTS : Stand alone document.	LOCALIZATION :

VOLUME : 1	TOTAL NUMBER OF PAGES : 16 INCLUDING PRELIMINARY PAGES : 5 NUMBER OF SUPPL. PAGES : 0	COMPOSITE DOCUMENT : N	LANGUAGE : EN

CONFIGURATION MANAGEMENT : No	CM RESP. :

REASONS FOR EVOLUTION : Updated to v1.0 of the processing chain: Updated values for masks

CONTRACT : None

HOST SYSTEM :
Microsoft Word 16.0 (16.0.4771)
\\bacchus\Gdoc\ModeleGDOC_en.dot
Version GDOC : v4.5.1
Base projet : \\to05res04\GdocBasesPartagees\Projets\CST\MUSCATE.accdb

INTERNAL DISTRIBUTION

Name	Entity	Internal Postal Box	Observations
KETTIG Peter	DNO/OT /IS	1501	
BAILLARIN Simon	DNO/OT /IS	1501	
DONADIEU Joëlle	DNO/OT /IS	1501	
PACHOLCZYK Philippe	DNO/OT /TA	612	
SELLE Arnaud	DNO/OT /TA	612	
HAGOLLE Olivier	DSO/SI / /CB	2801	

EXTERNAL DISTRIBUTION

Name	Entity	Observations
CLESSE Dominique	CAPGEMINI POUR DCT/PS/OT	
FAVOT Laurent	CAPGEMINI POUR DCT/PS/OT	

CHANGES

Issue	Rev.	Date	Reference, Author(s), Reasons for evolution
01	03	12/06/2019	intentionally empty KETTIG Peter DNO/OT /IS Updated to v1.0 of the processing chain: Updated values for masks
01	02	12-02-2019	intentionally empty KETTIG Peter DNO/OT /IS Updated concentration masks: One single concentration mask is created at the end of the chain instead of three over different bands. The mask is designated *_CON_ALL.tif in the product folder.
01	01	27-11-2018	intentionally empty KETTIG Peter DNO/OT /IS Added csv description
01	00	19-12-2017	intentionally empty KETTIG Peter DNO/OT /IS Creation of the document

CENTRE NATIONAL D'ETUDES SPATIALES	THEIA-ST-411-0477-CNES			
THEIA	Issue	**: 01**	Date	**: 01/03/2018**
	Rev.	**: 03**	Date	**: 12/06/2019**
Reference : intentionally empty	Page : i.5			

TABLE OF CONTENTS

GLOSSARY AND LIST OF TBC AND TBD ITEMS

PHOEBUS	Processing High Level Orchestration Engine and BUiness Services

List of TBC items:

List of TBD items:

1. OVERVIEW

1.1. SCOPE OF THE DOCUMENT

The aim of this document is to provide a definition for the Level-3A format in the scope of the Muscate project. To define this format, two existing ones are used: The Sen2Agri-L3 format, as well as the Muscate Level-2 one.

1.2. REFERENCE DOCUMENTS

RD1   GeoTIFF Format specification GeoTIFF revision 1.0 Specification Version 1.8.2 Last Modified 10 November , 1995
       Geotiff

RD2   Geospatial Data abstraction Library http://www.gdal.org
       Gdal

1.3. APPLICABLE DOCUMENTS

AD1   SPECIFICATION DE FORMAT DES PRODUITS (SPECIFICATIONS LOGICIELLES)
       CLESSE Dominique, 06-Jan-17, Issue 01, Rev. 09
       PSC-SL-411-0032-CG

AD2   Spécifications Technique de Besoin MUSCATE
       Issue 01, Rev. 00
       PSC-ST-40-0049-CN

2. NEW FORMAT DESCRIPTION

This section will describe the L2A-WATER format for MUSCATE.

2.1. MAIN FOLDER

The main folder of the format will be named after the Muscate Level-2B-WATER convention, which is shown in the following:

MISSION_YYYYMMDD-xxxxxx-xxx_L2B-WATER_TTILE#_C_Vx-x

MISSION: The mission name, which can be:

- SENTINEL2A

- SENTINEL2B

YYYYMMDD : The product date in the given format

xxxxxx-xxx: The Time in format HHmmSS-sss with sss being the milliseconds

TILE#: The Tile identifier

x-x : The Product version: Starting from 0-1 and increasing by 0-1 for a minor reprocessing or 1-0 for a major one.

2.1.1. Concentration rasters

The processing chain outputs rasters of the concentration of SPM estimated in the Bands B4 and B8.

The naming convention for it is:

MISSION_YYYYMMDD-xxxxxx-xxx_L2B-WATER_TTILE#_C_ Vx-x_CON_ALL.tif

The rest can be found above.

The concentration is given in **mg/L**. So a pixel value of 21.34 corresponds to 21.34 mg/L estimated at this point. The value -10000 signifies that there is no- or invalid data available.

The concentration is always calculated **only over the pixels classified as water.**

2.1.2. RGB rasters

An RGB raster for the given ROI is also included. The naming convention for it is:

MISSION_YYYYMMDD-xxxxxx-xxx_L2B-WATER_TTILE#_C_ Vx-x_RGB_ALL.tif

The rest can be found above.

The values correspond to reflectance TOC (Top-Of-Canopy), which is unitless. The value -10000 signifies that

there is no- or invalid data available.

2.1.3. Accompanying masks

The three masks generated by the Temporal-Synthesis are described in the following section. All masks are distributed in **R1 (10m)**. This means that for all surface reflectance files (see above), a mask of the corresponding resolution is available.

The respective files are located inside the MASKS folder of the root, with the filename being:

MISSION_YYYYMMDD-xxxxxx-xxx_L2B-WATER_TTILE#_C_ Vx-x_XXX_Rx.tif

XXX: The trigram for each mask as given above

X: The resolution of the image, being one of the two groups (ALL = 10m)

The rest can be found above.

2.1.4. How to use the masks

The masks have to be subtracted from the main concentration rasters in order to filter out adjacent effects on the water surfaces, such as:

- Clouds
- River banks
- Negative pixels
- Etc.

The concentration is always calculated **only over the pixels classified as water** (=The final 'CLU' mask). The common NODATA value for all masks is set to **-10000**.

2.1.4.1. PYN: Polygon mask

8bit mask

Defines the shape of the polygon used to clip the images.

- -10000 = Outside

- 1 = Inside

2.1.4.2. DIS: Cloud Disturbance mask

8bit mask

Indicates the clouds detected in the Level-2A produt for the given ROI.

- -10000 = Cloud/shadow; no data

- 1 = No cloud/shadow

2.1.4.3.   MND: MNDWI mask

8bit mask

Indicates the pixels identified as "no-water" using a threshold on the MNDWI.

- 1 = No water

- -10000 = No data

2.1.4.4.   NEG: Negative pixels mask

8bit mask

Indicates negative pixels due to the atmospheric correction of the images.

- 1 = Non-negative Pixel

- 0 = Negative Pixel

- -10000 = No data

2.1.4.5.   CLU: Clustering mask

8bit mask

Indicates the pixels detected as "water" using an unsupervised classification on the RED/NIR bands.

- -10000 = No water; no data

- 1 = Water

2.1.5.   Statistics file

The additional statistics file with history of the previous dates can be found in the DATA/ folder. One file exists for each polygon present in the raster.

Statistics describe the min,max,mean,median of each band as well as the concentration estimates.

Its filename is MISSION_YYYYMMDD-xxxxxx-xxx_L2B-WATER_TTILE#_ID_C_ Vx-x_WAT.csv

With the rest explained as above. Again, the NODATA value is set to -10000.

In the case of a possible sunglint, the column "Sunglint" for that specific date is **set to 1, otherwise 0.**

2.1.6.   Metadata-file

The main Metadata-file is located in the product root folder, called

MISSION_YYYYMMDD-xxxxxx-xxx_L2B-WATER_TTILE#_C_ Vx-x_MTD_ALL.XML

With all the name conventions explained above.

An example for a product of Sentinel-2A and Sentinel-2B is given in the Annex.

2.2.   METADATA INFORMATION TABLE

The following table illustrates where the information for each field comes from, similar to the table found in AD1

Field Name	Origin	Commentary
METADATA_FORMAT	Constant	=METADATA_MUSCATE
METADATA_PROFILE	Constant	=GENERIC
METADATA_INFORMATION	Constant	=EXPERT
PRODUCT_ID	Calculated following the format above	
AUTHORITY	Constant	=THEIA
PRODUCER	Constant	=MUSCATE
PROJECT	Calculated	
GEOGRAPHICAL_ZONE	Native metadata	
ACQUISITION_DATE	Calculated	Date for the temp. synthesis
PRODUCTION_DATE	Calculated	Current time of processing
PRODUCT_VERSION	Calculated	Starting at 1.0, then 1.1 ...
PRODUCT_LEVEL	Calculated	=L2B-WATER
PLATFORM	Native metadata	
ORBIT_NUMBER	Constant	From Level-2A

CENTRE NATIONAL D'ETUDES SPATIALES

THEIA

THEIA-ST-411-0477-CNES

Issue **: 01**   Date   **: 01/03/2018**

Rev.   **: 03**   Date   **: 12/06/2019**

Reference : intentionally empty

Page : 7

UTC_Acquisition_Range/MEAN	Native metadata	
Band_Global_List	Native metadata	With platform-id and count
Band_Group_List	Native metadata	Groups R1, R2 (if existing)
QUICKLOOK	Calculated	
IMAGE_FILE	Calculated	According to the naming convention above
MASK_FILE	Calculated	According to the naming convention above
DATA_FILE	Calculated	According to the naming convention above
GEO_TABLES	Constant	=EPSG
HORIZONTAL_CS_TYPE	Native metadata	
HORIZONTAL_CS_NAME	Calculated	
HORIZONTAL_CS_CODE	Native metadata	
RASTER_CS	Native metadata	
METADATA_CS	Native metadata	
Geopositioning	Native metadata	According to the synthesis-input
REFLECTANCE_QUANTIFICATION_VALUE	Constant	=10000
SPECIAL_VALUE	Constant	Nodata = -10000
Spectral_Band_Informations_List	Native metadata	According to the platform
SPATIAL_RESOLUTION	Native metadata	Only field inside the Spectral_Band_Informations_List to be filled out; Mandatory for MUSCATE
Product_Quality_List	Calculated	level=N2B-WATER (constant) PRODUCT_ID, ACQUISITION_DATE and PRODUCTION_DATE according to the same fields above Quality_Index is calculated
Contributing_Products_List	Calculated	level=N2

CENTRE NATIONAL D'ETUDES SPATIALES

THEIA

Reference : intentionally empty

THEIA-ST-411-0477-CNES

Issue **: 01**   Date   **: 01/03/2018**

Rev.   **: 03**   Date   **: 12/06/2019**

Page : 8

		PRODUCT_ID, SYNTHESIS_DATE, PRODUCTION_DATE and Quality_Index according to the input files
Processing_Information	Calculated	This section is added by PHOEBUS during the processing

3. GEOTIFF FORMATS

The new format will exist in two versions, which are specified in the following.

3.1. CLASSICAL GEOTIFF

This format is compliant to the GeoTIFF specification of RD1 and RD2. This implies that all images described above are stored in the GeoTiff format set by GDAL.

A GeoTiff is a .tif image with added geometrical information in its metadata-header, which sets the projection, resolution and position of it.

3.2. CLOUD-OPTIMIZED GEOTIFF

This is format is based on 3.1, with the addition of the following:

"A Cloud Optimized GeoTIFF (COG) is a regular GeoTIFF file, aimed at being hosted on a HTTP file server, with an internal organization that enables more efficient workflows on the cloud. It does this by leveraging the ability of clients issuing HTTP GET range requests to ask for just the parts of a file they need."

(cf. http://www.cogeo.org/)

The specifications and how to transform the regular GeoTiff to a CoG one can be found under:

https://trac.osgeo.org/gdal/wiki/CloudOptimizedGeoTIFF

The format incorporates the Gdal-Overviews as well as adds a compression to each image. Hence, the increase of size by adding the overview is partly equalized by the compression. This, however, results in a more varying product size, which will be described in 1.

A description when each format is used can be found in the next section.

3.3. FORMAT COMPATIBILITY

Because 3.2 is fully compatible to 3.1 as long as it is a fully defined GeoTiff as stated in RD1 , there is no difference to the user. To describe which of the two is used, the field Compression for each Image and Mask in the Product_Characteristics-Section are filled out as follows:

GeoTiff:

`<COMPRESSION>None</COMPRESSION>`

Co-GeoTiff:

```
<COMPRESSION>DEFLATE</COMPRESSION>
```

4. EXAMPLES

The following excerpt shows the resulting project structure for a complete L2B-WATER product.

4.1. PRODUCT STRUCTURE

```
├── DATA
│   └── SENTINEL2A_20160113-104010-950_L2B-WATER_T30PYT_001_D_V0-1_WAT_ALL.csv
├── MASKS
│   ├── SENTINEL2A_20160113-104010-950_L2B-WATER_T30PYT_D_V0-1_CLU_ALL.tif
│   ├── SENTINEL2A_20160113-104010-950_L2B-WATER_T30PYT_D_V0-1_DIS_ALL.tif
│   ├── SENTINEL2A_20160113-104010-950_L2B-WATER_T30PYT_D_V0-1_MND_ALL.tif
│   ├── SENTINEL2A_20160113-104010-950_L2B-WATER_T30PYT_D_V0-1_NEG_ALL.tif
│   └── SENTINEL2A_20160113-104010-950_L2B-WATER_T30PYT_D_V0-1_PYN_ALL.tif
├── SENTINEL2A_20160113-104010-950_L2B-WATER_T30PYT_D_V0-1_CON_ALL.tif
├── SENTINEL2A_20160113-104010-950_L2B-WATER_T30PYT_D_V0-1_MTD_ALL.xml
├── SENTINEL2A_20160113-104010-950_L2B-WATER_T30PYT_D_V0-1_QKL_ALL.jpg
├── SENTINEL2A_20160113-104010-950_L2B-WATER_T30PYT_D_V0-1_RGB_ALL.tif
└── SENTINEL2A_20160113-104010-950_L2B-WATER_T30PYT_D_V0-1_VIS_ALL.tif
```