

# Interactive and Flexible Environment for Climate Data Analysis

Christian Pagé

CERFACS, Toulouse, France



Abel Aoun

*CERFACS Toulouse, France*



Royal Netherlands  
Meteorological Institute  
*Ministry of Infrastructure  
and the Environment*

Alessandro Spinuso, Mats Veldhuizen, Ian van der Neut

*KNMI, De Bilt, Netherlands*



Klaus Zimmermann, Lars Barring

*SMHI, Norrköping, Sweden*



Phillip Kershaw

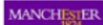
*NCEO/CEDA, STFC, Didcot, UK*

14th International Workshop on Science Gateways

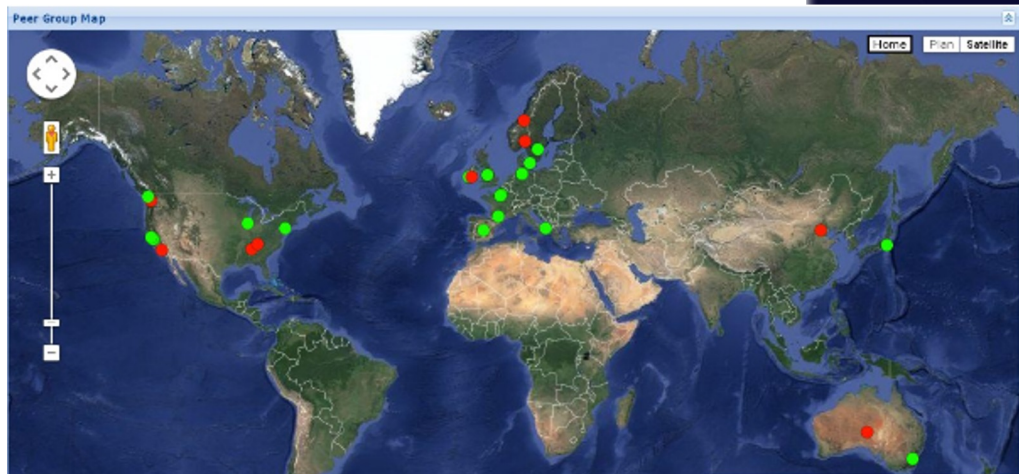
15th-17th June 2022, Trento, Italy



Max-Planck-Institut  
für Meteorologie

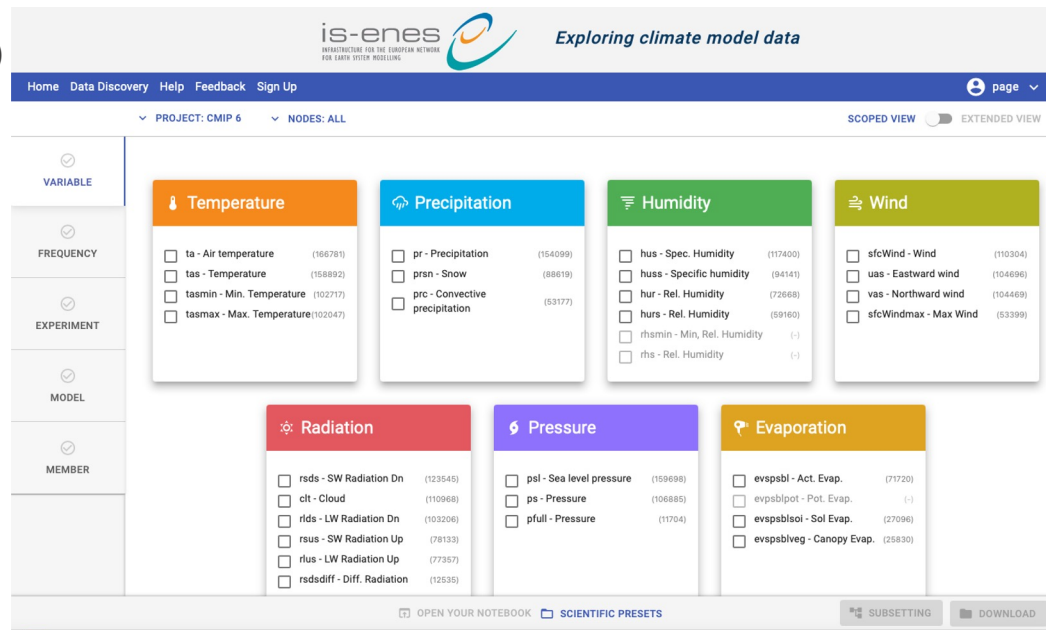


ESGF represents a **multinational** effort to securely **access, monitor, catalog, transport, and distribute** reference **data** for **climate** research experiments and observations.



## Platform for researchers to explore climate data and perform analysis

- Front-end to climate data infrastructure (ESGF)
- Tailored Search Interface with view modes
- Jupyter-Lab enhanced environment
- Notebooks gallery
- Flexible analysis features
  - Climate indices calculation: **icclim**
  - Data Staging/Reduction Workflows
  - Personal store for processing outcome
- Automated reproducibility mechanisms and documentation (Data/Analysis)
- Modular Deployment & Decoupled Architecture
- Pages for Models Performance Comparison (**ESMValTool**)



V2: Complete Redesign from V1

Beta version available <https://dev.climate4impact.eu>





## Search Parametrisation made easier

The screenshot shows the is-enes search interface with the following parametrisation options:

- Temperature**
  - ta - Air temperature (9)
  - tas - Temperature (6)
  - tasmin - Min. Temperature (6)
  - tasmax - Max. Temperature (6)
- Precipitation**
  - pr - Precipitation (6)
  - prsn - Snow (4)
  - prc - Convective precipitation (1)
- Radiation**
  - rads - SW Radiation Dn (5)
  - rflus - LW Radiation Up (4)
  - rsus - SW Radiation Up (2)
  - rlds - LW Radiation Dn (2)
  - radsdiff - Diff. Radiation (-)
  - clt - Cloud (2)
- Pressure**
  - ps - Pressure (-)
  - psl - Sea level (-)
  - pfull - Pressure (-)

## Nodes Selection by Service

The screenshot shows the is-enes search interface with an 'Available ESGF Nodes' dialog box open. The dialog box has the following content:

**Available ESGF Nodes**

Select & enable Rock WPS subsetting

Node	Subsetting Mode
<input checked="" type="checkbox"/> esgf1.dkrz.de	Rock WPS
<input checked="" type="checkbox"/> esgf3.dkrz.de	Rock WPS
<input type="checkbox"/> aims3.llnl.gov	Openpand
<input type="checkbox"/> cmip.dess.tsinghua.edu.cn	Openpand
<input type="checkbox"/> cmip.fio.org.cn	Openpand
<input type="checkbox"/> cordexesg.dmi.dk	Openpand
<input type="checkbox"/> crd-esgf-drc.ec.gc.ca	Openpand
<input type="checkbox"/> data.meteo.unican.es	Openpand
<input type="checkbox"/> dataserver.nccs.nasa.gov	Openpand
<input type="checkbox"/> dpegf03.nccs.nasa.gov	Openpand
<input type="checkbox"/> esg.cccr.tropmet.res.in	Openpand

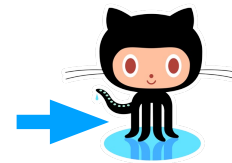
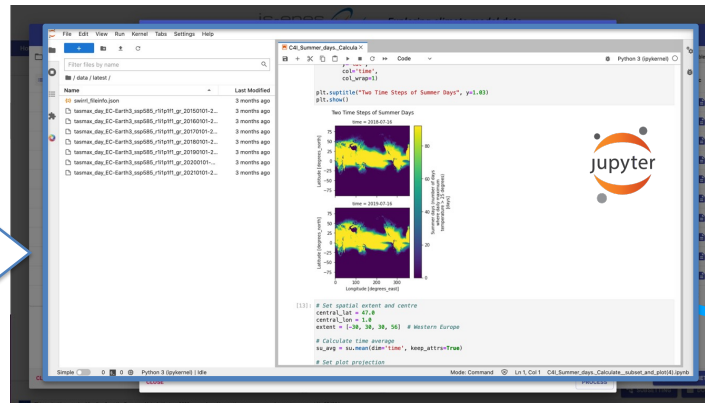
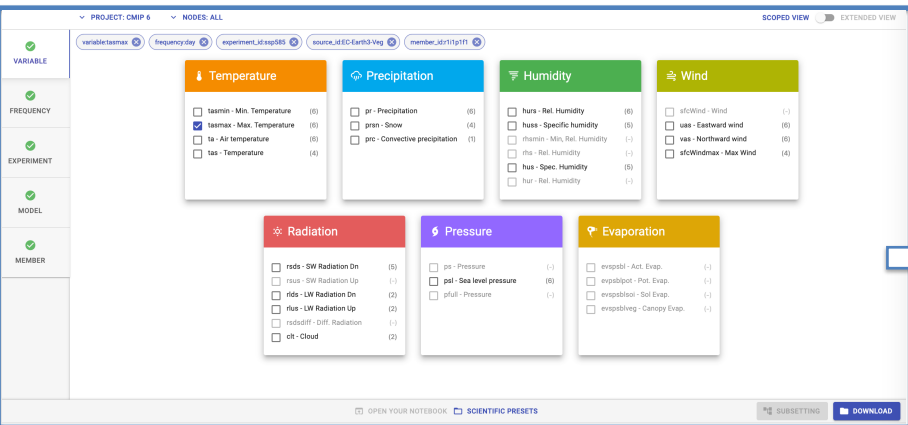
OK



Climate4Impact Search for CMIP5/6  
CORDEX Data

<https://dev.climate4impact.eu>

Workflows for data staging &  
reduction onto icclim powered Notebooks

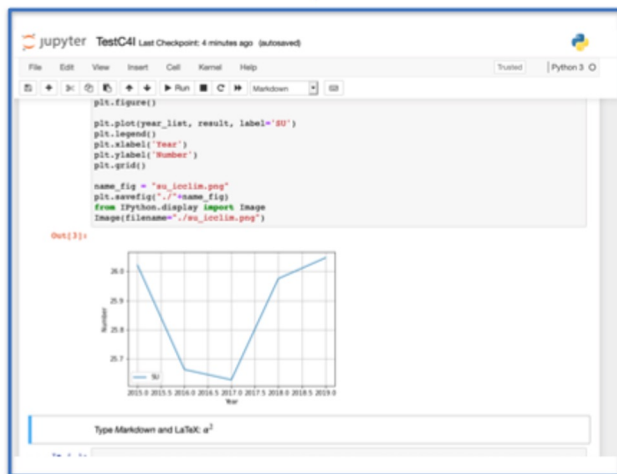
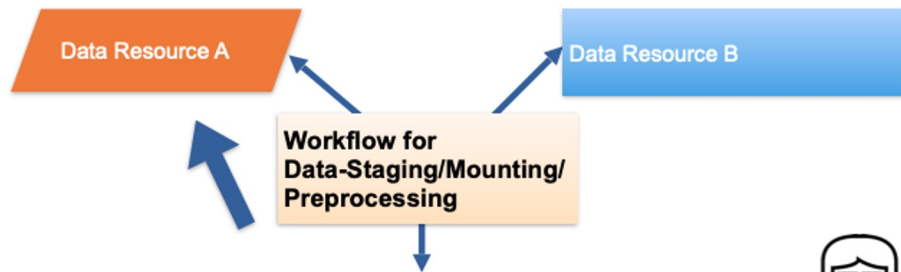


Save/Share  
Progress  
to Git

Reduced  
Data MyBinder  
Reproduce

- Trace Changes to Restore, Recover  
Software and/or Data





  
**a researcher wants**



- **access distributed raw data**
- **develop, document and reuse** methods for processing and visualisation.
- **update/extend** raw data and software
- **Track changes and rollback** (Traceability/Recovery)
- **keep old versions of the data** after updates (Reproducibility)
- **snapshot and restore** the state of a workspace software (Reproducibility)

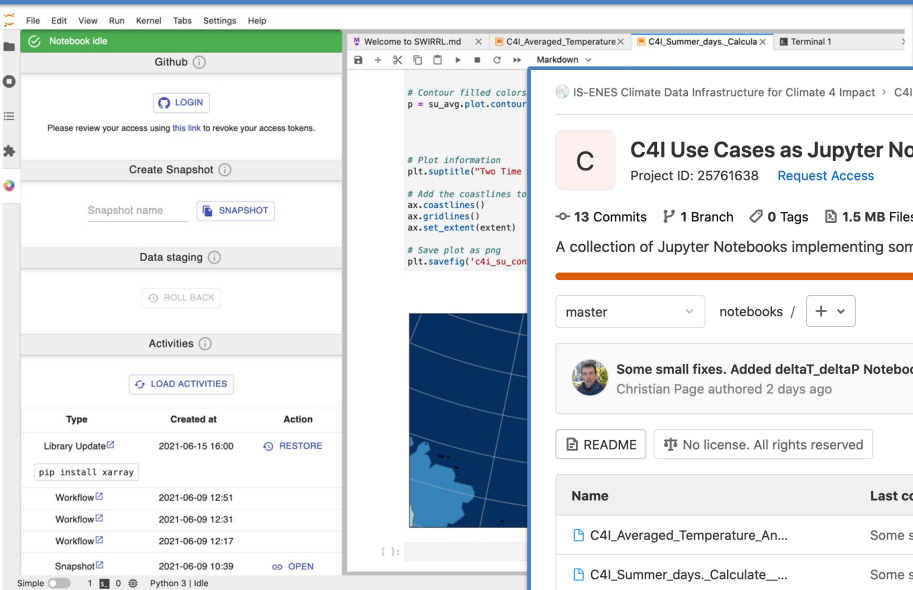
**Workflow Monitoring**

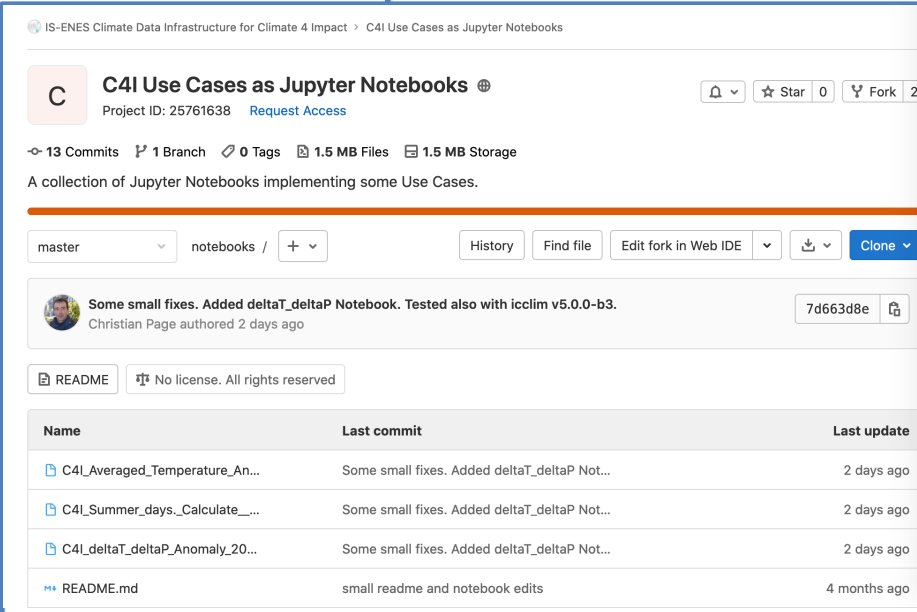
**GitHub Authentication**

**Snapshot Controls**

**Data Staging Rollback**

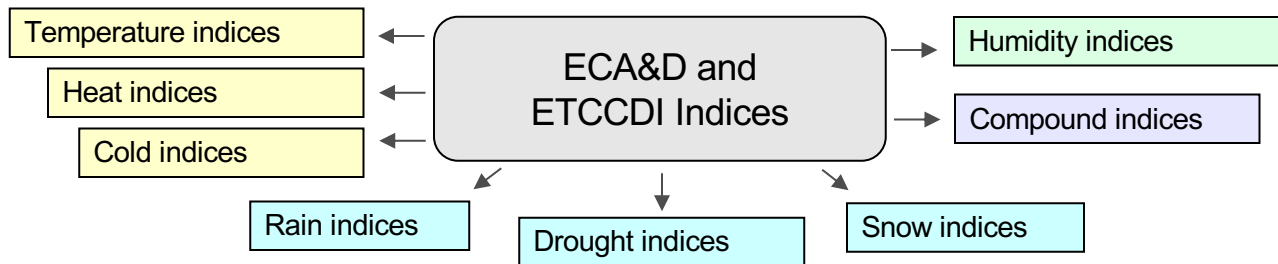
**Activities History and Provenance**





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

## Climate indices calculation in climate4impact: **icclim**



- Intra-period extreme temperature range [°C] - **ETR**
- Warm days (days with mean temperature > 90th percentile of daily mean temperature) - **TG90p**
- Summer days (days with max temperature > 25 °C) - **SU**
- ...

- Python code developed at Cerfacs since September 2013
- Funded by EU FP7 IS-ENES2, FP7 CLIPC and H2020 IS-ENES3
- Generic and modular approach, can be reused in other environments
- New V5 completely rewritten and using underlying xclim functions, based on xclim, xarray and dask
- I/O interface is structured for optimal performance
- Implement the proper percentile indices calculations when calculation period overlaps reference period (called bootstrapping method)



Documentation: [https://icclim.readthedocs.io/en/latest/python\\_api.html](https://icclim.readthedocs.io/en/latest/python_api.html)

Source code: <https://github.com/cerfacs-globc/icclim>

Current Version 5.3: <https://github.com/cerfacs-globc/icclim/releases/tag/5.3>

## `icclim.index(**kwargs)`

- Parameters:**
- **in\_files** (*str* | *list[str]* | *Dataset* | *DataArray*) – Absolute path(s) to NetCDF dataset(s), including OPeNDAP URLs, or path to zarr store, or `xarray.Dataset` or `xarray.DataArray`.
  - **index\_name** (*str*) – Climate index name. For ECA&D index, case insensitive name used to lookup the index. For user index, it's the name of the output variable.
  - **var\_name** (*str* | *list[str]* | *None*) – **optional** Target variable name to process corresponding to `in_files`. If `None` (default) on ECA&D index, the variable is guessed based on the climate index wanted. Mandatory for a user index.
  - **slice\_mode** (*str*) – Type of temporal aggregation: {"year", "month", "DJF", "MAM", "JJA", "SON", "ONDJFM" or "AMJJAS"}. Default is "year". See `slice_mode` for details.
  - **time\_range** (*list[datetime.datetime]*) – **optional** Temporal range: upper and lower bounds for temporal subsetting. If `None`, whole period of input files will be processed. Default is `None`.
  - **out\_file** (*str* | *None*) – Output NetCDF file name (default: "icclim\_out.nc" in the current directory). Default is "icclim\_out.nc". If the input `in_files` is a `Dataset`, `out_file` field is ignored. Use the function returned value instead to retrieve the computed value. If `out_file` already exists, icclim will overwrite it!
  - **threshold** (*float* | *list[float]* | *None*) – **optional** User defined threshold for certain indices. Default depend on the index, see their individual definition. When a list of threshold is provided, the index will be computed for each thresholds.

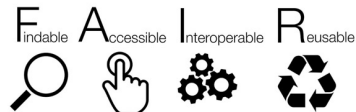


SWIRRL hides the complexity of orchestrating Workspaces in a target Cloud resource based on Kubernetes Cluster

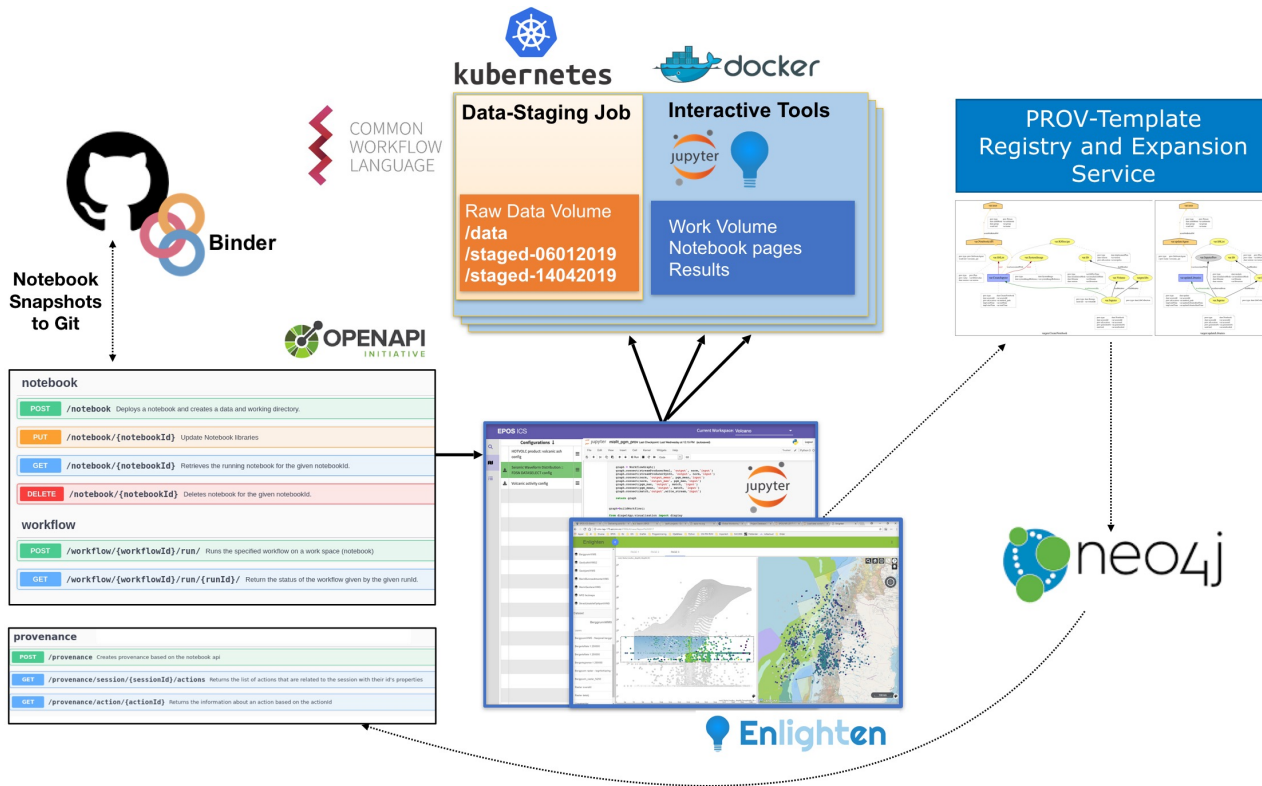
Integrates Interactive Tools (Notebooks, Workflows, GIT, Binder)

Offers a REST Web API

Manages Metadata (Provenance)



Supports Reproducibility



is-enes  
INFRASTRUCTURE FOR THE EUROPEAN NETWORK  
FOR EARTH SYSTEM MODELLING

Home Data Discovery Help Feedback Sign Up

PROJECT: CIMP 6 NODES: CUSTOM (2)

variable:ta variable:prsn variable:huss frequency:day experiment\_id:ssp585

VARIABLE

FREQUENCY

EXPERIMENT

MODEL

MEMBER

**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 Simulation
- AWI-CM-1-1-MR - AWI-CM 1.1 MR
- NorESM2-LM - NorESM2-LM (low atmosphere-medium ocean resolution)
- BCC-CSM2-MR - BCC-CSM 2 MR
- CMCC-CM2-SR5 - CMCC-CM2-SR5
- ERA5

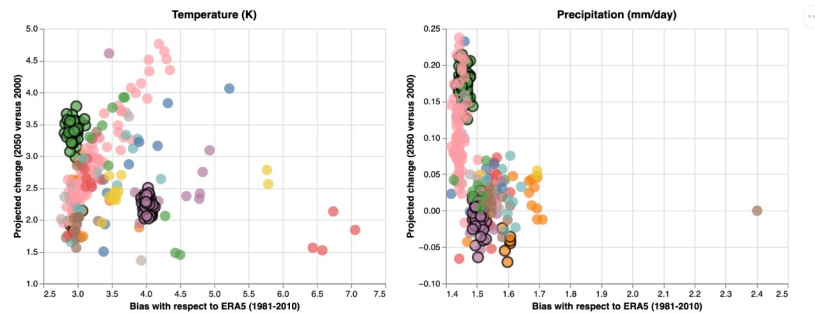
[COMPARE MODEL PERFORMANCE](#)

## Climate impact result viewer

This application shows 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 calculated 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.

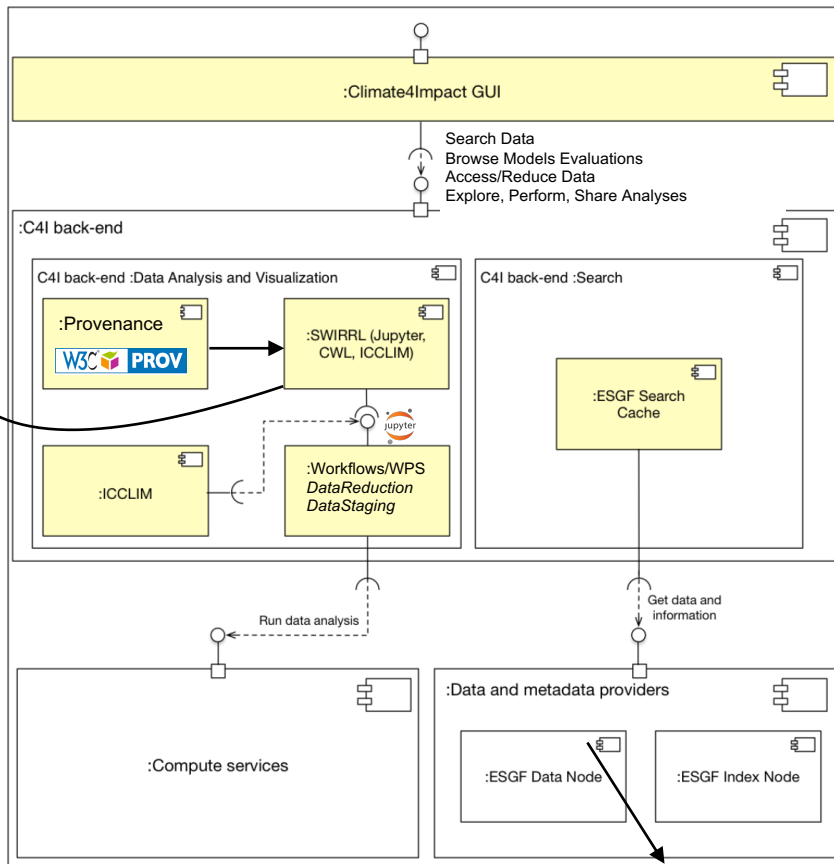


Project\_project CIMP6

Selected datasets:

- CNRM-CM6-1
- MIROC6
- CanESM5

[View static recipe output](#)



SSO with IDeA AAI

ESMValTool Evaluation pages

Link to DOI Pages ES-DOC

Git  
Store/Share  
Analyses



Binder  
Reproduce



# Thanks !

On behalf of the climate4impact and icclim teams

<https://dev.climate4impact.eu>

<https://icclim.readthedocs.io/en/stable/>

Christian Pagé  
Alessandro Spinuso  
Abel Aoun

christian.page@cerfacs.fr  
alessandro.spinuso@knmi.nl  
abel.aoun@cerfacs.fr



*This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N°824084*

**THE CONSORTIUM**

Coordinated by CNRS-IPSL, the IS-ENES3 project gathers 22 partners in 11 countries




Our website  
<https://is.enes.org/>



Follow us on Twitter !  
**@ISENES\_RI**



Contact us at  
[is-enes@ipsl.fr](mailto:is-enes@ipsl.fr)



Follow our channel  
**IS-ENES3 H2020**

