



PROGRAMME OF THE
EUROPEAN UNION



co-funded with



Access to EUMETSAT OC products via EDB and ThoMaS – a Tool to generate Matchups of OC products with Sentinel-3/OLCI

7th Sentinel-3 Validation Team Meeting 2022

18-20 October 2022 | ESA-ESRIN | Frascati (Rm), Italy

Juan Ignacio Gossn
EUMETSAT



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

These scripts are developed to

- i. Access EUMETSAT S3 OC products from EUMETSAT's Data Store (via eumdac) and perform extractions at queried lat/lon/times, and
- ii. Create matchups of bio-geophysical insitu data (in SeaBASS/OCDB-like format) with satellite S3 OC products.

The screenshot shows the GitHub interface for the repository 'juanchossn / ThoMaS'. At the top, there is a search bar and navigation links for Pull requests, Issues, Marketplace, and Explore. Below the repository name, there are buttons for Pin, Unwatch (1), and Fork (0). A navigation bar includes Code, Issues (2), Pull requests, Actions, Projects, Wiki, Security, Insights, and Settings. A notification banner indicates that 'insitu' had recent pushes 1 minute ago, with a 'Compare & pull request' button. Below this, there are buttons for 'Go to file', 'Add file', and 'Code'. A status bar shows 'This branch is 17 commits ahead, 2 commits behind master.' with a 'Contribute' dropdown. The commit history shows a recent commit by 'JuanIgnacio.Gossn' titled 'Slight modifications to readme' from 3 hours ago, with 62 commits. Below the commit list, there are folders for '.idea', 'AERONET', 'ANC', and 'BRDF', each with 'several updates' from 3 hours ago. The 'BRDF' folder has a more detailed commit message: 'Added IPF-IOP, added new fields in IDB_config, solved bug with convo...' from 24 days ago. On the right side, there is an 'About' section with the text: 'These scripts are matchups of bio-geophysical insitu data (in SeaBASS/OCDB-like format) with satellite S3 OC products from EUMETSAT's Data Store. To use the scripts, you need to inputted insitu data and geographic location. The scripts will download the necessary data and generate matchups of satellite OC products, extracted from EUMETSAT's Data Store.' Below the 'About' section, there are links for 'Readme', '3 stars', '1 watching', and '0 forks'.



EDB (Extraction Data Base) workflow

- 1) Searches **EUMETSAT's Data Store** with **eumdac** and downloads requested imagery.
- 2) Locates the **extraction window** and produces **minifiles** centred at this window (e.g. 5x5).
- 3) Calculates the extraction statistics following a user-defined extraction protocol (incl. flagging, outliers)

EDB (Extraction Data Base) workflow

- 1) Searches **EUMETSAT's Data Store** with **eumdac** and downloads requested imagery.
- 2) Locates the **extraction window** and produces **minifiles** centred at this window (e.g. 5x5).
- 3) Calculates the extraction statistics following a user-defined extraction protocol (incl. flagging, outliers)

EUMETSAT Data Store

The screenshot shows the EUMETSAT Data Store interface. The top navigation bar includes the EUMETSAT logo and 'DATA SERVICES'. Below it is a 'Product Navigator / Search results' section. A search filter for 'OLCI' is active, showing 36 results. The left sidebar contains filters for PLATFORM (Sentinel-3), SENSOR TYPE (Optical), SENSOR (OLCI, SLSTR), and PROCESSING LEVEL (Level 1 Data, Level 2 Data). The main content area displays two search results for OLCI Level 1B data, each with a satellite image thumbnail, a title, and a description.

EUMDAC Knowledge Base

The screenshot shows the EUMETSAT Data Access Client Knowledge Base page. The title is 'EUMETSAT Data Access Client'. A blue banner highlights 'EUMDAC - Python Client for EUMETSAT Data Access Services'. The main text explains that the page provides information on the EUMDAC, its capabilities, and how to use it. It also mentions a sidebar and a table of contents. A search box is located at the bottom right of the page.

To access data from the EUMETSAT Data Store, you need a consumer key (detail in ThoMaS readme file)

EDB (Extraction Data Base) workflow

- 1) Searches **EUMETSAT's Data Store** with **eumdac** and downloads requested imagery.
- 2) Locates the **extraction window** and produces **minifiles** centred at this window (e.g. 5x5).
- 3) Calculates the extraction statistics following a user-defined extraction protocol (incl. flagging, outliers)

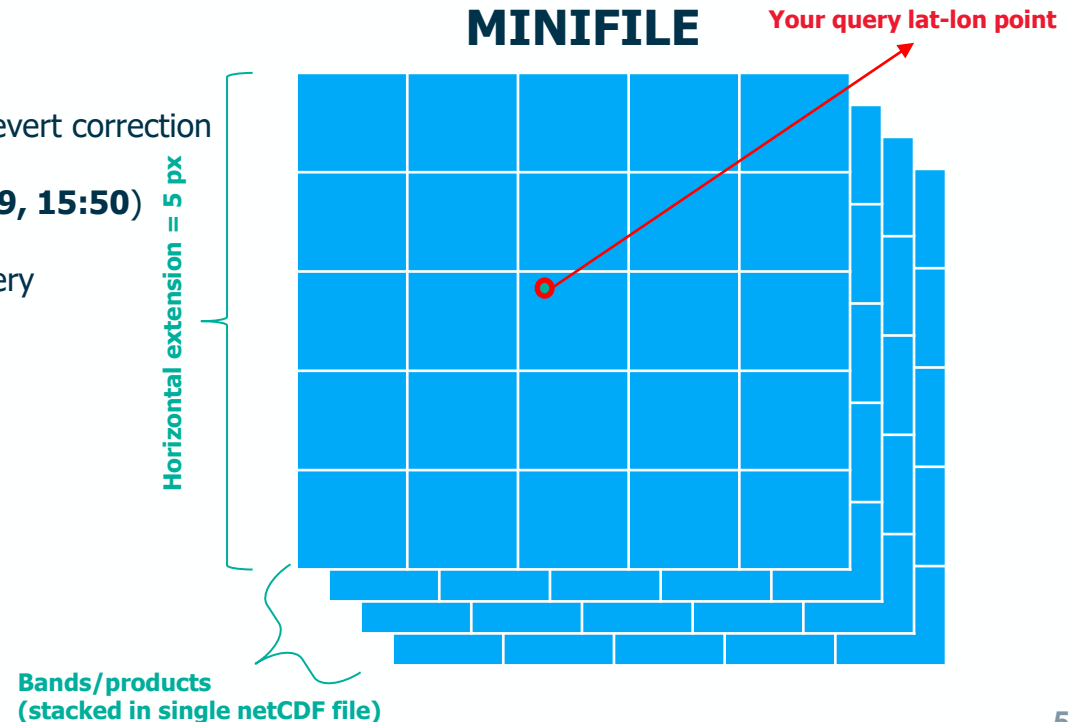
Additionally, BRDF factors are applied (if requested by user)

→ BRDF coefficients provided in MINIFILE in case user wishes to revert correction

→ Several possible schemes may be implemented

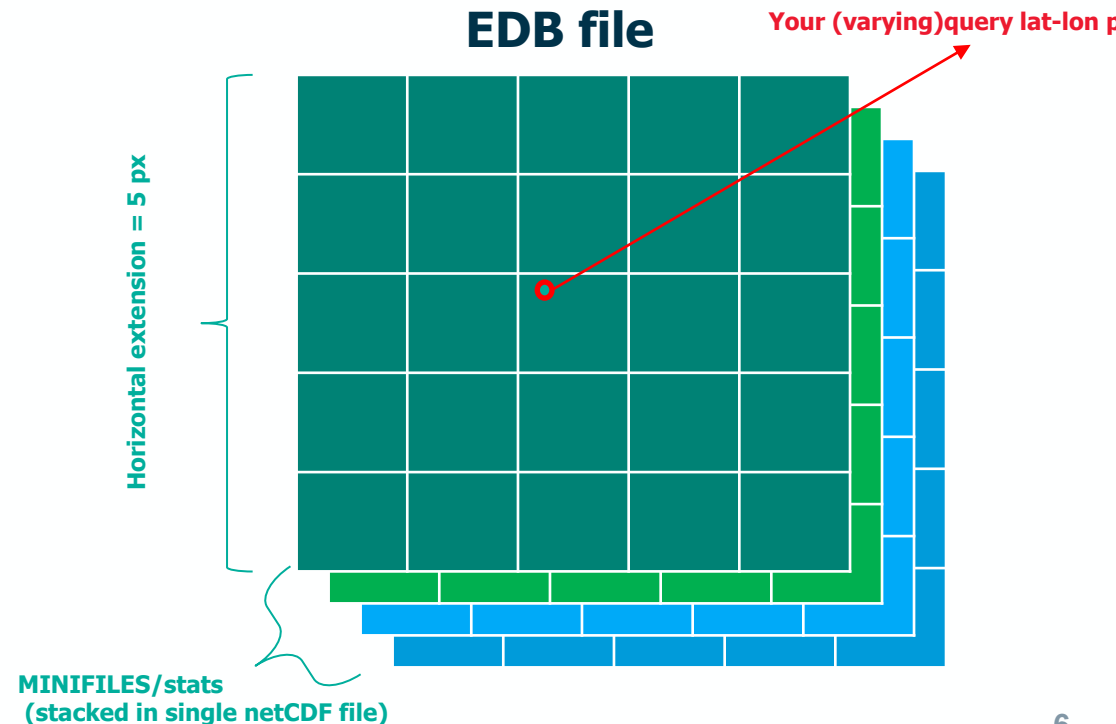
(see **Davide D'Alimonte's presentation, Wednesday 19, 15:50**)

Minifile format extended to SeaDAS-processed MODIS/VIIRS imagery

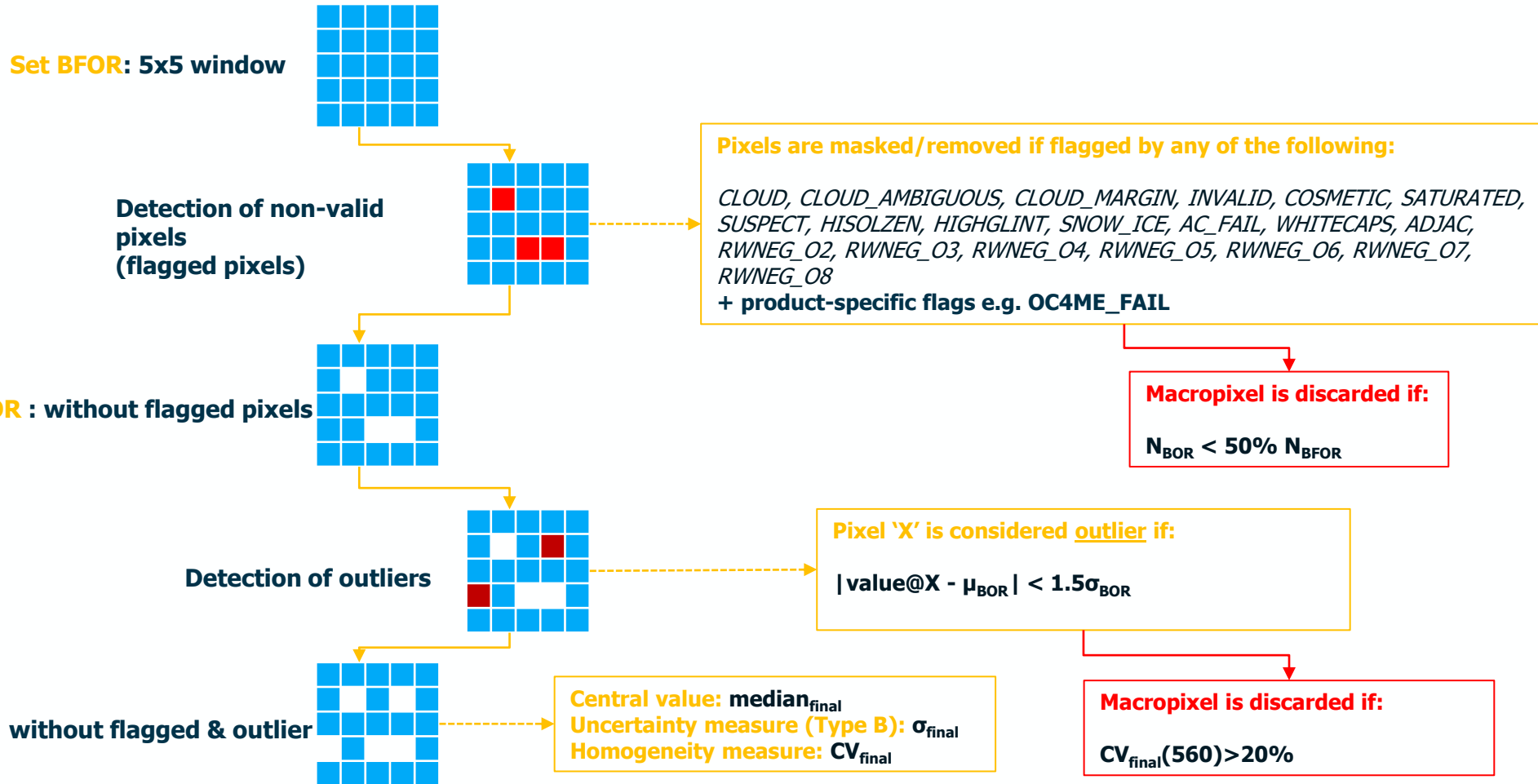


EDB (Extraction Data Base) workflow

- 1) Searches **EUMETSAT's Data Store** with **eumdac** and downloads requested imagery.
- 2) Locates the **extraction window** and produces **minifiles** centred at this window (e.g. 5x5).
- 3) Calculates the extraction statistics following a user-defined extraction protocol (incl. flagging, outliers)



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



Discussion on **EUMETSAT's matchup protocol**: after Davide's presentation
 Wednesday 19, 16:35

$\mu \rightarrow$ Mean
 $\sigma \rightarrow$ Standard deviation

MDB (Matchup Data Base) workflow

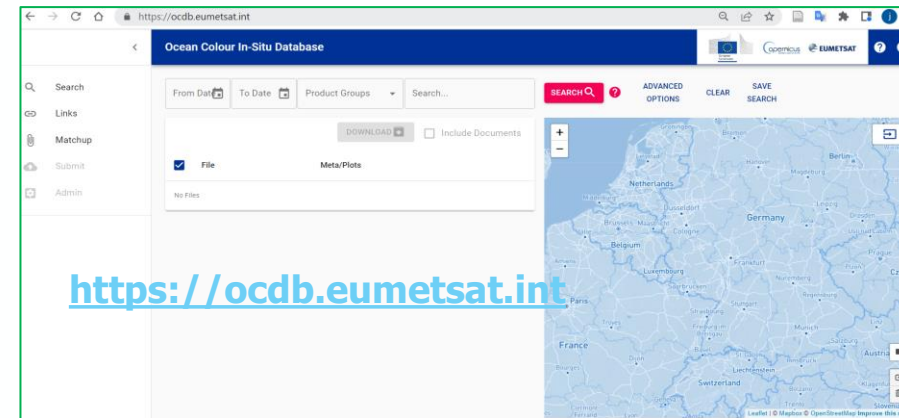
Create matchups of bio-geophysical insitu data (in **SeaBASS/OCDB**-like format) with satellite S3 OC

- 1) Ingest ancillary information (from ECMWF reanalysis datasets) at insitu lat/lon/times,
- 2) Apply convolution (hyperspectral) / band-shifting (multispectral) to satellite OLCI bands
- 3) Apply BRDF correction of in situ data following required scheme (work in progress)
- 4) Download the matching satellite products (from EUMETSAT's Data Store).
- 5) Extract "minifiles" from the satellite products at the desired locations and time periods,
- 6) Calculates the extraction statistics following a user-defined extraction protocol (incl. flagging, outliers)
- 7) Calculate the statistics of the insitu-satellite comparison.

EDB

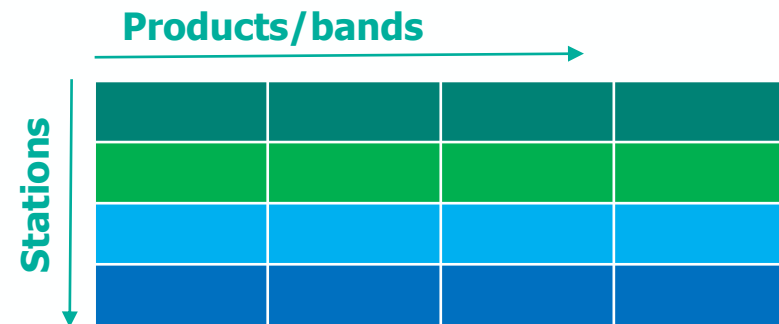
Tested for AERONET-OC data and TriOS data

OCDB – hosted by EUMETSAT



<https://ocdb.eumetsat.int>

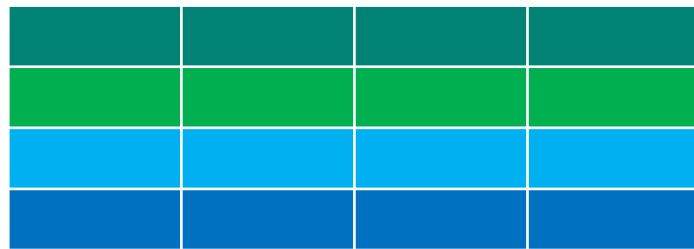
In situ data base (IDB) file



- In situ data must be in the OCDB/SeaBASS file format
- Very easy CSV format.
- Once insitu data is in this format, it's very easy to submit it to OCDB and execute this workflow! ;)

OCDB/SeaBASS file

Products/bands



ECMWF (ERA5/EAC4 datasets)

Products/bands



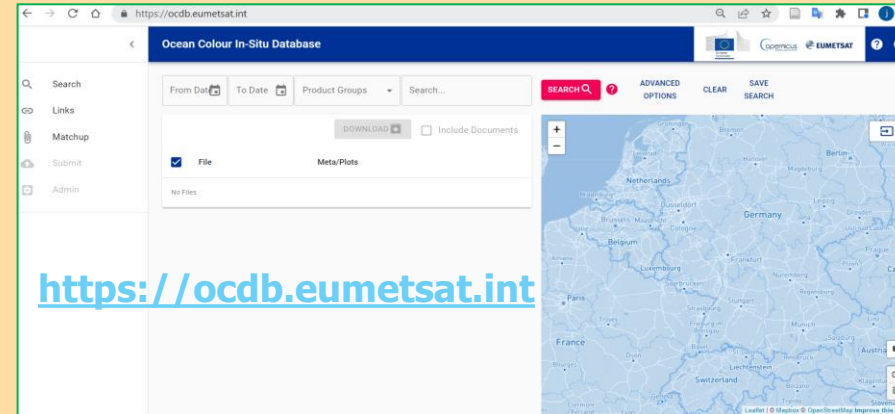
You must have a ECMWF user (very easy, , explained in the README)

Sample OCDB/SeaBASS file

```

/begin_header
/investigators=TestPI
/affiliations=TestAffiliation
/contact=TestContact@TestInstitutation.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,yyymmdd,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,...
    
```

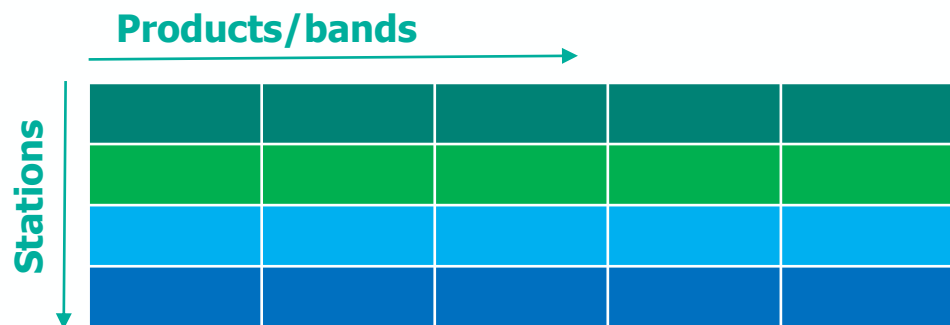
OCDB – hosted by EUMETSAT



MDB file

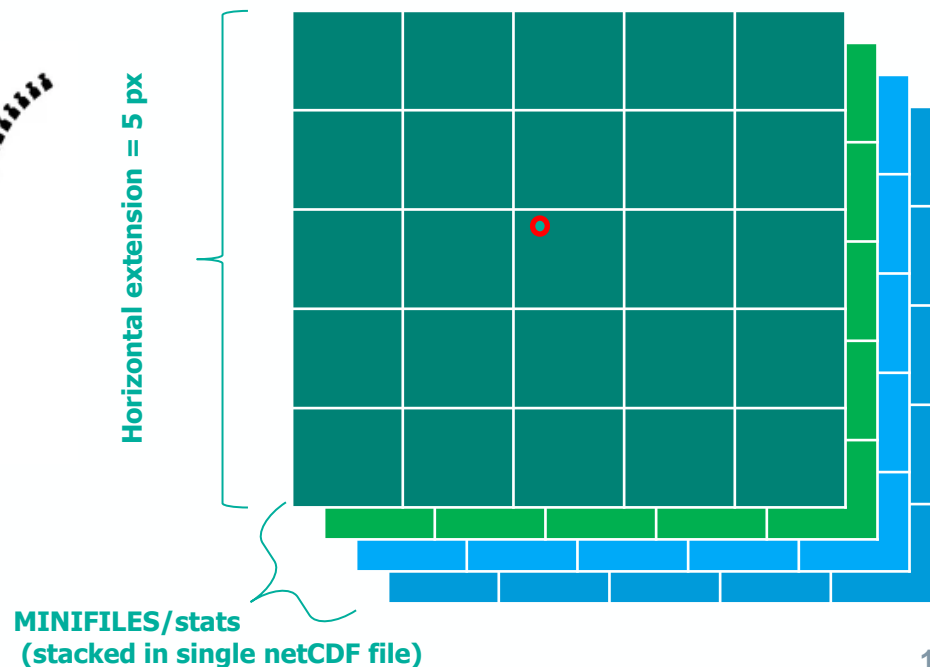
1. Merging insitu and extractions according to matchup pairs
2. Statistical metrics calculated + scatter/spectral plots

In situ data base (IDB) file



Horizontal extension = 5 px

Extraction data base (EDB) file

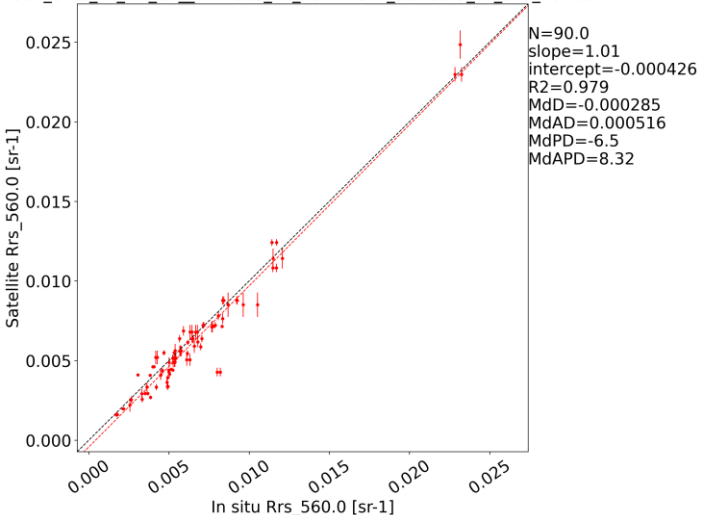


MDB: indexed by matchup pair ID
Still no space/time interpolation applied!

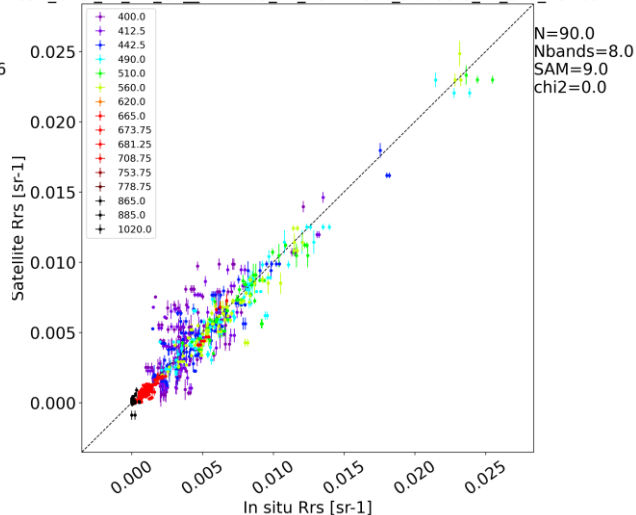


MDB: Automated output plots

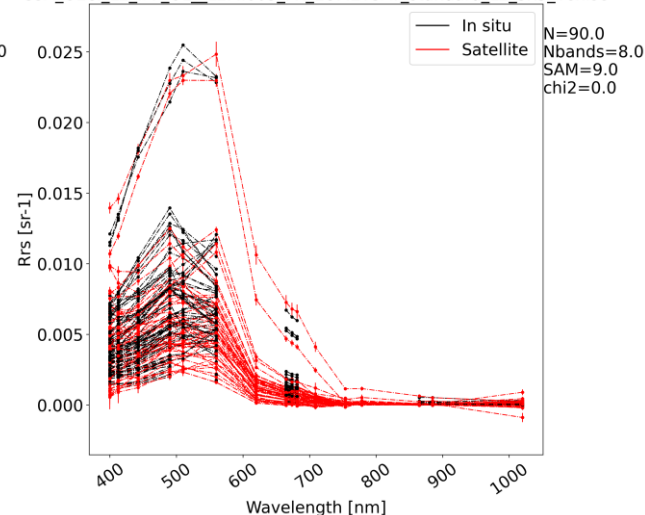
S3A_OLCI L2_IPF_OL_L2M.003_FR_EUMETSAT_standard_L2_5x5_Venise



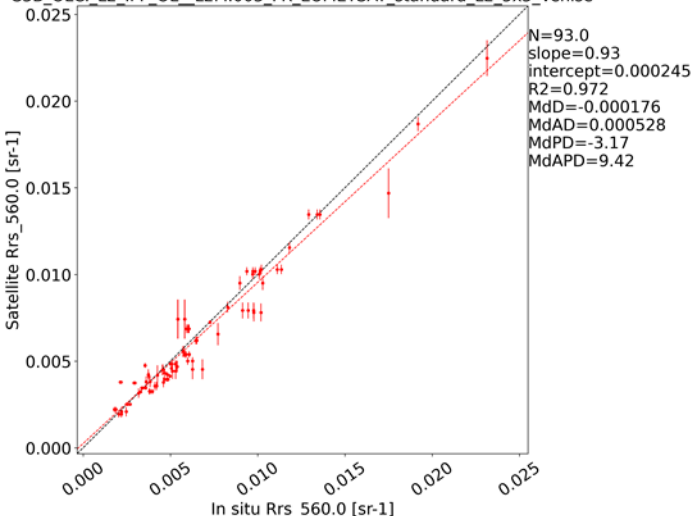
S3A_OLCI L2_IPF_OL_L2M.003_FR_EUMETSAT_standard_L2_5x5_Venise



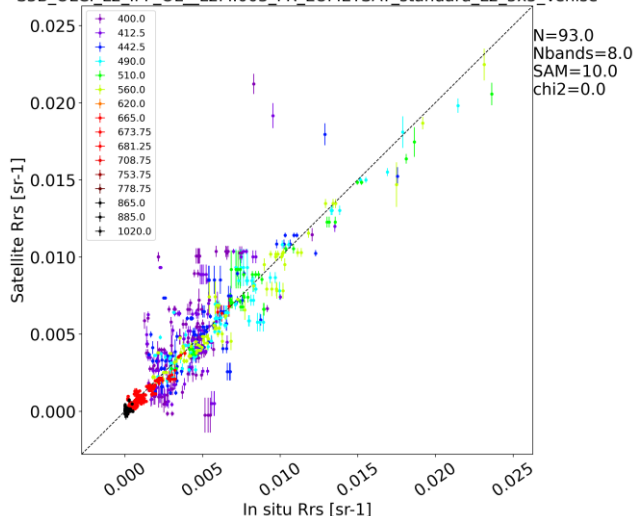
S3A_OLCI L2_IPF_OL_L2M.003_FR_EUMETSAT_standard_L2_5x5_Venise



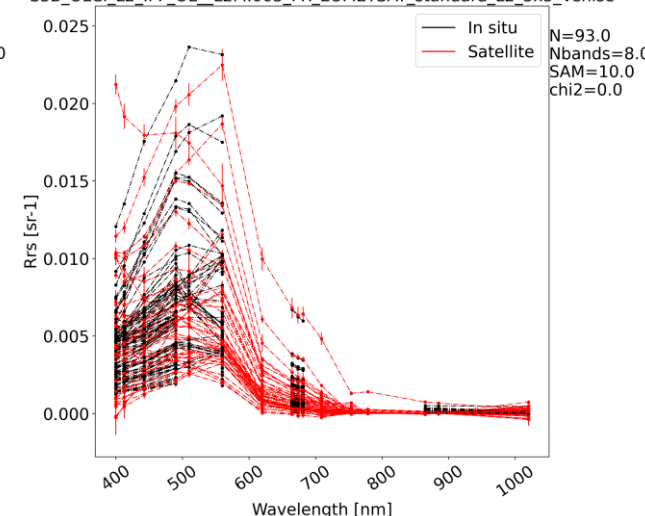
S3B_OLCI L2_IPF_OL_L2M.003_FR_EUMETSAT_standard_L2_5x5_Venise



S3B_OLCI L2_IPF_OL_L2M.003_FR_EUMETSAT_standard_L2_5x5_Venise



S3B_OLCI L2_IPF_OL_L2M.003_FR_EUMETSAT_standard_L2_5x5_Venise

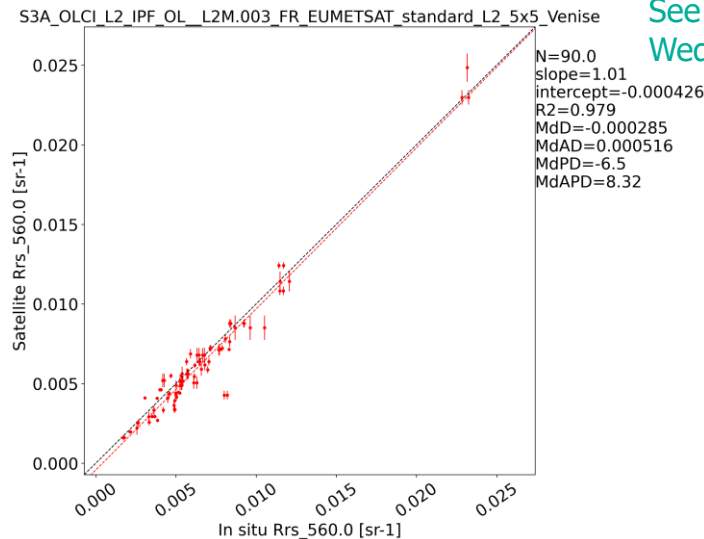
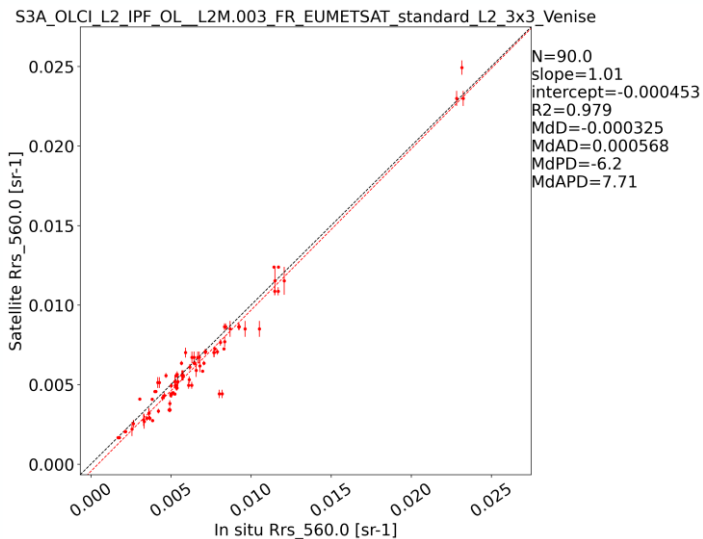


Comparison under various window sizes and extraction protocols

EUMETSAT's standard protocol

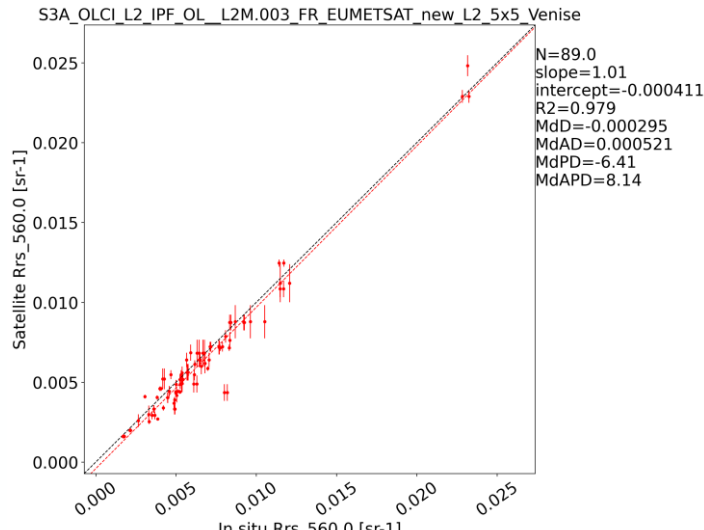
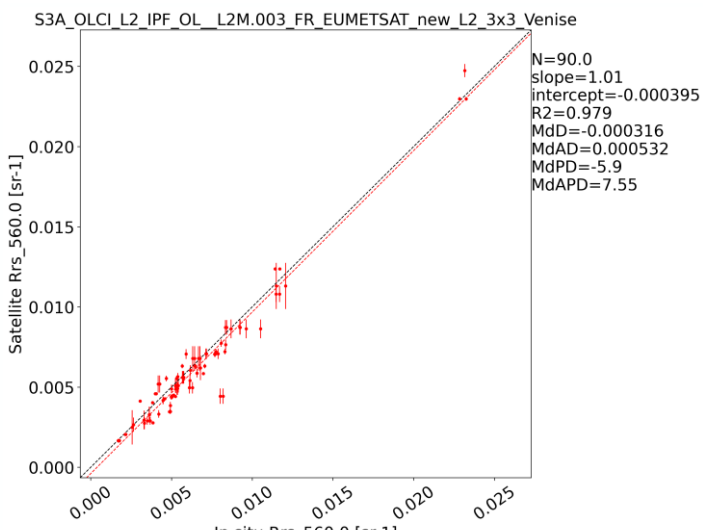
EUMETSAT's alternative protocol

Window size = 3x3



See discussion on matchup protocol
Wednesday ~ 17 hs

Window size = 5x5



Example of config file:

Process data from AERONET-OC/Venise station

Run by executing single command:

```
python main.py -cf config_file_name
```

If README of the code followed: No knowledge of Python programming is required

[global]

```
path_CODE_HOME: /home/myName/ThoMaS_repository  
path_output: /home/myName/Venise  
SetID: Venise
```

[insitu]

```
insitu_input: /home/myName/Venise/20200101_20221008_Venise.LWN_1ev20  
insitu_convert_OCDB: AERONET-OC  
insitu_satelliteTimeToleranceSeconds: 3600  
insitu_bandShifting: MelinSclep2015  
insitu_BRDF: M02
```

[satellite]

```
satellite_path-to-SatData: /home/myName/Venise/SatData  
satellite_source: EUMETSATdataStore  
satellite_collections: OL__L2M.003  
satellite_platforms: S3A, S3B  
satellite_resolutions: FR  
satellite_BRDF: M02
```

[workflow]

```
workflow: insitu, SatData, minifiles, EDB, MDB
```

[minifiles]

```
miniFile_winSize: 5
```

[EDB]

```
EDB_protocols_L1:  
EDB_protocols_L2: EUMETSAT_standard_L2  
EDB_winSizes: 5
```

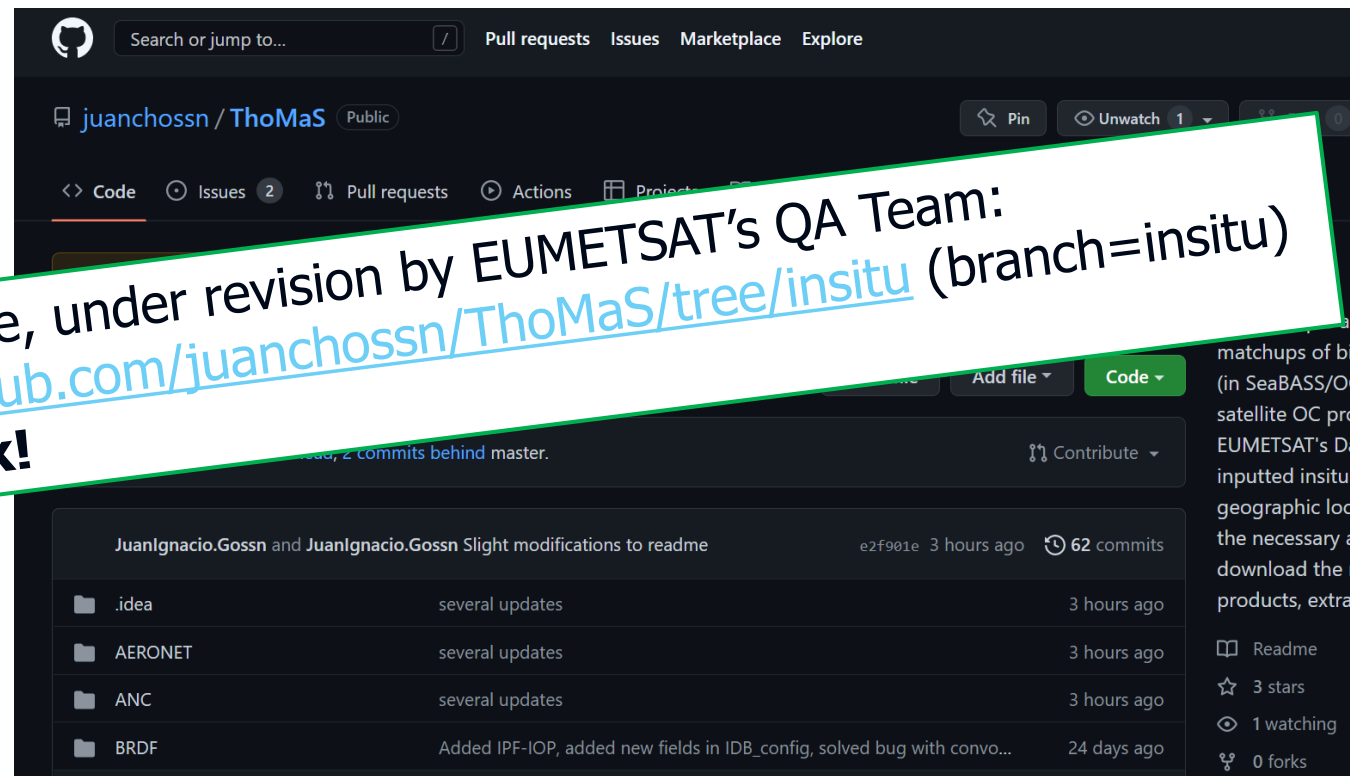
[MDB]

```
MDB_time-interpolation: noTimeInterp  
MDB_stats_protocol: EUMETSAT_standard_L2
```

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

These scripts are developed to

- i. Access EUMETSAT S3 OC products from EUMETSAT's Data Store (via eumdac) and perform extractions at queried lat/lon/times, and
- ii. Create matchups of bio-geophysical insitu data (in SeaBASS/OCDB-like format) with satellite S3 OC products.



The screenshot shows the GitHub interface for the repository 'juanchossn / ThoMaS'. The repository is public and has 2 issues, 1 pull request, and 62 commits. The commit history shows a recent commit by JuanIgnacio.Gossn and JuanIgnacio.Gossn titled 'Slight modifications to readme' with 62 commits, pushed 3 hours ago. The repository structure includes folders for '.idea', 'AERONET', 'ANC', and 'BRDF'. A green callout box is overlaid on the screenshot, containing text about the repository's status and a warning.

Scripts publicly available, under revision by EUMETSAT's QA Team:
Repo: <https://github.com/juanchossn/ThoMaS/tree/insitu> (branch=insitu)
Use at your own risk!

Thanks for listening!
Questions are welcome