Accessing and working with EUMETSAT Copernicus Marine Data

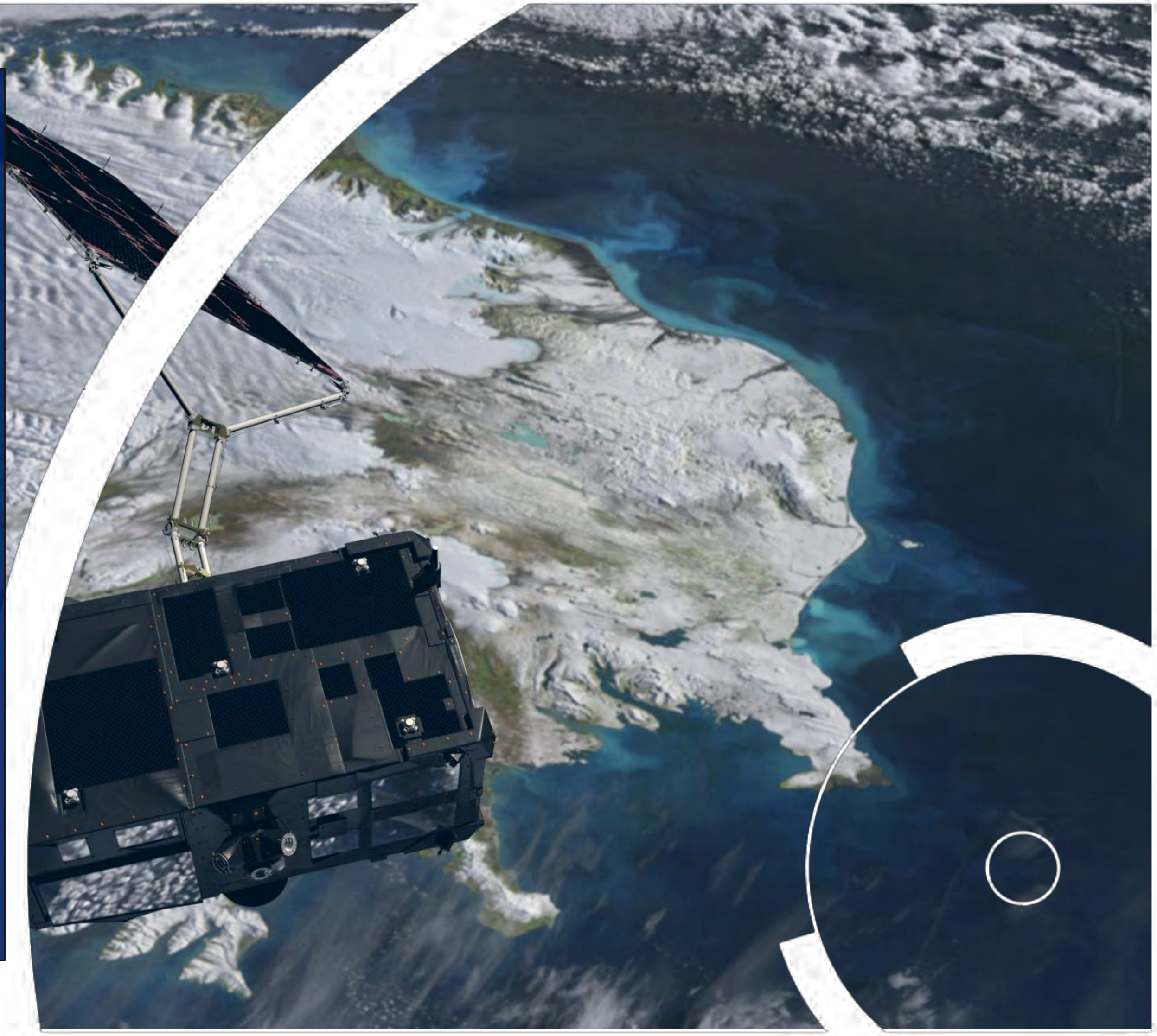
Hayley Evers-King[1], Juan Ignacio Gossn[1] , Benjamin Loveday[2]

Hayley.EversKing@eumetsat.int
JuanIgnacio.Gossn@eumetsat.int

Helpdesk: ops@eumetsat.int

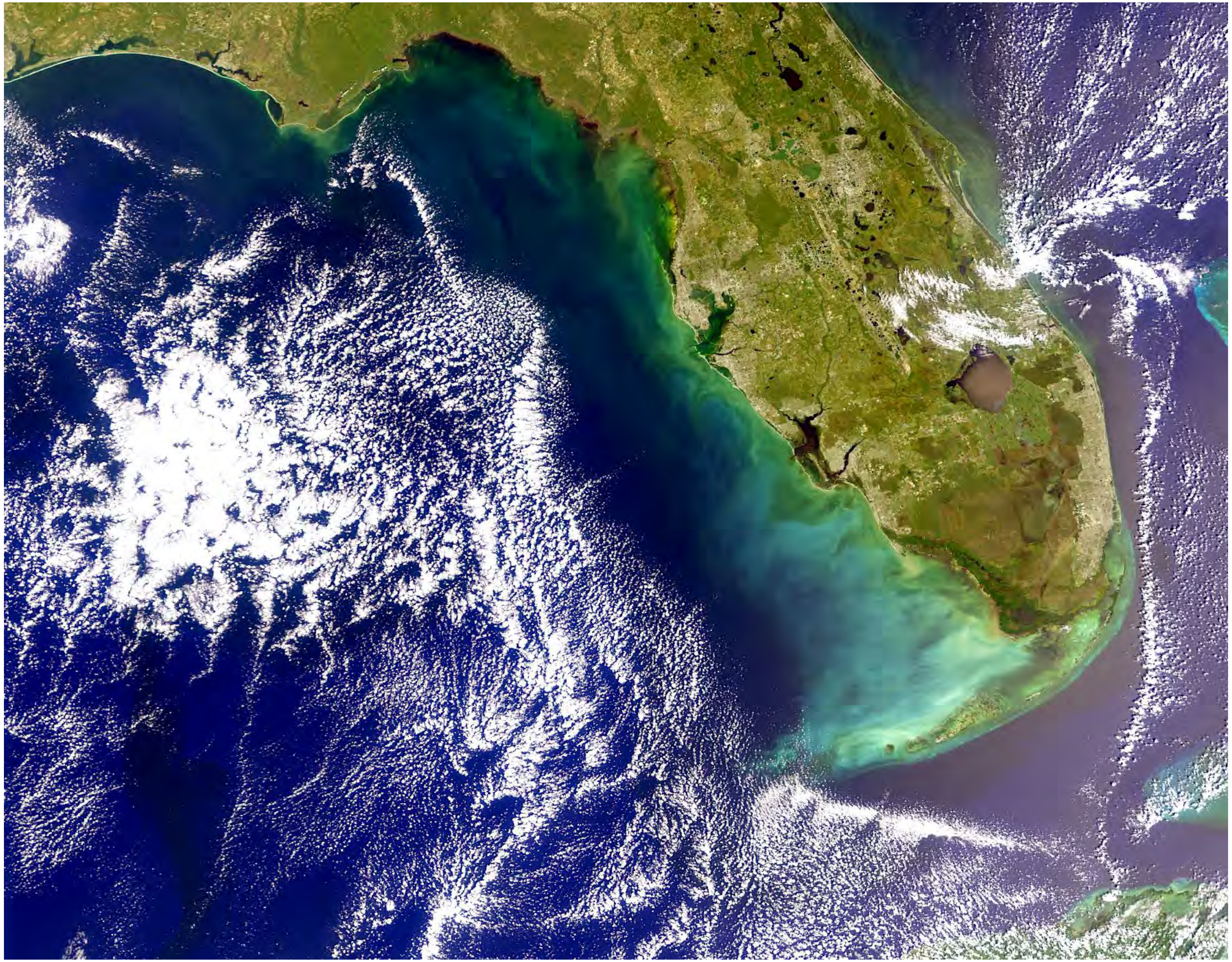
[1]EUMETSAT, Eumetsat Allee 1, Darmstadt, Germany
[2]Innoflair UG, Germany

IOCS, 11/2023



Welcome - a recent view of the region from Sentinel-3 OLCI

copernicus.eumetsat.int



What will we cover today?

- Part 1:
  - This introduction – EUMETSAT, Copernicus, Sentinel-3 and 6 (Hayley)
  - Accessing and working with Copernicus Data in Python (Hayley)
- Part 2:
  - ThoMaS – a Tool to perform Matchups with S3-OLCI (Juan)

Time to explore! Choose what you would most like to focus on from the tools presented. We will support you to work through the example workflows and answer any questions you have about our data and services.

Feedback always welcome – how can we improve the tools/training, chances to collaborate etc!

Your instructors

Dr Hayley Evers-King (EUMESAT): Lead Marine Applications Expert, User Support and Climate Division

Hayley provides user support and training on ocean applications of satellite data, and manages feedback between user(s) and operational satellite agencies. She has worked throughout the satellite data value chain; from the validation of satellite sensor measurements, to algorithm and application development. She is a keen programmer and a passionate science communicator. Prior to her current role, Hayley worked for 5 years as a Marine Earth Observation Scientist at Plymouth Marine Laboratory. She obtained a PhD in ocean remote sensing from the University of Cape Town, South Africa in 2014.

Dr Juan Ignacio Gossn (EUMETSAT): Remote Sensing Scientist; Ocean Colour

Juan Gossn was born in 1990 in San Isidro, Argentina. He is a remote sensing scientist specializing in Ocean Colour remote sensing. With a background in Physics from the University of Buenos Aires, he obtained his Ph.D. from the Argentinean Institute of Astronomy and Space Physics (IAFE – CONICET/UBA), focusing on the extremely turbid waters of Río de la Plata. In 2021, he joined EUMETSAT's Ocean Colour Services team, where he contributes to the improvement of operational ocean colour products through the analysis of in situ data and managing scientific studies.

Who are EUMETSAT?

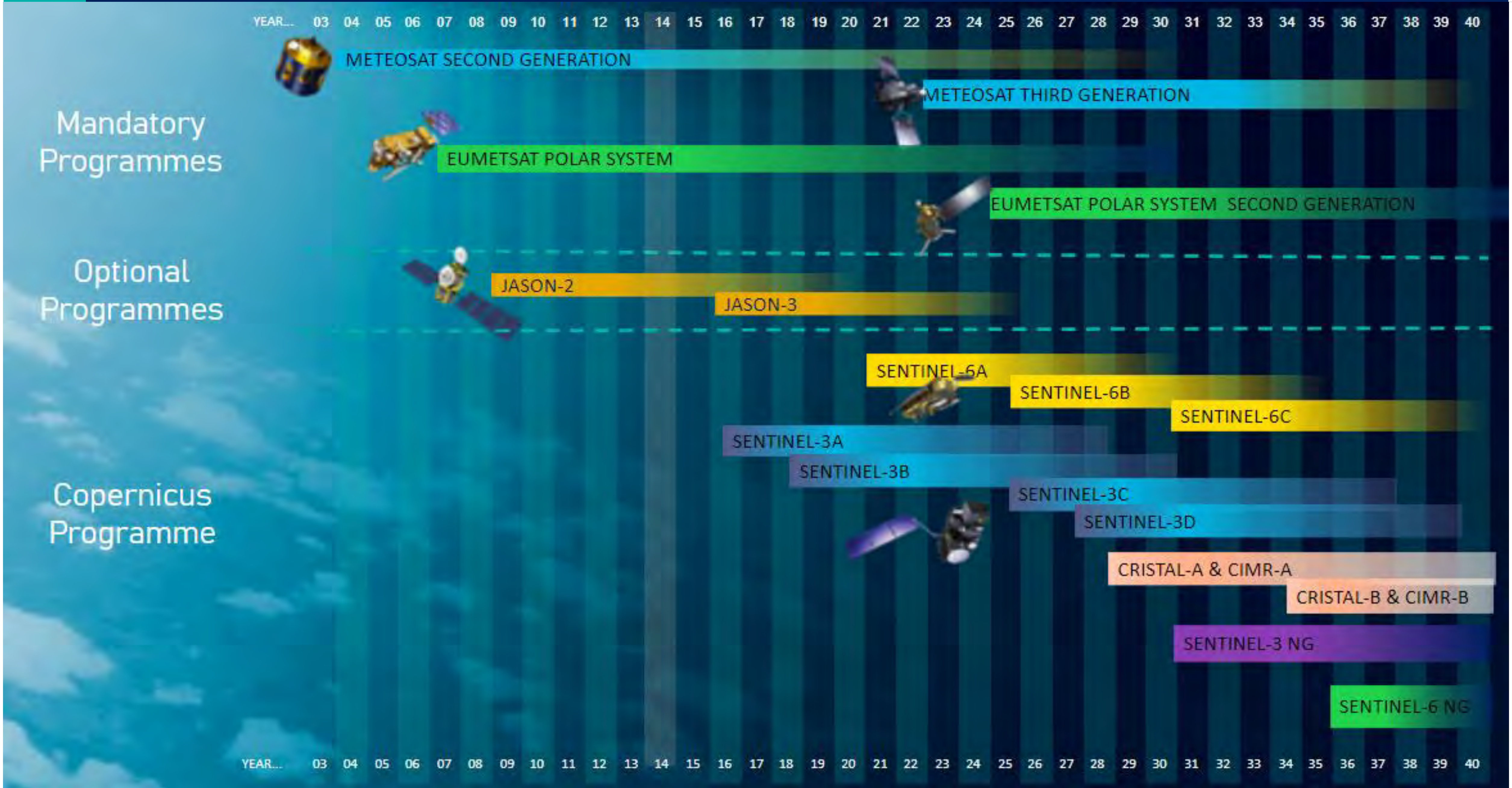
The European Organisation for the Exploitation of Meteorological Satellites

- Located in Darmstadt, Germany
- Founded in 1986, consists of 30 member states
- Two mandates:
  - Weather and Climate data for member states
  - Additional capabilities with EU and beyond:
    - Copernicus programme

Copernicus Europe's eyes on Earth

EUMETSAT OSI SAF OCEAN AND SEA ICE

EUMETSAT missions for marine – current and future

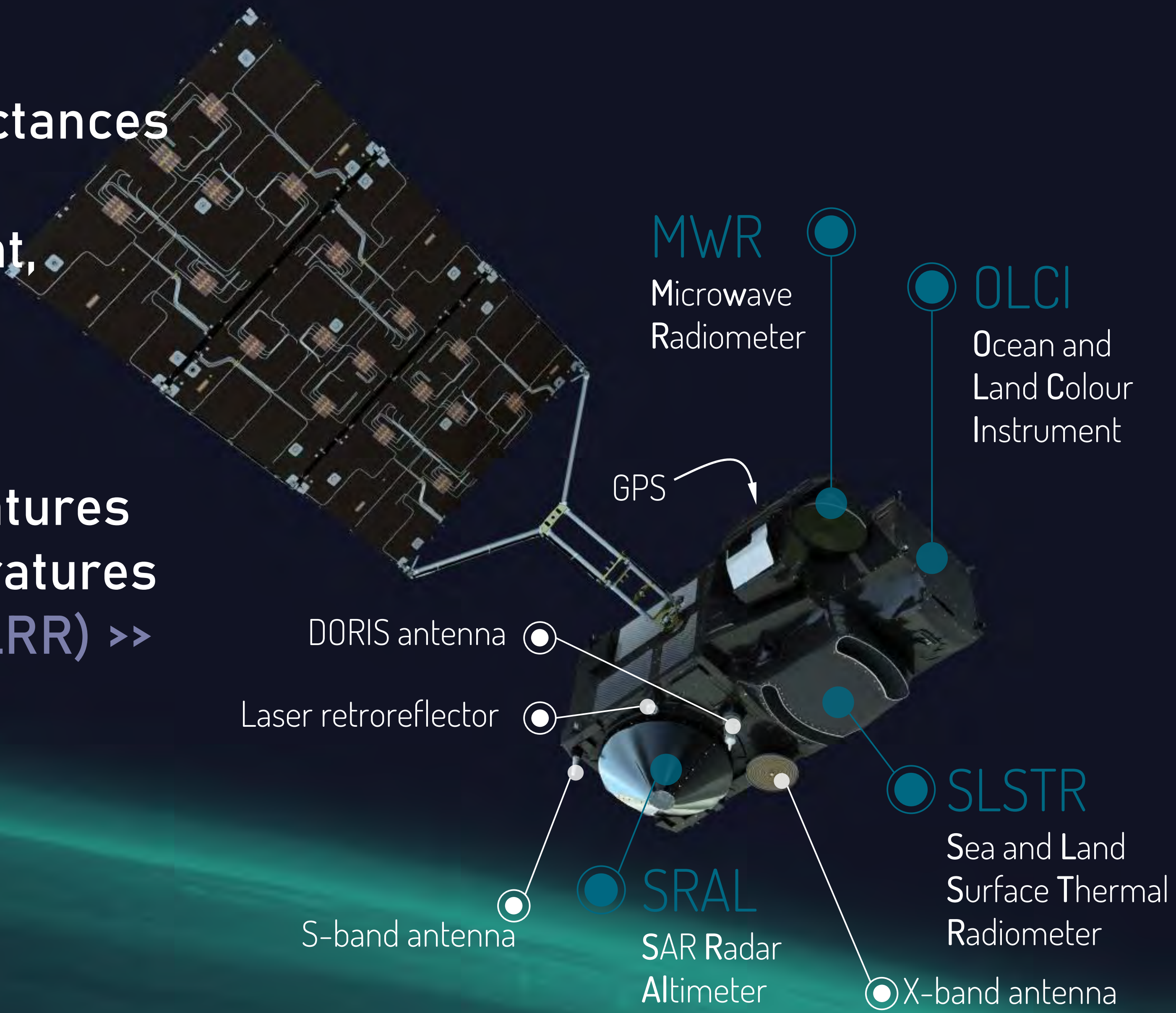


Marine missions: Sentinel-3

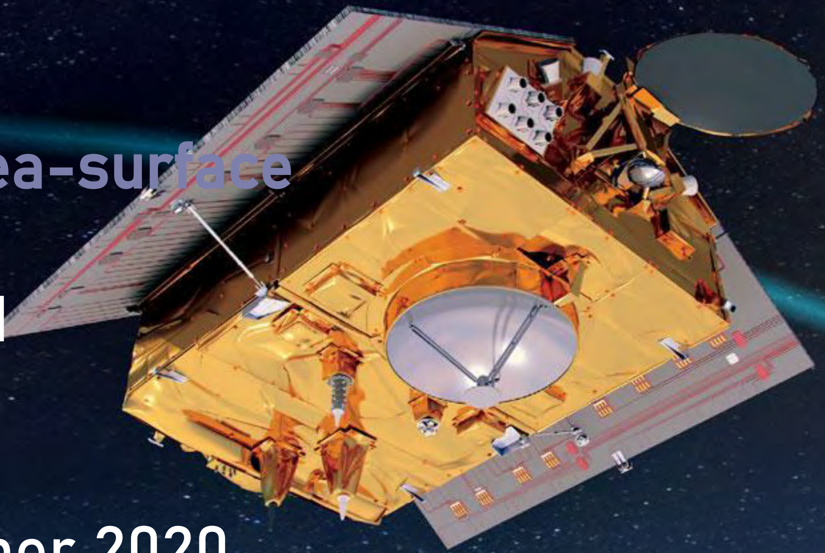
- The “blue” sentinel; suite of ocean observing instruments.
- Main objectives: acquire sea-surface topography, sea surface temperature and ocean colour data.
- Constellation of two platforms:
    - Sentinel-3A launched February 2016
    - Sentinel-3B launched April 2018
- Sun-synchronous 98.65º polar orbit, 27 day cycle.
- Near global coverage; <2 day revisit (optical) and <1 day (thermal).
- EUMETSAT operates the satellite & provides the marine data stream
    - Level-1 and level-2
    - Main user & provider of level-3, level-4 is Copernicus Marine Service
    - Redistributed (and used) by NOAA

Sentinel-3 instruments and variables

- OLCI >> visible radiometry
  - **ocean colour**: radiances & reflectances
    - 21 bands, 300m
  - chlorophyll, suspended sediment, CDOM
  - PAR / kd490
- SLSTR >> thermal radiometry
  - radiances & brightness temperatures
  - Sea and sea-ice surface temperatures
- SRAL / MWR / POD (DORIS/GNSS/LRR) >> surface topography mission
  - Sea surface height
  - Significant wave height
  - Wind speed



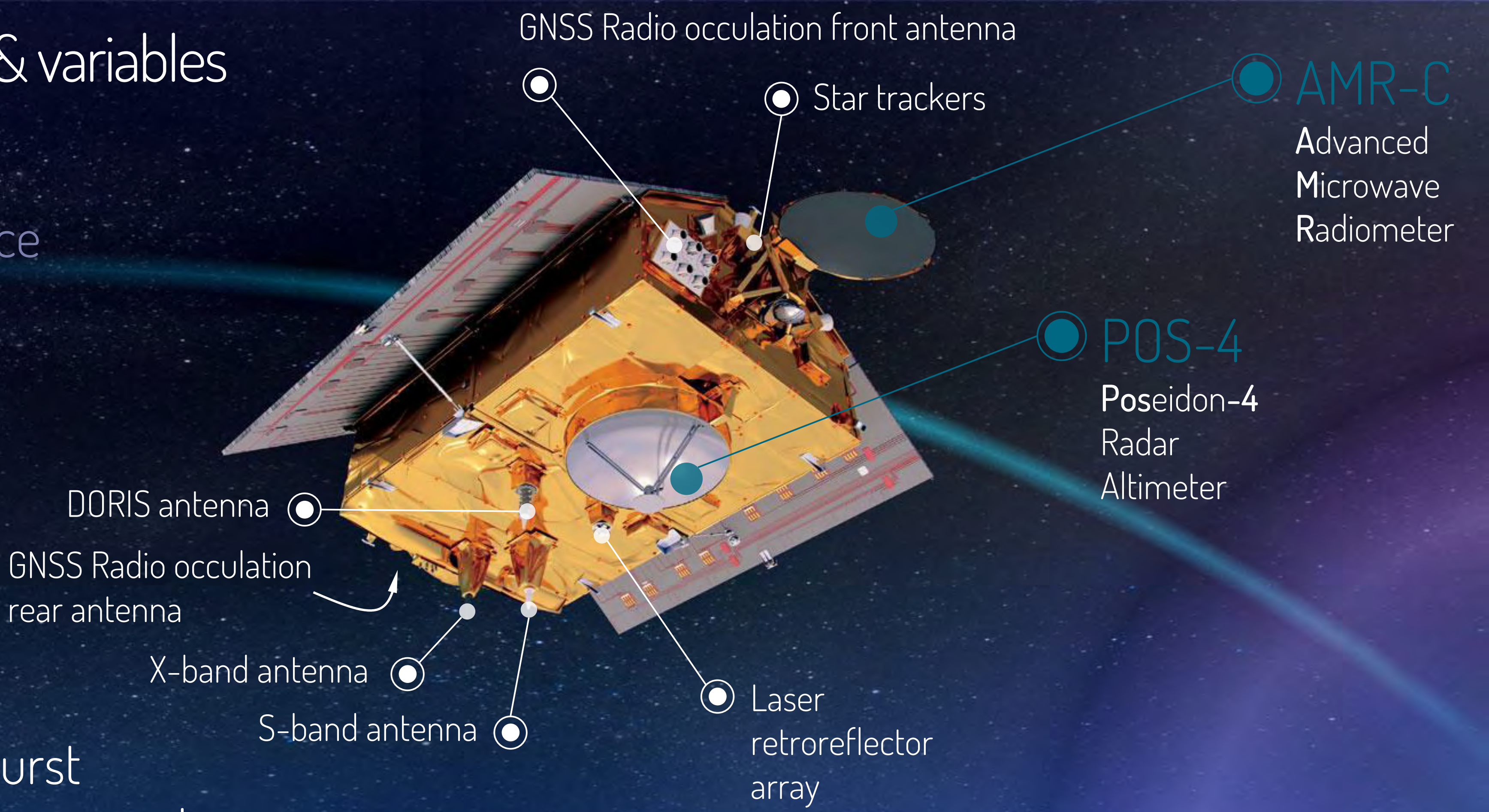
Marine missions: Sentinel-6



- The next **altimetry reference mission**
- Main objectives: high precision **sea-surface topography** measurements.
- Will continue the altimetry record into the next decade(s).
- Constellation of two platforms:
  - Sentinel-6A launched November 2020
  - Sentinel-6B launched planned 2025
- Non-sun sychronous 66º polar orbit, 10-day cycle.
- Collaborative mission: EUMETSAT / ESA / NASA / NOAA
- EUMETSAT operates the satellite & provides the data.
  - Level-1, level-2, level-2P/3 (through CNES)
  - Also available via PO.DAAC

Sentinel-6 instruments & variables

- POS-4 / AMR-C / POD (DORIS/GNSS/LRR) >> surface topography mission
  - Sea surface height
  - Significant wave height
  - Wind speed
- POS-4 :
  - back compatible (climate continuity)
  - state-of-the-art (open burst transmission, low noise, improved resolution)



More information: User Support and Training Resources

copernicus.eumetsat.int

EUMETSAT Helpdesk

OPS@eumetsat.int

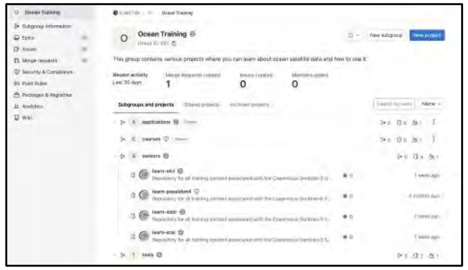
Contact the EUMETSAT helpdesk with any questions about EUMETSAT data products or services

User portal (coming soon)

Copernicus & mandatory missions
Data access

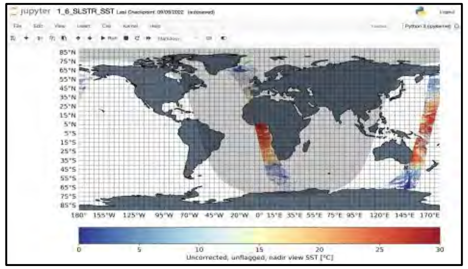
Code distribution

CONDA
GitLab



Jupyter Notebooks

jupyter



EUMETSAT User Support Resources

Cloud services

binder
WEkEO

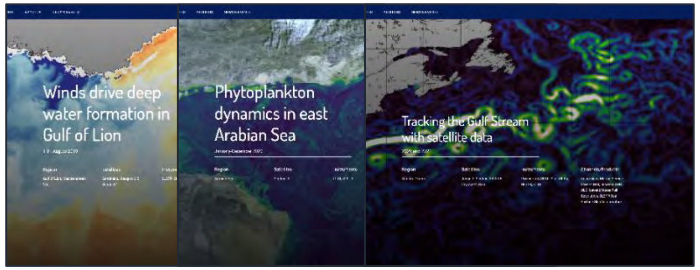
Courses

moodle

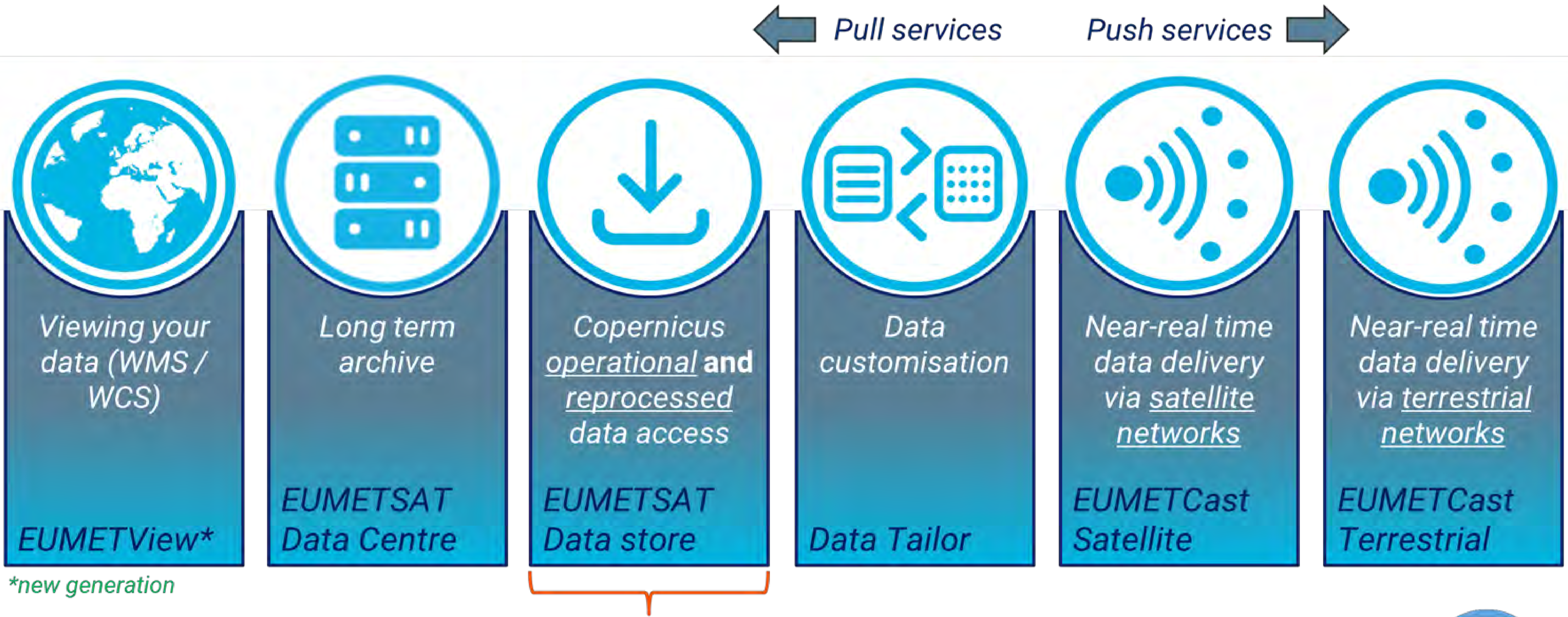
Video tutorials

YouTube

Case studies



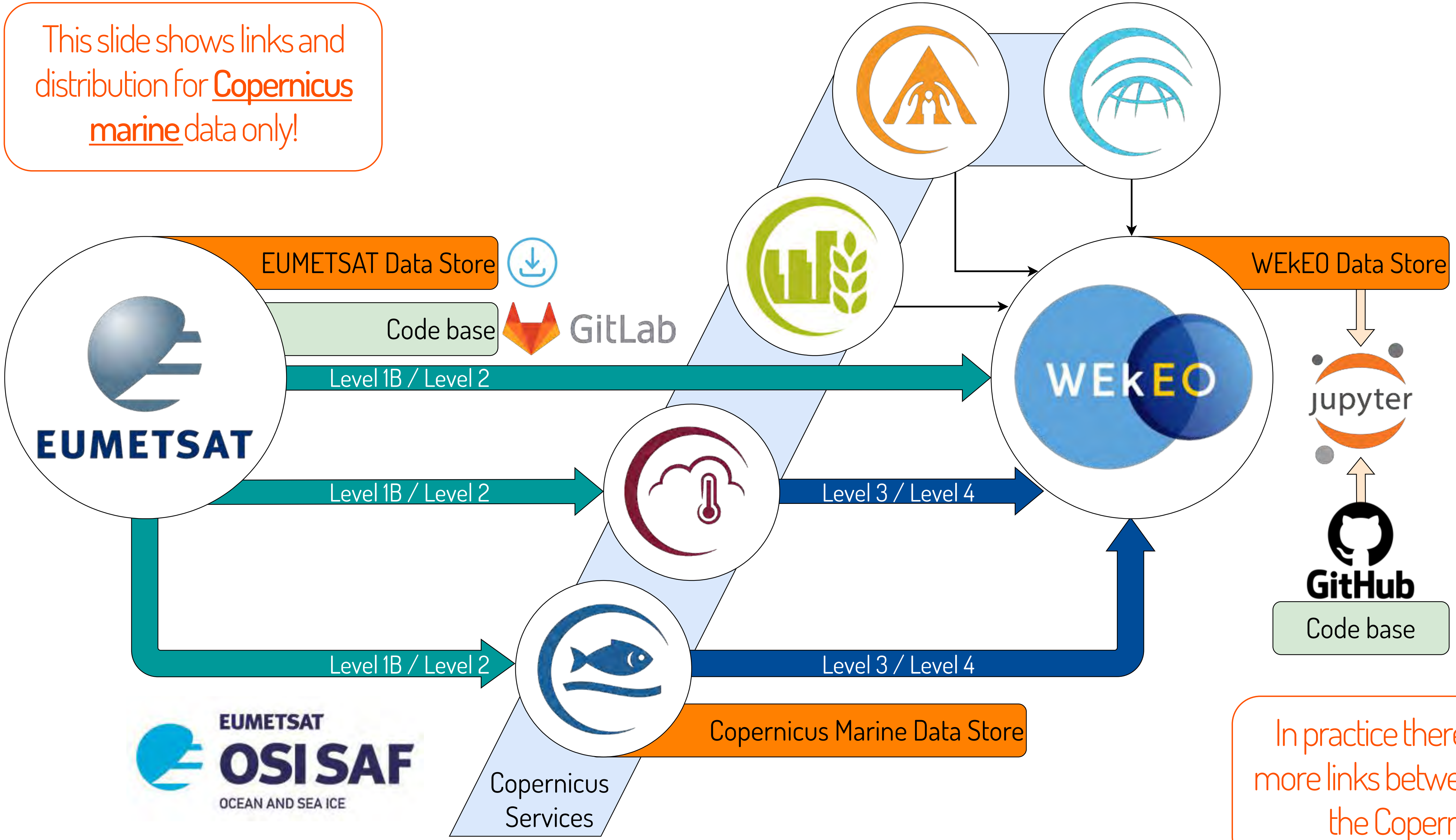
EUMETSAT Data Services



Data Store has replaced the CODA and CODAREP services used by many Sentinel-3 users, offering unified access to operational and reprocessed data. It will allow access long time-series of the most up to data products, via a single point without the use of Data Centre in most cases (including to WEkEO).

EUMETSAT Copernicus marine data distribution

This slide shows links and distribution for **Copernicus marine** data only!



In practice there are many, many more links between EUMETSAT and the Copernicus Services

The **EUMETSAT Data Store** provides single point of access to a growing catalogue of EUMETSATs meteorological, climate and ocean data.

- *All operational Sentinel-3 data can be accessed.*
- *Reprocessed data added as reprocessings complete*
- *For OLCI, a complete level-2 archive is already available.*
- *Feeds in to WEkEO harmonised data access*
- *Sentinel-6 reprocessing available, operational feed coming soon.*

Data collections

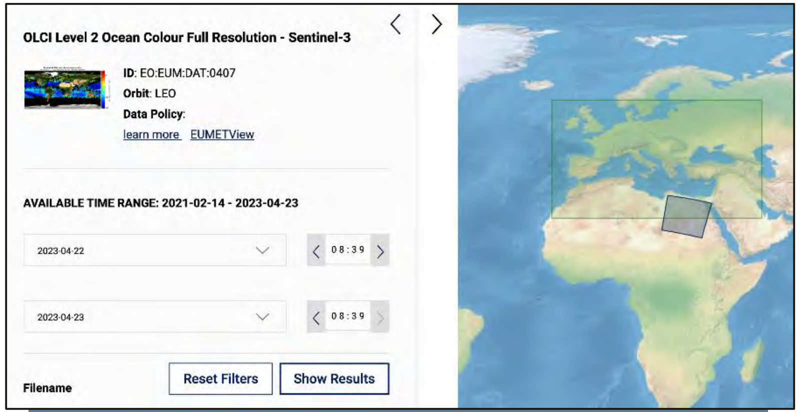
- Near real-time products
- Historic & reprocessed products
- Climate data records

EUMETSAT Data Access Client (eumdac)

Service available at: https://data.eumetsat.int

Data Store Interfaces

Online web user interface



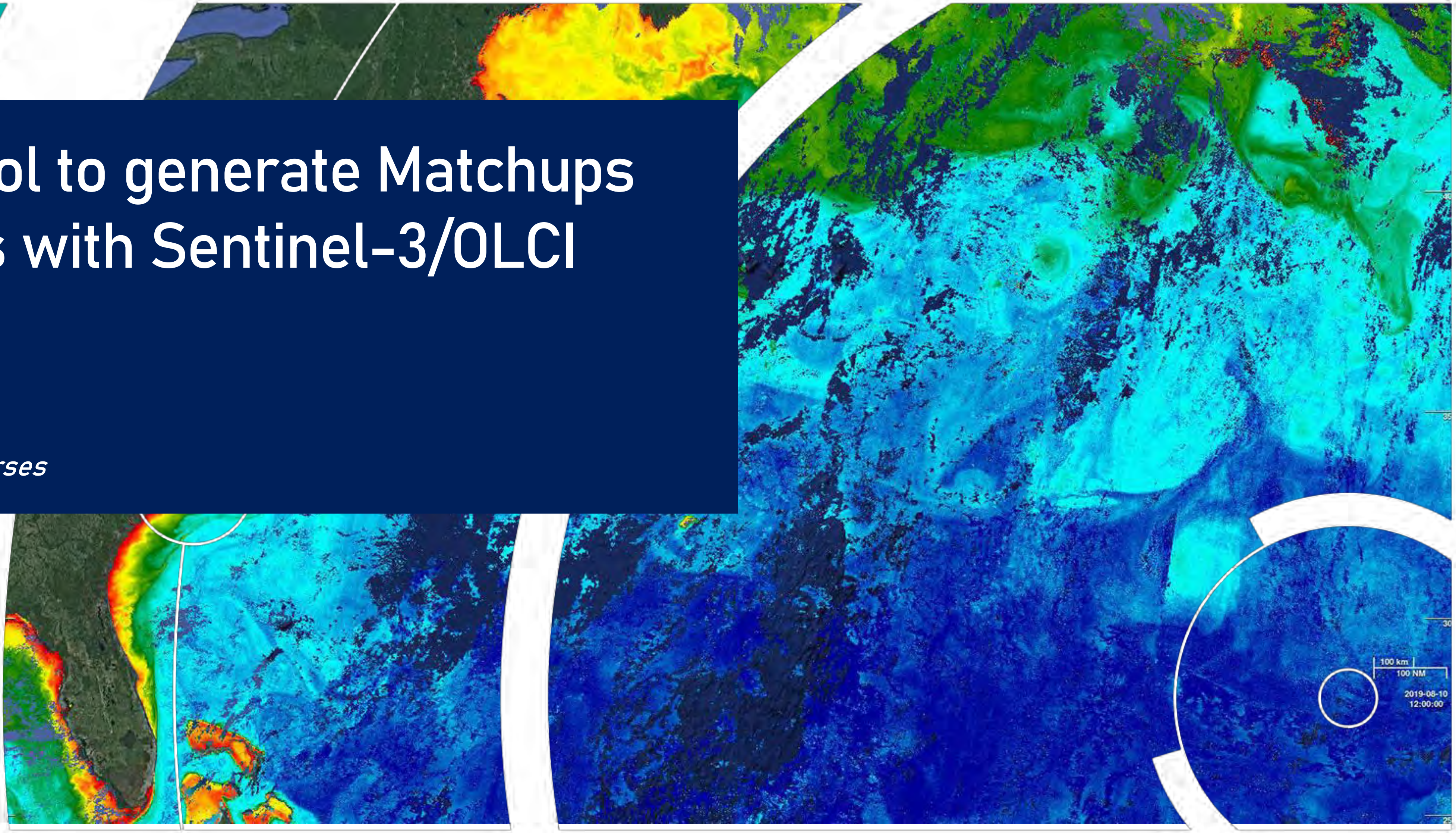
1. Browse API
2. OpenSearch API
3. Download API
4. Subscription API

Advanced data access with EUMDAC

- We offer a series of examples showing how to use the APIs of the EUMETSAT Data Store via EUMDAC.
  - Command line (see user guides)
  - Python library – supporting Jupyter notebooks available.

- Within the learn-olci repository, folder 2_OLCI_advanced
  - Advanced search and filtering options
  - Navigating operational and reprocessed collections to acquire time series

- EUMDAC embedded in ThoMaS toolkit



ThoMaS - a Tool to generate Matchups of OC products with Sentinel-3/OLCI

Juan Ignacio Gossn
EUMETSAT

EUMETSAT series of short courses

Today's tour with ThoMaS – a Tool to perform Matchups with S3-OLCI

1. What's ThoMaS? Scope
2. Usage
3. Some background
4. Pre-requisites
5. Getting the code
6. Setting the environment
7. Required inputs
8. Run the code
9. Set-up demo
10. Run the code: demo
11. Short tour around ThoMaS

1. What's ThoMaS? Scope

ThoMaS is a toolkit developed to create **matchups** of bio-geophysical **insitu data** with **satellite ocean colour products** from **Sentinel-3 OLCI** (S3/OLCI).

in SeaBASS format

Standard products from NASA's OBPG also supported
Others easily configurable, if netCDF or series of netCDFs

1. What's ThoMaS? Scope

After running **ThoMaS**, you will get:

→ **Insitu** data "transformed" to **match satellite** (spectral **convolution, band-shifting, BRDF...**).

→ **Satellite** data (L1B – TOA radiance - or L2 – BOA water reflectance) from **EUMETSAT Data Store** (reprocessed/operational) matching spatially/temporally your insitu.

→ **Extractions** of satellite data centred at lat/lon of insitu of user-defined size (3x3, 5x5..).

→ **Statistics of extractions** following EUMETSAT's or any user-defined **matchup protocol**.

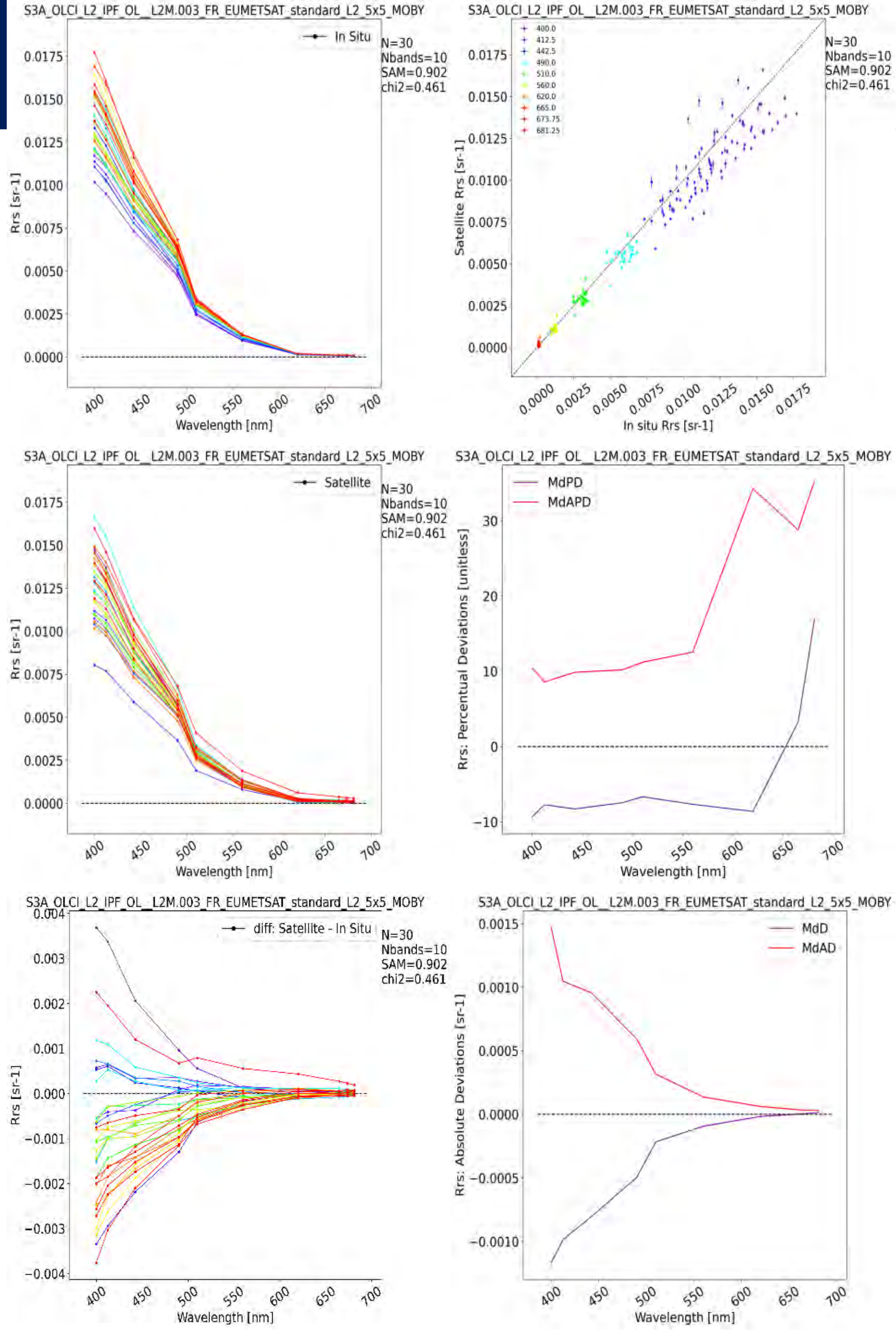
→ **Merging of simultaneous** (spatially-temporally) **insitu-satellite pairs, temporal interpolation**, and **statistics of matchups**.

→ **Outputs:**

→ **NetCDF 4** files: SatData, minifiles, Extraction Data Base files, In situ Data Base file, Matchup Data Base files.

→ **CSV**: summarizing satellite extraction statistics and matchup statistics.

→ **PNG**: Standardised output plots.



2. Usage

ThoMaS workflow is divided into **5** main steps:

The steps can be executed **sequentially** or **independently** in case the needed outputs of the previous steps are available.

1. Step **insitu**
    1. Ingest insitu data from **SeaBASS** input file
    2. Apply several transformations to make **insitu comparable to satellite** data (e.g. **spectral matching** with satellite, **BRDF** correction)
    3. Store them into standard **IDB** (**In situ Data Base**) netCDF4 file.

    → This step can optionally include the acquisition of **ancillary information** from **ECMWF** at the lat-lon-times of your insitu measurements.

2. Step **SatData**: Download and list the **satellite products (L1B and/or L2)** matching spatially-temporally your insitu data.

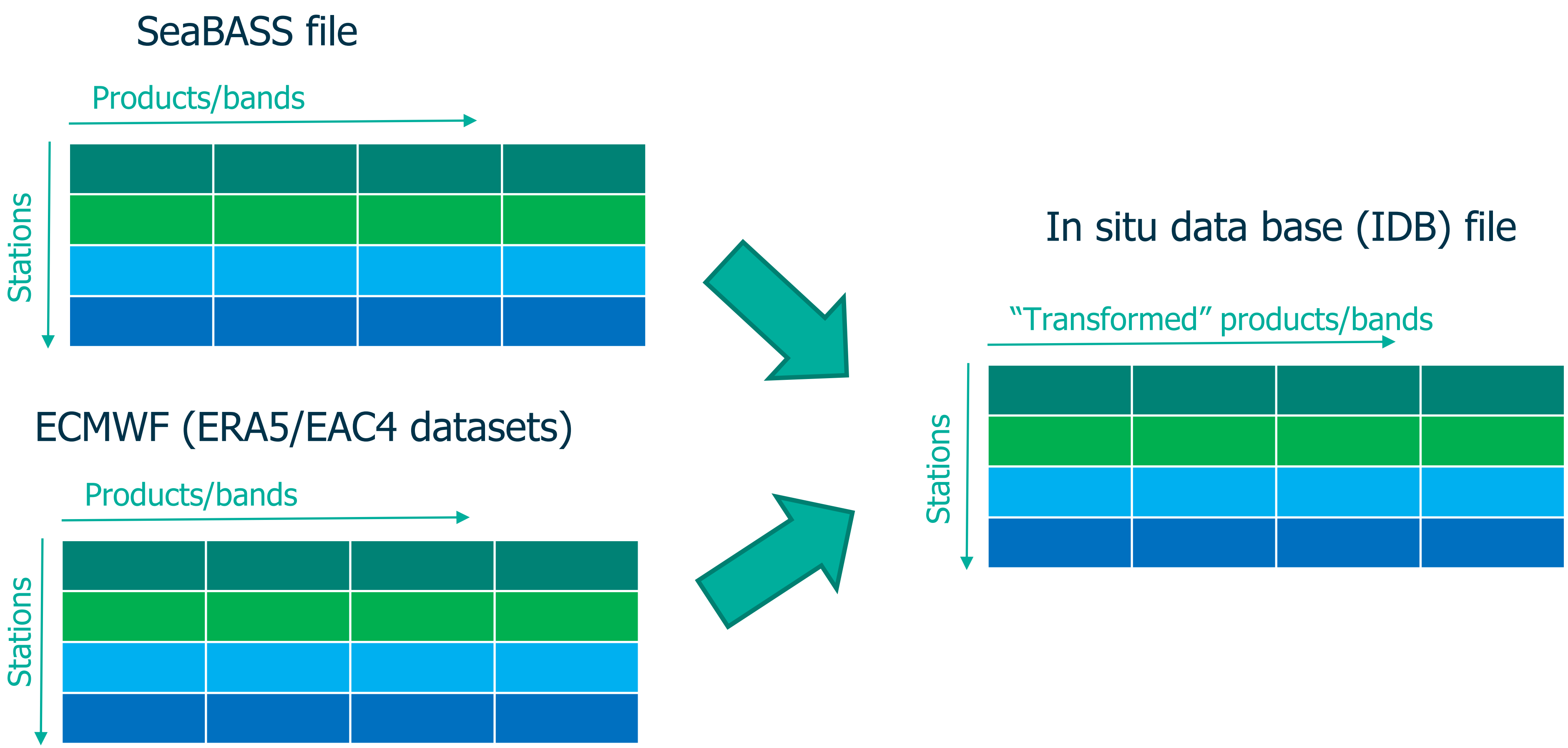
    → Download only for products available in EUMETSAT data store.

3. Step **minifiles**: SatData are grouped/unnested into single netCDF4 file, sliced in horizontal dimensions, centred at the desired (in situ) location.

4. Step **EDB**.
    1. Stack minifiles into single netCDF
    2. Apply transformations to SatData to make them comparable to in situ (e.g. **scale/unit conversion**, **BRDF** correction)
    3. Calculate extraction **statistics** over the extraction window following EUMETSAT's or any user-defined Matchup Protocol.
    4. Store into standard **EDB** (**Extraction Data Base**) netCDF4 and CSV files.

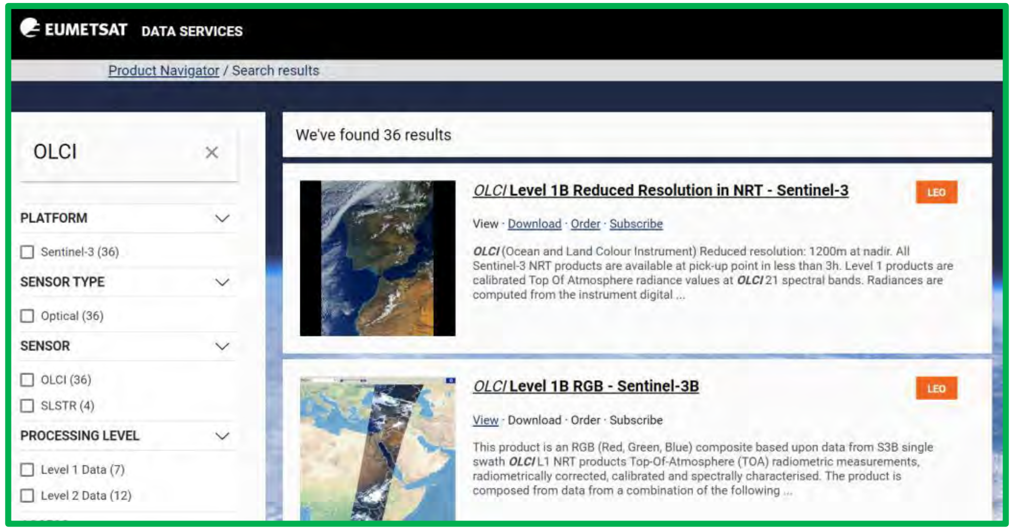
5. Step **MDB**.
    1. Combine insitu (**IDB**) and satellite (**EDB**) information indexed into insitu-satellite **matchup pairs**
    2. **Optionally apply time interpolation**
    3. Calculate **matchup statistics**
    4. Store into standard **MDB** (**Matchup Data Base**) netCDF4 and CSV files.

2. Usage. Step insitu



2. Usage. Step SatData

EUMETSAT Data Store

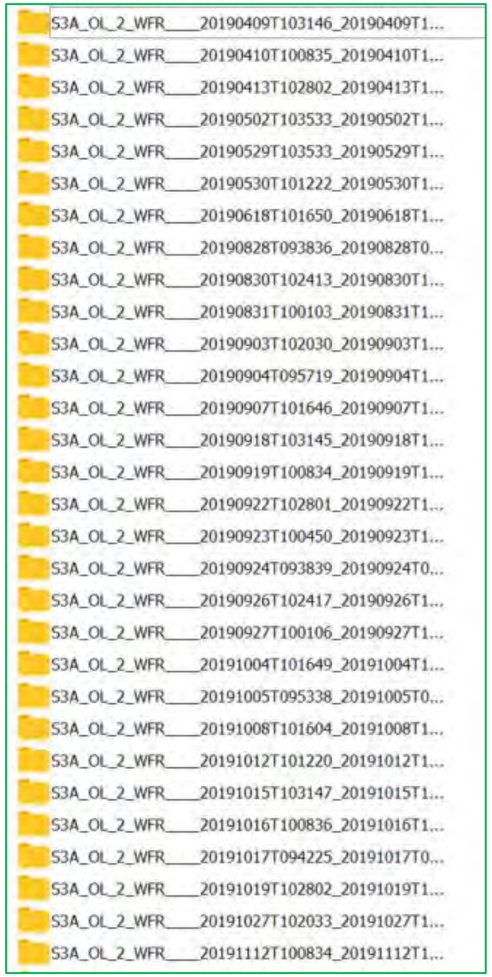


NASA's OBPG L2 standard products

...

...

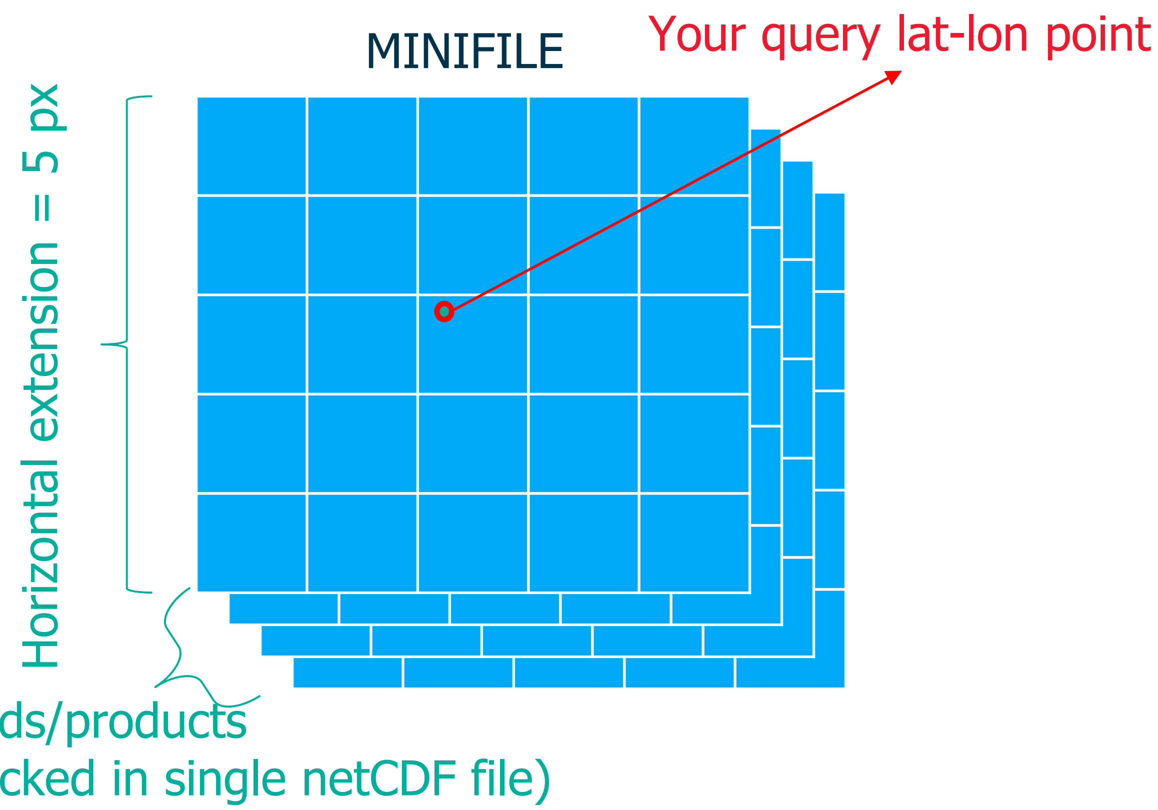
Your local system



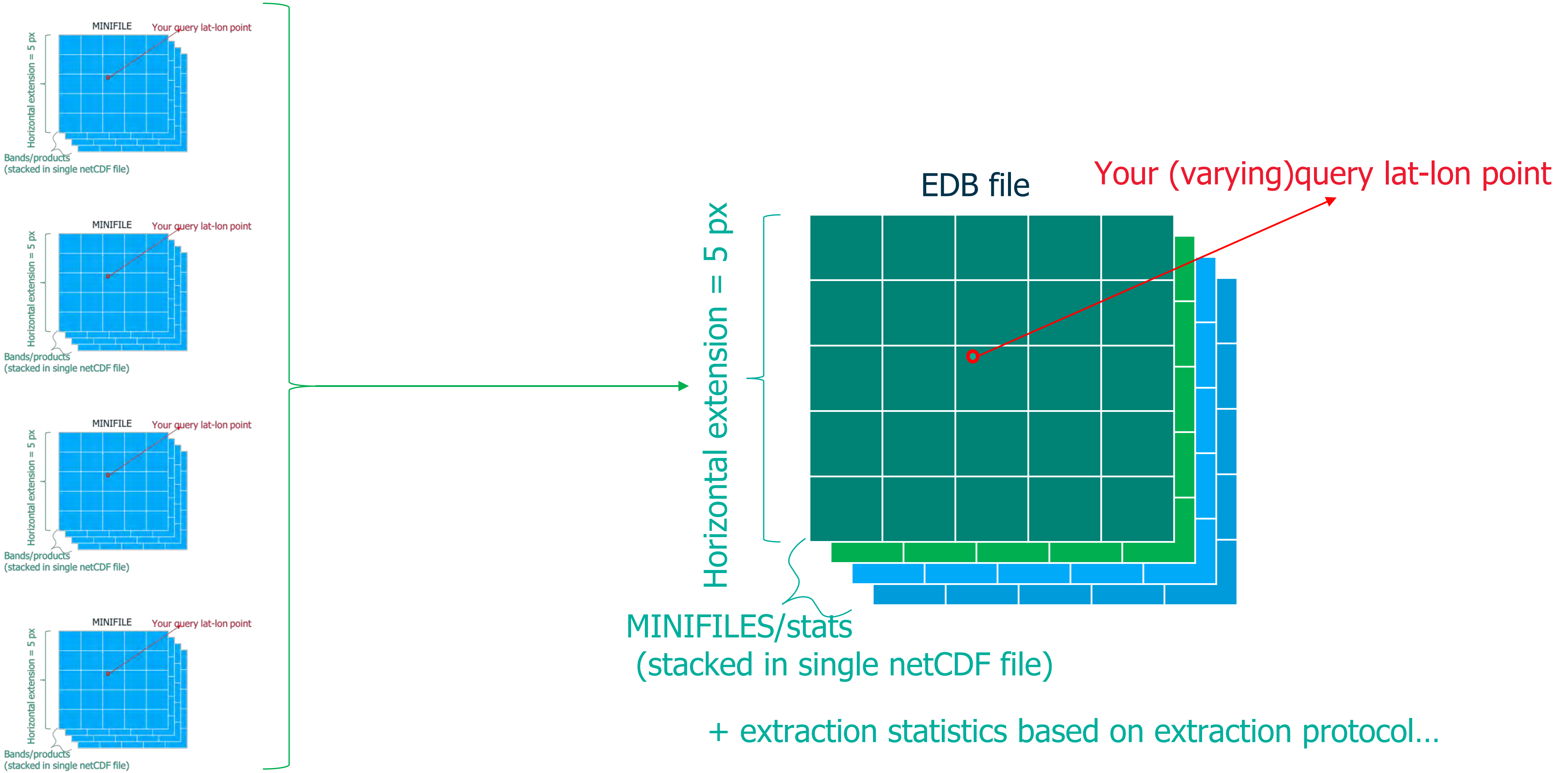
+ SatData Lists matching in situ

2. Usage. Step minifiles

- S3A_OL_2_WFR____20190409T103146_20190409T1...
- S3A_OL_2_WFR____20190410T100835_20190410T1...
- S3A_OL_2_WFR____20190413T102802_20190413T1...
- S3A_OL_2_WFR____20190502T103533_20190502T1...
- S3A_OL_2_WFR____20190529T103533_20190529T1...
- S3A_OL_2_WFR____20190530T101222_20190530T1...
- S3A_OL_2_WFR____20190618T101650_20190618T1...
- S3A_OL_2_WFR____20190828T093836_20190828T0...
- S3A_OL_2_WFR____20190830T102413_20190830T1...
- S3A_OL_2_WFR____20190831T100103_20190831T1...
- S3A_OL_2_WFR____20190903T102030_20190903T1...
- S3A_OL_2_WFR____20190904T095719_20190904T1...
- S3A_OL_2_WFR____20190907T101646_20190907T1...
- S3A_OL_2_WFR____20190918T103145_20190918T1...
- S3A_OL_2_WFR____20190919T100834_20190919T1...
- S3A_OL_2_WFR____20190922T102801_20190922T1...
- S3A_OL_2_WFR____20190923T100450_20190923T1...
- S3A_OL_2_WFR____20190924T093839_20190924T0...
- S3A_OL_2_WFR____20190926T102417_20190926T1...
- S3A_OL_2_WFR____20190927T100106_20190927T1...
- S3A_OL_2_WFR____20191004T101649_20191004T1...
- S3A_OL_2_WFR____20191005T095338_20191005T0...
- S3A_OL_2_WFR____20191008T101604_20191008T1...
- S3A_OL_2_WFR____20191012T101220_20191012T1...
- S3A_OL_2_WFR____20191015T103147_20191015T1...
- S3A_OL_2_WFR____20191016T100836_20191016T1...
- S3A_OL_2_WFR____20191017T094225_20191017T0...
- S3A_OL_2_WFR____20191019T102802_20191019T1...
- S3A_OL_2_WFR____20191027T102033_20191027T1...
- S3A_OL_2_WFR____20191112T100834_20191112T1...



2. Usage. Step EDB



+ extraction statistics based on extraction protocol…

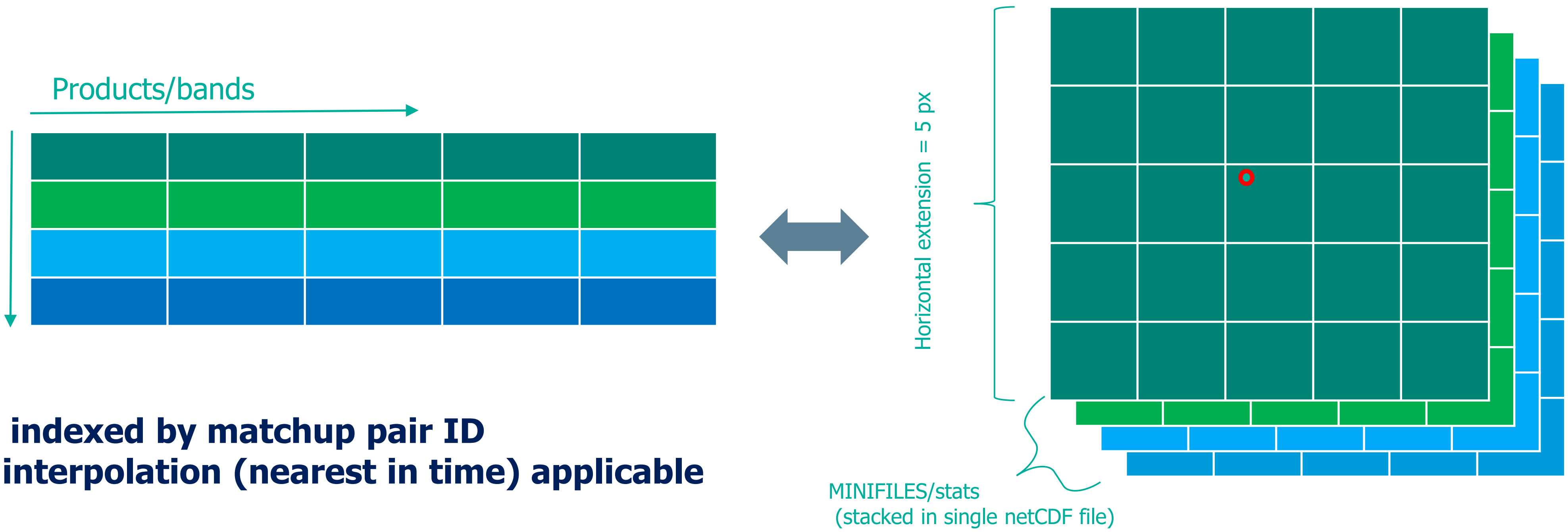
2. Usage. Step MDB

MDB file

1. Merging insitu and extractions according to matchup pairs
2. Statistical metrics calculated + scatter/spectral plots

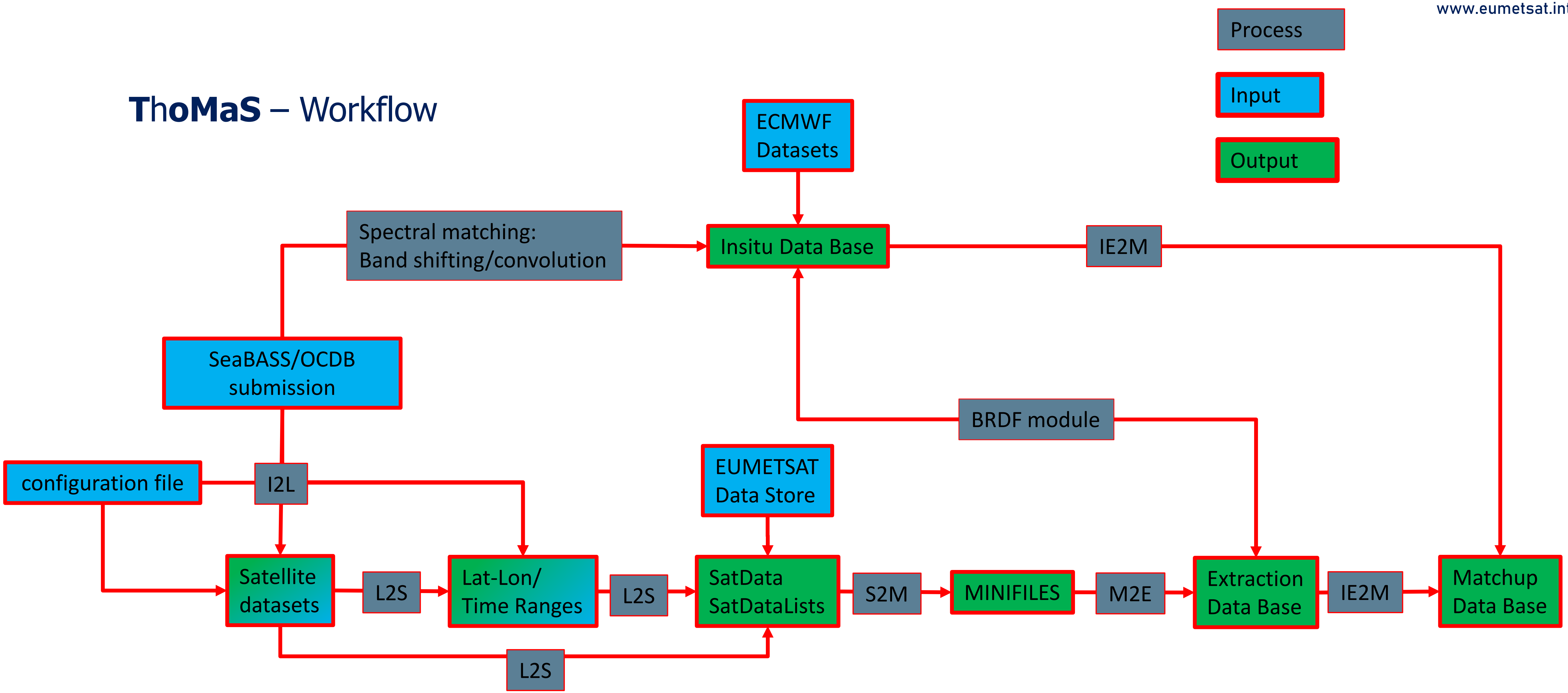
In situ data base (IDB) file

Extraction data base (EDB) file



MDB: indexed by matchup pair ID
Time interpolation (nearest in time) applicable

ThoMaS – Workflow



2. Usage: Summary on the terminology

- **IDB** (In situ Data Base): a netCDF file containing all the information related to the inputted insitu data. + (if requested) ancillary information from **ECMWF reanalysis datasets** at the insitu geographic location and time stamp.

- **SatData**: an image file/directory. In the case of standard L1/L2 OLCI products, it is composed of a directory containing several netCDF files, each containing one/several products + a manifest.xml file.

- **Minifile**: A single netCDF file containing all the relevant L1/L2 products from a single SatData, but only at the required location (and with a predefined window size).

- **EDB** (Extraction Data Base): All the statistical information (pixel-by-pixel flagging, outlier removal, central and dispersion values before/after outlier/mask removal, etc., details of the extraction protocol) is stored for all the extractions in one single netCDF file per extraction set.

- **MDB** (Match-up Data Base): All the information from IDB and EDB combined and re-indexed according to matchup pairs + matchup statistics.

Find **examples** of all these files (except SatData) in the **examples/example_files** directory.

3. Some background: match-ups

- What is a match-up according to **chatGPT?**

j — Hi ChatGPT

Hello! How can I assist you today?

j — Could you define me "match-up" in the realm of Earth Observation?

In the context of Earth Observation, "match-up" refers to the process of comparing satellite data to ground-based measurements or other reference data sources to assess the accuracy and quality of the satellite observations.

Match-up analysis is typically used to validate satellite measurements of various geophysical parameters, such as surface temperature, vegetation cover, sea surface temperature, and atmospheric aerosols, among others. The ground-based measurements can be obtained from various sources, such as ground-based sensors, buoys, weather stations, or other types of in-situ measurements.

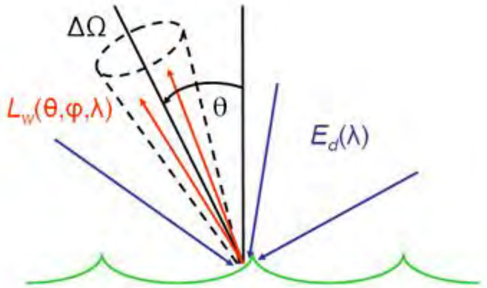
→ Ocean colour 👀 ←

Of course we have much more to define… and take care of…

3. Some background: match-ups

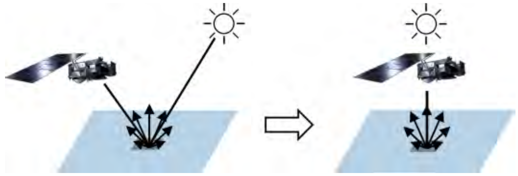
Definition of Rrs

$$R_{rs}(\theta, \phi, \lambda) \equiv \frac{L_w(\text{in air}, \theta, \phi, \lambda)}{E_d(\text{in air}, \lambda)} \quad (\text{sr}^{-1}).$$



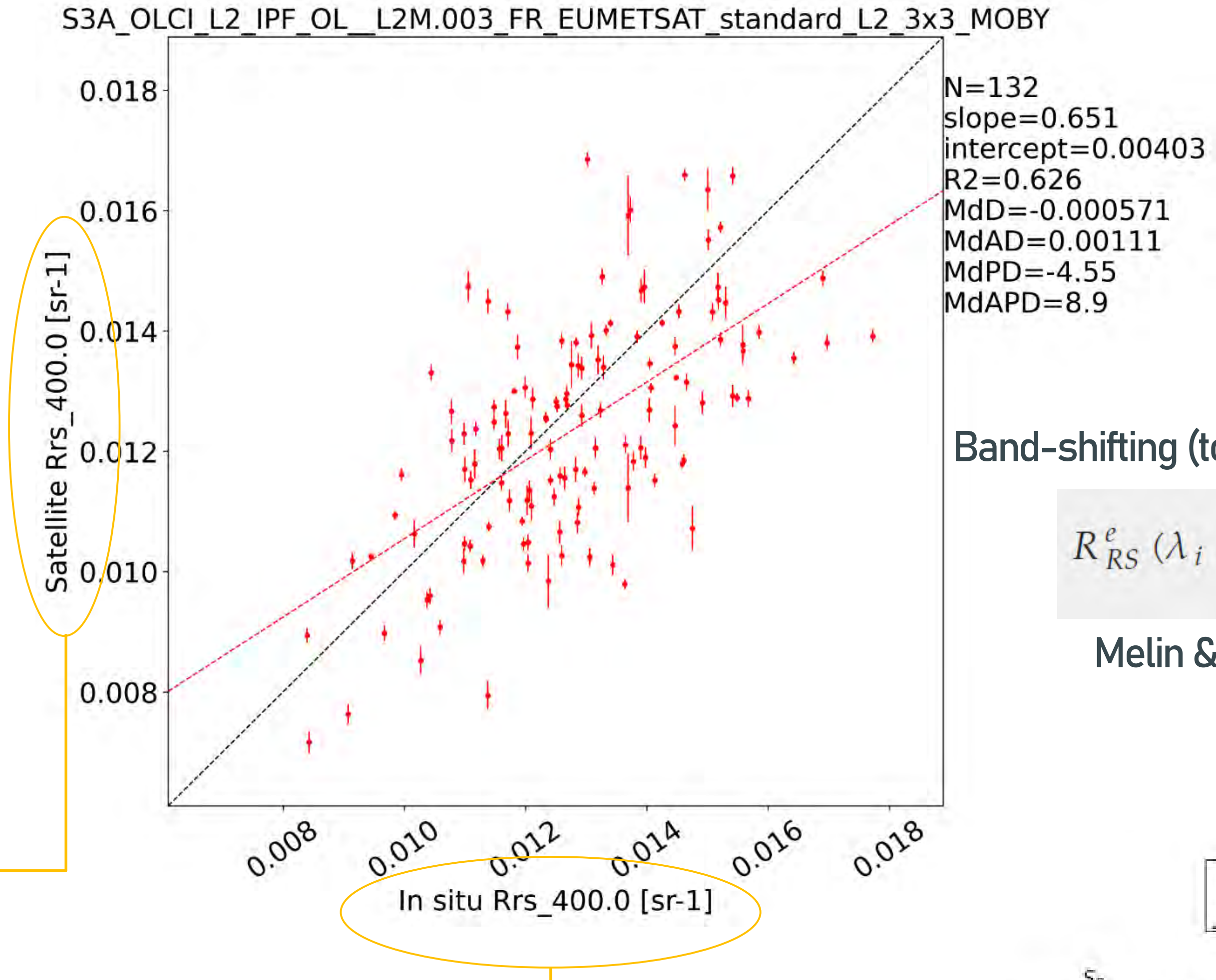
OO Web Book, Mobley, Boss & Roesler

BRDF correction:



D'Alimonte et al.

Morel et al. 2002 supported in ThoMaS

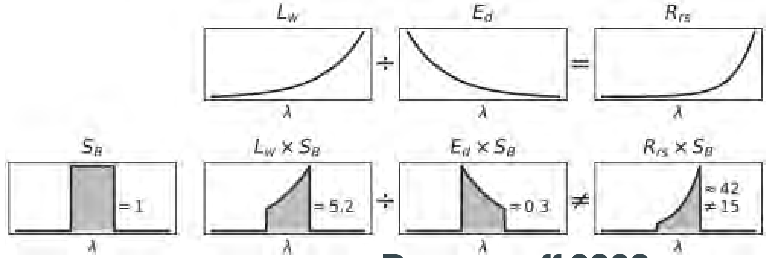


Band-shifting (to pair multispectral to multispectral)

$$R^{e}_{RS}(\lambda_i \rightarrow \lambda_t) = R^{f}_{RS}(\lambda_t) \frac{R_{RS}(\lambda_i)}{R^{f}_{RS}(\lambda_i)}$$

Melin & Sclep 2015 supported in ThoMaS

Spectral convolution



Burggraaff 2020

<u>**Are we comparing 🍏 with 🍏?**</u>

→ What is the <u>definition</u> of <u>Rrs</u>?

→ Are these two compatible "spectrally"? → convolution/band-shifting

→ Are these two compatible "directionally"? → BRDF correction

3. Some background: match-ups

www.eumetsat.int

Quality of insitu

ThoMaS still **<u>does not</u>** consider any quality flag to process insitu data… but stay tunned ☺…

fiducial reference measurements for satellite ocean colour


IOCCG Protocol Series

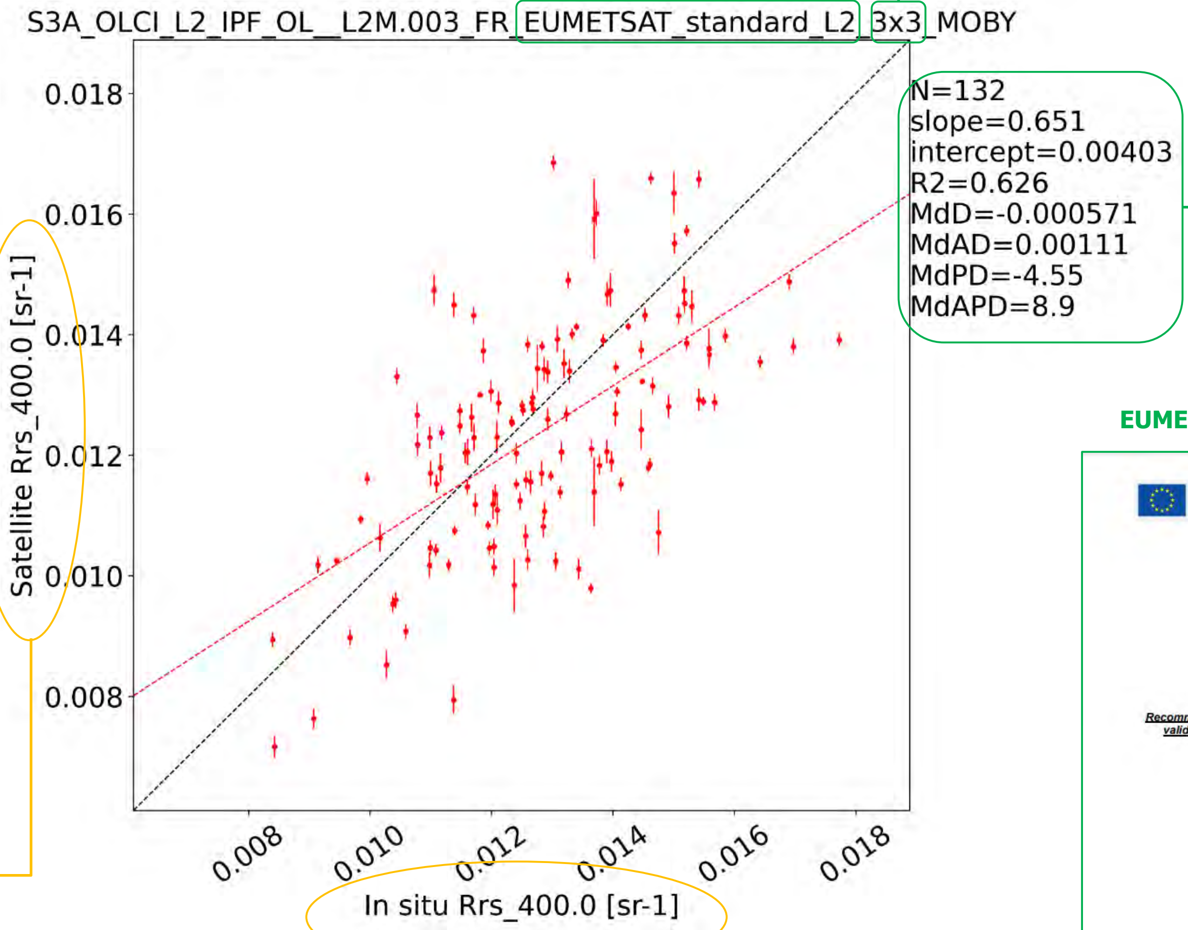
Ocean Optics & Biogeochemistry Protocols for Satellite Ocean Colour Sensor Validation

Volume 3: Protocols for Satellite Ocean Colour Data Validation: In Situ Optical Radiometry (v3.0)

December 2019

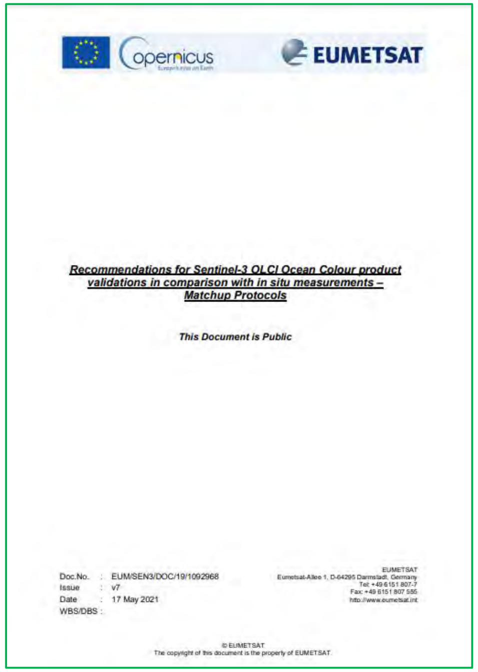
Define your extraction statistics!

Define your extraction window size!



Define the matchup statistics!

EUMETSAT's Matchup Protocols


Recommendations for Sentinel-3 OLCI Ocean Colour product validations in comparison with in situ measurements – Matchup Protocols

This Document is Public

Doc.No. : EUM/SEN3/DOC/19/1092968
Issue : v7
Date : 17 May 2021

© EUMETSAT
The copyright of this document is the property of EUMETSAT

<u>Are we comparing 🍏 with 🍏?</u>

→ Are the insitu measurements of sufficient quality?

→ Are insitu and satellite measurements temporally-spatially comparable?

→ What value (and uncertainty) shall I extract from the satellite data?

3. Motivation for having ThoMaS

1. Well documented, suited for a first approach to the matchup exercise for those who are new to the intricacies of the matchup steps.
2. Versatile: new types of satellite products can be easily added via configuration files (depending on mission, processor and processing baseline).

   → Currently supports Sentinel-3 (standard) L1B, L2, MODIS L2 (standard), VIIRS L2 (standard).

3. Versatile: new matchup protocols can be easily added via configuration files.

   → e.g. of existing ones: Bailey & Werdell 2006, EUMETSAT's standard, Copernicus SVC_VIS

4. It's published and open to scrutiny: it serves for the purpose of converging to a standard matchup practice.

EUMETSAT's Matchup Protocols: extraction of statistics at macropixel level

Window size recommended: 5x5 or 3x3
Tolerable insitu-satellite time difference: 1 hr or 3 hrs

Set BFOR: 5x5 window

Detection of non-valid pixels (flagged pixels)

Pixels are masked/removed if flagged by any of the following:

CLOUD, CLOUD_AMBIGUOUS, CLOUD_MARGIN, INVALID, COSMETIC, SATURATED, SUSPECT, HISOLZEN, HIGHGLINT, SNOW_ICE, AC_FAIL, WHITECAPS, ADJAC, RWNEG_O2, RWNEG_O3, RWNEG_O4, RWNEG_O5, RWNEG_O6, RWNEG_O7, RWNEG_O8

Macropixel is discarded if:

$N_{BOR} < 50\% \ N_{BFOR}$

Set BOR: without flagged pixels

Pixel 'X' is considered outlier if:

$|value@X - \mu_{BOR}| < 1.5\sigma_{BOR}$

Macropixel is discarded if:

$CV_{final}(560) > 20\%$

Set final: without flagged & outlier

Central value: $median_{final}$
Uncertainty measure (Type B): σ_{final}
Homogeneity measure: CV_{final}

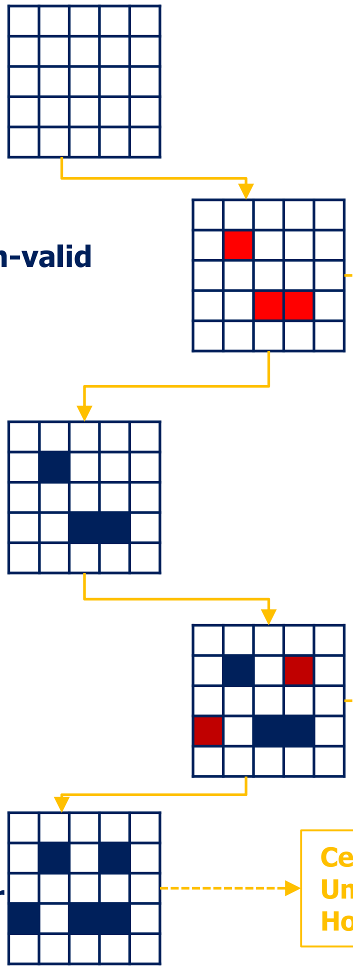
ThoMaS can be run with many other extraction protocols, that you define

μ → Mean
σ → Standard deviation

3. Some background: match-ups: Bailey & Werdell protocol

EUMETSAT's Matchup Protocols: extraction of statistics at macropixel level

Window size recommended: 5x5 or 3x3
Tolerable insitu-satellite time difference: 1 hr or 3 hrs



Set BFOR: 5x5 window

Detection of non-valid pixels (flagged pixels)

Pixels are masked/removed if flagged by any of the following:

ATMFAIL, LAND, HIGLINT, HILT, HISATZEN, STRAYLIGHT, CLDICE, COCCOLITH, HISOLZEN, LOWLW, CHLFAIL, NAVWARN, MAXAERITER, CHLWARN, ATMWARN, SEAICE, NAVFAIL, ABSAER, MODGLINT

Macropixel is discarded if:

$N_{BOR} < 50\% \ N_{BFOR}$

Set BOR: without flagged pixels

Pixel 'X' is considered outlier if:

$|value@X - \mu_{BOR}| < 1.5\sigma_{BOR}$

Macropixel is discarded if:

Median[CV(Rrs(410-551));CV(AOT(869))]>15%

Set final: without flagged & outlier

Central value: $median_{final}$
Uncertainty measure (Type B): σ_{final}
Homogeneity measure: CV_{final}

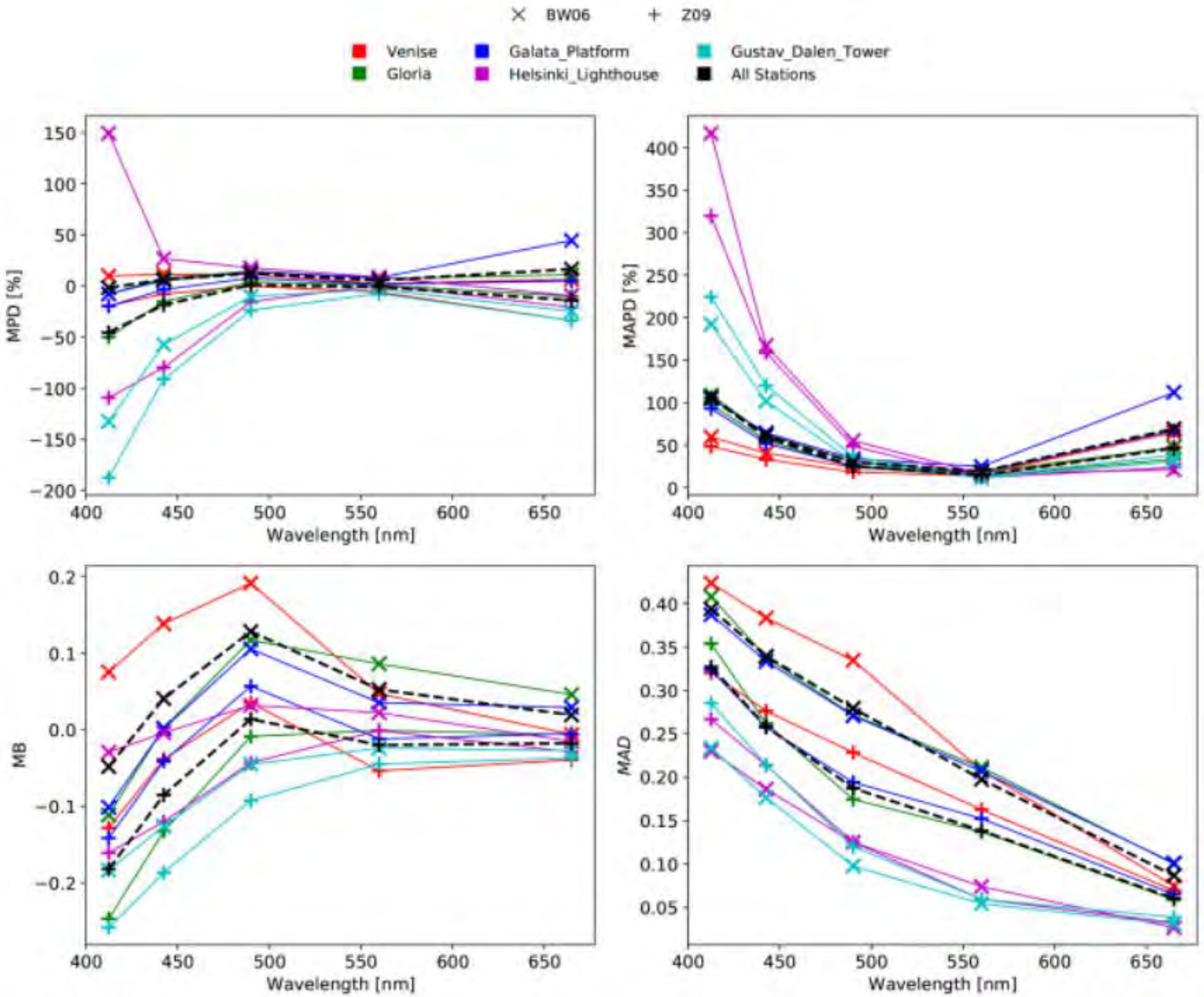
ThoMaS can be run with many other extraction protocols, that you define

μ → Mean
σ → Standard deviation

3. Some background: match-ups: what protocol to use?

www.eumetsat.int

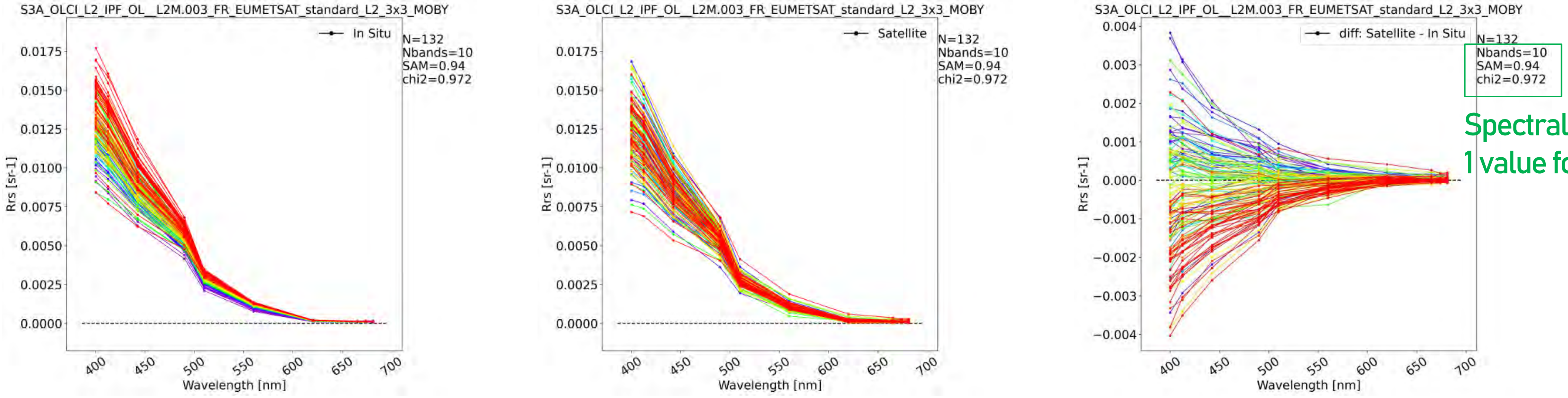
Should we care about wh



[Concha et al. 2021]

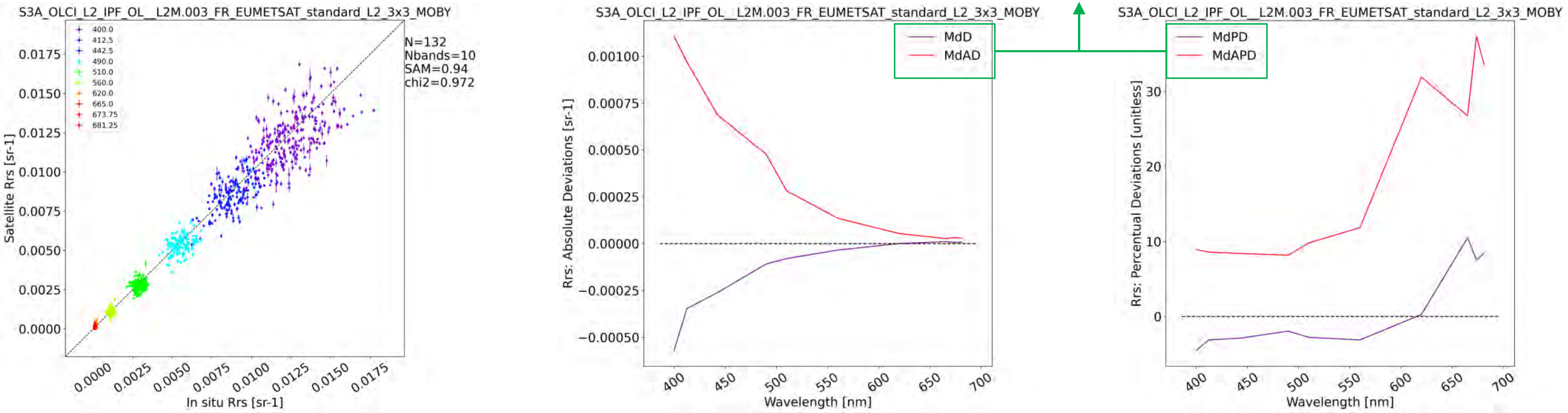
3. Some background: match-ups

Band-by-band plots and statistics are often not sufficient…



Spectral statistics
1 value for the whole set

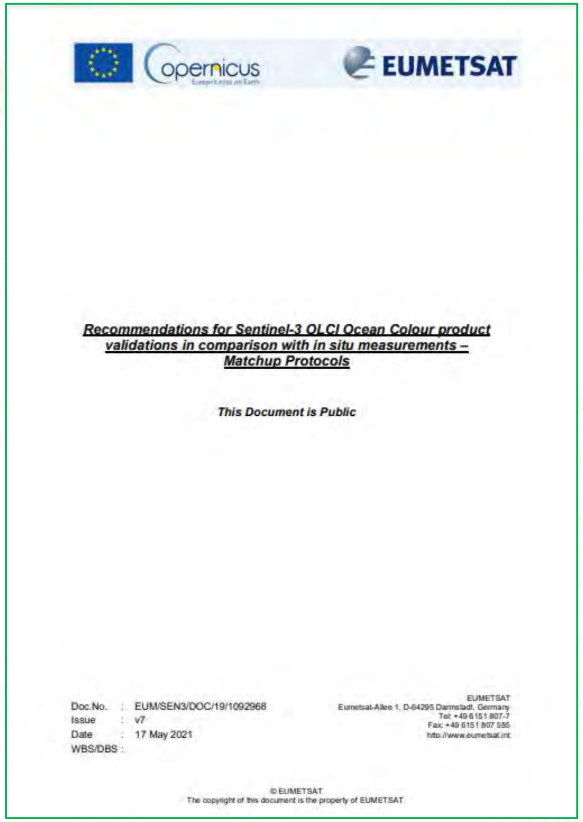
statistics band-by-band, plotted spectrally



3. Some background: match-ups

Band-by-band plots and statistics are often not sufficient...

Check recommended statistics and definitions at:
EUMETSAT's Matchup Protocols



Copernicus EUMETSAT

Recommendations for Sentinel-3 OLCI Ocean Colour product validations in comparison with in situ measurements – Matchup Protocols

This Document is Public

Doc.No. : EUM/SEN3/DOC/19/1092968
Issue : v7
Date : 17 May 2021
WBS/DBS :

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

Spectral statistics
1 value for the whole set

statistics band-by-band, plotted spectrally

ThoMaS: user can define it's own
- extraction statistics method
- window size
- time tolerance
- relevant statistics

3. Pre-requisites

1. Apart from that background knowledge…
2. Conda: Install the latest Anaconda Python distribution.
3. EUMETSAT Data Store: Create EO Portal user and get API consumer key and secret.
4. EUMETSAT Data Store: Save EO Portal API credentials under **~/.eumdac/credentials.txt**
5. ECMWF: Register to ADS/CDS and get url and key.
6. ECMWF: store ADS/CDS url/keys under **~/.ecmwf_api_config**

Dependencies

item	version	licence	package info
BeautifulSoup	4.6.0	MIT	https://anaconda.org/conda-forge/beautifulsoup4
cdsapi	0.1.6	Apache-2.0	https://anaconda.org/conda-forge/cdsapi
ephem	4.1.3	MIT	https://pypi.org/project/ephem/
eumdac	2.0.1	MIT	https://anaconda.org/eumetsat/eumdac
matplotlib	3.5.2	PSF-based	https://anaconda.org/conda-forge/matplotlib
netcdf4	1.5.8	MIT	https://anaconda.org/conda-forge/netcdf4
numpy	1.23.0	BSD-3-Clause	https://anaconda.org/conda-forge/numpy
pandas	1.4.3	BSD-3-Clause	https://anaconda.org/conda-forge/pandas
python	3.9	PSF	https://docs.python.org/3/license.html
scipy	1.8.1	BSD-3-Clause	https://anaconda.org/conda-forge/scipy
xarray	2022.3.0	Apache-2.0	https://anaconda.org/conda-forge/xarray
jupyter	1.0.0	Unspecified	https://anaconda.org/anaconda/jupyter

Conda will take care of this…

5. Getting the code

- Git way:

```
cd ~
mkdir ThoMaS
cd ThoMaS
git clone --depth 1 https://gitlab.eumetsat.int/eumetlab/oceans/ocean-science-studies/ThoMaS .
```

- Direct download:

https://gitlab.eumetsat.int/eumetlab/oceans/ocean-science-studies/ThoMaS

Recent updates were done on the code

```
git fetch
git pull
```

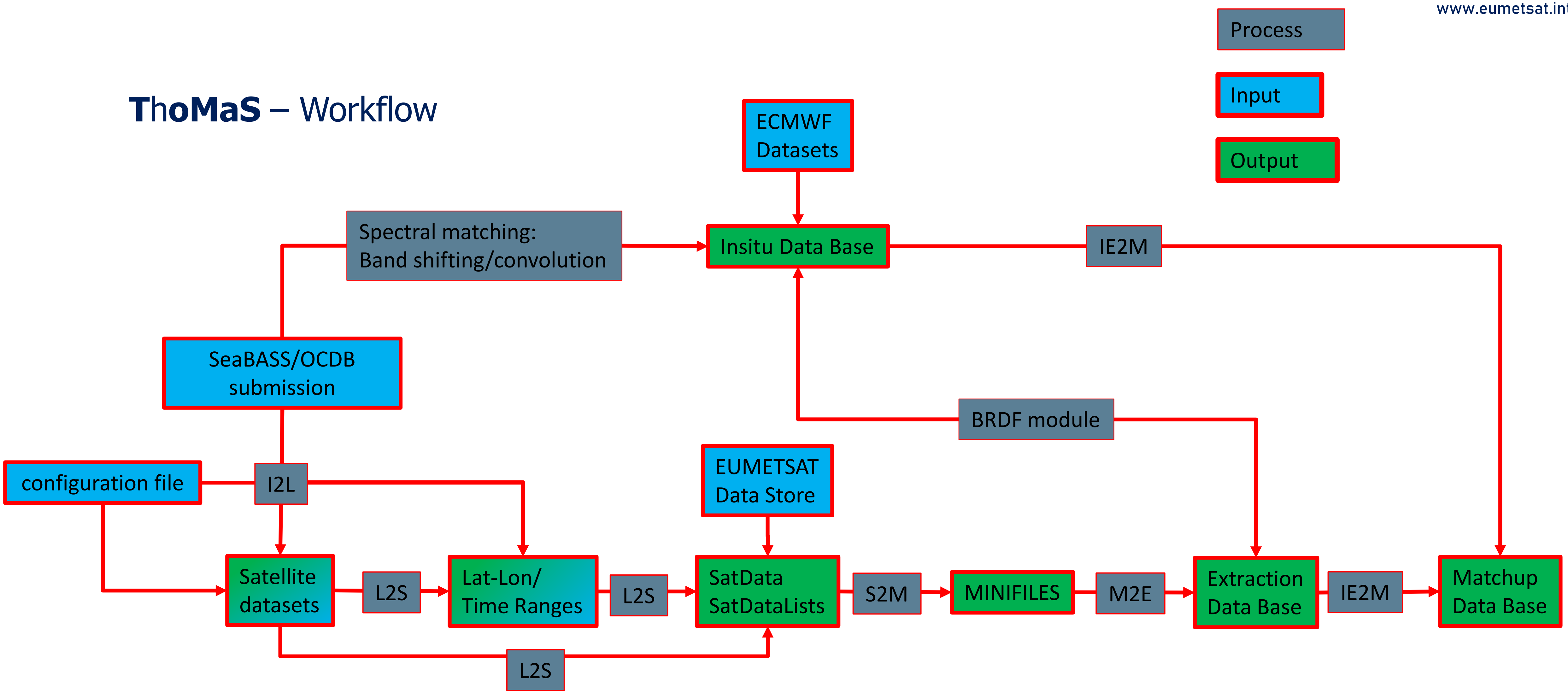
6. Setting the environment

- Once conda and ThoMaS are installed, create the thomas env:

```
cd ~
cd ThoMaS
conda env create -f environment.yml
conda activate thomas
```

libmamba is the best choice for those of you who are stuck in the **"Solving environment step"**

ThoMaS – Workflow



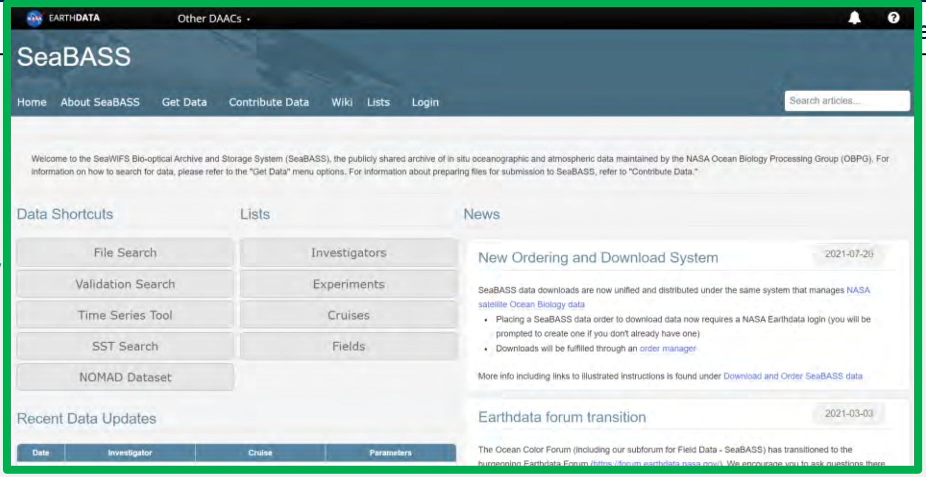
```
/begin_header
/investigators=TestPI
/affiliations=TestAffiliation
/contact=TestContact@TestInstitututution.org
/received=20190101
/processed=20190101
/experiment=TestExperiment
/cruise=TestCruise
/station=NA
/data_file_name=TestOCDBsubmission.txt
/documents=see_comments
/calibration_files=see_comments
/data_type=cast
/data_status=final
/water_depth=NA
/wavelength_option=hyperspectral
/BRDF_option=None
/instrument_model=RAMSES
/instrument_manufacturer=TriOS
/calibration_date=NA
!
! COMMENTS
!
! Citation: Cite your paper where your data are published
!
/missing=-9999.
/delimiter=comma
/fields=station,date,time,lon,lat,depth,cloud,RelAz,AOT,spm,Chla,Rrs355,Rrs360,Rrs365,Rrs370,Rrs375,Rrs380,Rrs385,Rrs390,Rrs395,Rrs400,Rrs405,Rrs410,…
/units=none,yyyymmdd,hh:mm:ss,degrees,degrees,m,%,degrees,unitless,mg/L,mg/m^3,1/sr,1/sr,1/sr,1/sr,1/sr,1/sr,1/sr,1/sr,1/sr,1/sr,1/sr,1/sr,1/sr,1/sr,…
/end_header
TestStation001,20210815,11:30:00,0,0,0.01,0,134.7,0.2444,0.1,0.1,0.001886946,0.002068008,0.002167035,0.002305759,0.002504616,0.002699149,0.002943716,0.003166,…
TestStation002,20201014,11:55:00,-32.6232,32.0859,0.01,0,134.8,0.2388,0.1,0.1,0.002158972,0.002363762,0.00247761,0.00264967,0.002880799,0.003116844,0.0034196,…
TestStation003,20201014,12:16:00,-32.0209,26.9584,0.01,0,134.7,0.2388,0.1,0.1,0.001918189,0.002097384,0.002191946,0.002353893,0.002567901,0.002792078,0.00307,…
TestStation004,20210910,09:40:00,3,-3,0.01,0,134.9,0.2388,0.1,0.1,0.001271305,0.001359205,0.00139314,0.001432681,0.001487187,0.00153414,0.00160502,0.00165816,…
TestStation005,20210910,09:54:00,-4,4,0.01,0,134.8,0.277,0.1,0.1,0.00123858,0.001313013,0.001336158,0.00136599,0.001412448,0.001450581,0.001508596,0.00155277,…
TestStation006,20210910,10:07:00,5,-5,0.01,0,134.8,0.277,0.1,0.1,0.001126482,0.001192422,0.001211226,0.001237373,0.001276362,0.001307301,0.00135953,0.0013975,…
```

7. Required inputs: SeaBASS/OCDB file

```
/begin_header
/investigators=TestPI
/affiliations=TestAffiliation
/contact=TestContact@TestInstitututuion.org
/received=20190101
/processed=20190101
/experiment=TestExperiment
/cruise=TestCruise
/station=NA
/data_file_name=TestOCDBsubmission.txt
/documents=see_comments
/calibration_files=see_comments
/data_type=cast
/data_status=final
/water_depth=NA
/wavelength_option=hyperspectral
/BRDF_option=None
/instrument_model=RAMSES
/instrument_manufacturer=TriOS
/calibration_date=NA
!
! COMMENTS
!
! Citation: Cite your paper where your data are published
!
/missing=-9999.
/delimiter=comma
/fields=station,date,time,lon,lat,depth,cloud,RelAz,AOT,spm,Chla,Rrs355,Rrs360,Rrs365,Rrs370,Rrs375,Rrs380,Rrs385,Rrs390,Rrs395,Rrs400,Rrs405,Rrs410,…
/units=none,yyyymmdd,hh:mm:ss,degrees,degrees,m,%,degrees,unitless,mg/L,mg/m^3,1/sr,1/sr,1/sr,1/sr,1/sr,1/sr,1/sr,1/sr,1/sr,1/sr,1/sr,1/sr,1/sr,1/sr,…
/end_header
TestStation001,20210815,11:30:00,0,0,0.01,0,134.7,0.2444,0.1,0.1,0.001886946,0.002068008,0.002167035,0.002305759,0.002504616,0.002699149,0.002943716,0.003166,…
TestStation002,20201014,11:55:00,-32.6232,32.0859,0.01,0,134.8,0.2388,0.1,0.1,0.002158972,0.002363762,0.00247761,0.00264967,0.002880799,0.003116844,0.0034196,…
TestStation003,20201014,12:16:00,-32.0209,26.9584,0.01,0,134.7,0.2388,0.1,0.1,0.001918189,0.002097384,0.002191946,0.002353893,0.002567901,0.002792078,0.00307,…
TestStation004,20210910,09:40:00,3,-3,0.01,0,134.9,0.2388,0.1,0.1,0.001271305,0.001359205,0.00139314,0.001432681,0.001487187,0.00153414,0.00160502,0.00165816,…
TestStation005,20210910,09:54:00,-4,4,0.01,0,134.8,0.277,0.1,0.1,0.00123858,0.001313013,0.001336158,0.00136599,0.001412448,0.001450581,0.001508596,0.00155277,…
TestStation006,20210910,10:07:00,5,-5,0.01,0,134.8,0.277,0.1,0.1,0.001126482,0.001192422,0.001211226,0.001237373,0.001276362,0.001307301,0.00135953,0.0013975,…
```

SeaBASS

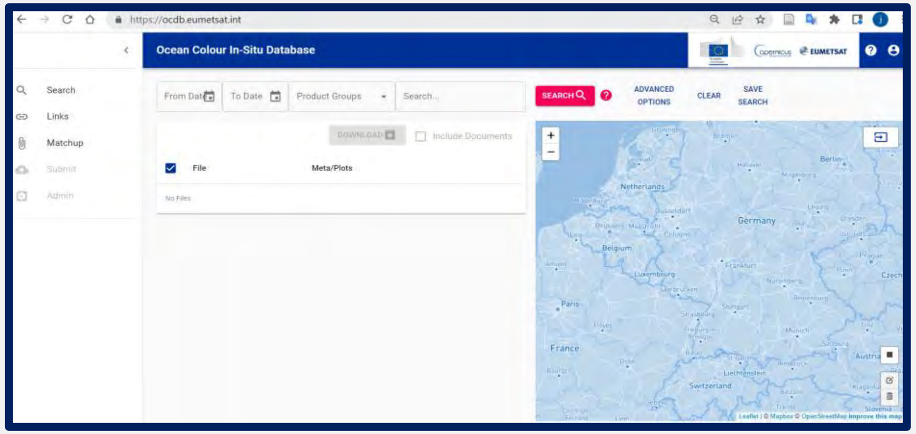
seabass.gsfc.nasa.gov



Inter-operable

OCDB

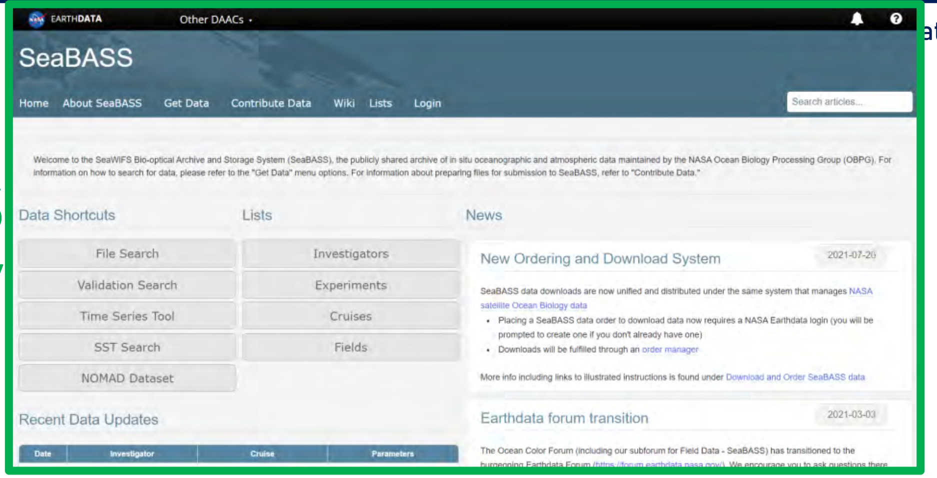
ocdb.eumetsat.int



7. Required inputs: SeaBASS/OCDB file

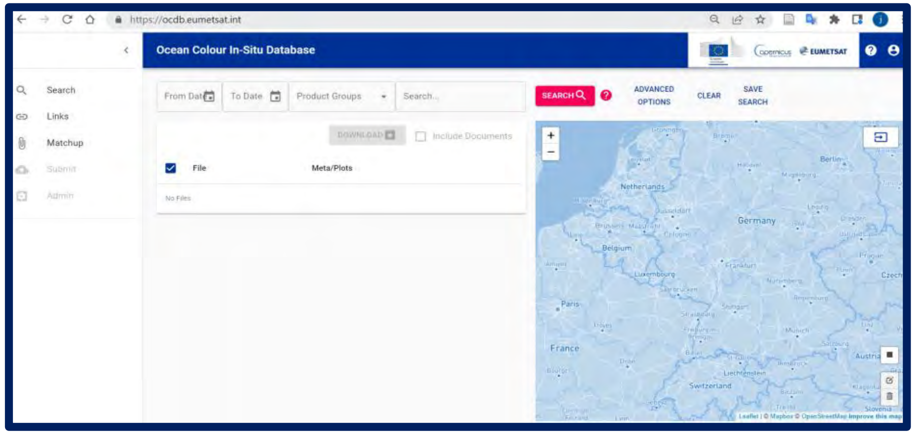
1. OCDB & SeaBASS offer documentation on how to get your in situ data in the correct format.
2. When submitting your data to OCDB, OCDB will guide you on the format.
3. ThoMaS repository contains example in situ files in this format and links to all the necessary resources

SeaBASS
seabass.gsfc.nasa.gov



Inter-operable

OCDB
ocdb.eumetsat.int




```
/begin_header
/investigators=TestPI
/affiliations=TestAffiliation
/contact=TestContact@TestInstitutution.org
/received=20190101
/processed=20190101
/experiment=TestExperiment
/cruise=TestCruise
/station=NA
/data_file_name=TestOCDBsubmission.txt
/documents=see_comments
/calibration_files=see_comments
/data_type=cast
/data_status=final
/water_depth=NA
/wavelength_option=hyperspectral
/BRDF_option=None
/instrument_model=RAMSES
/instrument_manufacturer=TriOS
/calibration_date=NA
!
! COMMENTS
!
! Citation: Cite your paper where your data are published
!
/missing=-9999
/delimiter=comma
/fields=station,date,time,lon,lat,depth,cloud,RelAz,AOT,spm,Chla,Rrs355,Rrs360,Rrs365,Rrs370,Rrs375,Rrs380,Rrs385,Rrs390,Rrs395,Rrs400,Rrs405,Rrs410,…
/units=none,yyyymmdd,hh:mm:ss,degrees,degrees,m,%,degrees,unitless,mg/L,mg/m^3,1/sr,1/sr,1/sr,1/sr,1/sr,1/sr,1/sr,1/sr,1/sr,1/sr,1/sr,1/sr,1/sr,1/sr,…
/end_header
TestStation001,20210815,11:30:00,0,0,0.01,0,134.7,0.2444,0.1,0.1,0.001886946,0.002068008,0.002167035,0.002305759,0.002504616,0.002699149,0.002943716,0.003166,…
TestStation002,20201014,11:55:00,-32.6232,32.0859,0.01,0,134.8,0.2388,0.1,0.1,0.002158972,0.002363762,0.00247761,0.00264967,0.002880799,0.003116844,0.0034196,…
TestStation003,20201014,12:16:00,-32.0209,26.9584,0.01,0,134.7,0.2388,0.1,0.1,0.001918189,0.002097384,0.002191946,0.002353893,0.002567901,0.002792078,0.00307,…
TestStation004,20210910,09:40:00,3,-3,0.01,0,134.9,0.2388,0.1,0.1,0.001271305,0.001359205,0.00139314,0.001432681,0.001487187,0.00153414,0.00160502,0.00165816,…
TestStation005,20210910,09:54:00,-4,4,0.01,0,134.8,0.277,0.1,0.1,0.00123858,0.001313013,0.001336158,0.00136599,0.001412448,0.001450581,0.001508596,0.00155277,…
TestStation006,20210910,10:07:00,5,-5,0.01,0,134.8,0.277,0.1,0.1,0.001126482,0.001192422,0.001211226,0.001237373,0.001276362,0.001307301,0.00135953,0.0013975,…
```

8. Run the code

Example 1: Indian Ocean

1. You want to download Sen-3A data overpassing the point location **(26 S, 100 E)** in the Indian Ocean, west of Australia, where I deployed an in situ instrument measuring continuously during the first 10 days of June 2022.

2. You want data from S3A, and L2 of the recent collection OL__L2M.003

3. Only full resolution (FR).

4. You just want to obtain the S3 files (SatData), minifiles and extractions, I have my own scripts to compute the statistics of the comparison with insitu.

5. In particular, you want to test EUMETSAT's standard protocol for window size: 3x3.

6. You want everything related to the run to be stored at /path/to/Indian_Ocean

Since insitu not provided, define **latLonTimeRanges.csv** and store it under ~/Indian_Ocean

StationID	Lat	Lon	time_start	time_stop
Indian_Ocean	-26	100	2022-06-01T00:00:00	2022-06-10T00:00:00

8. Run the code

Example 1: Indian Ocean

1. You want to download Sen-3A data overpassing the point location **(26 S, 100 E)** in the Indian Ocean, west of Australia, where I deployed an in situ instrument measuring continuously during the first 10 days of June 2022.

2. You want data from S3A, and L2 of the recent collection OL__L2M.003

3. Only full resolution (FR).

4. You just want to obtain the S3 files (SatData), minifiles and extractions, <u>I have my own scripts to compute the statistics of the comparison with insitu</u>.

5. In particular, you want to test EUMETSAT's standard protocol for window size: 3x3.

6. You want everything related to the run to be stored at /path/to/Indian_Ocean

Your configuration file must look like:

```
[global]
path_output: /path/to/Indian_Ocean
SetID: Indian_Ocean

[workflow]
workflow: SatData, minifiles, EDB

[satellite]
satellite_path-to-SatData: /path/to/Indian_Ocean/SatData
satellite_source: EUMETSATdataStore
satellite_collections: OL__L2M.003
satellite_platforms: S3A
satellite_resolutions: FR

[minifiles]
minifiles_winSize: 3

[EDB]
EDB_protocols_L2: EUMETSAT_standard_L2
EDB_winSizes: 3
```

Since <u>insitu not provided</u>, define **latLonTimeRanges.csv** and store it under ~/Indian_Ocean

StationID	Lat	Lon	time_start	time_stop
Indian_Ocean	-26	100	2022-06-01T00:00:00	2022-06-10T00:00:00

Example 6: MOBY

1. You have a prepared a set of hyperspectral Rrs insitu measurements from MOBY in SeaBASS format not corrected for BRDF effects.

2.You wish to get matchups between this MOBY subset and
- •S3A/OLCI standard FR L2,
- •From the current collection OL__L2M.003
- •using the standard extraction protocol from EUMETSAT,
- •an extraction window of 5x5,
- •an insitu-satellite time difference threshold of 1 hour (3600 seconds).

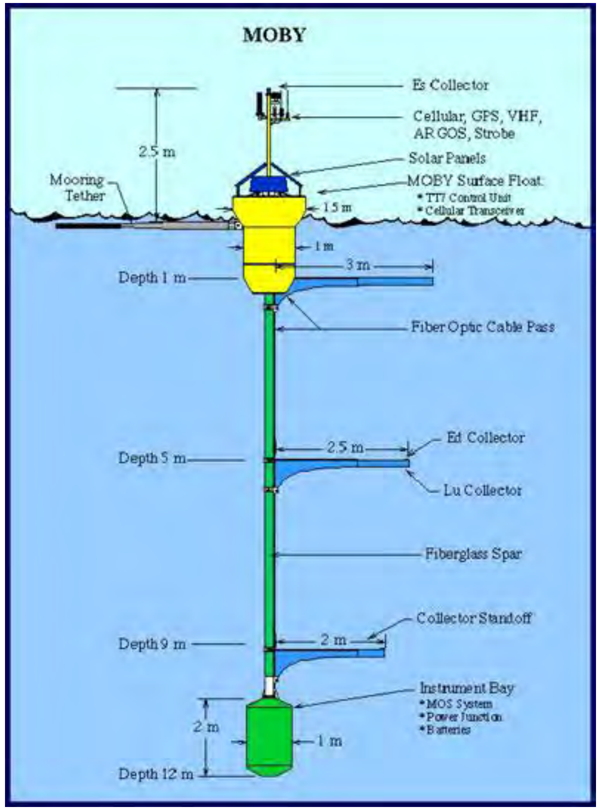
3.You are not interested in getting ancillary data from ECMWF for to the insitu data.

4.You want to apply the Morel et al. 2002 BRDF correction to both satellite and insitu.

5.You may have several insitu measurements corresponding to one single SatData within the time window that you selected, but you wish to keep only the closest in time with the satellite overpass.

6.You wish: SatData to be stored at /path/to/MOBY/SatData

7.all the other outputs (IDB, minifiles, EDB, MDB, etc.) to be stored at /path/to/MOBY



In this case, **latLonTimeRanges.csv** will be generated automatically by ThoMaS (based on your inputted insitu lat-lon and timestamps + your inputted time tolerance in config_file.ini) and stored under path_output

8. Run the code

Example 6: MOBY

1. You have a prepared a set of hyperspectral Rrs insitu measurements from MOBY in SeaBASS format not corrected for BRDF effects.

2.You wish to get matchups between this MOBY subset and
 - •S3A/OLCI standard FR L2,
 - •From the current collection OL__L2M.003
 - •using the standard extraction protocol from EUMETSAT,
 - •an extraction window of 5x5,
 - •an insitu-satellite time difference threshold of 1 hour (3600 seconds).

3.You are not interested in getting ancillary data from ECMWF for to the insitu data.

4.You want to apply the Morel et al. 2002 BRDF correction to both satellite and insitu.

5.You may have several insitu measurements corresponding to one single SatData within the time window that you selected, but you wish to keep only the closest in time with the satellite overpass.

6.You wish: SatData to be stored at /path/to/MOBY/SatData

7.all the other outputs (IDB, minifiles, EDB, MDB, etc.) to be stored at /path/to/MOBY

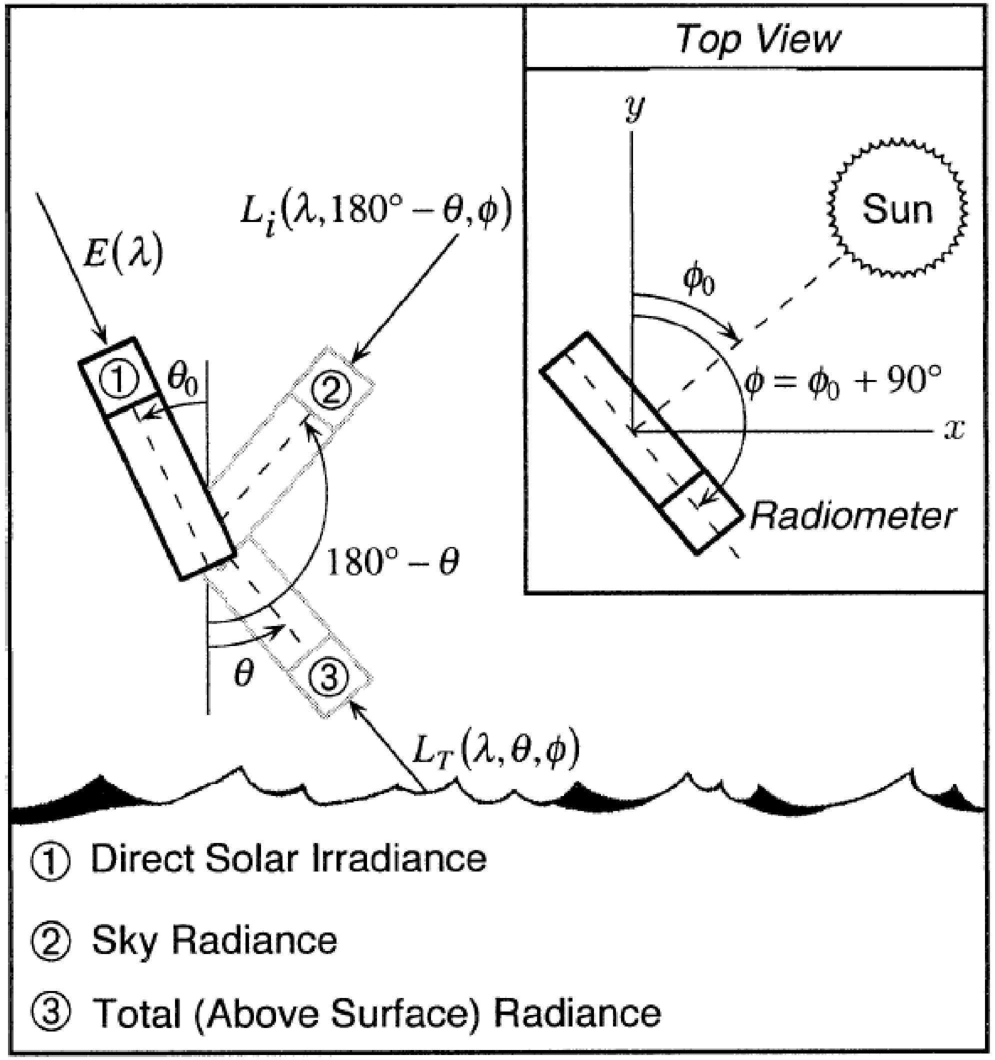
In this case, **latLonTimeRanges.csv** will be generated automatically by ThoMaS (based on your inputted insitu lat-lon and timestamps + your inputted time tolerance in config_file.ini) and stored under path_output

Your configuration file must look like:

```
[global]
path_output: /path/to/MOBY
SetID: MOBY

[workflow]
workflow: insitu, SatData, minifiles, EDB, MDB

[insitu]
insitu_input: /path/to/MOBY/MOBY_OCDB.csv
insitu_satelliteTimeToleranceSeconds: 3600
insitu_getAncillary: False
insitu_BRDF: M02

[satellite]
satellite_path-to-SatData: /path/to/MOBY/SatData
satellite_source: EUMETSATdataStore
satellite_collections: OL__L2M.003
satellite_platforms: S3A
satellite_resolutions: FR
satellite_BRDF: M02

[minifiles]
minifiles_winSize: 5

[EDB]
EDB_protocols_L2: EUMETSAT_standard_L2
EDB_winSizes: 5

[MDB]
MDB_time-interpolation: insitu2satellite_NN
MDB_stats_plots: True
MDB_stats_protocol: EUMETSAT_standard_L2
```

8. Run the code

Example 8: Socheongcho



[Zibordi et al. 2002]

etsat.int

Ongjin Socheongcho Ocean Research Station

8. Run the code

Example 8: Socheongcho

1. You want to test the performance of OLCI at the AERONET-OC station Socheongcho, West of Korea (Yellow Sea) during March 2021.

2.You wish to get matchups between this Socheongcho subset and
- •S3A/OLCI standard FR L2,
- •From the current collection OL__L2M.003
- •using the standard extraction protocol from EUMETSAT,
- •an extraction window of 5x5,
- •an insitu-satellite time difference threshold of 1 hour (3600 seconds).

3.You are not interested in getting ancillary data from ECMWF for to the insitu data.

4.You want to apply the Morel et al. 2002 BRDF correction to both satellite and insitu.

5.You may have several insitu measurements corresponding to one single SatData within the time window that you selected, but you wish to keep only the closest in time with the satellite overpass.

6.You wish: SatData to be stored at /path/to/Socheongcho/SatData

7.all the other outputs (IDB, minifiles, EDB, MDB, etc.) to be stored at /path/to/Socheongcho

In this case, **latLonTimeRanges.csv** will be generated automatically by ThoMaS (based on your inputted insitu lat-lon and timestamps + your inputted time tolerance in config_file.ini) and stored under path_output

```
[global]
path_output: /path/to/Socheongcho
SetID: Socheongcho

[workflow]
workflow: insitu, SatData, minifiles, EDB, MDB

[AERONETOC]
AERONETOC_pathRaw:/path/to/AERONET_OC_raw
AERONETOC_dateStart: 2021-03-01T00:00:00
AERONETOC_dateEnd: 2021-04-01T00:00:00
AERONETOC_dataQuality: 2
AERONETOC_station: Socheongcho


[insitu]
insitu_input: /path/to/Socheongcho/Socheongcho_OCDB.csv
insitu_satelliteTimeToleranceSeconds: 3600
insitu_getAncillary: False
insitu_BRDF: M02

[satellite]
satellite_path-to-SatData: /path/to/Socheongcho/SatData
satellite_source: EUMETSATdataStore
satellite_collections: OL__L2M.003
satellite_platforms: S3A
satellite_resolutions: FR
satellite_BRDF: M02

[minifiles]
minifiles_winSize: 5

[EDB]
EDB_protocols_L2: EUMETSAT_standard_L2
EDB_winSizes: 5

[MDB]
MDB_time-interpolation: insitu2satellite_NN
MDB_stats_plots: True
MDB_stats_protocol: EUMETSAT_standard_L2
```

int

8. Run the code

1. EUMETSAT Data Store credentials obtained and stored?
2. (optional) ECMWF ADS/CDS credentials obtained and stored?
3. ThoMaS code cloned?
4. thomas conda environment set up and activated?
5. Required inputs in place? (config_file.ini, insitu input file?, lat-lon-time ranges file?)

If **YES**.. run by executing this command:
python /path/to/ThoMaS/main.py –cf /path/to/config_file.ini

9. Set-up demo

9. Set-up demo: basic (Linux) terminal commands

Command	Description
`cd dir_name`	Change directory to “dir_name”
`cd ~`	Change directory to home dir
`mkdir dir_name`	Make new directory “dir_name”
`ls`	List files in current directory
`ls –a`	List files including hidden ones
`nano filename`	Check/edit file “filename”
`conda`	Check conda is installed
`conda env create –f env_file`	Create environment “env_name” from “env_file” file
`conda activate env_name`	Activate environment “env_name”
`git`	Check git is installed
`git clone --depth 1 url .`	Clone repository from URL to current directory

10. Run the code: Demo

11. Add support for new satellite sensors/processors/collections

Currently supported **sensors/processors/collections**

Collection	Sensor	Level	Processor	Resolution	EUMETSATdataStore_IDs	Platforms
OL__L2M.003	OLCI	L2	IPF	FR	EO:EUM:DAT:0556;EO:EUM:DAT:0407	S3A;S3B
OL__L2M.003	OLCI	L2	IPF	RR	EO:EUM:DAT:0557;EO:EUM:DAT:0408	S3A;S3B
OL__L1_.002	OLCI	L1	IPF	FR	EO:EUM:DAT:0577;EO:EUM:DAT:0409	S3A;S3B
OL__L1_.002	OLCI	L1	IPF	RR	EO:EUM:DAT:0578;EO:EUM:DAT:0410	S3A;S3B
OL__L2M.003-IOP	OLCI	L2	IPF-IOP	FR		S3A;S3B
OL__L2M.003-IOP	OLCI	L2	IPF-IOP	RR		S3A;S3B
OBPG	MODIS	L2	l2gen	FR		Aqua
OBPG	VIIRS	L2	l2gen	FR		Suomi-NPP
OBPG	VIIRS	L2	l2gen	FR		NOAA-20
GEOL1	SEVIRI	L1	IPF-GEO	FR		MSG1;MSG2;MSG3;MSG4

12. Add a new extraction protocol file

13. What comes next in ThoMaS?

- A GUI
- Add support of more satellite data (sensor, processor and processing baseline)
- Direct sat data download apart from those contained in EUMETSAT Data Store.
- Direct in situ data download for instruments other than MOBY, AERONET-OC.
- Ingest quality flags for in situ.
- Propagation of uncertainties to the estimation of matchup metrics.
- More BRDF correction schemes?
  - Morel et al. 2002
  - Lee et al. 2011
  - Twardowski and Tonnizzo 2018, Zaneveld 1995
  - Reverse already applied BRDF scheme + apply new one

  → BRDF4OLCI study

→For all this, we need your collaboration, let's make ThoMaS a community effort

Today's tour with ThoMaS

Hope you enjoyed it! Thank you!

Questions are welcome.

Getting started…

- Installing Anaconda
  - If you don't have Python…

- Cloning repositories
  - OR { **On windows:** Open Anaconda, find and launch the powershell.exe application. This will open a command line window.
  - **On Linux or OSx:** Open a terminal window
  - **Clone the learn-olci and ThoMaS repositories:** (you may wish to change directory first)
    - Copy and past the following lines into your terminal and hit "enter"
    - **learn-olci:** *git clone --recurse-submodules --remote-submodules https://gitlab.eumetsat.int/eumetlab/oceans/ocean-training/sensors/learn-olci.git*
    - **ThoMaS:** *git clone https://gitlab.eumetsat.int/eumetlab/oceans/ocean-science-studies/ThoMaS.git*

- Setting up Python environments
  - To create the required Python environments, you should copy and past the following lines into your terminal and hit "enter"
  - **learn-olci:** *conda env create -f learn-olci/environment.yml*
  - **ThoMaS:** *conda env create -f ThoMaS/environment.yml*

Practical session

- If you haven't yet – clone and set up the git repositories for learn-olci/ThoMaS as per the instructions in the README
  - Note the submoduling for learn-olci
  - Extensions are not needed
- Key elements of both include setting up the credentials for access (file with your consumer key and secret).
  - See 1_OLCI_introductory notebooks on data store access
  - Can also see API_authentication notebook for support
- Explore notebooks for data access depending on your experience level/needs
- Run examples **1**, **6** and **8** from ThoMaS (either command line or jupyter notebook ThoMaS_overview.ipynb)

Thank you!

Questions are welcome.

Contacts and further information

For information on our training programme

training@eumetsat.int

For information/support on EUMETSAT services

ops@eumetsat.int

For our training calendar

https://trainingevents.eumetsat.int/trui/