

Ease Access to Climate Simulations for Researchers: IS-ENES climate4impact

Christian Pagé

Research Engineer / Climate Research Domain

 CERFACS Toulouse, France

Xavier Pivan, Toulouse

Alessandro Spinuso, Maarten Plieger, Wim Som de Cerff, KNMI, Netherlands



THE UNIVERSITY
of EDINBURGH

 CERFACS
CENTRE EUROPÉEN DE RECHERCHE ET DE FORMATION AVANCÉE EN CALCUL SCIENTIFIQUE

 UNIVERSITY OF
LIVERPOOL

 Koninklijk Nederlands
Meteorologisch Instituut
Ministerie van Infrastructuur en Milieu

 grnet

 Fraunhofer


ATHENA
Research & Innovation
Information Technologies

Climate Data Distribution

ESGF Data Nodes 2019

- 31 worldwide
- 18 in Europe (17 institutions)
(coordinated by IS-ENES)



BC2DC'2019, 24 Sep 2019 San Diego, CA, USA

IS-ENES CDI climate4impact

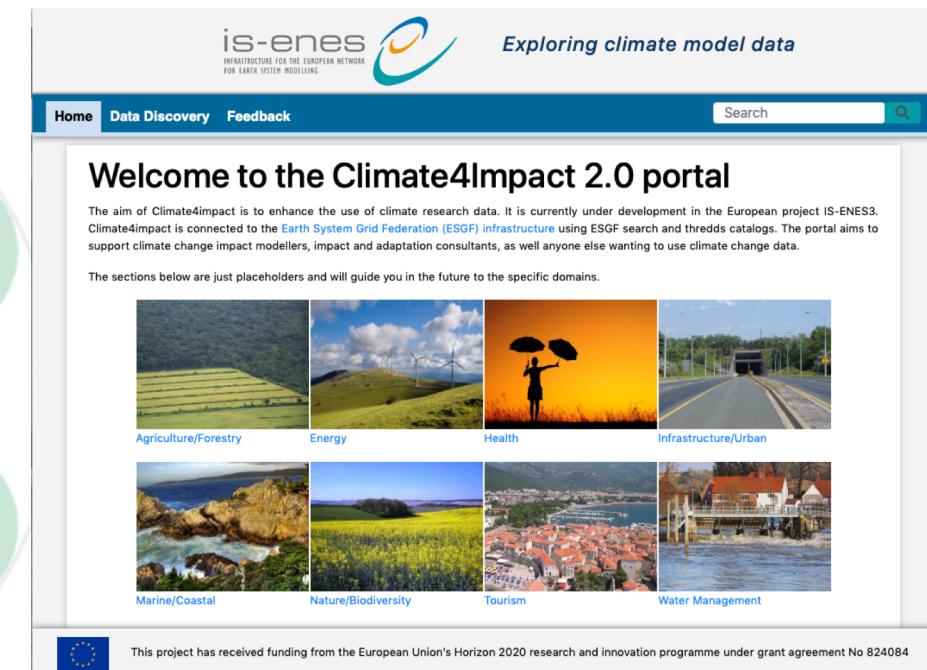
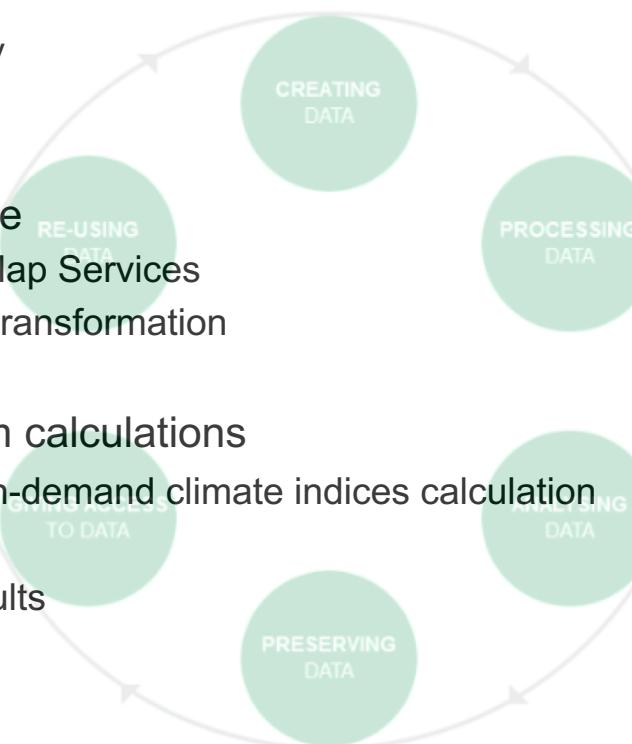
- Tailored for end-users
- Supports on-demand data processing

The screenshot shows two panels of the is-enes Data Discovery interface. The top panel displays a search bar and filters for Temperature, Precipitation, Humidity, Wind, Radiation, Pressure, and Evaporation. The bottom panel shows a preview of a climate dataset for Near-Surface Air Temperature (tas) on September 21, 1949, at 12:00 UTC, with a color scale from 223.0 to 311.0 K.

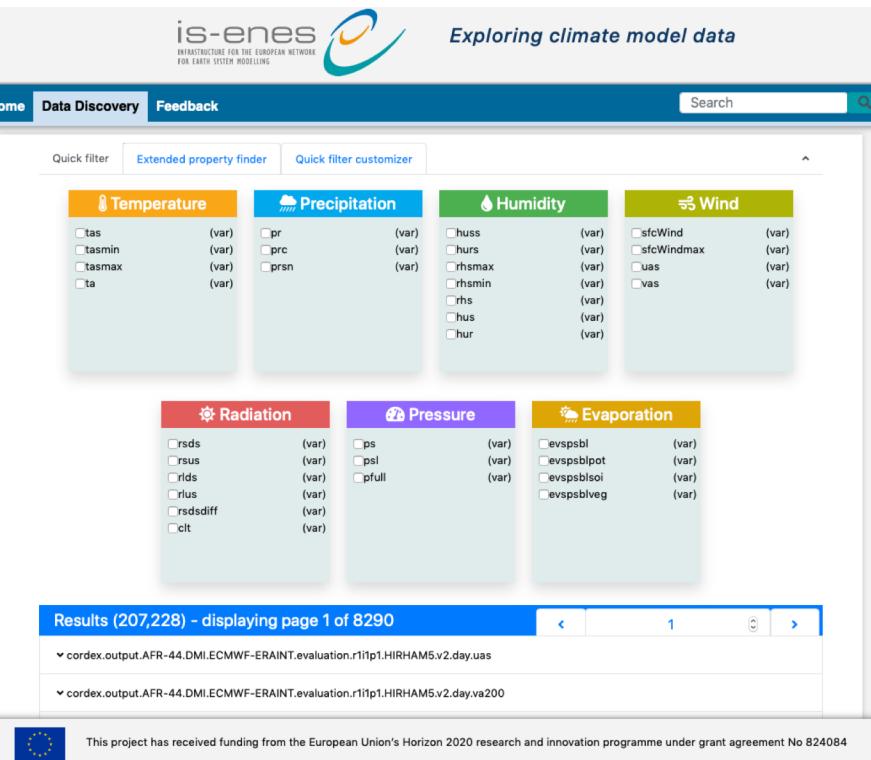
What is the climate4impact portal?

- Platform for researchers to explore climate data and perform analysis
- Connects to ESGF web services
 - Search, Catalog Support, Security
 - Several projects and experiments
- Visualization via ADAGUC Software
 - Visualization system using Web Map Services
 - Web Coverage Services for data transformation
- Analysis using (Py)WPS to perform calculations
 - icclim open-source software for on-demand climate indices calculation
 - Data sub-selection
 - Personal store for processing results
- In production
 - Deployed in the cloud
 - Is one of the official CMIP6 dissemination portals

Research data lifecycle

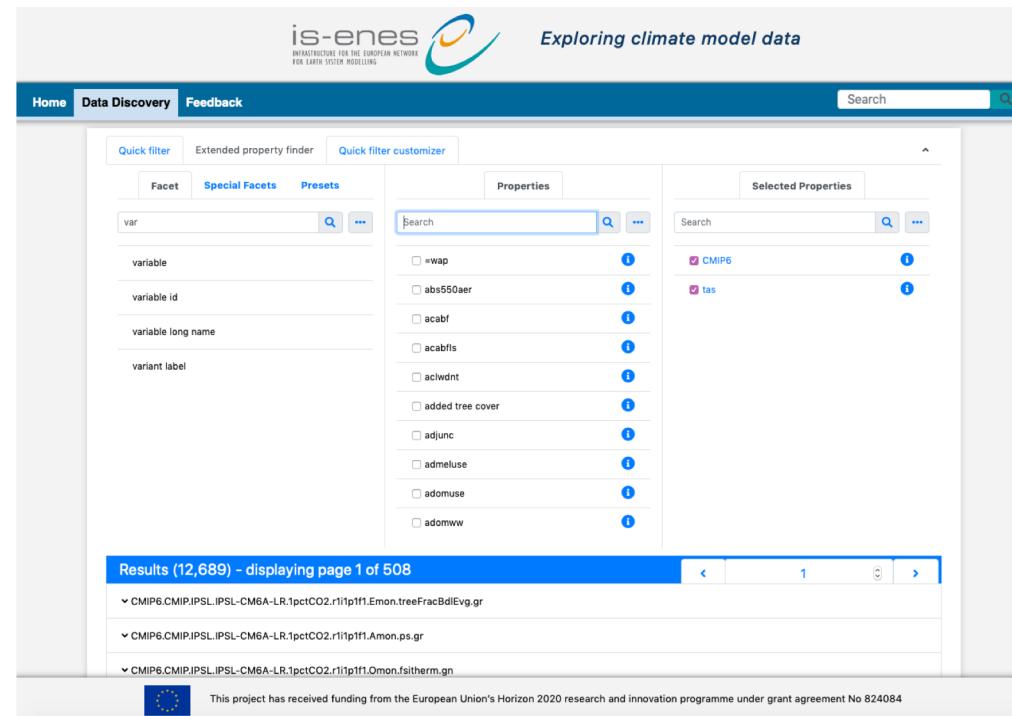


Web based faceted search



The screenshot shows the is-enes Data Discovery interface. At the top, there's a navigation bar with Home, Data Discovery (which is selected), and Feedback. A search bar is at the top right. Below the navigation is a 'Quick filter' section with tabs for Temperature, Precipitation, Humidity, Wind, Radiation, Pressure, and Evaporation. Each tab lists various variables like tas, pr, huss, sfcWind, etc., each with a checkbox and '(var)' next to it. Below this is a results summary: 'Results (207,228) - displaying page 1 of 8290'. A footer note at the bottom left states: 'This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824084'.

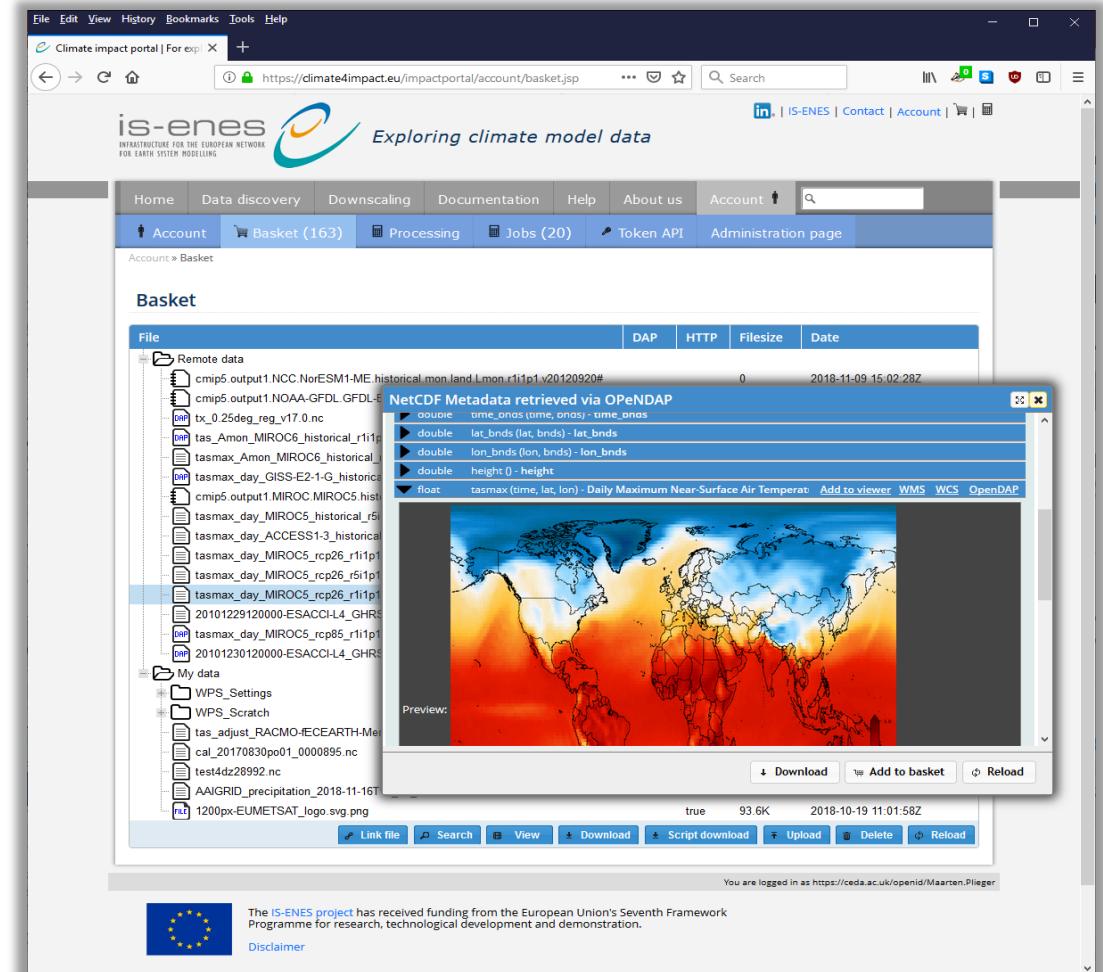
- ▶ Drill down search results
- ▶ Tooltips for acronyms
- ▶ Quick select menus, configurable
- ▶ ES-DOC integration
- ▶ Preview of data
- ▶ Save Search Parameters
- ▶ Export search list to CSV



The screenshot shows the is-enes Data Discovery interface with a different view. It features a 'Facet' section with dropdowns for 'variable', 'variable id', 'variable long name', and 'variant label'. To the right is a 'Properties' section with a search bar and a list of selected properties: CMIP6, tas, and wap. Below this is a results summary: 'Results (12,689) - displaying page 1 of 508'. A footer note at the bottom right states: 'This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824084'.

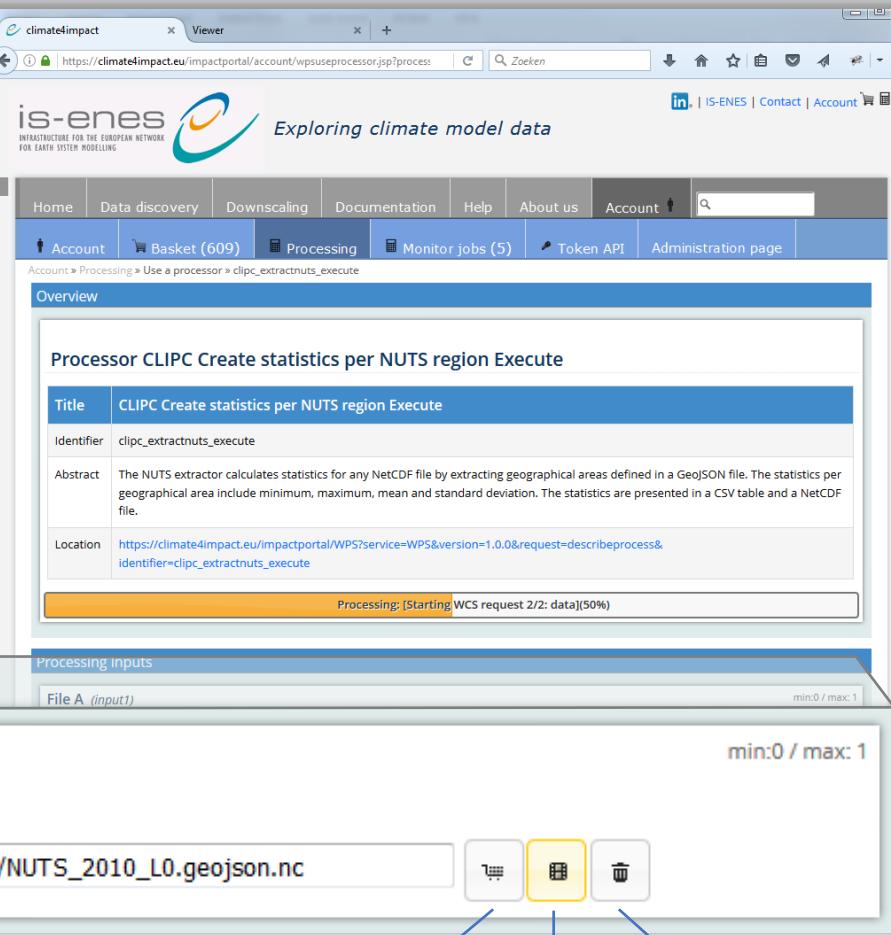
Personal User Space

- By default the basket contains:
 - “Remote data” for links
 - “My data” for your own data
- Script based download allows to select and download multiple files
- The basket allows for uploading your own files
 - Can be used in processing or visualization
 - NetCDF, CSV, GeoJSON, PNG
- Share your data located in your basket with others



Web processing interface for data analysis

- ▶ Generated user interface
- ▶ Lightweight
- ▶ Links to preview
- ▶ Links to basket / cart
- ▶ Get info from input files



The screenshot shows a web browser window displaying the **climate4impact** application. The URL is https://climate4impact.eu/impactportal/account/wpsuseprocessor.jsp?processid=clipc_extractnuts_execute. The page title is "Exploring climate model data". The main content area displays a processor titled "Processor CLIPC Create statistics per NUTS region Execute". The processor has the following details:

- Title:** CLIPC Create statistics per NUTS region Execute
- Identifier:** clipc_extractnuts_execute
- Abstract:** The NUTS extractor calculates statistics for any NetCDF file by extracting geographical areas defined in a GeoJSON file. The statistics per geographical area include minimum, maximum, mean and standard deviation. The statistics are presented in a CSV table and a NetCDF file.
- Location:** https://climate4impact.eu/impactportal/WPS?service=WPS&version=1.0.0&request=describeprocess&identifier=clipc_extractnuts_execute

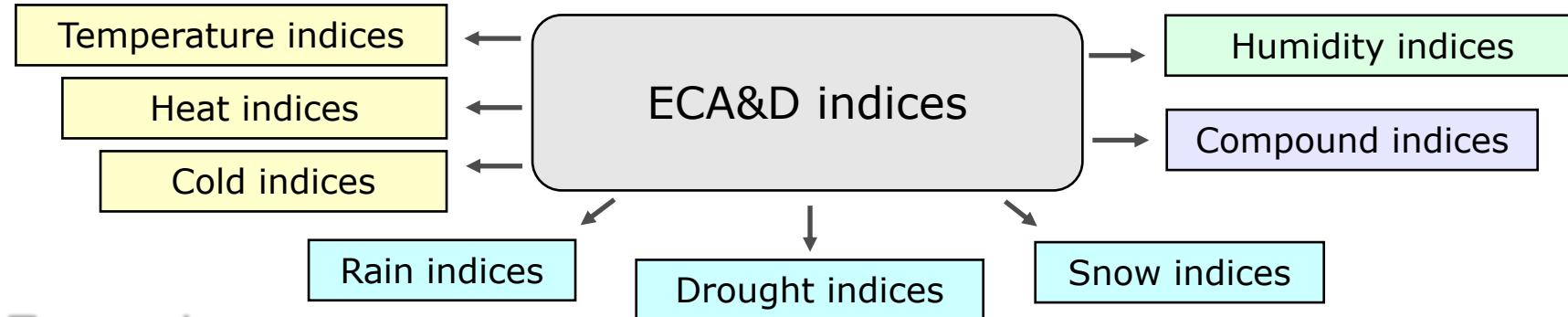
A progress bar at the bottom indicates "Processing: [Starting WCS request 2/2: data(50%)]". Below this, a "Processing inputs" section is shown with a single item:

File A (input1)
application/netcdf
http://opendap.knmi.nl/knmi/thredds/dodsC/CLIPC/storyline_urbanheat/geojson/NUTS_2010_L0.geojson.nc

Annotations with arrows point to specific elements:

- "title" points to the processor title.
- "identifier" points to the processor identifier.
- "abstract" points to the processor abstract.
- "value" points to the input file URL.
- "basket" points to the "Basket" icon in the processing inputs panel.
- "preview" points to the "Preview" icon in the processing inputs panel.
- "delete" points to the "Delete" icon in the processing inputs panel.

Climate Indices using icclim



Examples

- 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, started in September 2013
 - Generic and modular approach, can be reused in other environments
 - C functions called for optimization
- I/O interface is structured for optimal performance, with wrapper functions
- Some percentile-based indices (TG10p, TX10p, TN90p, etc) using bootstrap method

icclim source code and documentation is available via <https://github.com/cerfacs-globc/icclim>

An xarray/dask fork has been done and is now at an alpha stage.

Example: Calculating summer days (SU) 1/3

► Calculate number of days where maximum temperature is above 25 degrees per European country, based on experiment RCP 2.6 and climate model MIROC5

► Sign in

► Go to Search and select:

1. Project: CMIP5
2. Parameter: tasmax
3. Time frequency: daily
4. Experiment: rcp26
5. Model: MIROC5,
6. Ensemble: r1i1p1
7. Select the latest version

► Select a file from the dataset and add it to your basket

The screenshot shows the DARE Data Discovery interface. At the top, there's a navigation bar with links like Home, Data discovery, Downscaling, Documentation, Help, About us, Account, and a search bar. Below that is a toolbar with buttons for Search, Catalogs, Explore your own catalogs or files, Map & Plot, and Processing.

The main area is titled "Filters" and contains several filter panels:

- Quick select Parameter:** Includes Temperature (tas), Min. Temperature (tasmin), Max. Temperature (tasmax), and Air Temperature (ta).
- Precipitation:** Includes Precip. (pr), Conv. Precip. (prc), and Snow (prsn).
- Humidity:** Includes Specific Humidity (huss), Rel. Humidity (hurs), Max. Rel. Humidity (hrmax), Min. Rel. Humidity (hrmin), Rel. Humidity (rhs), Spec. Humidity (hus), and Rel. Humidity (hur).
- Wind:** Includes Wind (sfccwind), Max. Wind (sfccwindmax), E. Wind (uas), and N. Wind (vas).
- Radiation:** Includes SW Radiation Up (rsus), LW Radiation Up (rlus), LW Radiation Dn (rlrd), and Diff. Radiation Dn (rsdsdiff).
- Pressure:** Includes Pressure (ps), SL Pressure (psl), and Pressure (pfull).
- Evaporation:** Includes Act. Evap. (evpsblp), Pot. Evap. (evpsblp0t), Soil Evap. (evpsblsol), and Canopy Evap. (evpsblveg).
- Clouds:** Includes Clouds (clt).

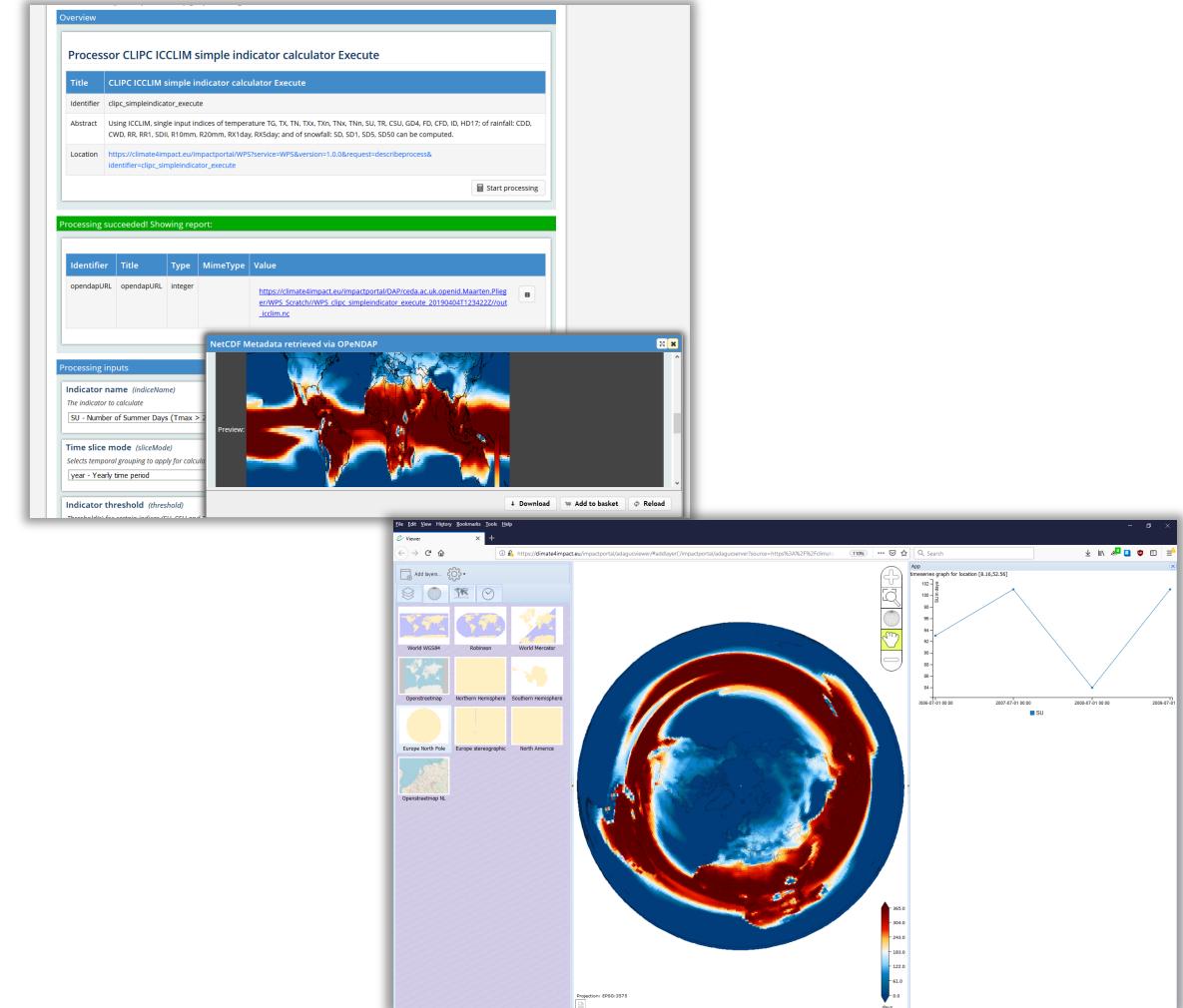
Below the filters is a "Selected filters" section with checkboxes for Project : CMIP5, Parameter : tasmax, Frequency : day, Experiment : rcp26, Model : MIROC5, Ensemble : r1i1p1, and Version : 20161012.

The results section shows "Found 1 datasets. Displaying page 1 of 1." with a preview of the dataset "cmip5.output1.MIROC.MIROC5.rcp26.day.atmos.day.r1i1p1.v20161012". It includes "Previous" and "Next" buttons and an "Export to CSV" link.

At the bottom right, there's a "NetCDF Metadata retrieved via OPeNDAP" panel showing a world map with temperature data and metadata details.

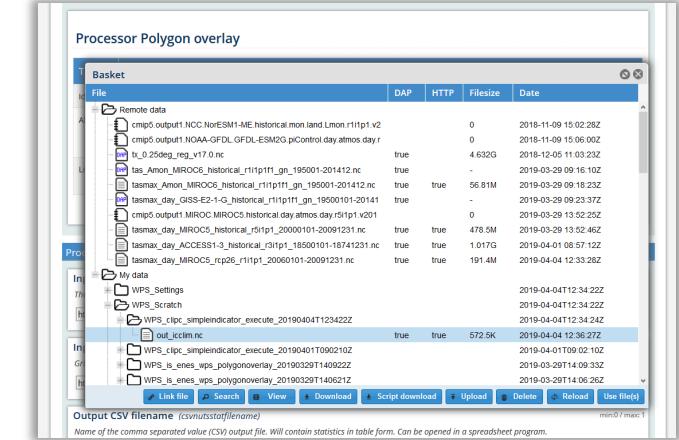
Example: Calculating summer days (SU) 2/3

- Go to Processing and select:
icclim simple indicator calculations
- Select SU, Summer days. Leave
the threshold to 25 degrees
Celsius
- Select the file from your basket
and click “Start processing”
- Visualize the output

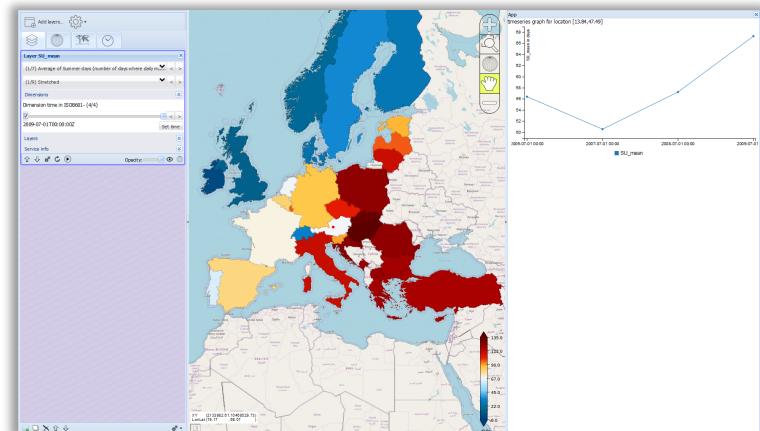


Example: Calculating summer days (SU) 3/3

- Go to Processing and select Polygon overlay
- For “Input File B - Gridded data”, choose the latest result with SU from your basket. This is the most recent folder under WPS_Scratch
- As variable select “SU”, as time range select “*”
- Click “Start processing”



Results: Summer days per European country for MIROC5 / RCP26 !



What can be improved?

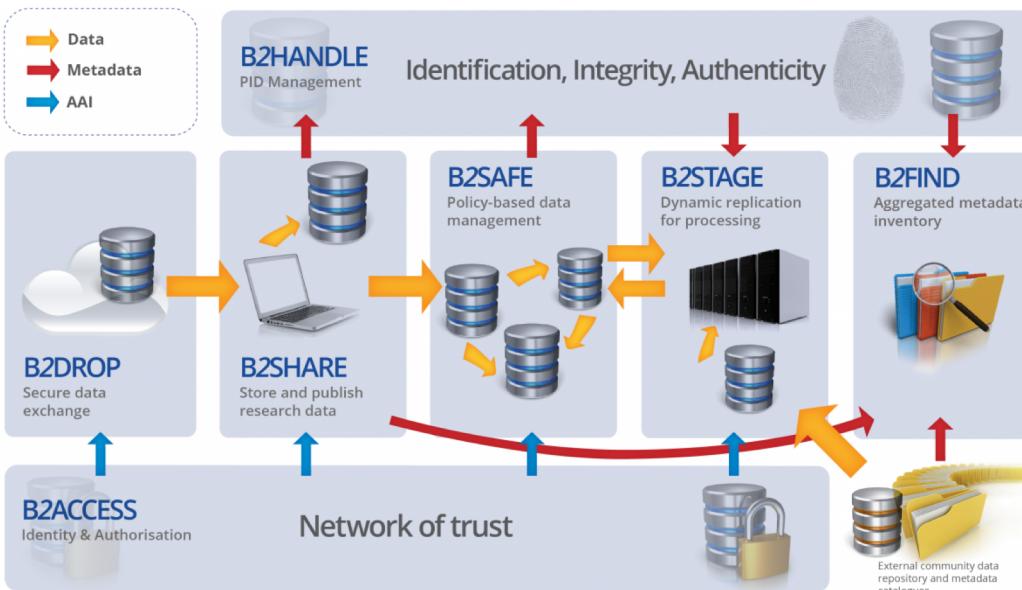
- ▶ Currently C4I handles ESGF data on file level
 - ▶ X Fragmentation of files is a barrier for many users and hurts user experience
 - ▶ → Hide file structure, work with datasets and search patterns
 - ▶ → Especially important, because now C4I is one of the official data distribution endpoints
- ▶ Currently the Processing services are on the same machine as the portal
 - ▶ X Currently not scalable and processing load effects the portal
 - ▶ → Make use of distributed Web Processing Services using delegation Currently in Progress!! (next slides)
- ▶ Currently the frontend uses old technologies (JSP, Jquery, ExtJS)
 - ▶ X Difficult to maintain, and it is difficult to re-use results from other work
 - ▶ → Migrate to ReactJS (Based on work done in the project C3S-Magic)
 - ▶ → Good moment to re-design the user interface in collaboration with users. Currently in Progress too!!
- ▶ Currently the viewer is running in a separate tab
 - ▶ By using ReactJS, it is easier to make use of an embedded viewer (adaguc-webmapjs)
- ▶ Currently provenance tracking is limited to a few processes
 - ▶ Enhance usage of W3C PROV-DM standard and WPS_PROV toolkit
 - ▶ We are looking for users who are willing to help to improve the platform!

European Landscape & Components

EUDAT & EOSC

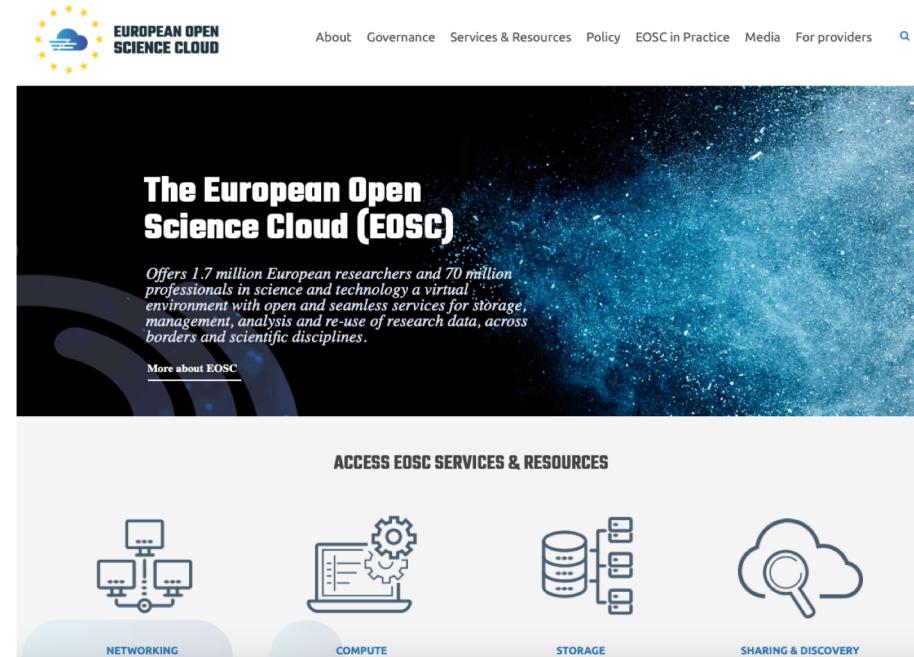
EUDAT CDI B2 Service Suite

- ▶ Integrated B2 Services
- ▶ B2ACCESS: Common AAI
- ▶ B2DROP: Secure Data Exchange
- ▶ Interface between EUDAT B2 Services and Communities infrastructures, such as Climate



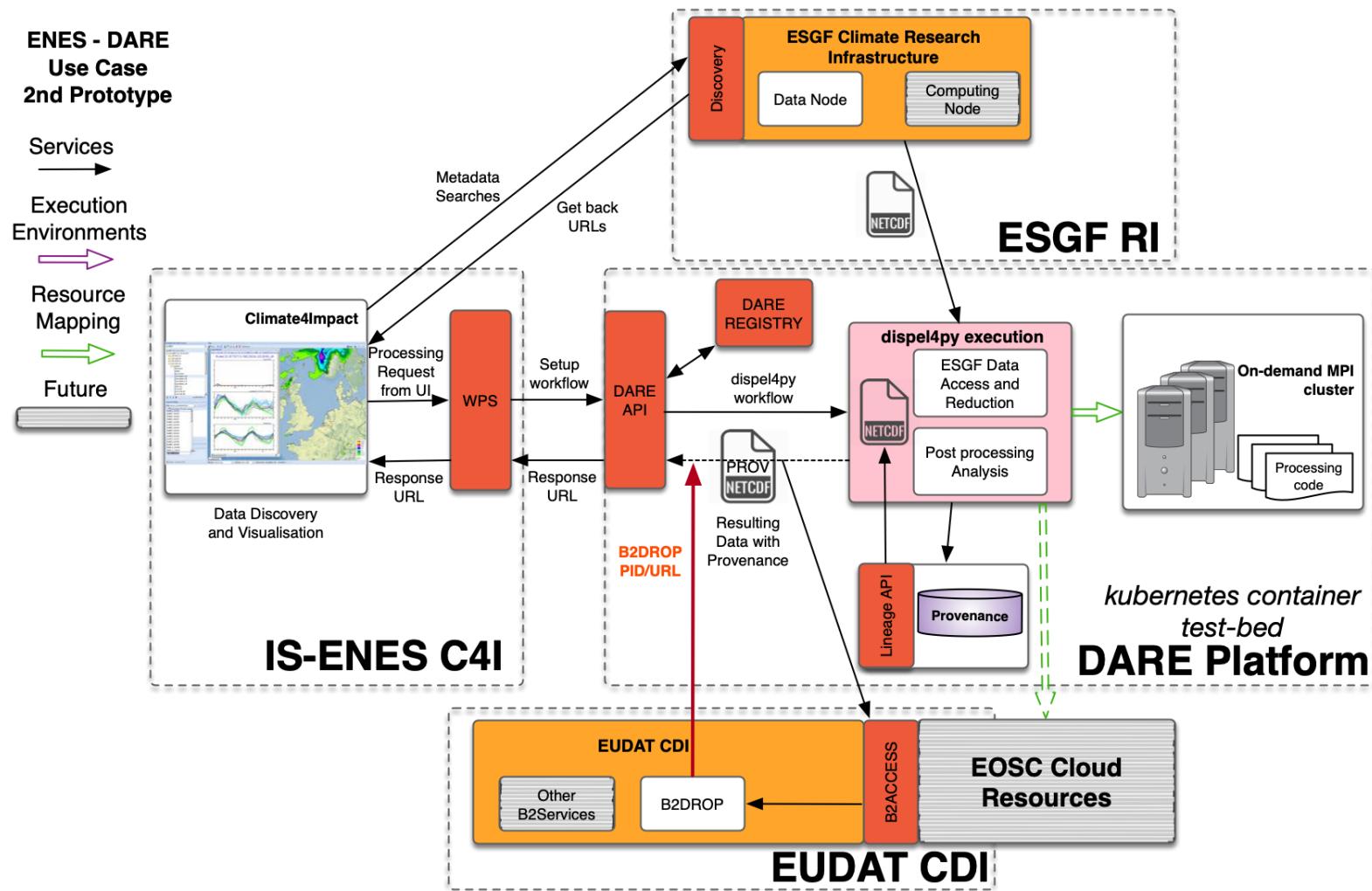
European Open Science Cloud (EOSC)

- ▶ Marketplace of Services
 - ▶ Compute
 - ▶ Storage
 - ▶ Sharing, etc.



DARE IS-ENES

Processing Delegation: Prototype Version



Summary

- C4I can provide Climate change impact researchers a better access to climate data
 - Handle very large datasets
 - Remote Data Processing with Provenance Information
 - Guidance and Help
- C4I is an official access point for new (CMIP6) climate scenarios
- A new C4I 2.0 is in active development
- Hiding e-infrastructures heterogeneity accelerate C4I developments
 - EUDAT and EOSC will provide needed services in the backend
 - The DARE Platform and its API will ease the use of different e-infrastructure and cloud services



Questions & Comments! ☺

<https://climate4impact.eu>



christian.page@cerfacs.fr

On Behalf of the climate4impact Team

**Gateways Poster 6A this evening for further discussions
on this topic!**