

Climate4Impact v2

Enhance the use of climate research data and methods

climate4impact.eu

Alessandro Spinuso, Mats Veldhuizen,
Ian van der Neut (KNMI),
Christian Pagé (CERFACS)



Climate4Impact v2



Online platform that combines discovery of climate research data of the Earth System Grid Federation (ESGF), with experimentation workspaces, offering reproducible and sharable environments for climate-impact analysis.



DAVID McNEWMAN/AGENCY IMAGES

Climate4Impact v2



Online platform that combines discovery of climate research data of the Earth System Grid Federation (ESGF), with experimentation workspaces, offering reproducible and sharable environments for climate-impact analysis.

Delivered in the context of the cooperation between IS-ENES and ESGF, which co-develop tailored tooling, infrastructure services and trainings



Climate4Impact v2



Online platform that combines discovery of climate research data of the Earth System Grid Federation (ESGF), with experimentation workspaces, offering reproducible and sharable environments for climate-impact analysis.

Delivered in the context of the cooperation between IS-ENES and ESGF, which co-develop tailored tooling, infrastructure services and trainings

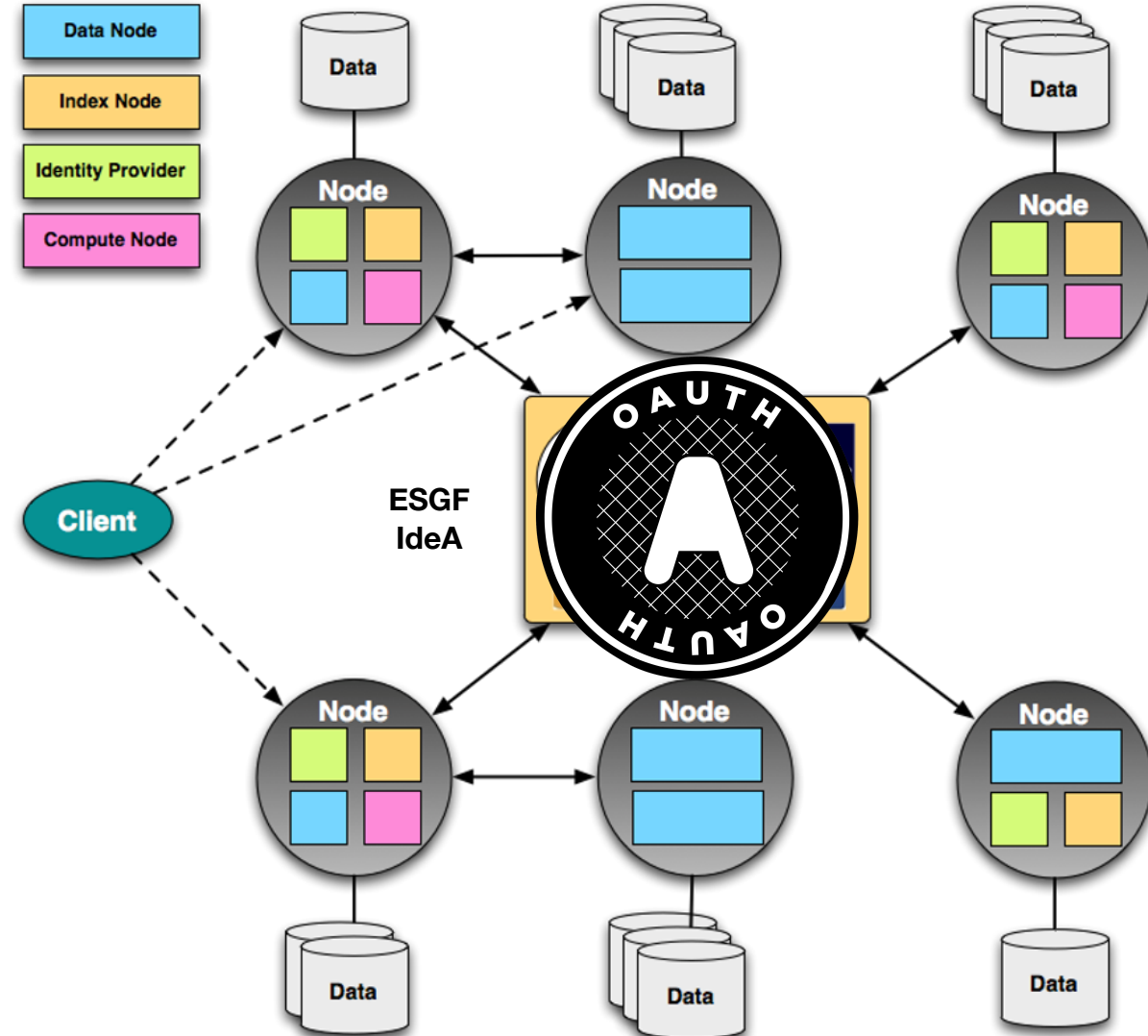
Aimed at supporting research on predicting and mitigating climate change effects on the whole Earth ecosystem





ESGF is a **multinational** effort that manages climate and environmental data.
A network of federated and **standardised data discovery, access and analysis services**.
Data Volume, in the order of tens of PB, is constantly increasing and distributed across different projects, e.g.

- CMIP5-6.. 7 coming up
- CORDEX
- Obs4MIPs

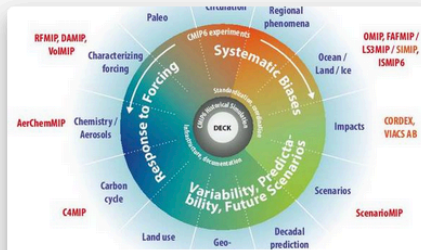




Climate4Impact v2 - Projects

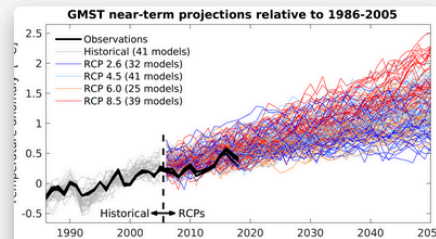
Select a project

You can search the Earth System Grid Federation for various kinds of climate data. Please select the project you are looking for.



CMIP6

Coupled model
intercomparison experiment
6.



CMIP5

Coupled model
intercomparison experiment
5.



CORDEX

Global



Climate4Impact v2 - Profiles and SSO



Guest Limited features	Registered User Full feature access
<ul style="list-style-type: none"><input checked="" type="checkbox"/> Browse ESGF Data<input checked="" type="checkbox"/> Download file links	<ul style="list-style-type: none"><input checked="" type="checkbox"/> Browse ESGF Data<input checked="" type="checkbox"/> Download file links<input checked="" type="checkbox"/> Access data in Jupyter workspace<input checked="" type="checkbox"/> Access scientific notebook presets<input checked="" type="checkbox"/> Subset large datasets<input checked="" type="checkbox"/> Create Binder snapshots of your work<input checked="" type="checkbox"/> Trace and rollback operations
BROWSE DATA	REGISTER

ESGF
Earth System Grid Federation

Login

Select your institution ▾

I don't have an account

or

Sign in with one of these identity providers

NEW

Two User Profiles

Guest: Basic Search & Local Download

Registered User: Data Analysis & Subsetting
(soon CORDEX support)


Registration procedure

- 1 - Send an email to c4i@knmi.nl specifying affiliation and motivation to become a Registered User
- 2 - Make sure the email is associated with at least one of your GitHub, ORCID, ESGF-CEDA accounts
- 3 - Once approved, login to C4I via the ESGF SSO

C4I v2 - Search and Nodes

Search Parametrisation made easier





Exploring climate model data

Home Data Discovery Help Feedback Sign Up
PROJECT: CMIP 6
NODES: ALL
SCOPED VIEW

VARIABLE

FREQUENCY

EXPERIMENT

MODEL

MEMBER

variable:tasmx
frequency:day
experiment_id:ssp585
source_id:EC-Earth3
member_id:r1i1p1f1
member_id:r2i1p1f1

Temperature

- ta - Air temperature (9)
- tas - Temperature (6)
- tasmin - Min. Temperature (6)
- tasmx - Max. Temperature (6)

Precipitation

- pr - Precipitation (6)
- prsn - Snow (4)
- prc - Convective precipitation (1)

Humidity

- hurs - Rel. Humidity (6)
- huss - Specific humidity (5)
- rhsm - Min, Rel. Humidity (-)
- rhs - Rel. Humidity (-)
- hus - Spec. Humidity (7)
- hur - Rel. Humidity (4)

Wind

- uas - Eastward wind (6)
- vas - Northward wind (6)
- sfcWind - Wind (5)
- sfcWindmax - Max Wind (4)

Radiation

- rsds - SW Radiation Dn (5)
- rlus - LW Radiation Up (4)
- rsus - SW Radiation Up (2)
- rlds - LW Radiation Dn (2)
- rdsd - Diff. Radiation (-)
- clt - Cloud (2)

Pressure

- ps - Pressure (-)
- psl - Sea level pressure (6)
- pfull - Pressure (-)

Evaporation

- evspsbl - Act. Evap. (-)
- evspsblpot - Pot. Evap. (-)
- evspsblsoi - Sol Evap. (-)
- evspsblveg - Canopy Evap. (-)

C4I v2 - Search and Nodes

Search Parametrisation made easier



Data Nodes Selection by Service

is-enes Exploring climate model data

Home Data Discovery Help Feedback Sign Up

PROJECT: CMIP 6 NODES: ALL

variable:tasmax frequency:day experiment_id:ssp585 source_id:EC-Earth3 member_id:r1i1p1f1

Temperature

- ta - Air temperature (9)
- tas - Temperature (6)
- tasmin - Min. Temperature (6)
- tasmax - Max. Temperature (6)

Precipitation

- pr - Precipitation
- prsn - Snow
- prc - Convective precipitation

Radiation

- rsds - SW Radiation Dn (5)
- rlus - LW Radiation Up (4)
- rsus - SW Radiation Up (2)
- rlds - LW Radiation Dn (2)
- rsdsdiff - Diff. Radiation (-)
- clt - Cloud (2)

is-enes Exploring climate model data

Home Data Discovery Help Feedback Register

PROJECT: CMIP 6 NODES & SUBSETTING

variable:tasmin frequency:day experiment_id:ssp585

Available ESGF Nodes


Select & enable Rook WPS subsetting

Node	Subsetting Mode
<input checked="" type="checkbox"/> esgf1.dkrz.de	Rook WPS
<input checked="" type="checkbox"/> esgf3.dkrz.de	Rook WPS
<input type="checkbox"/> aims3.lnl.gov	Not Available
<input type="checkbox"/> cmip.dess.tsinghua.edu.cn	Not Available
<input type="checkbox"/> cmip.fio.org.cn	Not Available
<input type="checkbox"/> cordexesg.dmi.dk	Not Available
<input type="checkbox"/> crd-esgf-drc.ec.gc.ca	Not Available
<input type="checkbox"/> data.meteo.unican.es	Not Available
<input type="checkbox"/> dataserver.nccs.nasa.gov	Not Available
<input type="checkbox"/> dpegf03.nccs.nasa.gov	Not Available

OK

ESMValTool in C4I Models' Performance Comparison





Exploring climate model data

Home
Data Discovery
Help
Feedback
Sign Up

PROJECT: CMIP 6
NODES: CUSTOM (2)

variable:ta

variable:prsn

variable:huss

frequency:day

experiment_id:ssp585

source_id:CanESM5

Model

- CanESM5 - CanESM5
- MPI-ESM1-2-LR - MPI-ESM1.2-LR
- UKESM1-0-LL - UKESM1.0-N96ORCA1
- CNRM-CM6-1 - CNRM-CM6-1
- MIROC6 - MIROC6
- CNRM-ESM2-1 - CNRM-ESM2-1
- MRI-ESM2-0 - MRI-ESM2.0
- GFDL-CM4 - GFDL-CM4
- HadGEM3-GC31-LL - HadGEM3-GC3.1-N96ORCA1
- MPI-ESM1-2-HR - MPI-ESM1.2-HR
- INM-CM4-8 - INM-CM4-8
- INM-CM5-0 - INM-CM5-0
- ACCESS-CM2 - Australian Community Climate and Earth System Simulator Climate Model
- AWI-CM-1-1-MR - AWI-CM 1.1 MR
- NorESM2-LM - NorESM2-LM (low atmosphere-medium ocean resolution, GHG concentrati
- BCC-CSM2-MR - BCC-CSM 2 MR
- CMCC-CM2-SR5 - CMCC-CM2-SR5
- FGOALS-G2.3 - FGOALS-G2.3

COMPARE MODEL PERFORMANCE

Model

Climate impact result viewer

This applications shows a results from CMIP5 and CMIP6 models, calculated with ESMValTool. It is intended to provide some guidance for climate impact researchers, to select one or more datasets that adequately sample the spread of the CMIP ensemble.

- Bias is calcated with respect to the ERA5 reanalysis dataset over the period 1981-2015.
- Future change is calculated for 2036-2065 as compared to 1986-2015.
- Area is set to Europe (lon 0-39; lat 30-76.25)
- All data are taken from the RCP/SSP 8.5 scenario

Hold ctrl to pan and zoom, hold alt to select a range (points will be highlighted in both graphs), then hold shift to select multiple points.

Temperature (K)

Precipitation (mm/day)

Project_project CMIP6

Selected datasets:

- CNRM-CM6-1
- MIROC6
- CanESM5

Netherlands eScience
Center

View static recipe output

C4I v2 Local Download



Climate4Impact Search for CMIP5/6
(soon CORDEX) Data
<https://climate4impact.eu>

Local Download via Metalink
<https://en.wikipedia.org/wiki/Metalink>

Search results

Dataset: 1 Single Files: 1

project:CMIP6 variable:tas variable:prc frequency:day frequency:mon experiment_id:ssp585 source_id:EC-Earth3 member_id:r11p1f1

ALL	CLEAR	Dataset ↑	Mirror	Version	Files	Size	DOI
<input type="checkbox"/>	^	CMIP6.ScenarioMIP.EC-Earth-Consortium.EC-Earth3.ssp585.r11p1f1.Amon.prc.gr	esgf-dn2.nsl.llu.se	2020/03/10	86	439.3 MB	
Select		File				Size	
<input checked="" type="checkbox"/>		prc_Amon_EC-Earth3_ssp585_r11p1f1_gr_201501-201512.nc				5.07 MB	
<input type="checkbox"/>		prc_Amon_EC-Earth3_ssp585_r11p1f1_gr_201601-201612.nc				5.02 MB	
<input type="checkbox"/>		prc_Amon_EC-Earth3_ssp585_r11p1f1_gr_201701-201712.nc				5.01 MB	
<input type="checkbox"/>		prc_Amon_EC-Earth3_ssp585_r11p1f1_gr_201801-201812.nc				5.05 MB	
<input type="checkbox"/>		prc_Amon_EC-Earth3_ssp585_r11p1f1_gr_201901-201912.nc				5.05 MB	
<input type="checkbox"/>		prc_Amon_EC-Earth3_ssp585_r11p1f1_gr_202001-202012.nc				4.99 MB	
<input type="checkbox"/>		prc_Amon_EC-Earth3_ssp585_r11p1f1_gr_202101-202112.nc				5.03 MB	
<input type="checkbox"/>		prc_Amon_EC-Earth3_ssp585_r11p1f1_gr_202201-202212.nc				5.02 MB	
<input type="checkbox"/>		prc_Amon_EC-Earth3_ssp585_r11p1f1_gr_202301-202312.nc				5.04 MB	
<input type="checkbox"/>		prc_Amon_EC-Earth3_ssp585_r11p1f1_gr_202401-202412.nc				5.06 MB	
Rows per page: 10 - 1-10 of 86 < >							
<input checked="" type="checkbox"/>	▼	CMIP6.ScenarioMIP.EC-Earth-Consortium.EC-Earth3.ssp585.r11p1f1.Amon.prc.gr	esgf-data1.llnl.gov	2020/03/10	86	439.3 MB	
<input type="checkbox"/>	▼	CMIP6.ScenarioMIP.EC-Earth-Consortium.EC-Earth3.ssp585.r11p1f1.Amon.prc.gr	esgf.nci.org.au	2020/03/10	1	5.18 MB	



```

1 <metalink xmlns="urn:ietf:params:xml:ns:metalink">
2   <file name="tasmin_day_EC-Earth3_ssp585_r11p1f1_gr_201501-20151231.nc">
3     <size>140066295</size>
4     <hash type="SHA256">6c54d27ecc06771fd0f593bba68df5cde31e1440581499142952088ae126213e</hash>
5     <url priority="1">http://esgf3.dkrz.de/thredds/fileServer/cmip6/ScenarioMIP/EC-Earth-Consortium/EC-E
6   </file>
7
8   <file name="tasmin_day_EC-Earth3_ssp585_r11p1f1_gr_201601-20161231.nc">
9     <size>140536230</size>
10    <hash type="SHA256">eb19230d74844f4b122ef144a51b8805c569f1f66e934824c8659bdbc09c0a2d</hash>
11    <url priority="1">http://esgf3.dkrz.de/thredds/fileServer/cmip6/ScenarioMIP/EC-Earth-Consortium/EC-E
12  </file>
13
14  <file name="tasmin_day_EC-Earth3_ssp585_r11p1f1_gr_201701-20171231.nc">
15    <size>140142064</size>
16    <hash type="SHA256">46420a196bd05aea989966bb49137ad0be93c784fef9aa3f64b815eb31984ed9</hash>
17    <url priority="1">http://esgf3.dkrz.de/thredds/fileServer/cmip6/ScenarioMIP/EC-Earth-Consortium/EC-E
18  </file>
19
20 </metalink>

```

- Scriptable (WGET/cURL)

C4I v2 Workflows and Workspaces



*Only for registered and validated accounts!
Cost-effective and secure*

Full Data Staging...

...Subsetting onto personal workspaces

Search results

Datasets: 0 Additional Files: 6 project:CMIP6 variable:tasmx

Dataset	Mirror (Node)	Version	Files	Size	Doc
CMIP6.ScenarioMIPDKRZ.MPI-ESM1-2-HR.ssp585.r11p1f1.day.tasmx.gn	esg3.dkzr.de	2019/07/10	18	3.77 GB	

Resume Notebook

Link collection ready (limited to 500 datasets). Found 0 datasets containing 0 file links and 6 single file(s).

Existing notebook found (<https://swirl.climae4impact.eu/swirl/jupyter/38a6f3fc-b689-4e5a-9fee-cea47e067630?token=4a4efb9798d8c50512bb2da76ea5a7ad6f5f7c58168d6e>) current data usage: 6.42 GB.

You have selected 1.32 GB of data. Some ESGF nodes perform worse than others and even the more performant ones can experience downtime occasionally. This means that your workflow can time out. If this occurs, we advise you to split your workflow up into smaller parts or to try a different node.

&

Select Parameters

Some ESGF nodes might not support subsetting. If you face issues with this functionality and no data becomes available, you can either try with a different node (mirror) or use the "Download" button to download the full data locally (via the "Link List") or to a C4I notebook.

Operation Selection
Operation to be used for rook subsetting
 Average No Op

Spatial Parameters
Coordinates of the bounding box in lat / lon format

Min. Latitude* Max. Latitude*
 Min. Longitude* Max. Longitude*

Temporal Parameters
Full years or specific dates in advanced mode

Advanced

Start year* End year*

Workspaces offer very large storage

The remote subsetting option, where available, helps reducing the requested data based on user's needs

Full Data staging allows 3GB per request
(prevents timeout, given the occasional high load on the nodes)

C4I v2 Workflows and Workspaces

Only for registered and validated accounts!
Cost-effective and secure

Full Data Staging...

The screenshot shows a web interface with a table of datasets. A modal dialog titled "Resume Notebook" is open, displaying a message: "Link collection ready (limited to 500 datasets). Found 0 datasets containing 0 file links and 6 single file(s)." Below this, it says "Existing notebook found (https://swirl.climate4impact.eu/swirl/jupyter/38a6f3fc-b689-4e5a-9fee-cea47e067630?token=4a4efb9798d8c50512bbb2da76ea5a7ad6f5f57c58168d6e) current data usage: 6.42 GB." A warning message follows: "You have selected 1.32 GB of data. Some ESGF nodes perform worse than others and even the more performant ones can experience downtime occasionally. This means that your workflow can time out. If this occurs, we advise you to split your workflow into smaller parts or to try a different node." There are buttons for "NEW NOTEBOOK" and "RESUME NOTEBOOK".

&

...Subsetting onto personal workspaces

The screenshot shows a map of Europe with a red rectangular box indicating a subsetting region. The map includes labels for various countries and cities. A "Select Parameters" dialog box is open over the map, showing a "Pseudo-Mercator" projection and a "Zoom" level of 5. The "Coordinate Format" is set to "Lat / Lon" and "GDAL".

Workspaces offer very large storage

Full Data staging allows 3GB per request
(prevents timeout, given the occasional high load on the nodes)

The remote subsetting option, where available, helps reducing the requested data based on user's needs

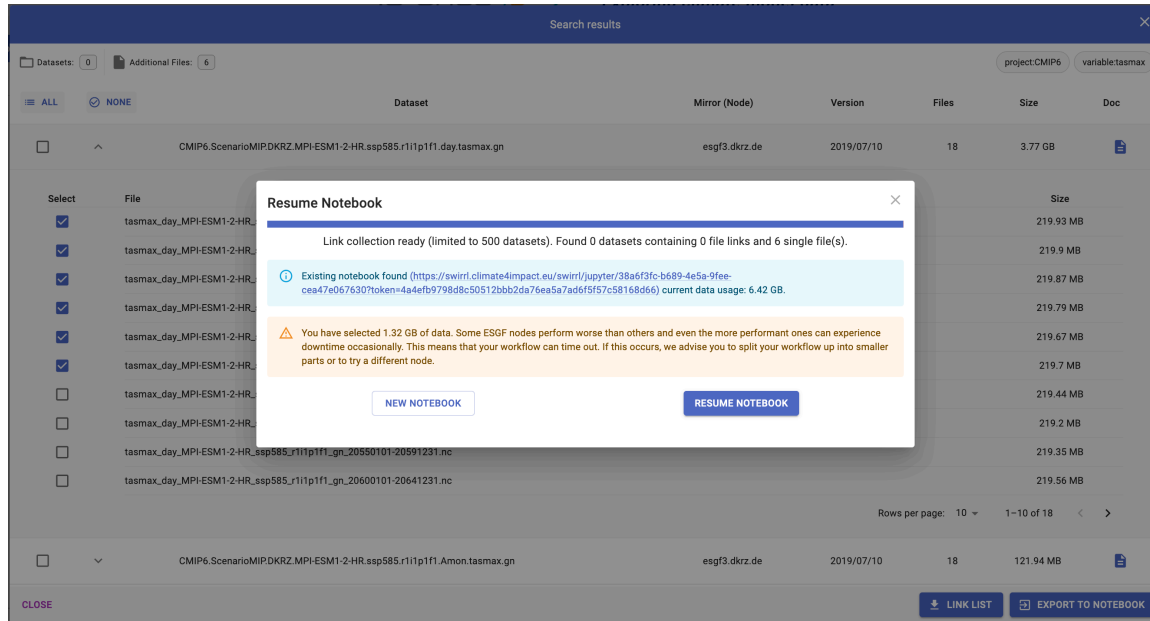
C4I v2 Workflows and Workspaces



Only for registered and validated accounts!
Cost-effective and secure

Full Data Staging...

...Subsetting onto personal workspaces



Resume Notebook

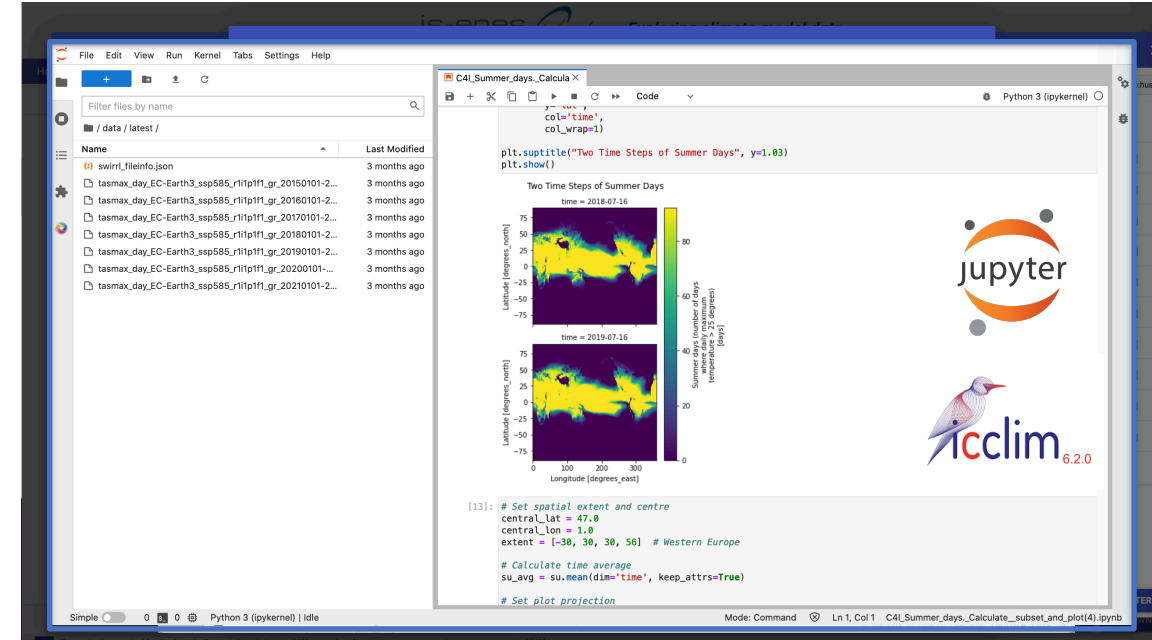
Link collection ready (limited to 500 datasets). Found 0 datasets containing 0 file links and 6 single file(s).

Existing notebook found (<https://swirl.climate4impact.eu/swirl/jupyter/38a6f3fc-b689-4e5a-9fee-cea47e067630?token=4a4efb9798d8c50512bbb2da76ea5a7ad6f5f57c58168d66>) current data usage: 6.42 GB.

You have selected 1.32 GB of data. Some ESGF nodes perform worse than others and even the more performant ones can experience downtime occasionally. This means that your workflow can time out. If this occurs, we advise you to split your workflow into smaller parts or to try a different node.

NEW NOTEBOOK RESUME NOTEBOOK

&



File Edit View Run Kernel Tabs Settings Help

Filter files by name

/ data / latest /

Name Last Modified

- swirl_fileinfo.json 3 months ago
- tasmax_day_EC-Earth3_ssp585_r11p1f1_gr_20150101-2... 3 months ago
- tasmax_day_EC-Earth3_ssp585_r11p1f1_gr_20160101-2... 3 months ago
- tasmax_day_EC-Earth3_ssp585_r11p1f1_gr_20170101-2... 3 months ago
- tasmax_day_EC-Earth3_ssp585_r11p1f1_gr_20180101-2... 3 months ago
- tasmax_day_EC-Earth3_ssp585_r11p1f1_gr_20190101-2... 3 months ago
- tasmax_day_EC-Earth3_ssp585_r11p1f1_gr_20200101-... 3 months ago
- tasmax_day_EC-Earth3_ssp585_r11p1f1_gr_20210101-2... 3 months ago

C4I_Summer_days_Calcula X

```

col='time',
col_wrap=1)

plt.suptitle("Two Time Steps of Summer Days", y=1.03)
plt.show()

```

Two Time Steps of Summer Days

time = 2018-07-16

time = 2019-07-16

Summer days (number of days where daily maximum temperature > 10 degrees)

Latitude (degrees_north)

Longitude (degrees_east)

jupyter

acclim 6.2.0

```

[13]: # Set spatial extent and centre
      central_lat = 47.0
      central_lon = 1.0
      extent = [-30, 30, 30, 56] # Western Europe

      # Calculate time average
      su_avg = su.mean(dim='time', keep_attrs=True)

      # Set plot projection

```

Simple 0 0 Python 3 (ipykernel) | Idle

Mode: Command Ln 1, Col 1 C4I_Summer_days_Calcula_subset_and_plot(4).ipynb

Workspaces offer very large storage

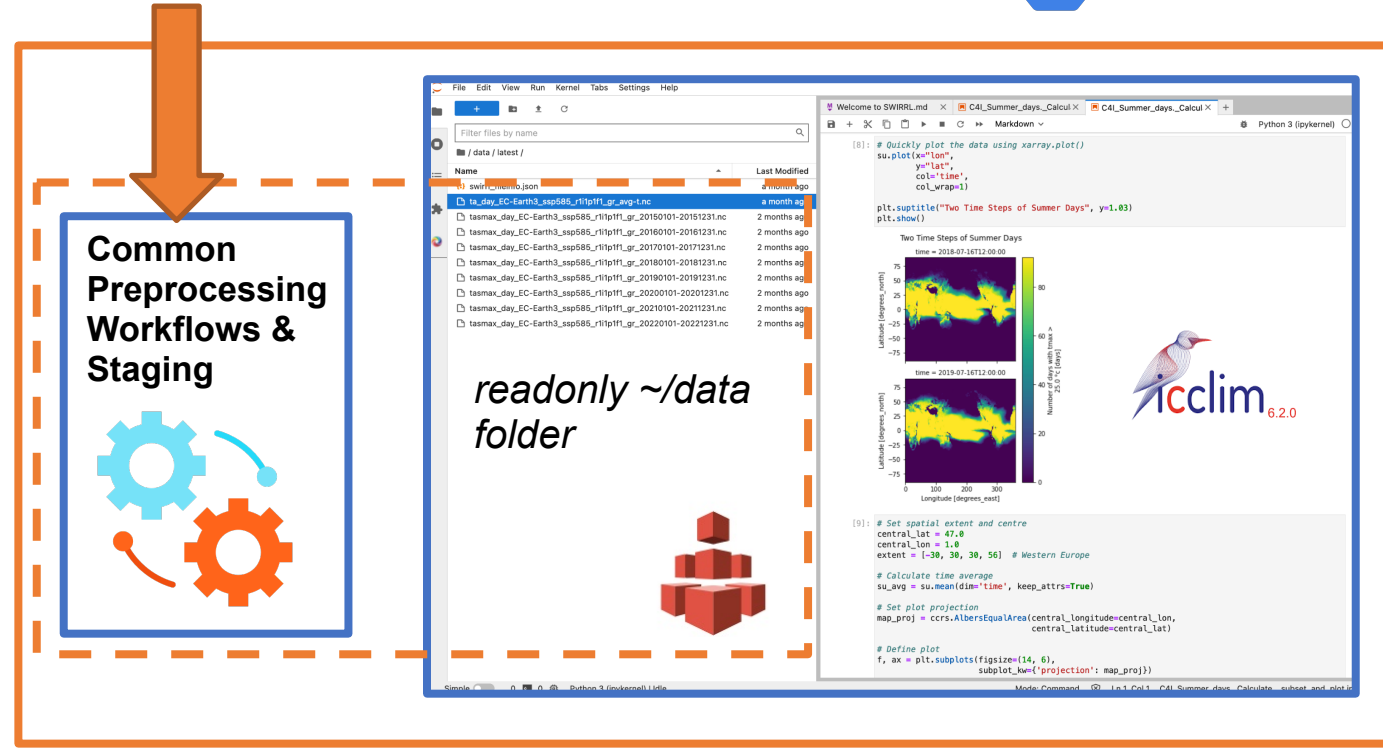
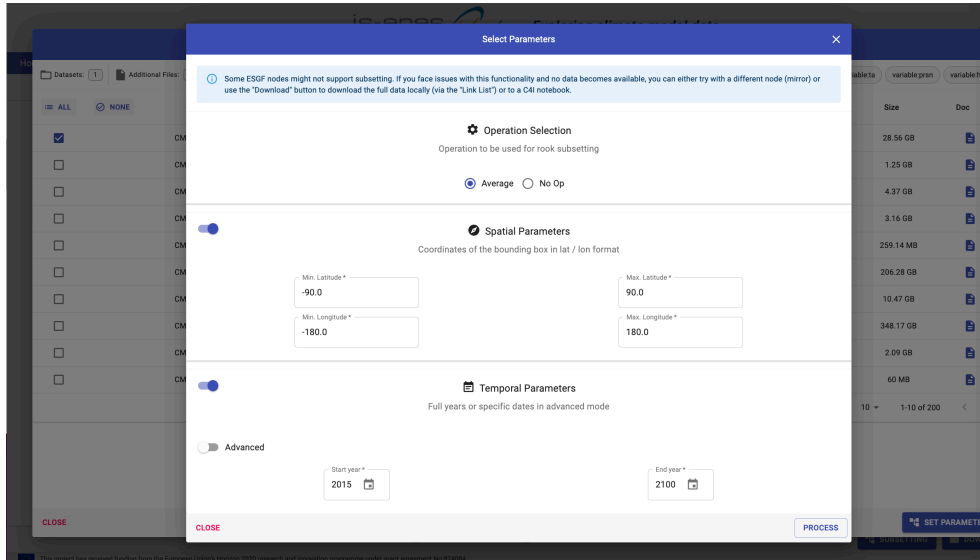
The remote subsetting option, where available, helps reducing the requested data based on user's needs

Full Data staging allows 3GB per request
(prevents timeout, given the occasional high load on the nodes)

SWIRRL - VRE API



Manage **data-driven collaborative workspaces**, combining reproducible **Notebooks, Workflows and Visualisation** tools on the Cloud




SWIRRL

https://doi.org/10.1162/dint_a_00129
<https://gitlab.com/KNMI-OSS/swirrl>



AWS Fargate

icclim sample notebooks



<https://gitlab.com/is-enes-cdi-c4i/notebooks>

File Edit View Run Kernel Tabs Settings Help

Notebook idle

Github LOGIN

Please review your access using [this link](#) to revoke your access tokens.

Create Snapshot SNAPSHOT

Snapshot name SNAPSHOT

Data staging ROLL BACK

Activities LOAD ACTIVITIES

Type	Created at	Action
Rookwps WF ✓	2022-11-29 14:09	Succeeded
Download WF ✓	2022-11-22 18:21	Succeeded
Create ✓	2022-11-22 18:17	RESTORE

Welcome to SWIRRL.md

ICCLIM C4I: Calculate the num

Example notebook that runs ICCLIM, which is pre-

The example calculates the number of summer da

The data is read using xarray and a plot of the tim

output types examples are shown.

The dataset that is expected for this notebook are

and experiment as well as one member. The time p

To keep this example fast to run, the following per

Preparation of the needed modules

```
[1]: import icclim

import sys
import glob
import os
import datetime
import cftime

import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import xarray as xr
import cartopy.crs as ccrs

print("python: ", sys.version)
print("numpy: ", np.__version__)
print("xarray: ", xr.__version__)
print("pandas: ", pd.__version__)
print("icclim: ", icclim.__version__)
print("cftime: ", cftime.__version__)

python: 3.10.6 | packaged by conda-forge
numpy: 1.23.5
xarray: 2022.11.0
```

IS-ENES Climate Data Infrastructure for Climate 4 Impact > C4I Use Cases as Jupyter Notebooks

C4I Use Cases as Jupyter Notebooks

Project ID: 25761638

27 Commits 1 Branch 0 Tags 6.7 MB Project Storage

A collection of Jupyter Notebooks implementing some Use Cases.

Updated to match icclim 5.0.2
Christian Pagé authored 1 year ago

5825906e Find file Clone

master notebooks

README

Name	Last commit	Last update
C4I_Averaged_Temperature_Anomaly_2...	Updated with icclim 5.0.2.	1 year ago
C4I_Summer_days__Calculate__subset_a...	Updated with icclim 5.0.2.	1 year ago
C4I_TX90p__Calculate__subset_and_plot...	Updated with icclim 5.0.2.	1 year ago
C4I_deltaT_deltaP_Anomaly_2081-2100...	Updated with icclim 5.0.2.	1 year ago
C4I_userindex_maxtemp_freezing.ipynb	Updated with icclim 5.0.2.	1 year ago
README.md	Updated to match icclim 5.0.2.	1 year ago
cordex_download.ipynb	Adjustments to dask config.	1 year ago

README.md



Reproducibility Controls

Workflow Monitoring

GitHub Authentication



Snapshot Controls



Data Staging Rollback

Activities History
and Provenance



The screenshot shows a JupyterLab interface with a sidebar on the left and a main workspace on the right. The sidebar contains several sections: 'Github' with a 'SWITCH USER' button, 'Create Snapshot' with a 'SNAPSHOT' button, 'Data staging' with a 'ROLL BACK' button circled in red, and 'Activities' with a 'LOAD ACTIVITIES' button. Below these is a table of activities with columns for 'Type', 'Created at', and 'Action'. The table lists several activities, including 'Library Update', 'Rookwps WF', 'Download WF', and 'Create', each with a 'RESTORE' button circled in red. The main workspace shows a code editor with Python code for time subsetting and plotting, and a plot titled 'Two Time Steps of Summer Days' showing two heatmaps of summer days over Western Europe for the years 2018 and 2019. The plot's y-axis is 'Latitude [degrees_north]' and the x-axis is 'Longitude [degrees_east]'. A color bar on the right indicates the 'Number of days with tmax > 25.0 °C [days]'.

Users create intermediate and shareable snapshots on GitHub, Binders, compatible with <https://mybinder.org/>

Rollback workflows and Restore environments to a previous state

Reproducibility Controls



The screenshot shows the SWIRRL interface with a table of activities and a notebook execution window. The table lists activities like 'Rookwps WF' and 'Download WF' with their status and timestamps. The notebook window shows Python code for plotting summer days data, resulting in two heatmaps for the years 2018 and 2019.

Type	Created at	Action
Rookwps WF 🔗	2022-11-29 14:09	🔄
Succeeded		
Download WF 🔗	2022-11-22 18:21	🔄
Succeeded		
Create 🔗	2022-11-22 18:17	🔄 RESTORE

```

[8]: # Quickly plot the data using xarray.plot()
su.plot(x="lon",
        y="lat",
        col='time',
        col_wrap=1)

plt.suptitle("Two Time Steps of Summer Days", y=1.03)
plt.show()

Two Time Steps of Summer Days
time = 2018-07-16T12:00:00
Latitude [degrees_north]
Longitude [degrees_east]
Number of days with tmax > 25.0 °C (days)

time = 2019-07-16T12:00:00
Latitude [degrees_north]
Longitude [degrees_east]
Number of days with tmax > 25.0 °C (days)

[9]: # Set spatial extent and centre
central_lat = 47.0
central_lon = 1.0
extent = [-30, 30, 30, 56] # Western Europe

# Calculate time average
su_avg = su.mean(dim='time', keep_attrs=True)

# Set plot projection
map_proj = ccrs.AlbersEqualArea(central_longitude=central_lon,
                               central_latitude=central_lat)

# Define plot
f, ax = plt.subplots(figsize=(14, 6),
                    subplot_kw={'projection': map_proj})
    
```

The screenshot shows a snippet of a PROV-O provenance record. A red box highlights the record, and an orange arrow points from the SWIRRL interface to it. The record includes details about the activity, its start and end times, and its association with other resources.

```

{
  "prov:used": [...],
  "provone:hadPart": [
    {
      "prov:wasAssociatedWith": [...],
      "@type": [
        "Resource",
        "prov:Activity",
        "provone:Execution"
      ],
      "rdfs:label": "orchestrate",
      "prov:startedAtTime": "2022-11-29T14:13:02Z",
      "@id": "urn:roocs:orchestrate_16ca3e1f-fee6-4419-a264-65d273a801bf",
      "prov:endedAtTime": "2022-11-29T14:13:45Z"
    },
    {
      "prov:wasAssociatedWith": [
        { ... }
      ],
      "prov:wasActivityOfInfluence": [...],
      "@type": [
        "Resource",
        "prov:Activity",
        "provone:Execution"
      ],
      "rdfs:label": "average_ta_1",
      "roocs:apply_fixes": false,
      "@id": "urn:roocs:average_ta_1_6eabac6b-444e-4e9a-a71b-bcc1b16b9fb1",
      "roocs:dims": "time"
    },
    { ... }
  ],
  "@type": [...],
  "@context": { ... },
  "prov:generated": { ... },
  "swirrl:sessionId": "13abfb95-fd4a-463f-ba61-fc96d40a9b6c",
  "swirrl:message": "Succeeded",
  "prov:wasAssociatedWith": [...],
  "@id": "urn:uuid:49f0d7f5-5cc8-4893-8e17-2a18ed870540",
  "swirrl:jobId": "49f0d7f5-5cc8-4893-8e17-2a18ed870540",
  "prov:endedAtTime": "2022-11-29T13:13:54.220Z",
  "prov:atLocation": "POST /workflow/rookwps/run/",
  "prov:startedAtTime": "2022-11-29T13:09:03.676Z"
}
    
```

automated provenance management and data-versioning, fostering traceability and reproducibility.

Help Material and Feedback Form



How do I search for data?

In the **Data Discovery Page** you select the project you want to access the data from (CMIP6 - CMIP5 - CORDEX).

Once you have selected the project, the portal will display a page where you can compose queries interactively by combining different properties. These are shown in separated tabs on the top of the page (*variable, frequency, experiment, model, member*) and can be configured in any order.



When configuring a specific property, to facilitate identifying the most relevant options, a subset of the possible values is presented, grouped by different topics. This is the *Scoped* view. This presentation mode can be changed by switching the knob at the top left of the screen, from *Scoped* to *Extended*. The *Extended* presentation mode will display, for each property, all the possible values available for the current query configuration.

In the image above we show that the summary of the current query is always displayed on the top of the screen. Each value can be interactively removed. Adding or removing a property value will always result in the live update of the available choices for each property.

When selecting the Model, by clicking on **Compare Model Performance** at the bottom of the Model panel, users will be prompted with a page showing how the selected models (and the others available) foresee a change on a variable (Temperature/Precipitation) compared to the past. The page is generated by using the [ESMValTool](#) and, in combination with the bias, gives some confidence on the performance.

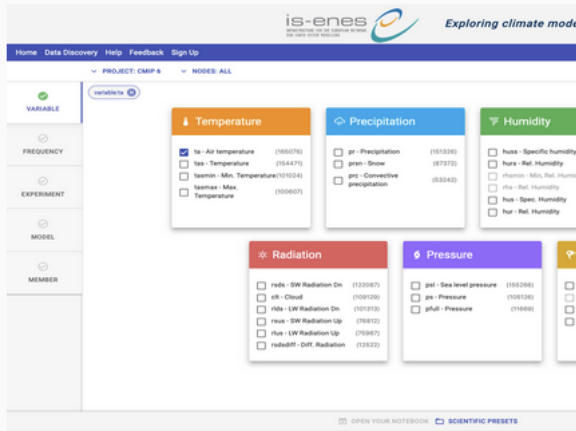
Help Material and Feedback Form



How do I search for data?

In the **Data Discovery Page** you select the project you want to access the data.

Once you have selected the project, the portal will display a page where you can select the **variable, frequency, experiment, model, member** and can be configured in the **SWIRRL Explorer**.



The screenshot shows the 'Data Discovery' page with a sidebar on the left containing filters for VARIABLE, FREQUENCY, EXPERIMENT, MODEL, and MEMBER. The main area displays a grid of variable categories: Temperature, Precipitation, Humidity, Radiation, and Pressure. Each category has a list of specific variables with checkboxes and IDs. For example, under Temperature, 'ta - Air Temperature' is selected with ID (165076).

When configuring a specific property, to facilitate identifying the most relevant data, the configuration can be changed by switching the knob at the top left of the screen, from **Scientific** to **SoC** configuration.

In the image above we show that the summary of the current query is always updated, and the live update of the available choices for each property.

When selecting the Model, by clicking on **Compare Model Performance** at the bottom, you can foresee a change on a variable (Temperature/Precipitation) compared to the other models.



The screenshot shows the 'Climate impact result viewer' interface. It includes the is-enes logo and navigation links. A message at the bottom states: 'This application shows a results from CMIP6 and CMIP5 model researchers, to select one or more datasets that adequately describe the climate impact result viewer'.

Data Management: the /data folder


Thanks to the support for workflows that perform data-staging operations from remote providers, workspace are populated with the data you are interested in.

SWIRRL manages the data in the **/data** folder, which is accessible from your ***JupyterLab** instance. To make sure that this folder is always consistent and your analysis reproducible, this is set as **'read only'**. However, you can always request to add more data.

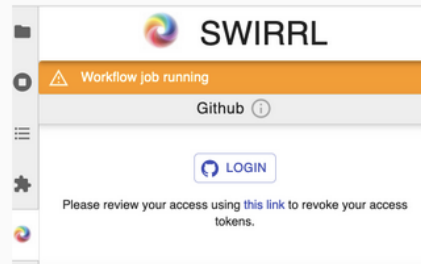
/data/latest

Is the data that you have requested already available?

Check your **/data/latest** folder! Here you will find the most updated version of the data available for you to use.

Not there yet? Click on the **SWIRRL Explorer** 

In this handy control panel you can check whether there are workflows running in the background. For instance, one might be copying data to your workspace just now.

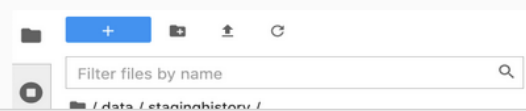


The screenshot shows the SWIRRL control panel. At the top, it says 'SWIRRL' with a logo. Below that, there's a section 'Workflow job running' with a warning icon. Underneath, it says 'Github' with a help icon. There is a 'LOGIN' button. At the bottom, a message says: 'Please review your access using this link to revoke your access tokens.'

You will read more about the functionalities of SWIRRL Explorer in this short guide!

/data/staginghistory

Each time data is staged to your workspace, SWIRRL keeps track of the changes by maintaining a staging history. You will find this in the **/data/staginghistory** folder.



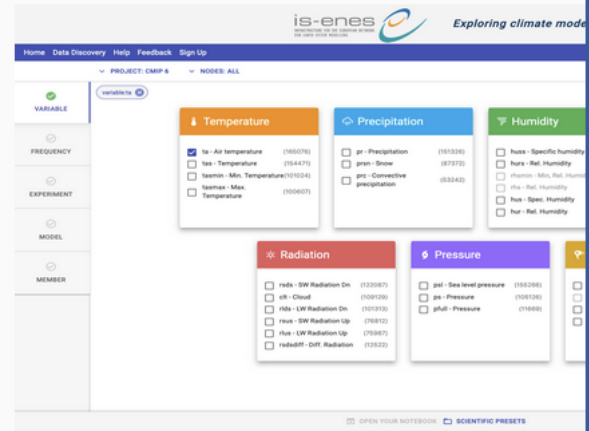
The screenshot shows a file explorer interface. At the top, there are icons for adding, uploading, and refreshing. Below that is a search bar labeled 'Filter files by name'. At the bottom, a file named '/data/staginghistory/' is visible.

Help Material and Feedback Form



How do I search for data?

In the **Data Discovery Page** you select the project you want to access the data. Once you have selected the project, the portal will display a page where you can select the **variable, frequency, experiment, model, member** and can be configured in the **SWIRRL Explorer**.



When configuring a specific property, to facilitate identifying the most relevant data, the configuration can be changed by switching the knob at the top left of the screen, from **Scientific** to **Configuration**.

In the image above we show that the summary of the current query is always updated with the live update of the available choices for each property.

When selecting the Model, by clicking on **Compare Model Performance** at the bottom, you can foresee a change on a variable (Temperature/Precipitation) compared to the other models.



Data Management: the /data folder

Thanks to the support for workflows that perform data-staging operations from remote providers, workspace are populated with the data. SWIRRL manages the data in the **/data** folder, which is accessible from your ***JupyterLab** instance. To make sure that this folder is accessible, you need to set the **'read only'**. However, you can always request to add more data.

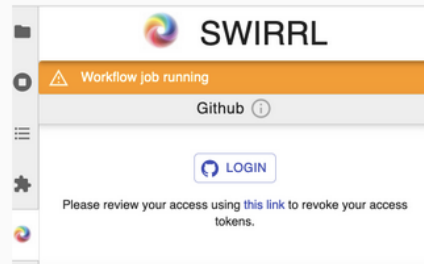
/data/latest

Is the data that you have requested already available?

Check your **/data/latest** folder! Here you will find the most updated version of the data available for you to use.

Not there yet? Click on the **SWIRRL Explorer** 

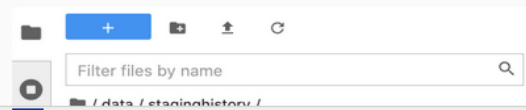
In this handy control panel you can check whether there are workflows running in the background. For instance, one might be copying data from a remote provider to the workspace.



You will read more about the functionalities of SWIRRL Explorer in this short guide!

/data/staginghistory

Each time data is staged to your workspace, SWIRRL keeps track of the changes by maintaining a staging history. You will find this in the **/data/staginghistory** folder.



Email *

Il tuo indirizzo email

What is your function? *

This helps us get a better understanding of whom our end users are.

Climatologist

Impact Researcher

Hydrologist

Geologist

Altro: _____

What kind of feedback to you have? *

This helps us organize the feedback we receive.

Feature Request

Bug Report

Web Analytics and Computational KPIs



GDPR Compliant Web analytics service (Matomo)

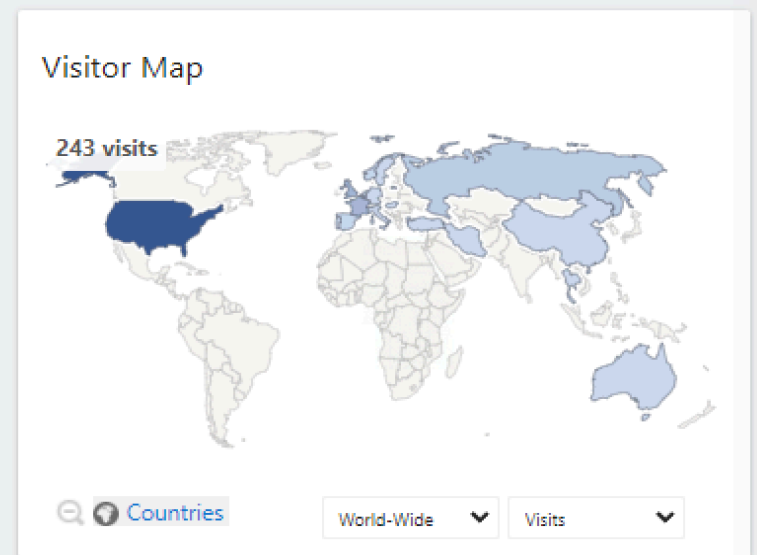
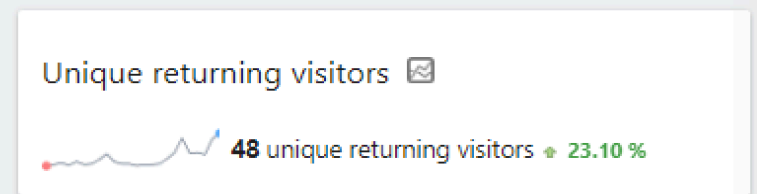


Actions: Event Categories

EVENT CATEGORY	EVENTS	EVENT VALUE
FacetSideBar	710	-
TailoredSearch	171	-
DownloadResultsWindow	85	-
DownloadMetalinkDialog	57	-
SubsetResultsWindow	13	-
OpenNotebook	12	-
ResumeNotebookDownloadDialog	7	-
NewNotebookSubsettingDialog	7	-

1-8 of 8

Secondary dimension is Event Action.
Switch to Event Name

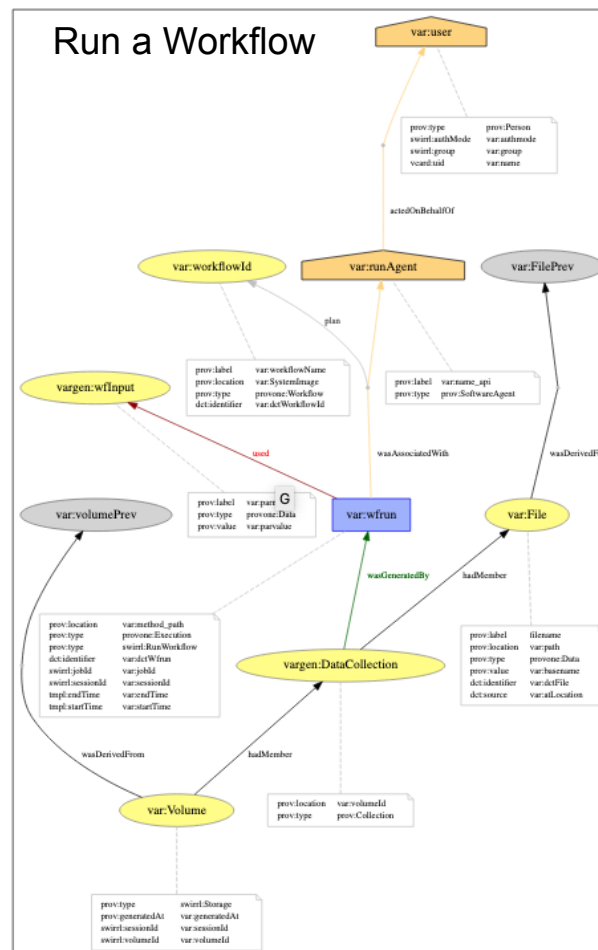
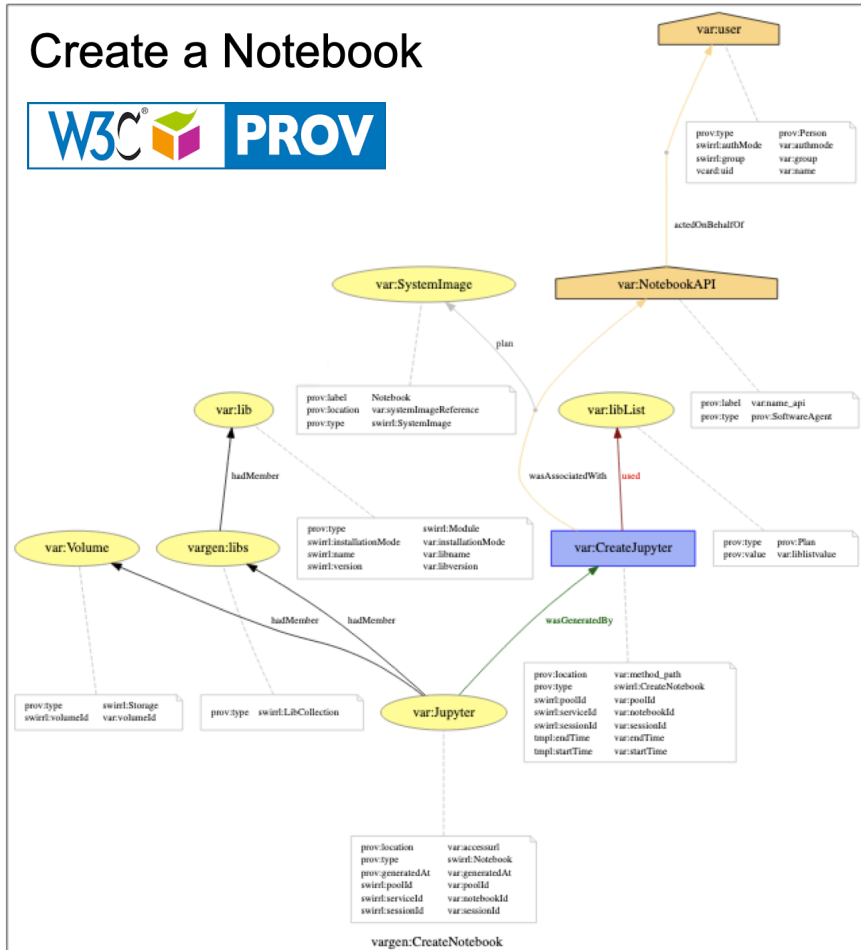


Tracking at Event Level allows to get useful insights at a very high granularity.

Web Analytics and Computational KPIs



Provenance used to Collect Computational KPIs (examples)



Notebook stats:

session_count	notebook_count
34	34

Workflow stats:

wf_types	run_count	success %
download	98	82%
rookswps	18	72%

Staged-Data

Total Data Files	Unique Data Files
1104	949

>90 Users Registered via the new ESGF SSO system at CEDA

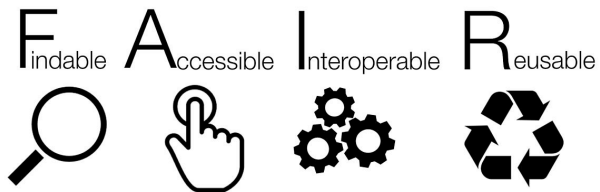
Climate4Impact v2



Improvements from v1

- **Search Usability**
- **Models Performance Comparison (ESMValTool)**
- **Single Sign On** with ESGF
- **Flexible analysis features** (Workspaces with ICCLIM Notebooks, Data Staging & Subsetting Workflows)

- **Automated reproducibility mechanisms fostering FAIR research**



- **Decoupled Architecture (SWIRRL/SSO/WPS)**

Ongoing and Future work.....



- **Access to CORDEX** with ESGF SSO supported at more CORDEX nodes (ongoing)
- **More ESGF nodes to implement Subsetting WPS** (e.g. CMCC coming soon)
- **More Notebooks Use Cases** using *icclim*
<https://gitlab.com/is-enes-cdi-c4i/notebooks>
- **Scientific Guidance produced by experts**
- **Use SWIRRL to create Collaborative Data Caches**
(frequently requested by lecturers for educational purposes)