

ESA CCI data and CCI toolbox

Carsten Brockmann, CCI Toolbox Science Lead

With contributions from Ed Pechorro, ESA ECSAT

ESA – Future Earth Lagoon Workshop, Cork 12.-14. September 2017.



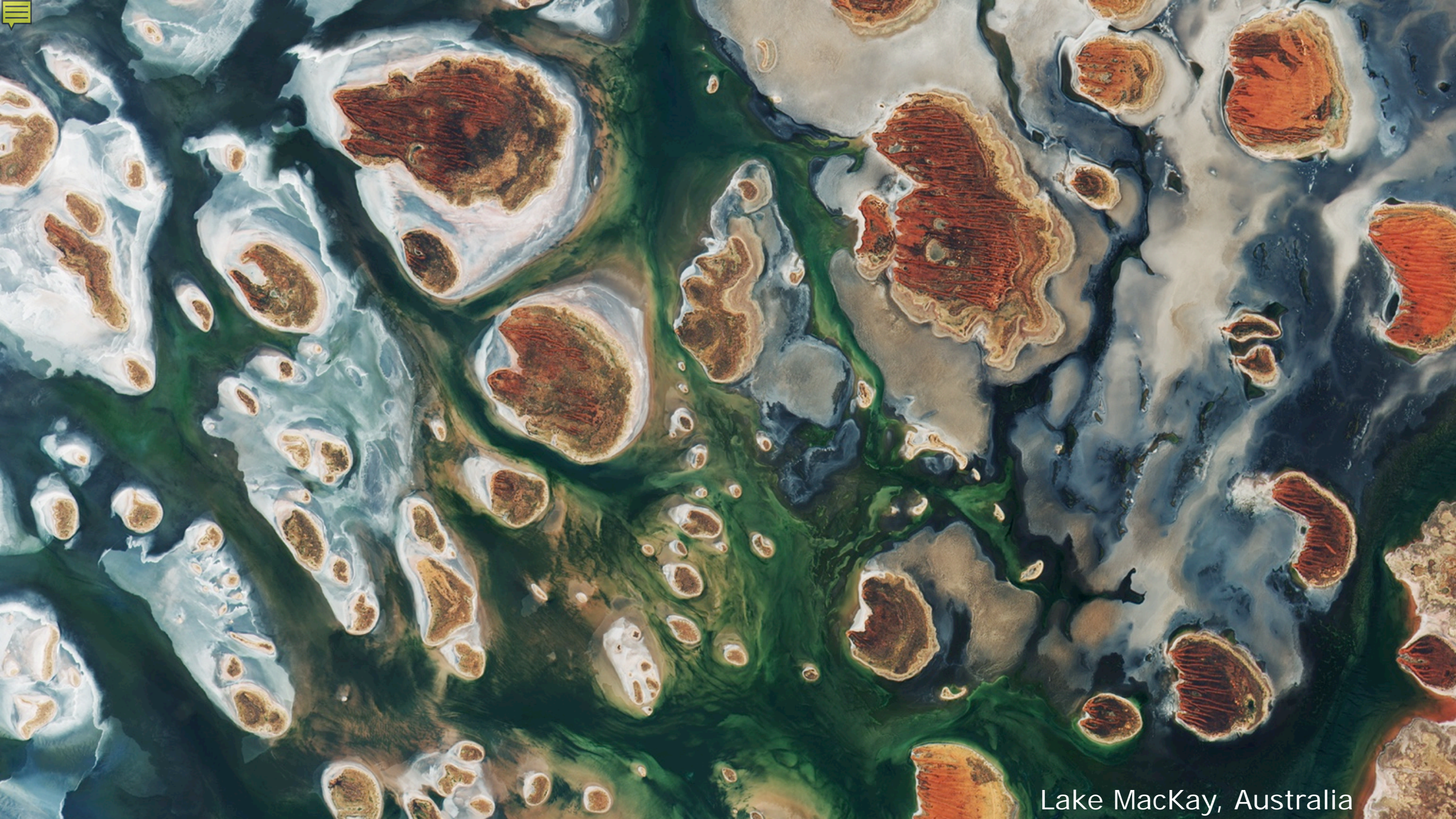
CCI Programme

Managing big climate data is increasingly complex.

But not just technically.

Collaboration. Culture. Organisation. Structure.





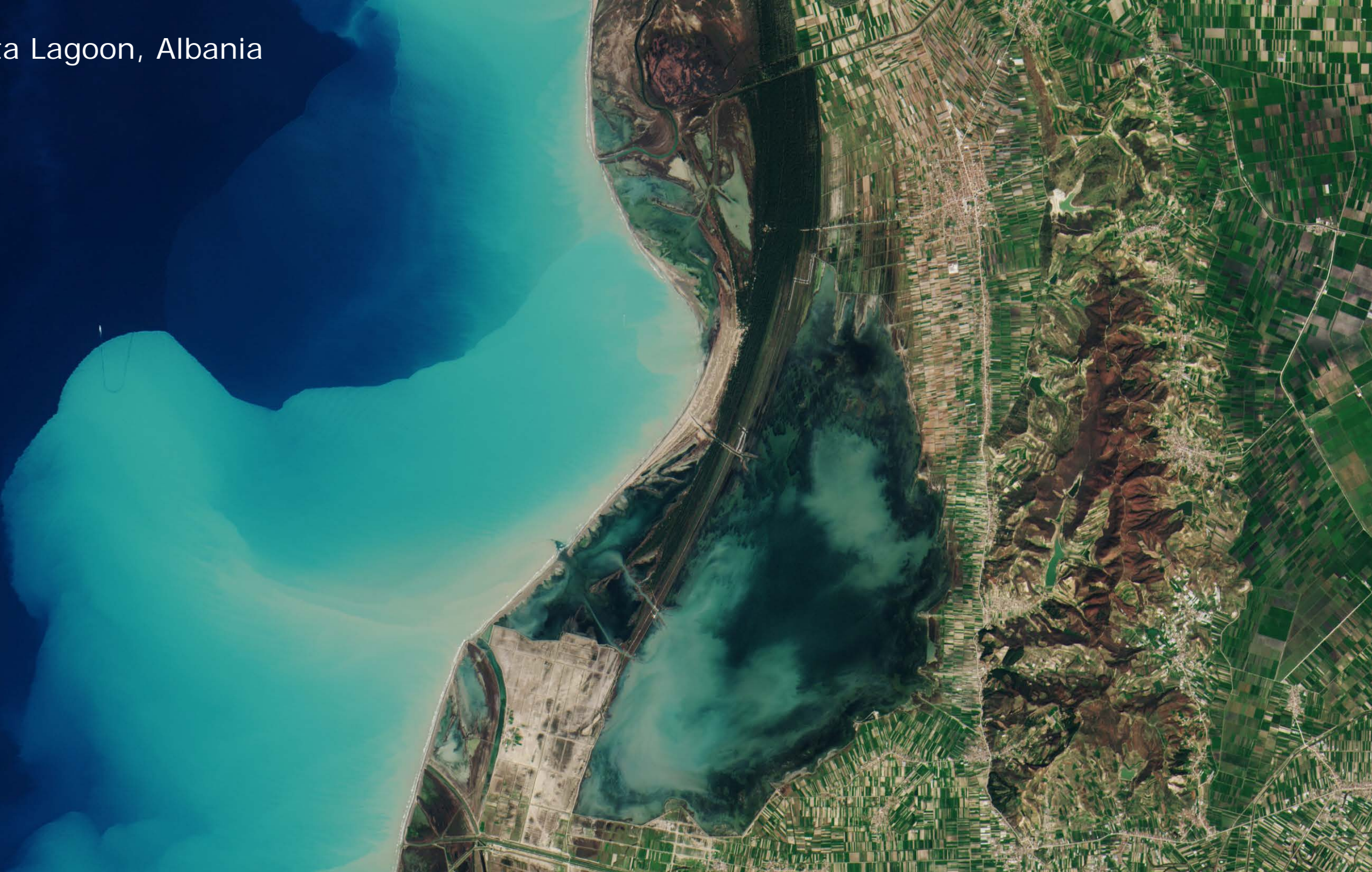
Lake MacKay, Australia

Tonga, Pacific





Karavasta Lagoon, Albania



Climate Monitoring Data

Climate change is a global challenge.

Open climate data is crucial.



ESA Climate Change Initiative (CCI)

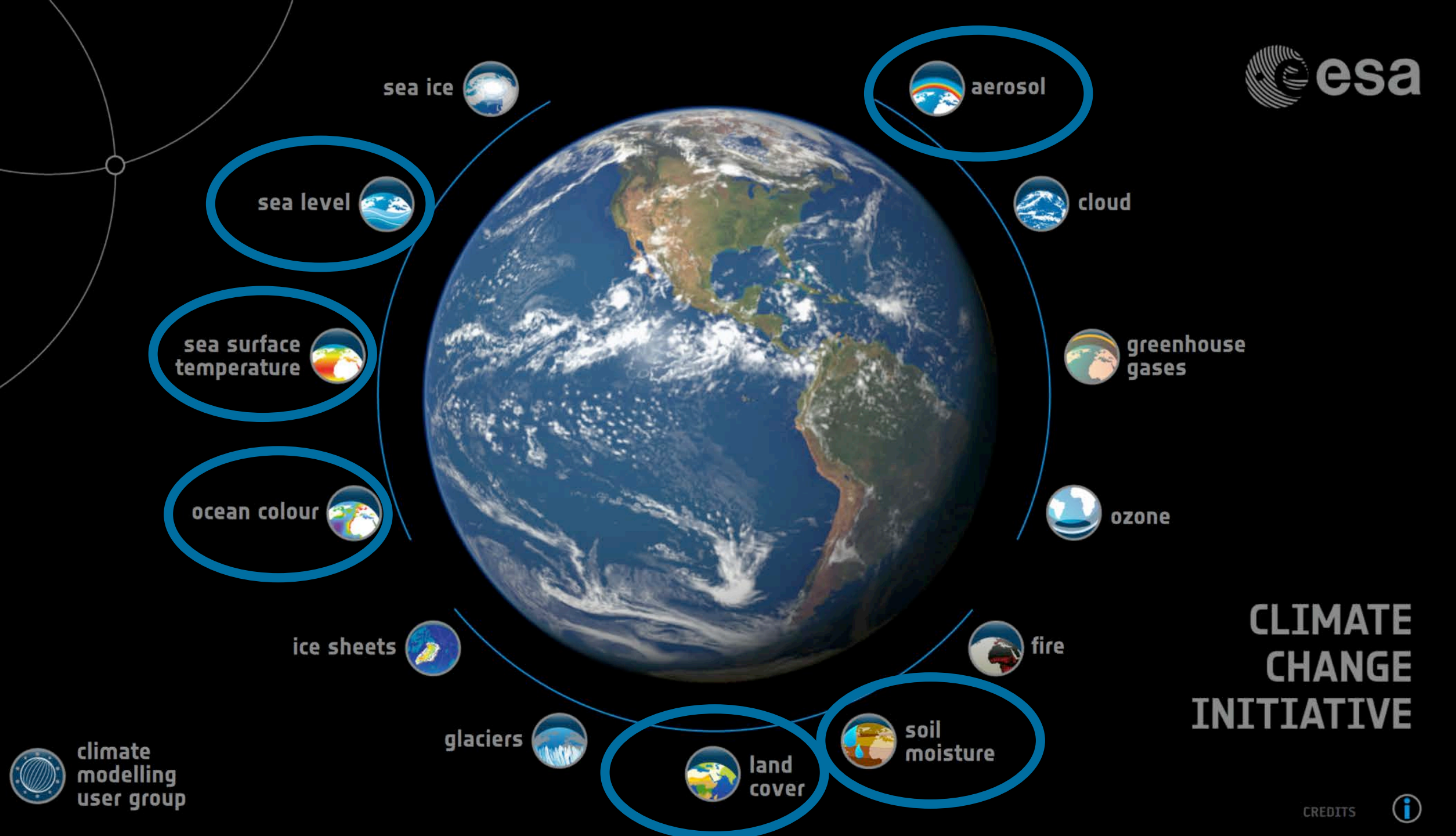
The objective of the Climate Change Initiative (CCI) is to realise the full potential of the long-term global Earth Observation archives that ESA together with its Member States have established over the last 30 years, as a significant and timely contribution to the Essential Climate Variable databases required by the United Nations Framework Convention on Climate Change (UNFCCC).



ESA Climate Change Initiative (CCI)

16 projects
>300 scientists
>100 organisations
18 countries
Since 2009





CLIMATE CHANGE INITIATIVE

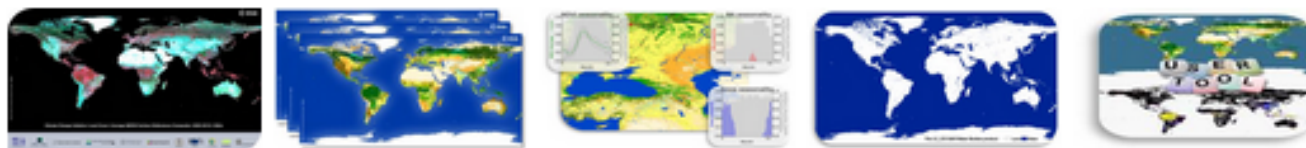
→ GCI DASHBOARD

Documents

- documents
- documents
 - Deliverables
 - Land_Cover_CCI_Ph1
 - Land_Cover_CCI_Ph2
 - Newsletters
 - Scientific communication

Name	Modified	Size	Owner
Land_Cover_CCI_ATBDv2_2.3.pdf	01/10/14 22:30	5 MB	dlederer
Land_Cover_CCI_CAR_1.1.pdf	01/10/14 22:33	16 MB	dlederer
Land_Cover_CCI_DARD_1.9.pdf	14/01/13 17:00	2 MB	dlederer
Land_Cover_CCI_DPMv2_2.3.pdf	01/10/14 22:34	3 MB	dlederer
Land_Cover_CCI_IODDv2_2.3.pdf	01/10/14 22:35	1 MB	dlederer
Land_Cover_CCI_IPVRv1_1.2.pdf	14/01/13 17:13	20 MB	dlederer
Land_Cover_CCI_PSD_1.11.pdf	01/10/14 22:44	5 MB	dlederer
Land_Cover_CCI_PUG-v2.4.pdf	02/10/14 14:16	7 MB	dlederer
Land_Cover_CCI_PVP_1.3.pdf	04/05/12 14:10	2 MB	dlederer
Land_Cover_CCI_RRP_v1.3.pdf	04/05/12 14:09	904 KB	dlederer
Land_Cover_CCI_SPD_1.1.pdf	14/01/13 17:10	6 MB	dlederer
Land_Cover_CCI_SRD_1.1.pdf	04/05/12 14:10	1 MB	dlederer
Land_Cover_CCI_SSD_1.0.pdf	14/01/13 17:15	2 MB	dlederer
Land_Cover_CCI_SVR_1.1.pdf	14/01/13 17:11	12 MB	dlederer
Land_Cover_CCI_URD_2.2.pdf	04/05/12 14:11	1 MB	dlederer

CCI Land Cover products



1980









2000



[hide legend](#), [hide header](#)[Land Cover Map 2015](#) | [MERIS surface reflectance](#) | [Water Bodies](#) | [Land Surface Seasonality](#) | [User tool](#)April 2017 new release [Download data](#)

Land cover legend

[view global \(level 1\)](#)

-  Cropland, rainfed
 -  - Herbaceous cover
 -  - Tree or shrub cover
 -  Cropland irrigated or post-flooding
 -  Mosaic cropland (>50%) / natural vegetation (Tree, shrub, herbaceous cover) (<50%)
 -  Mosaic natural vegetation (Tree, shrub, herbaceous cover) (>50%) / cropland (<50%)
 -  Tree cover, broadleaved, evergreen, closed to open (>15%)
 -  Tree cover, broadleaved, deciduous, closed to open
- Long=12.3047°, Lat=45.5742°

Documentation

- [Product User Guide v2](#)
- [Quick User Guide for Maps v2.0.7](#)
- [Quick user guide Land Surface Seasonality products](#)
- [Legend for LC Map v2.0.7](#)
- [Preview LC Map v2.0.7 for Year 2015](#)
- [Preview MERIS SR Composite](#)



Next Phase 2017 – 2026: CCI+

New ECVs:

Water Vapour

Salinity

Sea State

Lakes

Snow

Permafrost ECV

Land Surface Temperature

High Resolution Land Cover

Biomass

Lakes and Climate



The current effects of changing climate on lakes include:

- (i) an increase in the surface water temperature of lakes and streams, especially those at high altitudes and latitudes;
- (ii) an increase in temperature at lower levels of large deep lakes;
- (iii) a reduction in lake ice-cover;
- (iv) the melting of mountain glaciers and permafrost causing changes to discharge regimes in mountain streams.

It is probable that:

- (i) with changes in precipitation, there will be an increase in the transport of sediments and nutrients downstream to lakes;
- (ii) The lake size and spatial distribution may be altered as lakes appear and disappear or
- (iii) existing conditions may change abruptly, e.g., from freshwater to permanently saline lakes or from non-stratifying to stratifying lakes



Lakes: GCOS ECV Requirements

GCOS Requirements

	Lake water level	Lake Extent	LSWT	Lake Ice Thickness Cover	Lake Colour
Accuracy	3 cm for large lakes, 10 cm for the rest	10 % (relative) 5% (for 70 largest lakes)	1K	LIT: 1-2 cm LIC: 10 %	30 %
Stability	1 cm/decade	5%/decade	0.1 K/decade	LIT: N/A LIC: 1 % /decade	1 %/decade
Resolution	100 m	20 m	300 m	LIT: 100m LIC: 300 m	300 m
Frequency	Daily	Daily	Weekly	LIT: Monthly LIC: Daily	Weekly

HR Land Cover: Key Issues for CCI+ (1)



The objectives of HRLC_cci are restricted in area with a focus on:

- understanding regional change in vegetation manifest either in terms of land cover or in terms of vegetation state as it affects classification over time, as driven by climate.
- understanding classification variability across spatio-temporal scales from the moderate resolution previously tackled in LandCover_cci.
- understanding how the Land Cover ECV products at moderate resolution relate to the map products needed for mitigation and adaptation at local scales including forest monitoring and reporting activities as part of the REDD+ mechanism.



HR Land Cover: Satellite Instruments



Taking account of the LandCover_cci (MERIS, VGT, AVHRR, PROBA) heritage, the satellite focus of HR Land Cover is:

- **Optical High Resolution:** Landsat TM, ETM, OLI, SPOT HRV, HRVIR, HRG, S-2 MSI ...
- **Optical Very High Resolution:** IKONOS, QUICKBIRD-2, GEOEYE, WORLDVIEW, JSS 56, RAPIDEYE, PLANET ...
- **Microwave High Resolution:** ERS, Envisat, Sentinel 1, Radarsat, JERS-1, ALOS, ALOS-2, TerraSAR, Cosmo SkyMed,

Next Phase 2017 – 2026: CCI+

Other Programme Elements

1) New R&D on existing ECVs

2) Cross-ECV scientific exploitation

- CCI+ CMUG-type activity
- Cross-ECV targeted scientific studies

(E.g. Analysis of multiple ECVs for IMBIE (polar science), sea-level budget closure, carbon-cycle research, *etc*)

- Young Scientist Research Fellowship Scheme

3) Supporting activities on Knowledge Exchange

Climate Monitoring Data

An Overview of Climate Data Production.



Step 1. Deciding what to actually measure.

Essential Climate Variables have been defined by the global science community to support the United Nations Framework Convention on Climate Change (UNFCCC).



Criteria of Essential Climate Variables (ECV)

Relevance. Critical for climate monitoring.

Feasibility. Global measurement is feasible.

Cost effective. Using proven technology.

Satellites can help

Global. Observe entire Earth.

Uniformity. Same instrument everywhere.

Rapid Measurement. Constant watch.

Continuity. Long time series to monitor change in climate.

Step 2. Get the raw satellite data.

Current data.

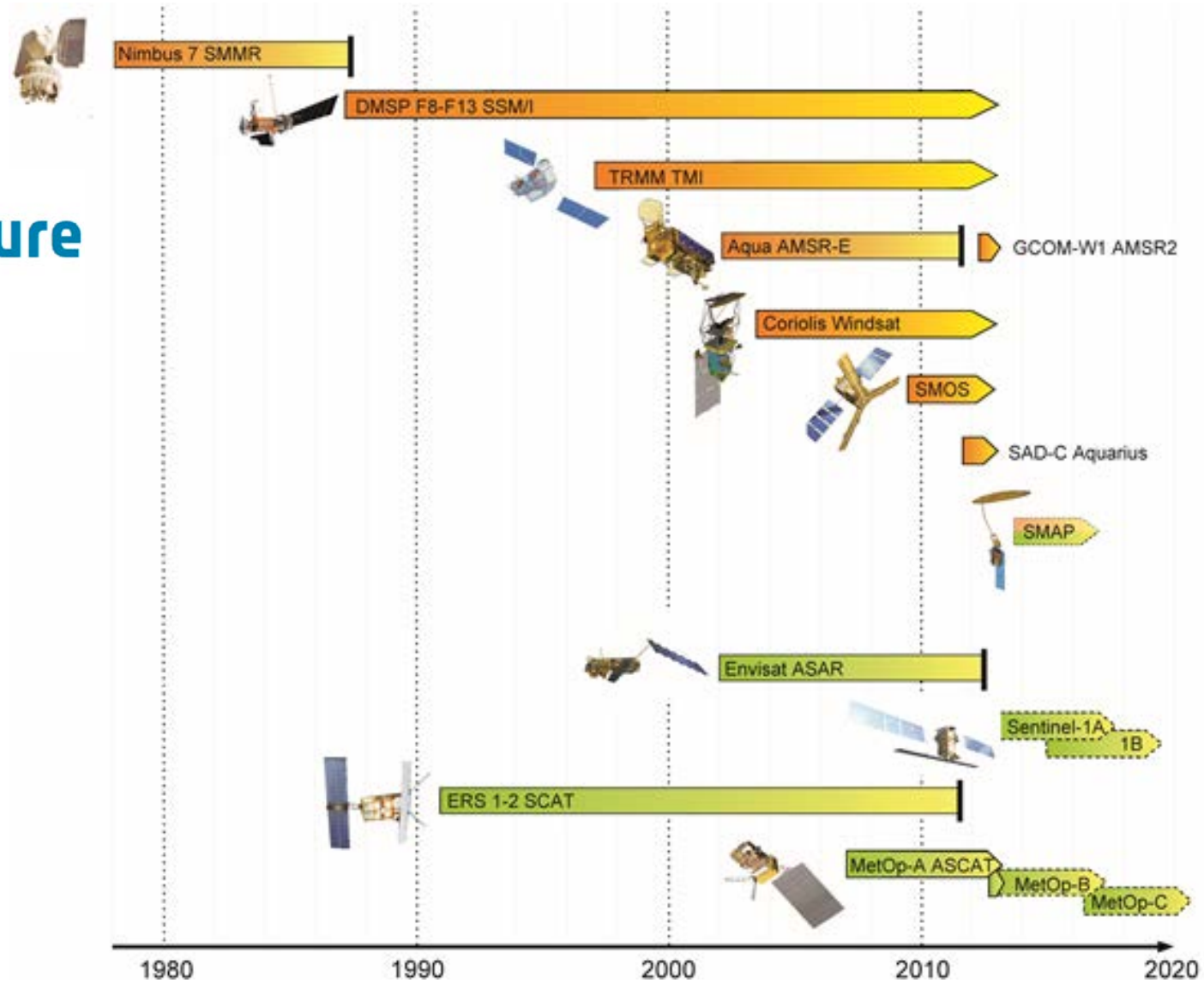
Archived data.

Planning for the future.

Example



soil moisture
cci



Step 3. Process the data.

Gridding, Homogenisation, Calibration & Validation, Quality.

Scientific processing - application of state-of-the-art algorithms distilled from the very latest scientific reasoning.

Step 4. Distribution of Climate Data Products.

“Just give me the data”.

ESA Climate Change Initiative (CCI)

Managing Complexity of Climate Data Production.

Open Data Challenges & Approaches



Managing Open Data Complexity

Meaningfulness & Community.

Managing Open Data Complexity

Ease of Data Access.







Home

Data

Submitted by Cat Downy on Wed, 13/04/2016 - 23:55



open data portal cci

Welcome to the **CCI Open Data Portal**.
A single point of entry to CCI data.
Open, free and easily accessible.

- ▶ About CCI
- Data
- ▶ Resources
- ▶ CCI Events
- Contact Us

Search



CCI Dashboard
The big picture.

CCI Search
Simple & faceted.

CCI FTP
Quick & simple.

CCI Viewer
View. Draw. Catch.

CCI Dashboard

- aerosol
cci
- cloud
cci
- fire
cci
- ghg
cci
- glaciers
cci
- antarctic
ice sheet
cci
- ice sheets
greenland
cci
- land cover
cci
- ocean colour
cci

Managing Open Data Complexity

Bespoke Open Standards.



Managing Open Data Complexity

Machine & Human Readable Standards.

Managing Open Data Complexity

Interoperability & Collaboration.



Managing Open Data Complexity

Open-source Tooling



toolbox
cci

CATE

“Climate Analysis Toolbox for ESA”

A software to facilitate processing and analysis of all the data products generated by the ESA Climate Change Initiative Programme (CCI).

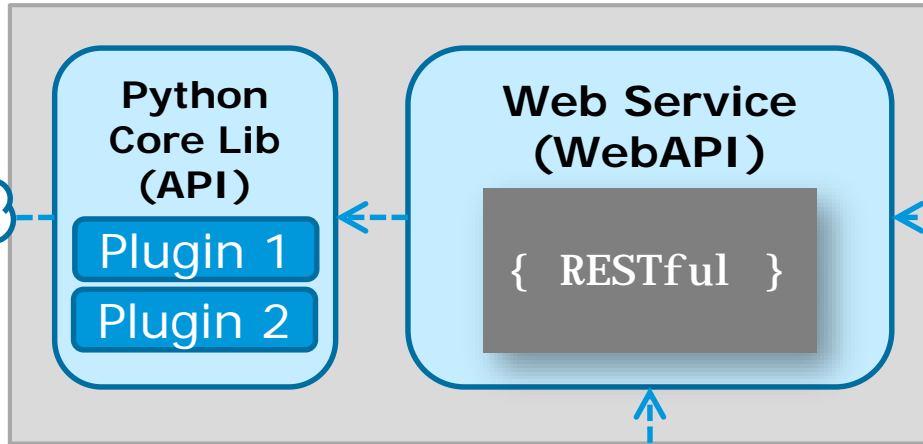
cci

→ **TOOLBOX**

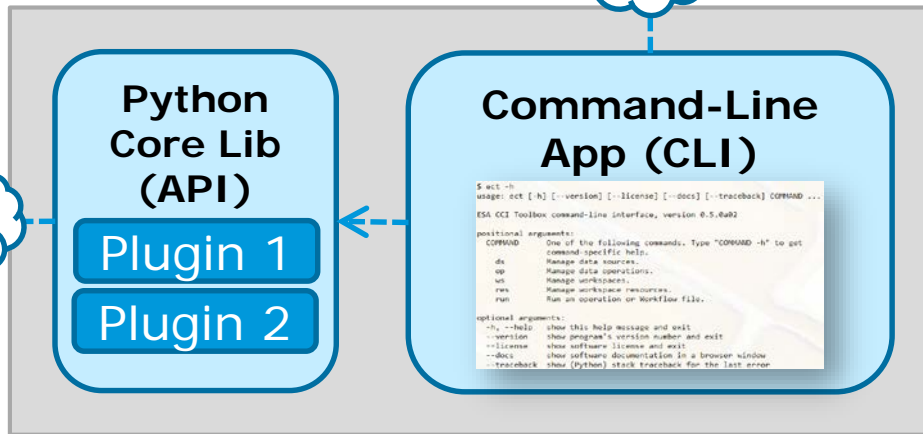
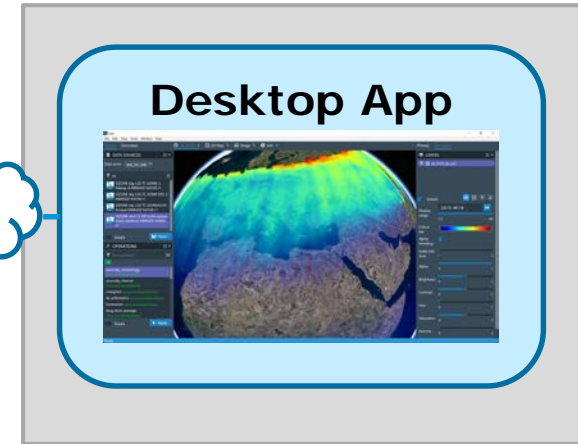
Cate Desktop (GUI)
Cate Programming (API)
Cate Command Line (CLI)

ESA Open Data Portal
and other data services

Process 2



Process 3



Process 1



Cate Desktop

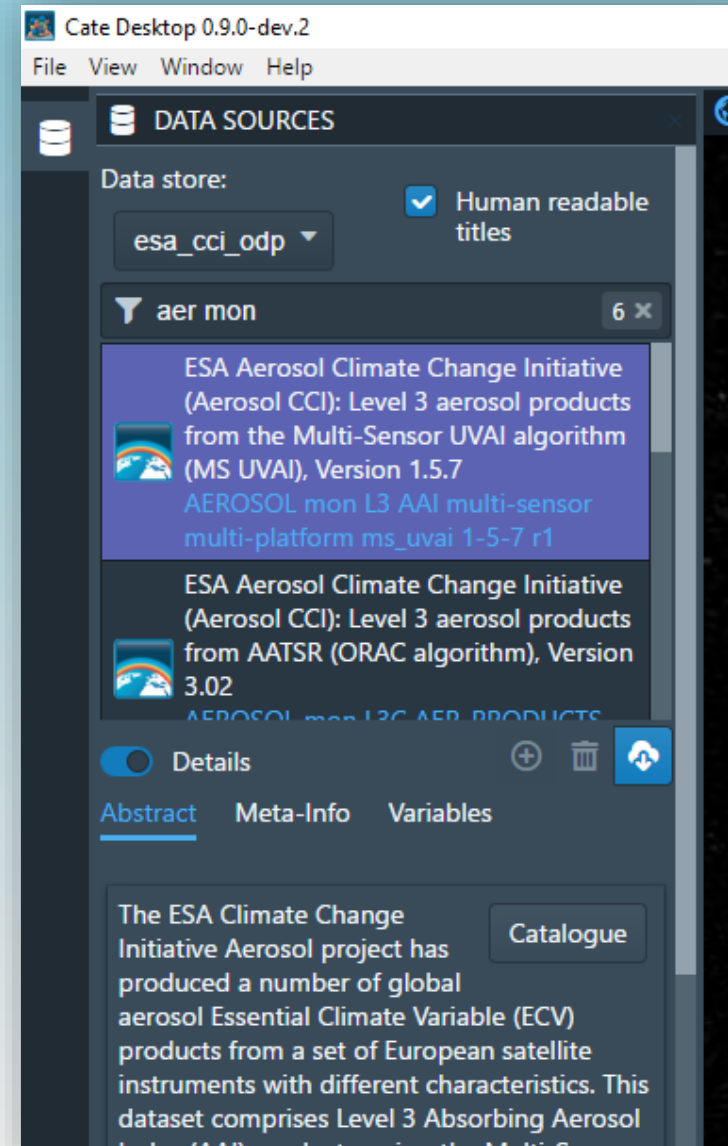


The screenshot displays the Cate Desktop application window titled "Cate Desktop 0.9.0-dev.2". The interface is divided into several panels:

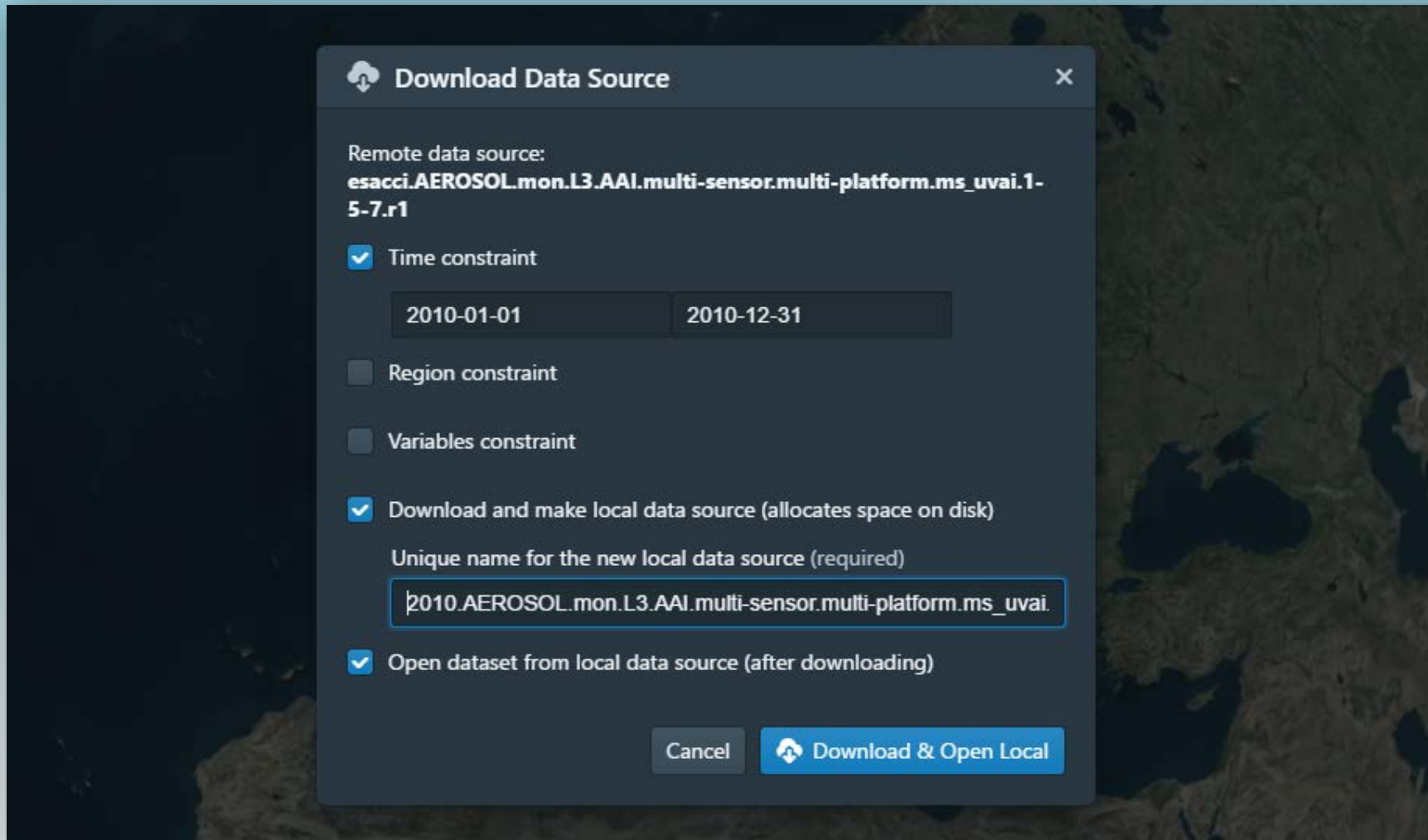
- DATA SOURCES:** Located on the left, it shows a search for "esa_cci_odp" with a "Human readable titles" checkbox checked. Two datasets are listed:
 - ESA Aerosol Climate Change Initiative (Aerosol CCI): Level 3 aerosol products from GOMOS (AERGOM algorithm), Version 2.19. AEROSOL 5-days L3C AEX GOMOS. Envisat_aergom_2-19_r1.
 - ESA Aerosol Climate Change Initiative (Aerosol CCI): Level 3 aerosol products from the Multi-Sensor UVAI algorithm (MS UVAI), Version 1.5.7. AEROSOL_cci_msa_uai_multi.
- OPERATIONS:** Below the data sources, it lists various operations such as "anomaly_external Dataset", "anomaly_internal Dataset", "coregister Dataset", "detect_outliers Dataset", "diff Dataset", "ds_arithmetics Dataset", "enso DataFrame", and "enso_nino34 DataFrame". The "anomaly_external" operation is selected, and its description is shown at the bottom: "Calculate anomaly with external reference data, for example, a climatology. The given reference dataset is expected to consist of 12".
- Workspace:** On the right, it shows a workspace named "<unnamed>" with "Workflow (0)" and "Resources (0)". It contains a message: "No workflow steps. Open a dataset in DATA SOURCES panel or apply a read_ operation from the OPERATIONS panel."
- Variables:** Below the workspace, it shows "No variables. Select a resource in the WORKSPACE panel first."
- Map:** The central part of the interface is a large globe showing a satellite view of Earth, centered on Europe and Africa.

At the bottom of the window, the status bar shows "Ready." and coordinates "lon=-58.03, lat=49.50".

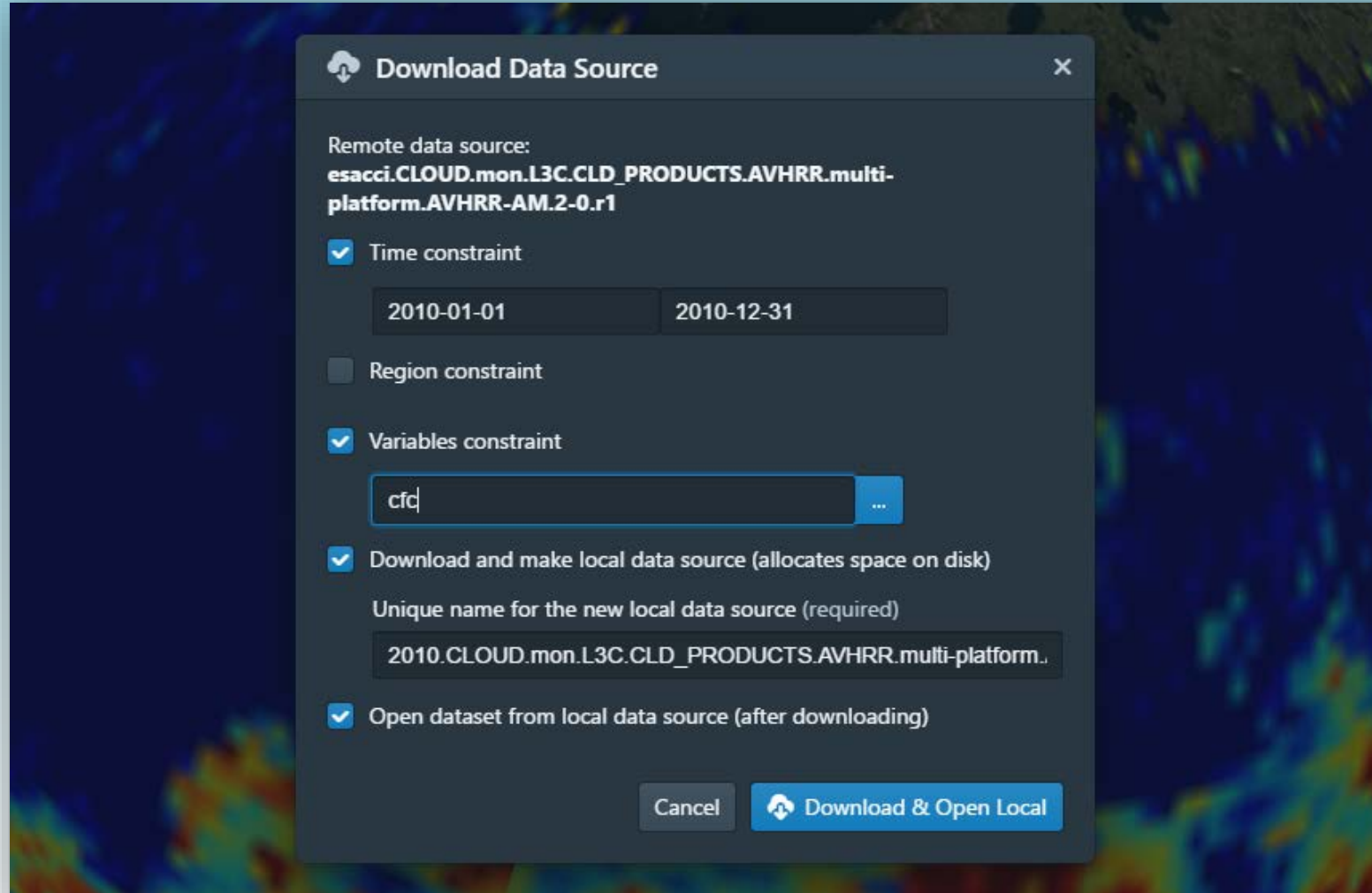
- Browse datasets published by **CCI Open Data Portal**
- Download full datasets or just subsets
 - Temporal subset
 - Spatial subset
 - Variable subset
- Manage also your **local data sources**



Access CCI Data directly from ODP



Constraining the Selection



Multiple Views – Compare ECVs



DATA SOURCES

Data store: local Human readable titles

Find data source 2 x

- ESA Aerosol Climate Change Initiative (Aerosol CCI): Level 3 aerosol products from the Multi-Sensor UVAI algorithm (MS UVAI), Version 1.5.7
local 2010 AEROSOL_mon L3 AAI multi-sensor multi-platform ms_uvai 1-5-7_r1
- ESA Cloud Climate Change Initiative (Cloud_cci): ATSR2-AATR monthly gridded cloud properties, version 2.0

OPERATIONS

Find operation 46 x

- anomaly_external Dataset
- anomaly_internal Dataset
- coregister Dataset
- detect_outliers Dataset
- diff Dataset
- ds_arithmetics Dataset
- enso DataFrame
- enso_nino34 DataFrame

Workspace

<unnamed> Modified

Workflow (2) Resources (2)

- res_1 Dataset
- res_2 Dataset

Attributes:

Name	Value
title	ESA Cloud CCI Ret...
project	Climate Change In...
product_version	2.0
Conventions	CF-1.6,ACDD-1.3
standard_name_vo...	NetCDF Climate F...
institution	RAL
source	AATR ENVISAT_V3.1
number_of_proces...	1191

VIEW

New World View

View title

World (2)

View mode

- 3D Globe
- 2D Map

Projection (for 2D Map only)

EPSG:4326

Show layer text overlay

Imagery layer credits

CESIUM | bing | Image courtesy of NASA • © 2017 Intermap • Earthstar Geographics SIO • © 2017 Microsoft Corporation

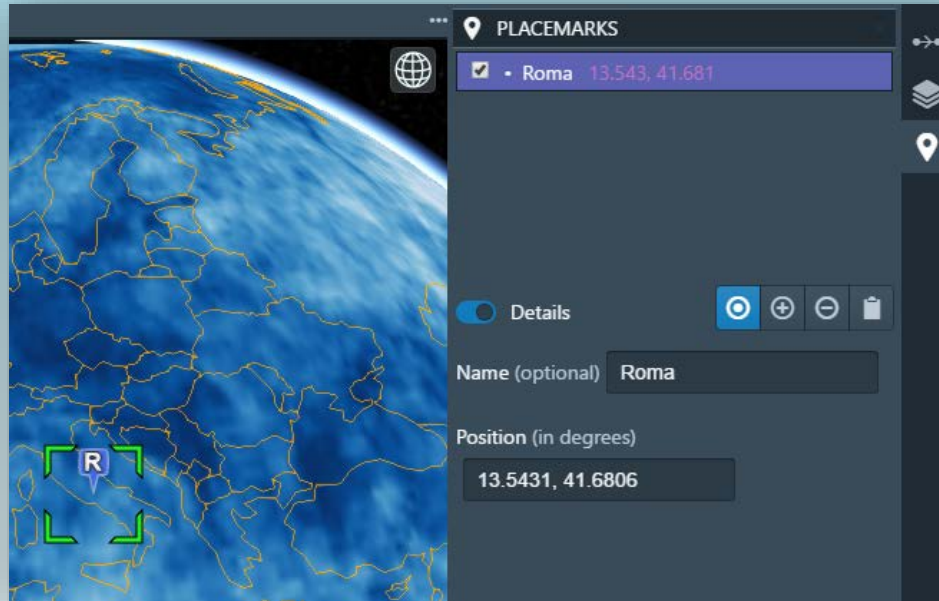
World (1) x World (2) x

Sel. var.: res_1/absorbing_aerosol_index: time = ...

Sel. var.: res_2/cfc: time = 2010-01-01

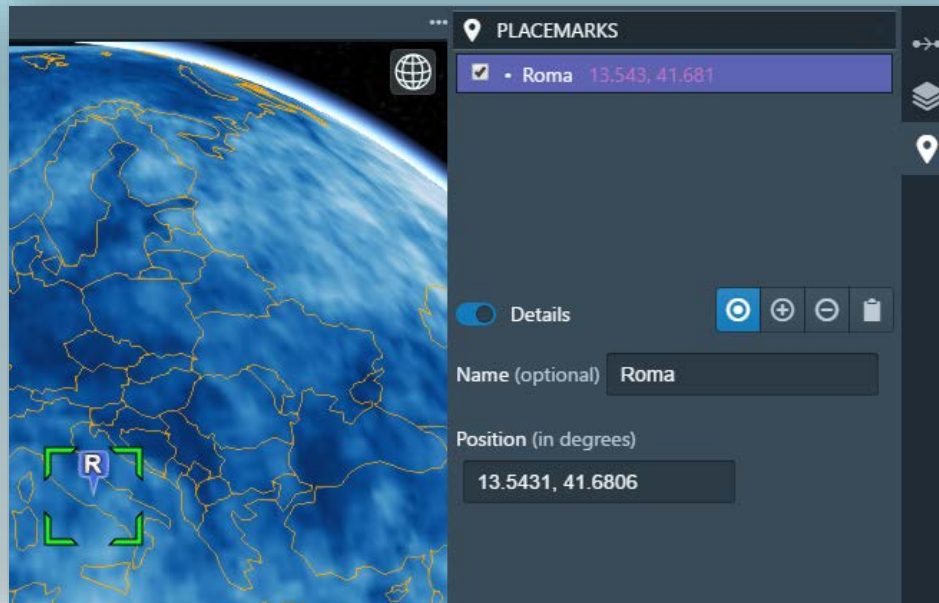
Ready. 1 failed task(s) lon=20.00, lat=61.89

- Create time series plot easily from **selected placemark**

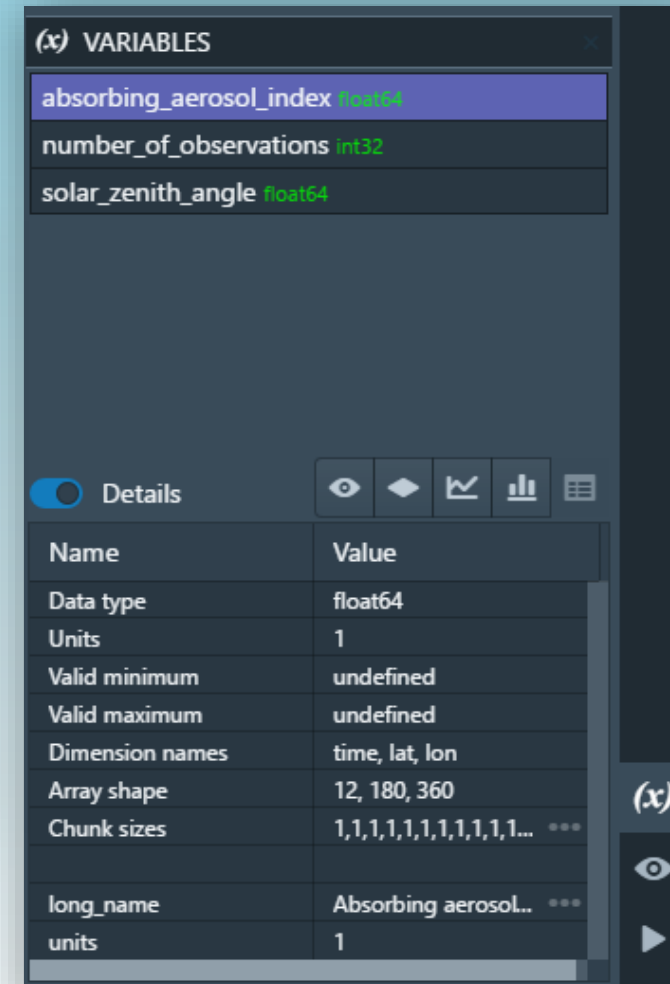


- and **selected variable**

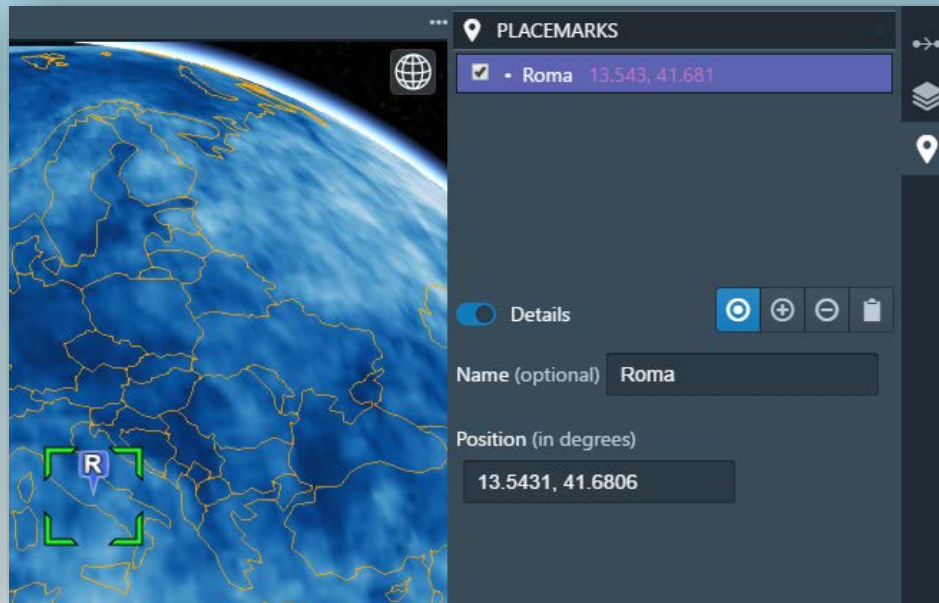
- Create time series plot easily from **selected placemark**



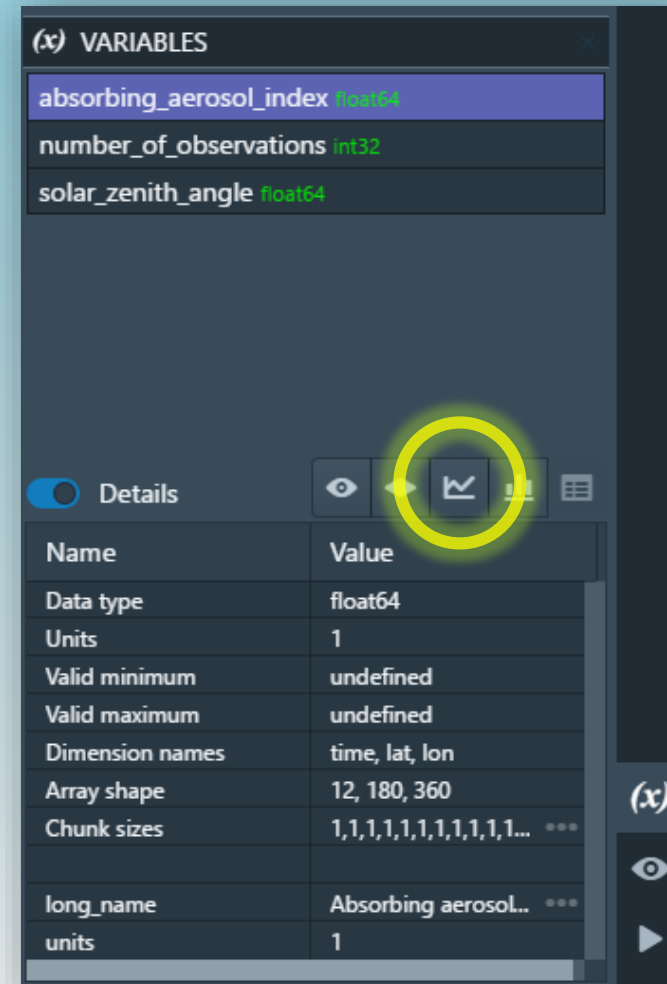
- and **selected variable**



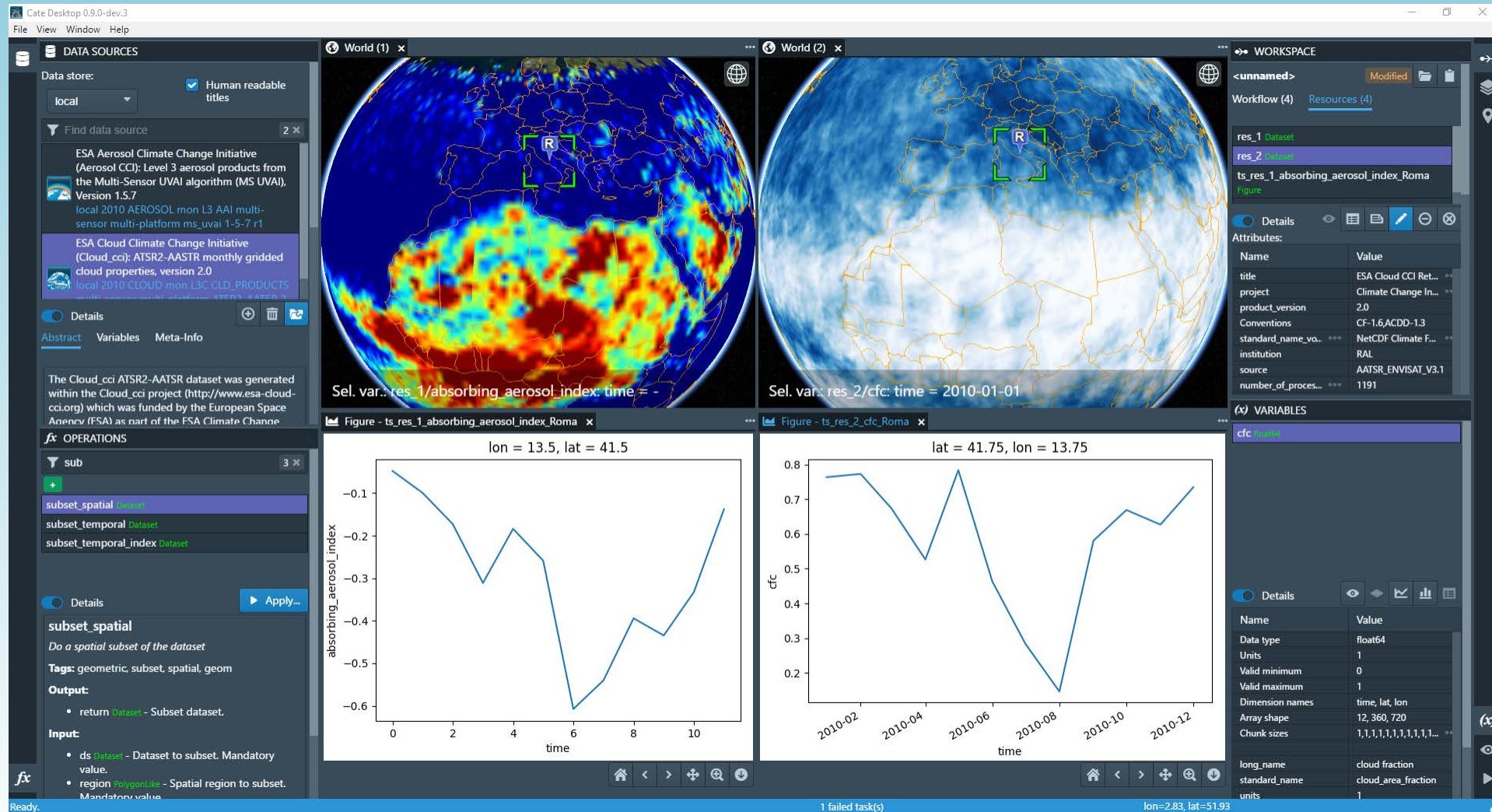
- Create time series plot easily from **selected placemark**



- and **selected variable**



Time Series



The screenshot displays the Cate Desktop 0.9.0-dev.3 interface. On the left, the 'DATA SOURCES' panel shows two datasets: 'ESA Aerosol Climate Change Initiative (Aerosol CCI): Level 3 aerosol products from the Multi-Sensor UVAI algorithm (MS UVAI), Version 1.5.7' and 'ESA Cloud Climate Change Initiative (Cloud_cci): ATSR2-AASTR monthly gridded cloud properties, version 2.0'. The 'OPERATIONS' panel shows a workflow with 'subset_spatial' and 'subset_temporal' operations. The 'WORKSPACE' panel on the right shows a workflow with resources 'res_1 Dataset', 'res_2 Dataset', and 'ts_res_1_absorbing_aerosol_index_Roma Figure'. Below the workspace, the 'ATTRIBUTES' and 'VARIABLES' panels provide metadata for the selected resources.

Two globe views are shown: 'World (1)' and 'World (2)'. The first globe shows the 'absorbing_aerosol_index' with a red box labeled 'R' over Rome. The second globe shows 'cfc' with a red box labeled 'R' over Rome. Below the globes are two time series plots:

- Plot 1: 'lon = 13.5, lat = 41.5'. The y-axis is 'absorbing_aerosol_index' ranging from -0.6 to -0.1. The x-axis is 'time' from 0 to 10. The plot shows a fluctuating line with a significant dip around time 6.
- Plot 2: 'lat = 41.75, lon = 13.75'. The y-axis is 'cfc' ranging from 0.2 to 0.8. The x-axis is 'time' from 2010-02 to 2010-12. The plot shows a fluctuating line with a peak around 2010-06 and a dip around 2010-08.

At the bottom, the status bar shows 'Ready.', '1 failed task(s)', and 'lon=2.83, lat=51.93'.

cci

→ **TOOLBOX**

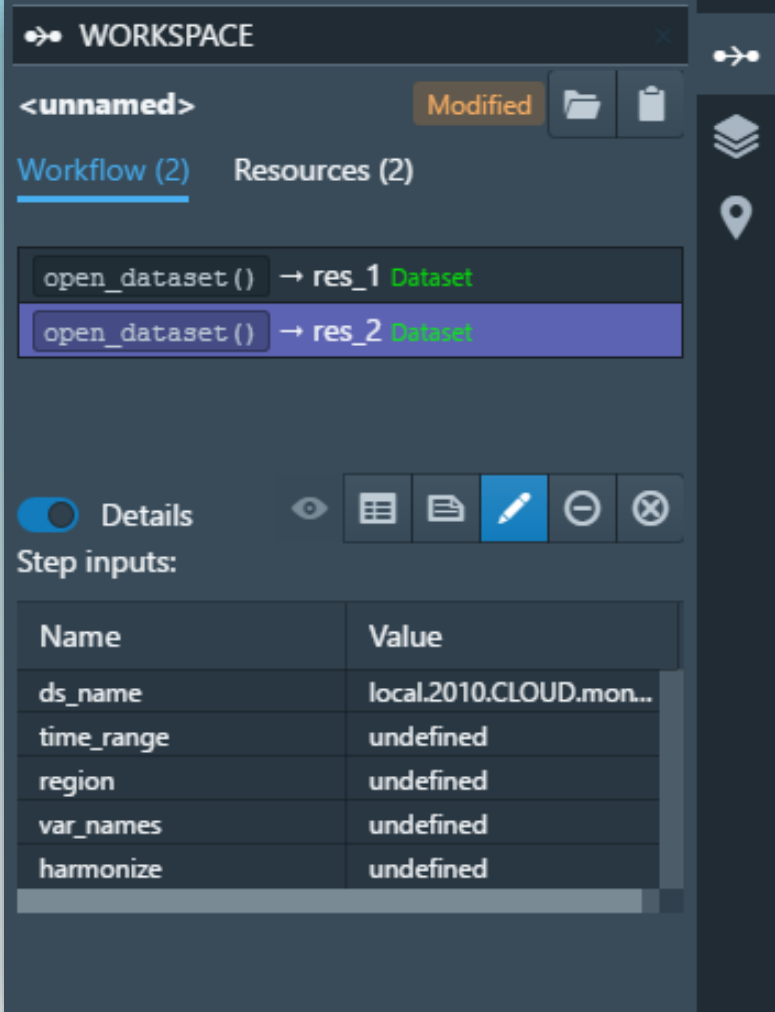
Cate Desktop (GUI)

Cate Programming (API)

Cate Command Line (CLI)

Workspace and Workflow

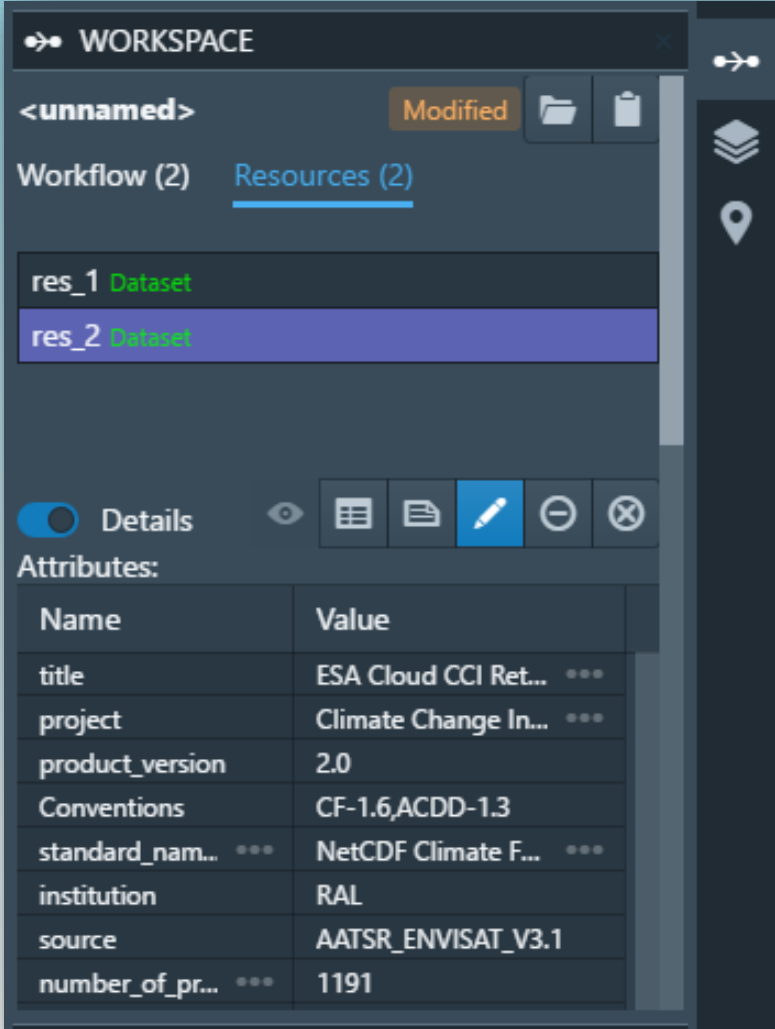
- Every step you perform, is a **step of a workflow** within your workspace



The screenshot displays the QGIS workspace interface. At the top, it shows 'WORKSPACE' and '<unnamed>' with a 'Modified' status. Below this, there are two tabs: 'Workflow (2)' and 'Resources (2)'. The 'Workflow (2)' tab is active, showing two steps: 'open_dataset () → res_1 Dataset' and 'open_dataset () → res_2 Dataset'. The second step is highlighted in blue. Below the workflow steps, there is a 'Details' section with a toggle switch and a table of 'Step inputs:'. The table has two columns: 'Name' and 'Value'. The inputs are: ds_name (local.2010.CLOUD.mon...), time_range (undefined), region (undefined), var_names (undefined), and harmonize (undefined).

Name	Value
ds_name	local.2010.CLOUD.mon...
time_range	undefined
region	undefined
var_names	undefined
harmonize	undefined

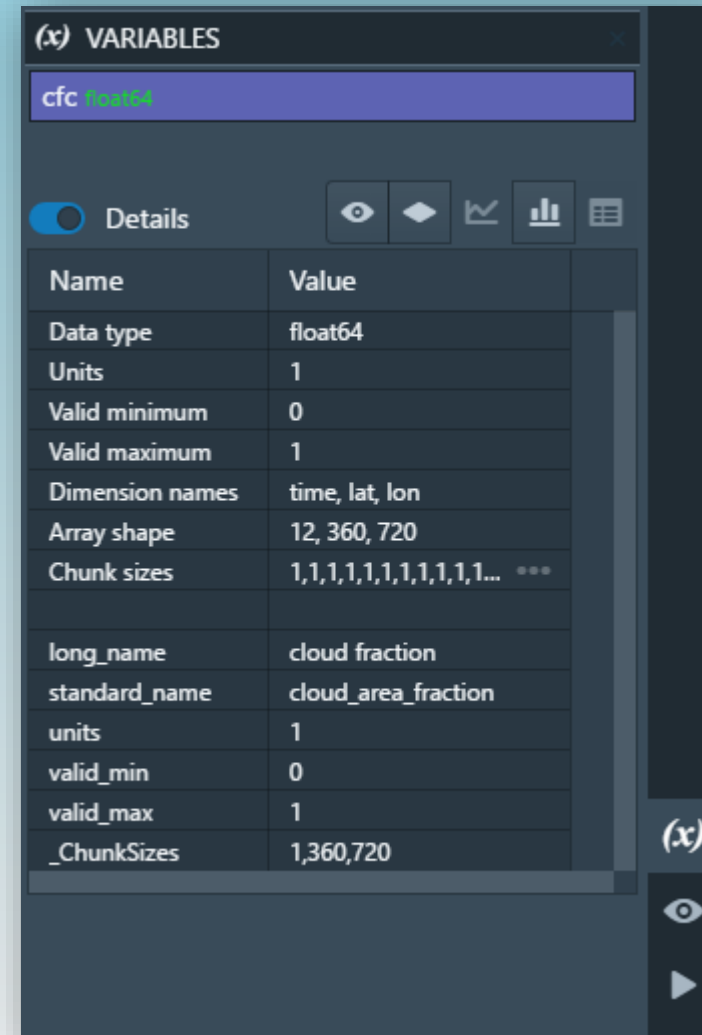
- Every step you perform, is a **step of a workflow** within your workspace
- The result of every step is a **new resource** which can be used as input to another step



The screenshot displays the 'WORKSPACE' interface. At the top, it shows '<unnamed>' with a 'Modified' status and icons for folder and trash. Below this, there are two tabs: 'Workflow (2)' and 'Resources (2)'. The 'Resources (2)' tab is active, showing a list of resources: 'res_1 Dataset' and 'res_2 Dataset'. The 'res_2 Dataset' is selected and highlighted in purple. Below the list, there is a 'Details' section with a toggle switch and several icons (eye, list, edit, zoom out, zoom in). The 'Attributes' table is visible below the details section.

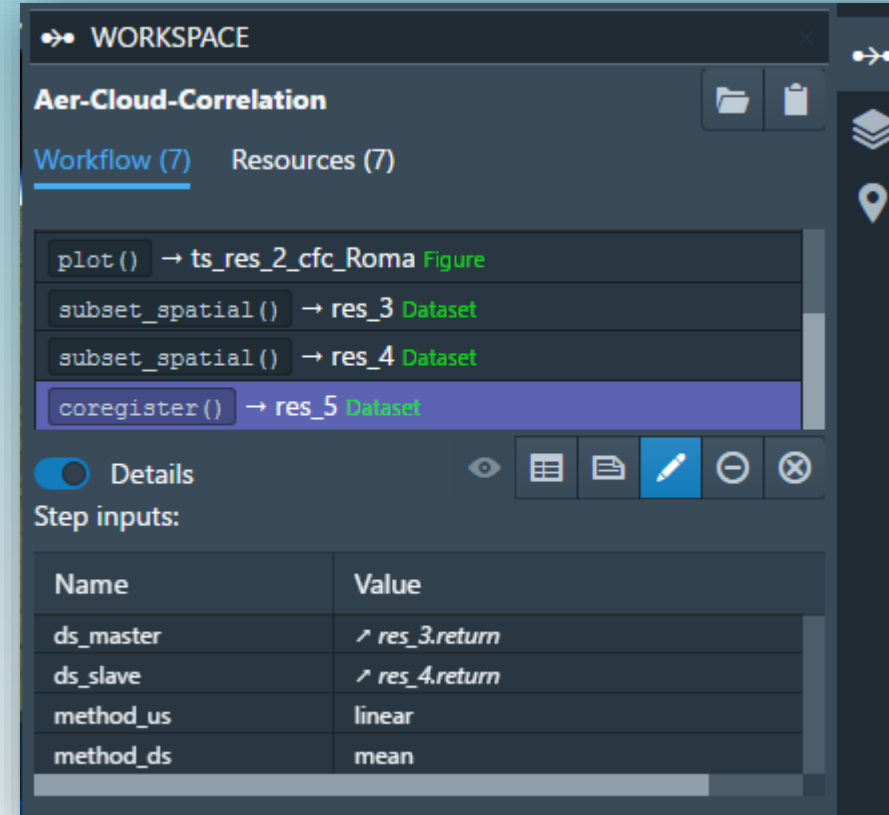
Name	Value
title	ESA Cloud CCI Ret... <small>...</small>
project	Climate Change In... <small>...</small>
product_version	2.0
Conventions	CF-1.6,ACDD-1.3
standard_nam... <small>...</small>	NetCDF Climate F... <small>...</small>
institution	RAL
source	AATSR_ENVISAT_V3.1
number_of_pr... <small>...</small>	1191

- Every step you perform, is a **step of a workflow** within your workspace
- The result of every step is a **new resource** which can be used as input to another step
- (NetCDF) dataset resources contain the actual **geo-physical ECVs**



Name	Value
Data type	float64
Units	1
Valid minimum	0
Valid maximum	1
Dimension names	time, lat, lon
Array shape	12, 360, 720
Chunk sizes	1,1,1,1,1,1,1,1,1,1... <small>...</small>
long_name	cloud fraction
standard_name	cloud_area_fraction
units	1
valid_min	0
valid_max	1
_ChunkSizes	1,360,720

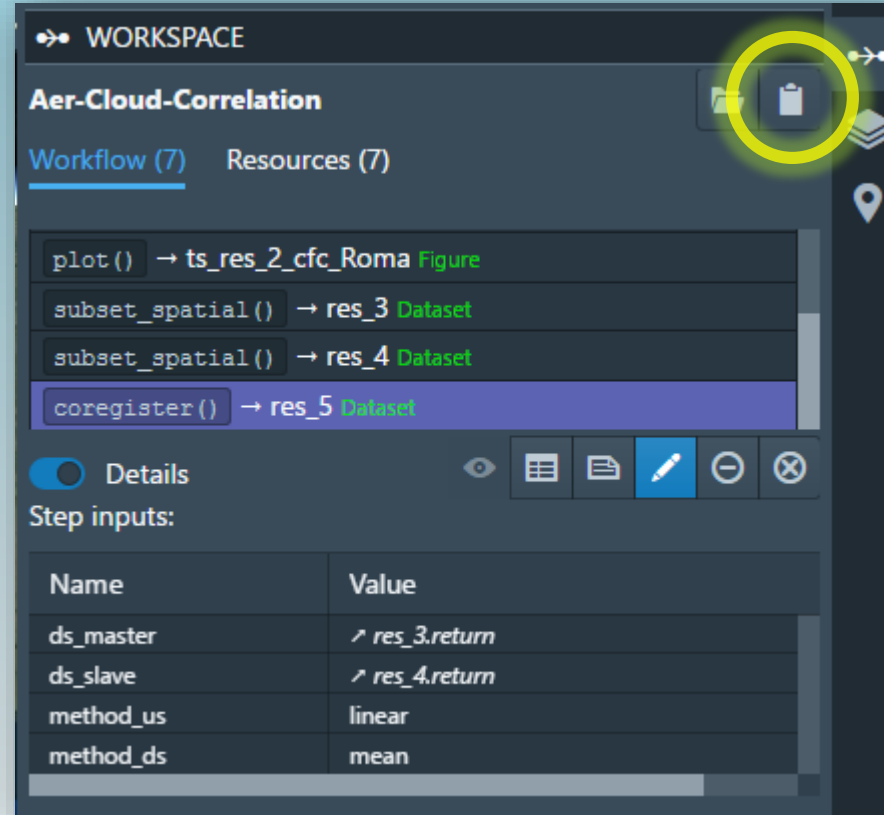
- Remember, every operation is a new **workflow step**



The screenshot shows the 'Aer-Cloud-Correlation' workspace in the ESA CCI Toolbox. The workflow consists of seven steps: 'plot()', 'subset_spatial()', 'subset_spatial()', 'coregister()', and three more 'subset_spatial()' steps. The 'coregister()' step is currently selected, and its details are shown below. The 'Step inputs' table is as follows:

Name	Value
ds_master	↗ res_3.return
ds_slave	↗ res_4.return
method_us	linear
method_ds	mean

- Remember, every operation is a new **workflow step**
- Same workflows can be executed from **Python** or from the **Command-Line Interface (CLI)**



WORKSPACE

Aer-Cloud-Correlation

Workflow (7) Resources (7)

plot() → ts_res_2_cfc_Roma Figure

subset_spatial() → res_3 Dataset

subset_spatial() → res_4 Dataset

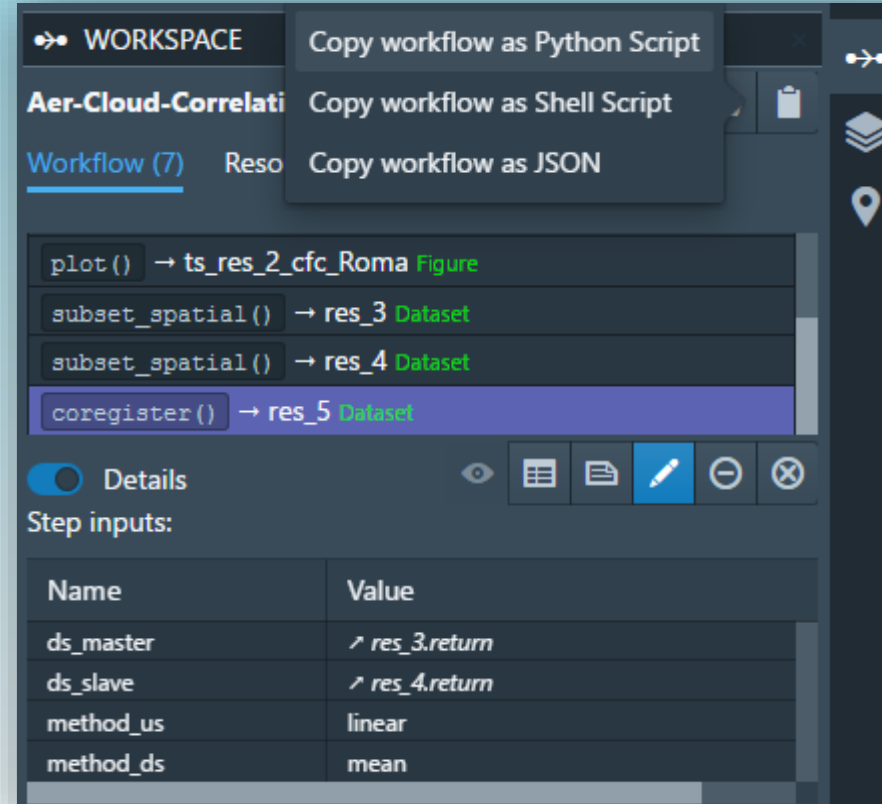
coregister() → res_5 Dataset

Details

Step inputs:

Name	Value
ds_master	↗ res_3.return
ds_slave	↗ res_4.return
method_us	linear
method_ds	mean

- Remember, every operation is a new **workflow step**
- Same workflows can be executed from **Python** or from the **Command-Line Interface (CLI)**



The screenshot shows the QGIS workspace with a workflow named "Aer-Cloud-Correlati" containing several steps. A context menu is open over the workspace, offering options to "Copy workflow as Python Script", "Copy workflow as Shell Script", and "Copy workflow as JSON". The workflow steps are:

- plot() → ts_res_2_cfc_Roma Figure
- subset_spatial() → res_3 Dataset
- subset_spatial() → res_4 Dataset
- coregister() → res_5 Dataset

The "coregister()" step is selected, and its details are shown in a panel below. The "Details" panel includes a "Step inputs:" table:

Name	Value
ds_master	↗ res_3.return
ds_slave	↗ res_4.return
method_us	linear
method_ds	mean

- Exported Python Code

```
1 from cate.ops import *
2
3 # Step 1
4 res_1 = open_dataset(ds_name="local.2010.AEROSOL.mon.L3.AAI.multi-sensor.
5
6 # Step 2
7 res_2 = open_dataset(ds_name="local.2010.CLOUD.mon.L3C.CLD_PRODUCTS.multi
8
9 # Step 3
10 ts_res_1_absorbing_aerosol_index_Roma = plot(ds=res_1, var="absorbing_aer
11
12 # Step 4
13 ts_res_2_cfc_Roma = plot(ds=res_2, var="cfc", indexers="lon=13.5431, lat=
14
15 # Step 5
16 res_3 = subset_spatial(ds=res_1, region="POLYGON ((30 20, 30 70, -20 70,
17
18 # Step 6
19 res_4 = subset_spatial(ds=res_2, region="POLYGON ((30 20, 30 70, -20 70,
20
21 # Step 7
22 res_5 = coregister(ds_master=res_3, ds_slave=res_4)
23
```

- Exported CLI Calls

```
1 #!/usr/bin/env bash
2 cate ws new
3 cate res set res_1 open_dataset ds_name="local.2010.AEROSOL.mon.L3.AAI.multi-sensor.
4 cate res set res_2 open_dataset ds_name="local.2010.CLOUD.mon.L3C.CLD_PRODUCTS.multi
5 cate res set ts_res_1_absorbing_aerosol_index_Roma plot ds=@res_1 var="absorbing_aer
6 cate res set ts_res_2_cfc_Roma plot ds=@res_2 var="cfc" indexers="lon=13.5431, lat=
7 cate res set res_3 subset_spatial ds=@res_1 region="POLYGON ((30 20, 30 70, -20 70,
8 cate res set res_4 subset_spatial ds=@res_2 region="POLYGON ((30 20, 30 70, -20 70,
9 cate res set res_5 coregister ds_master=@res_3 ds_slave=@res_4
10 cate ws exit
```

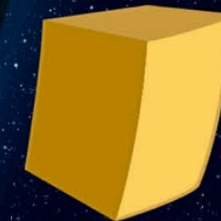
cci

→ **TOOLBOX**

Cate Desktop (GUI)

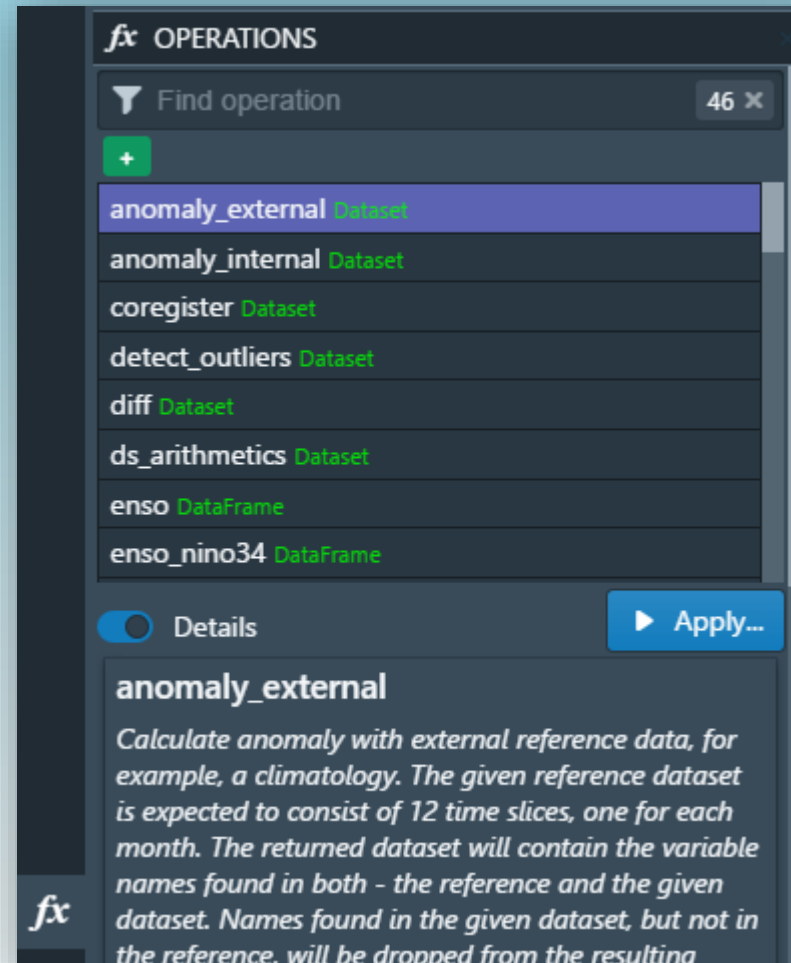
Cate Programming (API)

Cate Command Line (CLI)



Data Analysis

- CCI Toolbox comes with a **library of operations**



fx OPERATIONS

Find operation 46 x

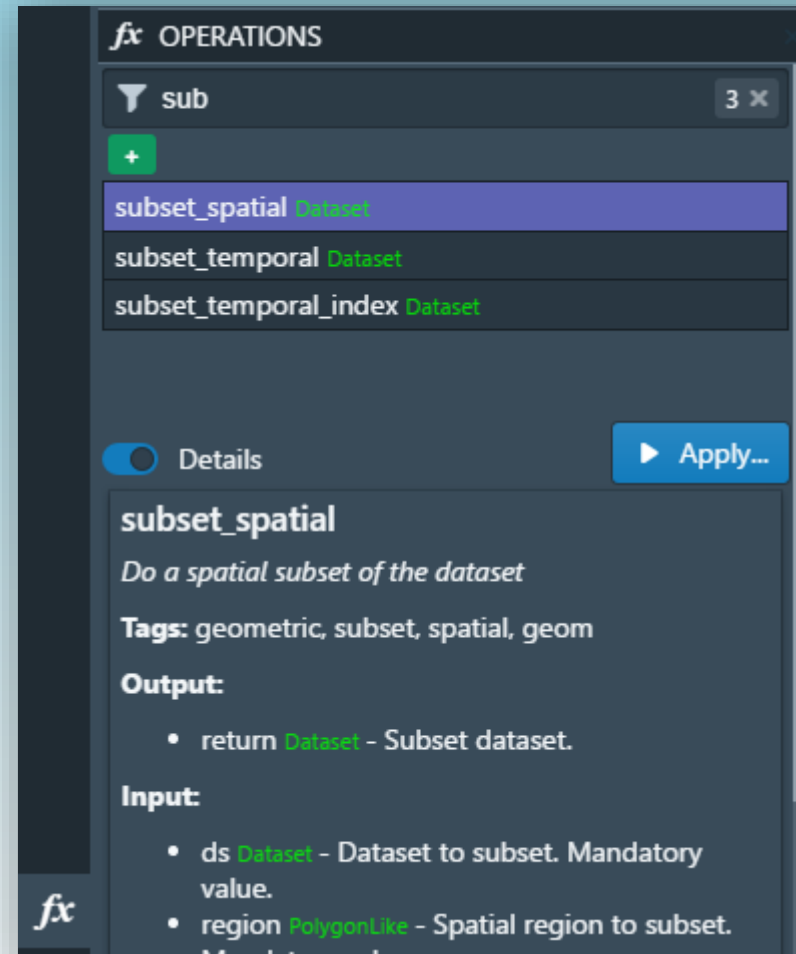
- + anomaly_external Dataset
- anomaly_internal Dataset
- coregister Dataset
- detect_outliers Dataset
- diff Dataset
- ds_arithmetics Dataset
- enso DataFrame
- enso_nino34 DataFrame

Details ▶ Apply...

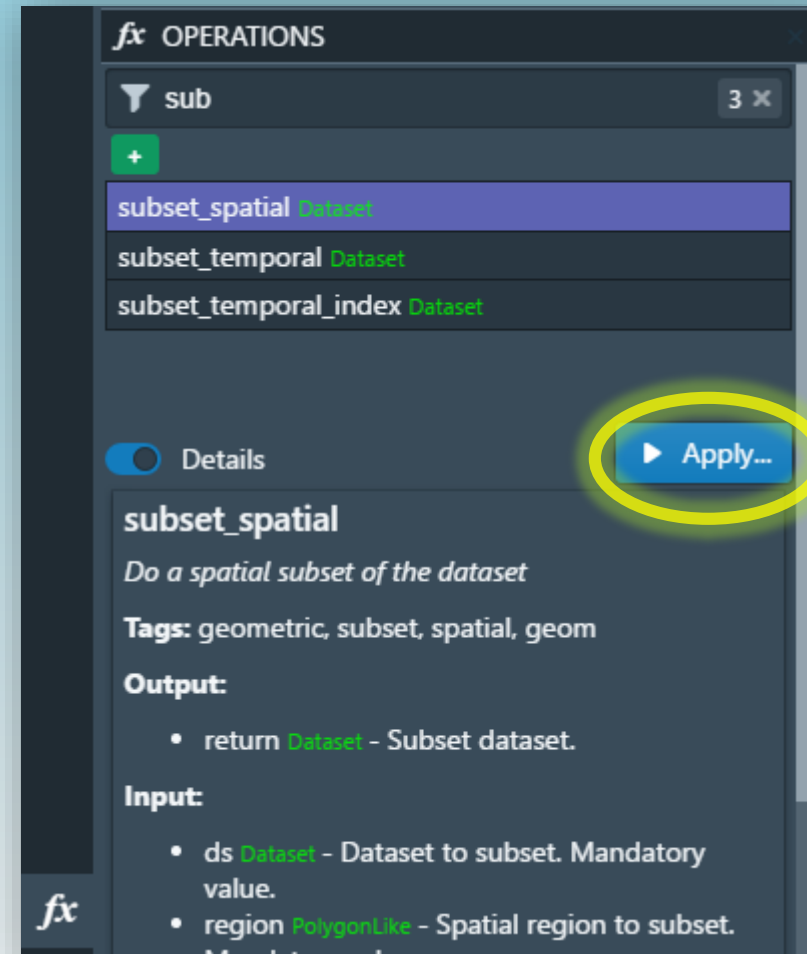
anomaly_external

Calculate anomaly with external reference data, for example, a climatology. The given reference dataset is expected to consist of 12 time slices, one for each month. The returned dataset will contain the variable names found in both - the reference and the given dataset. Names found in the given dataset, but not in the reference, will be dropped from the resultina

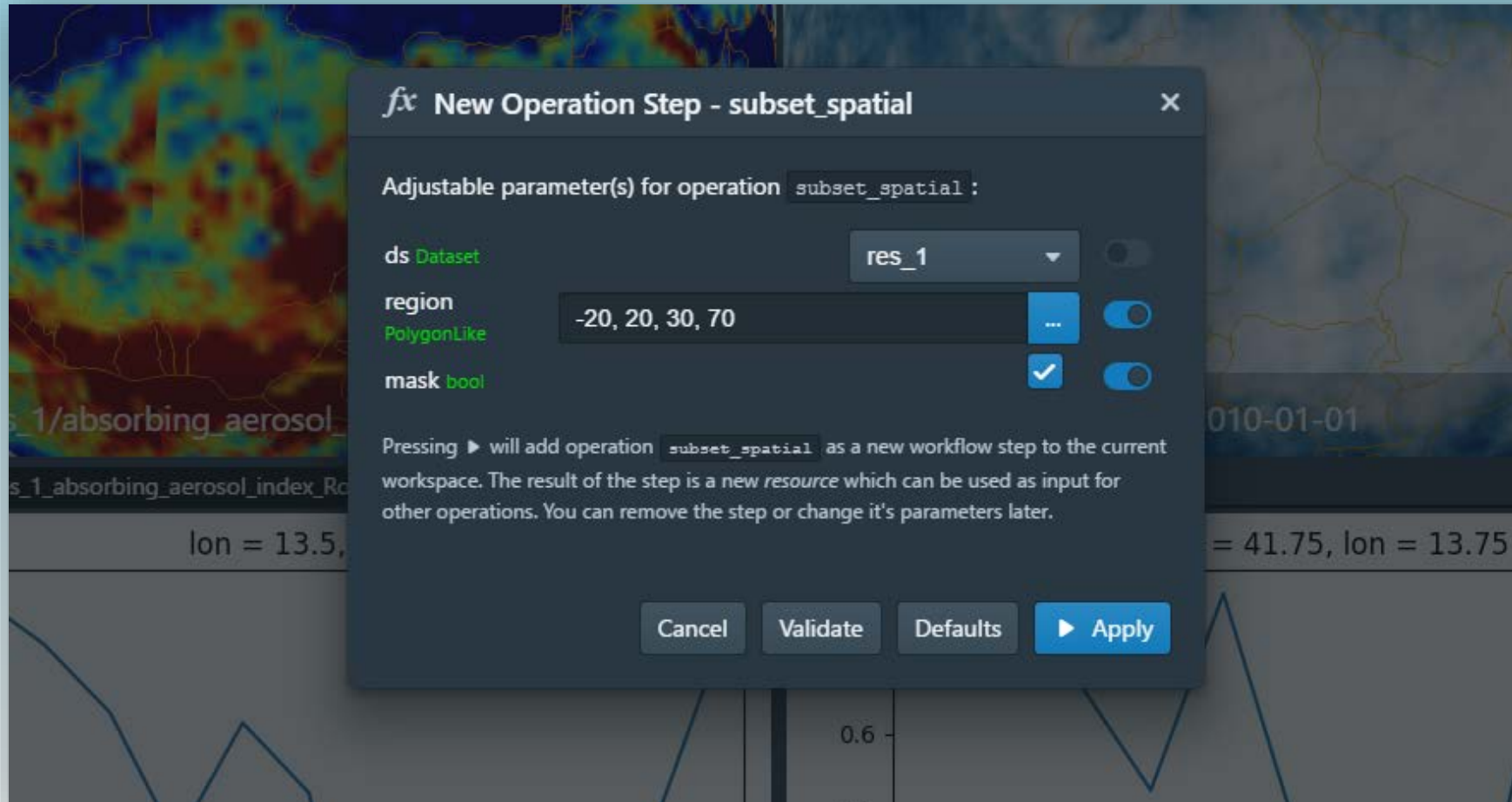
- CCI Toolbox comes with a **library of operations**
- Type to find operations, e.g. **subset**



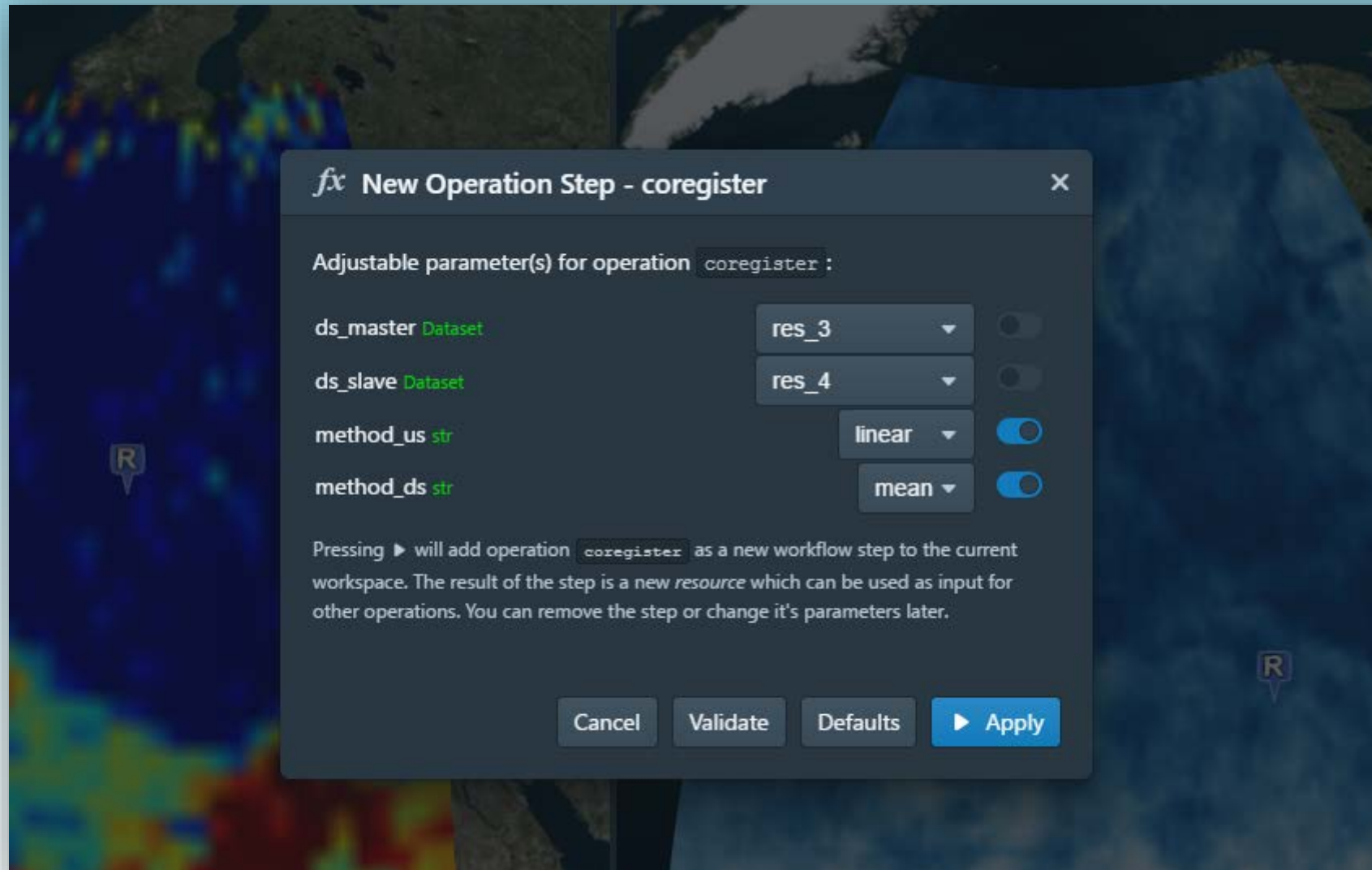
- CCI Toolbox comes with a **library of operations**
- Type to find operations, e.g. **subset**
- Operations are usual **Python functions** – it is very **easy to implement your own!**
- An operation's user interface is generated on-the-fly

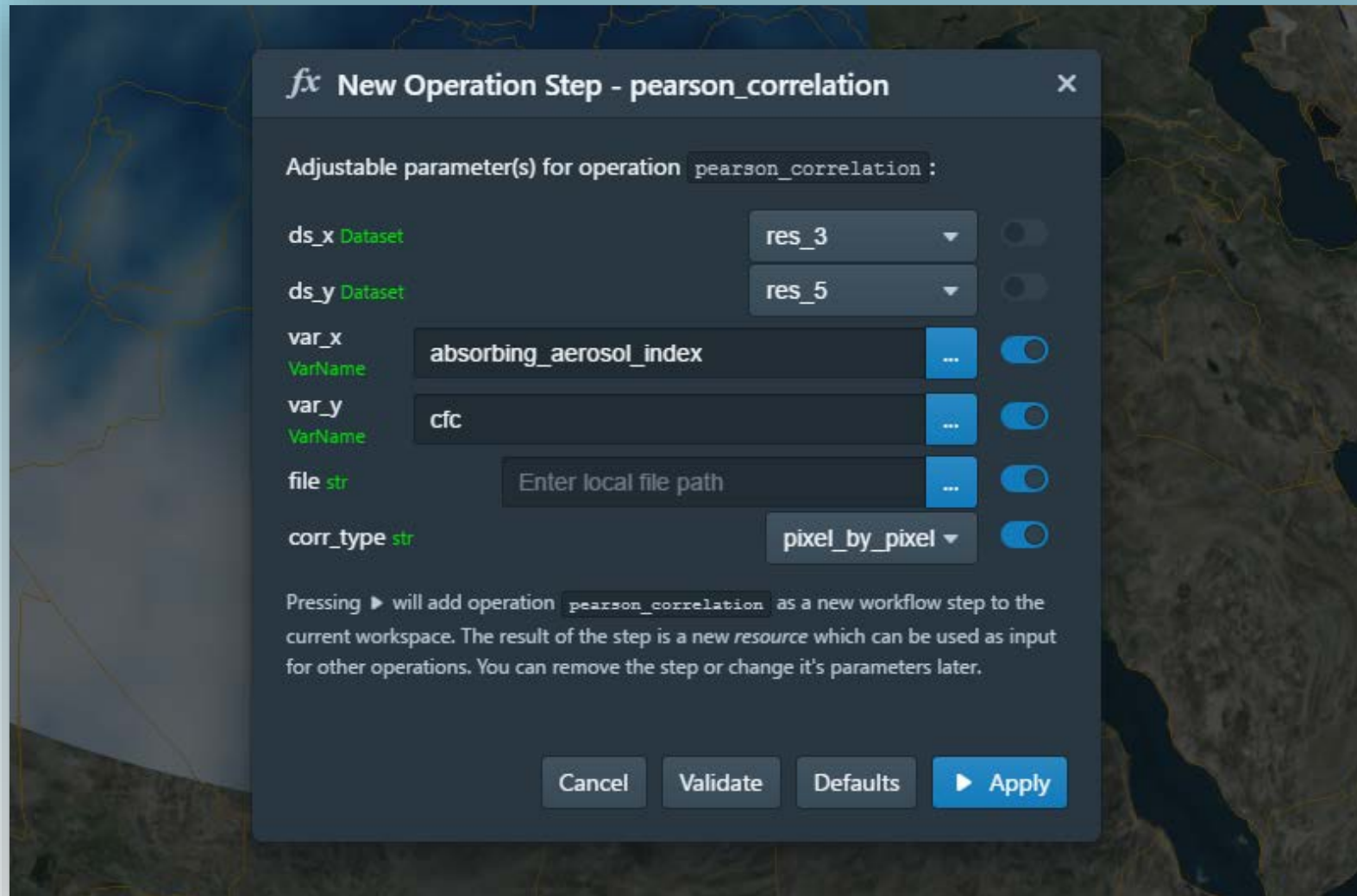


Spatial Subsetting

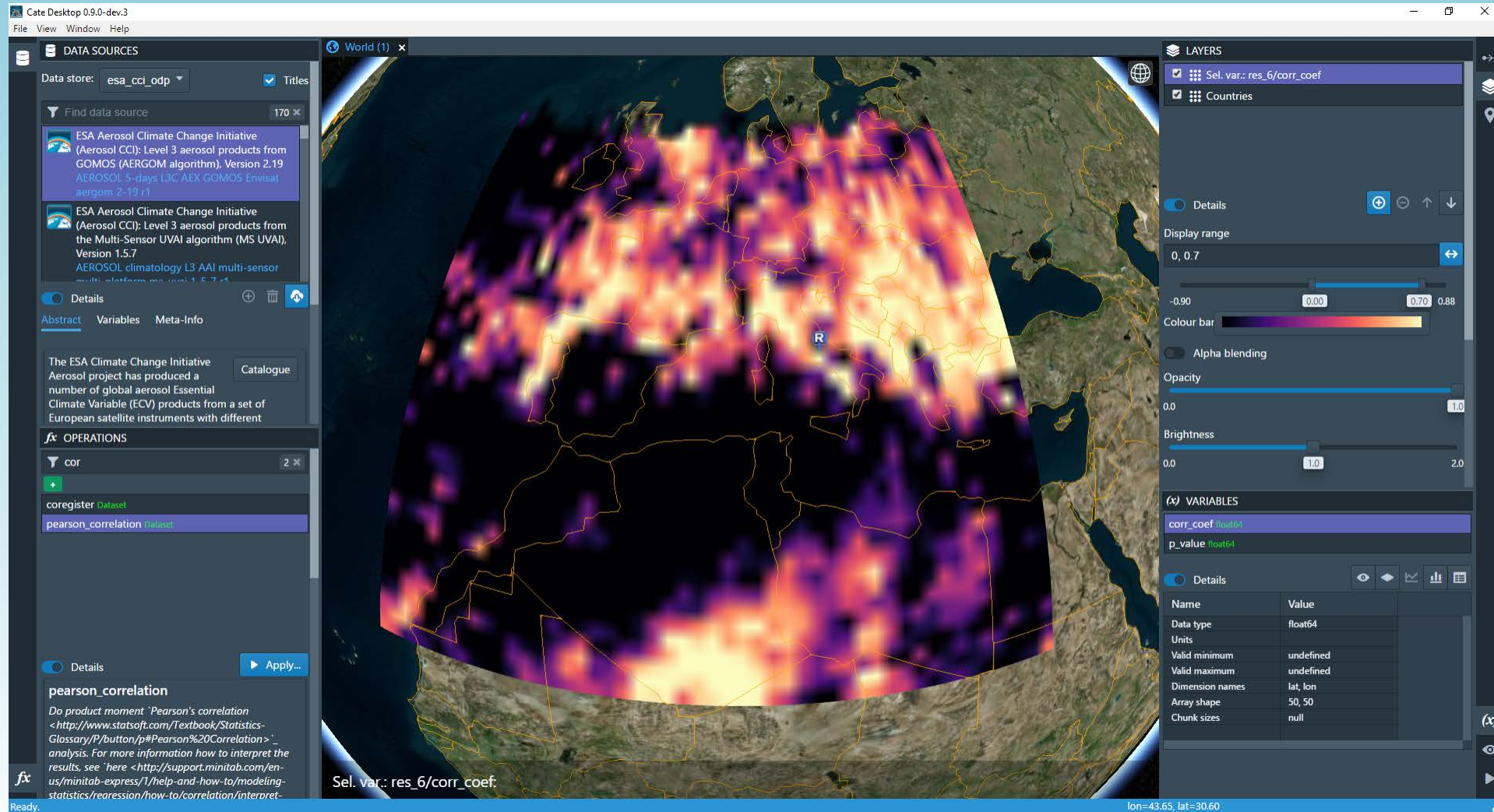


Dataset Co-registration

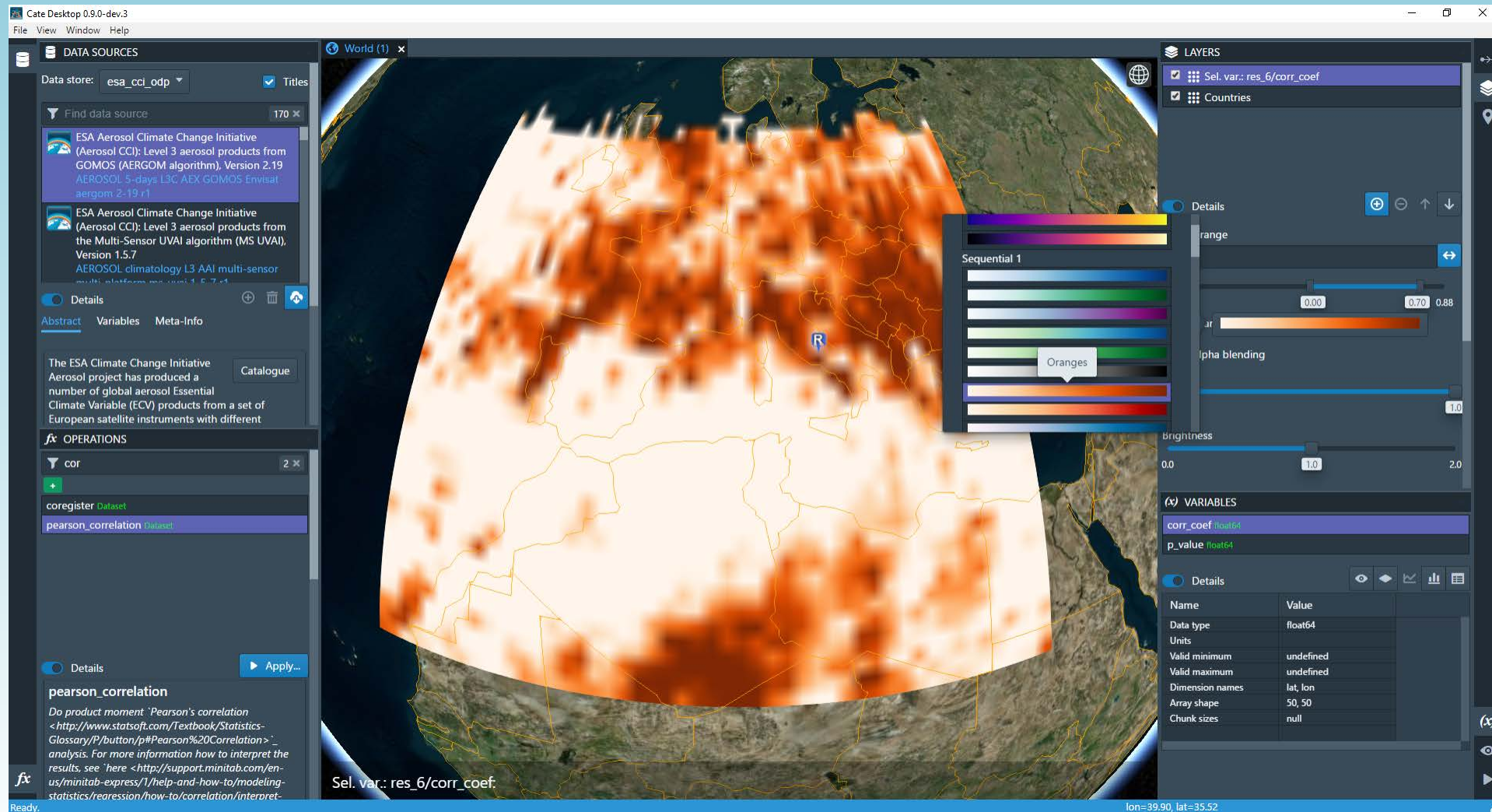




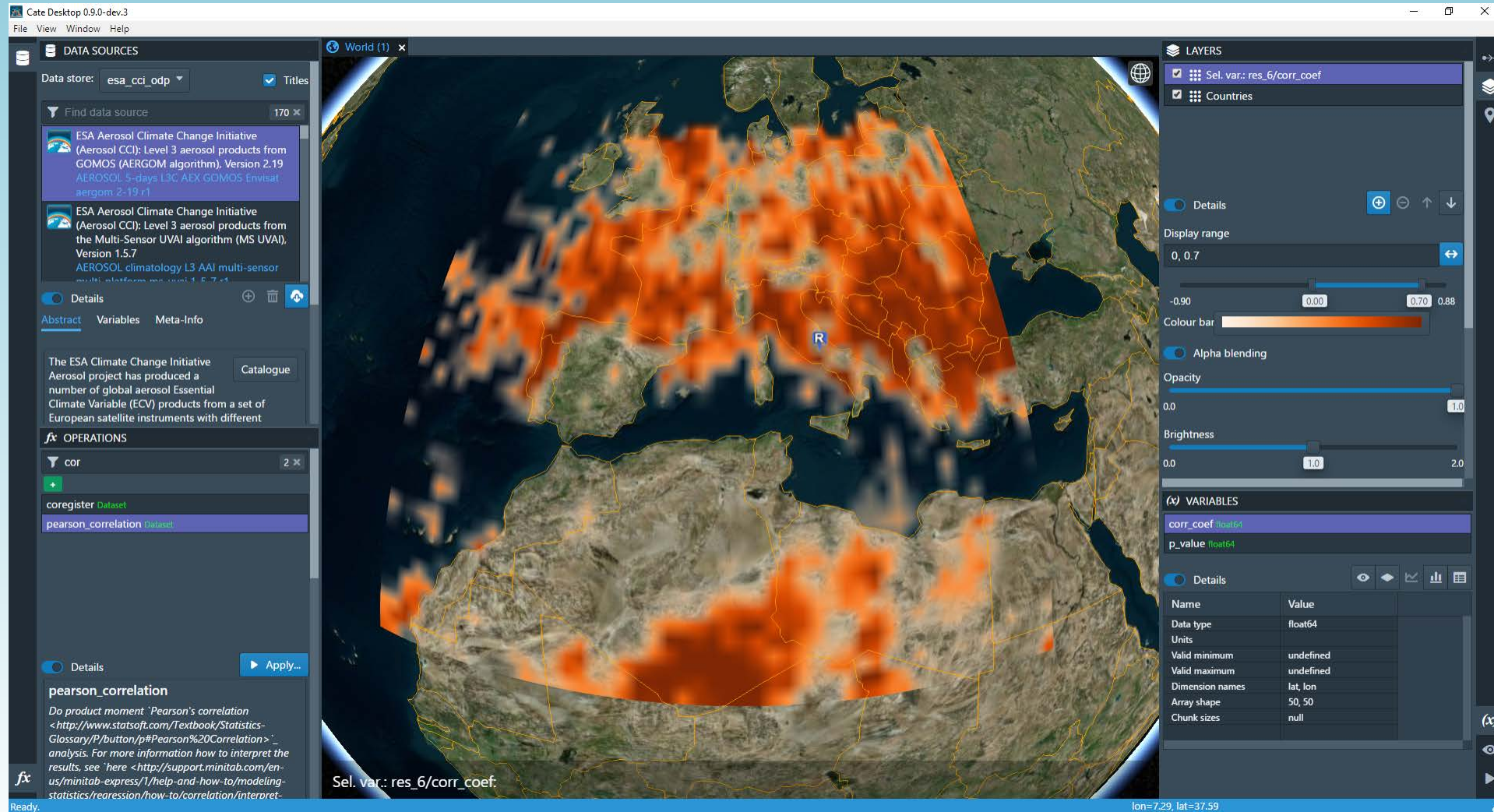
Aerosol/Cloud Correlation in 2010



Aerosol/Cloud Correlation in 2010



Aerosol/Cloud Correlation in 2010



Give it a Try!

- Releases (0.9.0.dev3)
 - <https://github.com/CCI-Tools/cate-core/releases>
- Source Code (MIT Open Source License)
 - <https://github.com/CCI-Tools/cate-core>
 - <https://github.com/CCI-Tools/cate-desktop> (GUI)
- Documentation
 - <http://ect-core.readthedocs.io/en/latest/>