Global Monitoring for Environment and Security (GMES) and Africa

eStation 3.0

User Manual

Version 1.2.0

October 2023

Date: 12/10/2023
Ref: GMES/DOC/eStation3.0/EN/UserManual

Abstract / Résumé
This document provides instructions for the exploitation of the eStation system, including the visualization/analysis component.

	Name	Position
Prepared by	Marco Clerici	JRC-EC responsible for ClimSA
Contributions/Reviews by	Christophe Lavaysse	JRC-EC thematic Expert for ClimSA
	Jurriaan Van't Klooster	IT-GIS Specialist
	Vijay Charan Venkatachalam	IT-GIS Specialist
	Dario Simonetti	IT-GIS Specialist
	Fabrizio Cappucci	IT-GIS Specialist

	Details	Date
Version history	Draft 1.1.0	June 2022
1.1.1	Updated with Portfolio, Processing, F4P and IMPACT sections.	August 2022
1.1.2	Updated with new screenshots.	March 2023
1.2.0	Updated with station data section	October 2023

CONTENTS

Table of Contents

ACRONYMS and DEFINITIONS

AMESD	African Monitoring of Environment for Sustainable Development
ACMAD	African Centre of Meteorological Applications for Development
AGRHYMET	Centre Régional de Formation et d'Application en Agrométéorologie et Hydrologie Opérationnelle
AU	African Union
EO	Earth Observation
EUMETSAT	European Organisation for the Exploitation of Meteorological Satellites
EUMETCast	EUMETSAT's primary dissemination mechanism for the near real-time delivery of satellite data and products
FTP	File Transfer Protocol
GIS	Geographical Information System
JRC	Joint Research Centre of the European Commission
REC	Regional Economic Communities
RIC	Regional Implementation Centre
TA	Technical Assistance
TAT	Technical Assistance Team
THEMA	Regional and Continental Thematic Actions

1. INTRODUCTION

1.1 SCOPE OF THE DOCUMENT

This document describes the functionalities of EStation (indicated as 'Station' in the following) application and explains how the final User can benefit from its features. It is meant mainly for the thematic expert making use of the system, and describes both the 'processing' and 'visualization/analysis' components.

1.2 DOCUMENT ORGANIZATION

The present document is structured into the following chapters:

- Chapter 2: Overview of the EStation

This is the basic introduction to the EStation, which provides application's rational, an overview of the GUI, some essential notions to understand the system functioning and an overview of the existing services.

- Chapter 3: EStation User Interface

It describes all functionalities that can be controlled through the GUI by the User, namely the Dashboard, Portfolio, Acquisition, Processing, Data Management, Analysis, Fitness for Purposes, IMPACT, Jupyter Notebooks, System Settings and Help panel.

This is intended for the Thematic User, in order to understand how to control and modify the operations of the application, and how to perform the analysis on the Climate and EO datasets.

- Chapter 4: Reference Guide

It contains a detailed description of the services and is meant for Advanced Users (the Basic User should be able to manage the system – without addition of customized treatments) e.g. the people in the RICs. Here we go in the detail of what the services do, and all the tables used for their customization.

2. OVERVIEW OF THE ESTATION

2.1 SYSTEM CONCEPT

The Station is meant to be a processing server for EO and Climate datasets, rather than a stand-alone GIS platform (like QGIS or similar solutions). The functioning of the application is organized around a number of Services, as displayed in Figure 1, namely:

- Data Acquisition: systematically acquire EO and Climate data from various sources, through FTP, HTTP protocols, APIs and locally available file systems, as a EUMETCast Receiving Station. Therefore, there are several 'get' services that can be configured and controlled independently:
  - Get EUMETCast
  - Get Internet (data from remote FTP and HTTP servers)
  - Get Data Store
- Data Ingestion: the 'Ingest' Service converts the files from the format under which they are acquired in (the so called 'native' format) into GeoTiff or NetCDF, also called 'pivot' formats. Optionally, this operation includes geographical re-projection and clipping to a specific region of interest;
- Processing Service: to derive from the input data additional products, like long term statistics, anomalies, and other added-value indicators;
- System Service: to run a number of 'house-keeping' and background tasks, including managing the local file system and cleaning temporary directories.

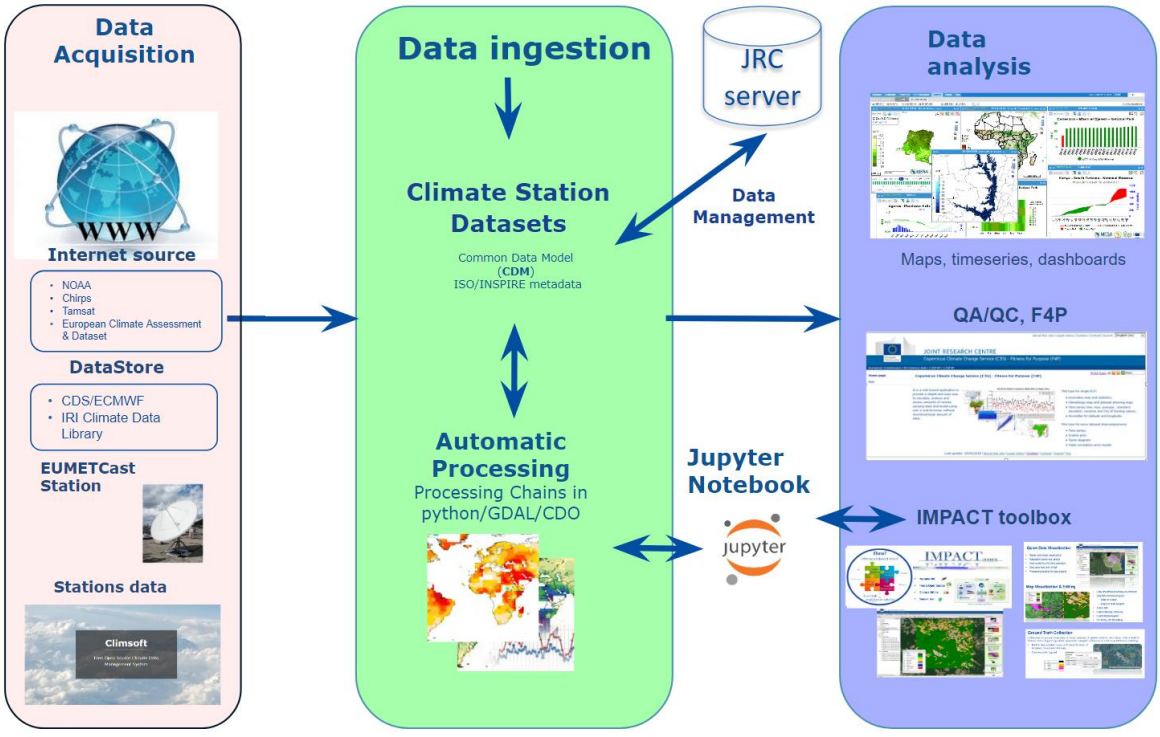


Figure 1: Overview of the Services running on the Station

2.2 System Structure and configurations

The Station's services and the processing components are written in python; the visualisation and GUI component are written in JavaScript (using Ext JS library). A PostgreSQL database stores the EO and Climate products definition, settings for accessing server storing data through the 'get' services and all relevant user configurations.

Several software and libraries are required, including the GDAL library and its python wrapper for geo-processing, Mapserver for maps rendering, and a series of non-standard python modules, such as ruffus, a computation pipeline library, for the processing engine.

All the components of EStation such as database, GUI, IMPACT tool, Mapserver are containerized and deployed in the server using the Docker.

2.3 Application Overview

A web interface, found at the address http://**localhost:8080** allows controlling the Station**.** It folds different controls in separate tabs, namely:

- **Dashboard**: it presents the overall status of the Station, and offers the control of all the enabled services (see Figure 2)
- **Portfolio**: to configure the list of datasets in the Station, mainly to activate products for acquisition and processing.
- **Acquisition**: to view and control the status of the services for retrieving and ingesting the EO and Climate data; it represents and gives access to the 'Get' (internet and datastore) and 'Ingestion' services, represented in Figure 1: Overview of the Services running on the Station .
- **Processing**: to start and stop the processing of new products, i.e. to control the 'Processing' Service.
- **Data Management**: it lists all available data sets, both those acquired and processed by the EStation, showing the completeness of time series (and thus listing missing data).
- **Analysis**: to perform the data analysis and generate images for bulletins/reports.
- **Fitness for Purposes**: it offers the capability to perform comparison among different products. The comparison can be conducted in terms of statistical analysis and/or using global (or regional) maps.
- **IMPACT toolbox**: it offers a combination of remote sensing, photo interpretation and processing technologies in a GIS environment, allowing non specialist users to easily accomplish all necessary pre-processing steps while giving a fast and user-friendly environment for visual editing and map validation.
- **Jupyter Notebooks**: it proposes a working environment of python (ipykernel notebook) where user can code their process to create the indicators from the existing data and visualize them.
- **System Settings**: to control the application settings (including the region, log level, and working directories paths), and to generate system reports.
- **Help**: to read system help files: allows downloading the pdf files and access reference web sites.

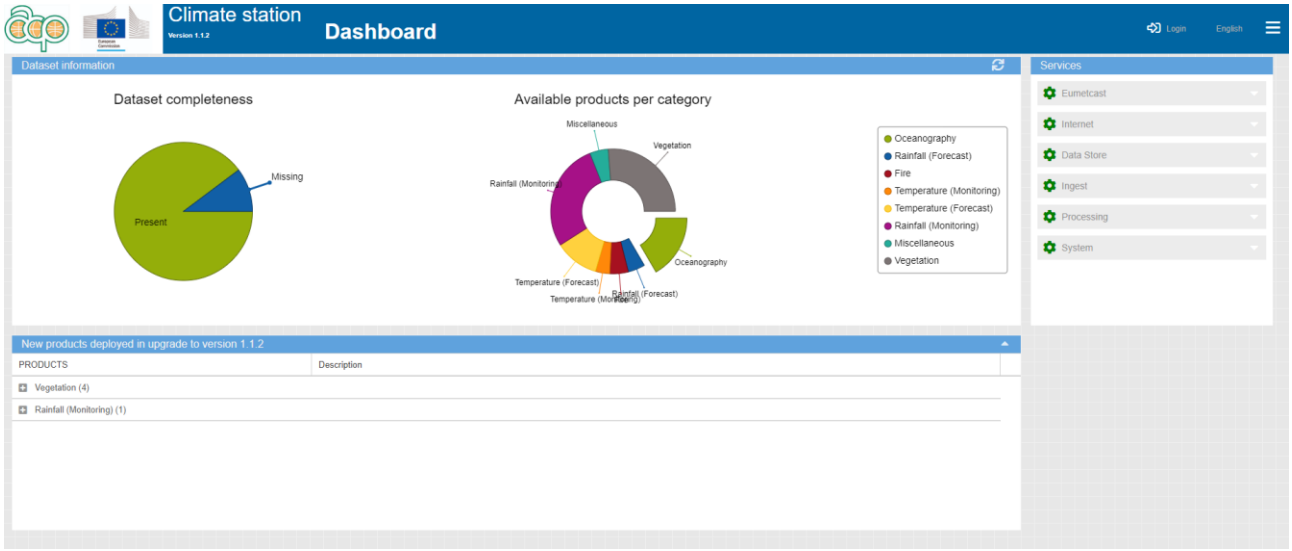


Figure 2: View of EStation Dashboard

2.4 ESTATION ESSENTIAL CONCEPTS

This section presents concepts and definitions underpinning the design and implementation of the Station; whenever possible, definitions make use of examples, and reference to international standards is made.

2.4.1 Product

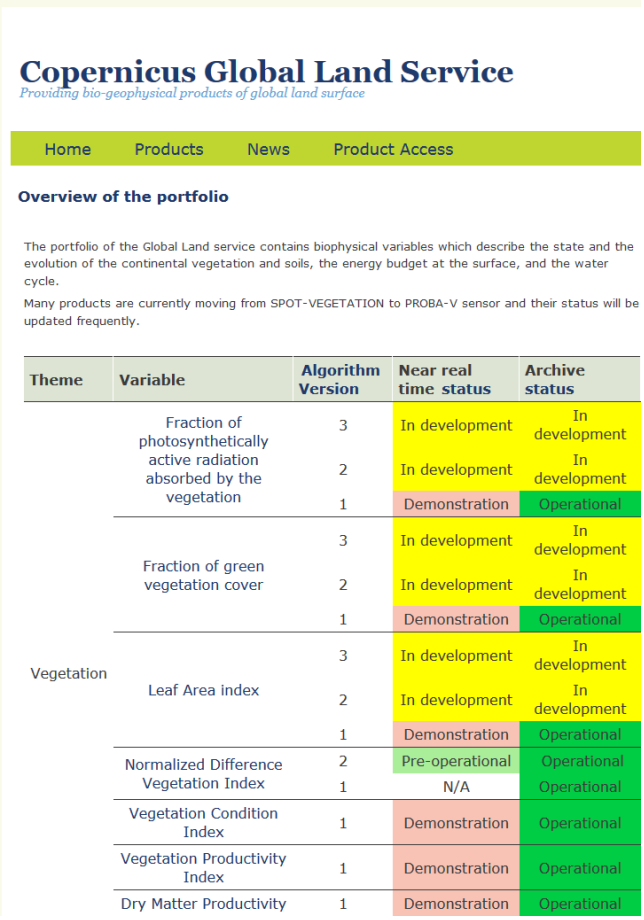
A 'Product', either from EO or models, is characterized by its geographical footprint (the Earth area covered by the image), the image's geographic or projected co-ordinate system, the time frame for synthesis images (daily, 10-day or longer synthesis) and the product update frequency (15-minute, daily, 10-day). These parameters generally depend on the satellite type and the on-board sensor characteristics.

Other characteristics must also be considered, such as the file format (grib2, NetCDF, HDF4 or HDF5), the data distribution policy, or the distribution mean (digital radio broadcast such as EUMETCast or the Internet). In the picture above the 'Products' from Copernicus GLS are displayed as an example.

Figure 3: 'Products' on the Copernicus web site

2.4.2 Version

A product's 'Version' identifies a specific collection of images, depending mainly on the algorithm used for their computation; the various versions of the Copernicus products are displayed in Figure 4 under the column 'Algorithm version'. For the 'incoming' products on the EStation, the version is the one defined by the data provider.



Copernicus Global Land Service

Providing bio-geophysical products of global land surface

Home | Products | News | Product Access

Overview of the portfolio

The portfolio of the Global Land service contains biophysical variables which describe the state and the evolution of the continental vegetation and soils, the energy budget at the surface, and the water cycle.

Many products are currently moving from SPOT-VEGETATION to PROBA-V sensor and their status will be updated frequently.

Theme	Variable	Algorithm Version	Near real time status	Archive status
Vegetation	Fraction of photosynthetically active radiation absorbed by the vegetation	3	In development	In development
		2	In development	In development
		1	Demonstration	Operational
	Fraction of green vegetation cover	3	In development	In development
		2	In development	In development
		1	Demonstration	Operational
	Leaf Area index	3	In development	In development
		2	In development	In development
		1	Demonstration	Operational
	Normalized Difference Vegetation Index	2	Pre-operational	Operational
		1	N/A	Operational
	Vegetation Condition Index	1	Demonstration	Operational
	Vegetation Productivity Index	1	Demonstration	Operational
	Dry Matter Productivity	1	Demonstration	Operational

Figure 4: Version of Copernicus GL products

2.4.3 Subproduct

An acquired product might contain more than a single variable or layer. For example, Spot-Vegetation NDVI has 2 layers: the NDVI itself and a Status Map showing a quality flag for each pixel. Furthermore, starting from the incoming product several added-value products and indicators can be generated, such as long term statistics and anomalies. Therefore, we decided to have a 'two-level' identification approach so that more than one **subproduct** can be associated to the same **product**. More specifically, for each product on the EStation there will be:

- One 'native' product, which has exactly the same name of the product with the '_native' suffix, and refers to the 'incoming' EO and climate product. This 'subproduct' cannot be visualized in the EStation, it is actually a 'reference' for the incoming dataset, i.e. to the files received on the EStation before their ingestion;
- One or more 'ingested' products, according to the number of layers that are extracted from the incoming dataset;
- Zero, one or more 'derived' products, computed on the EStation by the processing Service.

For example (Figure 5) the **subproducts** for 'fewsnet-rfe' (version 2.0) are displayed: a single 'native' is defined - as it is always the case – named 'fewsnet-rfe_native'; a single 'ingested' product exists, as the incoming dataset has no any ancillary information (and the rainfall estimate), while several subproducts are generated on the system.

productcode [PK] character varyir	subproductcode [PK] character varying	version [PK] character varying	defined_by character vai	activated boolean	category_id character vai	product_type character varying
fewsnet-rfe	fewsnet-rfe_native	2.0	JRC	TRUE	rainfall	Native
fewsnet-rfe	10d	2.0	JRC	FALSE	rainfall	Ingest
fewsnet-rfe	1monmin	2.0	JRC	FALSE	rainfall	Derived
fewsnet-rfe	1monmax	2.0	JRC	FALSE	rainfall	Derived
fewsnet-rfe	1mondiff	2.0	JRC	FALSE	rainfall	Derived
fewsnet-rfe	10dmax	2.0	JRC	FALSE	rainfall	Derived
fewsnet-rfe	10dmin	2.0	JRC	FALSE	rainfall	Derived
fewsnet-rfe	1moncum	2.0	JRC	FALSE	rainfall	Derived
fewsnet-rfe	1monavg	2.0	JRC	FALSE	rainfall	Derived
fewsnet-rfe	10dnp	2.0	JRC	FALSE	rainfall	Derived
fewsnet-rfe	10davg	2.0	JRC	FALSE	rainfall	Derived
fewsnet-rfe	10dperc	2.0	JRC	FALSE	rainfall	Derived
fewsnet-rfe	1monperc	2.0	JRC	FALSE	rainfall	Derived
fewsnet-rfe	10ddiff	2.0	JRC	FALSE	rainfall	Derived
fewsnet-rfe	1monnp	2.0	JRC	FALSE	rainfall	Derived

Figure 5: Subproducts existing for the 'fewsnet-rfe' product, 2.1 version.

2.4.4 Dataset

A ***dataset***, in Station's jargon, is an ensemble of images related to the same ***subproduct,*** for a given geographic extent and temporal window. The concepts of 'dataset' and 'subproduct' are close; a 'dataset' stresses the idea of having an ensemble of files generated for the same product for a given region and a given period of time.

2.4.5 Mapset

The '**mapset'** is the ensemble of information defining a map representation of a raster product, using the following characteristics:

1. Spatial Reference System (SRS) - defined through the SRID[1] and referring, by default, to the EPSG authority. It logically includes:

    1.1 Geographic Coordinate System (GCS), including Datum.
    1.2. Map Projection (if any - name and parameters).

2. Pixel size (unit, value)
3. Boundary Box (ULx/y, LRx/y) or 'Extent' or 'Origin'
4. Raster size (Xsize, Ysize)

The '**mapset**' replaces the concept of 'ROI' (existing on EStation 1.0) and offers the possibility to have all geo-referencing information in a single object, which is stored in a single table of the database, and is convenient for direct re-projection of an image from an original to a target 'mapset'.

[1] see http://en.wikipedia.org/wiki/Spatial_reference_system

mapsetcode [PK] character varying	defined charact	descriptive_name character varying
CHIRP-Africa-5km	JRC	Global Chirp 0.05 deg
CHIRP-Global-5km	JRC	Global Chirp 0.05 deg
default	JRC	Default mapset
FEWSNET-AEA-Africa-8km	JRC	Africa 8km (FEWSNET) - Native Proj since 1/1/2009
FEWSNET-AEA-Africa-8km-old	JRC	Africa 8km (FEWSNET) - Native Proj until 21/12/2008
FEWSNET-Africa-8km	JRC	WSGS Africa for FEWSNET
MODIS-Africa-4km	JRC	Africa 4km (MODIS)
MODIS-CEMAC-500m	JRC	CEMAC 500m
MODIS-Global-4km	JRC	Global 4km
MODIS-IOC-4km	JRC	IOC 4km
MODIS-UoG-4km	JRC	UoG 4km
MSG-satellite-3km	JRC	MSG disk 3km
PROBAV-Africa-1km	JRC	Africa 1 km from 10x10 tiles (VITO)
SPOTV-Africa-1km	JRC	Africa 1km
SPOTV-Africa-500m	JRC	Africa geografic 500m
SPOTV-CEMAC-1km	JRC	CEMAC 1km
SPOTV-ECOWAS-1km	JRC	ECOWAS 1km
SPOTV-IGAD-1km	JRC	IGAD 1km
SPOTV-IOC-1km	JRC	IOC 1km
SPOTV-SADC-1km	JRC	SADC 1km
SPOTV-Sahel-1km	JRC	Sahel 1km
SPOTV-UoG-1km	JRC	UoG 1km
TAMSAT-Africa-4km	JRC	Africa 4km (TAMSAT)

Figure 6: example of mapsets defined on the station

In Figure 6 the mapsets defined in the first release of Station are listed. As you can see, they have different spatial resolution (pixel size from 30m to 8 km), and they refer to the full continent or for one of the Regions, which are identified after normally the name of the Regional Economic Communities (ECOWAS, IGAD, IOC and so on).

2.4.6 Sources of datasets (or Datasource)

Datasets of the Station are retrieved from the EUMETCast dissemination (*i.e.* from EUMETCAST receiving Station), from remote servers by internet (HTTP or FTP) and from APIs (as C3S Data Store & IRI data library). Therefore, each of them represents a source of Datasets, and a set of variables have to be indicated to retrieve a dataset from them (e.g. the server URL, the access credential or the naming rules).

The 'EUMETCast' datasources are the simplest, since only few properties are essential: a unique identifier of the datasource itself (eumetcast_id) and a regular expression to match the filenames associated to a dataset (filter_expression_jrc). The service is described in full detail in the 4.1.1 paragraph.

The 'internet' datasources include some additional elements needed to access the remote servers and repositories, such as the base URL of the data provider, login credentials (username/password), the regular expression to filter the datasets names, and a time frame defined through 'start' and 'end' dates.

The 'DataStore' datasources make use of the configuration files which are added by the user understanding the data retrieval mechanism from the CDS and IRI portal. For CDS, configurations taken from the API request within their dataset retrieval portal whereas for IRI data library, configurations taken from the expert mode of specific datasets from their portal

Additional information for managing the datasource in defined in the '***datasource_description***' table, which contains elements to describe the 'incoming' files in terms of:

- File naming rules and file extension, which allows the 'get' services' to identify and manage the incoming files, e.g. by extracting the date/time of the image.
- The geo-reference, i.e. geographic coverage (e.g. global) and native mapset; note that in some cases the information on the geo-reference is coded in the incoming file itself, and there is no need to fill this field.
- The pre-processing type to be applied during the ingestion, which varies according to the file format (HDF, GTiff, NetCDF) and the tiles organization (see 4.1.4 for more detail)

2.4.7 Services

The following Services are implemented on the system (see also Figure 1):

- Get EUMETCast: to copy data from EUMETCast PC to the Station.
- Get Internet: to copy data from remote servers (any data provider available online) to the Station
- Get Data Store: to copy data from the Climate Data Store(CDS) & IRI data library portal to the Station
- Ingestion: to convert the incoming products into GeoTiff or NetCDF format, and optionally re-project to the defined 'mapset'.
- Processing: to derive from the incoming products additional indicators.
- System: to manage all background operations, such as clean the temporary files, check the metadata & sanity check, data/database synchronization between the Station components (if more than one) or other servers, database dump, system diagnostic.

The Services are executed on the Station as 'daemons', e.g. as a detached process that runs in the background: they should, in principle, be permanently activated.

The normal operations to be executed by the User to control the Services are:

- Define/verify the settings on the EStation for each of the services (*e.g.* control that the 'get' and 'ingest' for a given product is activated)
- Activate the Service from the GUI and monitor its execution.
- Optionally modify the Settings on the fly (no need of re-starting the Service)
- Optionally stop or restart a Service in case of unforeseen circumstances.

2.4.8 EStation standard format conventions

The files ingested on the Station, or generated by the processing service, present some common characteristics that are been defined to facilitate the exploitation of the Users, not only on the Station but also in third-party software. There characteristics include a common file format (GeoTiff or NetCDF), a unique convention for scaling the physical values in digital number and for encoding the 'no-data', and a list of metadata, written as 'tags' in the files.

Data coding

All raster files contain 'digital values' coded over bytes, integers or float types. In order to convert these numbers in physical quantities, the following equation is applied:

$$Phys.Value = DN * scale_{factor} + scale_{offset}$$

On the Station the following convention is (mostly) adopted:

scale_factor= 10^N when N is an integer, positive or negative
scale_offset= 0

This convention facilitates the User in understanding the contents of the raster files, while looking at them from third-party software (like QGIS).

Nodata encoding

The 'nodata' value is a numeric value used to indicate that no observations are available over some pixels of the image. It is coded in the metadata of the image. On the Station the nodata encoding is standardized, so that for the simplest data types (BYTE, INT16 and UINT16) it depends on the data type, as shown in the following table:

Data type	Nodata value
BYTE	255
INT16	-32767/8
UINT16	65536

Table 1: Nodata default coding for data type

For the other data types (INT32, UINT32, FLOAT32 and FLOAT64), which are by the way rarely used in the system, there is more freedom in the nodata coding, and we normally adopt the same value as in the incoming images.

List of metadata

The images in the Station's standard format contain a list of metadata that are produced during the ingestion or the processing as GeoTiff r NetCDF tags. They are visible, e.g., by doing 'gdalinfo' command followed by the name of the image.
The role is twofold:

- For the User: to extract information from the image, e.g. in order to see when it has been produced, on the basis of which input files, or what are the scaling factor and offset.

- For the Station itself: to extract information (e.g. the product, version, subproduct, mapset) needed to place the image in the correct director.

Name	Typical value	Comments
eStation_category	rainfall	Category of the product (as in the GUI, in Acquisition, Processing and Data Management Tabs).
eStation_comp_time	2015-04-23 14:17:21	Time of computation of the image (local time, taken from the PC)
eStation_conversion	Phys = DN * scaling_factor + scaling_offset	Fixed (see 'data coding' above)
eStation_date	20100101	Time of the EO and Climate product
eStation_date_format	YYYYMMDD	Format of the date
eStation_defined_by	JRC	Who has defined the product (JRC or User)
eStation_descr_name	TAMSAT RFE	Descriptive name of the product
eStation_description	TAMSAT Rainfall estimates	Longer description of the product
eStation_es2_version	2.1.0	EStation.1 version
eStation_frequency	e1dekad	Frequency of the product
eStation_input_files	/data/ingest/rfe2010_01-dk1.nc	List of input files used to generate the product. It might be a list of 'native' products (as in the example here) or EStation.1 products
eStation_mac_address	6c:ae:8b:52:77:d2	Unique identifier of the machine where the product has been generated.
eStation_mapset	TAMSAT-Africa-4km	Product mapset (see 2.4.5).
eStation_nodata	*-32768*	Nodata value coding (see above)
eStation_product_version	2.0	Product Version
eStation_provider	TAMSAT - JRC	Provider of the data. It might be a Space Agency, a Project, JRC or a RIC.
eStation_parameter	Parameters	Parameters for deriving indicatore
eStation_scaling_factor	1.0	Scaling factor for DN to physical value conversion (see above)
eStation_scaling_offset	0.0	Scaling offset for DN to physical value conversion (see above)
eStation_subdir	tamsat-rfe/2.0/TAMSAT-Africa-4km/tif/10d/	Subdirectory where the product is located on EStation, to be added to the data processing base dir (/data/processing by default).
eStation_subProduct	10d	Name of the subproduct
eStation_unit	mm	Physical Unit of the product (once the conversion to physical value is applied)

Table 2: List of standard Station metadata

2.4.9 Importing datasets

On the Station several mechanisms exist to acquire data. In normal operations, as described above, the 'Get' (from EUMETCast, Internet and Data Stores) and 'Ingestion' services feed the system (see 2.4.9.1). This is meant for near real time reception of data, not archives or historical datasets.

In fact, at the installation of the Station, normally the file system is populated with 'historical archives', i.e. datasets representing the most recent years for the various subproducts, in order to allow comparative analysis (see 3.8).

When some data are missing on the station, e.g. for an interruption in the acquisition services, it is possible to identify the gaps in the Data Management page, and create a 'request' to fill the gap (see 3.6).

2.4.9.1 Acquisition of 'native' datasets

The 'standard' mechanism is to 'get' the data from EUMETCAST receiving Station (through the 'Get EUMETCast' service) and from http/ftp servers ('Get Internet' service) and DataStore service. The files are retrieved on the machine in their original ('native') format, as generated by the data provider, and subsequently ingested. This mechanism is monitored and controlled from the 'Acquisition' interface (see 3.5).

2.4.9.2 Installation of Historical Archives

At the moment of the installation of the Station, the machine has to be fed with historical datasets in order to provide the thematic User with time series for its analysis. These datasets represent several tenths of GB, and they are a collection of files in Station format that are 'packed' into a zipped tar file (.tgz) or tiff files.

The archive has to be mounted to the Station computer on a location that corresponds to the 'Archive Dir' path specified in the system setting. Then a specific routine is activated to parse the archive and copy to the local machine the datasets of interest for the User, according to the 'Acquisition' and 'Processing' settings.

2.4.9.3 Completing datasets from Data Management page

When a gap is existing, or additional temporal windows are needed, a request can be generated from the Data Management tab (see 3.6). This request is basically a list of images missing/desired on the local machine. The request is sent to the JRC reference Station, which creates and transmit to the Station the missing files.

3. USER'S GUIDE

In the Station most of implemented functionalities are accessed through the user interface (GUI) and no integration in the hosting OS menu has been implemented. The current chapter describes all the operations that can be performed from the User Interface.

3.1 ACCESSING THE USER INTERFACE

The GUI can be opened from http://localhost:8080, each with its own IP address or hostname depending on the configurations during the installation.

- To access the user interface locally from EStation open the installed browser Mozilla Firefox and go to the following address:

  http://localhost:8080 (User interface locally on EStation)

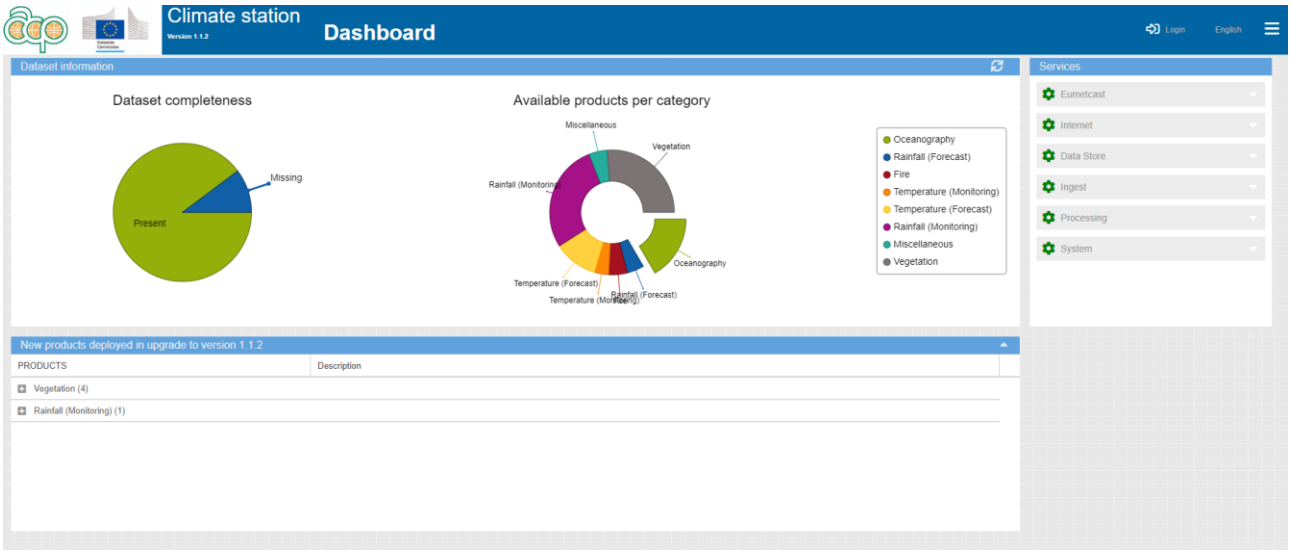


Figure 7: Dashboard of the system

3.2 MAIN MENU

The functionalities available on the system are presented in the following pages (tabs):

- **Dashboard**: presents an overview of the Station and gives control over the services;
- **Portfolio:** to configure the list of datasets in the Station (mainly to activate products for acquisition and processing)**;**
- **Acquisition**: to view and control the status of the services for retrieving and ingesting the EO and Climate data;
- **Processing**: to start and stop the processing of new products;
- **Data Management**: lists available data sets, both those acquired and processed by the Station, showing the completeness of time series, and thus allowing the retrieval of missing datasets;
- **Analysis**: to perform the data analysis and generate images for bulletins/reports.
- **Fitness for Purposes**: It offers the capability to perform comparison among different products. The comparison can be conducted in terms of statistical analysis and/or using global (or regional) maps.
- **IMPACT**: It offers a combination of remote sensing, photo interpretation and processing technologies in a GIS environment, allowing non specialist users to easily accomplish all necessary

pre-processing steps while giving a fast and user-friendly environment for visual editing and map validation.

- **Jupyter Notebooks**: It offers user, a working environment of python (ipykernel notebook) where Users can code their process to create the indicators from the existing data and visualize them.
- **System**: to control the application settings (log level, Ethernet IPs, and working directories paths), and to generate system reports;
- **Help**: to read system help files: allows downloading the pdf files and access reference web sites.

The primary role of the Station is the **Acquisition*, Ingestion and Processing service*** and ***Analysis*** of gathered and locally generated data products.

3.3 DASHBOARD

The dashboard presents an overview of a Station. It shows the status and gives control over the various services on each machine.

When opening the User Interface of the Station in a web-browser, the Dashboard page is shown first, and it is meant to be accessed for the verification of basic functionality and initial diagnostic operations.

3.3.1 Services control

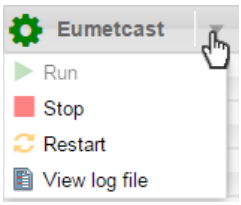
As displayed in Figure 7, there are 6 services running on the EStation:

- Get EUMETCast: Get (i.e. copy locally) data from the receiving station (EUMETCast Receiving Station) for all activated products that have an activated EUMETCast data source defined.
- Get Internet: Get data from internet sources (FTP or HTTP) for all activated products that have one or more activated Internet data sources defined.
- Get DataStore: Get data from Climate Data Store (CDS) & IRI data library for all activated products that have one or more activated data sources defined.
- Ingest: pre-process incoming raw data, by transforming the file format into the pivot format (GeoTiff or NetCDF), and optionally re-projecting.
- Processing: run all defined processing chains that are activated to derive from the incoming products additional indicators.
- System: manage all background operations, such as automatic data and database synchronization.

Services

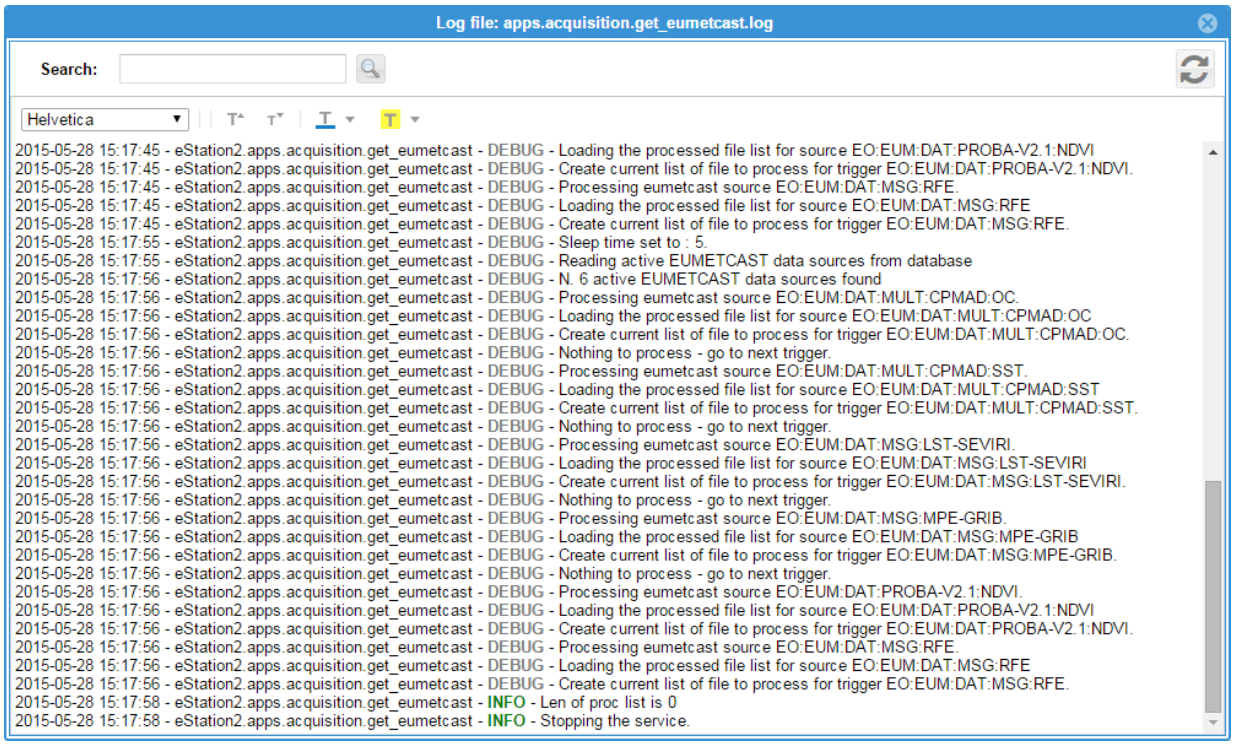
The six services (EUMETCast, Internet, Datastore, Ingest, Processing and System) can individually be started, stopped or restarted and for each of them its log file can be viewed.

- By clicking on the arrow next to the name of a service, a menu will drop down with the items Run, Stop, Restart and “View log file”.



- If the service is running, the ‘wheel’ icon is green and red when the service is not running.
- You can refresh the status of a service by clicking on the title of any service menu button.

- When choosing “View log file” of a service, a window will be shown with the content of all the log files (.log, .log1, .log2, etc...) of the service, in reversed order with the latest date on top. Searching within the log file and text highlighting functionality are available.



3.4 PORTFOLIO

After the first installation of the Station, all products coded in the application (see the Product Report for the updated list) are deactivated. The Portfolio is the place to manage the list of activated products and their ingestion and related processing chains, over a given region. It is the place where you can:

- See the list of products activated (and the related regions).
- Search for products to activate and add ingestions/processing[2] for those products, over a selected region.

When opening the Portfolio page from the main menu, you will see the **Current Portfolio** (on the right panel), showing a categorised list of all activated products and their ingestions.

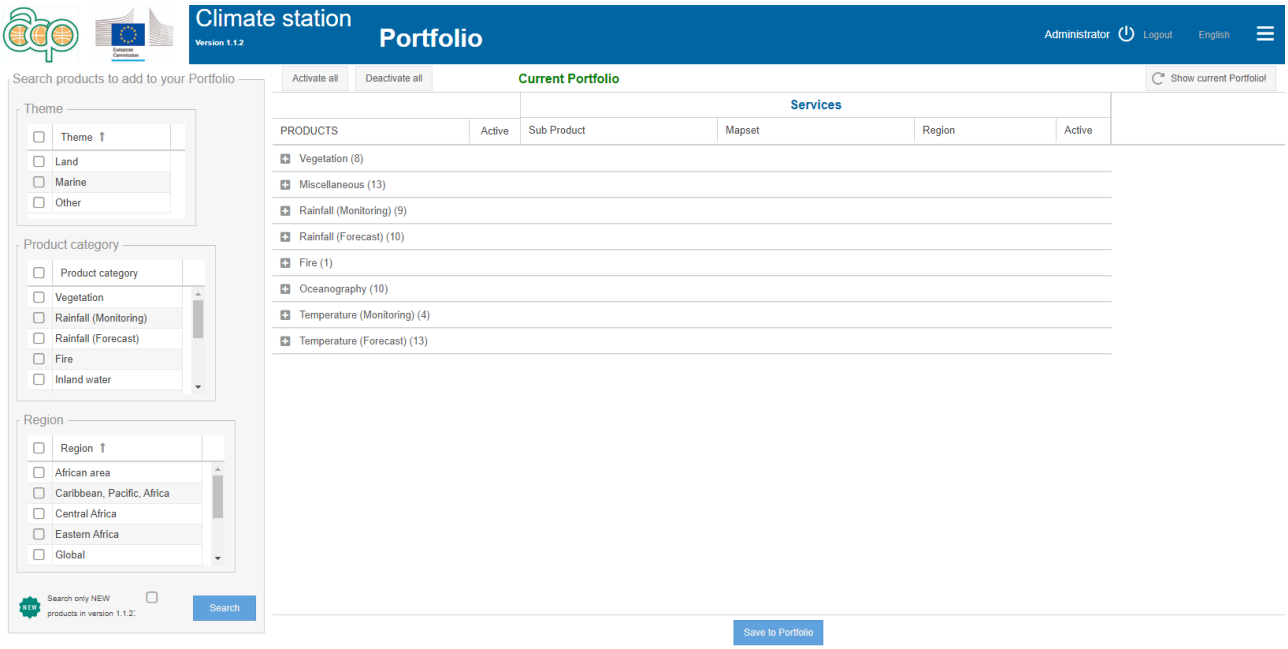


Figure 8: Overview of the Portfolio page.

On the left side you can search for products to activate. You will have to select one or more product categories to search in and one or more regions of interest. Selecting a theme (Land, Marine or Other) will select the product categories that are within the selected theme.
Once you selected a category and a region, you can click on the "Search" button, and the search result will be displayed on the right. In Figure 9 you see the result of the search for 'rainfall' products defined in the Station for either 'Africa' or 'ACP Area'.
The query returns the list of products and regions matching the selection criteria: the status of activation of the product itself and of the ingestion are highlighted in Figure 9 by the green rectangles, on the left and right, respectively. The status of activation is initially false, but

[2] The 'Portfolio' page displays only an 'ingestion' columns, but the standard processing chains are also activated for the same regions defined for Ingestion. The result is visible in the 'Processing' page.

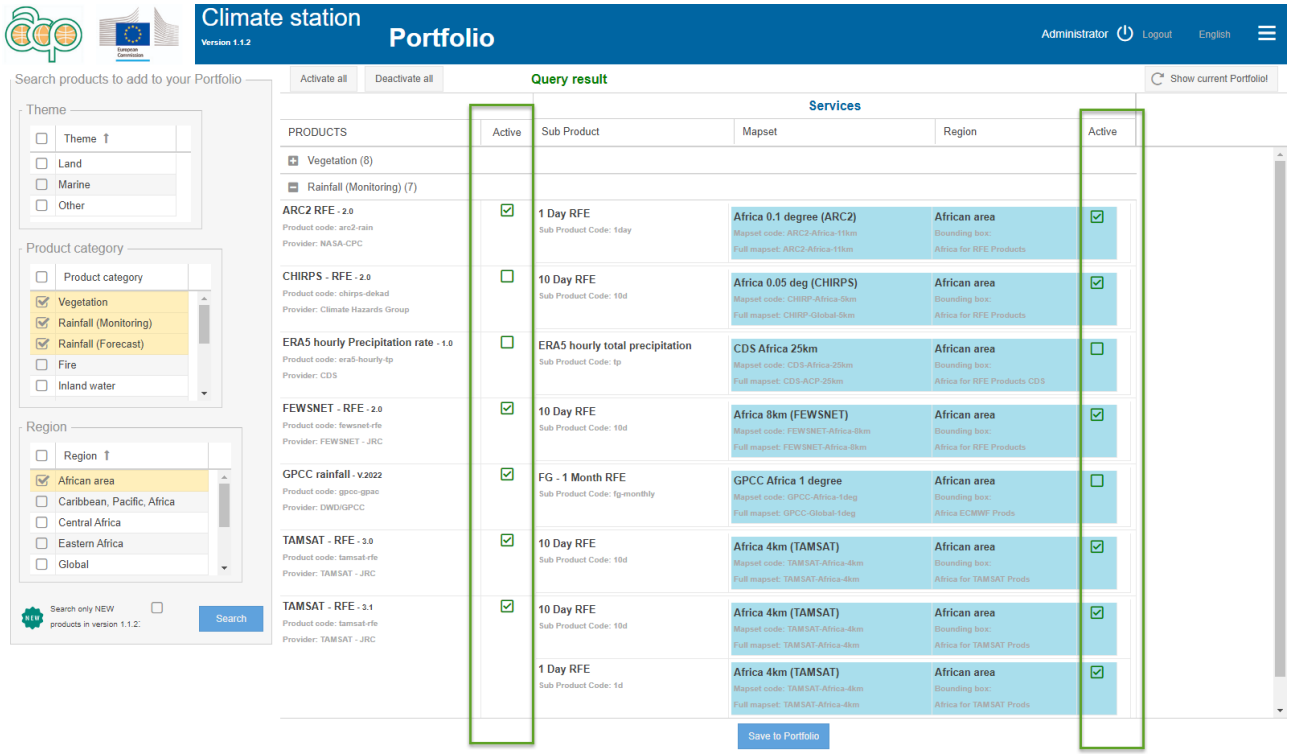


Figure 9: result of the search for products

You can modify the activation a product, or of an ingestion, by clicking on the related checkbox. You can also activate or deactivate all products and ingestions that are in the query result by clicking on the “Activate all” or “Deactivate all” buttons.

When you are happy with your selections, you can save the result in your portfolio by clicking on the “Save to Portfolio” button. When saved, a confirmation message appears and your current portfolio is shown.

3.5 ACQUISITION

Under 'Acquisition' tab two main services are merged:

1. The **Get** data, with its three implementations (from EUMETCast, Internet and Data Stores).

2. The **Ingestion**, i.e. the pre-processing that ingests/converts the incoming raw data for a product to the pivot (GeoTiff or NetCDF) format, and optionally re-projects and subsets them to a specific defined 'mapset'.

The Acquisition tab shows all activated products, grouped by product categories, and allows the user to:

- See and control the status of the 4 services (**Eumetcast, Internet, Data Store** and **Ingest)**
- Check the status of completeness of the ingested datasets
- Activate or deactivate each individual Get and Ingest defined for a product
- Display the log files associated to each individual Get and Ingest.

As advanced features, the acquisition page also gives the user the possibility to:

- Activate or deactivate a product
- Add a new User defined product
- Assign an existing or newly created Get and Ingest definition.

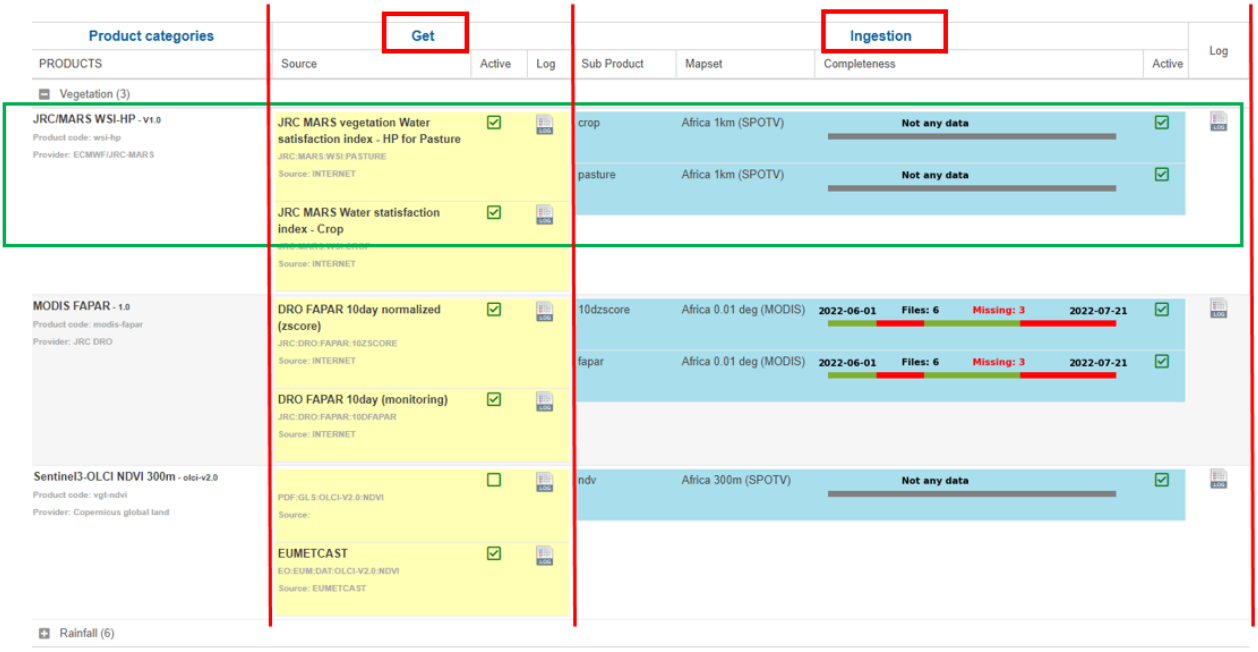


Figure 10: Organization of the Acquisition tab

As if can be seen in Figure 10, the Acquisition display is organized as a table, with a product on each row (see JRC/MARS WSI-HP on the first one, highlighted in green) and the 'Get' and 'Ingestion' attributes in adjacent columns (left to right – highlighted in red).

3.5.1 Functionality

The Acquisition page allow to fully control both get and ingest, ranging from basic operations to more advanced one. The basic features are accessible from the page as it originally displays, while the advances requires to 'unlock' the page to access additional features.

Unlocking is by clicking on the  icon, at the left top corner (see Figure 10).

Basic Operations in Acquisition

The following operations can be executed without 'unlocking' the page and entering the advanced mode.

Control the status of the individual Services and their log files

- Like in the Dashboard, the current status of the services is **Eumetcast, Internet, Data Store** and **Ingest** is displayed, and can be individually changed (started, stopped or restarted) and the related log file can be viewed.

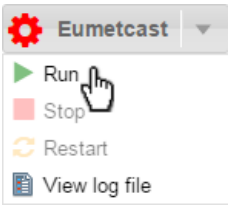


Figure 11: Controlling a Service

See the list of active products

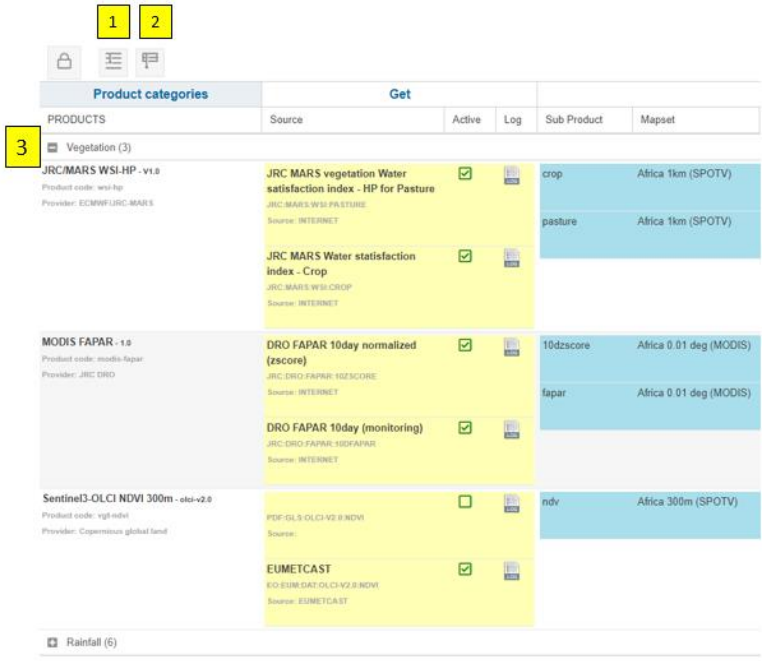


Figure 12: Expand/Collapse products and categories

The products are grouped in categories, and the User can change their visualization by:

- Expanding all product categories to see all their activated products by clicking on the "**Expand all**" button (icon near [1] in Figure 12).
- Collapse all product categories by clicking on the "**Collapse All**" button (icon near [2] in Figure 12).

- Expand or collapse each product category individually by respectively clicking on categories title bar or on the + and – sign on the left of the title of a product category (icon near [3] in Figure 12).

  The number between the brackets next to the title/name of a product category indicates the amount of active products within that product category.

- To refresh/reload the Acquisition page click on the refresh button , at the top-right corner of the page.

<u>Operate on the single product</u>

As described above, the Acquisition tab allows controlling 'get' and 'ingest' of the products. Figure 13 represent the ARC2 RFE product, version 2.0, for which we get the '1day' subproduct over the Africa 0.1 region.

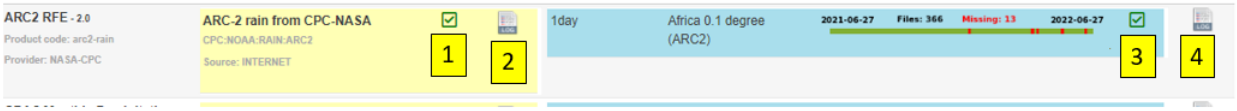


Figure 13: Operate on the single product

From the Basic interface, the User can:

- Activate/deactivate a single 'Get' source:

  The check-box near the [1] in Figure 13 represents the status of activation, and can be changed simply clicking on it.

- Activate/deactivate a single 'Ingestion':
  The check-box near the [3] in Figure 13 represents the status of activation, and can be changed simply clicking on it.

- View the log file for an individual 'Get' or 'Ingest' defined for a product:
  The LOG icon near the [2] (for Get) and the [4] (for Ingest) numbers. A new window will be shown with the content of all the log files (.log, .log1, .log2, etc...) of the Get or Ingest, in reversed order with the latest date on top (see Figure 14).

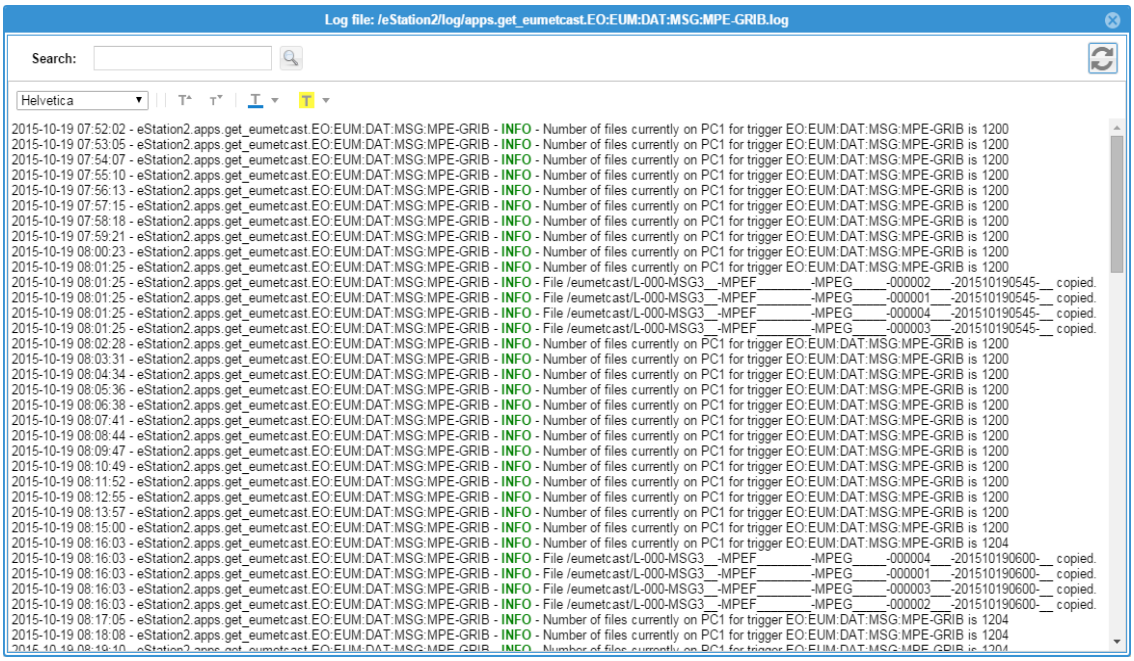
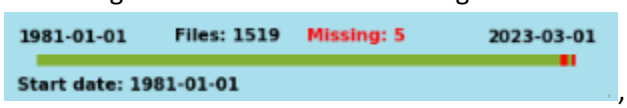


Figure 14: log file for a get EUMETCast source

- Check the status of completeness of ingested datasets. For each 'Ingestion' the Acquisition page shows a completeness chart 

, indicating the first date and the expected last date of the dataset, the total expected files for the dataset and the total of missing files.

By clicking on a dataset completeness chart, a list of all the periods of present, missing and permanent missing files pops up (see Figure 15)

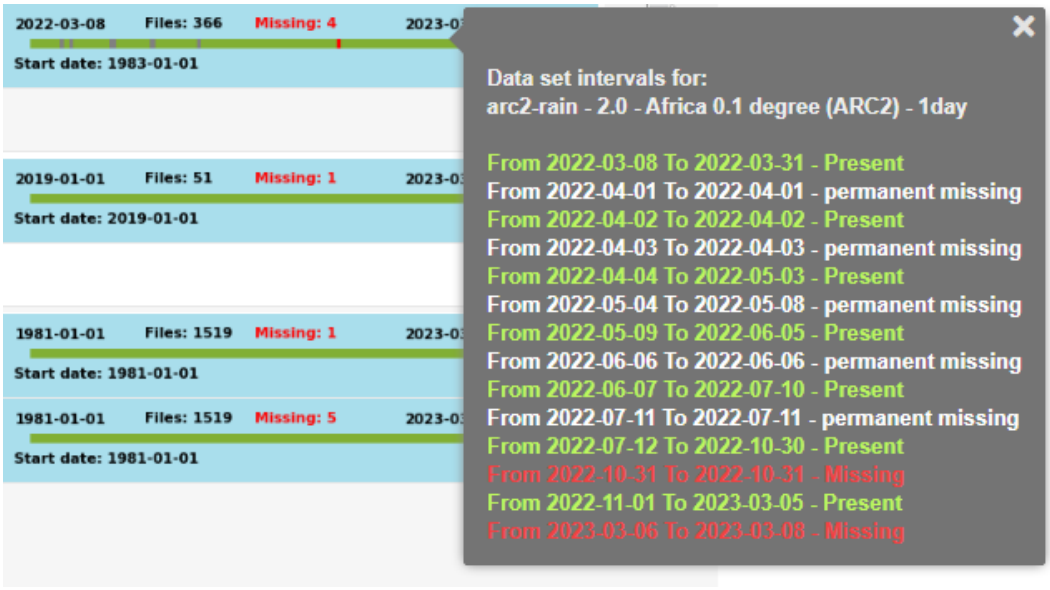
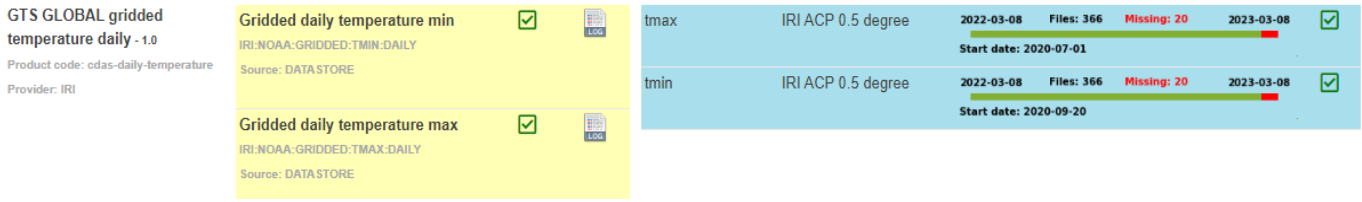


Figure 15: Detail of the completeness chart view

NOTE: for products with a high frequency, for example every 15 or 30 minutes, the completeness chart is created for a shorter period (see example below, with the interval limited to the most recent year for a daily product).



Advanced Operations

In order to perform advanced operations, the User has to unlock the Acquisition page. The icon in the top-left corner indicates that the Acquisition page is locked. Click on the icon to unlock the page. Hidden functionality and information will be shown and the lock icon becomes an unlocked icon (see Figure 16 – top left).

The following paragraphs present the additional features accessible on the unlocked page.

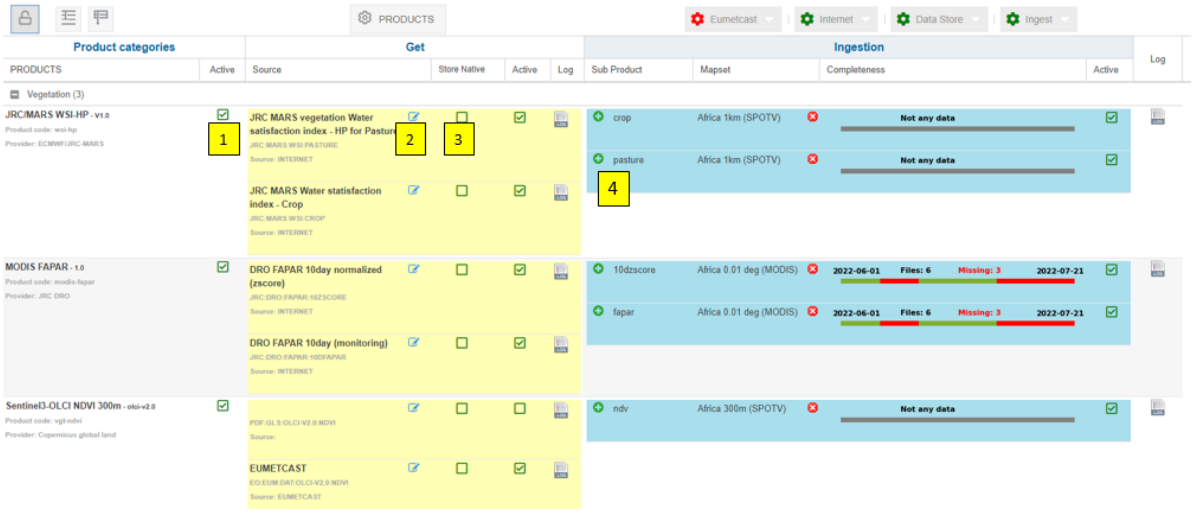


Figure 16: Unlocked Acquisition tab

Change treatment of a product

- Deactivating a product.
  The User might decide that an initially defined product is of no interest anymore, and decide he wants to stop its acquisition and processing. By unchecking the icon under the second leftmost column (see [1] -Figure 16) the product is removed from the Acquisition list, and will disappear from the current page. The Station services will stop treating this product (including processing) but the files in the local file system are not deleted, and still visible under Data Management and in the Jupyter Notebooks.

- Change the storing of native products.
  As previously described, the various datasets are retrieved to the station in their 'native' format, and converted to GeoTiff or NetCDF by ingestion. The original files might be kept, or deleted, at the User's choice. By default this is activated for the products having a low temporal resolution (10 days).

The icon near the [3] in Figure 16 means the storing of the native data is activated. The icon means the storing of the native data is deactivated.

Change the properties of the data-sources associated to a product.

As explained in Section 2, the 'data sources' define all properties for retrieving datasets from external sources to the Station. A number of sources are defined by JRC for the distributed portfolio, but the User might want or need to modify some of them, e.g. for a change in the filenames (or credentials) done at the data provider level.

- Change the settings of a EUMETCAST data source. Click on the icon next to a EUMETCAST data source (see [2] in Figure 16) and a window will be opened where its settings can be changed.

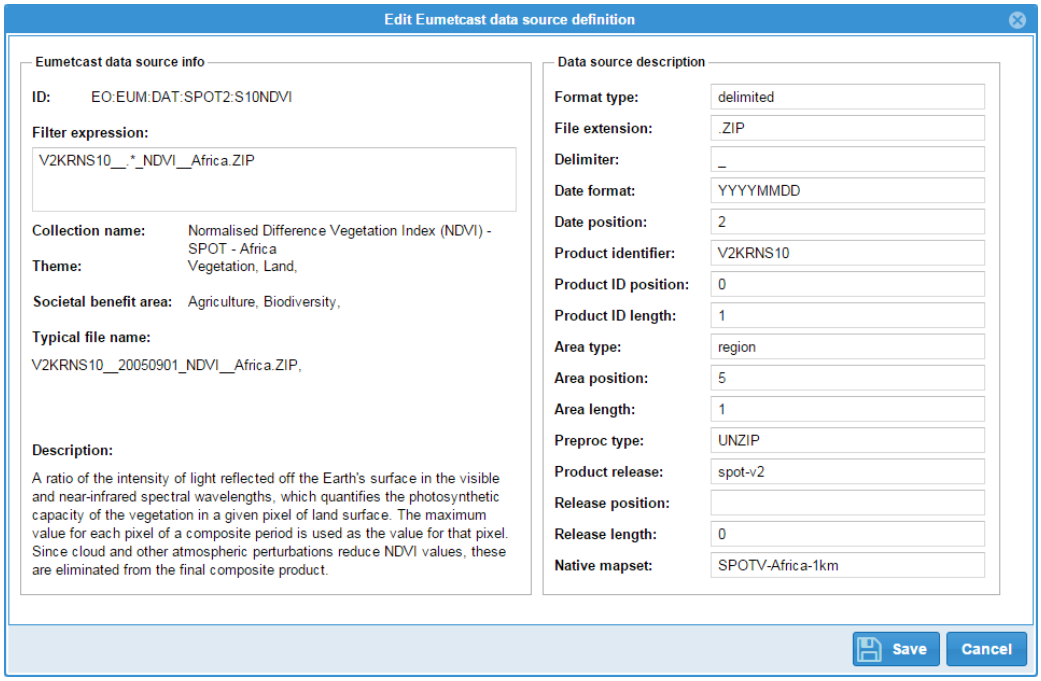


Figure 17: Eumetcast data source definitions

Click on the "Save" button to save any changes made.

- Change the settings of an INTERNET data source. Click on the icon next to an INTERNET data source (see [2] in Figure 16) and a window will be opened where its settings can be changed.

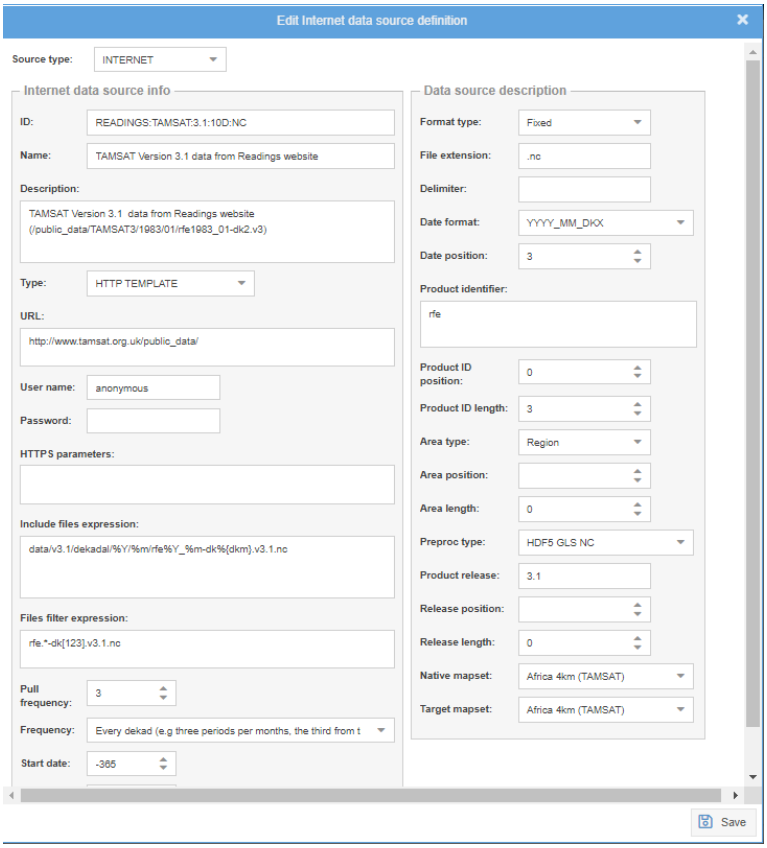


Figure 18: Internet data source definitions

Click on the "Save" button to save any changes made.

Change the Mapset (ROI) of an ingestion

In some cases, you might want to ingest an existing product over an additional region of interest, or modify the existing one.

- Adding a new Mapset[3] for a Sub Product, to be ingested:
  Click on the icon next to a Sub Product (see [4] in Figure 16) to open a new window where a Mapset can be selected.

[3] See 2.4.5 for the definition of Mapset.

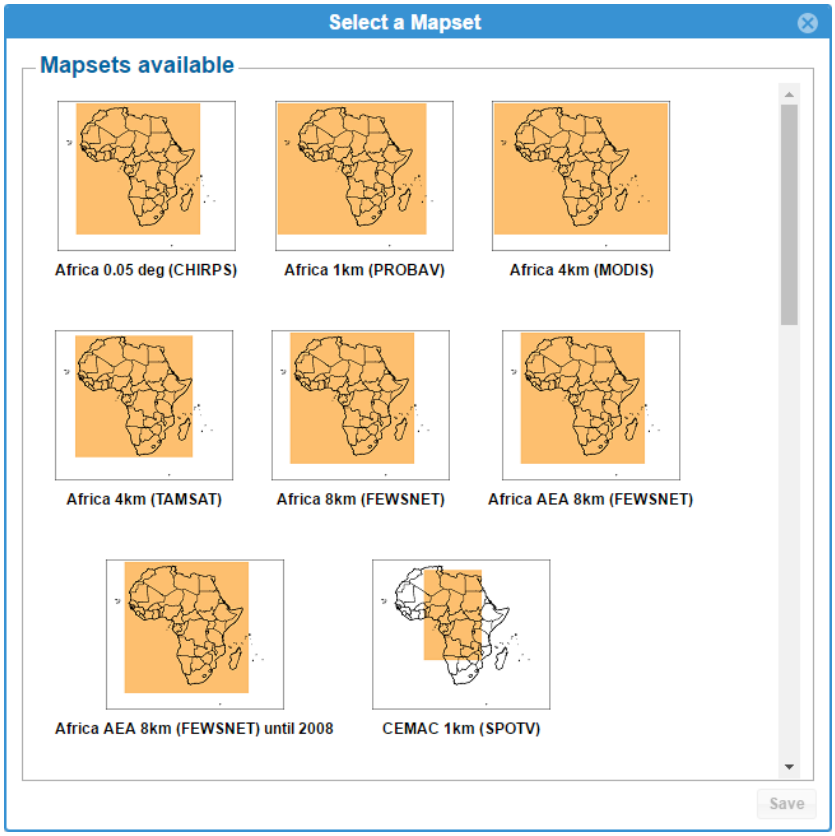


Figure 19: list of available datasets

Select a Mapset and click on the "Save" button. The window will be closed and the selected Mapset will be added to the Acquisition page as a User defined Mapset ingestion for the sub product in question.

Click on the delete icon to delete the Mapset from the Sub Product. A confirmation window pops up, asking to confirm the deletion.

Add a product from the predefined list

- All products listed in the Acquisition page are 'Activated' products, already defined in the system with assigned data sources and ingestions. For most products, especially those broadcasted through Eumetcast, the JRC has defined their definition, their GET (data source) and Ingestion definition and assignment. To activate an already defined deactivated product, click on the PRODUCTS button. A window will be shown with a categorized list of deactivated products.

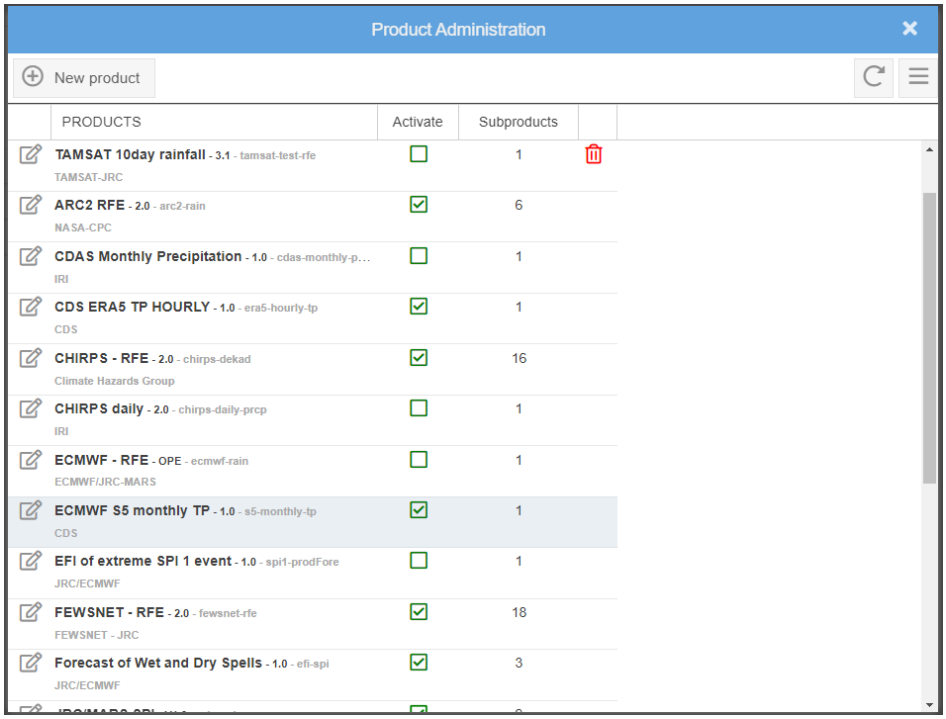


Figure 20: Product Administration view

The ☐ icon means the product is deactivated. Click on the ☐ icon to activate the product. The product will be added to the list of Active products in the Acquisition page.

Close the Product Administration window by clicking on the “Close” button.

NOTE: The current presentation is not exhaustive of all features of the Station. For the more complex operations, video tutorials will be released to help communication to the Users.

3.6 DATA MANAGEMENT

The data management page is intended for having an overview and control over the completeness of the datasets existing on the EStation. The datasets shown are grouped by category, as in the previously discussed tabs, and all products 'enabled' on the station are displayed, together with their ingested and derived subproducts. It is important to note that here also the locally derived subproducts, generated by the Processing service, are represented, unlike in the Acquisition tab.

The data management page gives the possibility to complete datasets with missing files, by creating a request job that will automatically download the missing data from a cloud server managed by the JRC.

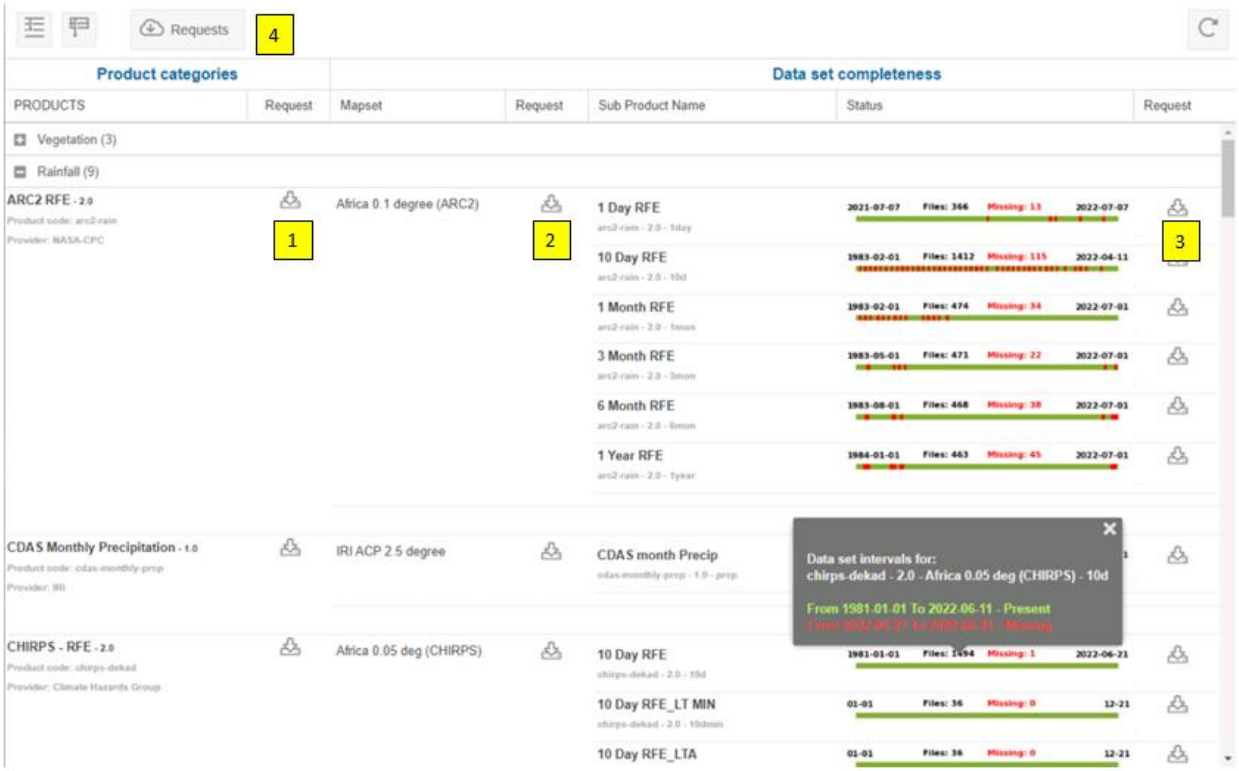


Figure 21: Data Management View

3.6.1 Functionality

As in Acquisition, the products are grouped by category, and the User can:

- Expand or collapse each product category individually by respectively clicking on the + or – sign on the left of the title of a product category.
- Collapse all product categories, clicking on the button and to expand all product categories, click on the button.
- Refresh/reload the Acquisition page, clicking on the refresh button.

There are three possible ways to create a request job to complement the local data (see Figure 21):

a. Create a request to complete all datasets of a product [1].
b. Create a request to complete all datasets of a mapset of a product [2].
c. Create a request to complete a single dataset [3].

When you click on one of these download icons a popup window is shown, listing all the datasets with the number of missing files that will be in the request.

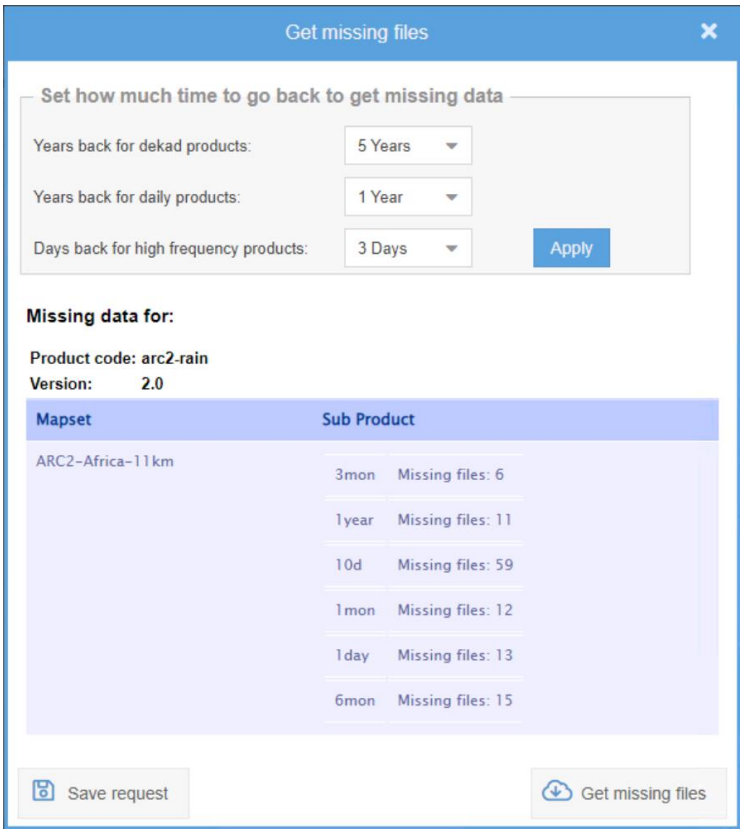


Figure 22: List of datasets for a Request

When your Station is offline, you can click on the **'Save Request"** button to save/download the request file. Once downloaded, send the file by email to JRC-ESTATION@ec.europa.eu. The EStation team will generate and send you an archive with the missing files, stated in the request file.

You can directly download the missing files by clicking on the **"Get missing files"** button. The missing files will be downloaded from a cloud server, which is maintained by the JRC.

When there is no internet connection or your EStation is in a network with a proxy, a message will popup stating:

> *"Error connecting to the server, please check if your network is connected to the internet or uses a proxy. Set your proxy settings under the system page!"*

If your EStation is connected to the internet and, if necessary, you have set your proxy settings under the system page, a request job is created that starts downloading the missing files.

After the creation of a new Request, the administration tool[4] will be shown, as in Figure 23.

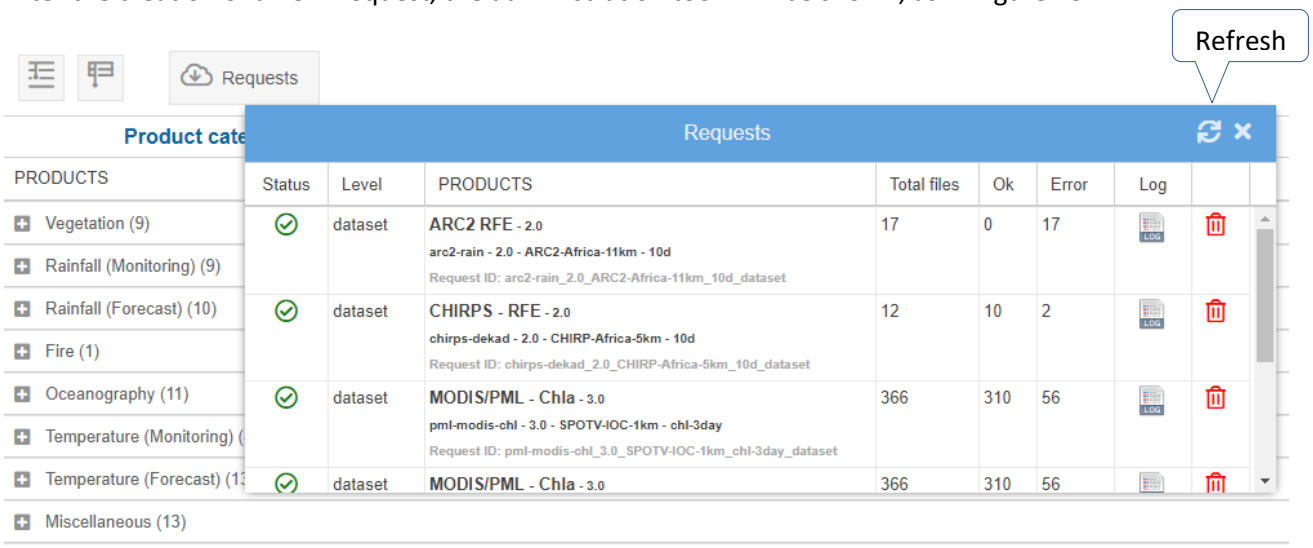


Figure 23: Requests' administration tool

The status of all requests is listed and each request can be:

- Finished
- Running

To see the log file of a request, click on its LOG icon.

To delete a request job, click on its icon.

Refresh the list to update the totals (see figure above).

Pausing and restarting a request job, might give an error message because of no internet connection. The request job will have a red play icon, indicating an error, and the job can be restarted.

- Check the status of completeness of datasets. For each dataset the Data management page shows a completeness chart 

, indicating the first date and the expected last date of the dataset, the total expected files for the dataset and the total of missing files. By going over a dataset completeness chart with the mouse pointer, a list of all the periods of present, missing[5] and permanent missing files pops up:

[4] The administration tool can be opened at any time by clicking on the icon near [4] in Figure 21

[5] An image might be missing for various reasons, including the station outage periods. Though, it is normal the latest (or two latest) datasets are displayed as missing.


2021-07-07 Files: 366 Missing: 13
1983-02-01 Files: 1412 Missing: 115
1983-02-01 Files: 474 Missing: 34
1983-05-01 Files: 471 Missing: 22
1983-08-01 Files: 468 Missing: 38
1984-01-01 Files: 463 Missing: 45
Data set intervals for:
arc2-rain - 2.0 - Africa 0.1 degree (ARC2) - 1day
From 2021-07-07 To 2022-01-04 - Present
From 2022-01-05 To 2022-01-06 - Missing
From 2022-01-07 To 2022-03-31 - Present
From 2022-04-01 To 2022-04-01 - Missing
From 2022-04-02 To 2022-04-02 - Present
From 2022-04-03 To 2022-04-03 - Missing
From 2022-04-04 To 2022-05-03 - Present
From 2022-05-04 To 2022-05-08 - Missing
From 2022-05-09 To 2022-06-05 - Present
From 2022-06-06 To 2022-06-06 - Missing
From 2022-06-07 To 2022-07-04 - Present
From 2022-07-05 To 2022-07-07 - Missing

3.7 PROCESSING

The processing page is the interface to the 'processing' service, e.g. to the generation of derived products from the locally available datasets.

The processing page allows the User to control the status of the processing service, and to activate/deactivate a single processing 'chain'. All processing chains are created and defined by the JRC. In the processing page only the enabled processing chains are shown. The relevant processing chains are automatically enabled and activated based on the products activated from the portfolio.

As illustrated in the figure below, a processing 'chain' is based on an Algorithm ('Type + Options' in the middle section) and has one or more subproducts as Input (left section) and one or more derived subproducts as Output (right section). For a description of the different algorithms see 4.1.6.

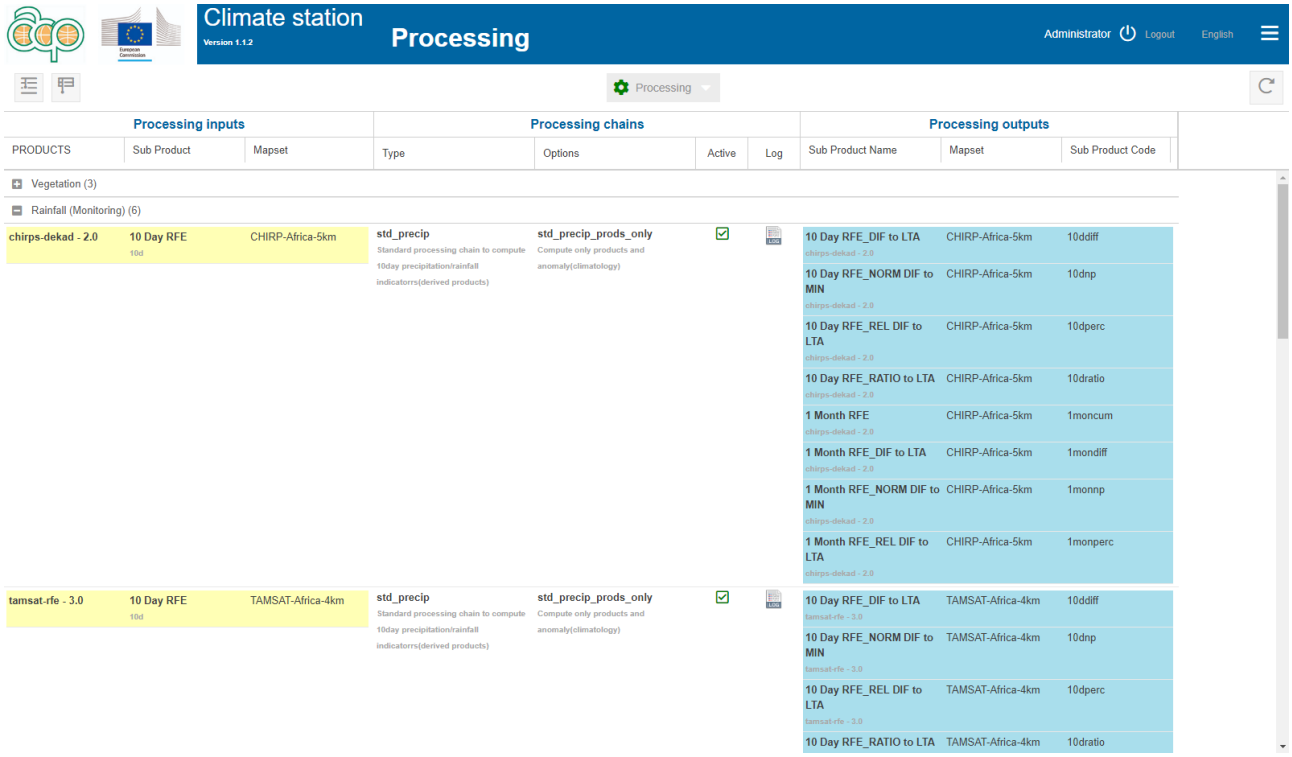
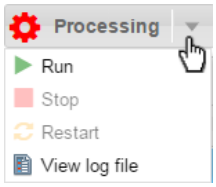


Figure 24: Processing view

3.7.1 Functionality

- Expand all product categories to see all the products with one or more processing chains defined, by clicking on the button and to collapse all product categories by clicking on the button.
  1. Expand or collapse each product category individually by respectively clicking on categories title bar or on the + and – sign on the left of the title of a product category.

  2. Like in the Dashboard, the Processing service can be started, stopped or restarted and its log file can be viewed. By clicking on the arrow next to the title, a menu will drop down with the items Run, Stop, Restart and "View log file".



If the service is running, then the cog icon is green and red when the service is not running.

You can refresh the current status of the Processing services by clicking on the title of the service menu button.

3. To refresh/reload the Processing page click on the refresh button.
4. Activate/deactivate a processing chain.

   The icon means the processing chain for the derived sub products is activated. Click on the icon to deactivate the processing.

   The icon means the processing chain is deactivated. Click on the icon to activate the processing chain.

5. View the log file for an individual processing 'chain', click on the LOG icon. A new window will be shown with the content of all the log files (.log, .log1, .log2, etc...) of the processing 'chain', in reversed order with the latest date on top.

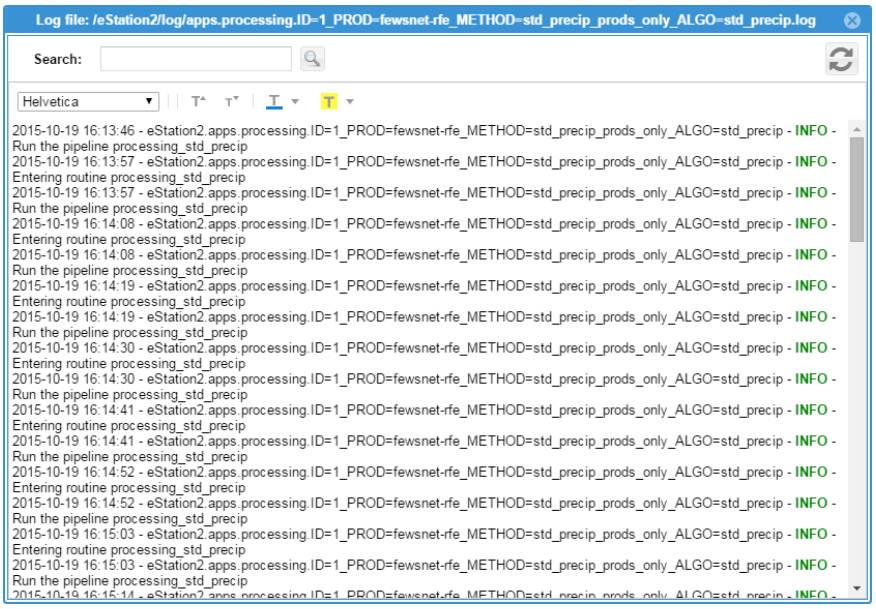


Figure 25: Processing service log file of specific chain

3.8 ANALYSIS

The Analysis tool is the entry point for the EStation data visualization and analysis.

It is a user-friendly interface to display any available products:

- as map over space and time with appropriate legend and vector overlay or
- as graphs to visualise the time series data aggregated from any polygon of interest.

The registered user can save any MAP and GRAPH composition as templates. A powerful functionality is the creation of a WORKSPACE where the user can display multiple maps and graphs and save the multiple object combinations as workspace templates and retrieve them at a later stage.

The management of the different objects is made through five components:

- Workspace
- Map
- Graph
- Legend
- Layers

3.8.1 Workspace functionalities

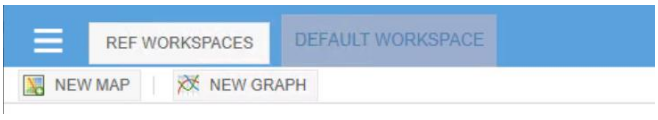
To facilitate the analysis and interpretation of the EO data a new way to organise the different maps and graphs has been introduced through the development of a workspace environment. This work panel can be populated with various existing or new maps and graphs. The user can distribute and resize the different objects as needed and then save this workspace as a template. The user can generate various workspace templates dedicated to any specific thematic and focus on geographic area.

It is recommended to log in to be able to save workspaces, maps, graphs, colour palettes and layers.

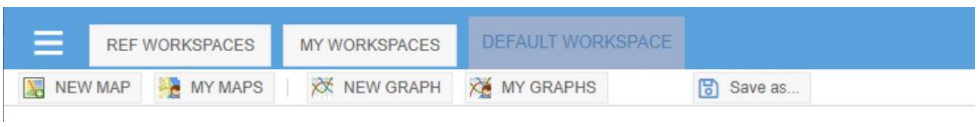
3.8.1.1 Workspace interface description

When opening the Analysis page, you will see a blank page of the “DEFAULT WORKSPACE”, with on the top a toolbar.

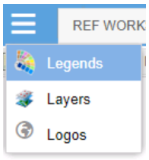
For a not logged in user:



For a logged in user:



1. When logged in, the top tab bar contains a menu and the two buttons “REF WORKSPACES” and “MY WORKSPACES” and a tab for the “DEFAULT WORKSPACE”.
   a. The “REF WORKSPACES” button opens a list of Reference Workspaces from which you can choose and check it to be opened within the “DEFAULT WORKSPACE”.
   b. The “MY WORKSPACES” button opens a list of your saved Workspaces, where you can manage your Workspaces. You can select one or more Workspaces and open them.
   c. The menu gives access to the Layers, Legends, and Logos administration tools.



   - The “Legends” item opens the legend administration window to copy, edit, delete any colour palette and associated annotation attributes.
   - The “Layers” item give access to the layers administration interface to add, delete, display in a menu/submenu, and define the layer to load by default.
   - The “Logos” item opens the logos administration window to add logos and administer the present logos.

2. Below there are the buttons “NEW MAP”, “MY MAPS”, “NEW GRAPH”, “MY GRAPHS” and “Save as”.
   - The “NEW MAP” button opens a Map window from where the user can visualize any available subproducts and make a composite with a specific colour palette, window and map size, vector layer, title, disclaimer, logo, and scale, digitize polygon/line/points and export the composite as PNG file. Any new map window will be by default align to the background layer scale.
   - The “MY MAPS” button opens a list of the different map templates prepared and saved by the user, which you can open in any workspace.
   - The “NEW GRAPH” button opens the interface to display the time series interface to generate various type of graphs showing the aggregated subproducts over a period of interest from any selected polygon/point into the map viewer. The operator can define interactively the subproducts, period, window dimension, colour, scale, font size, title and logo and export the composite as PNG or the resulting aggregated data set as EXCEL file.
   - The “MY GRAPHS” button opens a list of the different graph templates prepared and saved by the user, which you can open in any workspace.
   - The “Save as” button gives the logged in user the possibility to save the “DEFAULT WORKSPACE” under a different name.

3.8.1.2 Adding a map or graph object in a Workspace

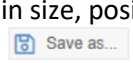
When you want to create a new workspace, start by using the “DEFAULT WORKSPACE”.

You can add a new map by selecting “NEW MAP” icon and create your own map window or select an existing template in “MY MAPS”.

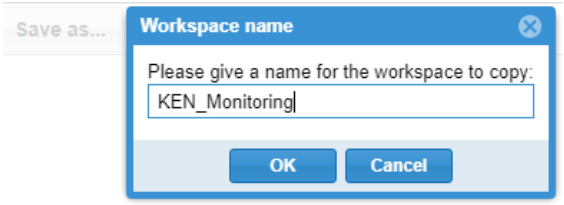
You can also add a new graph by clicking on “NEW GRAPH” or add an existing Graph template from “MY GRAPHS” button.

In your workspace, you can add several maps and or graphs considering that the limiting factor is the screen size and resolution.

3.8.1.3 Saving a Workspace

Any change in size, position, content of the different maps and graphs objects can be saved in a Workspace. Click on the 

button and enter the name you want to give to the template then click OK.



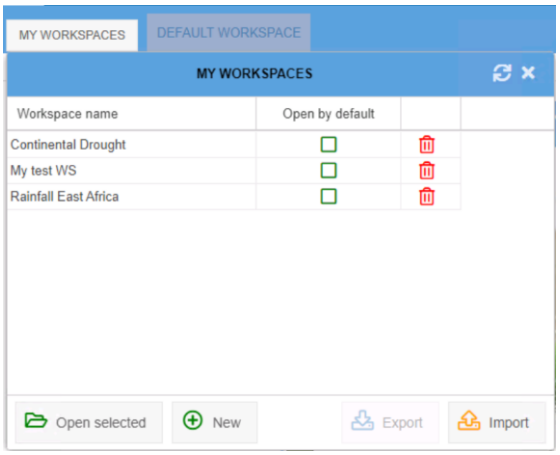
A message “Workspace created!” will be displayed.

When saving the template, you do not update the map and graph templates possibly opened to populate the workspace. They are considered as graphic objects composition specific for each workspace.

You can find and open your saved Workspace under “MY WORKSPACES”.

3.8.1.4 Opening/deleting a Workspace

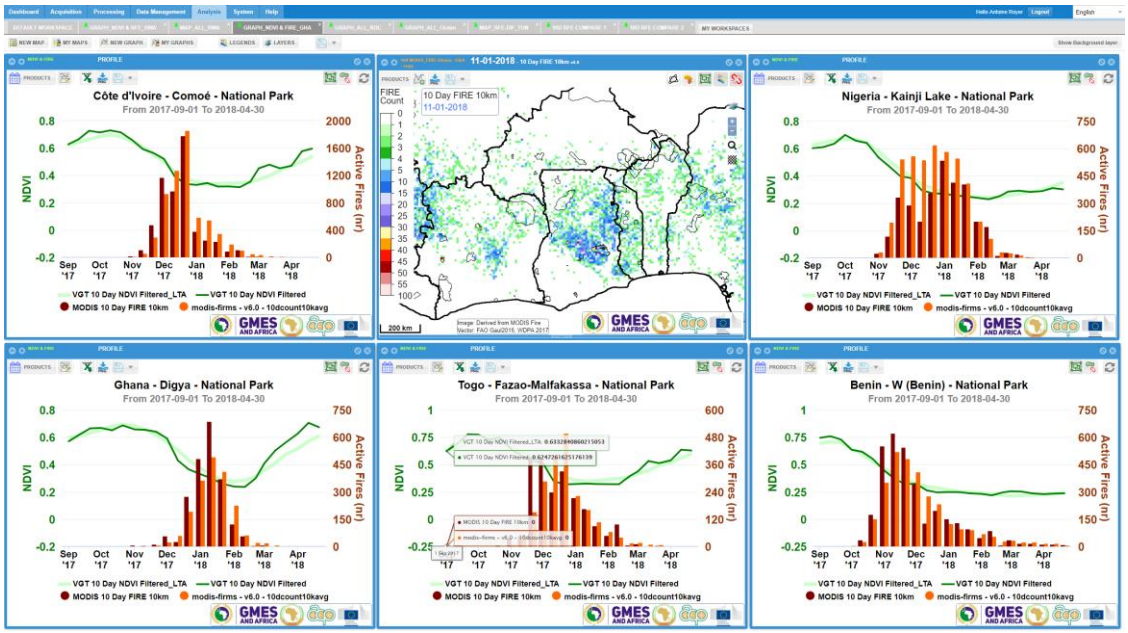
Each workspace is saved within the specific user session. Any saved workspace can also be retrieved within the same user session at a later stage. Click on “MY WORKSPACES” select the workspace of interest and use the button “Open selected” to open and display the workspace.

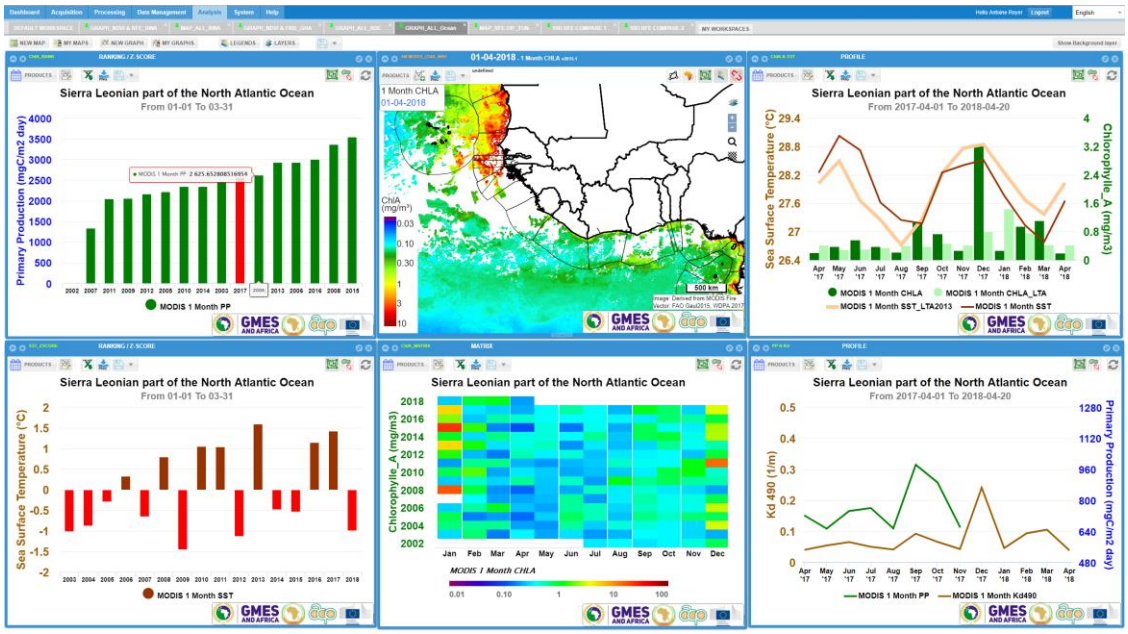


If you want to delete a workspace, you must click on the corresponding red bin icon.

You can also export selected workspaces, downloaded in a .json file, which then can be imported on the same or on another EStation, under a different user.

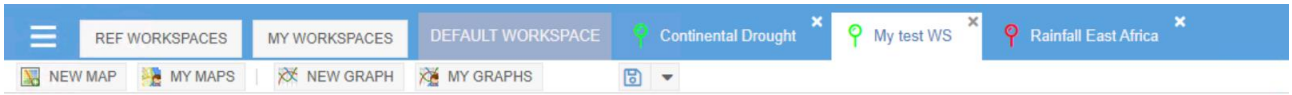
Two workspaces examples:





3.8.1.5 Pin/Unpin a Workspace template

You can pin (in GREEN) or unpin (in RED) a workspace template by double-clicking on the pin icon of a loaded workspace (see below). Any pinned workspace (with a GREEN pin) will be loaded when logged in with the same user session.



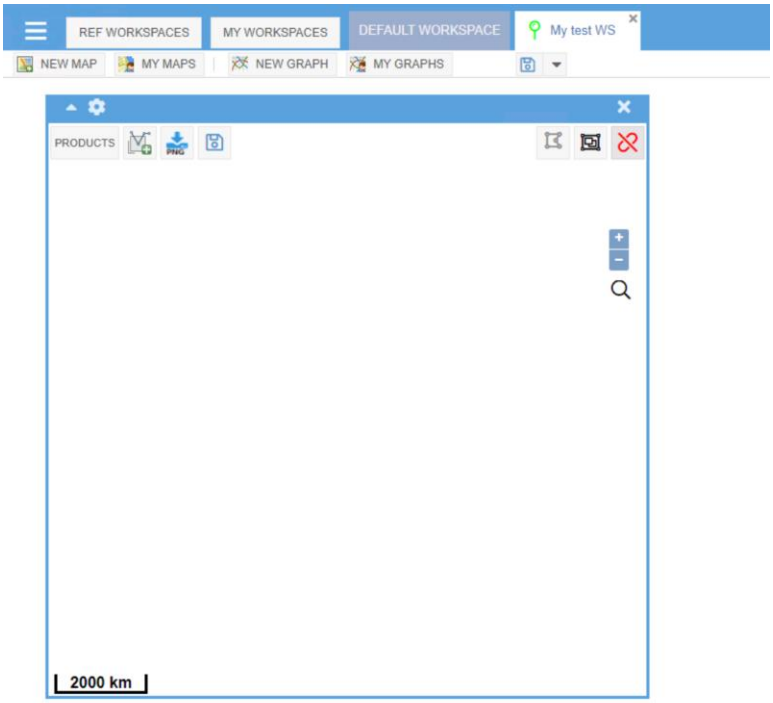
You can select any workspace to display its content. The graphs and maps in a workspace will be updated with the latest data when opening the workspace.

By double clicking on the workspace tab name, you can edit the workspace name which is saved automatically.

3.8.2 Mapview functionality

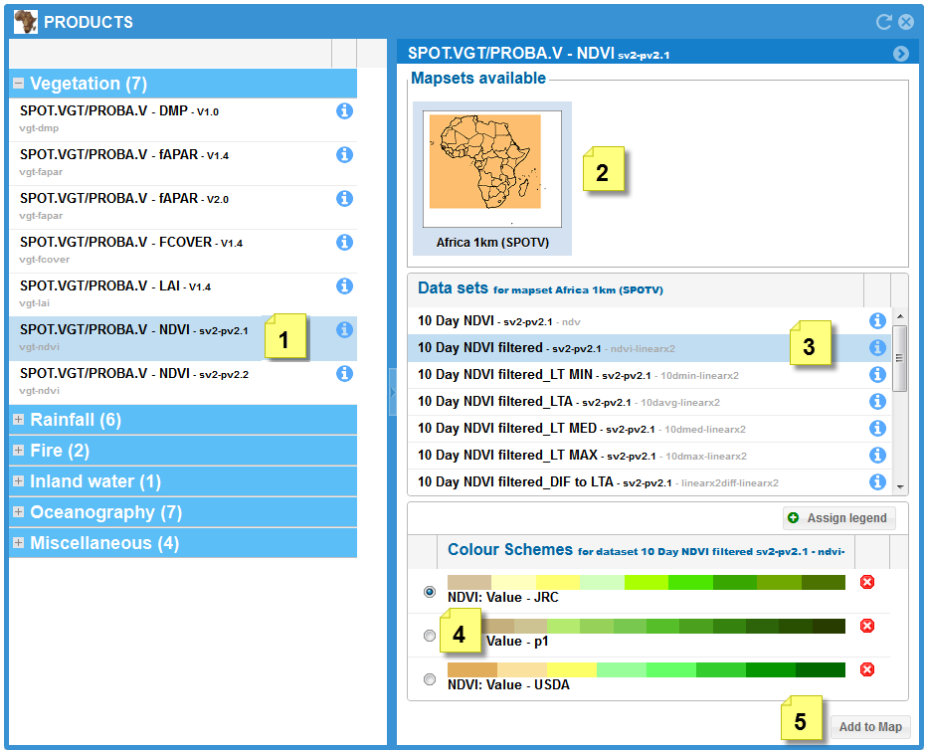
3.8.2.1 Viewing a product in a Mapview

Open a new Mapview window by clicking on the “NEW MAP” button. A new empty Mapview window will be opened.



To view a Product in a Mapview a product has to be chosen from the Product Navigator.

From each Mapview the Product Navigator can be invoked by clicking on the “PRODUCTS” button in the Mapview.



A product navigator window will be opened showing the list of available products for each category (1). Click on the product you are interested in.

On the right of the available products list, the available Mapsets for the chosen product are shown (2).

Select a Mapset and you will see the list of available sub products (3).

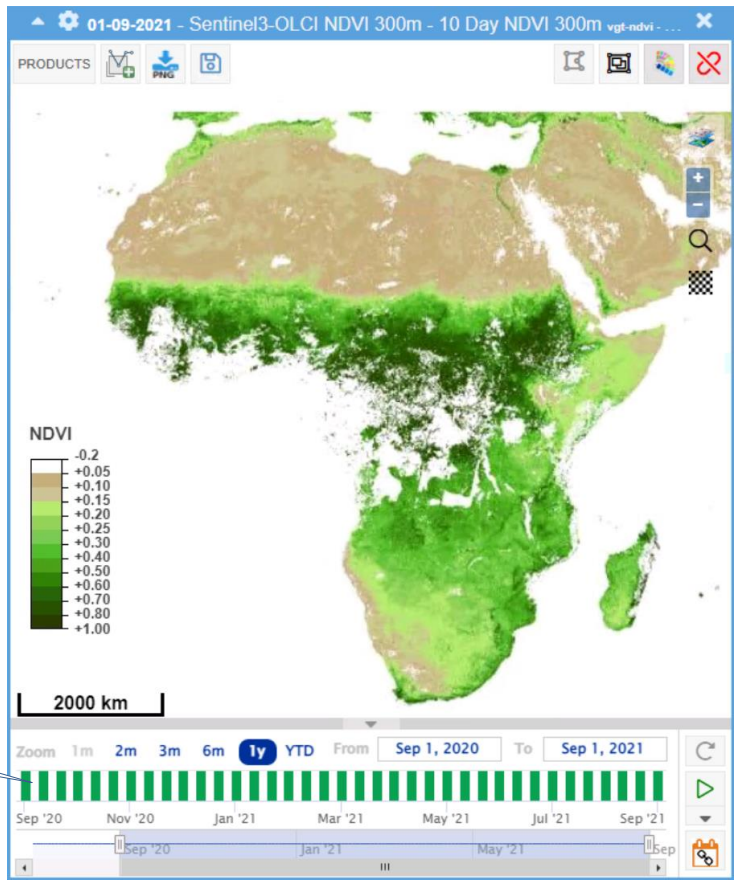
Click on the sub product you are interested in to see its defined colour schemes (4) with the default colour scheme selected.

Choose the colour scheme you prefer and finally click on the “Add to map” button.

Steps to follow:

1. Select a product
2. Select a mapset
3. Select one of the available sub products
4. Chose a colour scheme (legend)
5. Click on the “Add to map” button

The Product navigator will close, and the last available date of the selected product dataset will be shown in the Mapview window with the legend colour scheme and on the bottom of the Mapview its available timeline.



Product timeline

3.8.2.2 Product timeline

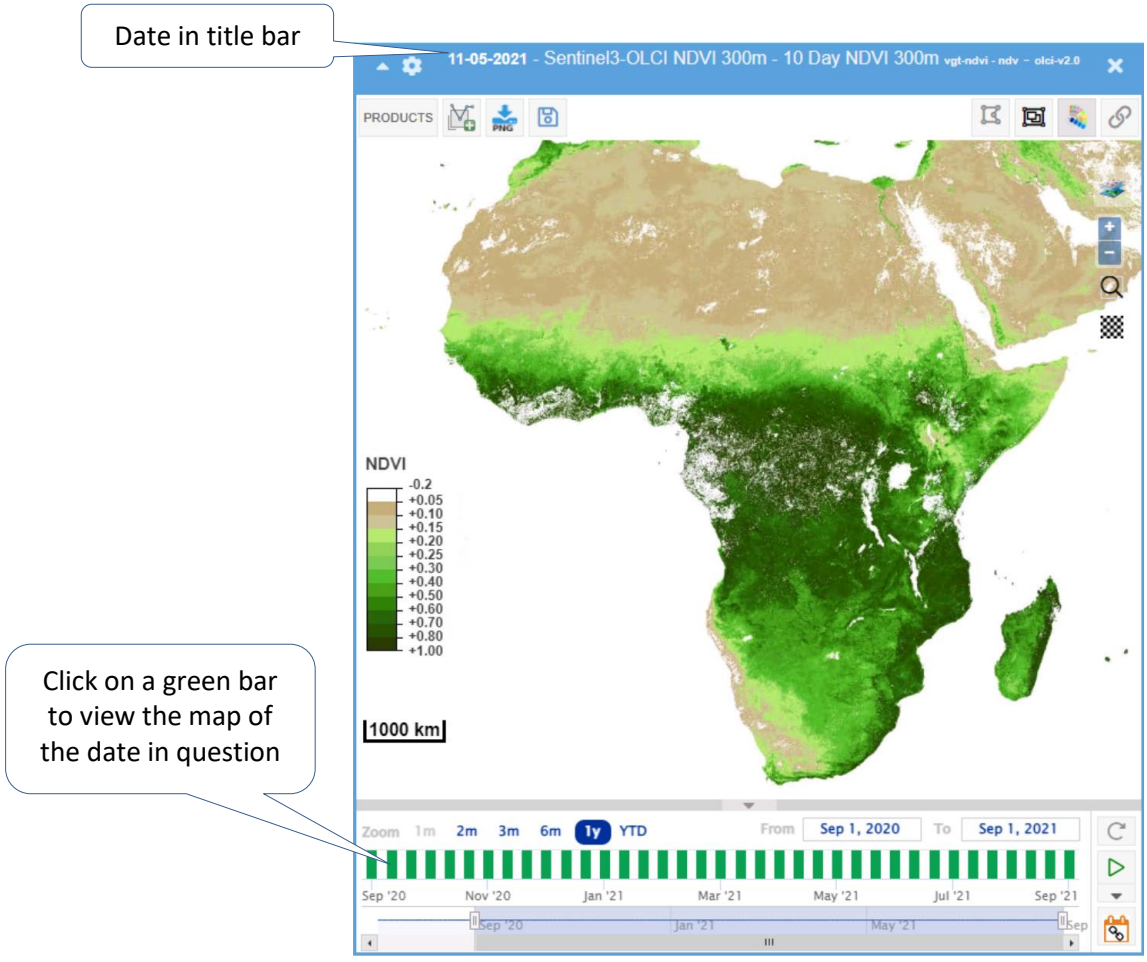
When a product has been added to a Mapview, the timeline of the (sub) product is shown on the bottom of the Mapview. All the possible dates within the products frequency are visualized chronologically as small, coloured bars.

- A **green bar** for every date physically present in the file system.
- A **red bar** for every date missing.
- A **grey bar** for every date permanently missing.

Viewing a sub product's date in the map area of a Mapview

Only the present dates with the **green bar** are clickable. When clicked, the map for the clicked date of the in the Product navigator selected (sub) product, will be rendered in the Mapview.

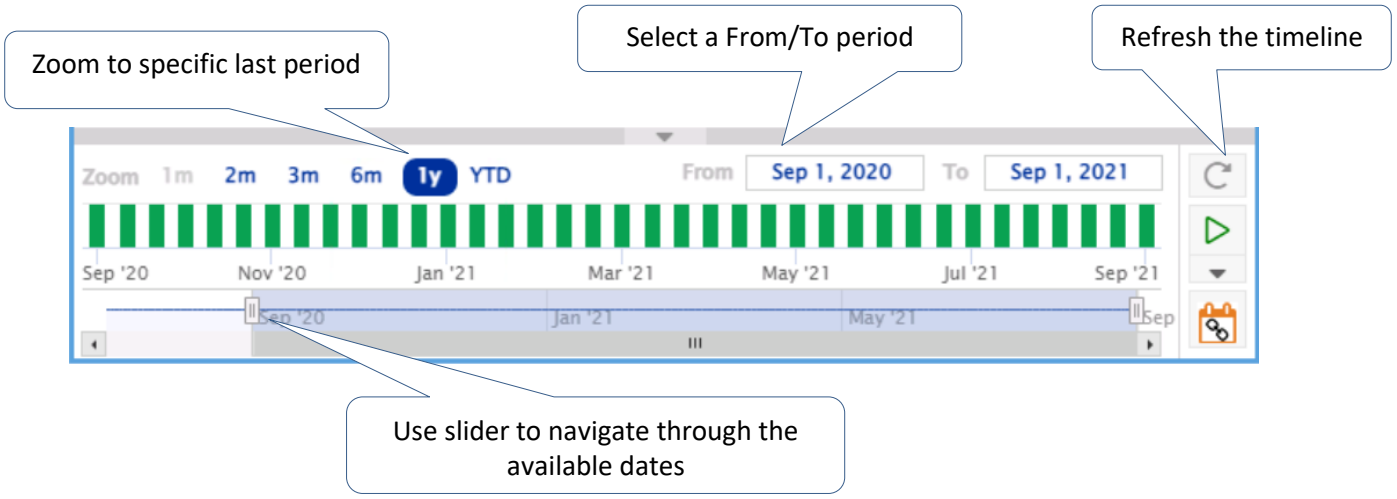
In the title bar of the Mapview the current date is shown.



Navigating through the timeline

Navigating through the timeline is mainly done using the slider on the bottom, which can be moved from right to left and vice versa and also made smaller or larger, to zoom in or out in the timeline.

Zooming can also be done through the 1m, 2m, 3m, 6m, 1y and YTD buttons in the left upper corner of the timeline or through the From/To fields in the upper right corner of the timeline. YTD will show the last available year. 1y will show the whole year that you currently are in on the timeline. The default zoom button depends on the product's frequency. For example, for a daily product, the 2m button is the default.



"Video" play the visible timeline of a Mapview

Showing all the dates visible in the timeline, like a video, can be done by clicking on the play icon  . When clicked, the first date visible in the timeline of the sub product will be displayed in the Mapview. After the default interval of 3 seconds (3000 ms), the next date in the timeline will be displayed, etcetera, until the last visible date in the timeline, and then starts again with the first visible date.

When playing the visible timeline, the play icon changes into a stop icon  . Click the stop icon to stop playing the timeline.

The interval can be changed by clicking on the small down arrow below the play icon. Using the small arrows will add or diminish 500 milliseconds to the interval.

Link or unlink the timeline of multiple Mapviews

Clicking on the green bar of a date in the timeline will show the clicked date of the sub product in its Mapview, but also in all other open Mapviews. By default, all open Mapviews are "timeline linked".

The same clicked date or the nearest available date of the sub product in other Mapviews will be taken.

Unlink a Mapview's timeline by clicking on the timeline link button  . The icon will change into the red icon with an unlinked chain  , meaning that the timeline of the Mapview is unlinked.
When you "play" the timeline of a Mapview, also all open and timeline linked Mapviews are "Played".

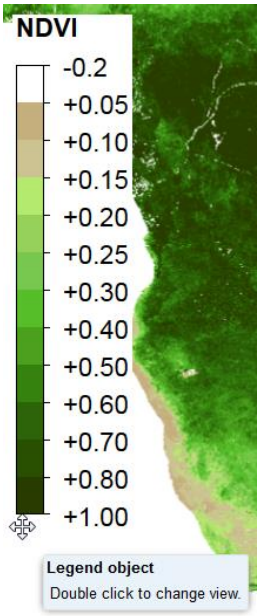
3.8.2.3 Link/unlink Mapview window from background layer

All Mapview windows are by default linked to the background layer (which is hidden).

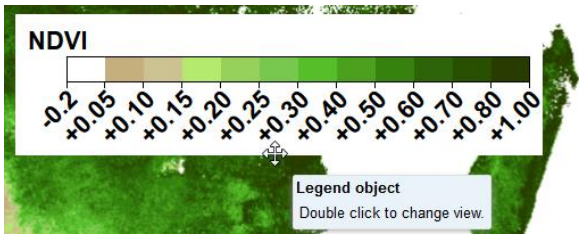
Maps are linked means that when zooming or panning within a Mapview or the background layer, all the linked Mapview windows will equally zoom or pan. This default behaviour you can **turn ON or OFF** for each Mapview by clicking respectively on the un-link  or link  toggle button.

3.8.2.4 Legend colour scheme

When a product is added to a Mapview, the legend of the chosen colour scheme is shown within the Mapview in vertical layout.



To view the legend in horizontal layout, double-click on the legend. Toggle between horizontal and vertical view of the legend by double-clicking on the legend.



To hide or show the legend, click the toggle button  in the "Mapview" toolbar.

Moving the legend frame within the Mapview can be done when you click, hold and drag the legend.

3.8.2.5 Navigation: zoom in/out and panning

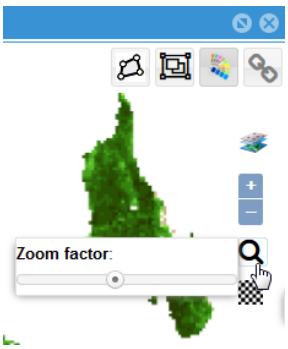
The zoom tool  gives the possibility to zoom in (+) or out (-).

The fasted way to zoom in or out is by using the mouse scroll wheel. Click within a Mapview and then use the scroll wheel on the mouse.

Panning is also done using the mouse. Click and hold the mouse button within a Mapview and then move the mouse to pan the map.

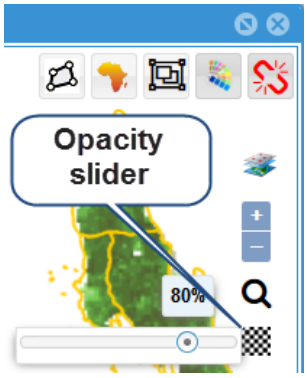
3.8.2.6 Zoom factor

The zoom factor slider is shown by clicking on the icon  , found in the right corner of the map area of a Mapview window. Use the slider to change the zoom factor. Values go from 1 to 10, with value 1 meaning slow/smooth zooming and value 10 fast zooming. The default is 5. Changing the zoom factor in one Mapview will change the zoom factor also in all the other linked Mapviews.



3.8.2.7 Opacity slider

The opacity slider is shown by clicking on the opacity icon  , found in the right corner of the map area of a Mapview window with a sub product. Use the slider to change the opacity of the product layer in steps of 10%.



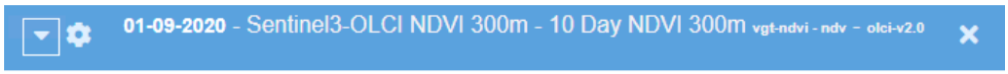
3.8.2.8 Show/hide tool bar

To create more space in the map area of a Mapview, you can hide the tool bar with the tool buttons by clicking on the  icon in the left corner of the header.

3.8.2.9 Collapse or expand a Mapview window

Because you can open many Mapview windows, it is good to be able to collapse Mapview windows and see only its header. Handy for organizing better the opened Mapview windows.

Click on the  icon in the left corner of the header to collapse a Mapview window.

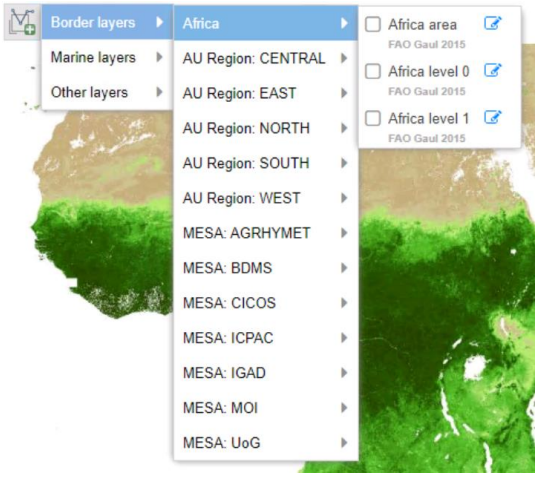


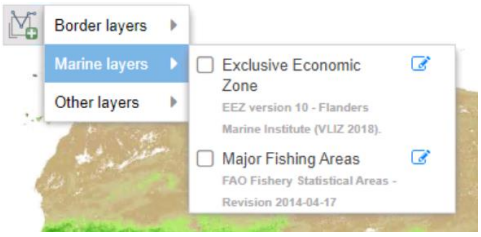
Move the header by clicking on it and hold the mouse button to be able to drag the collapsed Mapview window around the Analysis page in the browser. You can do this also when the Mapview window is expanded.

To expand a Mapview window, click again on the  icon.

3.8.2.10 Add vector layer

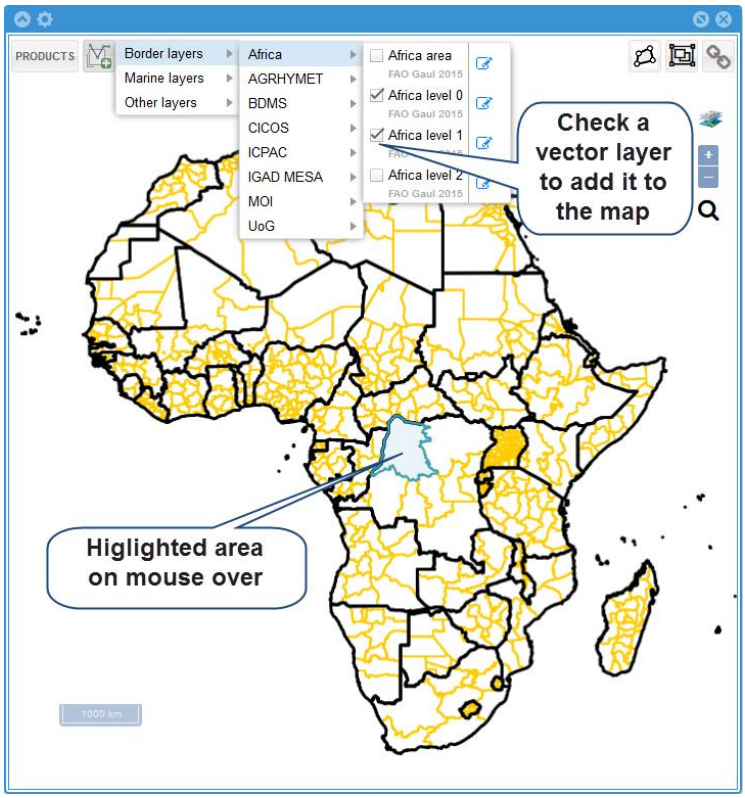
In the tool bar of a Mapview, the layer menu button  shows, when clicking on it, a menu of all the active vector layers included in the EStation. Vector layers can be activated/de-activated through the Layer administration.





Supplied vector layers are Gaul 2015-2014 Border layers, Marine layers (Fishing Areas and EEZ) and Other layers like Protected Areas.

Select a vector layer's checkbox from the menu to add the layer to the map area of the Mapview. De-select a vector layer's checkbox from the menu to remove the layer from the map area.



All polygons (area geometries) are interactive. Going over a polygon (or multi polygon) of an area, will highlight the area and will show the area name in the tool bar.
Clicking on an area will select the area and the geometry of the area becomes **red** (the default is red for the selected feature outline colour but can be changed). Also, the name of the selected area will be shown in the Time series tool under "Selected region". The time series tool is only working if an area has been selected.

When more than one vector layer has been added, priority is given to the layers for selecting an area with the highest priority (lowest priority number in the vector menu). Protected Areas have the highest priority, then level 2, then level 1, then level 0 and finally level 00. This means that when you have all 4 levels added, you will have to hide layers with higher priority to be able to select areas within a lower priority layer.

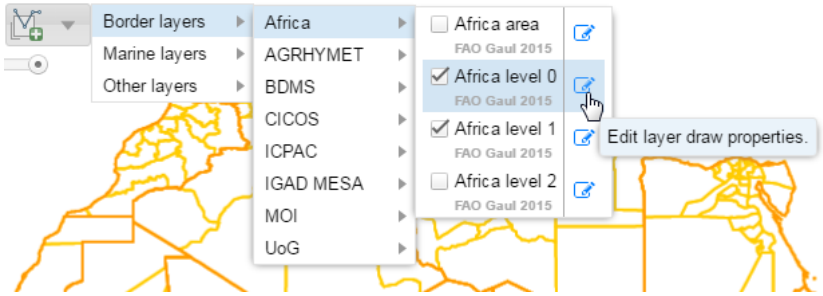
- Make level 0 areas selectable then hide the Protected areas layer and all level 1 and level 2 layers.
- Make level 1 regions selectable then hide the Protected areas layer and all level 2 layers.

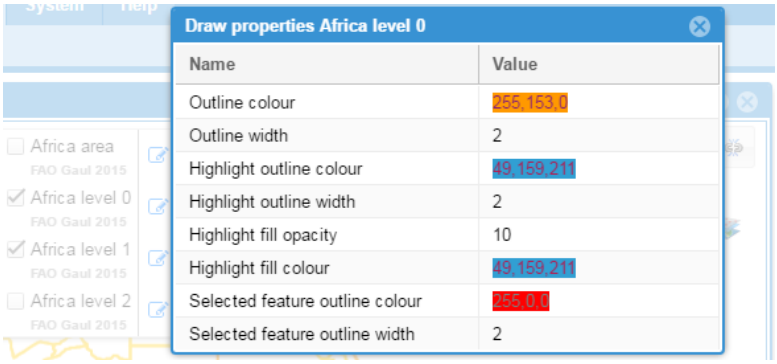
Hide layers though the Layer switcher.

Double clicking on a polygon/feature will zoom to its extent.

3.8.2.11 Edit layer draw properties from menu

The draw properties of each vector layer can be edited individually through the layer menu by clicking on the  icon next to the layer in the menu, which invokes the Edit layer draw properties tool.

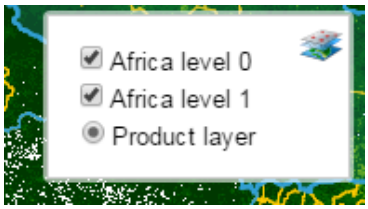




The changes will not directly apply to the layer if it has already been added to the Mapview. If so, remove the layer from the Mapview and then add the layer again to see the changes made to the layer's draw properties.

3.8.2.12 Layer switcher

Vector layers added to a Mapview can be hidden through the Layer switcher  , present in the top right corner of the map area. Clicking on the Layer switcher will show a list of all the layers added. The product layer (if present) is the base layer and cannot be hidden.

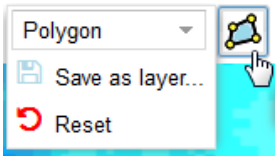


Click on the checkbox next to a vector layer to hide or show the layer.

3.8.2.13 Draw geometries

Activating draw mode

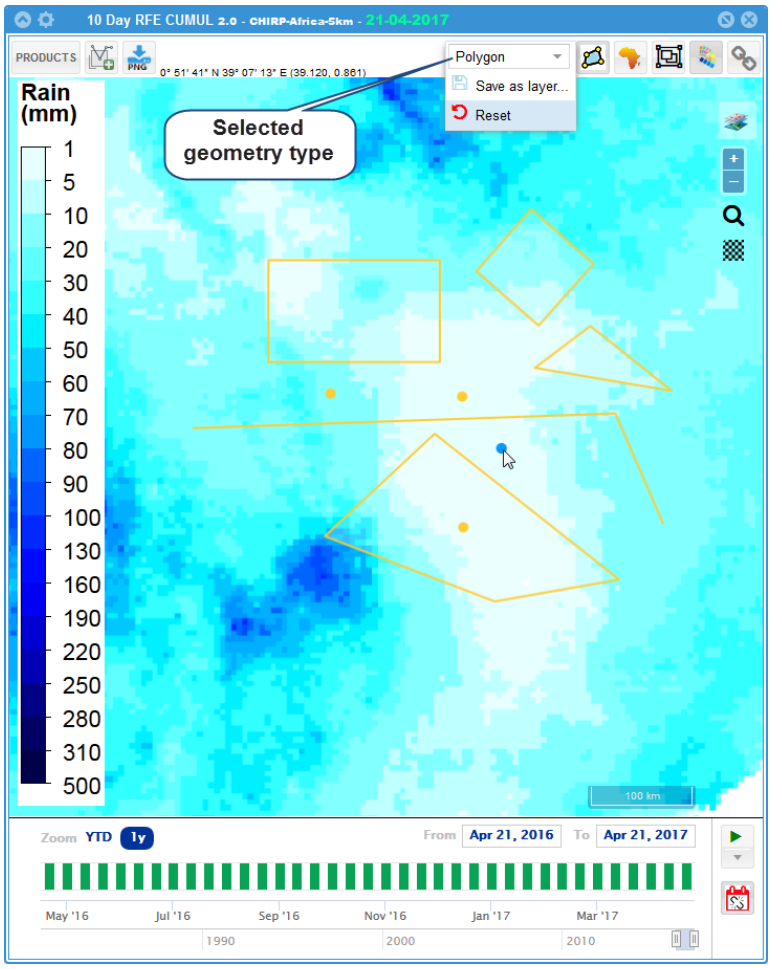
Every Mapview has the draw geometry functionality. Clicking on the button  , will show a menu and put the map area in draw mode.



From the menu you can select the geometry type you want to draw, save the drawn geometries as a layer and reset the draw layer by removing all the drawn geometries.

Geometry types are:

- Polygon
- LineString
- Point
- Square
- Box

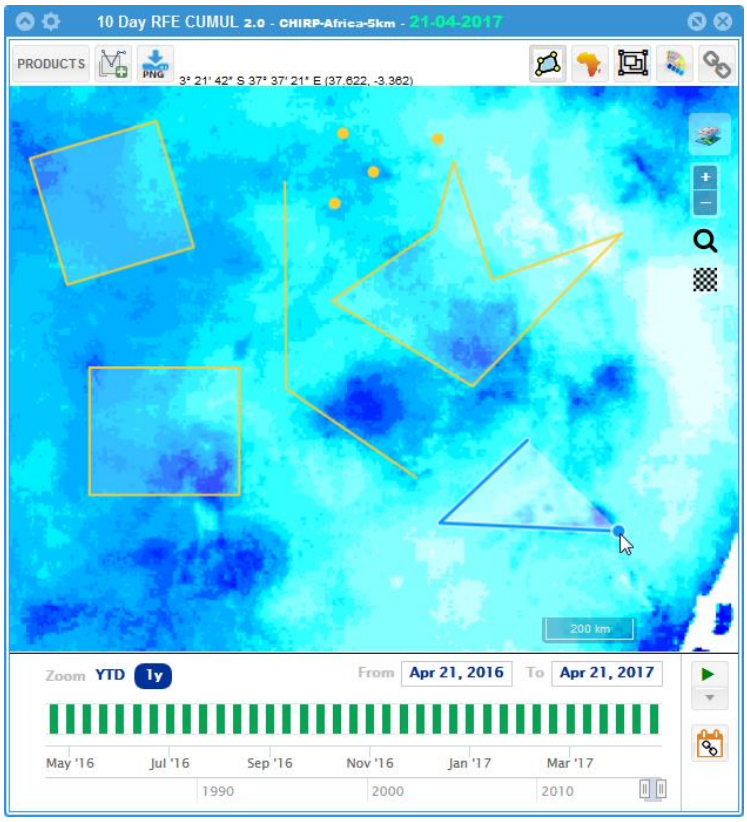


The map area is in draw mode when you see a light blue dot following the cursor  and the draw geometry button looks like this  .

Turn off drawing mode by clicking on the  button.

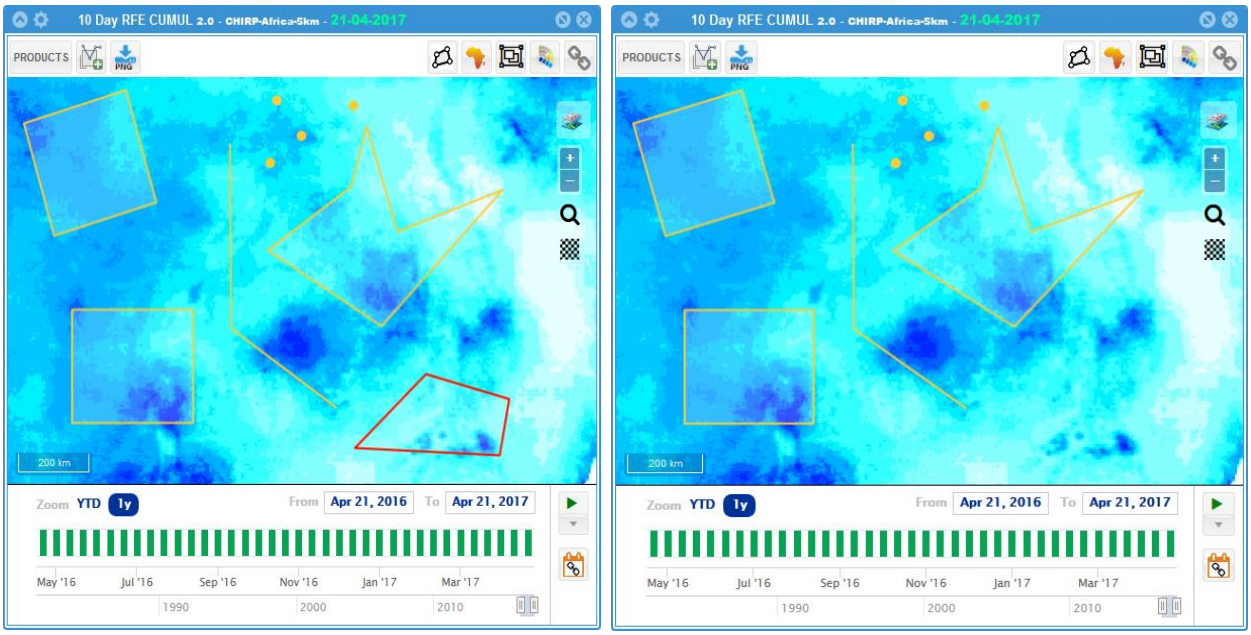
Drawing geometries

In draw mode click somewhere in the map area to start drawing the selected geometry type from the menu. Depending on the geometry type, click or double click in the map area to finish the drawing of the geometry in question.



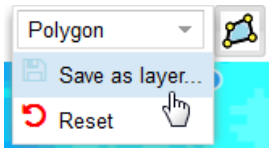
Deleting a drawn geometry

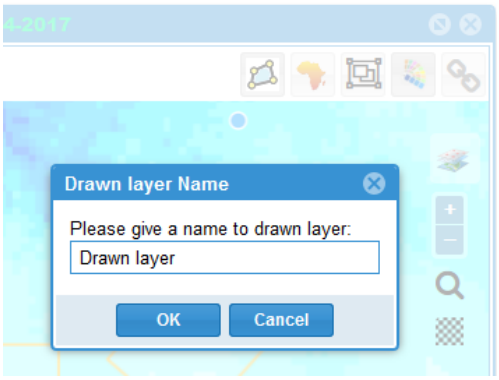
You can delete a drawn geometry by turning off drawing mode and select the geometry to be deleted. Once selected, hit the “Delete” key on your keyboard to delete the geometry.



Save drawn geometries as a layer

Selecting “Save as layer…” from the menu will show a dialog asking to give a name to the layer.





Give a suitable name for your layer and click on OK.
All the drawn geometries are saved in a file in geojson format and added as a user layer on the EStation with default settings. The saved layer is also opened, directly in the Mapview.

To share your drawn layer, the file is also downloaded to the download directory of your browser.

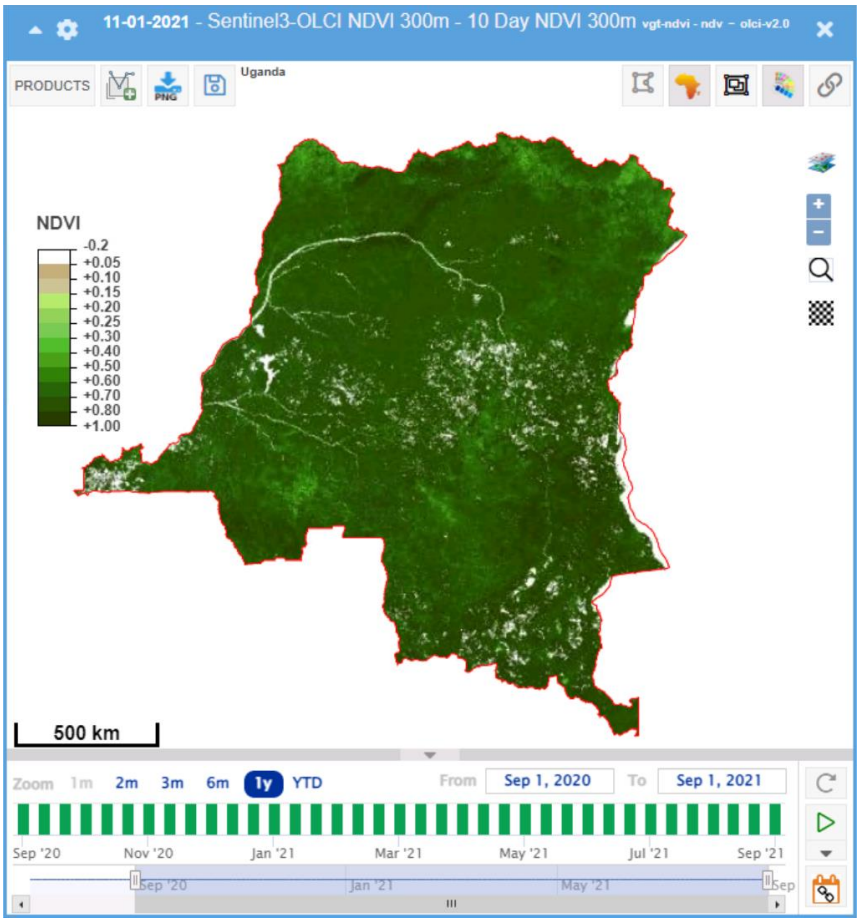
To change the default settings of the layer, go to the Layer administration.

3.8.2.14 Out masking a product

When a product and one or more vector layers have been added to a Mapview, the Out-mask toggle button  appears in the Mapview tool bar. When clicked, the toggle button becomes darker grey  , which means that Out masking is turned on.

Selecting the feature/geometry of a region in the Mapview will now out mask the selected region.

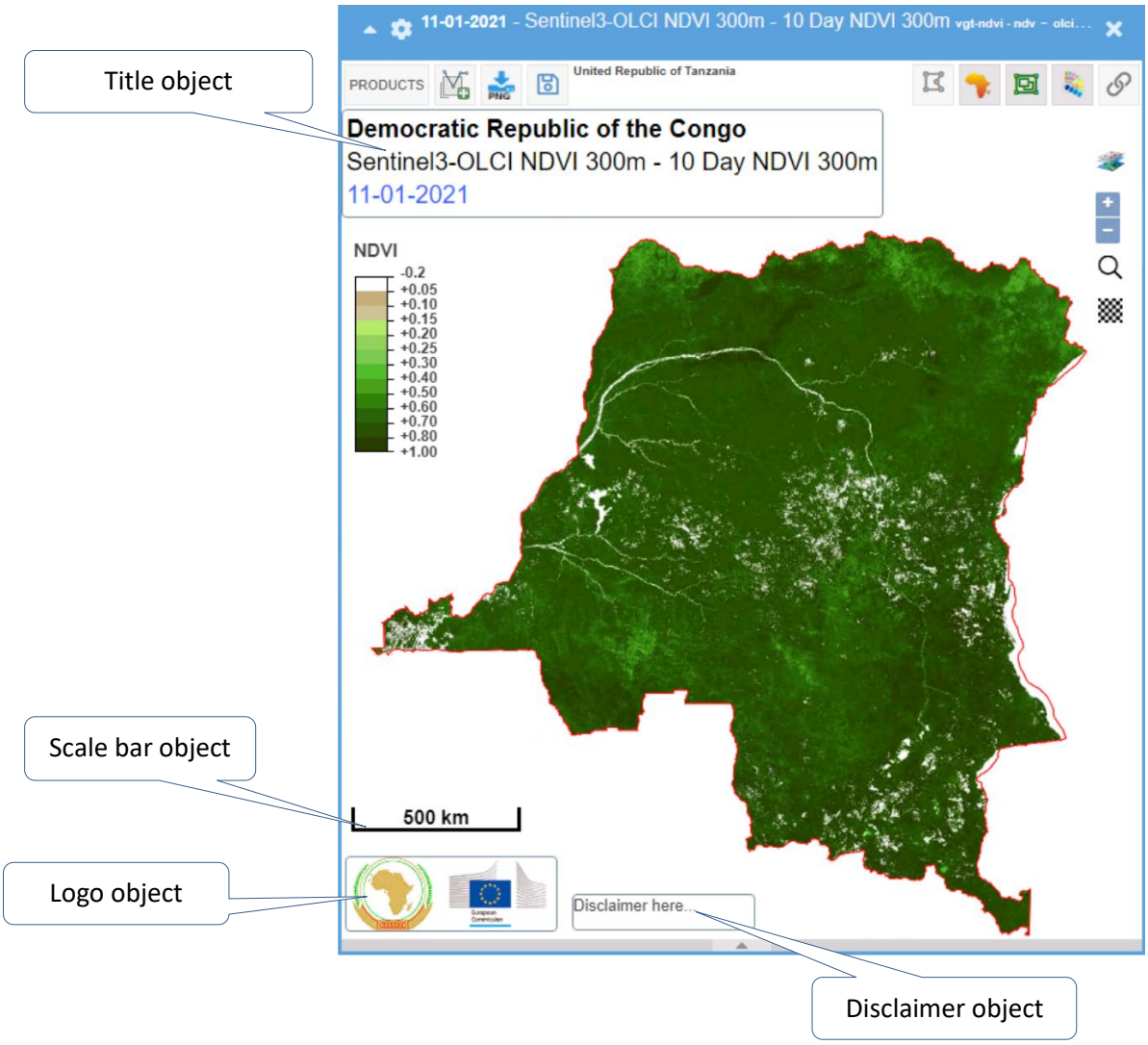
Turn off Out masking by clicking again on the toggle button.



3.8.2.15 Title, logo and disclaimer objects

There are three objects you can show and hide in the map area, the Title, Logo and Disclaimer objects.

To show the objects, click on the button. The button becomes green and the three objects are shown in the map area.



You can move and reposition all three objects by click and hold on an object and drag the object to reposition it. Also, the scale object is repositionable.

Edit Title object

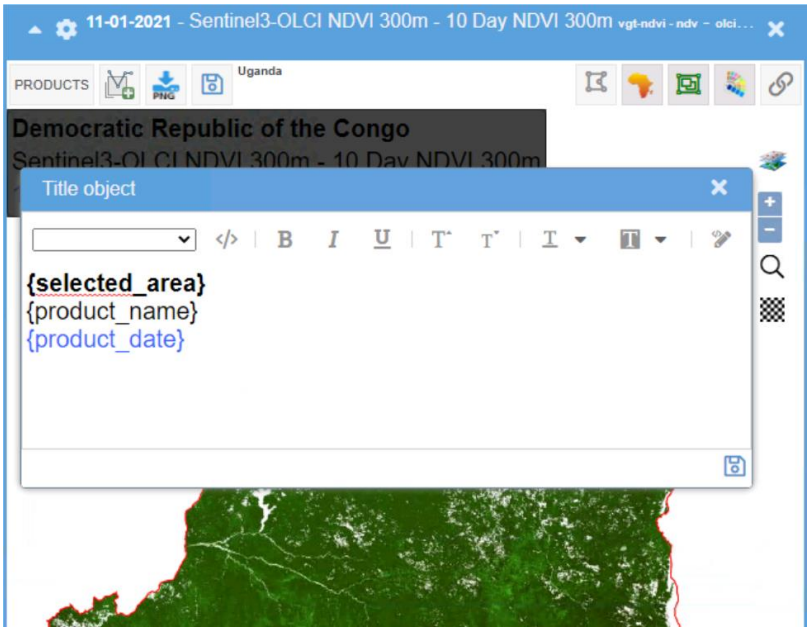
Double-click the Title object to edit it. The title editor is shown whereby default the 3 dynamic fields are added. You can type any text around the dynamic fields or delete the dynamic fields if they are not desired. Format the text (and dynamic fields) using the available text formatting tools in the editor.

Dynamic fields are fields that are automatically filled in when used. There are 3 fields:

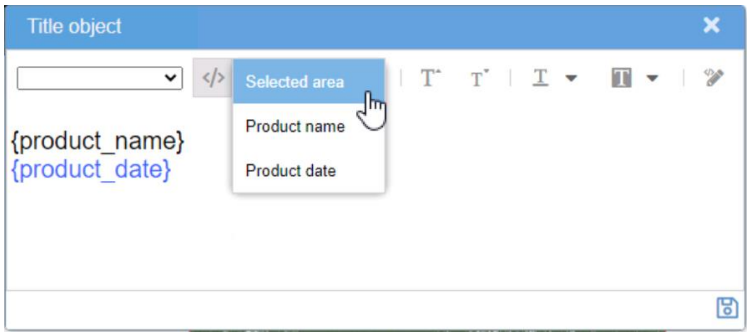
{selected_area} – The name of the selected area/geometry of a layer in the Mapview.

{product_name} – The name of the sub product shown in the Mapview.

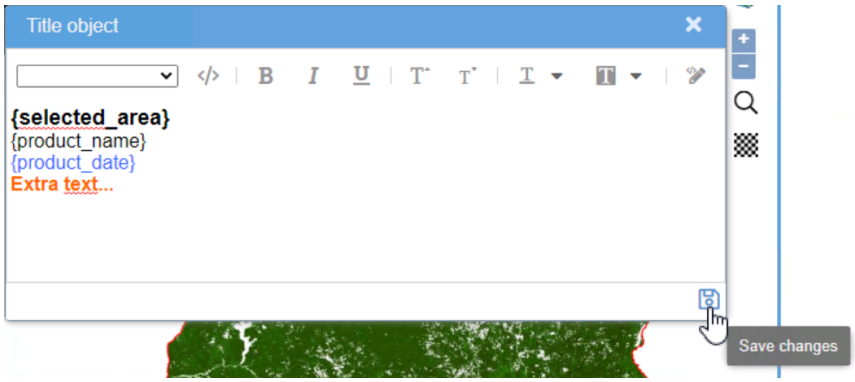
{product_date} – The date of the sub product shown in the Mapview.



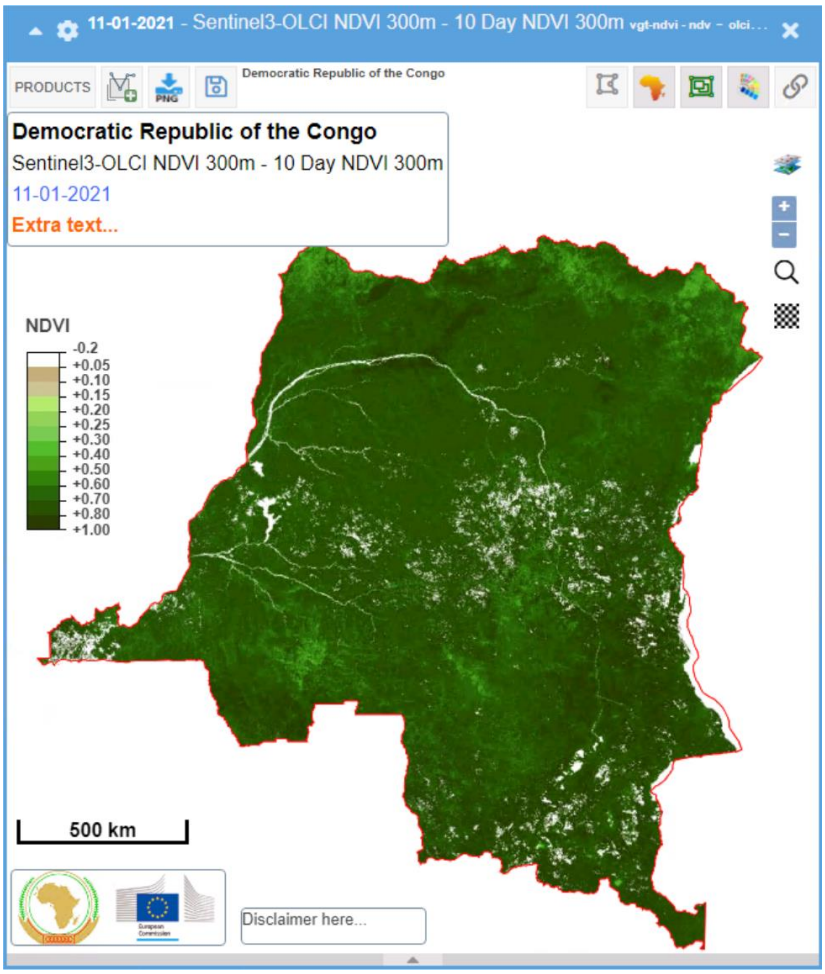
When you have deleted dynamic fields, you can add them by typing exactly the field names in brackets as shown above or by using the dynamic field selection tool </> in the editor's toolbar.



- To save your changes, click on the save button found in the header of the editor.



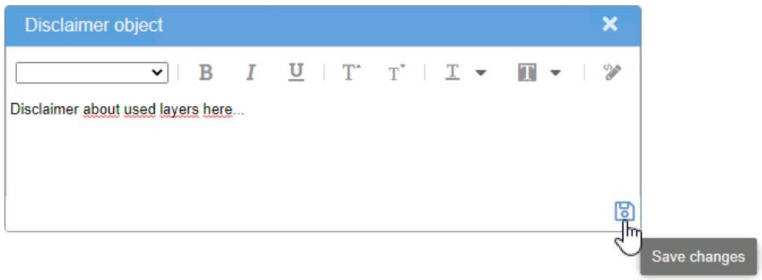
- The editor will close, and the changes reflect in the title object of the Mapview.



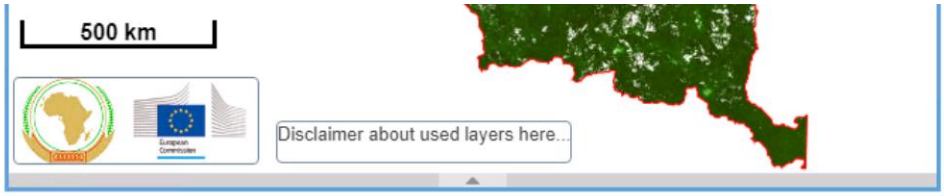
Edit Disclaimer object

Double-click the Disclaimer object to edit it. The disclaimer editor is shown. You can type any text and format the text using the available text formatting tools in the editor.

- To save your changes, click on the save button found in the header of the editor.



- The editor will close, and the changes reflect in the disclaimer object of the Mapview.

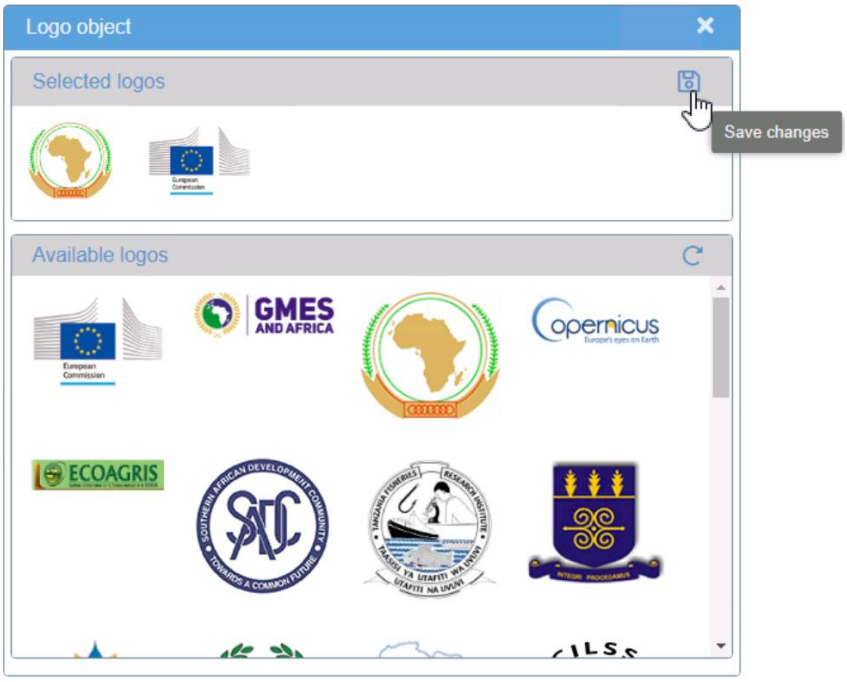


Edit Logo object

Double-click the Logo object to edit it. The logo editor is shown.

The upper box contains the selected logo's that will appear in the logo object. The lower box contains the available logos.

- Double-click on a logo in the selected logo's box to remove the logo.
- Double-click on a logo in the available logo's box to add the logo to the selected box.
- To save your changes, click on the save button found in the header of the editor.



3.8.2.16 Save map as PNG image

To save a map rendered in a Mapview click on the PNG button present in the tool bar in the top of the Mapview. A snapshot of the all the visible layers in the map, the product legend (as shown in the map) and the title, disclaimer and logo objects will be made in PNG format and automatically downloaded in the download directory of the browser.

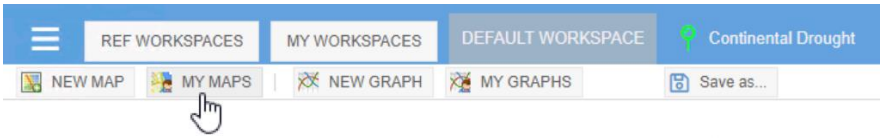
3.8.2.17 Map templates

Save Mapview as template

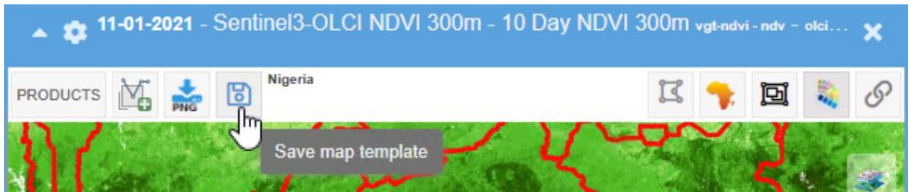
A map template is a by the user prepared Mapview saved in the database with a given name. A saved map template can be reopened exactly as the user prepared the Mapview (currently the saving of the opened vector layers.

To save a prepared Mapview as a map template, the user must first login.

When logged in you will see a “MY MAPS” button in the toolbar of the Analysis tool.



Also, you will see a “Save” button in each Mapview you open.

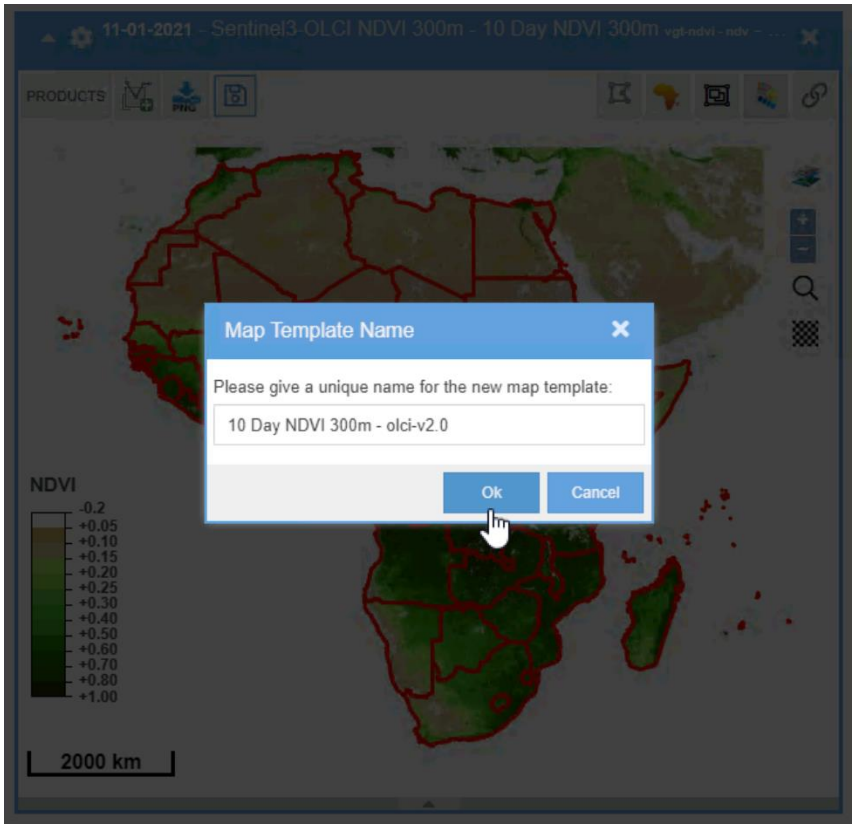


What will be saved in the map template?

- Mapview window size and position in the Analysis tool.
- Sub product added to the Mapview.
- Legend id.
- Legend visible or not.
- Legend position in the Mapview.
- Legend orientation (vertical or horizontal).
- Title object, Logo object and Disclaimer object visible or not.
- Title object, Logo object and Disclaimer object content and position in the Mapview.
- Scale bar object position in the Mapview.
- Vector layers opened in the Mapview.
- Zoom extent.
- The selected geometry.
- Out mask on or off.

When you have prepared a Mapview, save it as a map template by clicking on the “Save” button.
You will be asked to give the map template a name. If a product is opened in the Mapview, the product name and version is proposed as the map template name.

Give the map template a suitable name and click on Ok.



The name given will appear in **orange** in the header of the Mapview, indicating that the Mapview is a map template.

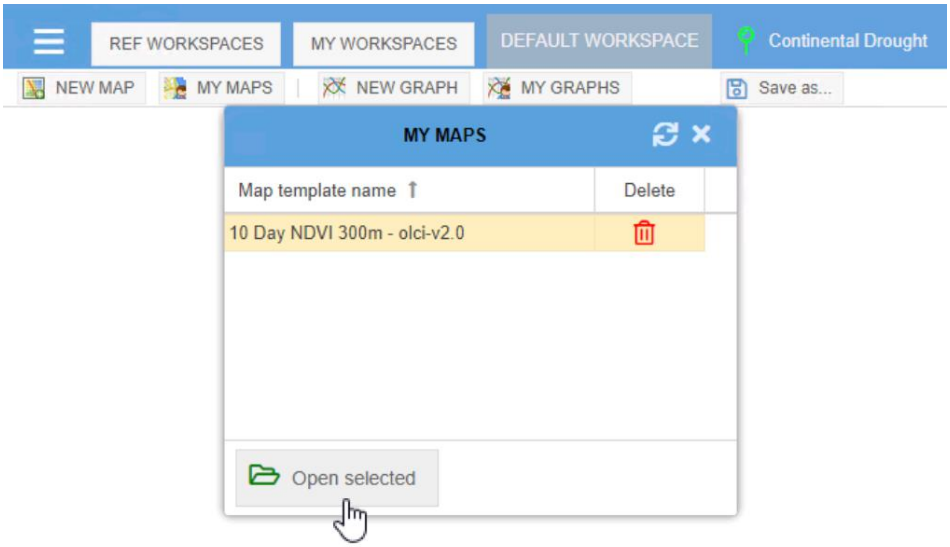
Click on the "Save button to save any changes made to the map template.

Save a map template under another name by choosing "Save as…" under the "Save" button.

Open a map template

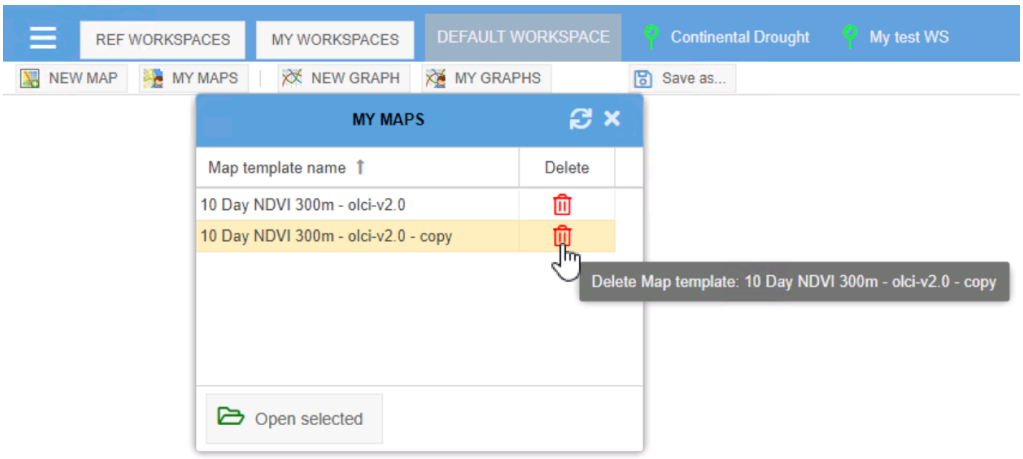
Click on the "MY MAPS" button in the toolbar of the Analysis tool. The list of all the user's map templates will appear next to the button.

Select one or more (using the Ctrl key) map templates and click on "Open selected".



The selected map template(s) will be opened exactly the way it was saved.

Delete a map template by clicking on the icon next to the name of the map template.

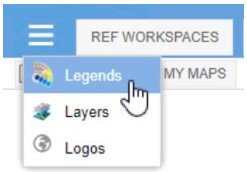


3.8.3 Legends administration

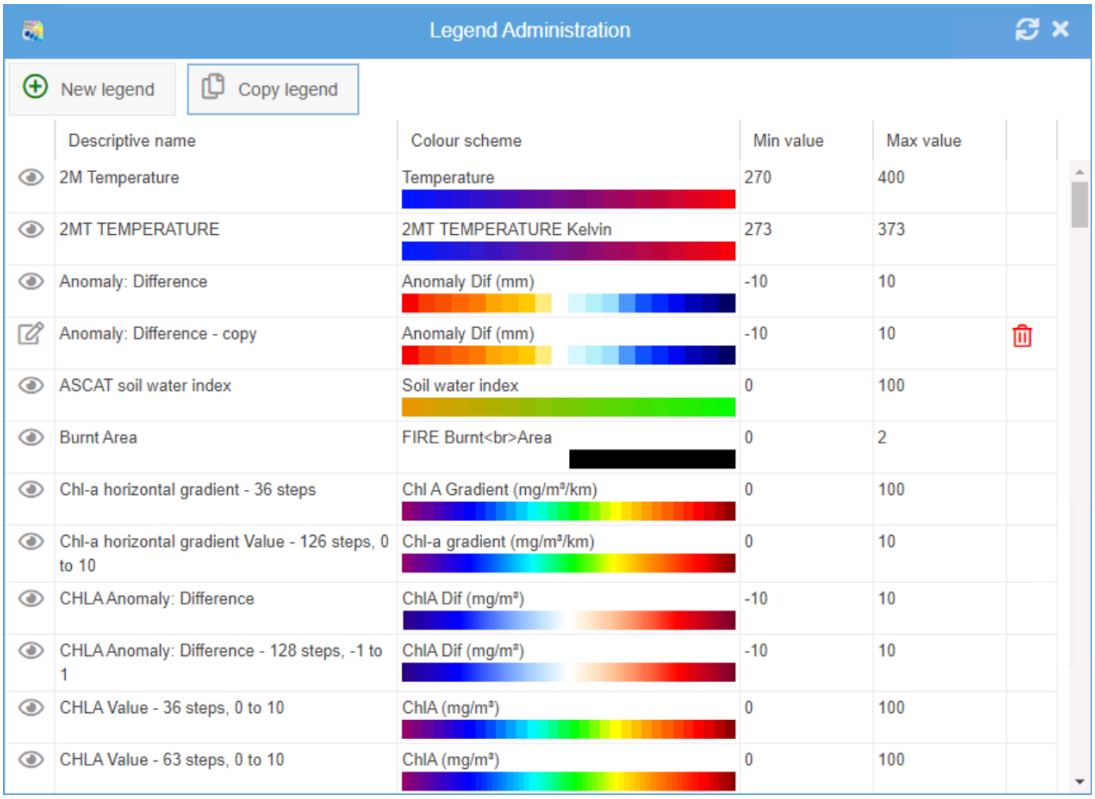
The Legends administration tool is available to:

- View predefined legend definitions.
- Create a new legend.
- Copy a legend.
- Edit a user defined legend definition.

Open the Legends administration window by clicking on the menu on the top left corner of the Analysis page.



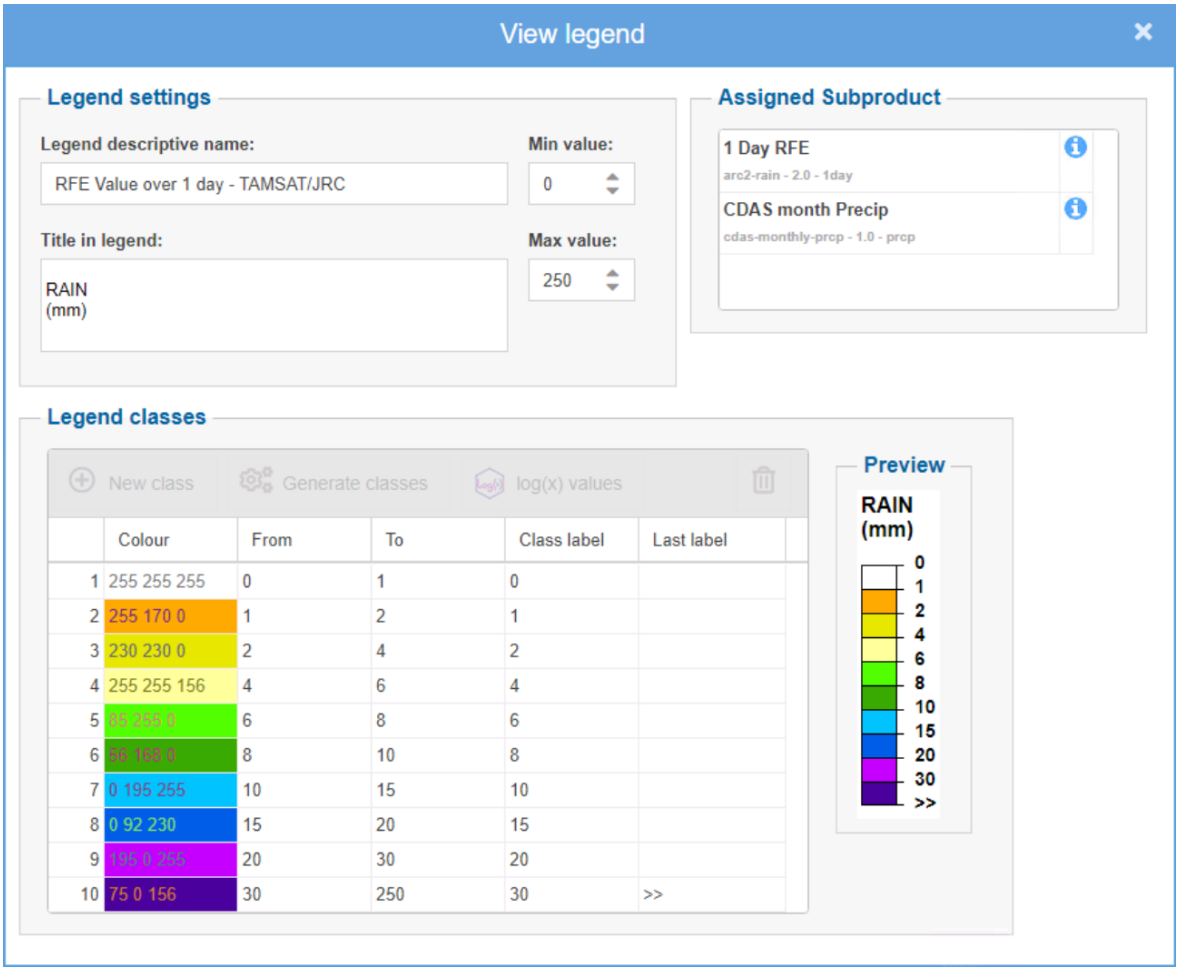
A full list is presented of the legends defined on the EStation, of which most are supplied by the JRC.



3.8.3.1 View JRC defined legend definitions

Legends defined by the JRC cannot be changed or deleted. You see an eye icon in front of JRC defined legends. To view a legend defined by the JRC, double click on the legend or click on the eye icon.

The legend definition window is opened in view mode.



A legend definition contains 4 sections:

1. Legend settings
2. Legend classes
3. Preview – the preview of the vertical legend definition
4. Assigned sub products – the list of sub products to whom the legend is assigned.

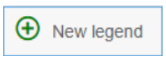
Legend settings

	Description
Descriptive name	The descriptive name of the legend.
Title in legend	The title of the legend, which appears above the legend.
Min value	The min value, which is the “From” value of the first legend class.
Max value	The max value, which is the “To” value of the last legend class or a higher value.

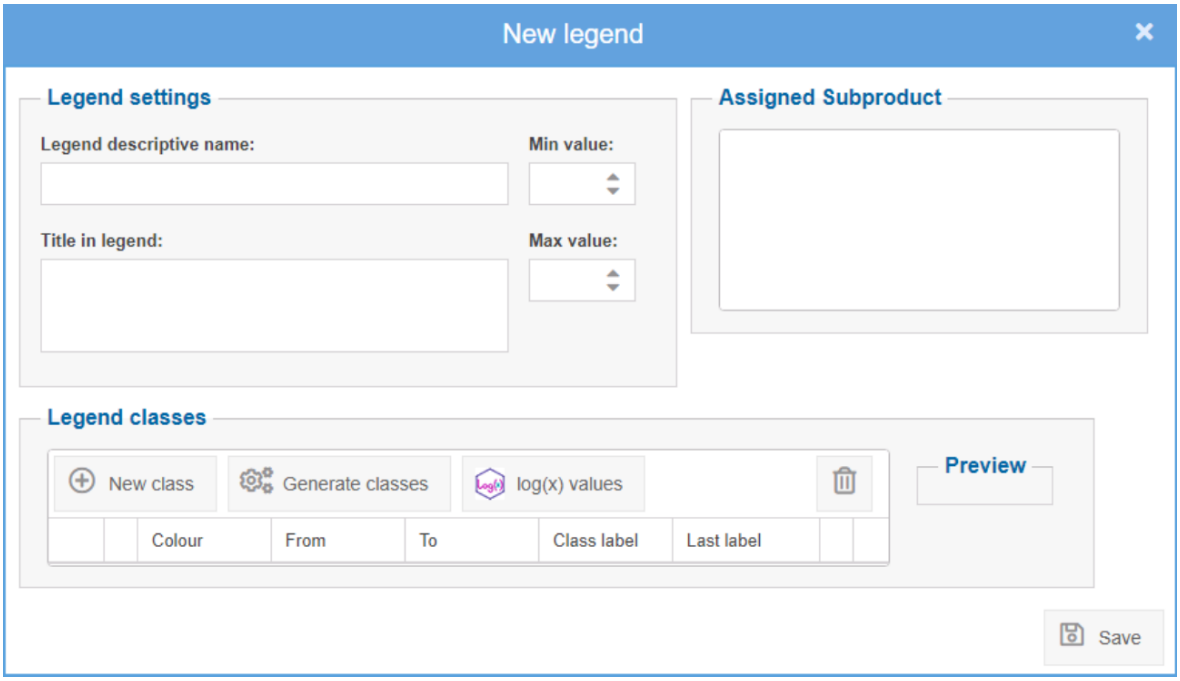
Legend classes

	Description
Colour	The RGB colour of the legend class.
From	The “From” value of the legend class.
To	The “To” value of the legend class.
Class label	The label of the legend class.
Last label	The label of the last/bottom tick in the legend.

3.8.3.2 Create a new legend

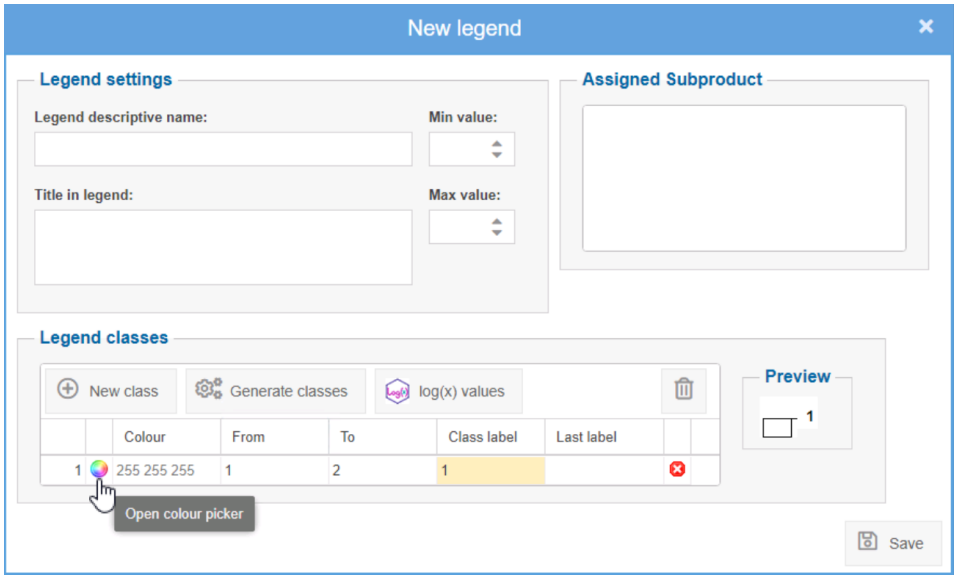
To start creating a new legend, click on the 

button. The legend definition window is opened in “new legend” mode.

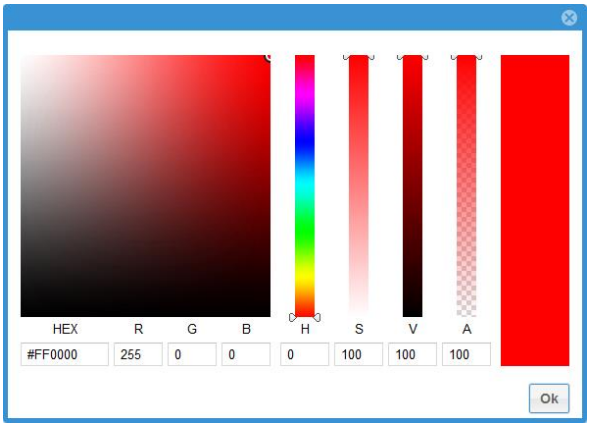


Give the legend a descriptive name and fill in the fields “Min value”, “Max value” and “Title in legend”.

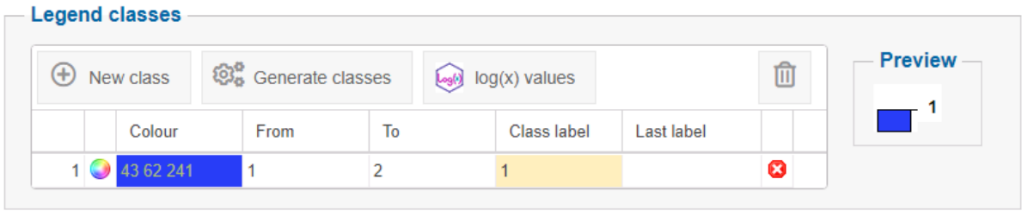
To create classes, click on the “New class” button as many times the legend needs.
You will see a new class record appearing in the legend classes grid for each new class created, with the default values. To change the default values, click on a class field to edit the value manually.
The classes are automatically sorted by the “From” field in ascending order and every change made will be reflected in the “Preview” area.
For the “Colour” field of each class, a colour picker is available. To open the colour picker, click on the icon.

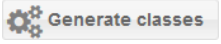


Choose a colour from the colour picker and click on OK.

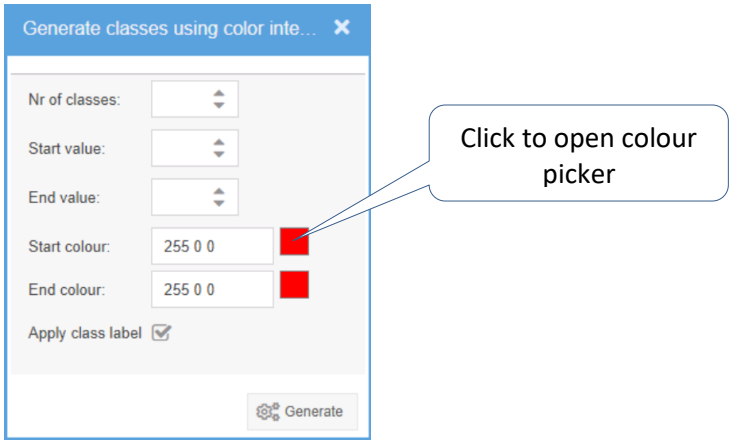


The colour picker will close, and the selected colour will be filled in the colour field of the class in RGB format (with a space delimiter).



Instead of adding new classes by clicking on the “New class” button, you can also generate classes. Click on the 

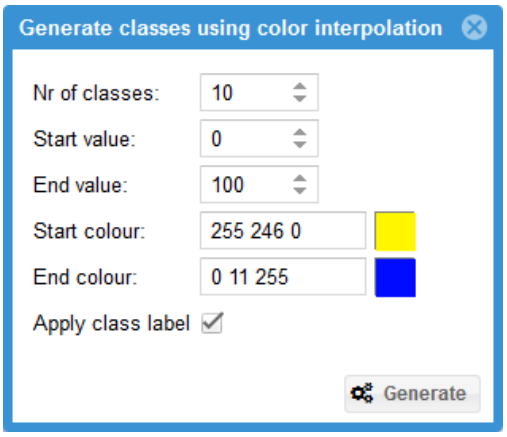
button to open the window to setup the generation of classes.



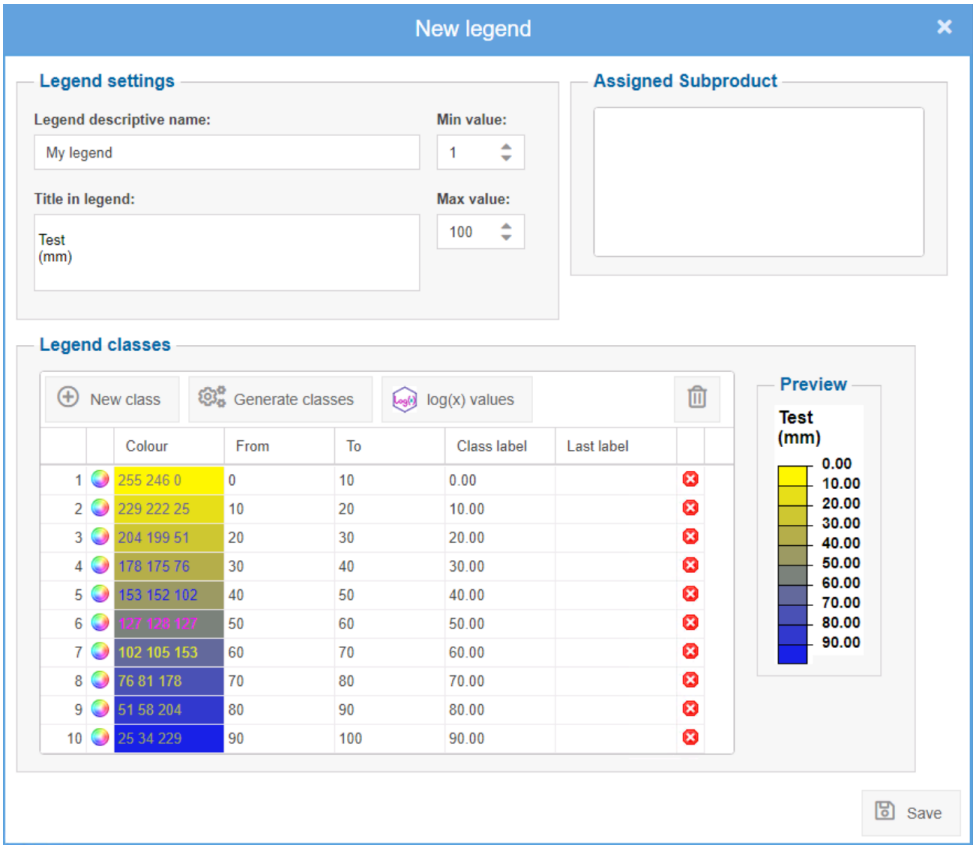
Classes generation fields

	Description
Nr of classes	The number of classes to generate.
Start value	The "Start" value of the first legend class.
End value	The "End" value of the last legend class.
Start colour	The colour of the first class, from which colour interpolation starts.
End colour	The colour of the last class, from which colour interpolation starts.
Apply class label	When checked a label for all classes will automatically be generated from the class "From" value.

For example:



Click on the "Generate" button to generate the classes. The "Generate classes" window will close, and you will see the generated classes in the "New legend" or "Edit legend" window.

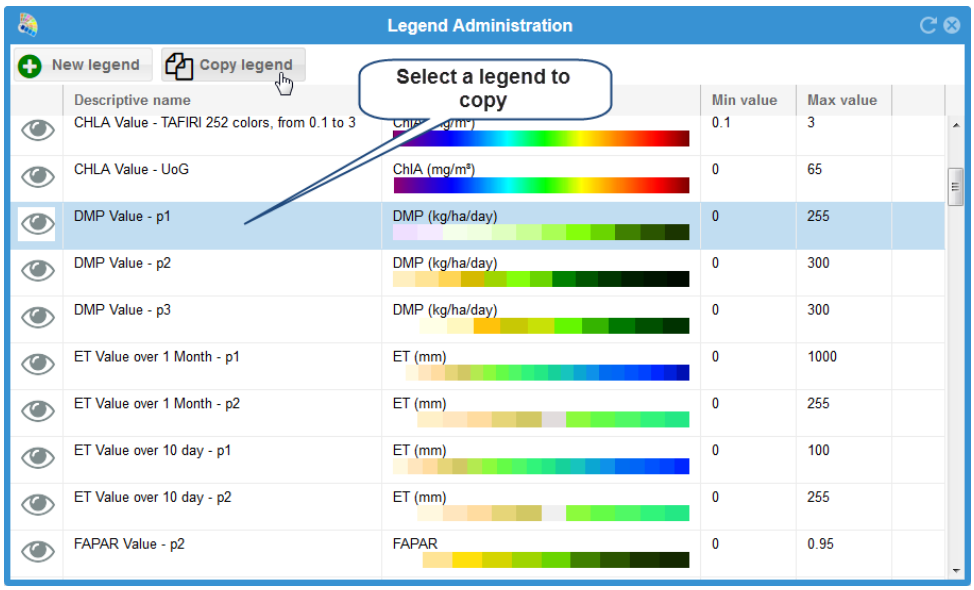


Click on the button to delete all classes.

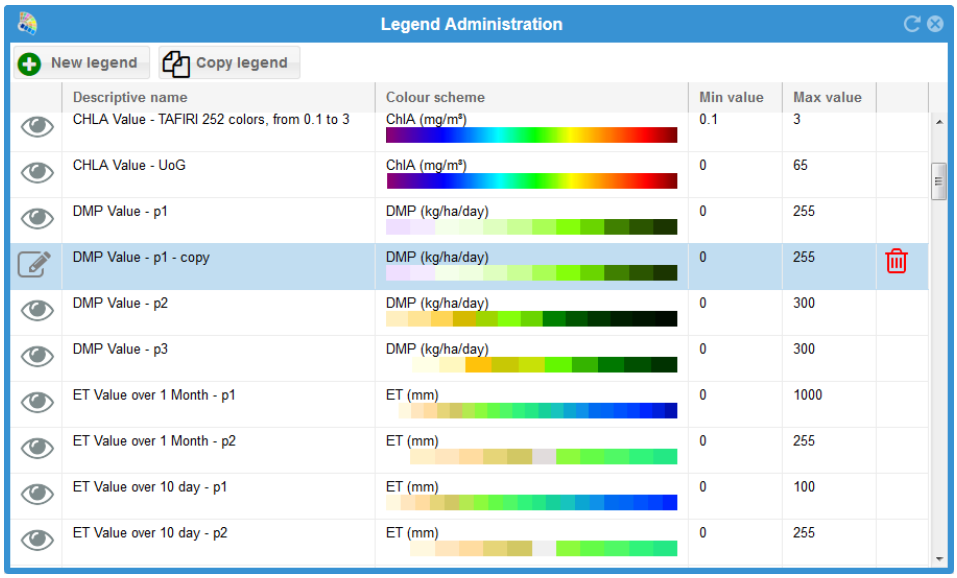
To save the legend, click on the “Save” button. Depending on the number of classes, the saving might take some seconds. When the new legend is saved, a message will appear from the top of the browser.
When the new legend is saved, the window title will change in “Edit legend”.

3.8.3.3 Copy a legend

Select a legend from the list of existing legends and click on the Copy legend button.



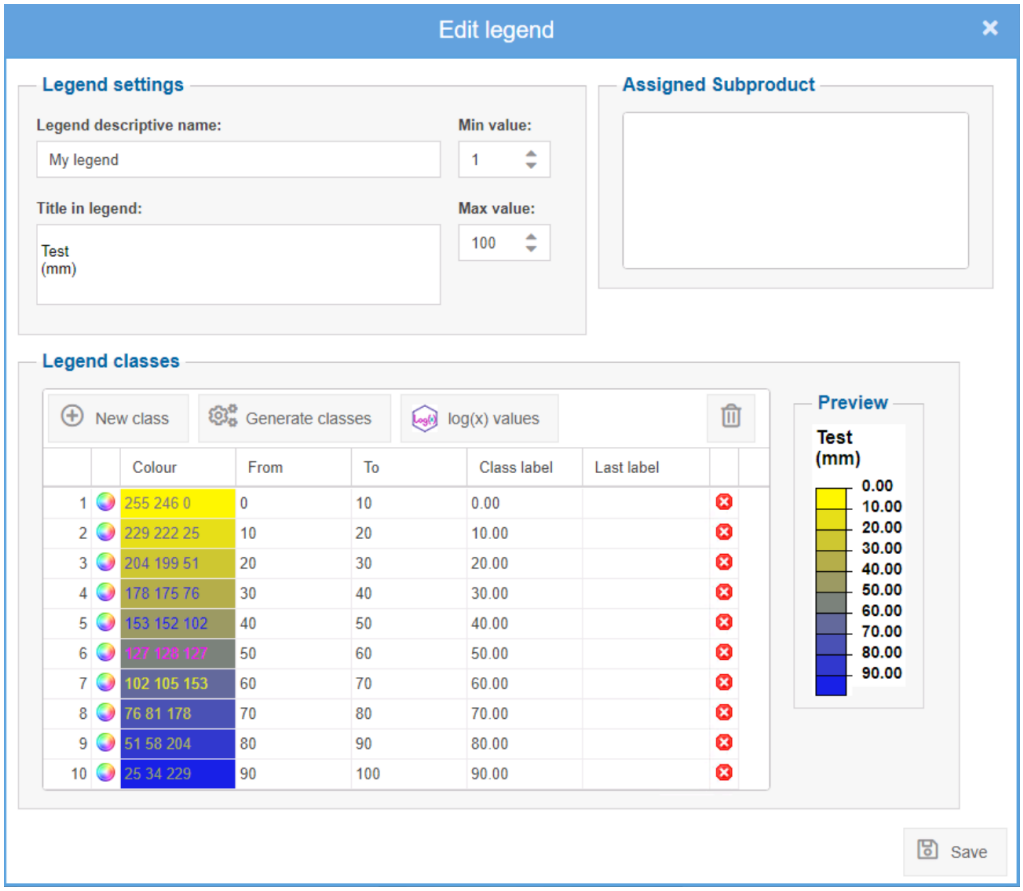
The legend is copied and given the same name ending with – copy.



If you copied a JRC defined legend, the edit icon will appear next to the copied legend name, and you can now edit the copy. Also, the delete icon appears to delete the legend.

3.8.3.4 Edit a user defined legend definition

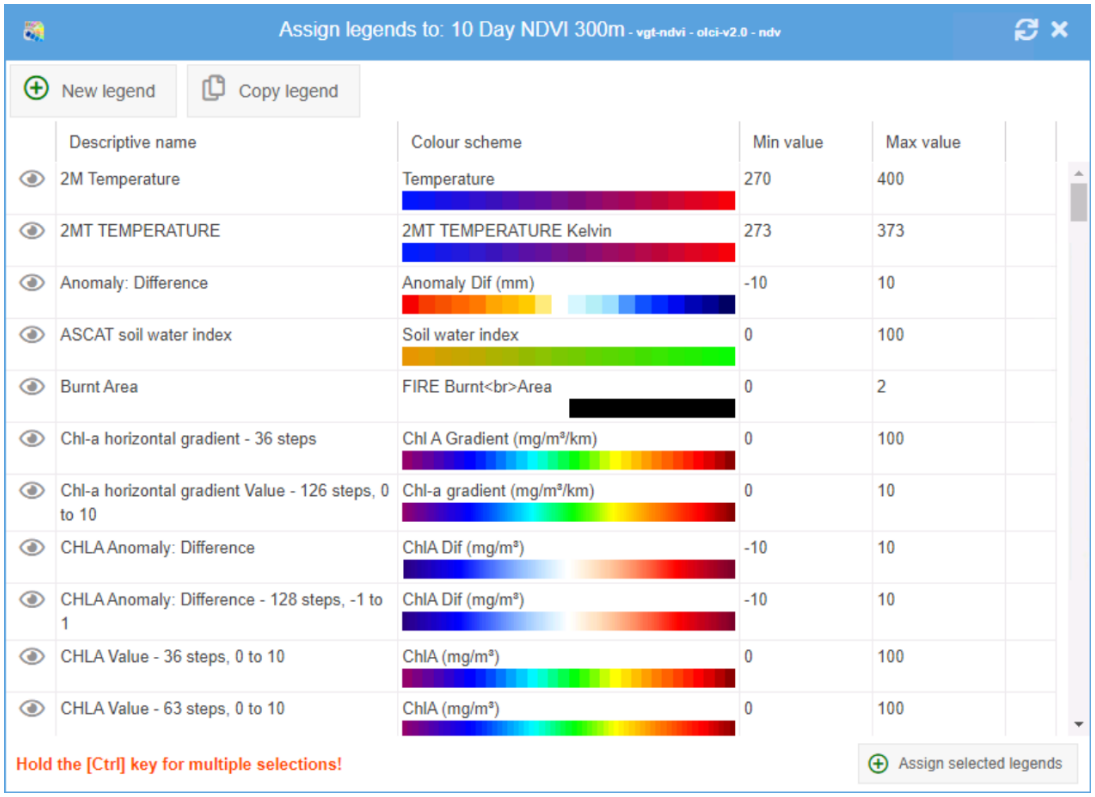
To edit a legend (not defined by the JRC), click on the icon next to the legend name in the legend administration window. The “Edit legend” window will be opened, where you can make any changes to the legend definition.


Edit legend
Legend settings
Legend descriptive name:
My legend
Min value:
1
Title in legend:
Test
(mm)
Max value:
100
Assigned Subproduct
Legend classes
New class
Generate classes
log(x) values
Colour
From
To
Class label
Last label
1 255 246 0 0 10 0.00
2 229 222 25 10 20 10.00
3 204 199 51 20 30 20.00
4 178 175 76 30 40 30.00
5 153 152 102 40 50 40.00
6 127 128 127 50 60 50.00
7 102 105 153 60 70 60.00
8 76 81 178 70 80 70.00
9 51 58 204 80 90 80.00
10 25 34 229 90 100 90.00
Preview
Test
(mm)
0.00
10.00
20.00
30.00
40.00
50.00
60.00
70.00
80.00
90.00
Save

3.8.3.5 Assign legends to products

To assign a legend to a product, open the product navigator from any “Map view”.

Go to the product to which you want to assign a legend and click on the Assign legend button.
A list of available legends will appear.



Select the legend you want to assign by clicking on the legend.
To select more legends, hold the {Ctrl} key and select other legends.
When you have selected the legends, you want to assign to the product, click on “Assign selected legends”.
The new assigned legend(s) will appear in the product navigator under the product’s colour schemes.

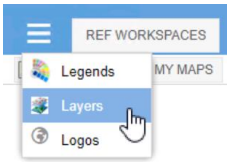
3.8.4 Layer administration

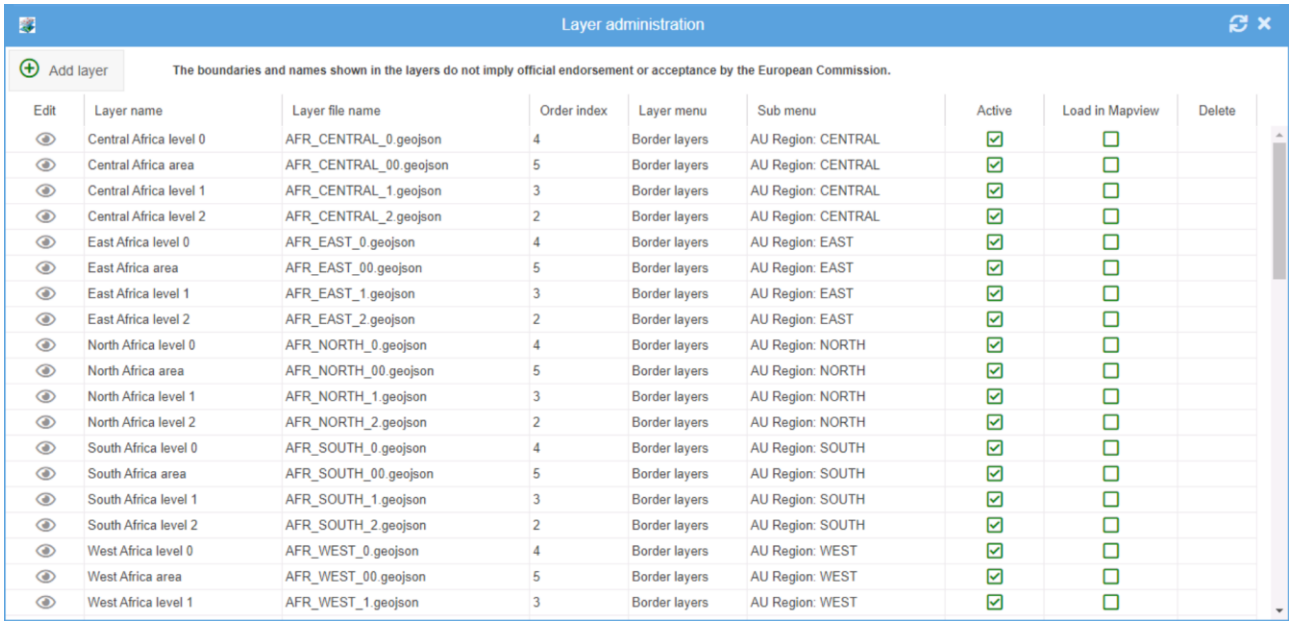
The Layer administration tool is available to:

- Edit a layer's properties.
- Create a new layer.
- Upload a vector layer in .geojson format.
- Reorganize the layers under the layer menu button  in Mapviews.
- Set layers to be automatically opened in a Mapview.

3.8.4.1 Open a layer

Open the Layer administration window by clicking on the menu button on the top left corner of the Analysis page.





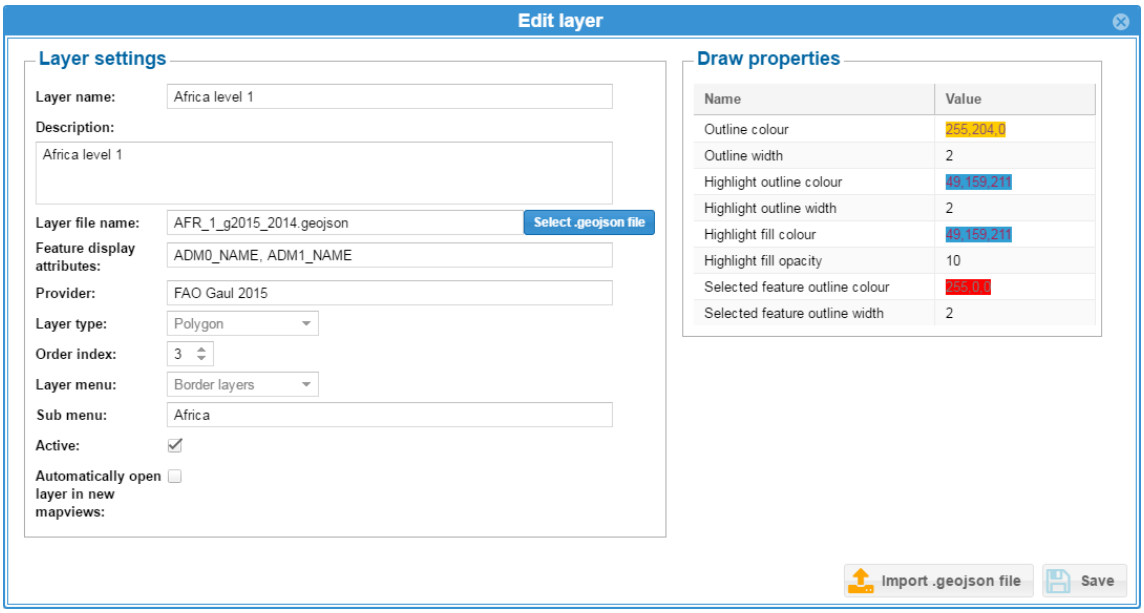
Edit	Layer name	Layer file name	Order index	Layer menu	Sub menu	Active	Load in Mapview	Delete
	Central Africa level 0	AFR_CENTRAL_0.geojson	4	Border layers	AU Region: CENTRAL	☑	☐	
	Central Africa area	AFR_CENTRAL_00.geojson	5	Border layers	AU Region: CENTRAL	☑	☐	
	Central Africa level 1	AFR_CENTRAL_1.geojson	3	Border layers	AU Region: CENTRAL	☑	☐	
	Central Africa level 2	AFR_CENTRAL_2.geojson	2	Border layers	AU Region: CENTRAL	☑	☐	
	East Africa level 0	AFR_EAST_0.geojson	4	Border layers	AU Region: EAST	☑	☐	
	East Africa area	AFR_EAST_00.geojson	5	Border layers	AU Region: EAST	☑	☐	
	East Africa level 1	AFR_EAST_1.geojson	3	Border layers	AU Region: EAST	☑	☐	
	East Africa level 2	AFR_EAST_2.geojson	2	Border layers	AU Region: EAST	☑	☐	
	North Africa level 0	AFR_NORTH_0.geojson	4	Border layers	AU Region: NORTH	☑	☐	
	North Africa area	AFR_NORTH_00.geojson	5	Border layers	AU Region: NORTH	☑	☐	
	North Africa level 1	AFR_NORTH_1.geojson	3	Border layers	AU Region: NORTH	☑	☐	
	North Africa level 2	AFR_NORTH_2.geojson	2	Border layers	AU Region: NORTH	☑	☐	
	South Africa level 0	AFR_SOUTH_0.geojson	4	Border layers	AU Region: SOUTH	☑	☐	
	South Africa area	AFR_SOUTH_00.geojson	5	Border layers	AU Region: SOUTH	☑	☐	
	South Africa level 1	AFR_SOUTH_1.geojson	3	Border layers	AU Region: SOUTH	☑	☐	
	South Africa level 2	AFR_SOUTH_2.geojson	2	Border layers	AU Region: SOUTH	☑	☐	
	West Africa level 0	AFR_WEST_0.geojson	4	Border layers	AU Region: WEST	☑	☐	
	West Africa area	AFR_WEST_00.geojson	5	Border layers	AU Region: WEST	☑	☐	
	West Africa level 1	AFR_WEST_1.geojson	3	Border layers	AU Region: WEST	☑	☐	

In the Layer administration window, you can quickly activate or de-activate a layer to be shown as an item under the layer menu button  in Mapviews.

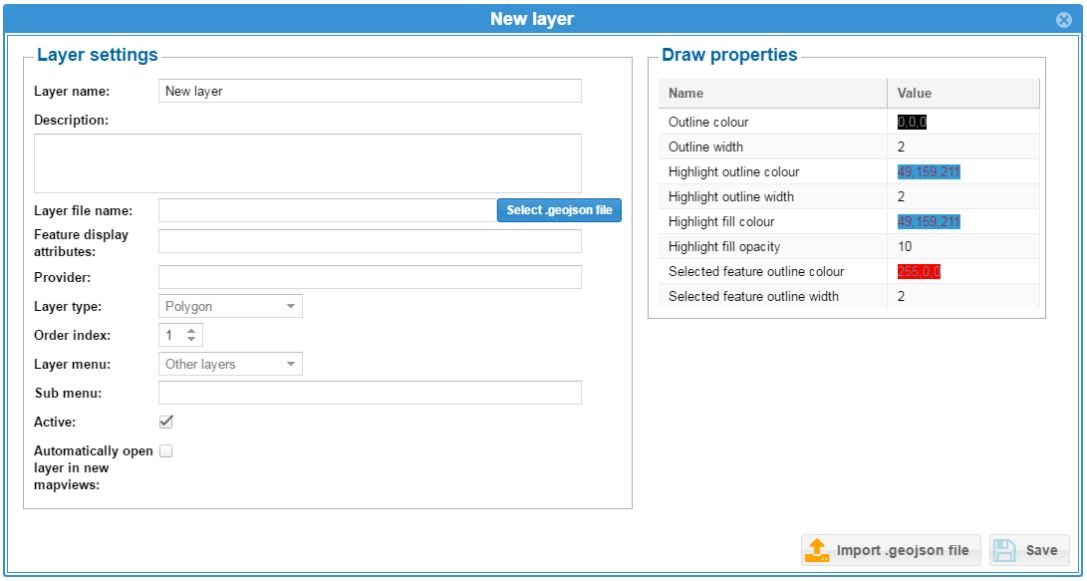
Under the "Load in Mapview" column you can quickly set a layer to be automatically loaded in a new Mapview. Setting too many (big) layers to be automatically loaded in a new Mapview, could crash the Browser! So, it is advised to set maximum 3 layers to be loaded automatically.

3.8.4.2 Edit and add a layer

Edit a layer's properties by clicking on the icon next to the layer.



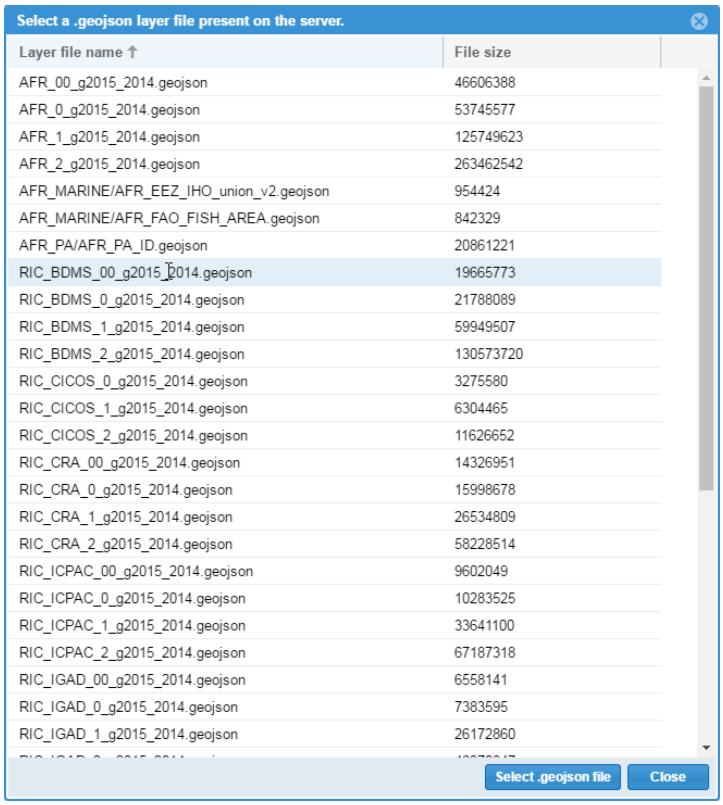
Add a new layer by clicking on the Add layer button.



Layer settings:

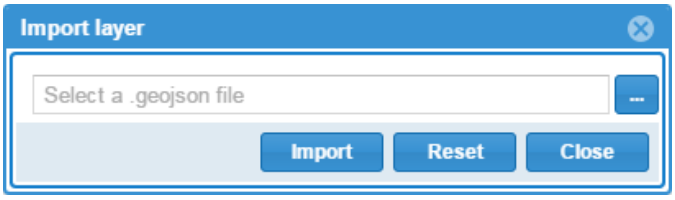
	Description
Layer name	The name of the layer as it appears in the menu.
Description	A description of the layer.
Layer file name	The name of the .geojson file present on the EStation under the directory /EStation/layers/ can only be selected from the list shown when clicking on the Select .geojson file button (see below).
Feature display attributes	A comma delimited list of feature attribute names that are present in the .geojson layer file. These attributes are shown in the tool bar area of a Mapview when going over a feature and in the Selected region under the time series area when selecting a feature. To find out which attributes are available in a vector layer file, open the file in QGIS and then open its attribute table.
Provider	The provider/creator of the layer. For example, FAO Gaul 2015.
Layer type	The type of layer: Polygon, Line or Point. Only support for Polygon layers has been implemented for now.
Order index	The priority given to the layer for selecting an area/feature. The highest priority is 1 and the lowest priority is 5 (e.g., for level 00 border layers).
Layer menu	The layer menu has 3 main menu items, Border layers, Marine layers and Other layers. A layer must belong to one of these main menu items.
Sub menu	Under the 3 main menu items sub menus can be defined. In this field you can give a name of the sub menu under which the layer will fall.
Active	A layer will appear in the layer menu when it is activated and will not appear in the layer menu when the layer is de-activated.
Automatically open in new mapviews	If this field is checked, the layer will automatically be loaded in a new Mapview.

Select a layer present on the EStation by clicking on the Select .geojson file button. A list is shown of all the .geojson files present under the directory /EStation/layers/ on the EStation.



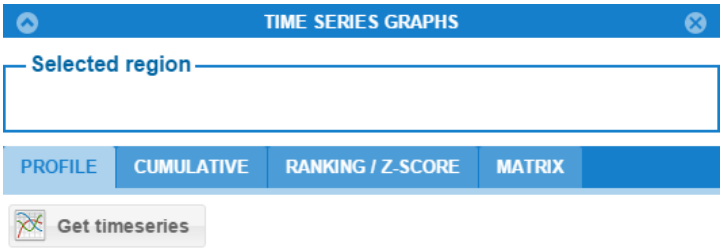
3.8.4.3 Import a layer

Import a layer file by clicking on the Import .geojson file button. The layer has to be in .geojson format, which is the only format supported in a browser by the JavaScript library Openlayers.



3.8.5 Graph view functionalities

In this version there are 4 types of graphs available.



Profile	**To display the evolution (X axis) profile of a product value (Y axis) for a given period**
Cumulative	To visualize the cumulated value (Y axis) of a product/indicator along the time (X axis)
Ranking / Z-Score	Sort the historical values of a product for a given period/year Display the normalize value (y axis) of a product for a given period/year along the historical years (X axis)
Matrix	To display the seasonal(X-axis) value of a product along the historical years(Y-axis)

Each type has its own purpose, options and restrictions.

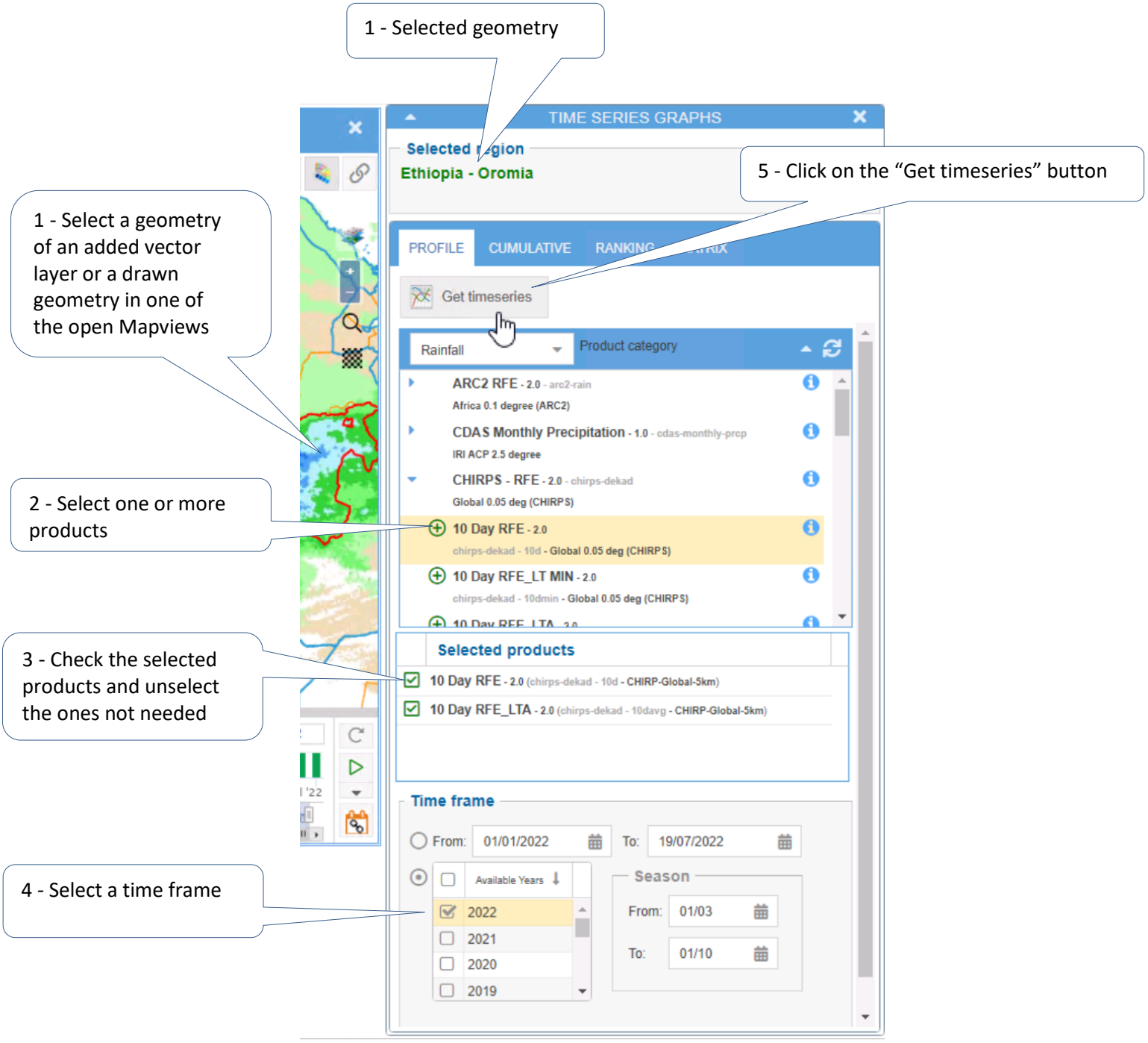
IMPORTANT!

To generate time series, visualized in a chart/graph, the user will have to select an area of interest from an added vector layer or a drawn geometry in one of the open Mapviews. The geometry of the selected area will be highlighted in **red** in the Mapview and its name shown in the "Selected region" box in the TIME SERIES GRAPHS panel.

3.8.5.1 General steps to generate a time series graph

General steps to generate a time series graph, to be taken in the TIME SERIES GRAPHS panel on the right of the analysis tool:

1. Select an area of an added vector layer or a drawn geometry in one of the open Mapviews.
2. Select one or more sub products.
3. Check the selected products and unselect the ones not needed.
4. Select a time frame
5. Click on the "Get timeseries" button



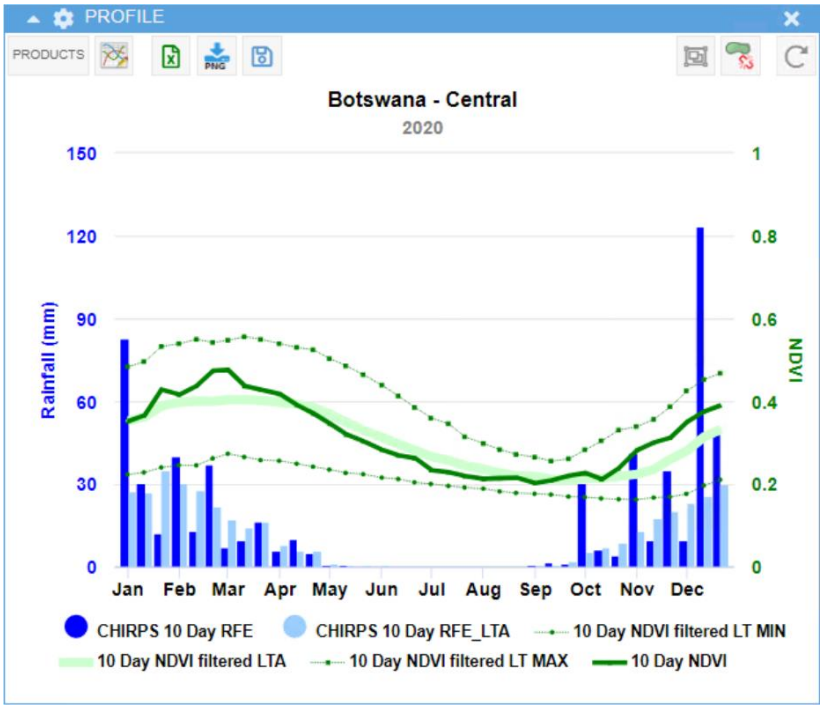
A new time series graph window will be opened, showing the chart type with the time series of the selected data sets, over the selected region of the chosen time frame.

All graphs are interactive and will show the time series values of each time series when moving the mouse of the chart.

3.8.5.2 Profile graph

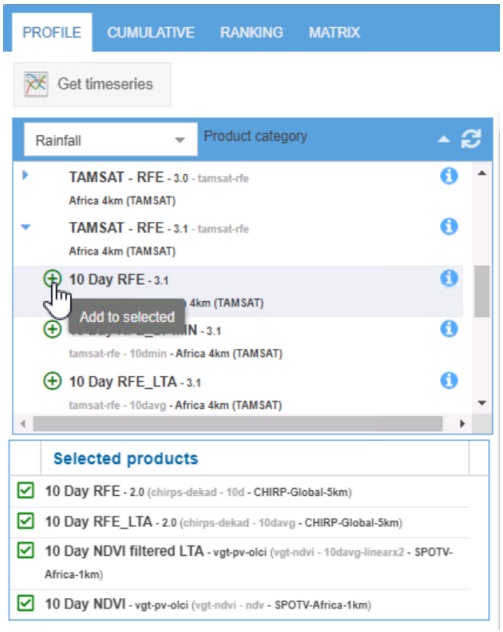
Description

The profiles graph is used to display along the X axis time period from one to several Y axes products variables with different units and scales. The data can be displayed as a bar or line with different line thicknesses and colours…The font and styles of the titles and annotation can be interactively edited. The title can use the attributes name index of the entity (point/line/polygon) selected and from which the indicators value is calculated (count, average, cumulate, percentage, surface) as shown below.

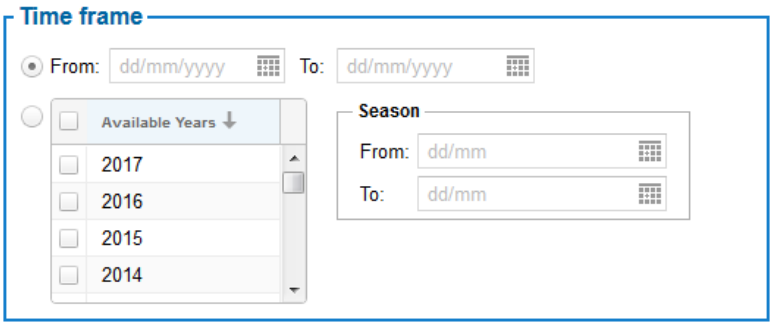


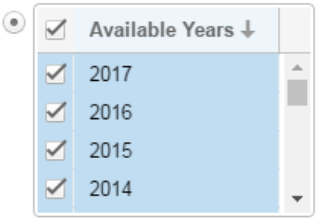
Product selection

Multiple products can be selected from multiple product categories.



Time frame selection

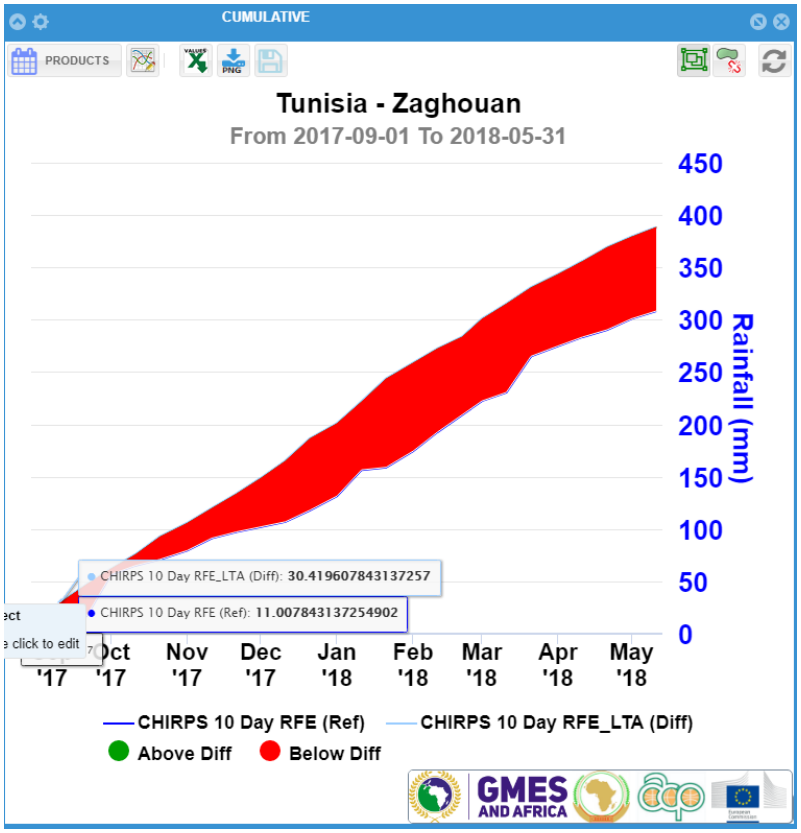


From – To	The "From-To" time frame can span over multiple years. Format dd/mm/yyyy.
Available Years	Multiple years can be selected from the list of available years of the selected sub products. In case of multiple years, the graph will show the time series of each sub products of each selected year. Multiple years can be selected from the list of available years of the selected product. The checkbox near the "Available Years" title will select all the years. 
Compare seasons	For the selected year(s) of interest give a season period in the "From-To" fields. The season period can span over 2 years, starting for example from 01/09 of the selected year(s) and ending on 01/05 of the following year. 

3.8.5.3 Cumulative graph

Description

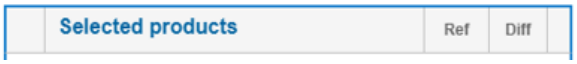
The graph displays the accumulation of the Y axis product variable from the first to the last dates selected on the X axis. It can be compared to another variable typically a specific year compared to the Long-Term Average (or minimum or maximum) of the same variable. The positive difference is shown in green when the negative is red. Like for the profile graph it is possible to edit the line thickness and colour, the titles and axis fonts and colours.



Product selection

Multiple products can be selected from multiple product categories.

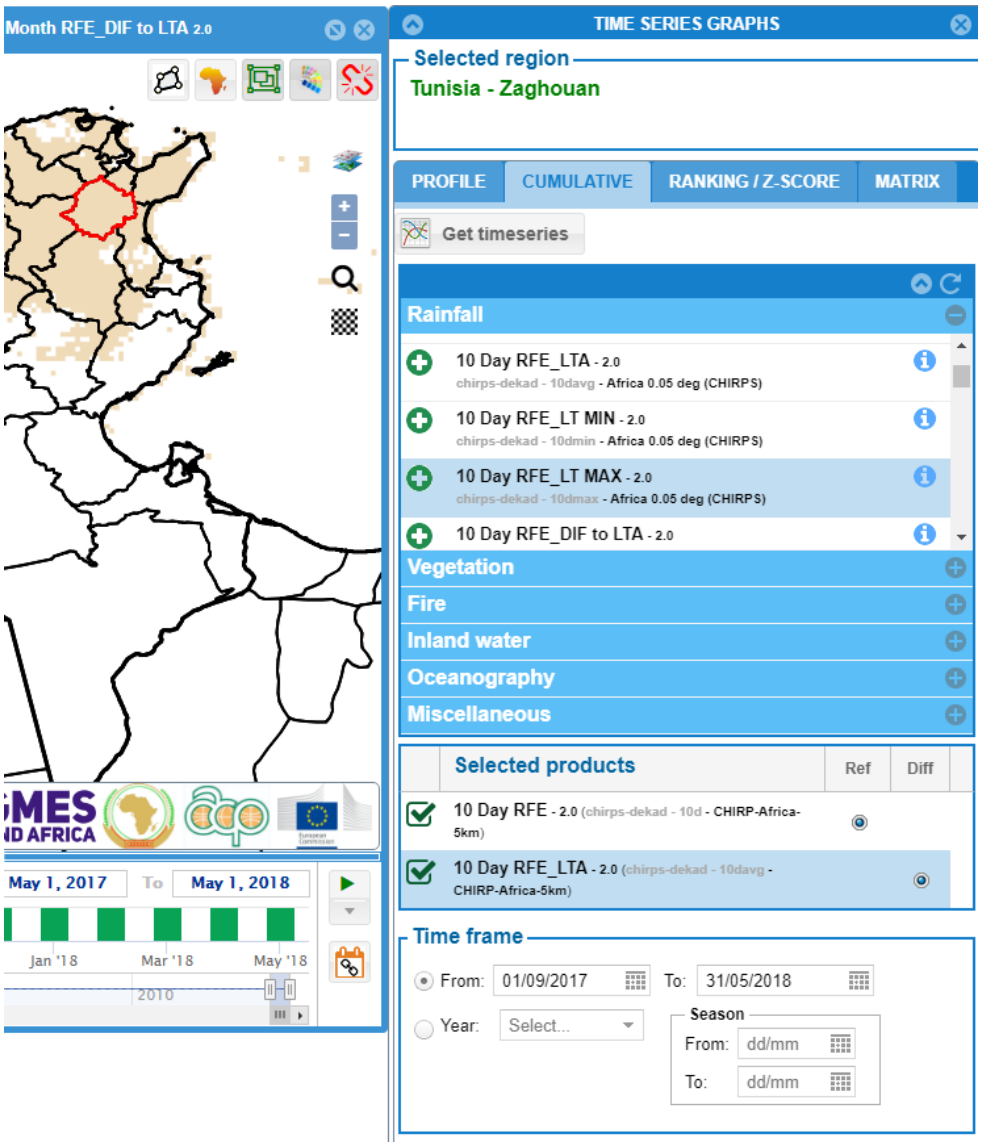
All selected products will be a cumulative line in the time series graph.

In 

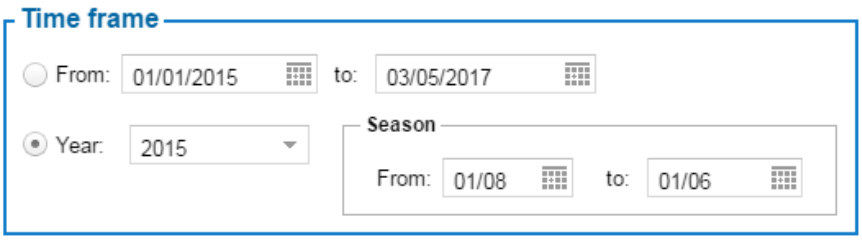
you have the option to indicate what is the **'Reference' (Ref)** product or the **'Difference' (Diff)** product:

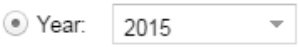
- when 'Ref' is **below** 'Diff' a **RED** colour will be shown.
- when 'Ref' is **above** 'Diff' a **GREEN** colour will appear.

Only one product is the Reference and one product the Difference.



Time frame selection

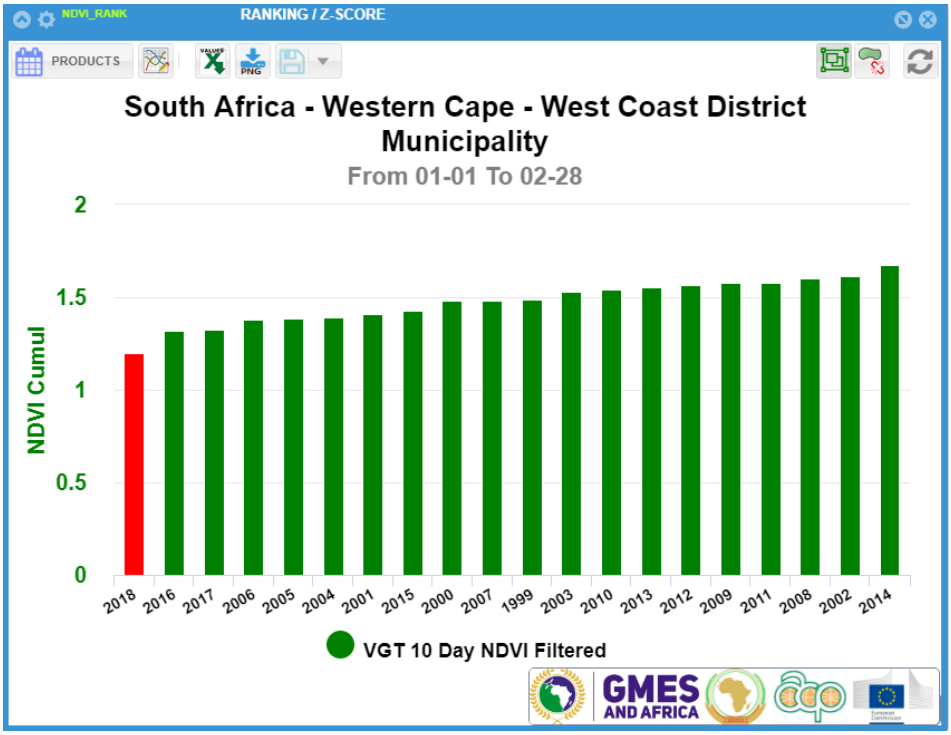


From – To	The "From-To" time frame can span over multiple years. Format dd/mm/yyyy.
Year	Only one year can be selected from the list of available years of the selected products. 
Compare seasons	For the selected year give a season period in the "From-To" fields. The season period can span over 2 years, starting for example from 01/08 of the selected year(s) and ending on 01/06 of the following year. 

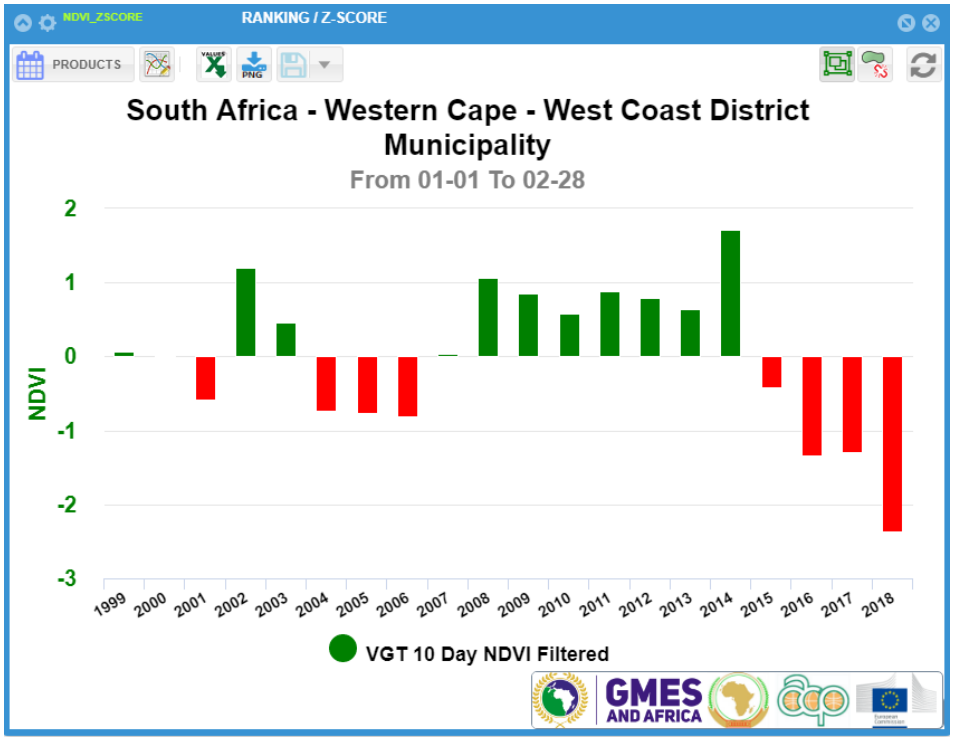
3.8.5.4 Ranking / Z-Score graph

Description

To better compare the product value of a given period for a specific year with the historical time series it is possible to apply a ranking from the lowest to the highest corresponding value. It clearly shows the exceptional years with the highest or lowest values of the selected products.



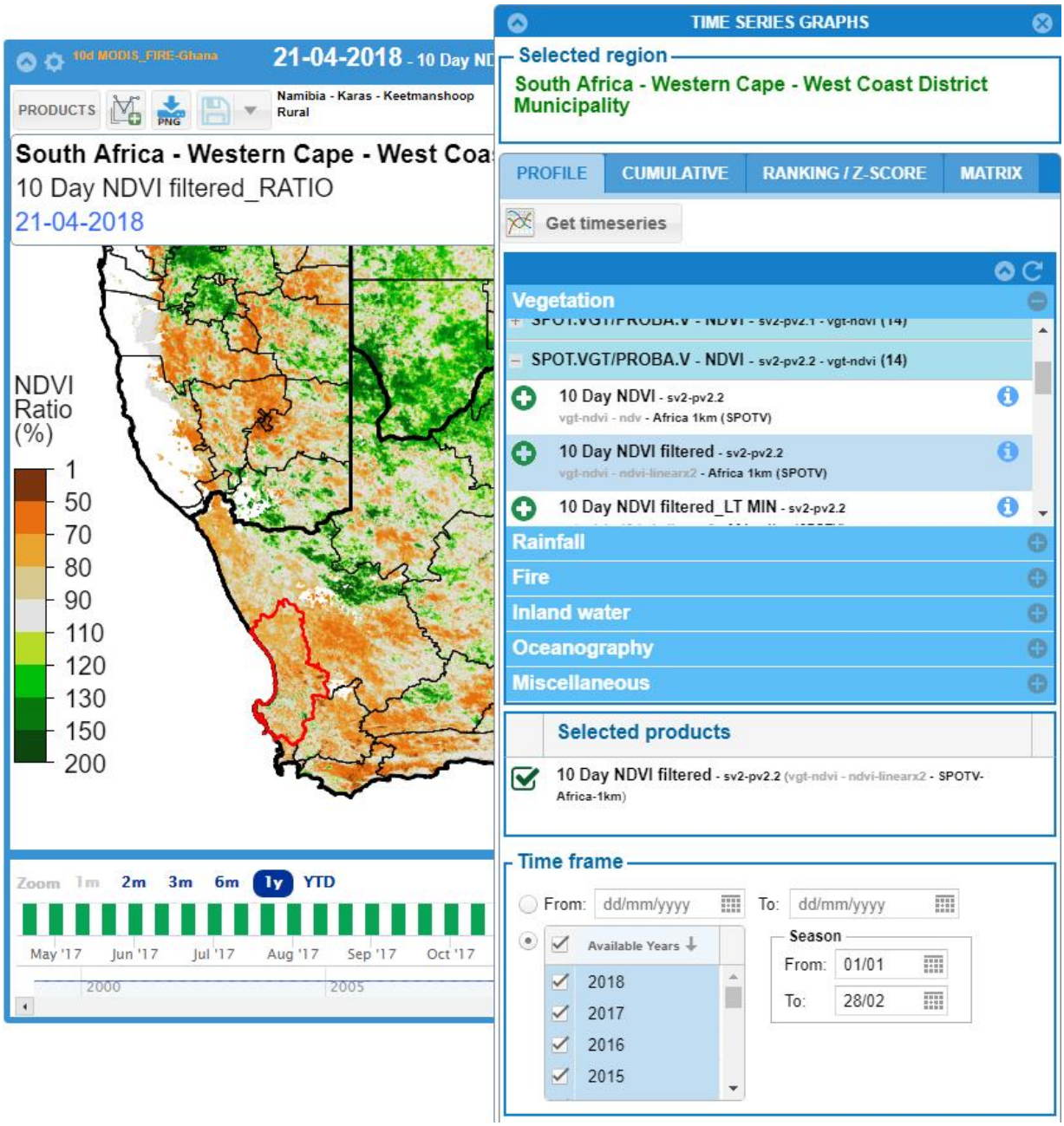
Moreover, it is possible to display on the Y axis the Z-Score value [(Year value–Mean value)/Standard deviation] along the years X axis in a chronological order as shown below.



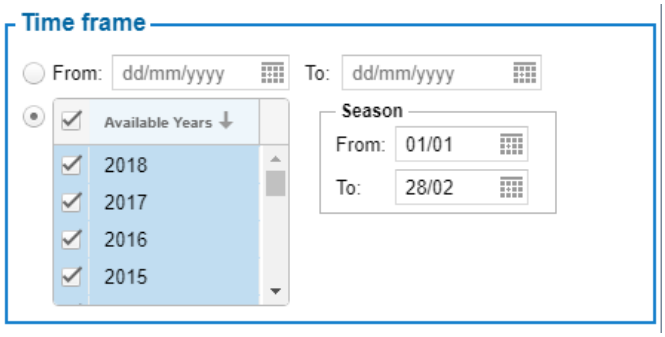
Product selection

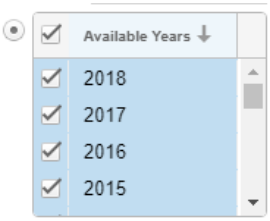
Only one product can be selected for the Ranking / Z-Score.

For the selected product you have the option to visualize the Z-Score instead of Ranking.



Time frame selection

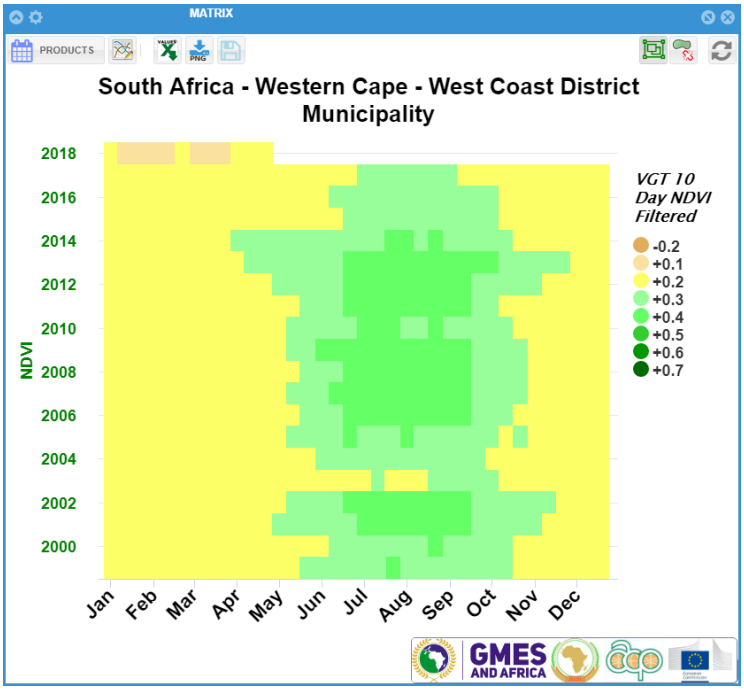


Available Years	Multiple years can be selected from the list of available years of the selected product. The checkbox near the "Available Years" title will select all the years. With a Ranking and Z-Score graph you normally select all the years. 
Compare seasons	For the selected years give a season period in the "From-To" fields. The season period can span over 2 years, starting for example from 01/01 of the selected year(s) and ending on 28/02 of the following year. 

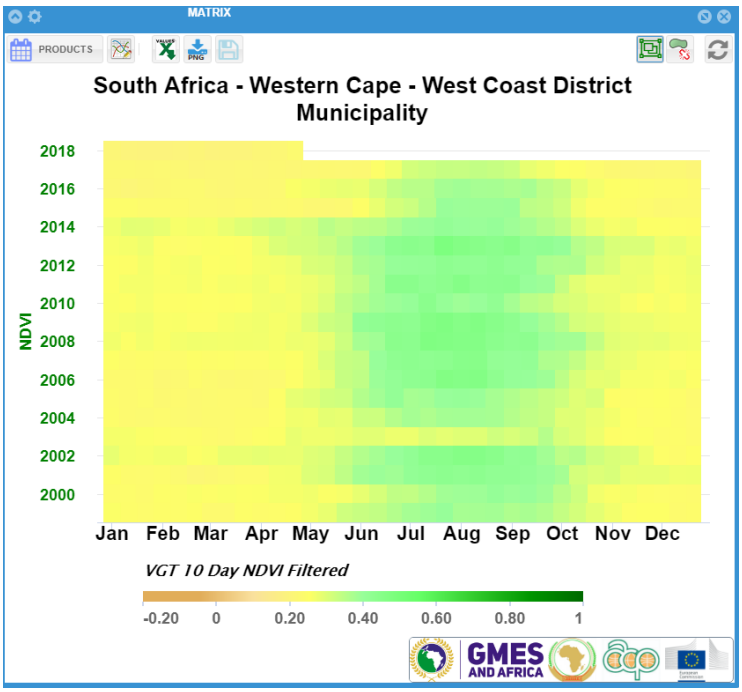
3.8.5.5 Matrix graph

Description

The Matrix graph (or Heatmap) is used to get a synoptic overview of the historical evolution of a specific product (like rainfall, temperature, NDVI…). It facilitated the identification of exceptional years and the inter years comparison. It displays on the X axis a specific seasonal value over the different historic year on the Y axis. To be consistent with the map the user can interactively select any existing map colour palette of a given product.



Full color palette display



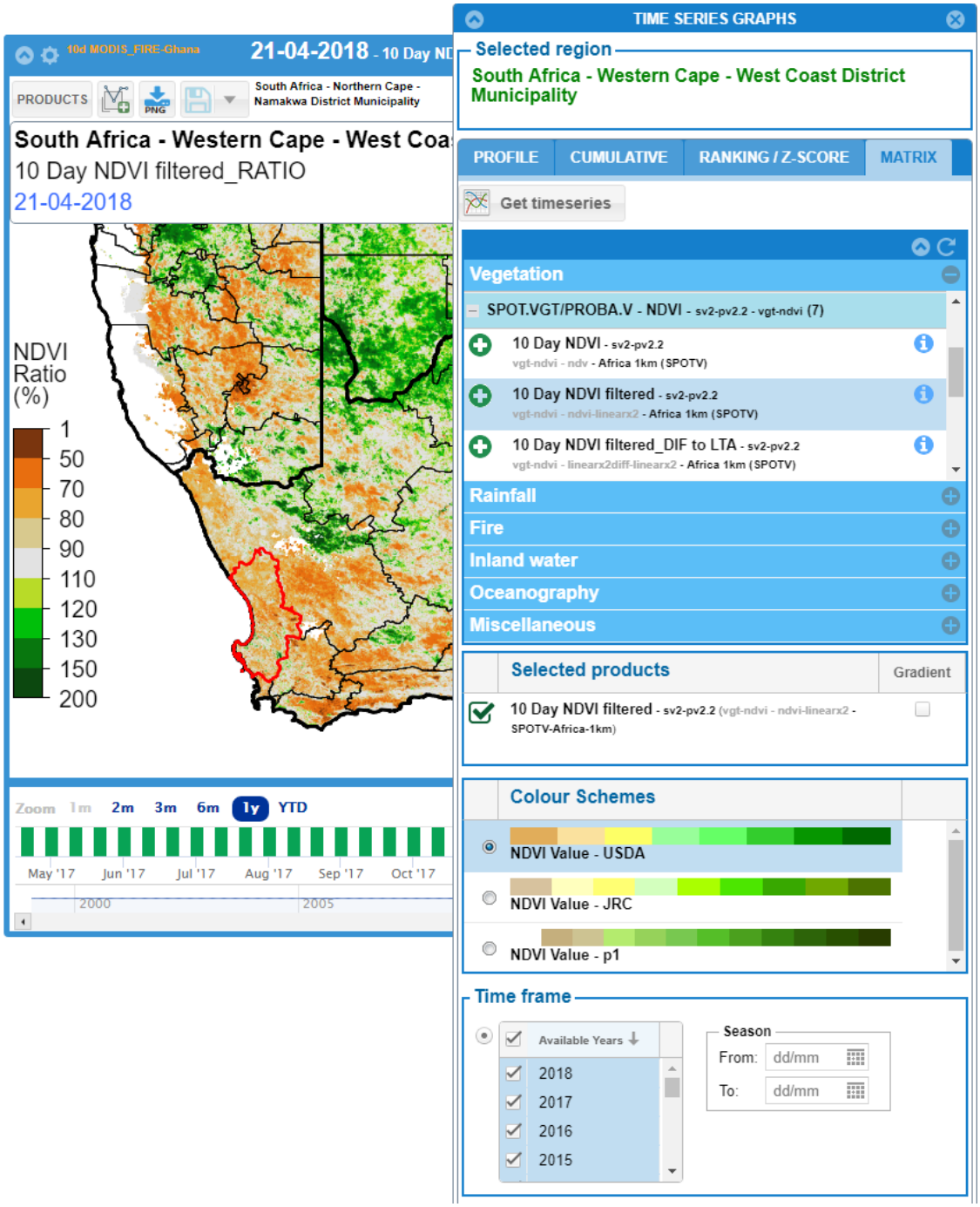
Gradient color palette display

Product and colour scheme selection

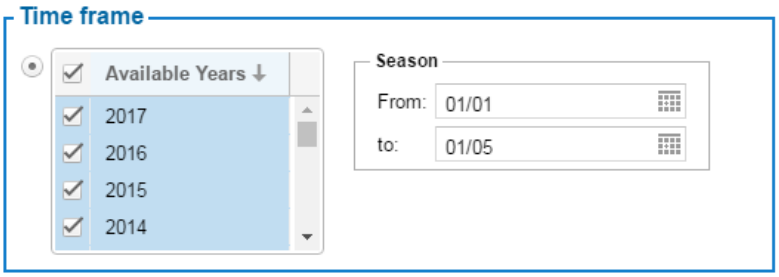
Only one product can be selected for the Matrix graph.

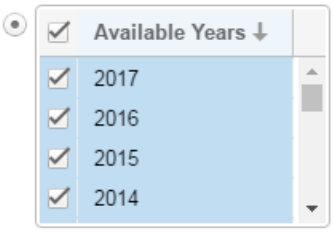
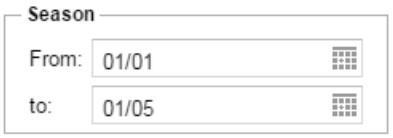
Under the selected product you will find its available colour schemes, where the default colour scheme is selected. Choose the colour scheme you want to apply to the Matrix graph.

For the selected product you have the option to visualize the Matrix in gradient, which means that apart from the colour steps of the selected colour scheme, intermediate colours are generated. For colour schemes with more than 25 steps, the gradient still will automatically be taken.



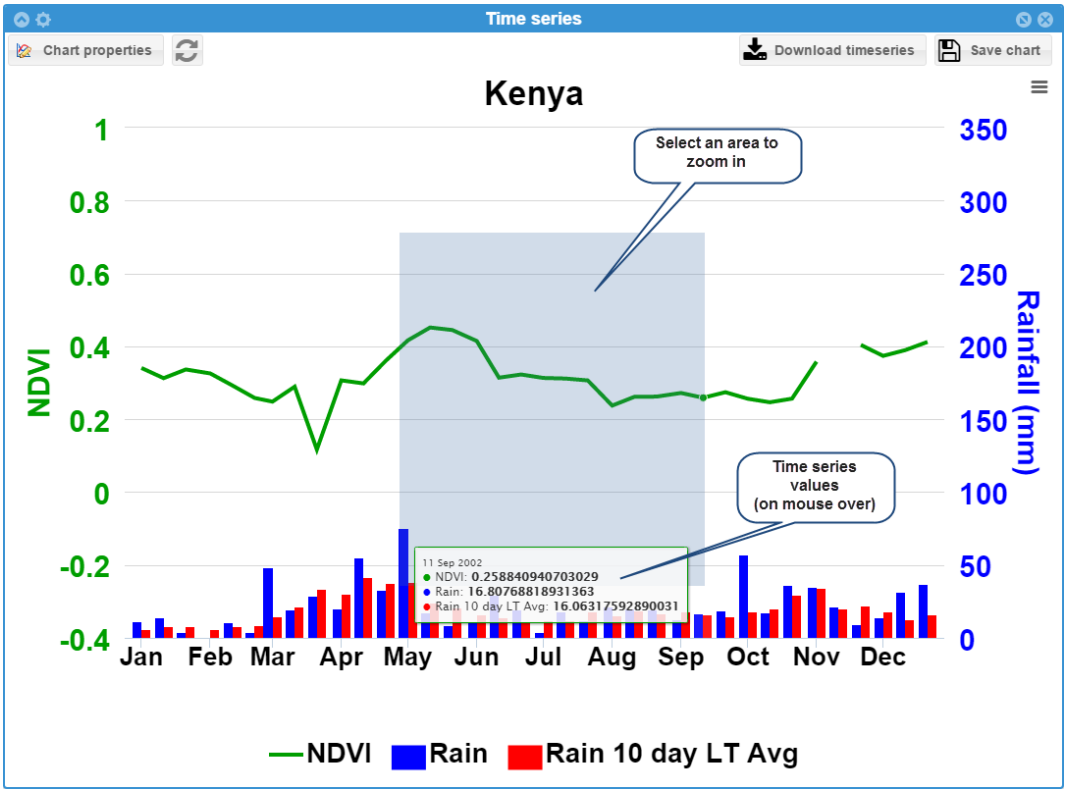
Time frame selection



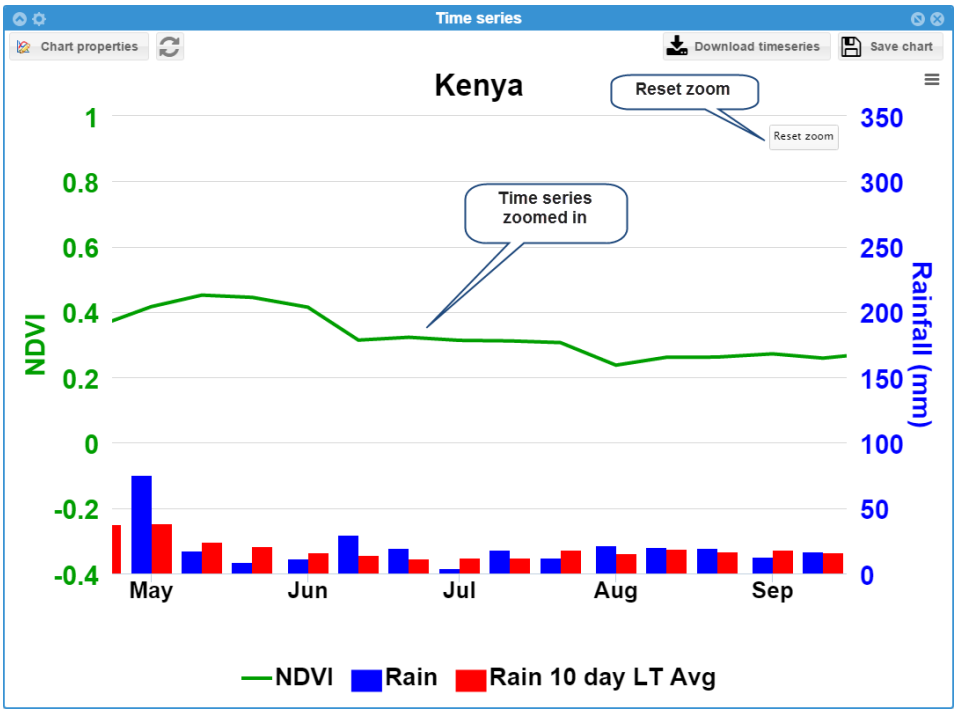
Available Years	Multiple years can be selected from the list of available years of the selected product. The checkbox near the "Available Years" title will select all the years. 
Compare seasons	For the selected years give a season period in the "From-To" fields. The season period can span over 2 years, starting for example from 01/09 of the selected year(s) and ending on 01/04 of the following year. 

3.8.5.6 Chart zooming

You can also zoom in by selecting an area in the chart. The chart will change and show only the plot over the selected zoom.

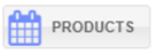


To zoom out, click on the Reset zoom button in the chart.



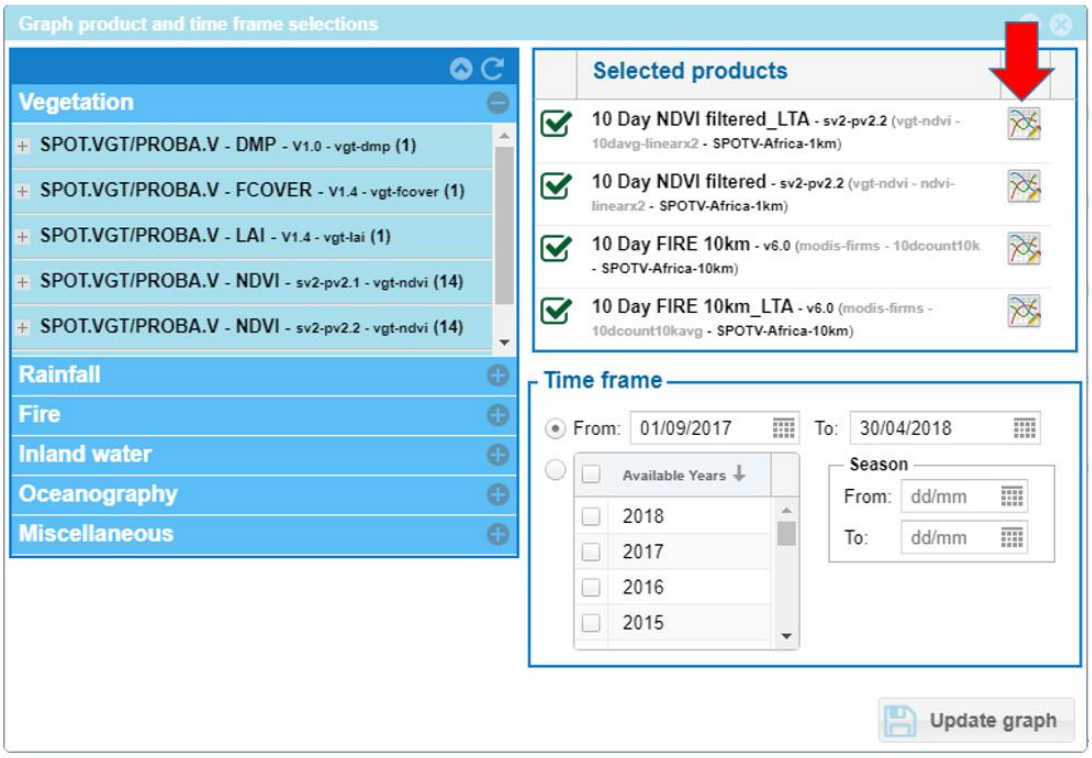
3.8.5.7 Draw properties

To modify a graph, you can edit the draw and the chart properties through two specific icons.

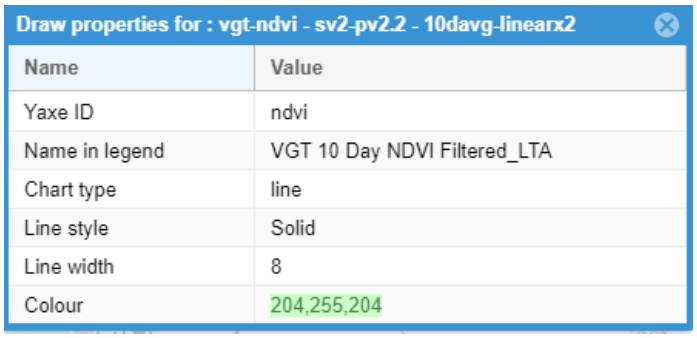
For all the available time series datasets default draw properties are given. To change these draw properties, click on the 

icon on the top left of the graph window.

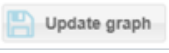
A new window will appear and to modify the draw properties of a specific products click on the corresponding icon  .



The time series draw properties window will pop up for the dataset in question.



Yaxe ID	**The Yaxe ID defines to which Yaxe the time series dataset belongs. When 2 or more time series datasets belonging to the same Yaxe ID are selected, then these datasets will fall under the same Yaxe in the chart. If the selected time series datasets belong to different Yaxe IDs, then for each Yaxe ID a Yaxe is created in the chart. Each Yaxe ID has common properties that can be changed in the chart properties (see paragraph 3.9.4.2).**
Name in legend	The name of the time series in the legend of the chart.
Chart type	The chart type, Line or Column.
Line style	The style of the line if chart type is Line. Styles are for example: Solid, Dot, Dash, DashDot, etc.
Line width	The width of the line if chart type is Line.
Colour	The colour of the time series dataset in the chart.

To apply the modification on the graph, click on the update button 

.

3.8.5.8 Chart properties

To open the chart properties, click on the  button, found in the tool bar of a chart window.

Chart properties are the properties that apply to all charts generated, plus the properties of each X and Y Axis in the generated chart. Changing a property value will be applied immediately to the chart, except for the aggregation fields.

To apply the aggregation field changes, refresh the chart by clicking on the  button.

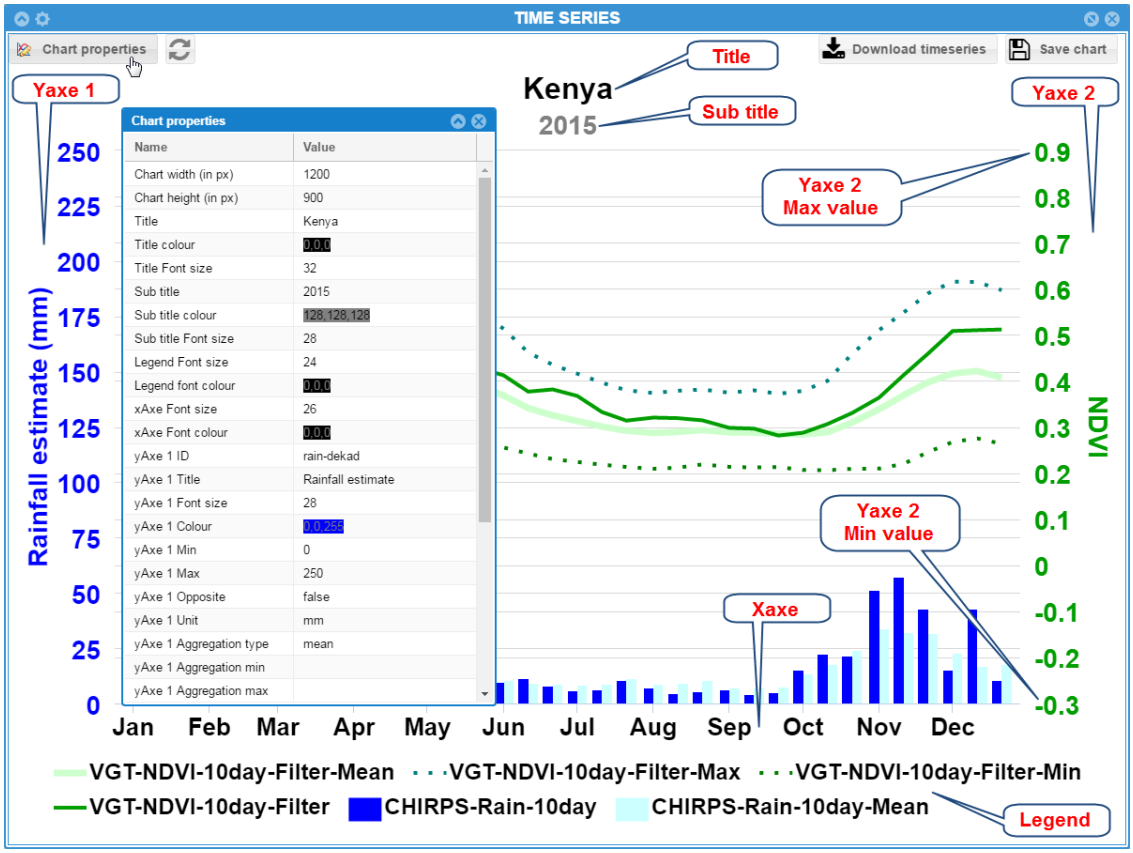


Chart width (in px)	**The width of the chart in pixels.**
Chart height (in px)	The height of the chart in pixels.
Title	The chart title that by default is the name of the selected region.
Title colour	The colour of the chart title.
Title Font size	The font size of the chart title.
Subtitle	The chart subtitle that by default is the selected time period.
Subtitle colour	The colour of the chart subtitle.
Subtitle Font size	The font size of the chart subtitle.
Legend Font size	The font size of the chart's legend items.
Legend Font colour	The font colour of the chart's legend items.
xAxe Font size	The font size of the xAxe.
xAxe Font colour	The font colour of the xAxe.
yAxe # ID	The yAxe ID uniquely identifying the yAxe fields below. One or more time series datasets can have the same yAxe ID assigned, indicating that these datasets will appear on the same yAxe in the chart, when selected. This field cannot be edited here in the chart properties. You can change the yAxe ID for each time series dataset individually editing the time series draw properties (see paragraph 3.9.5.1 Time series draw properties).
yAxe # Title	The yAxe title.

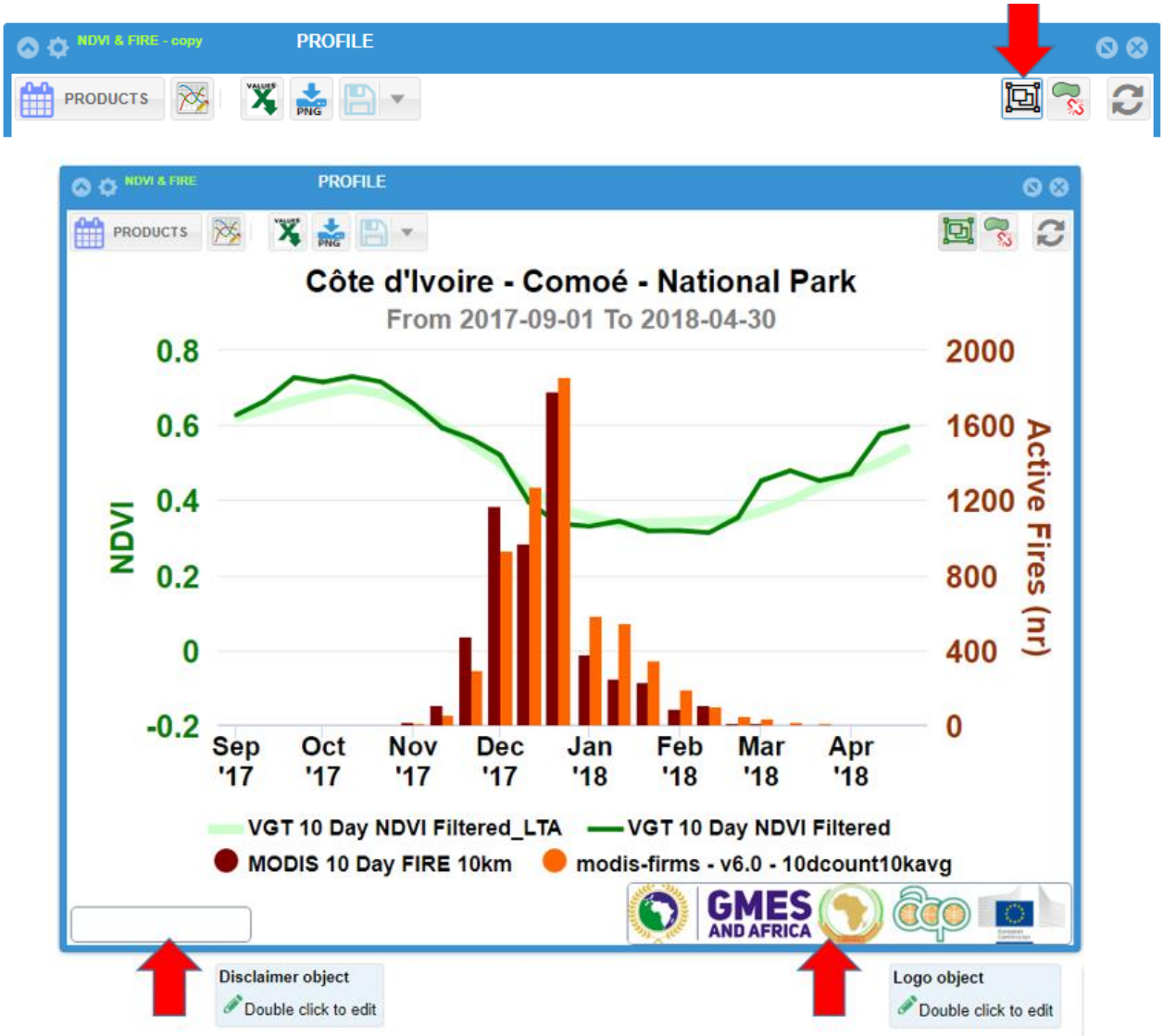
yAxe # Font size	The font size of the yAxe title and values.
yAxe # Colour	The colour of the yAxe title and values.
yAxe # Min	Set the yAxe minimum value to start with.
yAxe # Max	Set the yAxe maximum value.
yAxe # Opposite	If TRUE, the yAxe will be placed on the opposite/right side of the chart. If FALSE, the yAxe will be placed on the left side of the chart.
yAxe # Unit	The unit of the time series yAxe, shown in the yAxe title.
yAxe # Aggregation type	Type of aggregation applied to compute the value at polygon level from pixels. Possible values are: • mean: average of all valid[6] values in the polygon. 'Nodata' values are excluded, while aggregation min/max are not considered. Example: vegetation indicators (vgt-ndvi) • cumulate: arithmetic sum of all valid values in the polygon; aggregation min/max are not considered. Example: precipitation (fewsnet-rfe). • count: returns the number of valid pixels in the polygon having values in the range aggregation min to aggregation max. Example: Active Fires (modis-firms) • percent: returns the percent of valid pixels in the polygon having values in the range aggregation min to aggregation max. Example: in vegetation anomaly indicators, to identify strong anomalies. • surface: geographic extension of the area – within the polygon – where pixels are in range aggregation min to aggregation max. Example: water bodies (wd-gee)
yAxe # Aggregation min	Used for aggregation type 'count', 'percent' and 'surface' as described above.
yAxe # Aggregation max	Used for aggregation type 'count', 'percent' and 'surface' as described above.

[6] In the current context, 'valid' means that we exclude the 'nodata' values.

3.8.5.9 Logo and disclaimer objects

Like for the map, the graph has Logo and Disclaimer objects that you can show and hide in the graph viewer.

To show the objects, click on the button. The button becomes green and the two objects are shown in the graph area.

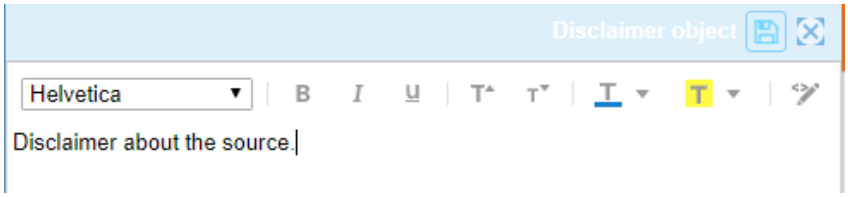


You can move the two objects by click and hold on an object and drag it to reposition.

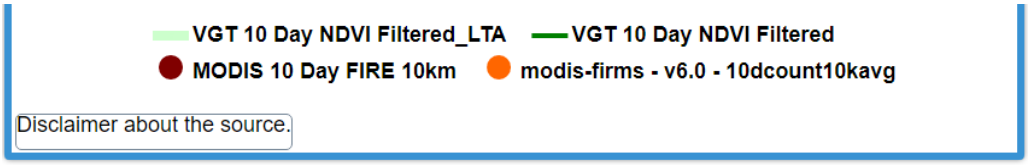
3.8.5.9.1 Edit Disclaimer object

Double-click the Disclaimer object to edit it. The disclaimer editor is shown. You can type any text and format the text using the available text formatting tools in the editor.

- To save your changes, click on the save button found in the header of the editor.



- The editor will close, and the changes reflect in the disclaimer object of the Chart view.

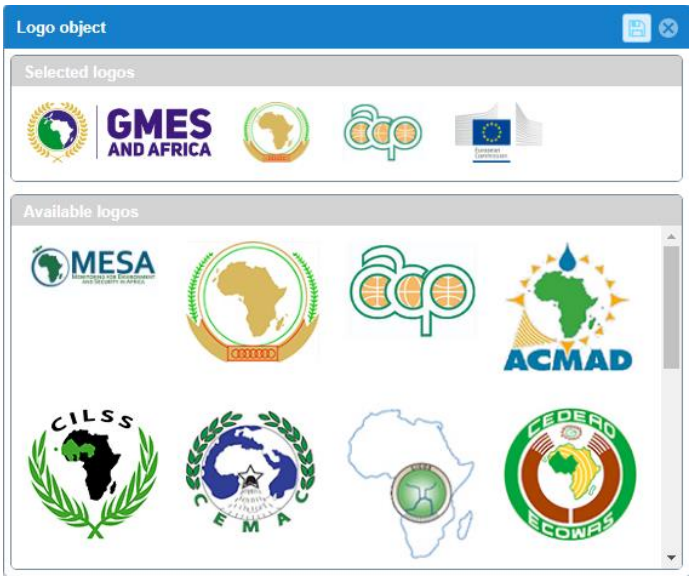


3.8.5.9.2 Edit Logo object

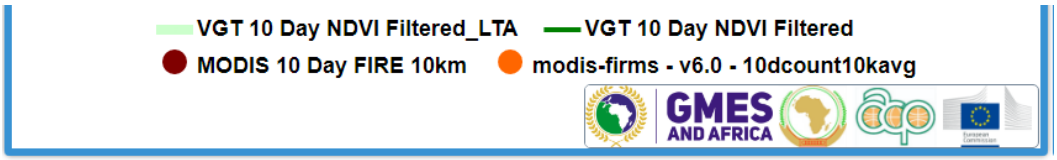
Double-click the Logo object to edit it. The logo editor is shown.

The upper box contains the selected logo's that will appear in the logo object. The lower box contains the available logos.

- Double-click on a logo in the selected logo's box to remove the logo.
- Double-click on a logo in the available logo's box to add the logo to the selected box.
- To save your changes, click on the save button found in the header of the editor.



- The editor will close, and the changes reflect in the disclaimer object of the Chart view.



3.8.5.10 Dynamic/Static graph values

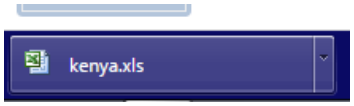
To explore the different values aggregated from the layer polygons/points the user can activate the exploratory mode icon from OFF to ON in the graph window tool bar.

When activated the user can select any polygon/point from the layer display in the map view and the corresponding products values will be displayed interactively.

3.8.5.11 Download timeseries

To download the time series, click on the "Download timeseries" button present in the tool bar of the "Time series chart" window.

The time series will be saved in XLS format and automatically downloaded in the download directory of the browser.



A	B	C	D
DateTime	NDVI	Rain	Rain 10 day LT
1/1/2002 0:00	0.34	11.25	6.05
1/11/2002 0:00	0.31	13.81	8.38
1/21/2002 0:00	0.34	4.37	8.05
2/1/2002 0:00	0.33	0.62	6.12
2/11/2002 0:00	0.29	10.59	7.80
2/21/2002 0:00	0.26	3.72	8.78
3/1/2002 0:00	0.25	48.48	14.56
3/11/2002 0:00	0.29	19.57	21.84
3/21/2002 0:00	0.12	28.63	33.64
4/1/2002 0:00	0.31	20.46	30.13
4/11/2002 0:00	0.30	54.86	41.44
4/21/2002 0:00	0.36	32.98	37.50
5/1/2002 0:00	0.42	75.63	38.13
5/11/2002 0:00	0.45	17.41	24.51
5/21/2002 0:00	0.44	9.05	21.01
6/1/2002 0:00	0.41	11.33	16.05
6/11/2002 0:00	0.31	29.43	13.91

3.8.5.12 Save graph as PNG image

To export a time series chart, click on the “PNG” button  present in the tool bar of the “Time series chart” window.

A snapshot of the chart will be made in PNG format and automatically saved in the download directory of the browser.

3.8.5.13 Graph template

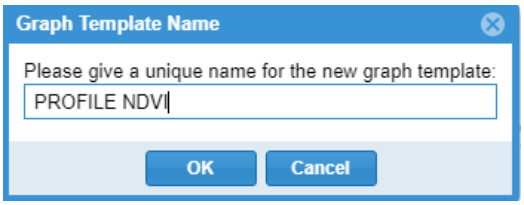
In this new EStation version, the user as the possibility to save a graph as a template and manage a list of graph templates. The resulting template will keep the window size, the logo and disclaimer objects size and position, the chart and draw properties. There are two possibilities of saving a template: when it is a new graph or an existing one.

Like for the maps you **must be logged in** to have the possibility **to save your graph templates**.

- Save a new graph

When it is a new graph, the user can save the graph with the button in the graph window toolbar.

A new window popup to type the name of the template as below:



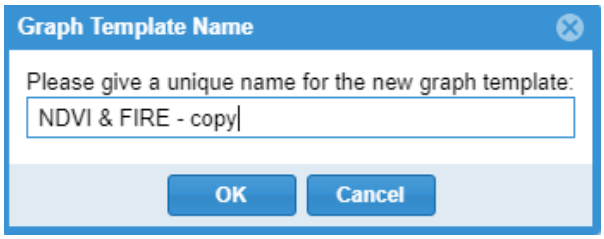
The template will be saved in the template repository.

- Save an existing graph

If the graph exists already, the toolbar will show an extra button as below:

To just update the existing graph with the same name, click on the “Save » icon

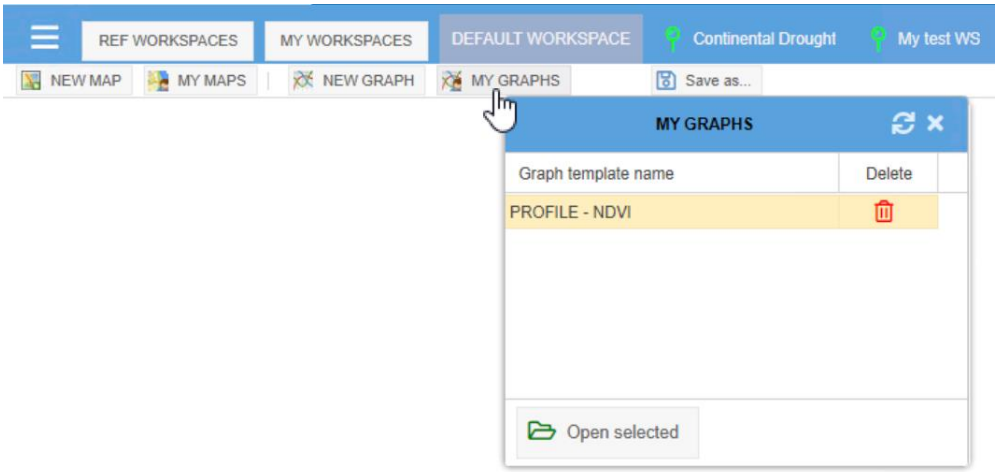
To use the display graph as template for further editing “Save as” icon  in the tool bar. Then a window will pop up to enter a new graph template name.



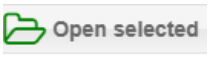
And will be saved in the graph template repository.

3.8.5.14 Manage the graph templates

Like for the maps there is now a specific repository for the different graph templates call “MY GRAPHS”.



By selecting the line of interest, the user can

- load any existing template by using  in the current workspace
- or delete it from the repository with  icon.

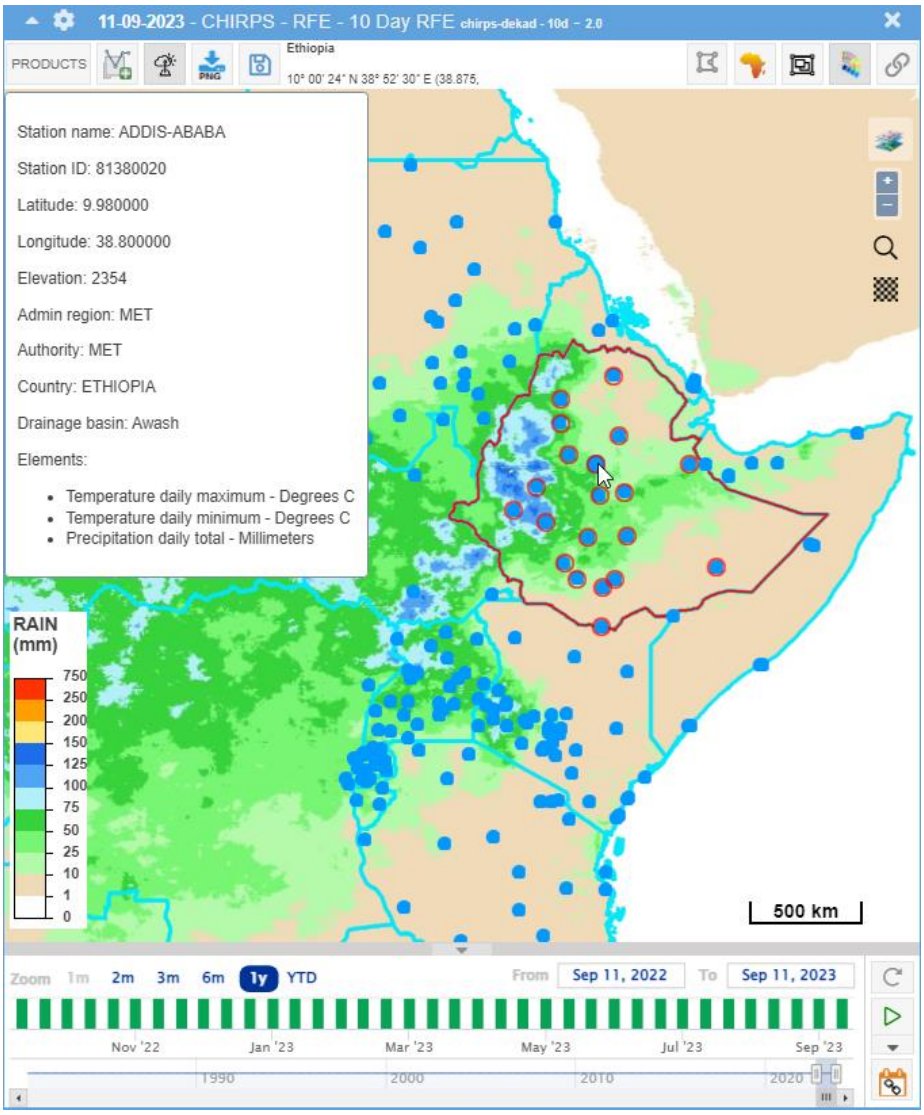
3.8.6 Weather Station data

The analysis tool contains the functionality to view the location of the imported stations within a Mapview as point data, and to generate time series graphs for the observation elements of the selected stations.

3.8.6.1 Viewing weather station location in a Mapview

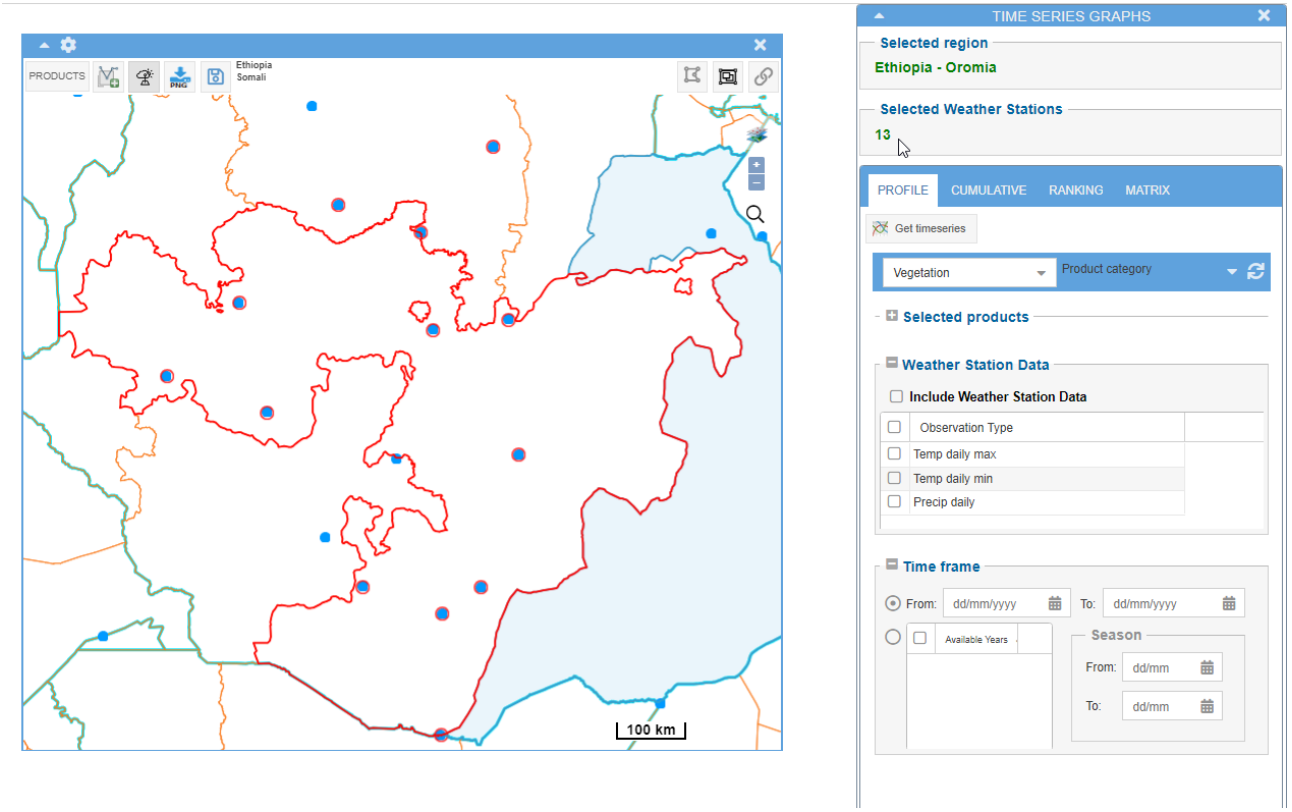
In the toolbar of a Mapview, you will find a toggle button to show or hide the station point data within the map. The station point data can only be shown in one Mapview.

When the stations are shown in the map, you can see a station's information by hovering over a station point, as shown below.



Selecting multiple stations can be done in two ways.
The first is by selecting a region from an in the map added vector layer, like a border layer or saved drawn layer. When you select a region, automatically all stations within the selected region are also selected.

The second method is by holding the CTRL key on your keyboard and select multiple station points.



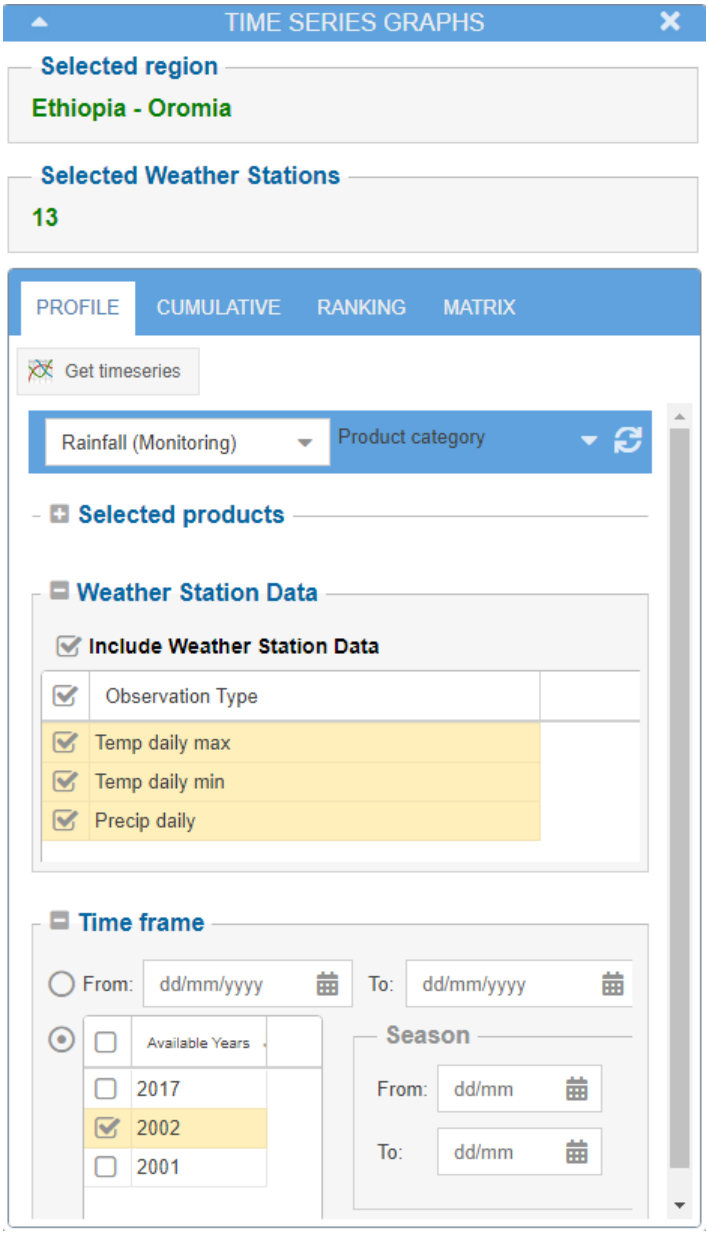
3.8.6.2 PROFILE graph with weather station time series

When the station point data is shown in a Mapview, you will see under the PROFILE graph new sections. A section “Selected Weather Stations” is added, where you will see the **total selected stations**.

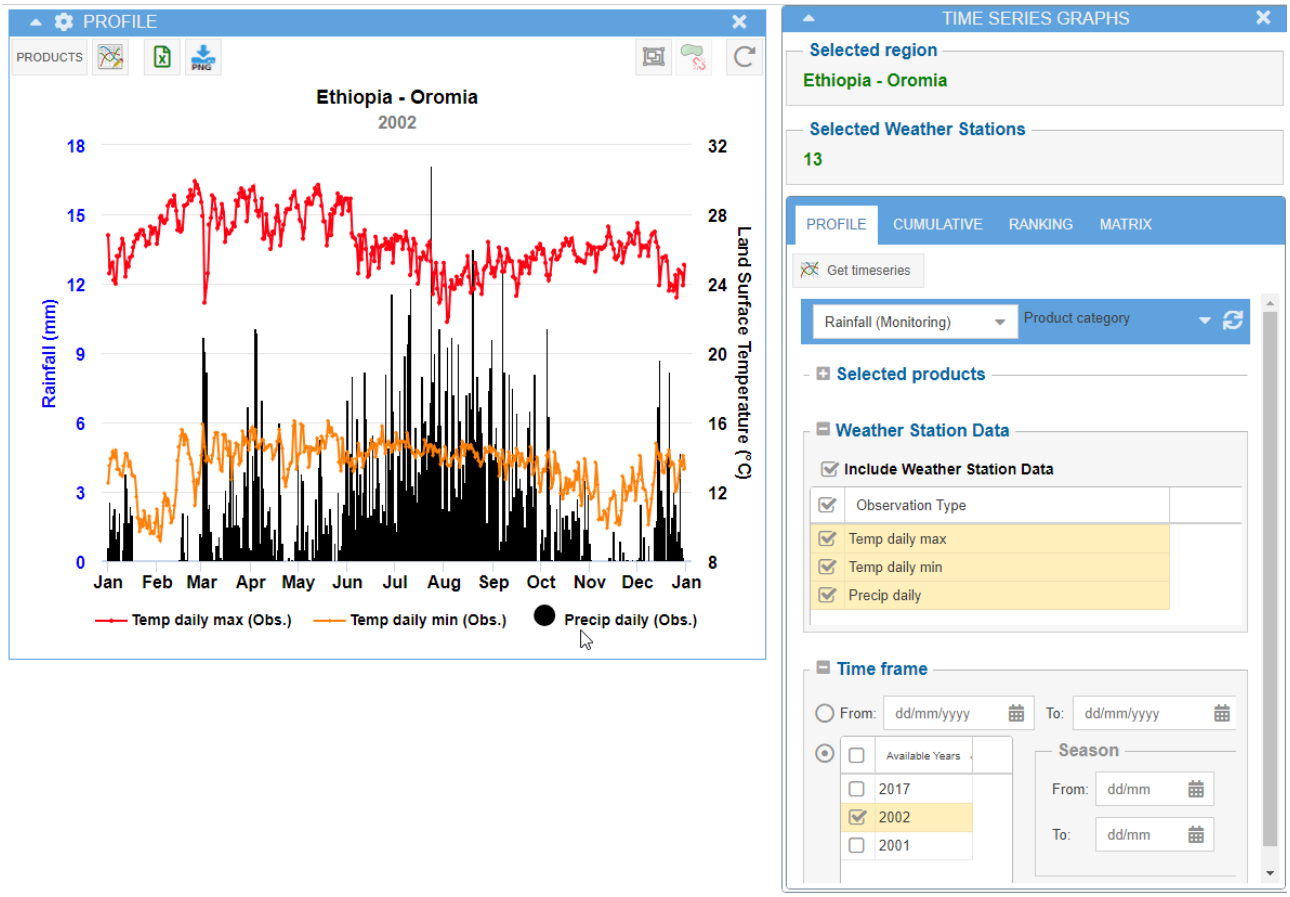
Under the “Selected products” section the **“Weather Station Data”** section is added, where you will see the list of observation types of the selected stations, for which data is present.

To generate a time series profile graph for the observation types of the selected stations, you have to first check “Include Weather Station Data”, and then select one or more observation types, as illustrated in below figure.

The available years of the selected observation types, are added to the years in the “Time frame” section.

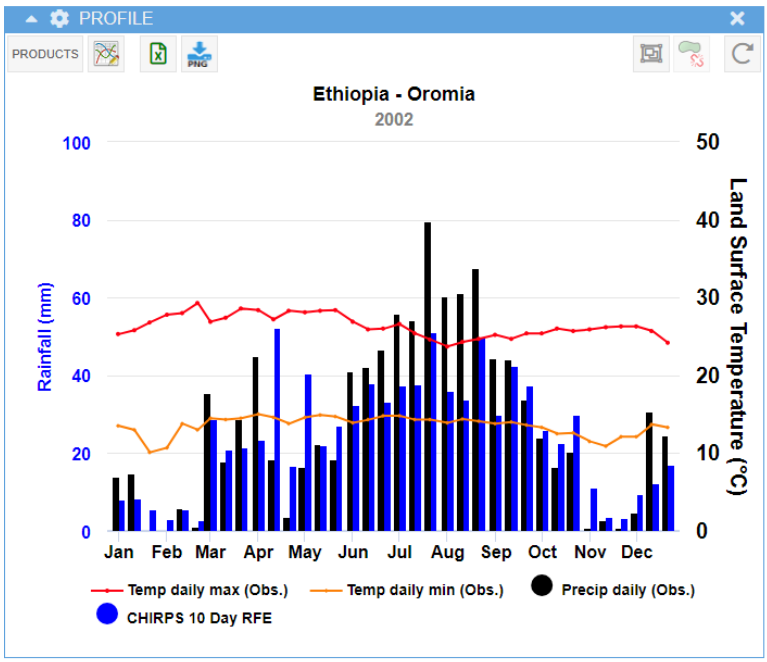


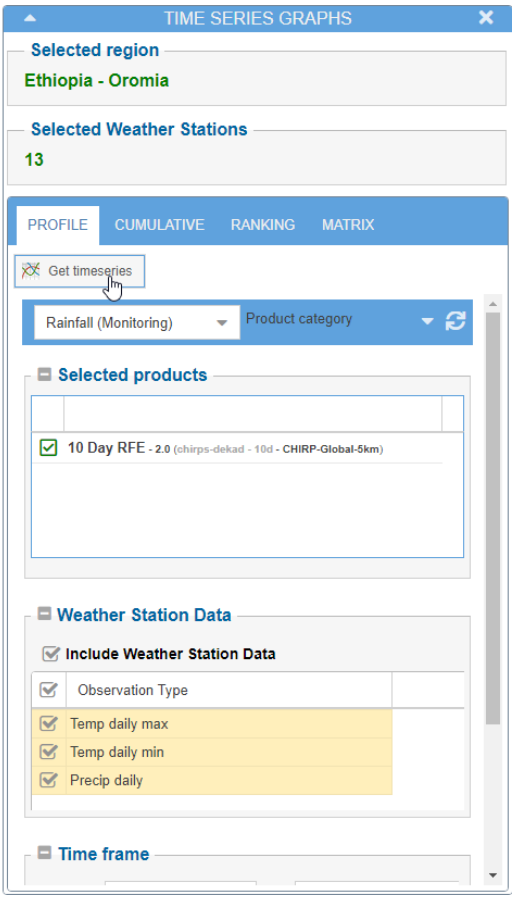
When you have checked “Include Weather Station Data”, selected one or more observation types and selected a time frame, you click on the “Get timeseries” button to generate the profile graph, as illustrated in below figure.



Weather Station time series data can also be combined with selected products time series.
The station data is automatically adjusted to the lowest time frequency of the selected products.

For example if you select a product with a 10 day time frequency the time series of the selected observation type with a daily frequency, will be will automatically accumulated to a 10 day frequency, as shown in below figure.





3.9 FITNESS FOR PURPOSES

This page incorporates into the Station part of the features available in the Copernicus Climate Change Service (C3S) Fitness for Purpose (F4P) Benchmark Platform, developed by JRC as part of its activities in support of C3S[7]. Only some features of the original toolbox have been, integrated, since some others were already available under the Analysis page, and the remaining deemed as less relevant for the Station's Users.

Figure 26: Overview of the F4P page

3.9.1 Description of the available F4P methods

The F4P methods that are implemented in the Station are briefly described below:

- **Hovmöller Diagram**: is a commonly used way of plotting a large amount of climatic data in a meaningful and understandable form (Hovmöller, 1949). The axes of a Hovmöller diagram are typically longitude or latitude (ordinate or y-axis) and time (abscissa or x-axis) with the value of some field represented through color or shading. In general this diagram is great for displaying large amounts of data in a meaningful and understandable form.

- **Scatter Plot:** it is also named as Scatter Density Diagram, and it allows to highlight any potential bias tendency within data sets by means of a bi-dimensional density histogram comparison. This functions expects that only two datasets are selected (the selection is performed acting on the include checkbox, section 4). The computation can be also queried for temporal series, in this case a canvas is produced for every pair of dataset.

[7] See for reference the dedicated, JRC Technical Report 114432-2019, available online at https://fapar.jrc.ec.europa.eu/_www/pubs.php#pubs_papers.

- **Trend Analysis:** this function takes as input the selected time-series of the selected product and computes the trend over the selected time period. The temporal extension is functional to the desired study and should be selected carefully. The presence of monotonic upward or downward is assessed using the non-parametric seasonal Kendall test.

- **Histogram and CDF**: This function allows to display both histograms and cumulative distribution functions CDFs of the selected samples. The figures are produced at monthly step, i.e. all the selected datasets which belong to the selected months are displayed in the same canvas. In the case that a _reference_ dataset is indicated, the GUI performs and shows the Kolmogoros-Smirnov statistics is calculated with respect to it (Stephens, 1974). This latter is a nonparametric test of the equality of continuous, one-dimensional probability distributions that can be used to compare two samples. The null hypothesis (H0) of this test is that the two samples are derived from the same common distribution. The test provides two indicators, the *KS-Statistic* and the corresponding *P-Value*. By definition, if the *KS-Statistic* term is small, meaning that the maximum distance of two CDFs is close to zero and the *P-Value* is higher than the significance level (assuming equal to 0.05), the evidence in support of H0 is strong.

- **Gamma Index:** it is a quality estimator of the consistency between two distributions, in which one is taken as a _reference_ and the other is the distribution to compare (to test) against. The selection of the reference can be performed explicitly via the dedicated switch or implicitly, in this latter case the reference distribution is calculated as the average distribution considering all the involved datasets, and the gamma index would express the consistency of the two distribution with respect their mean.

3.9.2 Running the F4P methods

The UI interface to access the F4P features on the Station is displayed in Figure 27. After selecting the type of analysis to perform, among the 5 available, the boxes for selection of the Product, Time and ROI are updated.

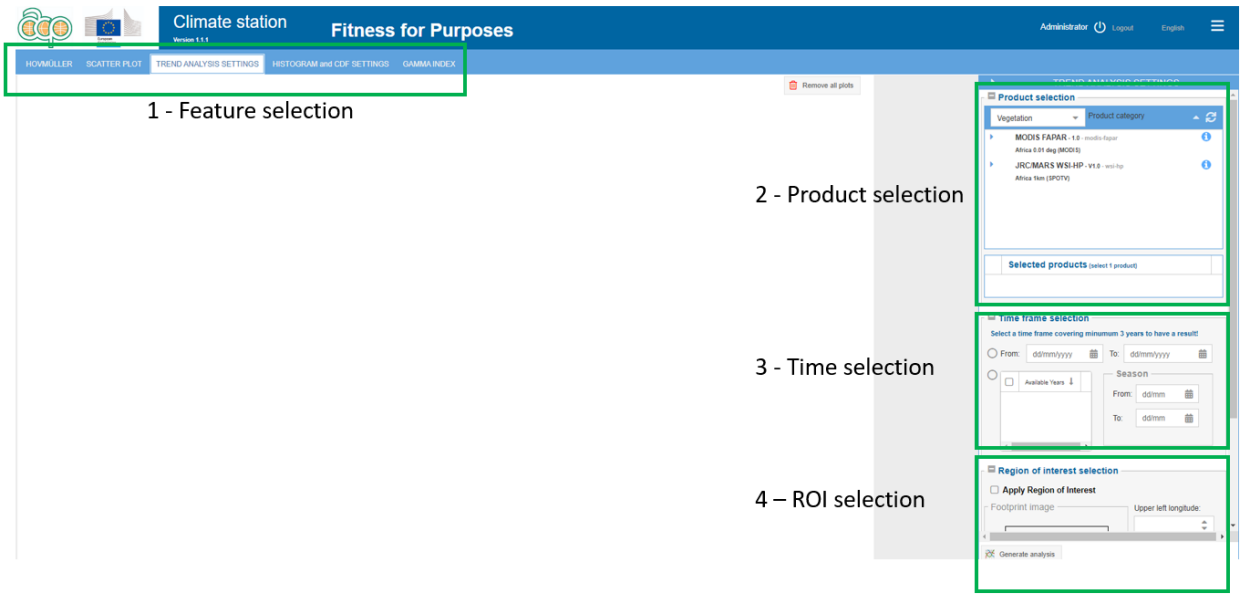


Figure 27: Procedure to create a F4P plot

Product Selection: the number of products that can be selected depends on the type of analysis, as displayed in the table below, in the second column. For the **Hovmöller** and **Trend Analysis**, a single product is processed at a time. For the Scatter Plot and the gamma index, two product are required and they need to have the same projection[8]. For Histogram and CDF two products can be selected, without any constraints on their mapset[9]

Type of Analysis	N. of Products	Constraints on Product Selection	Time Selection Options	Output Plots
Hovmöller	**1**	-	From-To Period 1+ Years (with Seasons)	Single Plot
Scatter Plot	**2**	The two products need to have the same grid (mapset)	Single date From-To Period 1+ Years (with Seasons)	1 Plot per Time
Trend Analysis	**1**	-	From-To Period 1+ Years (with Seasons))	Single Plot
Histogram and CDF	**2**	None	Single date From-To Period 1+ Years (with Seasons)	1 Plot per Time
Gamma Index	**2**	The two products need to have the same grid (mapset)	Single date From-To Period 1+ Years (with Seasons)	1 Plot per Time

[8] Re-projection on the fly in not implemented in the current release of the station. Can be done in the Jupyter Notebooks environment.

[9] Creation of the Histogram and CDF does not require pixel-to-pixel comparison.

Time Selection (and outputs): the options of selecting the time frame, and the number of output plots created, also depends on the type of analysis. This is reported in columns 4 and 5 of the table. For **Hovmöller** and **Trend Analysis,** the selected time period leads to the creation of a single graph. The time frame can be a single interval defined through a start and end-date, or a number of full years, or even a season defined over multiple years.
For **Scatter Plot, Histogram and CDF,** and **Gamma Index,** a plot is created for each time selected. The selection of the times includes therefore the single data option as well.

ROI Selection: the F4P algorithms operates by default on the full extension of the datasets. The User can restrict the analysis to a sub-region (or region of interest) modifying the values presented in the corresponding box, and checking the 'Apply Region of Interest' box.

3.10 IMPACT TOOLBOX

High resolution Earth Observation imagery has been identified as a key source of information for mapping and monitoring land cover changes or degradation over time. However, data extraction, layer stacking, radiometric calibration, normalization, mosaicking, automatic classification and segmentation are only some of the pre-processing steps that a user has to undertake to obtain a basic land cover map.

IMPACT Toolbox offers a combination of remote sensing, photo interpretation and processing technologies in an easy to use WEB-GIS environment, allowing non specialist users to easily accomplish all necessary pre-processing steps while giving a fast and user-friendly environment for visual editing and map validation.

The rational for this development has been the strong link between the JRC projects to monitoring tropical forests (TREES, FOROBS, ReCaREDD, and REDD-Copernicus) in collaboration with the forest national services.

Originally developed as a portable Windows application, IMPACT has been recently integrated into the Station (in a dedicated Docker container) offering multi-OS support and direct access to datasets automatically retrieved by the latter. Figure 28 represent the link between the two systems, with IMPACT working on the data locally retrieved by the station.

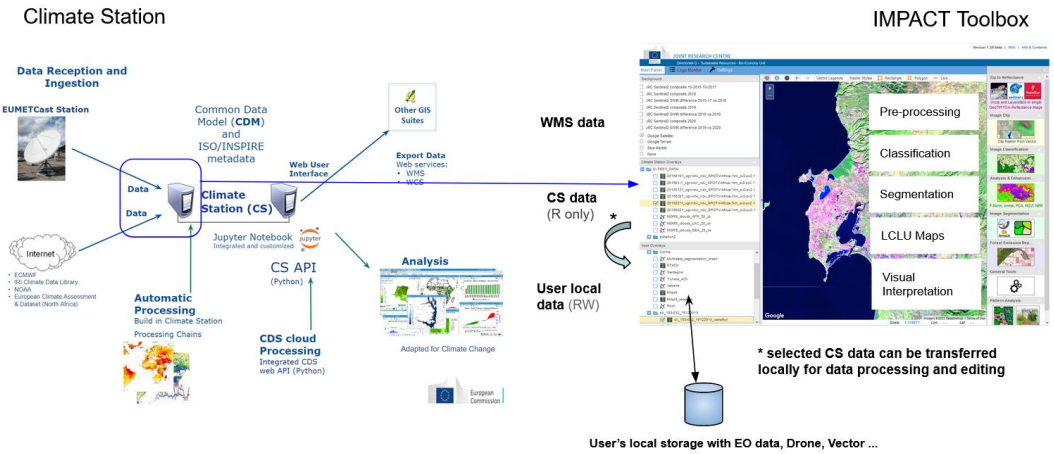
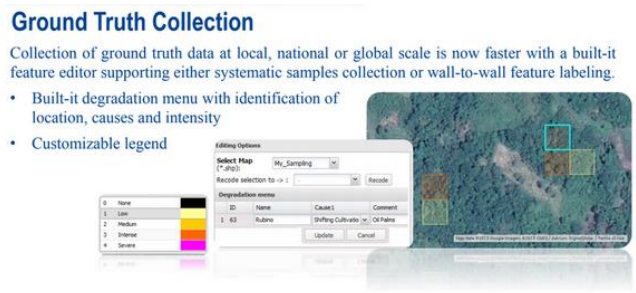


Figure 28 : view of the logical link between the Station and IMPACT

The IMPACT web-interface mimic the classic and well known GIS software offering a central panel containing background maps and user overlays, a left panel with access to local and remote (e.g. the Station) raster and vector layers and a right panel with a set of processing modules spanning from automatic ingestion of raw (i.e. zipped) Landsat 4-5-7-8-9, Sentinel-2 and Rapid Eye imagery to image classification, segmentation, zonal statistics, morphological patter analysis, forest degradation detection and carbon emission reporting. Is worth mentioning advanced tools for raster and vector editing and label assignment, key steps in the land cover/land use change mapping activities as well as in ground truth collection activities.

Figure 29: Map Visualization & Editing features

Figure 29 and Figure 30 represent same of the features available in IMPACT for the visualization and editing features of a map and the editing of ground-truth collected data, respectively.



More information, user manual and test cases are available at:

https://forobs.jrc.ec.europa.eu/products/software/impact.php

3.11 JUPYTER NOTEBOOKS

The main objective of the Jupyter notebooks in the Station is to give the flexibility for the user to exploit its power to customize the visualization, derive indicators of his/her choice using the available libraries, create the chains of processing to derived multiple variables and also more customized way of using IMPACT toolbox, Analysis tool, Fitness for purpose tool.

The usage diagram is shown in Figure 30, which highlights the components of the Station can be accessed through the Notebooks: as you can see, they are not suggested for Acquisition, which we propose to exploit only through the Station services. On top of scientific work, one can use the Jupyter notebooks to create reports and bulletins with texts and graphs, which can be shared with others for visualization.

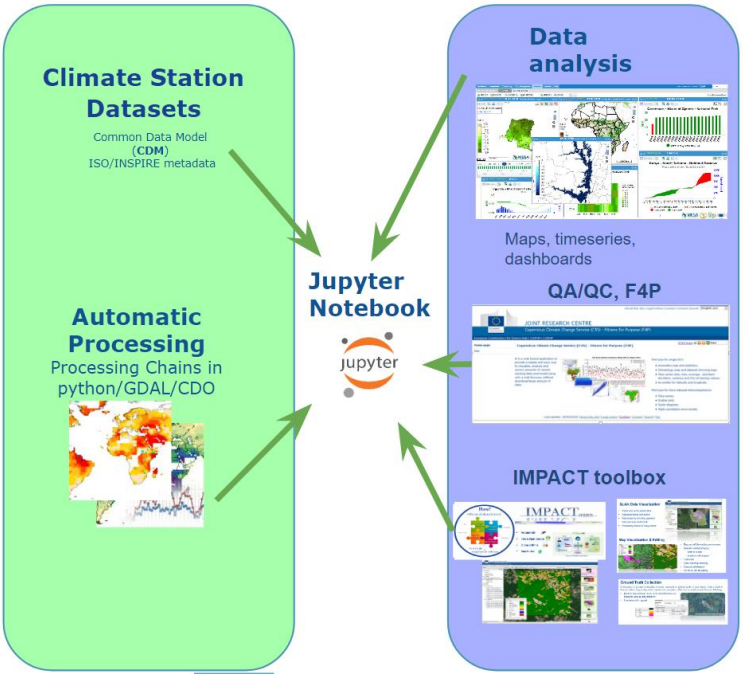


Figure 30: Jupyter Notebook usage in the Station

In the notebook you can make use of Python 3 kernel, Python 3 console, command line terminal, text file, markdown file and python file as a notebook. We also provide pre-cooked notebooks which can be used by the user.

In this chapter you will learn about four main topics

- Basics of Jupyter notebook
- Collections & Datasets in notebook
- Analysis in notebook
- Processing in notebook

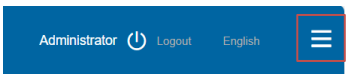
3.11.1 Basics of Jupyternotebook

A notebook integrates code and its output into a single document that combines visualizations, narrative text, mathematical equations, and other rich media. In other words: it's a single document where you can run code, display the output, and also add explanations, formulas, charts, and make your work more transparent, understandable, repeatable, and shareable.

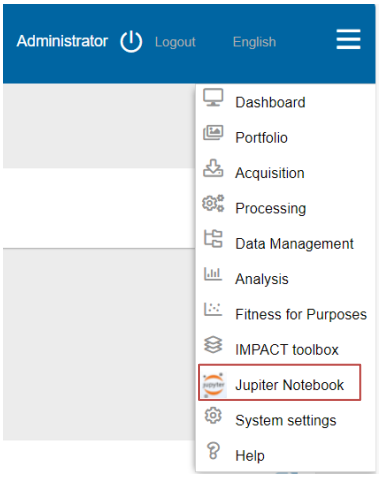
Using Notebooks is now a major part of the data science workflow at companies across the globe. If your goal is to work with data, using a Notebook will speed up your workflow and make it easier to communicate and share your results.

The procedure to open Jupyter notebook in the station

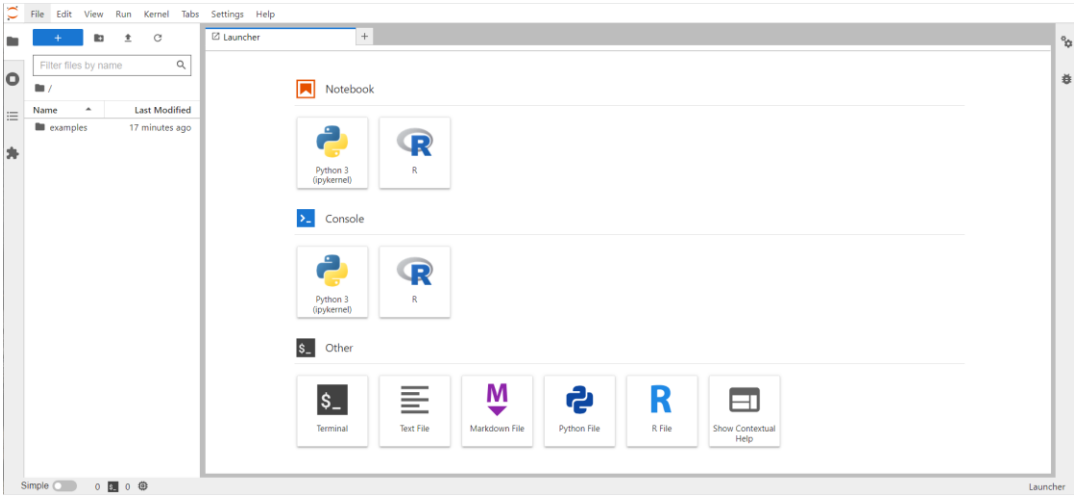
1. First open the Station GUI.
2. Login to the station using your personal credentials.
3. Once you login, you can click on the menu item from the top right corner.



4. Then click on the “Jupyter Notebook” option.



5. A new Jupyter notebook will be created for your username and open in a new browser as shown below.



An Overview of the most important features that you can find over this dashboard is given in the following illustration.

Menu Bar

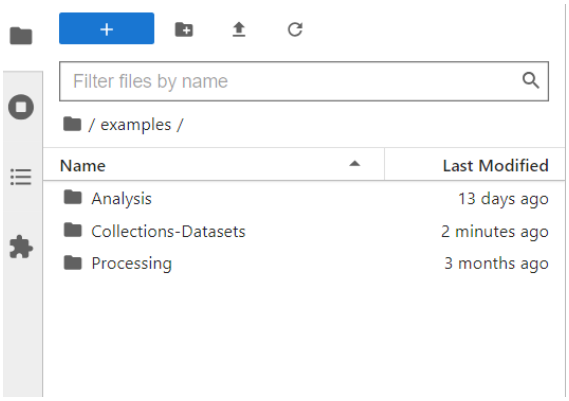
The menu bar at the top of JupyterLab has top-level menus that expose actions available in JupyterLab with their keyboard shortcuts. The default menus are:

- **File**: actions related to files and directories
- **Edit**: actions related to editing documents and other activities
- **View**: actions that alter the appearance of JupyterLab
- **Run**: actions for running code in different activities such as notebooks and code consoles
- **Kernel**: actions for managing kernels, which are separate processes for running code
- **Tabs**: a list of the open documents and activities in the dock panel
- **Settings**: common settings and an advanced settings editor
- **Help**: a list of JupyterLab and kernel help links

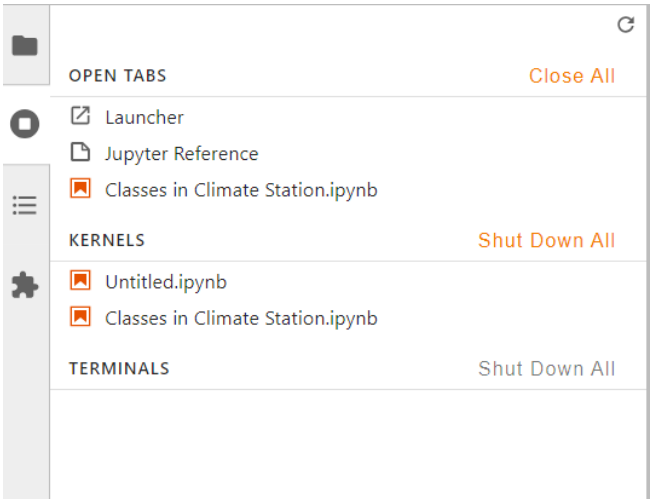
File Explorer & Interface Menu

The left sidebar contains a number of commonly-used tabs including:

- File Explorer: Where the working files and folders are present (In the figure below you will see the examples folder where the list of template ipynb files are structured in 3 different folders)



- a list of tabs in the main work and of running kernels and terminals,

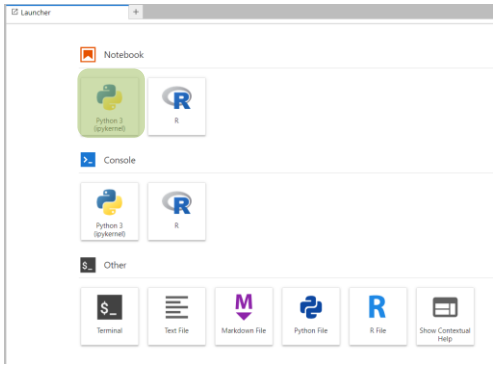


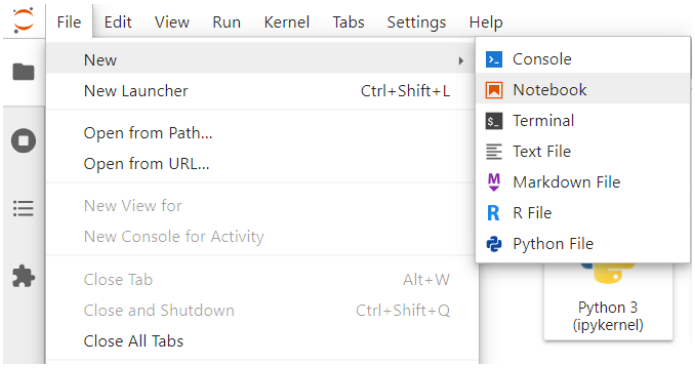
- the command palette (in 3.0+ moved to a modal window accessible with a keyboard shortcut),
- the table of contents,
- the extension manager.

Main Work Window

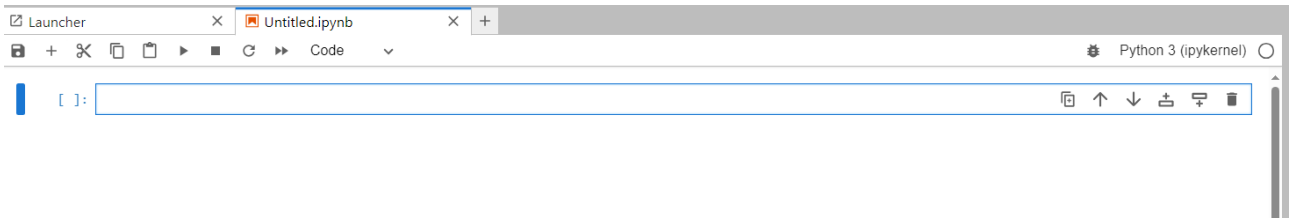
The main work area in JupyterLab enables you to arrange documents (notebooks, text files, etc.) and other activities (terminals, code consoles, etc.) into panels of tabs that can be resized or subdivided. Drag a tab to the center of a tab panel to move the tab to the panel. Subdivide a tab panel by dragging a tab to the left, right, top, or bottom of the panel:

From the launcher → Notebook, create Python3 (ipykernel) by clicking on it or by click on File → New → Notebook



or


A new Untitled.ipynb is created as shown below. Its interface will hopefully not look entirely alien. After all, Jupyter is essentially just an advanced word processor.



In addition, you find the fast feature access bar, integrated in your new created notebook, inside the main working window:
A summary of the meanings of the features is given below.

Toolbar Symbol	Meaning
	Save current notebook file
+	Add a new cell to the current notebook
	Cut one or more selected cells
	Copy one or more selected cells
	Paste cells from clipboard into the current notebook file
	Execute one or more selected cells
	Stop execution
	Restart kernel
	Restart kernel and run all cells of the current notebook file
Code	Change cell type

There are two fairly prominent terms that you should notice, which are probably new to you: cells and kernels are key both to understanding Jupyter and to what makes it more than just a word processor.

Fortunately, these concepts are not difficult to understand.

- A kernel is a “computational engine” that executes the code contained in a notebook document.
- A cell is a container for text to be displayed in the notebook or code to be executed by the notebook’s kernel.

Cells

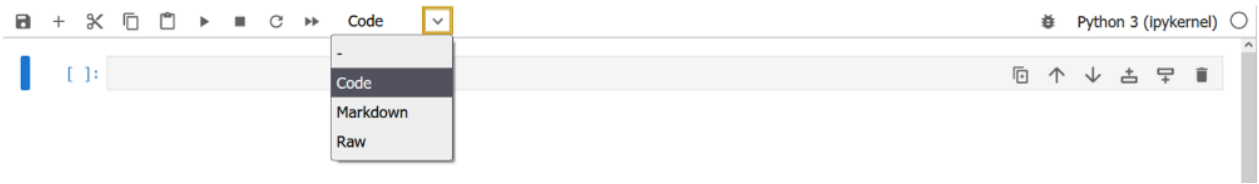
Cells form the body of a notebook. As mentioned above you can choose between different cell types in Jupyter Notebooks according to your content type. The different cell types are:

- Code cells: These cell types allow you to write and run code snippets interactively. They were executed asynchronously. Every new cell in a notebook file is created as code cell by default.
- Markdown cells: These cells are used to add narrative text of any kind to your notebook file. As stated before you can also add equations, images, videos or other contents to these cells. The text formatting and content bindings are interpreted over the markdown markup language. An Overview of the options that exist inside these cells is given in Jupyters online documentation about markdown cells.
- Raw cells: These cells are not interpreted in any way. You could use these for instance to share code snippets that should not be executable or in general for any information that should not be interpreted and rendered by the computer. Hence the cell content will be display originally as you wrote it in this cell.

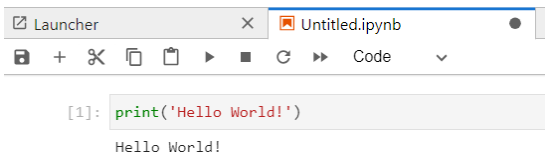
The first cell in a new notebook is always a code cell.

If you click on a cell inside of a Jupyter Notebook the notebooks toolbar will show you the specific cell type. For example, the cell in the following screenshot is a code cell:

You can change the type of one or more selected cells by clicking on the little arrow besides the cell type. This will open a context menu which lists the cell types described above:



Let's test it out with a classic hello world example: Type *print('Hello World!')* into the cell and click the run button Notebook Run Button in the toolbar above or press Ctrl + Enter.



The result should look like this:

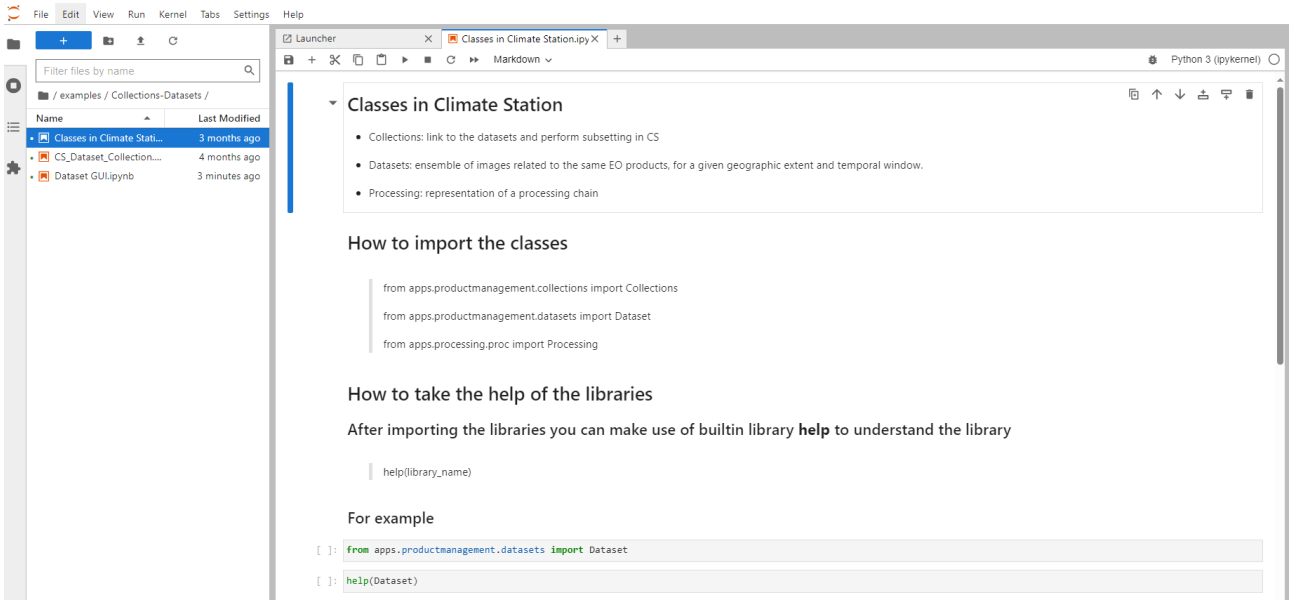
When we run the cell, its output is displayed below and the label to its left will have changed from In [ ] to In [1].

3.11.2 Collections, Datasets in Notebook

From the portfolio you would have activated the list of products you are interested in. In order to work with these products it is necessary to learn about some of the libraries like Collections & Datasets. So lets open the pre-cooked notebook in order to understand how to work with some of the internal libraries.

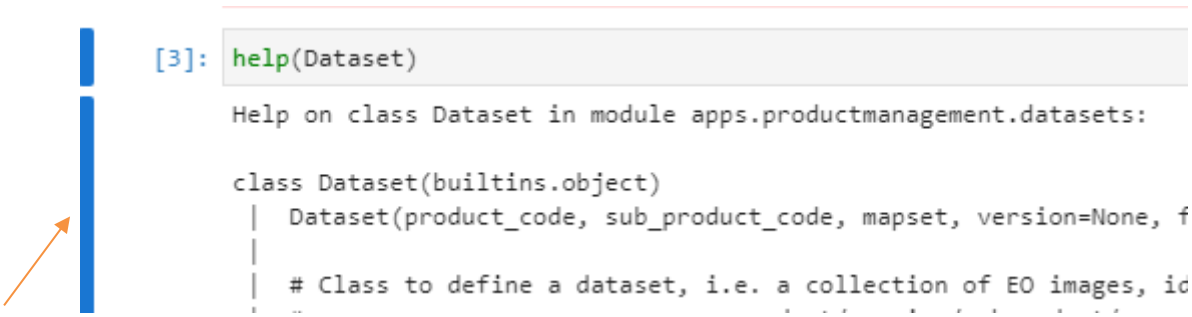
Classes_in_Climate_Station.ipynb

From example → Collections-Datasets folder, open the Classes_in_Climate_Station.ipynb notebook by double clicking on it.

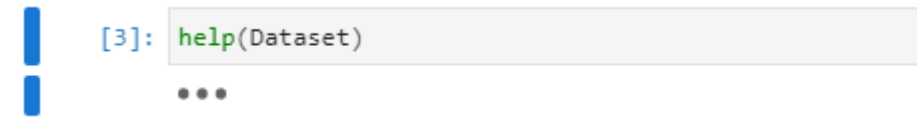


In the initial part of the notebook, the first few lines are written as simple text. There is Collections, Datasets and Processing classes which can be used and how it can be imported. In order to take help of these libraries we can make use of builtin *help* function of python.

Lets execute the line `[ ]: from apps.productmanagement.datasets import Dataset` and take the help `[2]: help(Dataset)` to understand how it can initialized and list of functions, methods present within it. There should be long list of functions and methods. In order shrink the results, click the mouse over result and on the left side you will notice blue vertical bar. Click on the blue vertical bar in order to close the result.

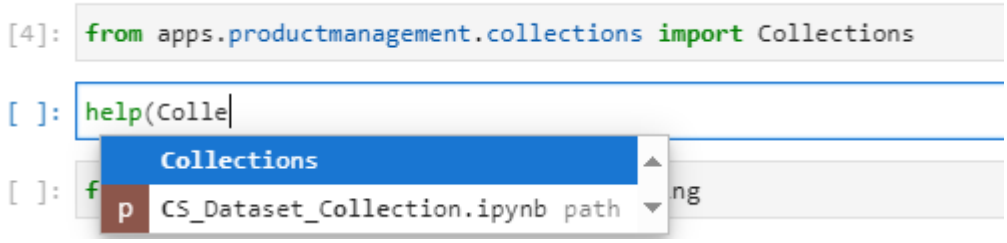


So the results are closed like this



Similarly you can write a line by yourself to import Collections class and execute it. You will notice that there are so many method and functions.

- Click on + button to create a new cell below and copy “from apps.productmanagement.collections import Collections” from the above text and click on run button ▶
- Add a new line and type “help(Colle” and press the tab you will automatically get the suggestion to fill the Collections since the library is imported



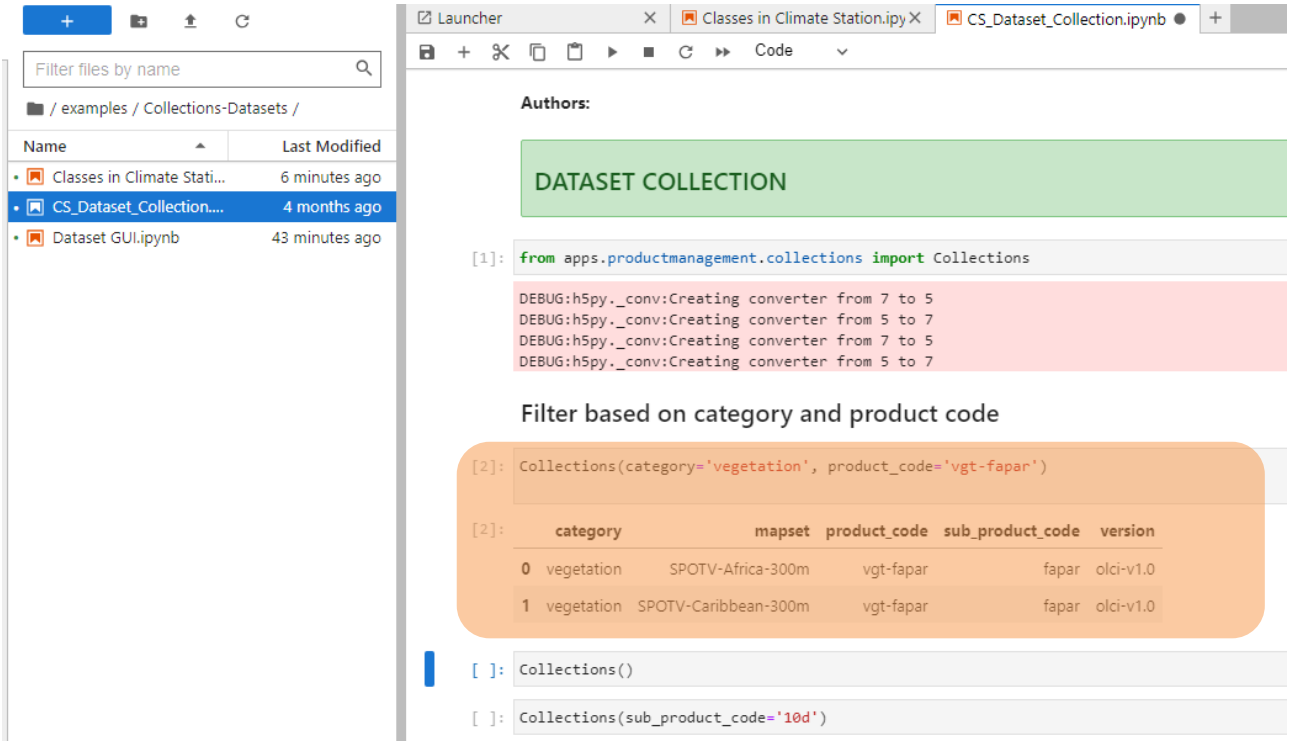

```
[5]: help(Collections)
Help on function Collections in module apps.productmanagement.collections:

Collections(category=None, product_code=None, version=None, mapset=None, sub_product_code=None, mobile_app=False)
    # ########################################################################################################################
    # Collections --> list the datasets and perform subsetting in CS #
    # Arguments:
    #   category[optional]: Filter by category passed (Can be used independently or along with other arguments)
    #   product_code[optional]: Filter by Product code (Can be used independently or along with other arguments)
    #   sub_product_code[optional]: Filter by Subproduct Code (Can be used independently or along with other arguments)
    #   version[optional]: Filter by version (Can be used independently or along with other arguments)
    #   mapset[optional]: Filter by mapset (Can be used independently or along with other arguments)
    # ########################################################################################################################
```

By now you should have got familiarized with the usage of the notebook. Let open another notebook to learn about Collections library.

CS_Dataset_Collection.ipynb

From example → Collections-Datasets folder, open the CS_Dataset_Collection.ipynb notebook by double clicking on it and execute some lines as shown below. From the help of Collections library you should have understood it is the library to list all the datasets which are activated in the system. Also you will the find the list of argument which you can pass to filter or subset the datasets.



When you execute the line 2, we are sub setting the datasets by passing the category='vegetation' and product_code='vgt-fapar'. In the above picture we got two result but if you haven't activated the FAPAR product it returns the empty list.

The line `[ ]: Collections()` lists all the datasets which are activated in your system without sub setting them. In our case we have 391 rows.

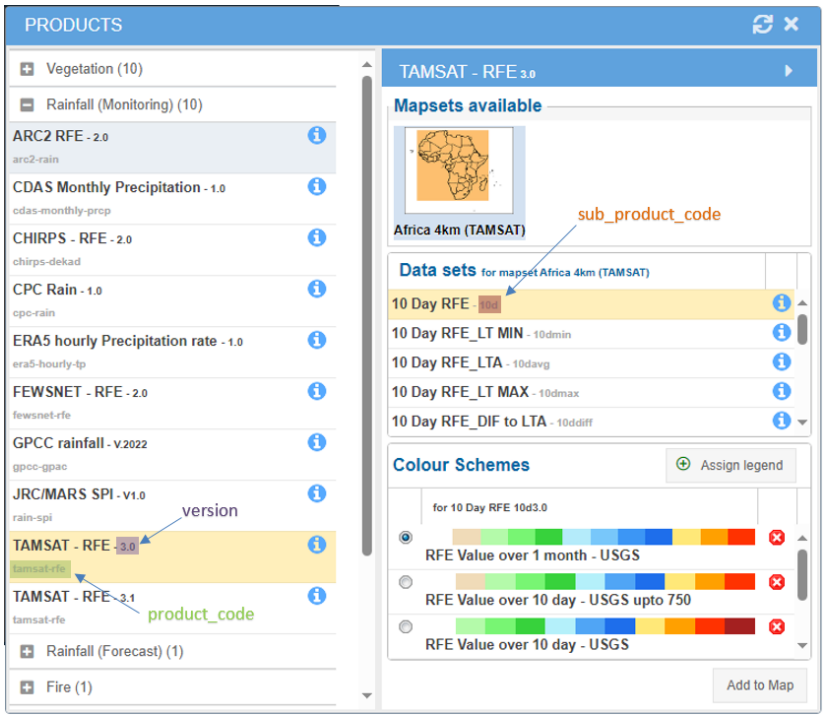
```
[3]: Collections()
```

[3]:

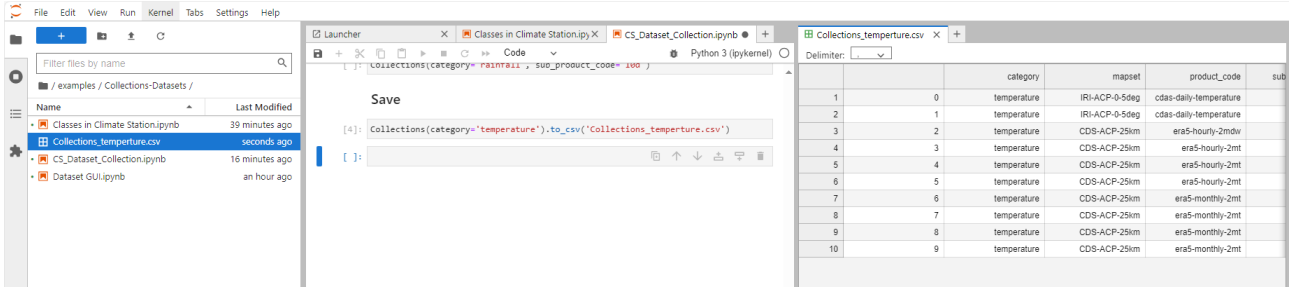
	category	mapset	product_code	sub_product_code	version
0	vegetation	GPCC-Global-1deg	gdo-rdri	rdria	V2.3.2
1	vegetation	SPOTV-Africa-1km	modis-ndvi	ndvi	1.0
2	vegetation	SPOTV-Africa-1km	modis-ndvi	zndvi	1.0
3	vegetation	SPOTV-Africa-1km	modis-ndvi	ndvid	1.0
4	vegetation	MARS-Global-1km	modis-ndvi	ndvi	1.0
...	...	...	...	...	...
386	atmosphere	ADS-ACP-75km	cams-pm10	daily	1.0
387	atmosphere	ADS-ACP-75km	cams-pm10	subdaily	1.0
388	atmosphere	ADS-ACP-75km	cams-pm2p5	monthly	1.0
389	atmosphere	ADS-ACP-75km	cams-pm2p5	daily	1.0
390	atmosphere	ADS-ACP-75km	cams-pm2p5	subdaily	1.0

391 rows × 5 columns

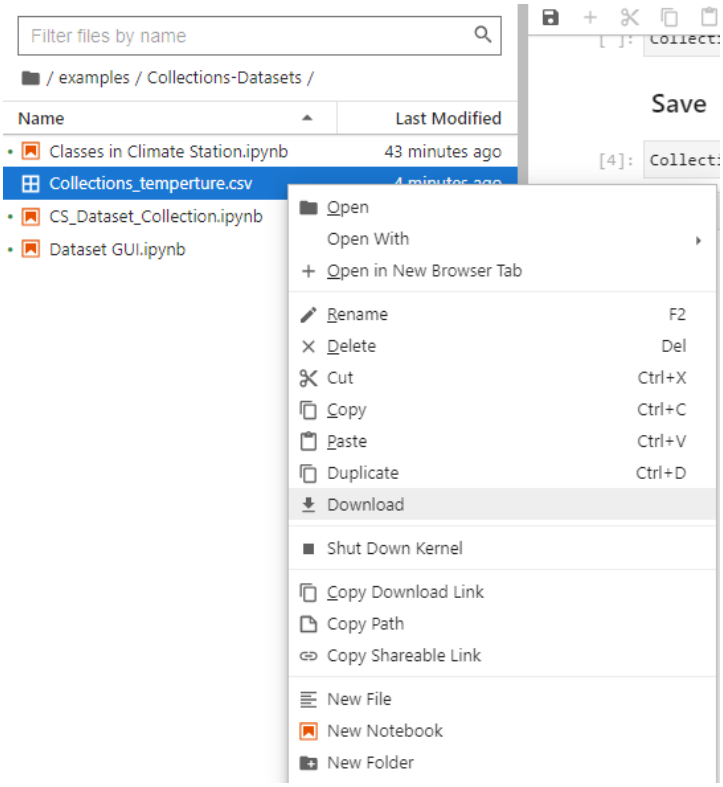
These listed information of the datasets(product_code, sub_ product_code, mapset, version) are very important in order to work with other template notebooks. You can also find these information from the Analysis Tool→ Product Navigator



You can also save this collections list as the csv using `Collections(category='temperature').to_csv('Collections_temperture.csv')`. Once you execute those command, it saves the files in your file explorer. You can open it directly in notebook by double clicking on it.



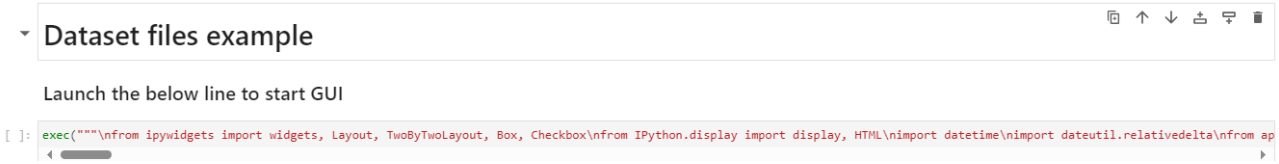
Once you are happy with the result you can also save it in your local file system for reference by right click on the file and selecting download.



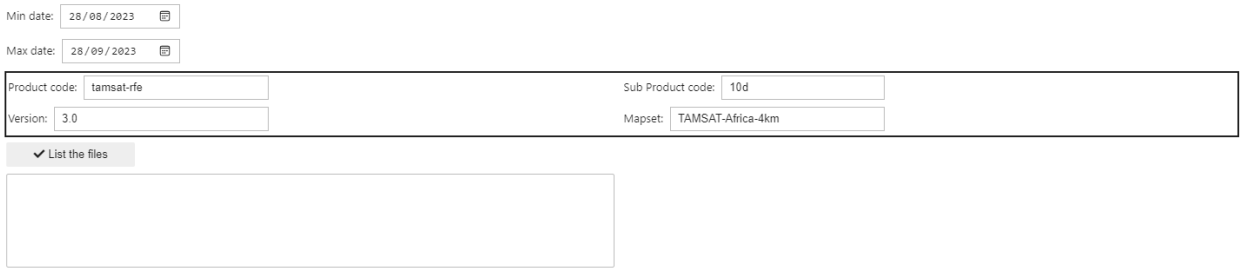
Dataset_GUI.ipynb

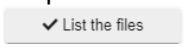
Now let's open another notebook in order to practise how to feed these information to retrieve the list of files.

From example → Collections-Datasets folder, open the Dataset_GUI.ipynb notebook. In this notebook you will see all the codes are compressed in a single line. Run the single line so that it loads a GUI.

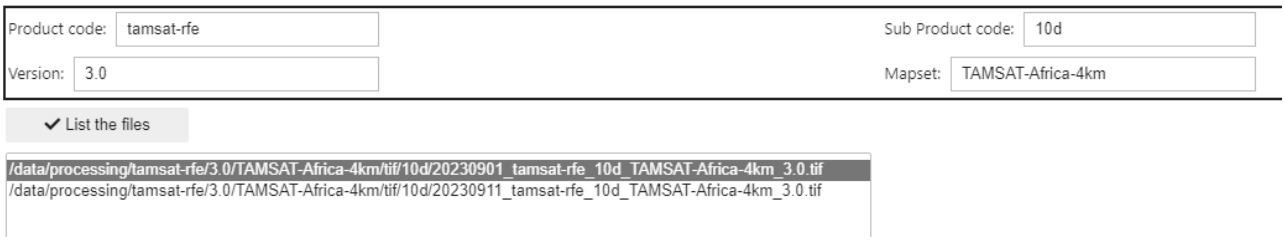


Once you execute the above line you will see the GUI like this

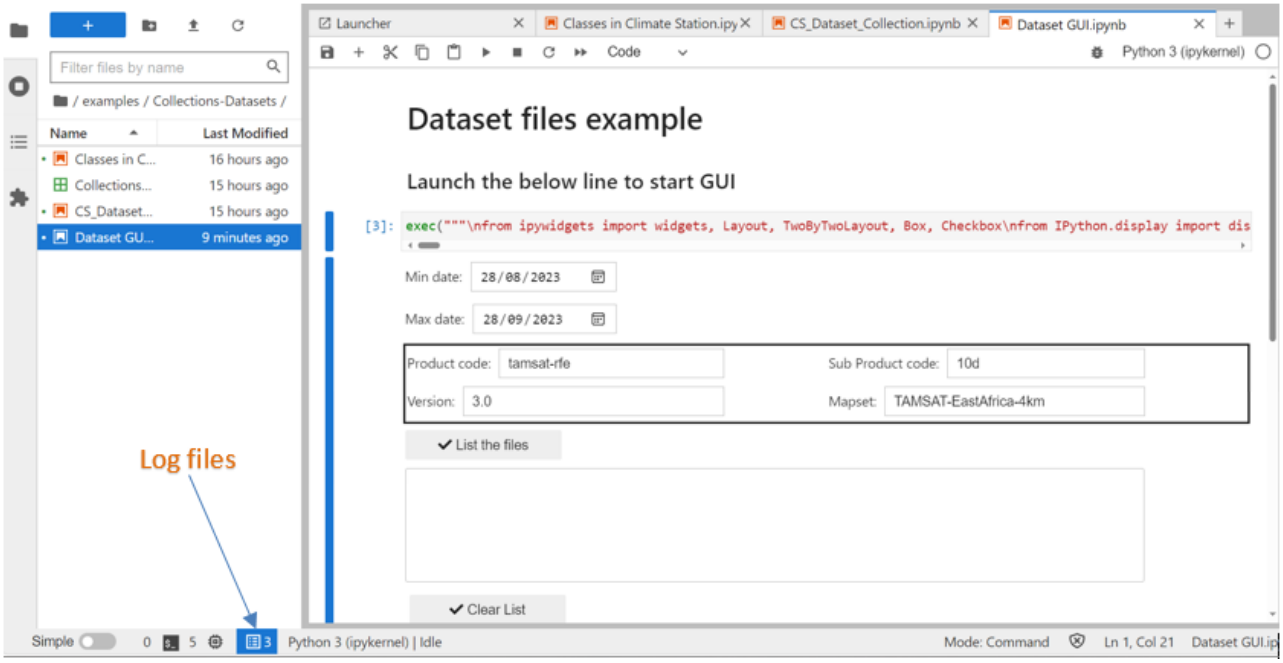


You will notice you can edit the start date & end date as well as 4 parameters Product code, Sub product code, Version, Mapset in order to list the files within your system. By default we provided some parameters but you have to check you Collections result to fill these values. Once filled correctly click on 

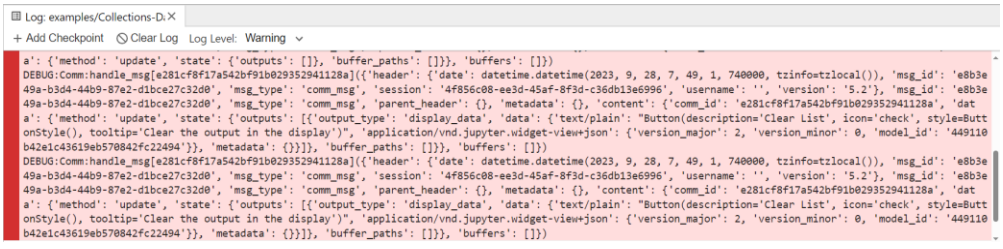
to see list of files available in your system.



In case if it is empty, you check the bottom left corner of the notebook where the log files are present.



Click on the blue icon described in the above picture to open the log files as shown below



You can change the log level in order to visualize the error.

Important Note: Currently you are working on the template notebooks within the example folder, so any changes or modification made on this folder will be reverted if the system in updated. So it is necessary to copy the notebooks outside the example folder so that all the changes are preserved. Each user will have their own personal workspace outside the example folder where can drag and drop any files to work in the system.

3.11.3 Analysis in Notebook

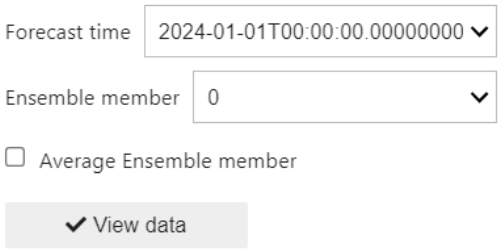
In the section 3.8 you would have learnt about the Analysis Tool and its usage. Those predefined set of Analysis which cannot be alerted by the users. Whereas in this section we will discuss about the Analysis which can be customized according to the user needs.

MultiDimensional_GUI.ipynb

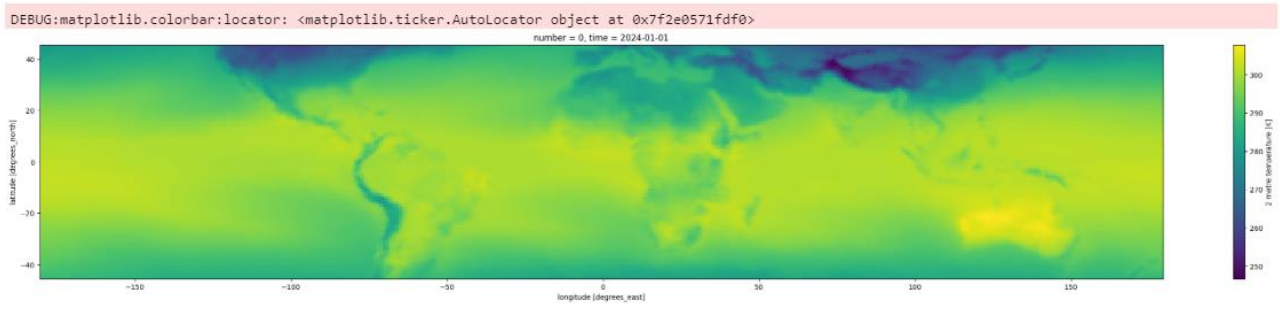
You will find this notebook in example/Analysis folder. This notebook is the tool to visualize the seasonal and subdaily forecast data. Once you open it you will see the codes are compressed in the single line, execute it in order to open the Graphical User Interface. Provide the seasonal forecast dataset parameters which are activated in your system in order to list the files

You will find the list of files available in your system. You will see the full path filenames where /data/processing/s51-monthly-2mt/1.0/CDS-ACP-1deg/tif/2mt/ is the folder and 20230901_s51-monthly-2mt_2mt_CDS-ACP-1deg_1.0.nc is the filename in the format <date>_<product_code>_ <sub_ product_code>_<mapset>_<version>.nc

Once you select the file automatically dropdown menu to select Forecast time and Ensemble member appears.

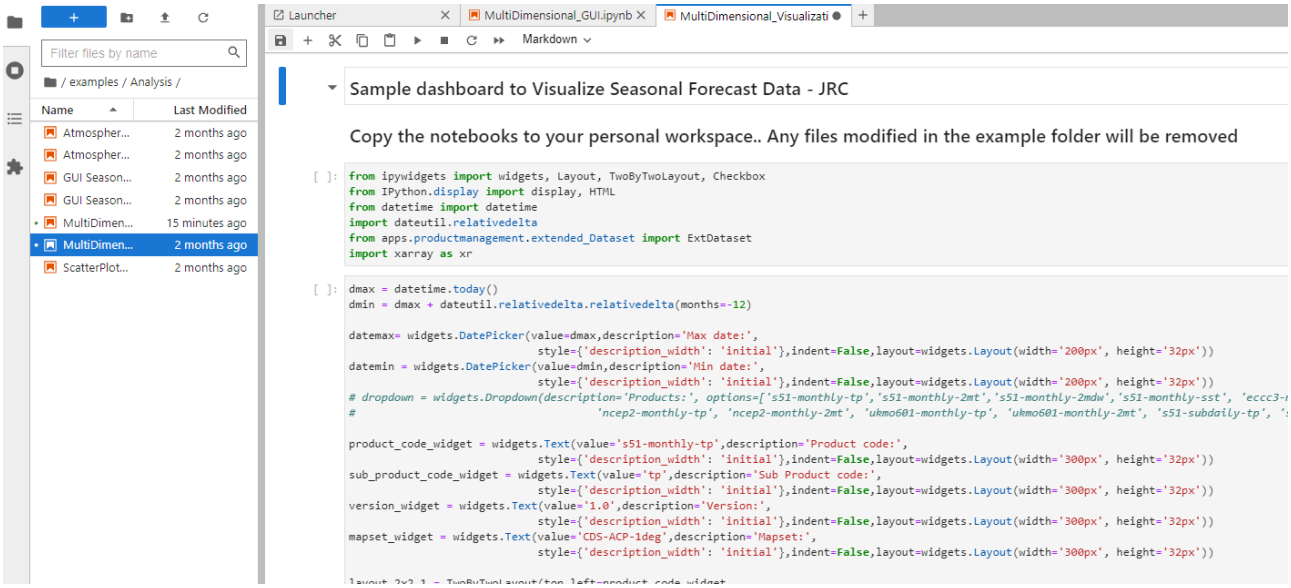


After selecting the options click on View data button to visualize the image



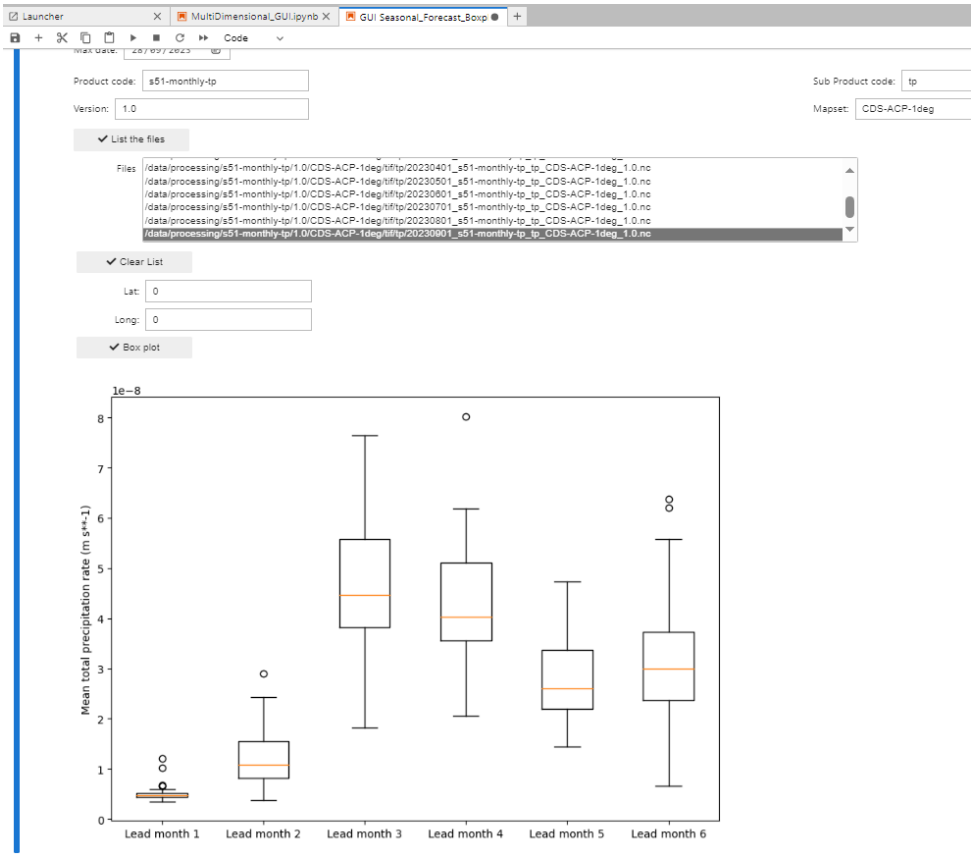
MultiDimensional_Visualization.ipynb

In case if you want to edit or customize the previous notebook you can make use of this uncompressed notebook with full codes.



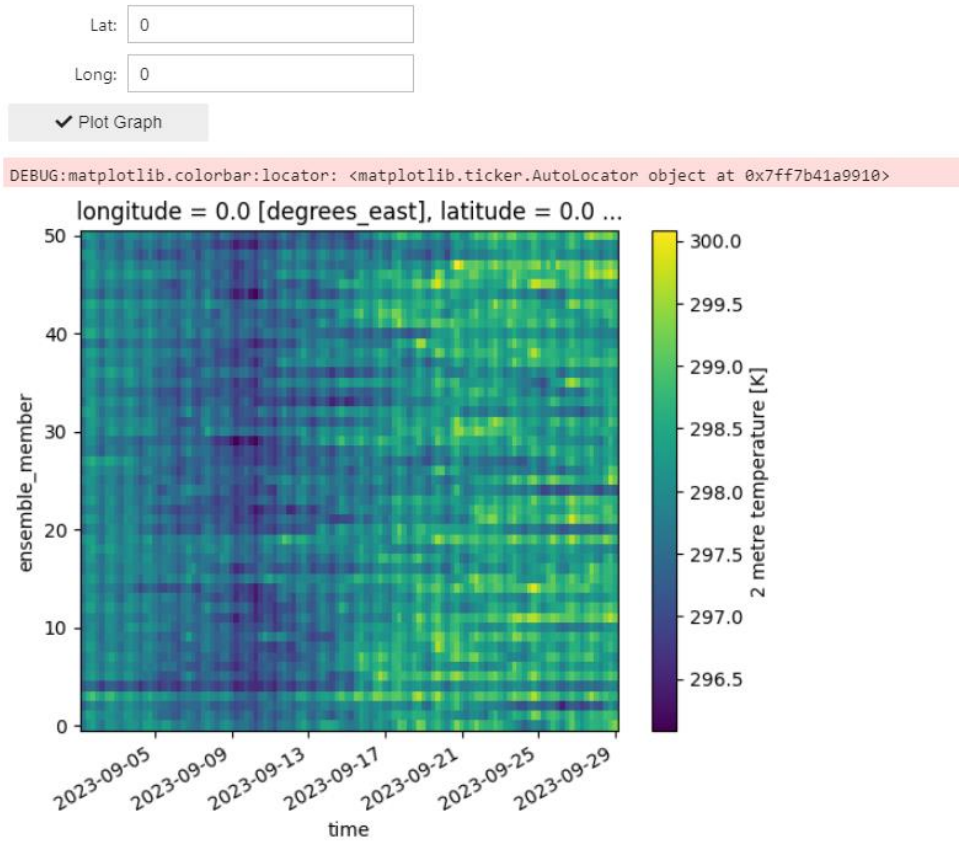
GUI Seasonal_Forecast_Boxplot_Graph.ipynb

The previous notebooks just visualize the seasonal and subdaily forecast products but in this notebook we will plot the box plot graph in order to visualize all forecast leadtimes in the x-axis and indicator value plotted as distribution of all ensemble member for particular latitude and longitude location.



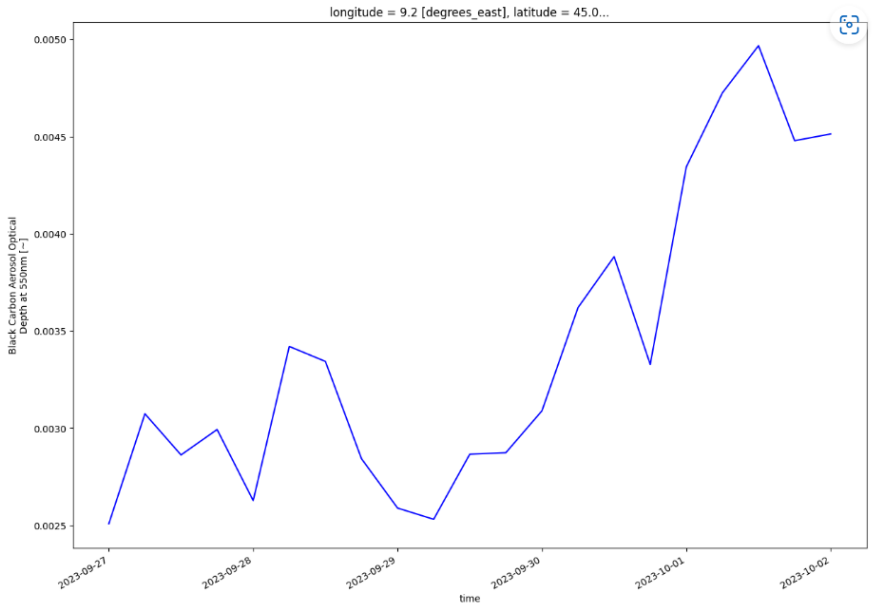
GUI Seasonal_Forecast_Mesh_Graph.ipynb

Similarly you can also plot the mesh graph for the seasonal and sub daily forecast product



Atmosphere composition forecasts graph.ipynb

If you want to work with Copernicus Atmospheric data then you can make use of this notebook to plot the forecast graph of any atmospheric indicator for particular latitude and longitude location.

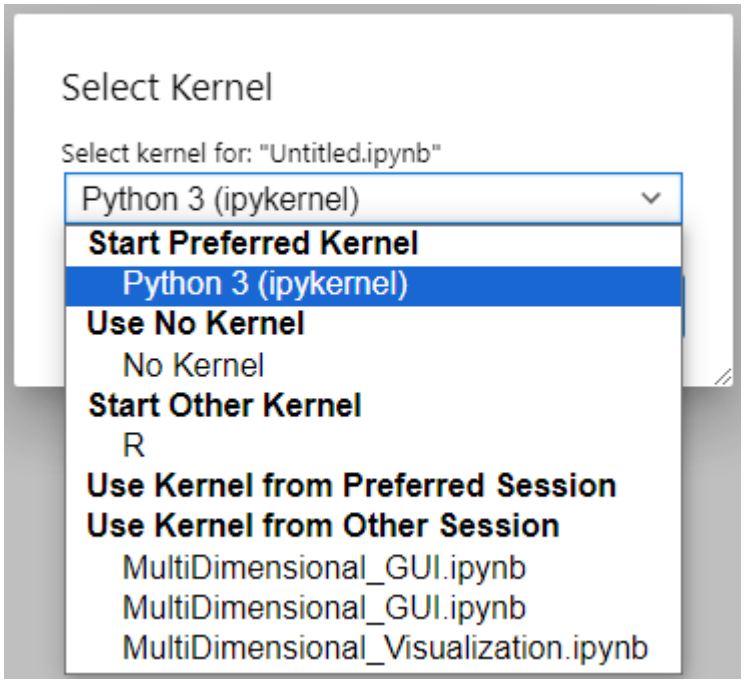


3.11.4 Processing in Notebook

In the section 3.7, you would have learnt about the predefined set of processing chains which are activated automatically when a product is activated from the portfolio. If you wanted to customize those chains or create your own processing chains then it is necessary you have to use the notebook to do so.

There are list of processing templates provided within examples/Processing folder. If you wanted to use any of those to customize, then it is necessary to copy those notebooks outside the examples folder and start modifying it.

In the section 3.11.2 we learnt how to use the Classes in the system using Classes_in_Climate_Station.ipynb notebook. We have a separate Class or library to do processing in the system. Open a new notebook (File→New→Notebook) in your workspace and select the kernel you wanted to work with.



Import the Processing class by copy pasting the line from Classes_in_Climate_Station.ipynb notebook.

```
[ ]: from apps.processing.proc import Processing
```

Once executed this line we can use the built-in help function to understand about the class

```
[2]: help(Processing)

Help on class Processing in module apps.processing.proc:

class Processing(Proc)
 |  Processing(dataset)
 |
 |  # ########################################################################################################################
 |  # Processing --> Processing Class, Contains several methods to derive indicators from the input products
 |  # Arguments:
 |  #   Dataset(class): The input dataset for which you wanted to derive indicators
 |  #       eg. Dataset(product_code='modis-sst', sub_product_code='sst-day', version='v2013.1', mapset='MODIS-Africa-4km')
 |  # ########################################################################################################################
 |
 |  Method resolution order:
 |      Processing
 |      Proc
 |      builtins.object
 |
 |  Methods defined here:
 |
 |  __init__(self, dataset)
 |      Initialize self.  See help(type(self)) for accurate signature.
 |
 |  ----------------------------------------------------------------------
```

In order to initiate the Processing it is necessary to pass the dataset object. Let have a look at an example from the processing example to understand better.

Processing - Computation of Long Term Average (LTA).ipynb

This is the notebook example to compute long term climatology average for any products available in the system. The first and foremost step is to import the classes and libraries in order to work with this computation.

```
[ ]: # Import dataset class and Proc class
     from apps.productmanagement.datasets import Dataset
     from apps.processing.proc import Processing
```

Here we imported two classes Dataset and Processing and now let's initialize the dataset object in order to pass to the processing class. As we know from the previous example, it is necessary to pass four mandatory parameters in order to initialize the dataset object. These parameters can be taken from the Collections result which is described in the section 3.11.2

```
[ ]: # Initializes the dataset object by passing the product code, subproduct code, version and mapset
     tamsat = Dataset(product_code='tamsat-rfe', sub_product_code='10d', version='3.1', mapset='TAMSAT-Africa-4km')
```

In your case you would have activated the same product over your region of interest (eg. East Africa) so check the collections result to fill the mapset information accordingly. You can assign any name to this object, in the example we used "tamsat". Once initialized now let's filter the time period for which you wanted to create the climatology. In this example we filter from 2017 until 2020

```
[ ]: # Filter the dates (Here 2017 to 2020)
     tamsat.filter(start_date='20170101', end_date='20201221')
```

Once filtered you can also check the list of files using *get_filenames_range()* function in a new cell. It displays all the files from 2017 to 2020.

```
[4]: tamsat.get_filenames_range()
```

```
[4]: ['/data/processing/tamsat-rfe/3.1/TAMSAT-Africa-4km/tif/10d/20170101_tamsat-rfe_10d_TAMSAT-Africa-4km_3.1.tif',
      '/data/processing/tamsat-rfe/3.1/TAMSAT-Africa-4km/tif/10d/20170111_tamsat-rfe_10d_TAMSAT-Africa-4km_3.1.tif',
      '/data/processing/tamsat-rfe/3.1/TAMSAT-Africa-4km/tif/10d/20170121_tamsat-rfe_10d_TAMSAT-Africa-4km_3.1.tif',
      '/data/processing/tamsat-rfe/3.1/TAMSAT-Africa-4km/tif/10d/20170201_tamsat-rfe_10d_TAMSAT-Africa-4km_3.1.tif',
      '/data/processing/tamsat-rfe/3.1/TAMSAT-Africa-4km/tif/10d/20170211_tamsat-rfe_10d_TAMSAT-Africa-4km_3.1.tif',
      '/data/processing/tamsat-rfe/3.1/TAMSAT-Africa-4km/tif/10d/20170221_tamsat-rfe_10d_TAMSAT-Africa-4km_3.1.tif',
      '/data/processing/tamsat-rfe/3.1/TAMSAT-Africa-4km/tif/10d/20170301_tamsat-rfe_10d_TAMSAT-Africa-4km_3.1.tif',
      '/data/processing/tamsat-rfe/3.1/TAMSAT-Africa-4km/tif/10d/20170311_tamsat-rfe_10d_TAMSAT-Africa-4km_3.1.tif',
```

Now let's initialize the processing class by passing tamsat dataset object to it and naming the initialization as proc. You can give any desired name but make sure to use the same in all the places below

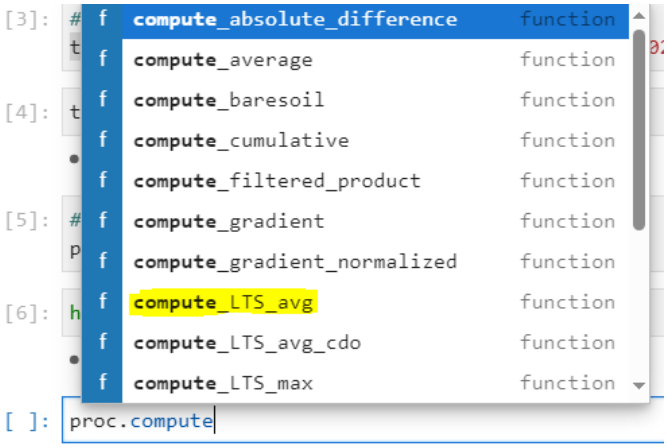
```
[5]: # Initialize proc class
     proc = Processing(tamsat)
```

Now we wanted to know which method to use in order to compute the long term climatology average, so we can either use the help function to search for appropriate method

```
[6]: help(proc)
```

```
compute_LTS_avg(self, output_sprod=None)
    ##########################################################################################################################
    # Compute Long Term Statistics average
    #       -- By default all the files are computed in the temporary file directory
    # Arguments:
    #   output_sprod '' [Optional]: Output subproduct code
    ##########################################################################################################################
```

Or you can also use "TAB" button to get the suggestion from the processing class as shown below



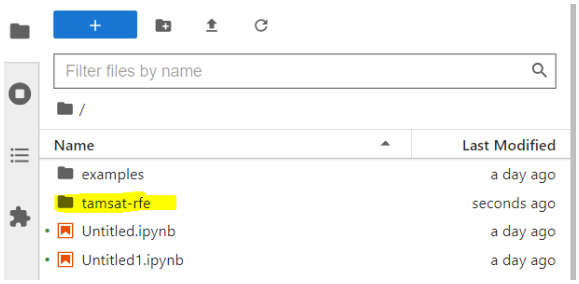
For the method compute_LTS_avg there is no mandatory arguments to pass which we noticed from the description so we can execute the method as it is by just assigning a value to it. In this example we used lts_avg_dataset.

The processing takes a while to complete depending on the computation type.

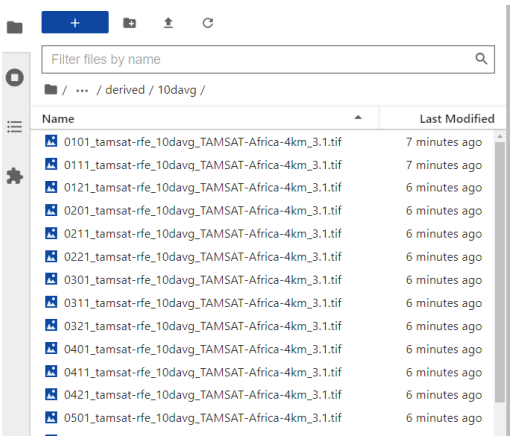
While process is running you will see the cell with [*]→

Once the process is over you see a number [7] →

If the process is completed without any error then you will find the generated files in your file explorer base directory. In our example since we used tamsat-rfe product you will notice the folder created with the same name.



You can navigate within the folder to find the files. The folder structure is in this format <product_code>/<version>/<mapset>/derived/<sub_ product_code>/<files>



If you wanted to download these files to use it in another software you can download it by right clicking on the file and choosing download. You can also list all files using *get_filenames() method.*

```
[6]: #Get all the processed file
     lts_avg_dataset.get_filenames()
```

```
[6]: ['/var/lib/jupyter-notebook/tamsat-rfe/3.1/TAMSAT-Africa-4km/derived/10davg/0101_tamsat-rfe_10davg_TAMSAT-Africa-4km_3.1.tif',
      '/var/lib/jupyter-notebook/tamsat-rfe/3.1/TAMSAT-Africa-4km/derived/10davg/0111_tamsat-rfe_10davg_TAMSAT-Africa-4km_3.1.tif',
      '/var/lib/jupyter-notebook/tamsat-rfe/3.1/TAMSAT-Africa-4km/derived/10davg/0121_tamsat-rfe_10davg_TAMSAT-Africa-4km_3.1.tif',
      '/var/lib/jupyter-notebook/tamsat-rfe/3.1/TAMSAT-Africa-4km/derived/10davg/0201_tamsat-rfe_10davg_TAMSAT-Africa-4km_3.1.tif',
      '/var/lib/jupyter-notebook/tamsat-rfe/3.1/TAMSAT-Africa-4km/derived/10davg/0211_tamsat-rfe_10davg_TAMSAT-Africa-4km_3.1.tif',
      '/var/lib/jupyter-notebook/tamsat-rfe/3.1/TAMSAT-Africa-4km/derived/10davg/0221_tamsat-rfe_10davg_TAMSAT-Africa-4km_3.1.tif',
      '/var/lib/jupyter-notebook/tamsat-rfe/3.1/TAMSAT-Africa-4km/derived/10davg/0301_tamsat-rfe_10davg_TAMSAT-Africa-4km_3.1.tif',
      '/var/lib/jupyter-notebook/tamsat-rfe/3.1/TAMSAT-Africa-4km/derived/10davg/0311_tamsat-rfe_10davg_TAMSAT-Africa-4km_3.1.tif',
      '/var/lib/jupyter-notebook/tamsat-rfe/3.1/TAMSAT-Africa-4km/derived/10davg/0321_tamsat-rfe_10davg_TAMSAT-Africa-4km_3.1.tif',
      '/var/lib/jupyter-notebook/tamsat-rfe/3.1/TAMSAT-Africa-4km/derived/10davg/0401_tamsat-rfe_10davg_TAMSAT-Africa-4km_3.1.tif',
      '/var/lib/jupyter-notebook/tamsat-rfe/3.1/TAMSAT-Africa-4km/derived/10davg/0411_tamsat-rfe_10davg_TAMSAT-Africa-4km_3.1.tif',
```

In order to view these files, it is necessary to use some of the python libraries. Check Visualize_processed_tif_files.ipynb notebook in the same folder for reference. First lets import some important libraries like numpy, matplotlib etc

```
[1]: from PIL import Image
     import matplotlib.pyplot as plt
     import numpy as np
```

Once imported select the image which you want to view and assign to a variable “image_path” and use Image.open to open the images.

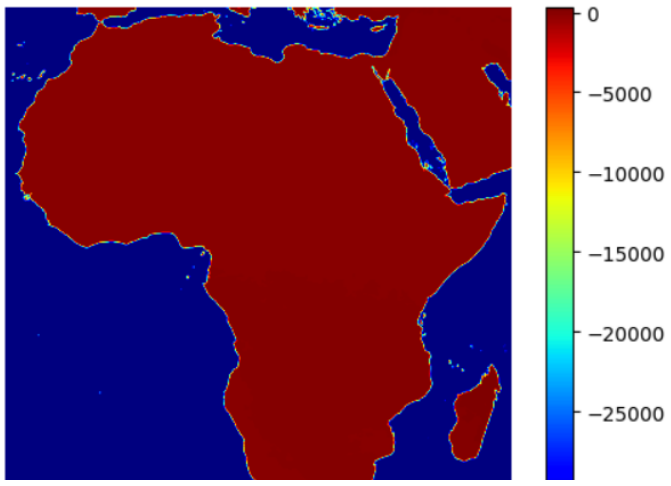
```
[2]: # Load the TIF image
     image_path = "/var/lib/jupyter-notebook/tamsat-rfe/3.1/TAMSAT-Africa-4km/derived/10davg/0101_tamsat-rfe_10davg_TAMSAT-Africa-4km_3.1.tif"
     image = Image.open(image_path)
```

Once the image is opened use numpy to read its array

```
[3]: # Convert the image to a NumPy array
     image_array = np.array(image, dtype='float')
```

In order to display the image you can use pyplot like shown below

```
[4]: # Display the image
     plt.imshow(image_array, cmap="jet")
     plt.axis("off")
     plt.colorbar()
     plt.show()
```



In this color bar, you will notice there are negative values displayed in blue. Lets see the array in order to understand what is the nodata value.

```
[13]: image_array.min()
```

```
[13]: -32768.0
```

This is the not realistic value of precipitation, so it is necessary to understand the nodata value of the input file before processing. Let's take one of the input data.

```
[22]: tamsat.get_filenames_range()[0]
```

```
[22]: '/data/processing/tamsat-rfe/3.1/TAMSAT-Africa-4km/tif/10d/20170101_tamsat-rfe_10d_TAMSAT-Africa-4km_3.1.tif'
```

Now we can make use of gdalinfo in order to view the file

```
[23]: !gdalinfo "/data/processing/tamsat-rfe/3.1/TAMSAT-Africa-4km/tif/10d/20170101_tamsat-rfe_10d_TAMSAT-Africa-4km_3.1.tif"
```

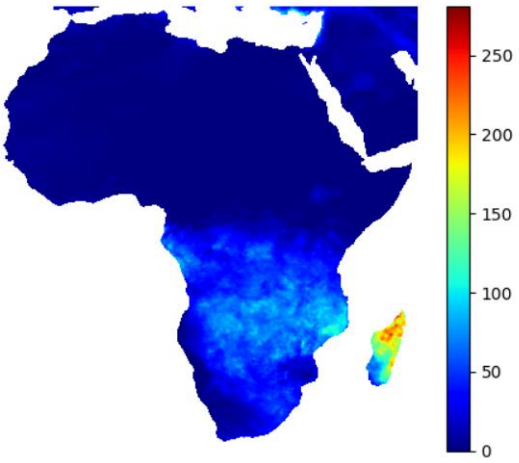
Let check the metadata part and you will notice the nodata to be -32768

```
Metadata:
  AREA_OR_POINT=Area
  eStation2_category=rainfall
  eStation2_comp_time=2021-03-15 10:37:50
  eStation2_conversion=Phys = DN * scaling_factor + scaling_offset
  eStation2_date=20170101
  eStation2_date_format=YYYYMMDD
  eStation2_defined_by=JRC
  eStation2_description=SUM composite of the RainFall Estimate (RFE) over a 10 DAY time interval
  eStation2_descr_name=10 Day RFE
  eStation2_es2_version=2.0.0
  eStation2_frequency=e1dekad
  eStation2_input_files=rfe2017_01-dk1.v3.1.nc;
  eStation2_mac_address=8f:14:db:61:8a:06
  eStation2_mapset=TAMSAT-Africa-4km
  eStation2_nodata=-32768
  eStation2_parameters=None
  eStation2_product=tamsat-rfe
  eStation2_product_version=3.1
  eStation2_provider=TAMSAT - JRC
  eStation2_scaling_factor=1.0
  eStation2_scaling_offset=0.0
  eStation2_subdir=tamsat-rfe/3.1/TAMSAT-Africa-4km/tif/10d/
  eStation2_subProduct=10d
  eStation2_unit=mm
```

so let's replace nodata value to nan and plot the image. From the metadata we also know the unit is mm

```
[15]: image_array[image_array==-32768] = np.nan
```

```
[16]: # Display the image in pseudo-color
      plt.imshow(image_array, cmap="jet")
      plt.axis("off")
      plt.colorbar()
      plt.show()
```

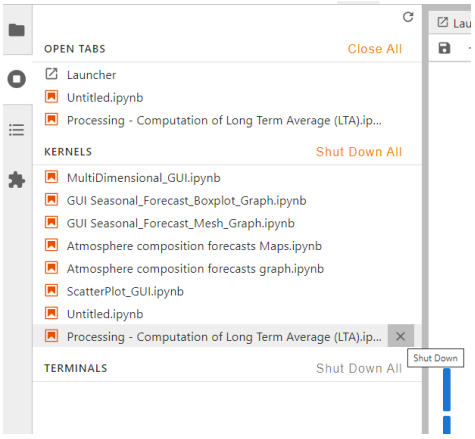


The final and most important thing after using the processing class is to close it using the method *close()* which clear all the temporary directory.

```
[7]: # Finally clear the tmporary directory where the computed files are stored
     proc.close()
```

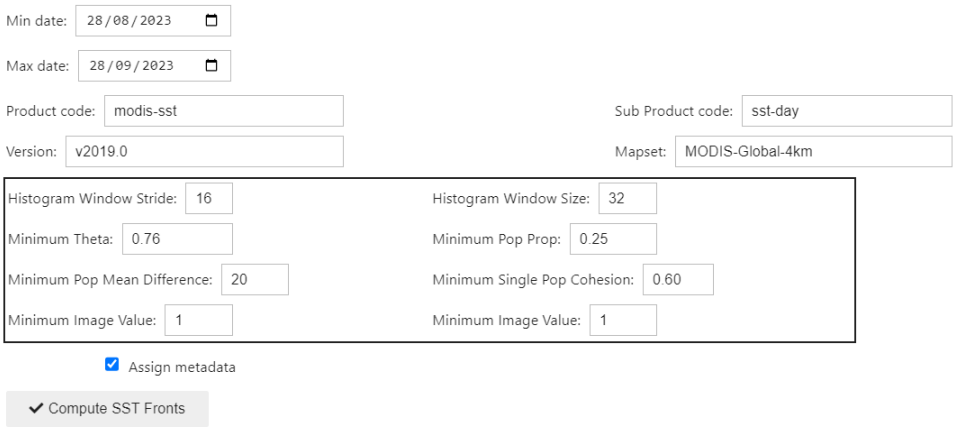
Similarly there are lot of templates to compute other indicators in the same folder examples/Processing which you can make use of.

Important Note: While moving to different notebook it is necessary to shut down the current notebook in order to avoid unwanted memory usage by the system. To shutdown you can go to interface menu in the left side bar and choose the kernel tab where you will find all the running kernel which can be closed one by one or everything using Shut Down All button.



Processing GUI - Computation SST front.ipynb

This is the notebook with graphical user interface in order to compute SST fronts. There is predefined processing chain to compute it when MODIS sst is activated using the default parameters but if you want to alter parameters depending on your location you can make use of this notebook. Run the notebook and you will get the GUI like below.



As we did in other GUI notebooks you can set the start & end date, datasets parameters and valid parameters according to the ocean you are looking in and compute it. You will find the results in the File explorer base directory once it is completed with success. In case of any error you can have look at the log files.

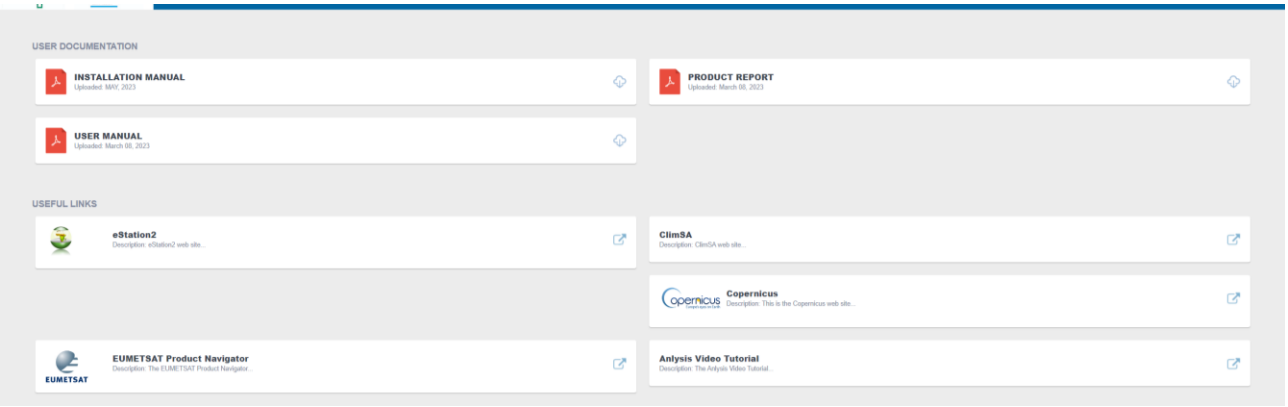
Processing Chain - Standard Precipitation.ipynb

This notebook is one of the example of predefined processing chains of precipitation where number of indicators such as absolute difference, percentage difference, and normalized precipitation are computed for 10days as well as monthly products.

3.12 HELP

The Help page gives the user access to the user documentation, e.g. this 'User Manual', and Product descriptions.

In addition, some links to useful web sites and specific Notes documentation are available on the Help page.



Clicking on a User document or Note will open the file in a new browser tab.
Clicking on a Useful link will open the link in a new browser tab.

4. REFERENCE GUIDE (ADVANCED USERS)

The current chapter provides more insights on the functioning of the EStation, and it is especially meant for Advanced Users aiming at strongly customize the application, e.g., for generating additional derived products.

4.1 DETAIL OF SERVICES

As explained in section 2.1, the EStation is mainly an EO and Climate data processor, and three main types of Services exist for processing data: the data retrieval (or 'Get' services), the format conversion ('Ingestion' service) and the generation of derived products ('Processing' service).
For each service we describe in the current section what is its role and goal ('*What it does*'), its functioning mechanism ('*How it works*') and how its configuration is managed ('*Configuration*'). The last part refers to the database tables containing the settings related to the service. An additional paragraph ('*References*') contains specific references to the filesystem components (input and directories, temporary working repositories, lock files and so on) managed by the service.

As already mentioned in section 2.4.7, there are two distinct services for retrieving external datasets on the EStation, from the EUMETCAST receiving Station (Get EUMETCast) and from remote servers (Get Internet). They have some commonalities, being the latter slightly more complex for the need of accessing various and different directories structures.

4.1.1 Get EUMETCast Service

What it does

The service copies files that are made available by Tellicast and FTS services in a directory of EUMETCAST receiving Station, which is accessed through ftp by EStation. This is a 'pull' approach and it does not take care of the house-keeping of the files in the original directory (see Figure 31)

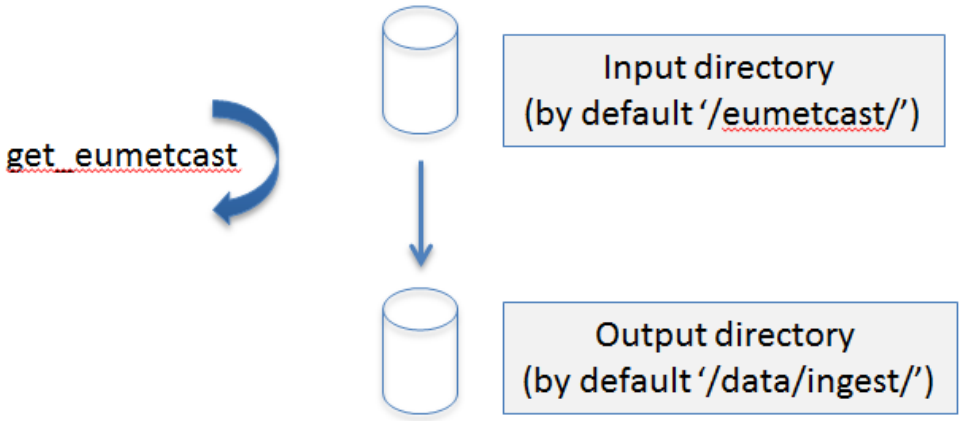


Figure 31: Get EUMETCast Service

As anticipated in paragraph 2.4, a location containing files belonging to the same EO and Climate product is defined as a 'data source'. The 'data sources' are processed independently from each other, and they can be activated/deactivated while the Service is running, so that they are taken into consideration at the next process cycle without need of a re-start.

How it works

The service should also ensure that a file in the input directory is copied to the output directory only once, and not continuously overwritten; therefore, a list[10] of the already copied files is created and maintained for each of the data sources. Note that the files removed from the input directory are also removed from the list

[10] This file is called 'get_eumetcast_processed_lists'.

and, consequently, if a file is re-disseminated after having been deleted from input dir[11], it will be copied again by the Service.

The overall organization of the Get EUMETCast service can be represented as below:

Loop over all active EUMETCast data sources and, for each of them:

- Create a list of files in input directory that match the EUMETCast source (i.e. a regular expression[12]).
- Generate list of files to be copied, i.e. the existing ones not yet copied.
- For each of the files to be copied:
  - Copy from 'input' to 'output' directory
  - Add the file to the list of copied files
- Check if all files in the list still exists in the filesystem (and clean the list accordingly)
- Save the list

Configuration

The Service is configured and controlled by two tables in the postgresql database (see also Chapter 4 of the Administration Manual).

The *eumetcast_source* table contains the description of all products disseminated by EUMETCast: Figure 32 displays a subset of the table, including the columns *eumetcast_id* (used to identify uniquely the source) and the *filter_expression_jrc*, used for associating the source the input files.

Edit Data - eStation2DB (localhost:5432) - estationdb - products.eumetcast_source

No limit

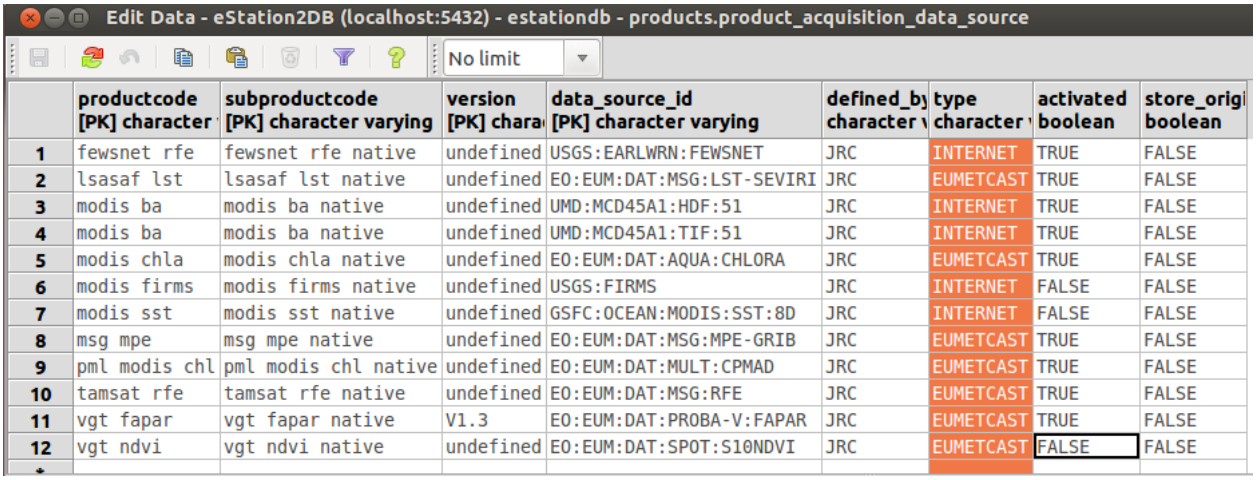
	eumetcast_id [PK] character varying	filter_expression_jrc character varying	collection_name character varying	statu: boole	internal_identifier character varying
1	EO:EUM:DAT:MSG:MSG:W:ABBA	''	MSG Active Fires - Africa	FALSE	EO:EUM:DAT:MSG:MSG:W:ABBA
2	EO:EUM:DAT:NWPMODEL:RETIM	''	M??t??o-France RETIM	FALSE	EO:EUM:DAT:NWPMODEL:RETIM
3	EO:EUM:DAT:MULT:MODIS:TC	''	MODIS True Colour Image - Multimission -	FALSE	EO:EUM:DAT:MULT:MODIS:TC
4	EO:EUM:DAT:MULT:MODIS:LFDI	''	Lowveld Fire Danger Index - Multimissior	FALSE	EO:EUM:DAT:MULT:MODIS:LFDI
5	EO:EUM:DAT:MULT:DWDSAT	''	DWDSAT	FALSE	EO:EUM:DAT:MULT:DWDSAT
6	EO:EUM:DAT:FENGYUN:PREO	''	Precipitation Estimation Product - 1 & 3	FALSE	EO:EUM:DAT:FENGYUN:PREO
7	EO:EUM:DAT:FENGYUN:DMG	''	Dust Monitoring - FengYun 2E	FALSE	EO:EUM:DAT:FENGYUN:DMG
8	EO:EUM:DAT:MULT:MODIS:BA	''	MODIS Burned Area product - Multimissior	FALSE	EO:EUM:DAT:MULT:MODIS:BA
9	EO:EUM:DAT:METOP:MGR-SST	20130521121903-OSISAF-L2P GHRSST-SSTsubs	Full Resolution Sea Surface Temperature	FALSE	EO:EUM:DAT:METOP:MGR-SST
10	EO:EUM:DAT:MULT:AOCM	8a-super.gif,	Animated Ozone Contour Map	FALSE	EO:EUM:DAT:MULT:AOCM
11	EO:EUM:DAT:NWPMODEL:ECGTS	A HRXE70ECMF090000 C ECMF 20120809000006	ECMWF NWP data for the GTS (essential ar	FALSE	EO:EUM:DAT:NWPMODEL:ECGTS
12	EO:EUM:DAT:NWPMODEL:ECACMAD	A HVXE92ECMF090000 C ECMF 20120809000006	ECMWF NWP data for ACMAD members - NWP M	FALSE	EO:EUM:DAT:NWPMODEL:ECACMAD
13	EO:EUM:DAT:AQUA:CHLORA	A.*L3m DAY CHL chlor a 4km.bz2	Chlorophyll Alpha (MODIS, Mapped 4km) -	FALSE	EO:EUM:DAT:AQUA:CHLORA
14	EO:EUM:DAT:MULT:AGRICMASK	AMESD SADC AGRIC MASK 20100901 SAfri v1.	Agriculture Mask - Multimission - Southe	FALSE	EO:EUM:DAT:MULT:AGRICMASK
15	EO:EUM:DAT:MODEL:RAIN:FCST	AMESD SADC augMM.ndjRAIN AN.fcst.txt, A	Seasonal Rainfall Forecast (LRF01) - Mod	FALSE	EO:EUM:DAT:MODEL:RAIN:FCST
16	EO:EUM:DAT:MODEL:TN:FCST	AMESD SADC augMM.ndjTN AN.fcst.txt, AME	Seasonal Minimum Temperature Forecast (L	FALSE	EO:EUM:DAT:MODEL:TN:FCST

Figure 32: *eumetcast_source* table.

A specific EUMETCast source is associated to a Station product though the *product_acquisition_data_source*, which is displayed in Figure 33. In order to uniquely identify the native subproduct, the triplet *product/subproduct/version* is specified in the table. Note also that this table is common between the two 'get' services (EUMETCast and Internet).

[11] Retention time of input directory is normally 1 week.

[12] A 'regular expression' (see https://en.wikipedia.org/wiki/Regular_expression) is a string containing characters that can match several strings, e.g. several similar filenames.



	productcode [PK] character	subproductcode [PK] character varying	version [PK] chara	data_source_id [PK] character varying	defined_by character	type character	activated boolean	store_origi boolean
1	fewsnet rfe	fewsnet rfe native	undefined	USGS:EARLWRN:FEWSNET	JRC	INTERNET	TRUE	FALSE
2	lsasaf lst	lsasaf lst native	undefined	EO:EUM:DAT:MSG:LST-SEVIRI	JRC	EUMETCAST	TRUE	FALSE
3	modis ba	modis ba native	undefined	UMD:MCD45A1:HDF:51	JRC	INTERNET	TRUE	FALSE
4	modis ba	modis ba native	undefined	UMD:MCD45A1:TIF:51	JRC	INTERNET	TRUE	FALSE
5	modis chla	modis chla native	undefined	EO:EUM:DAT:AQUA:CHLORA	JRC	EUMETCAST	TRUE	FALSE
6	modis firms	modis firms native	undefined	USGS:FIRMS	JRC	INTERNET	FALSE	FALSE
7	modis sst	modis sst native	undefined	GSFC:OCEAN:MODIS:SST:8D	JRC	INTERNET	FALSE	FALSE
8	msg mpe	msg mpe native	undefined	EO:EUM:DAT:MSG:MPE-GRIB	JRC	EUMETCAST	TRUE	FALSE
9	pml modis chl	pml modis chl native	undefined	EO:EUM:DAT:MULT:CPMAD	JRC	EUMETCAST	TRUE	FALSE
10	tamsat rfe	tamsat rfe native	undefined	EO:EUM:DAT:MSG:RFE	JRC	EUMETCAST	TRUE	FALSE
11	vgt fapar	vgt fapar native	V1.3	EO:EUM:DAT:PROBA-V:FAPAR	JRC	EUMETCAST	TRUE	FALSE
12	vgt ndvi	vgt ndvi native	undefined	EO:EUM:DAT:SPOT:S10NDVI	JRC	EUMETCAST	FALSE	FALSE

Figure 33: product acquisition datasource table

The configuration of the Get EUMETCast Service consists in two main steps:

1. Defining (modifying/adding) the sources in *eumetcast_source table*.
2. Associating the source to the products in the *product_acquisition_data_source* table.
3. Activating/Deactivating the single source in the *product_acquisition_data_source* table.

The first two steps are pre-set in the EStation by JRC for all the products proposed to the Users, and need to be done only for additional products (Advanced Users). The activation/deactivation might be modified by the User according to its thematic needs and some specific operational constrains. The procedure for performing these operations though the GUI is described in Section 3.

References

Table 3 contains all the elements in the filesystem, both files and directories, that are relevant for the Service implementation, and is meant as a Reference for advanced Users.

Element	Directory	File	Example/Default	Description
Input Dir	/eumetcast/	-	-	Input directory of the service.
Output Dir	/data/ingest/	-	-	Output directory of the service.
Process	<BASE_DIR[13]>/EStation/apps/acquisition/	get_eumetcast_py	-	Python module that implements the Service.
Process pid file	/tmp/EStation/services/	get-eumetcast.pid	-	Stores the pid[14] of the service.
Processed list	/EStation/get_lists/get_eumetcast/	get_eum_processed_list_<source_id>.list	get_eum_processed_list_EO:EUM:DAT:MSG:MPE-GRIB.list	List of the files already copied for a specific source.

[13] The base directory of the eStation installation.

[14] Process Identifier, a unique integer number associated to the Linux process.

Ancillary Info	/EStation/get_lists/get_eumetcast/	get_eum_processed_list_<source_id>.info	get_eum_processed_list_EO:EUM:DAT:MSG:MPE-GRIB.info	Information on the execution of the Service for the specific source (displayed in the GUI).

Table 3: Filesystem elements relevant for Get EUMETCast service

4.1.2 Get Internet

What it does

The service copies files available on remote ftp/http serves to the local machine, into a directory that is by default the input directory of the ingestion service.

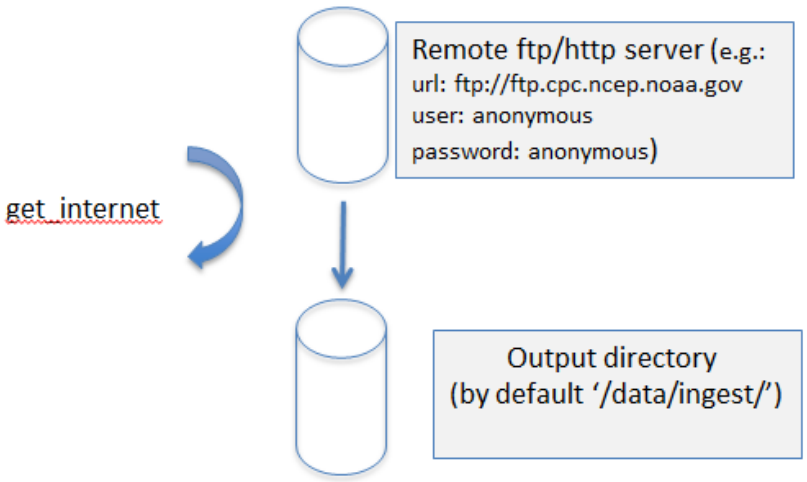


Figure 34: Get Internet Service

This service is similar to Get EUMETCast, for the principle of having different 'data sources' for the various datasets, and for keeping a list of the already copied files. Some differences exist for the need of specifying more elements to access the remote server, and to take account of the directory and file naming and organization.

How it works

The main differences in the mechanism of the Get Internet Service with respect to the Get Internet are:

1. In addition to the location of the remote files (a URL address), user credentials have to be provided (username and password).
2. On the remote server a complex directory structure can exist (see examples below).
3. Unlike for the Near Real Time (NRT) dissemination of EUMETCast, a full collection of data normally exists on the remote server, which potentially covers a large time-window; therefore time sub-setting has normally to be considered.

Note also that when a file is removed from the remote source, its name is not removed from the list of the processed file. As a consequence, files removed and re-inserted on the remote server are not downloaded a second time: the only option for re-downloading files is to manually clean the processed list.

The overall process for the Get Internet service is described below:

Loop over active Internet sources

- Create a list of files on the remote server that match the internet source definition (either type 1 or 2).
- Compute list of files to be copied, i.e. the ones available but not yet copied
- For each file to be copied:
  - Download the file to the local target dir.
  - Add the file to the list
- Save the list

Configuration

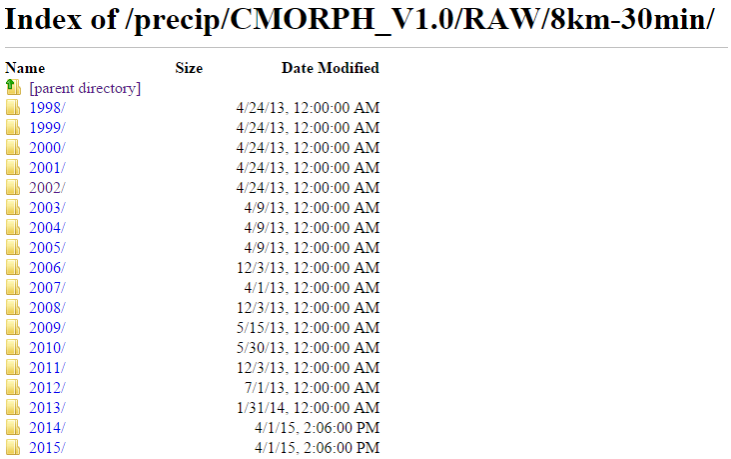
The Get Internet service takes into account the existence of both ftp servers, whose directory tree can be 'navigated' to search for the requested files, and http servers, where the exact location and naming of the files has to be known in advance. The two cases are described separately hereafter.

Type 1: ftp servers

Let's start from an example and consider the NOAA ftp server at the address ftp://ftp.cpc.ncep.noaa.gov, which is possibly a very rich and articulated site we access for retrieving, e.g., the CMORPH V 1.0 dataset. The specific dataset we are interested in (8 km resolution, 30 minute repeat cycle raw data) is located under the directory:

ftp://ftp.cpc.ncep.noaa.gov/precip/CMORPH_V1.0/RAW/8km-30min/

This address is therefore the starting point of our search in the server: as displayed in Figure 35, data are organized in subdirectories named after the year. The filename is like CMORPH_V1.0_8km-30min_201103.tar.



Index of /precip/CMORPH_V1.0/RAW/8km-30min/

Name	Size	Date Modified
[parent directory]		
1998/		4/24/13, 12:00:00 AM
1999/		4/24/13, 12:00:00 AM
2000/		4/24/13, 12:00:00 AM
2001/		4/24/13, 12:00:00 AM
2002/		4/24/13, 12:00:00 AM
2003/		4/9/13, 12:00:00 AM
2004/		4/9/13, 12:00:00 AM
2005/		4/9/13, 12:00:00 AM
2006/		12/3/13, 12:00:00 AM
2007/		4/1/13, 12:00:00 AM
2008/		12/3/13, 12:00:00 AM
2009/		5/15/13, 12:00:00 AM
2010/		5/30/13, 12:00:00 AM
2011/		12/3/13, 12:00:00 AM
2012/		7/1/13, 12:00:00 AM
2013/		1/31/14, 12:00:00 AM
2014/		4/1/15, 2:06:00 PM
2015/		4/1/15, 2:06:00 PM

Figure 35: Example of ftp server (CMORPH dataset)

The idea is to identify the files to be downloaded through a regular expression, composed by two parts:

- A fixed prefix that represent the starting point of the search. It is called 'url' and in our example is:

  url = ftp://ftp.cpc.ncep.noaa.gov/precip/CMORPH_V1.0/RAW/8km-30min/

- A variable part for identifying all subdirectories and filenames we are interested in. In our example is:

  include_files_expression = [12][0-9][0-9][0-9]/CMORPH_V1.0.*

  Note that the part '[12][0-9][0-9][0-9]/' corresponds to the 'year' subdirectory, and CMORPH_V1.0.* matches all files, regardless to their date.

These two variables are sufficient to define the 'internet source' for the ftp servers.

Type 2: http servers

Unlike for the ftp servers, on the http servers there might be restrictions in reading the contents of a directory, so that it is possible to access and download a given file, but not to 'walk' the directory tree down to that file. Consider as an example the Ocean Colour datasets distributed by GSFC-NASA at the address http://oceandata.sci.gsfc.nasa.gov/cgi/getfile/: entering in a browser this address the remote directory is not displayed, while it is possible to download the files by indicating the full name, e.g.:

http://oceandata.sci.gsfc.nasa.gov/cgi/getfile/A2015048.L3m_DAY_CHL_chlor_a_4km.bz2

As a consequence, we cannot read the contents of the remote directories and match it with some regular expression; the full filename has instead to be known in advance, including the variable part related to the observation date. The approach adopted for the http servers is therefore slightly more complicated than for the ftp servers, and is based on 3 elements, as described below.

- A fixed 'url' is provided as the initial part of the URL address, e.g.:

  url = http://oceandata.sci.gsfc.nasa.gov/cgi/getfile/

- A 'template' is provided to define the remaining part of the path (subdirectories and filename), which depends on the date. This part is still called 'include_files_expression', as for ftp servers, but its definition is different. It is not anymore a 'regular expression', rather a template containing *%type* elements that represents part of a date field (e.g. year, month, day of the month), and for the CHL source of our example will be:

  include_files_expression = A%Y%j.L3m_DAY_CHL_chlor_a_4km.bz2

  Note that the same notation as in UNIX *date* function is adopted[15] : in the following table we reproduce the format fields most commonly used:

Format	Description	Example
%Y	4-digit year	2015
%m	2-digit month	12
%d	2-digit day of	01
%H	2-digit hour	23
%M	2-digit minute	59

[15] See e.g. http://www.cyberciti.biz/faq/linux-unix-formatting-dates-for-display/

%j	3-digit day-of-year (from 1 to 366)	121
%{dkm}[16]	1 digit dekad of month	1,2 or 3

Table 4: Date formats

- Three fields are defined to define all dates in a given period, namely:
  - start_date, end_date of the period
  - frequency, i.e. the repeat cycle of the product (every day, every 30 minute)

On the basis of the 'start_date', 'end_date' and frequency, all possible dates are computed and for each of them the corresponding filename is derived by using the expression defined through the *include_file_expression*. This filename is added to the initial path specified at point a, in order to have the full path.

Configuration

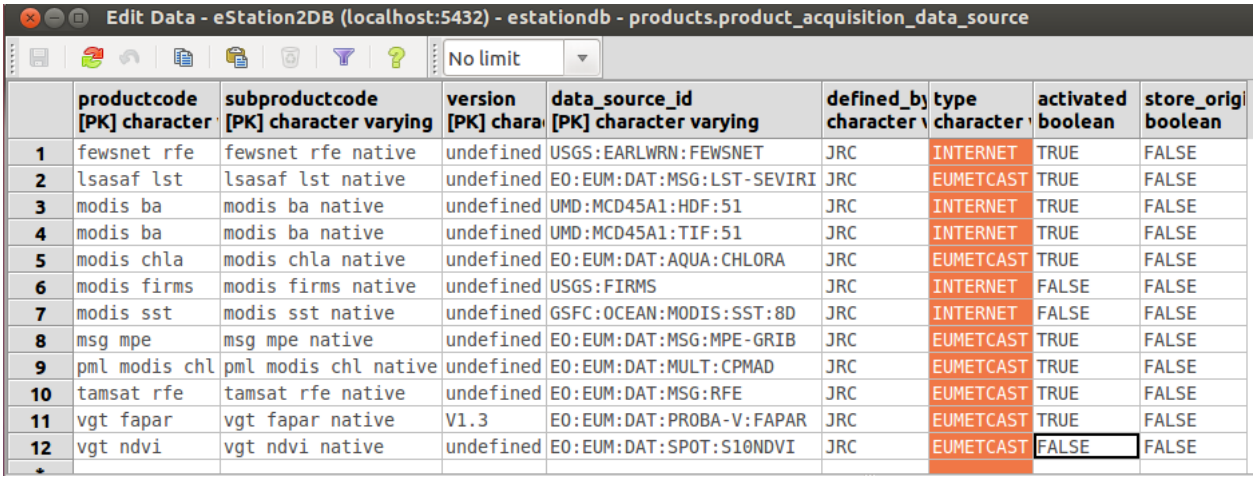
The Service is configured and controlled by two tables in the postgresql database (see also chapter 4 of the Administrator Manual). The 'internet_source' table contains the information necessary to identify the remote source and the files to be retrieved from it; in Table 5 a description of the columns relevant for the current discussion is provided.

[16] This format is a specific implementation on the eStation, and represent the 'decade' in the range 1..36.

Column	Description	Example/List
internet_id	Unique Identifier, user defined	GSFC:CGI:MODIS:CHLA:1D
defined by	Who has defined the field (either JRC or the user)	JRC
descriptive Name	A descriptive name to identify the source	MODIS 4km Chla Daily
description	A (possibly more detailed) description	MODIS 4km Chla Daily
url	Url address of the ftp or http server. It includes the 'fixed' part of the full path.	http://oceandata.sci.gsfc.nasa.gov/cgi/getfile/ ftp://ftp.cpc.ncep.noaa.gov
username	User name for server access	anonymous
password	Password	anonymous
type	Type of server to be accessed (ftp or http)	ftp: type 1 above http_tmpl: type 2
include_filter_expression	Expression to match the variable part of the full path. It is a 'regular expression' for ftp servers and a 'template' for http ones.	A%Y%j.L3m_DAY_CHL_chlor_a_4km.bz2
files_filter_expression	Expression for matching the downloaded files in the ingestion phase. It is always a regular expression (also for the http) and refers only to the filename (not subdirectories).	.*.L3m_DAY_CHL_chlor_a_4km.bz2
status	Status of activation of the source: it should always be on, unless the source is obsolete or still under test. Note that the activation of the get for the source is done in pads table.	True
pull_frequency		
frequency	The string identifying the repeat cycle (or frequency) of the dataset to be retrieved. It applies only to http server.	e1dekad (i.e. every 'dekad) e1month (i.e. every month)
start_date	Start date of the period to be considered, in format YYYYMMDD. It applies only to http server.	20150101
end_date	End date of the period to be considered, in format YYYYMMDD. It applies only to http server.	20150631

Table 5: Contents of Internet source table (partial)

One, or more, specific sources can be associated to a 'native' subproduct though the 'product acquisition data table', which is displayed in Figure 33. Note that the table is common between the two 'get' services.



Edit Data - eStation2DB (localhost:5432) - estationdb - products.product_acquisition_data_source

No limit

	productcode [PK] character	subproductcode [PK] character varying	version [PK] chara	data_source_id [PK] character varying	defined_by character	type character	activated boolean	store_origi boolean
1	fewsnet rfe	fewsnet rfe native	undefined	USGS:EARLWRN:FEWSNET	JRC	INTERNET	TRUE	FALSE
2	lsasaf lst	lsasaf lst native	undefined	EO:EUM:DAT:MSG:LST-SEVIRI	JRC	EUMETCAST	TRUE	FALSE
3	modis ba	modis ba native	undefined	UMD:MCD45A1:HDF:51	JRC	INTERNET	TRUE	FALSE
4	modis ba	modis ba native	undefined	UMD:MCD45A1:TIF:51	JRC	INTERNET	TRUE	FALSE
5	modis chla	modis chla native	undefined	EO:EUM:DAT:AQUA:CHLORA	JRC	EUMETCAST	TRUE	FALSE
6	modis firms	modis firms native	undefined	USGS:FIRMS	JRC	INTERNET	FALSE	FALSE
7	modis sst	modis sst native	undefined	GSFC:OCEAN:MODIS:SST:8D	JRC	INTERNET	FALSE	FALSE
8	msg mpe	msg mpe native	undefined	EO:EUM:DAT:MSG:MPE-GRIB	JRC	EUMETCAST	TRUE	FALSE
9	pml modis chl	pml modis chl native	undefined	EO:EUM:DAT:MULT:CPMAD	JRC	EUMETCAST	TRUE	FALSE
10	tamsat rfe	tamsat rfe native	undefined	EO:EUM:DAT:MSG:RFE	JRC	EUMETCAST	TRUE	FALSE
11	vgt fapar	vgt fapar native	V1.3	EO:EUM:DAT:PROBA-V:FAPAR	JRC	EUMETCAST	TRUE	FALSE
12	vgt ndvi	vgt ndvi native	undefined	EO:EUM:DAT:SPOT:S10NDVI	JRC	EUMETCAST	FALSE	FALSE

Figure 36: product acquisition datasource table

References

Table 6 contains all the elements in the filesystem, both files and directories, that are relevant for the Service implementation, and is meant as a Reference for advanced Users.

Element	Directory	File	Example/Default	Description
Input Dir	/eumetcast/	-	-	Input directory of the service
Output Dir	/data/ingest/	-	-	Output directory of the service
Process	<BASE_DIR>/EStation/apps/acquisition/	get_eumetcast_py	/srv/www/EStation/apps/acquisition/get_eumetcast.py	Python module in charge of the service
Process pid file	/tmp/EStation/services/	get-eumetcast.pid		Stores the pid[17] of the service.
Processed_list	/EStation/get_lists/get_eumetcast/	get_eum_processed_list_<source_id>.list	get_eum_processed_list_EO:EUM:DAT:MSG:MPE-GRIB.list	
Ancillary Info	/EStation/get_lists/get_eumetcast/	get_eum_processed_list_<source_id>.info	get_eum_processed_list_EO:EUM:DAT:MSG:MPE-GRIB.info	

Table 6: Filesystem elements relevant for get_eumetcast service

[17] Process Identifier, a unique integer number associated to the Linux process.

4.1.3 Data Store Service

What it does

The service copies files available on CDS & IRI climate library to the local machine based on the configuration files which are added by the user(predefined files by JRC) understanding the data retrieval mechanism from the CDS and IRI portal. For CDS, configurations taken from the API request within their dataset retrieval portal whereas for IRI data library, configurations taken from the expert mode of specific datasets from their portal

On top of coping the files it also ingest them directly as netcdf file based on the pre-process type defined. In some case it just assign the metadata and move the file to appropriate locations.

How it works

The main differences in the mechanism of the Data Store Service with respect to the Get Internet are:

- Get Internet just download the data in /data/ingest folder whereas Data Store service also ingest them
- Another main difference is Data Store service ingest the data as NetCDF file

The overall process for the Data Store service is described below:

Loop over active datastore sources

- Create a jobs based on the configuration available for the specific datasource
- Once the jobs are processed by the portal (In case IRI data library data is downloaded instantly)it downloads the data
- The downloaded files are ingested in the similar way described in Ingestion service
- Save the list of downloaded files

Figure 37 gives the detailed process of data flow within the DataStore service

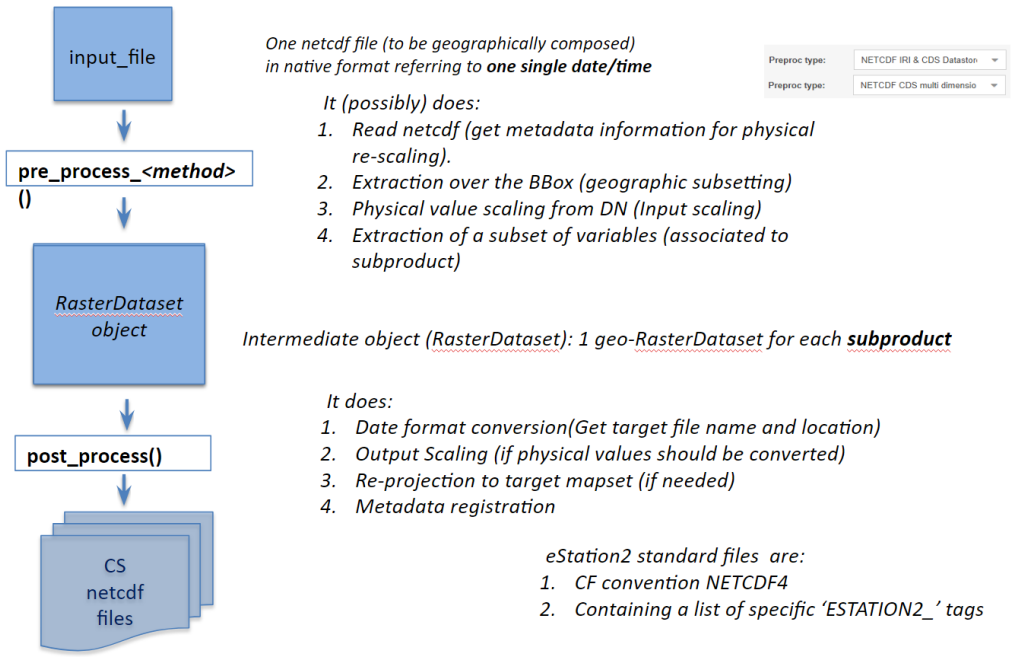


Figure 37: Datastore data flow

Configuration

The two cases are described separately here.

Type 1: CDS

The configuration from the CDS are converted to configuration file with the name of the datasource ID and loaded into the folder /data/static_data/config_cds/.

In the Figure 38 you will how configurations are transformed in the CS config file. Later this config files are processed by filling the year, month,day & time and jobs are created to retrieve the data

Figure 38: Example of ERA5 configuration transformed into configuration file in EStation

Type 2: IRI data library

The IRI data library expert mode configurations of specific dataset is converted to the configuration file in the CS as shown in Figure 39. The time range is alone processed from the Datastore service and feed into the config file to process the data

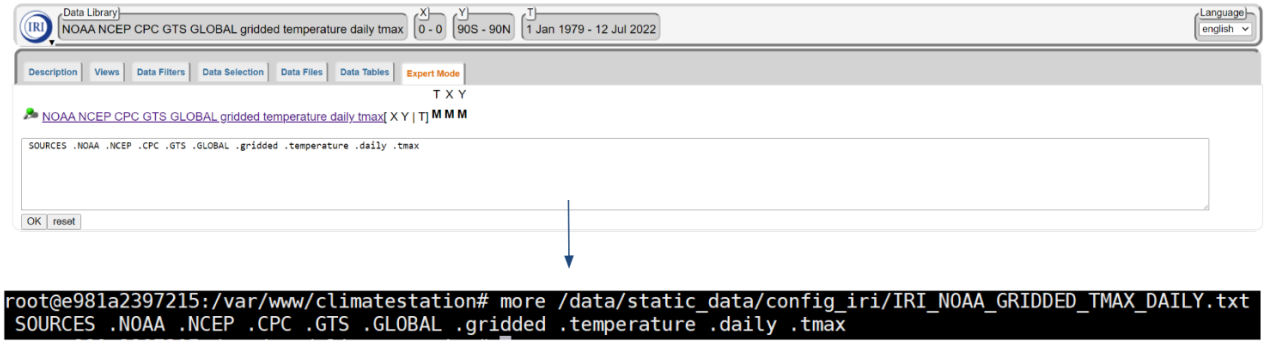


Figure 39: Example of GTS Tmax configuration from IRI transformed into configuration file in EStation

4.1.4 Ingestion Service

What it does

The main goal of the ingestion Service is to extract from the retrieved files the *subproducts* needed by the thematic User, for the specific *mapsets* he has defined. These *subproducts* are therefore stored in the standard EStation format (GTIFF containing specific tags), and ready on the system for visualization and further processing.

The complexity of the ingestion service, with respect to the *Get* services, relies, a part from the geo-processing of various formats, on the fact that several *subproducts* can be extracted from the same files, and

for more than one *mapset*. Furthermore, the same *product* might have been retrieved from difference *sources* (e.g. EUMETCast and Internet, or different internet servers), having each source a different file naming and format.

How it works

To deal with the above described complexity, the overall service is organized in two main steps:

- Identify the files from a *source* to be processed for a specific *product.*
- Process the files to extract the *subproducts* for the defined *mapset (*or *mapsets).*

Step 1: select the files for a product/source

The overall mechanism of the ingestion loop is described in Figure 40. Its role is to select a list of files existing in the input directory to be passed to a specific routine that extracts from them the defined *subproducts* for the active *mapsets*.

As first action, the list of all *products* whose ingestion is active (see also Figure 10) is created. For each of them, the *sources* it has been retrieved from are identifying, and each *source* is treated separately, as the files coming from the various sources of the some product might be in a different format, or at least have a different filename.

Once a *source* is selected, it is possible to identify in the input directory all the files retrieved from that source, and to group them by date[18]. These files can subsequently be treated in order to extract from them one or more *subproducts*, for one or more *mapsets.*

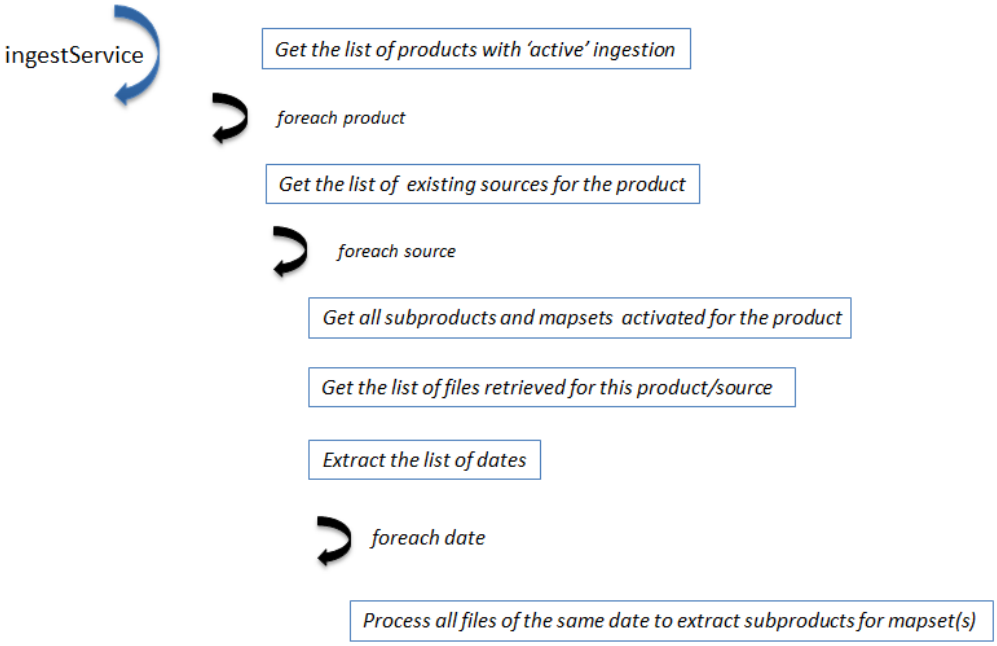


Figure 40: Ingestion Service structure

[18] For each date, one or more file can exist, depending on the policy adopted by the data provider. In same case, namely the low resolution products, a single file exist, having continental or global coverage. With higher spatial resolutions, the information is stored in various files, also known as tiles or regions.

Step 2: process the files to extract subproducts

A single file, or a list of files covering adjacent geographic areas, are passed to a routine to process them and extract the numerical value to be converted into physical values and stored in EStation format. Each of the file can be in various file format (e.g. GTIFF, HDF4, HDF5, netcdf, HRIT, grib) and can contain one or more layers. The operations performed in this step are listed in Figure 41.

A *pre-processing* is applied in order to have, as intermediate step, a set of GTIFF and geo-referenced files containing a single *subproduct.* According to the *native* format, the series of performed operation is different. In the most general case, the *pre-processing* does the following:

- Unzip the files (from .gzip, .tgz, .bz2)
- Extract the physical values for each subproduct, and do mosaicking
- Write the values in a GTIFF format
- Geo-reference the file.

This pre-processing highly depends on the nature of the input files: a set of pre-processing routines are defined in order to deal with the most common cases.

Once the intermediate files are generated in a temporary working space, the generation of the EStation standard files is performed by:

- Converting the digital numbers to physical values, and convert back to digital number with a standard convention (see 2.4).
- Apply the geographic clipping/re-projection to generate the output with the defined *mapset* (i.e. for a specific boundary box and projection)
- Write to the files the EStation metadata (see 2.4)

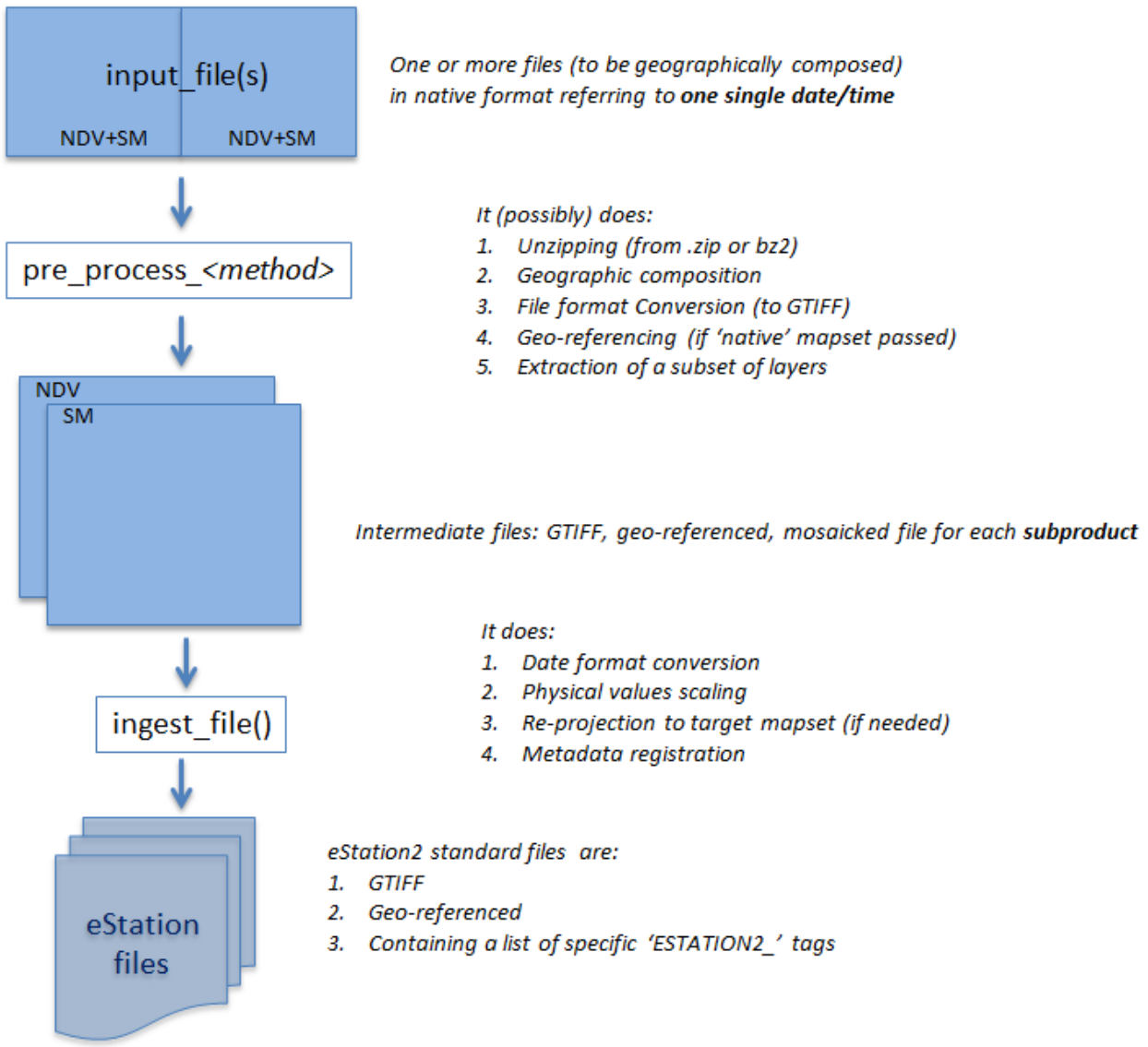


Figure 41: Ingestion data flow

Configuration

Several tables are involved in the configuration of the ingestion mechanism, and their relationship is represented in Figure 42, in a simplified manner. The tables used for the Get services are also involved because, as already specific, the format and naming of files containing the some product can vary according to the source.

The Product table is the pivot table for the EStation: therein all retrieved/processed and visualized products have to be defined. The Ingestion table establish the relationship between a *product/subproduct*[19]and the mapset we will apply in the ingestion. The product_acquisition_data_source table associate a *product* and a source, which can be of 'EUMETCast' or 'Internet' type. This table has been already described in previous paragraphs, as well as the Datasource_Description (see 2.4.6), which mainly defines the rule adopted for the file naming. The Sub_Datasource_Description table contains the description of the contents of the input files (i.e. how many layers are present, what are the scale factor and offset, nodata coding). This table is directly

[19] Always identified by the *product/version/subproduct* triplet.

linked to the Product table, to establish a relationship between the various layers in the files, and the associated *subproducts.*

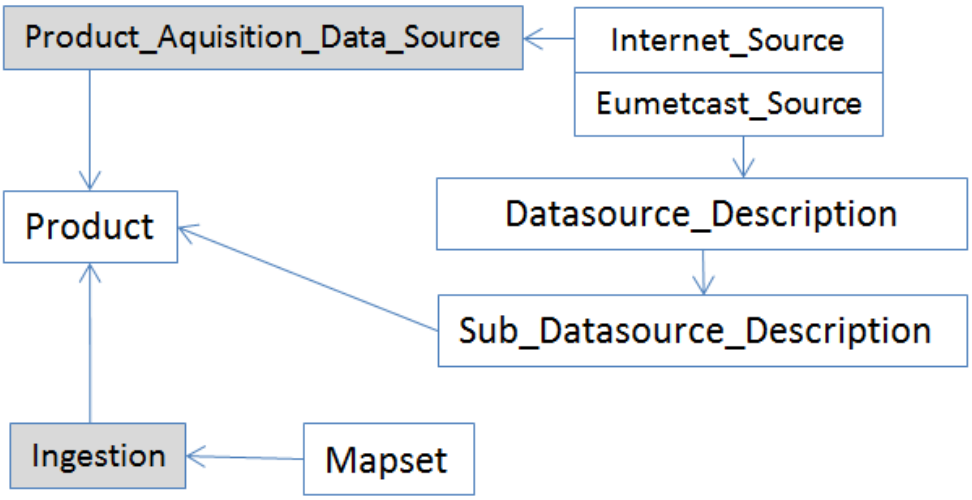


Figure 42: DB tables for ingestion

4.1.5 Processing Service

What it does

The 'processing' service is devoted to compute EO and Climate products and indicators from the ones already existing in the system, i.e. to implement some algorithms and put them in operations. These algorithms can include re-projection functionalities, computation of temporal composition (e.g. from 10d to 1 month precipitations), computation of long term statistics and anomalies, or more complicated operations. For the implementation of the service we base on the 'ruffus' python library[20], which is part of the EStation installation.

How it works

As a first step, the 'processing' chain to be implemented is defined by detailing the starting product (or products) and all computed products generated from them, which can represent either an intermediate result or a final indicator to be used.

In Figure 43 the processing chain for 'standard precipitation' products is displayed. The starting product is the 10-day Rainfall estimate (RFE) from which we compute the inter-annual statistics: min, max, average and median (step 1). From the product and these statistics several 'anomalies' are computed in Step 2 (10ddiff, 10dperc, 10dNPcum). Then, in step 3.a, we restart from the 10d product and we compute the cumulated precipitation over the month (1monCum); subsequently, statistics (step 3.b) and anomalies (step 3.c) are derived at the monthly level.

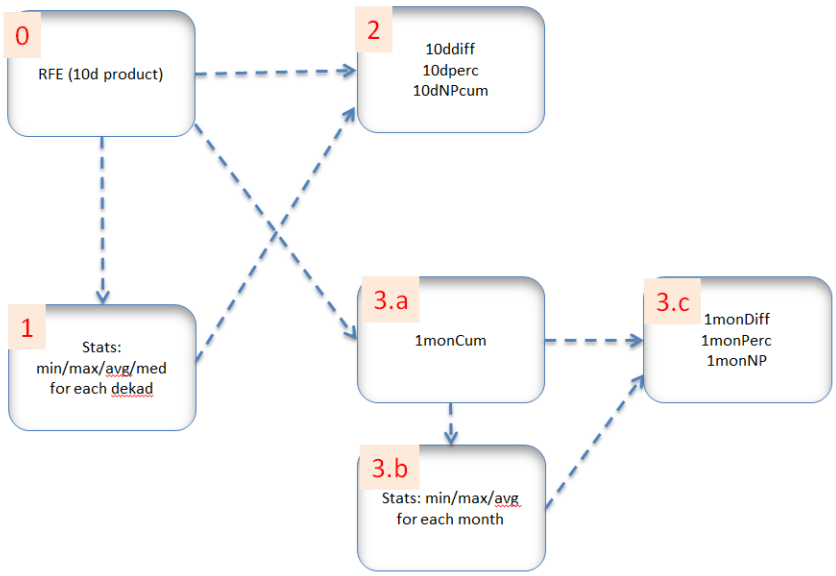


Figure 43: example of a processing chain

This processing chain is coded in python by following the ruffus approach, i.e. by creating a 'pipeline' that receives in input the 10d product, and defines all products depending on it. The added value of the 'ruffus' library is that, once all dependencies between input and output products are defined, the code is able to determine autonomously which outputs are missing with respect to the available inputs, and to trigger the computation of these images only.

[20] See http://www.ruffus.org.uk/

On the EStation release, a number of chains are defined, as described in Table 7; additional processing chains will be provided in the following releases.

Algorithm name	Options	Inputs	Outputs (by group)	Notes
std_precip	std_precip_prods_only std_precip_stats_only	A single precipitation product (e.g. 10d RFE)	10d stats 10d anomalies 1mon cumulate 1moncum stats 1moncum anomalies	The 'stats_only' option compute 10d-stats and 10dcum-stats only. The 'prods_only' option compute 10d-anomalies and 1moncum and 1moncum-anomalies only. The 'all' stats computes everything. This processing chain is displayed in Figure 43)
std_ndvi	std_ndvi_prods_only std_ndvi_stats_only	A single ndvi product (e.g. vgt-ndvi).	See	
merge_versions	-	Two or more products whose timeseries have to be joined (e.g. NDVI from SPOT-v1, SPOT-v2 and PROBA-V)	A single output (the merged timeseries from the inputs)	For each input product, the temporal window to be considered has to be provided.
proc_fronts	-	A Sea Surface Temperature (SST) product.	Sst-fronts	
proc_precip_1day		Daily rainfall product	10d, 1mon, 1year, 3mon, 6mon	
proc_opfish		Chorophyll data	opfish	
proc_gradient		Chorophyll data	gradient	

Table 7: List of available processing chains

The full processing chain of ndvi is graphically represented in Figure 44 and Figure 45. The starting point is a single 'NDV' product, and the following steps are defined:

1. Compute the statistics (1.a) and anomalies (1.b) from NDVI.

2. Apply a two-steps filtering procedure to filter the cloud contaminated pixels (2.a), and on the basis of the final result (ndvi_filterx2):

    2.b Computed statistics

    2.c Compute a 'baremask' product (to filter out the non-vegetated areas)

    2.d Compute the anomalies (e.g. ICN, VCI)

3. From the filtered NDVI, compute the monthly product (3.a), and from it:

    3.b Compute statistics

    3.c Compute the monthly 'baremask' product

    3.d Compute the monthly anomalies

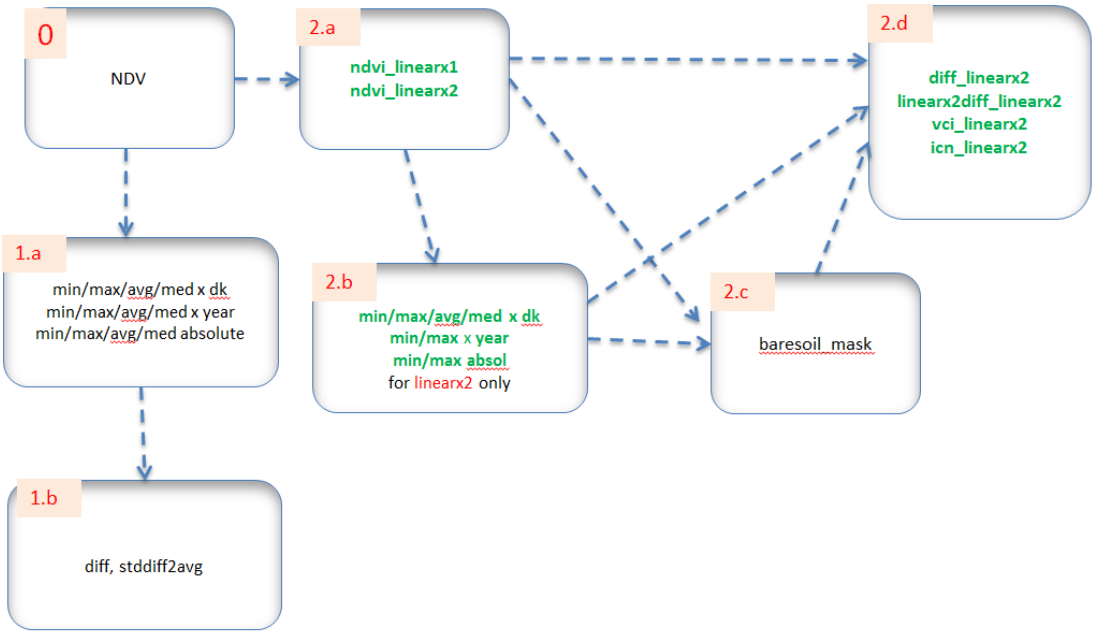


Figure 44: std_ndvi processing chain

The detailed description of this chain (the most complex implemented in the application) goes beyond the scope of this Manual, and is treated in the training sessions.

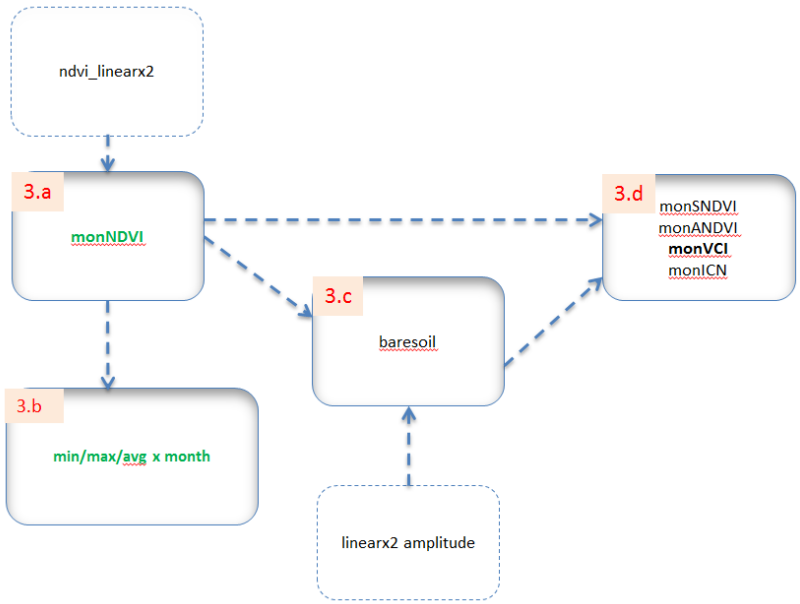


Figure 45: std_ndvi processing chain - 2

4.1.6 System Service

What it does

The System service is in charge of all 'background' operations that are not directly involved in the other Services, but essential for the good functioning of the machine, e.g. the synchronization of the data and database. In particular the service is in charge of the following operations:

1. Data synchronization
2. Database synchronization
3. Database dump
4. Cleaning temporary directories
5. Checking metadata information and check the corrupted files.

How it works

The System service works as a continuous loop and iterates over the above-described operations and, for each of them:

- Determines if the operation has to be executed.
- Determines if the 'time' requirements for the operations are met: some operations are executed at a given hour, or every number of hour or minutes.
- Executes the operation.

Configuration

The service is not based on database settings, rather on configuration files, namely the 'System Settings' and 'Factory Settings'.).

The 'System Settings' file is parsed to determine the Role and Mode of the computer, which are defined at the installation (role) or modified in case of malfunctioning of one of the 2 computers.

The 'Factory Settings' that impact the System service are displayed in Table 8. On the 2.1 release they are not reported in the 'System' interface, and therefore not modifiable by the User.

Parameter name	Default value	Role
system_delay_data_sync_min	10	Delay time between executions of data sync
system_time_db_dump_hhmm	00:00	Time of the day for execution of the DB dump
system_time_spirits_conv	00:10	Time of the day for execution of the DB dump
system_sleep_time_sec	10	Delay time between execution of system loop

Table 8: Factory settings influencing System Service