CERFACS

CENTRE EUROPÉEN DE RECHERCHE ET DE FORMATION AVANCÉE EN CALCUL SCIENTIFIQUE

is-enes

INFRASTRUCTURE FOR THE EUROPEAN NETWORK FOR EARTH SYSTEM MODELLING

Accessing Climate Data and Tools for End Users

A Practical Overview for Users

Christian Pagé, CERFACS, Toulouse, France

IS-ENES3 virtual Autumn School on Climate data use for impact assessments

4th Nov. - 11th Dec. 2020

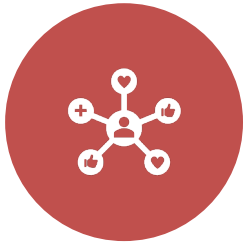
The IS-ENES3 project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824084.

www.cerfacs.fr

ESGF Infrastructure

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

Introduction



The Earth System Grid Federation (ESGF) is a globally distributed e-infrastructure for the hosting and dissemination of climate-related data.

ESGF was originally developed to support CMIP5 (5th Coupled Model Intercomparison Project) ...

Provide a means for climate data users to access and analyse the data output

For 5th Assessment report made by the IPCC (Intergovernmental Panel on Climate Change).

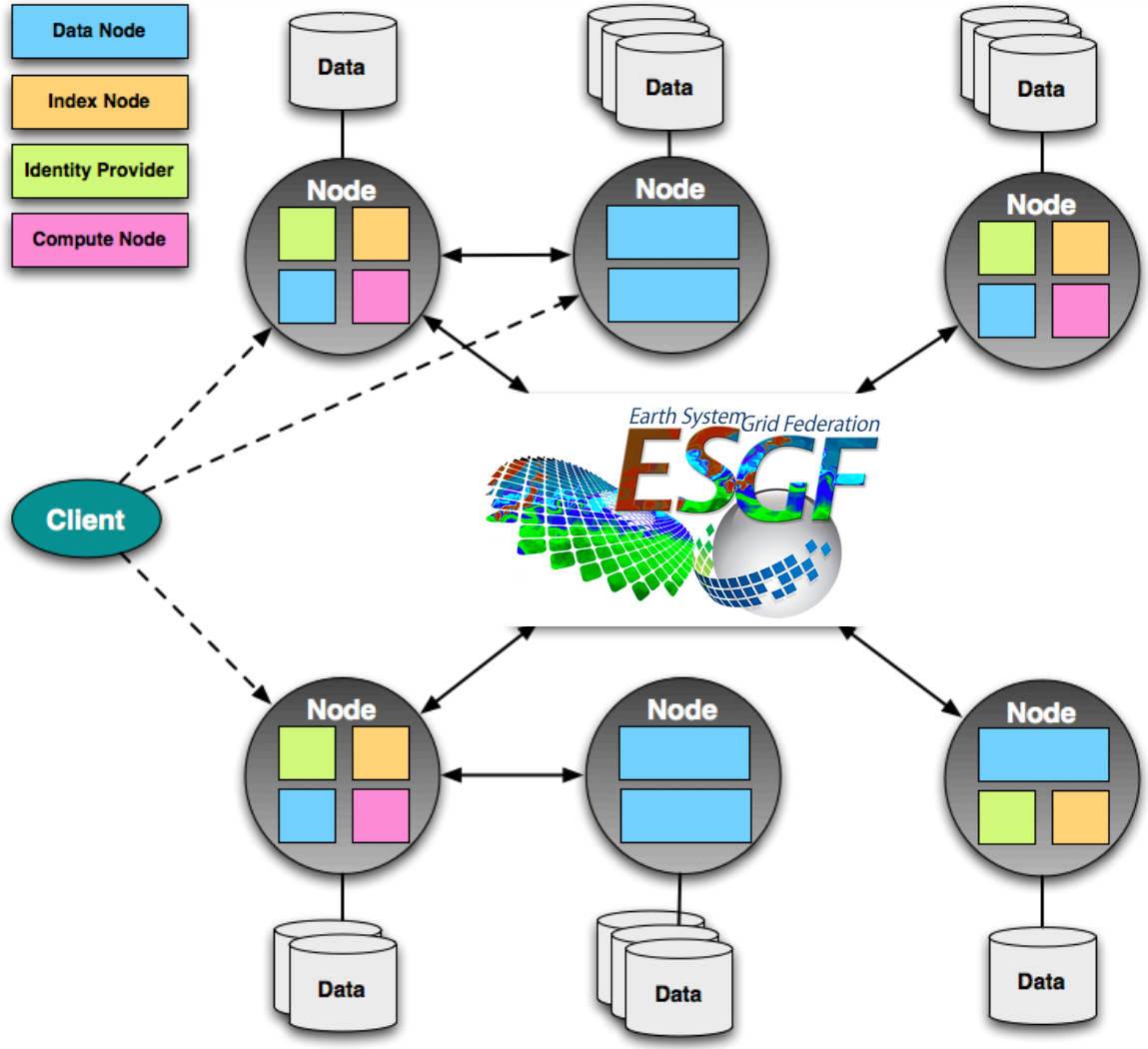
Ten years of Operations: History and Evolution

ESGF has grown to support over
25000 registered users

Besides the CMIPs, supports a range of other projects such as:

- Energy Exascale Earth System Model
- Obs4MIPS
- CORDEX
- European Space Agency's Climate Change Initiative Open Data Portal.

Important experience gathered over the years about community collaboration for a distributed infrastructure - operational procedures and governance

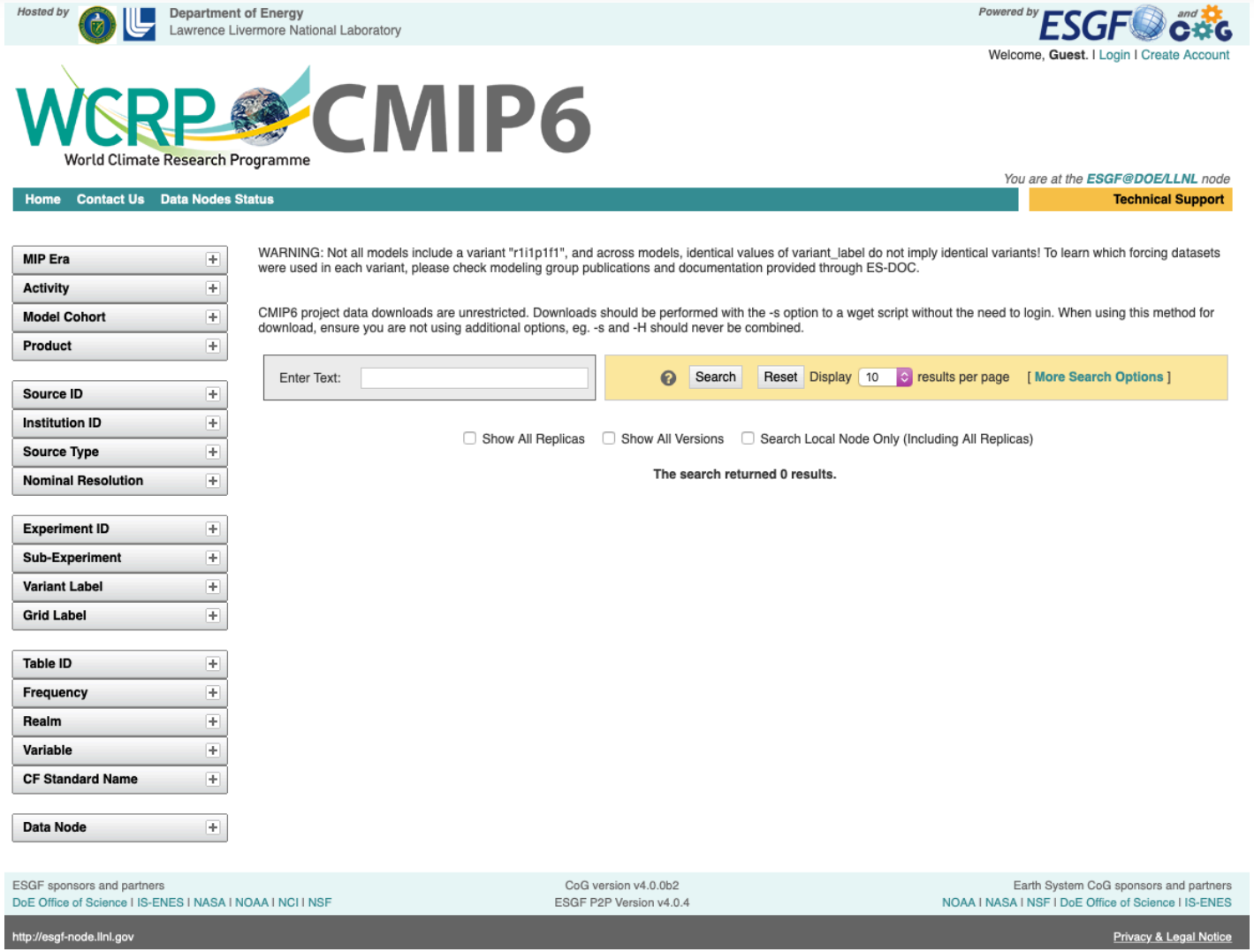


ESGF: Current Data Holdings

ESGF opened for CMIP6 data in June 2018

Data holdings:
~ 218 Experiments
~ 37 Institutions
~ 94 CMIP6 Models
~ 30 Data Nodes

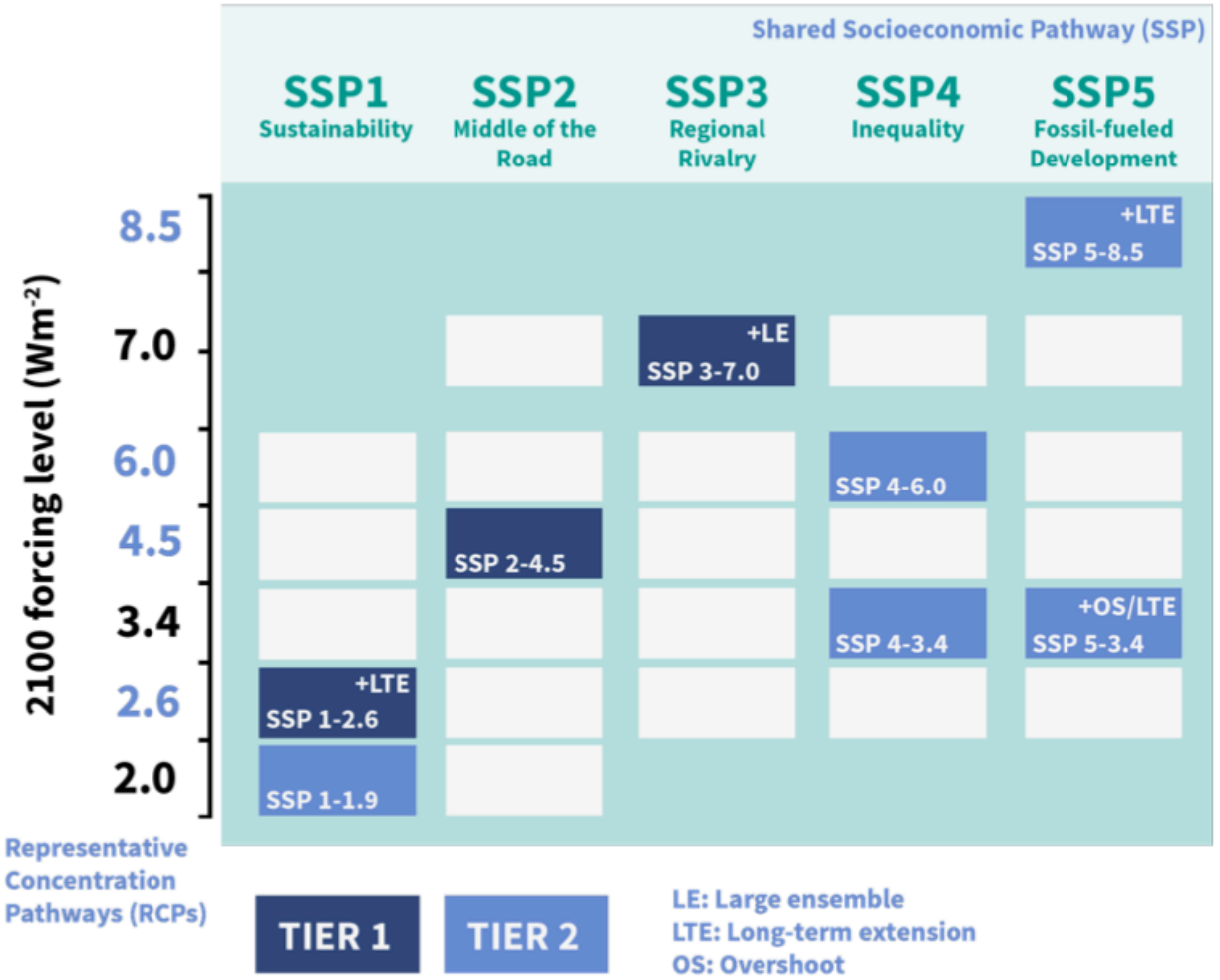
~ 18 Petabytes Expected



CMIP6 GES Scenarios

CMIP6

- Now we have SSPs instead of CMIP5 RCPs and CMIP3 SRES
- Here we have the correspondence between RCPs and SSPs



See: O'Neill et al., The Scenario Model Intercomparison Project (ScenarioMIP) for CMIP6, GMD, 2016

Accessing CMIP6 data using ESGF

CMIP6

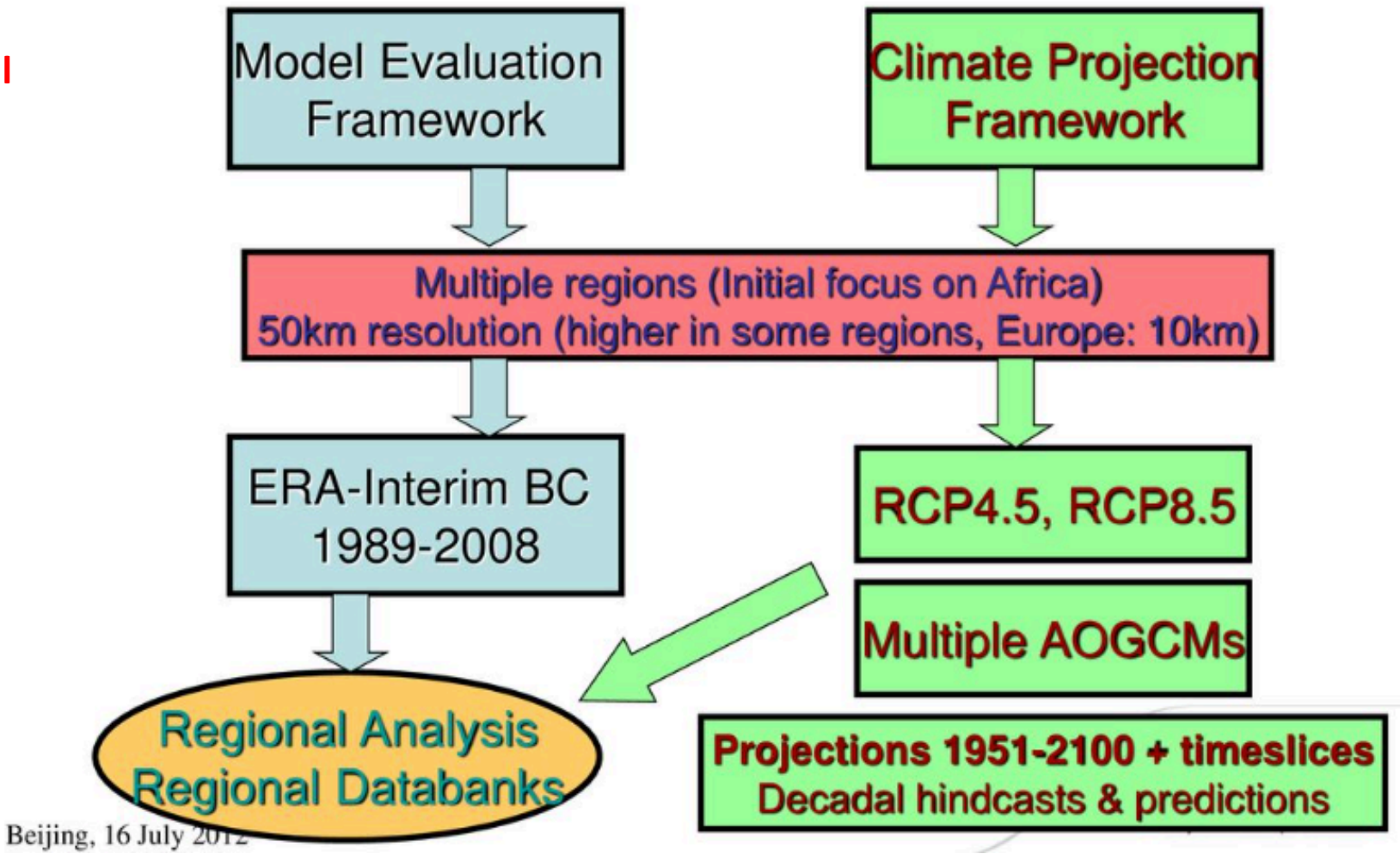
- Accessing datasets using the ESGF Data Nodes

https://esgf-node.llnl.gov/search/cmip6/

CORDEX Experiments

⚠ ⚠ ⚠

CORDEX datasets still CMIP5-based only



CORDEX Data

⚠ ⚠ ⚠

CORDEX datasets still CMIP5-based only

https://esgf-node.llnl.gov/projects/esgf-llnl/

Hosted by Department of Energy, Lawrence Livermore National Laboratory

Powered by ESGF and CoG

Welcome, **Guest**. | Login | Create Account

Lawrence Livermore National Laboratory

*You are at the **ESGF@DOE/LLNL** node*

Home | Contact Us | Data Nodes Status | Technical Support

ESGF@DOE/LLNL

Search & Download Data

Simple Text Search | More search options

The Earth System Grid Federation (ESGF) is a collaboration that develops, deploys and maintains software infrastructure for the management, dissemination, and analysis of model output and observational data. ESGF's primary goal is to facilitate advancements in Earth System Science. It is an interagency and international effort led by the Department of Energy (DOE), and co-funded by National Aeronautics and Space Administration (NASA), National Oceanic and Atmospheric Administration (NOAA), National Science Foundation (NSF), and international laboratories such as the Max Planck Institute for Meteorology (MPI-M) German Climate Computing Centre (DKRZ), the Australian National University (ANU) National Computational Infrastructure (NCI), Institut Pierre-Simon Laplace (IPSL), and the British Atmospheric Data Center (BADC). The ESGF mission is to:

- Support current CMIP activities, and prepare for future assessments
- Develop data and metadata facilities for inclusion of observations and reanalysis products for CMIP use
- Enhance and improve current climate research infrastructure capabilities through involvement of the software development community and through adherence to sound software principles
- Foster collaboration across agency and political boundaries
- Integrate and interoperate with other software designed to meet the objectives of ESGF: e.g., software developed by NASA, NOAA, ESIP, and the European IS-ENES
- Create software infrastructure and tools that facilitate scientific advancements

ESGF is a component architecture expressly designed to handle large-scale data management for worldwide distribution. The team of computer scientists and climate scientists has developed an operational system for serving climate data from multiple locations and sources. Model simulations, satellite observations, and reanalysis products are all being served from the ESGF P2P distributed data archive.

ESGF portals like this one provide access to the output of the climate models contributing to the next assessment report of the Intergovernmental Panel on Climate Change IPCC through the Coupled Model Intercomparion Project CMIP. The WGCM Infrastructure Panel is the official CMIP document home.

Federated ESGF-CoG Nodes

CoG-CU
ESGF@CEDA
ESGF@DKRZ
ESGF@IPSL
ESGF@NASA/NCCS
ESGF@NCI
ESGF@NOAA/GFDL
ESGF@NSC/LIU
ESGF@PIK

Browse Projects

This | All | My | Tags

Parent projects (0)

Peer projects (7)

ESGF-CEDA
ESGF-DKRZ
ESGF-GFDL
esgf-gsfc
ESGF-IPSL
ESGF-LIU
ESGF-PIK

Child projects (8)

CMIP3
CMIP5
cmip6
create-ip
E3SM
input4MIPs
obs4MIPs

Enter Tag

Reset | Go

Start typing, or use the 'Delete' key to show all available tags.

***ESGF-LLNL Tags:** None*

Search Data

The following projects are hosted on LLNL data or index nodes, and these are unrestricted. Downloads should be performed with the -s option to a wget script without the need to login. When using this method for download, ensure you are not using additional options eg. -s and -H should never be combined.

CORDEX-Adjust: be careful

High CAUTION

Bias-adjusted RCM data

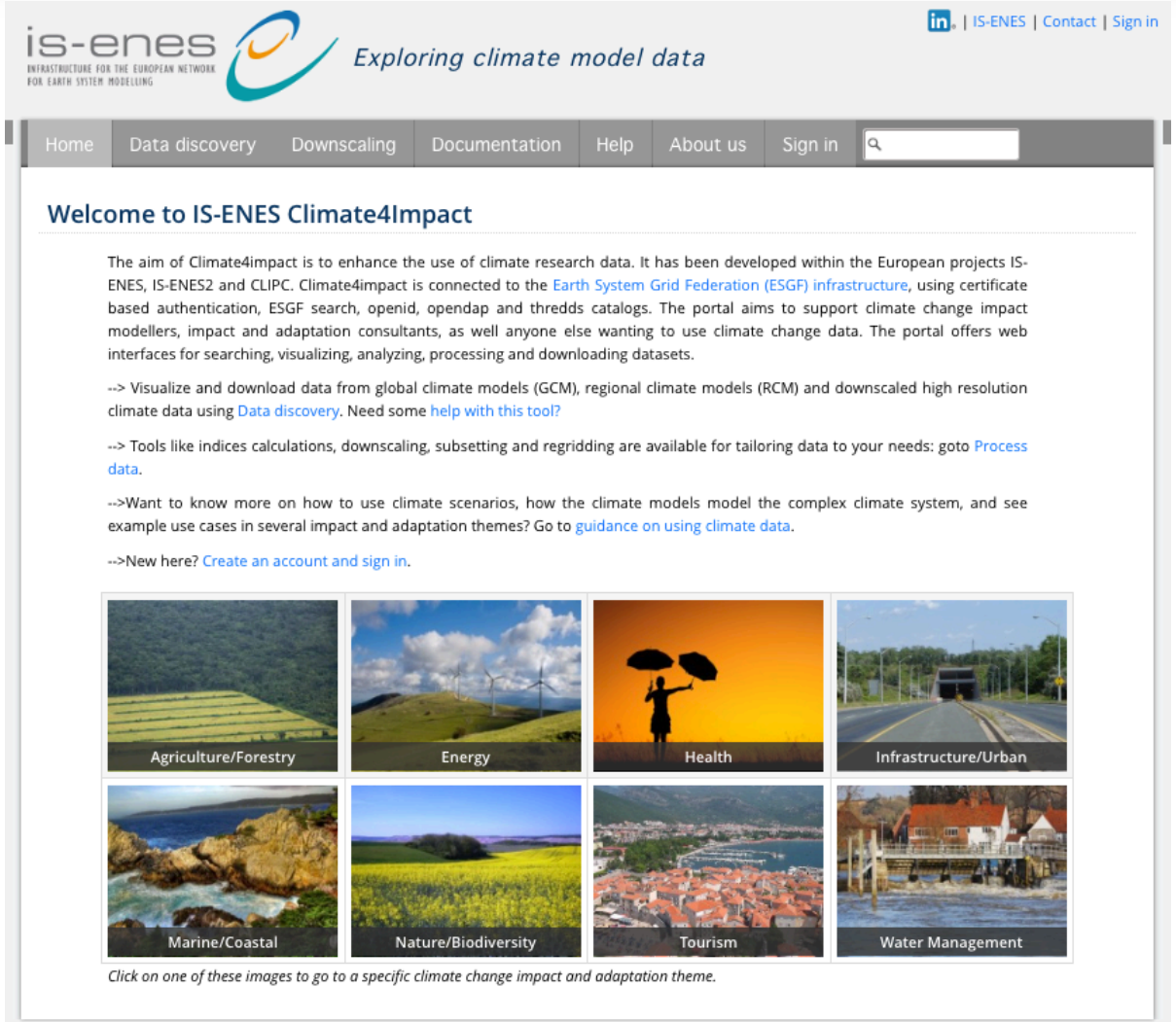
Note that even if bias adjustment (or bias correction) is widely used it is still a controversial approach with its own pros and cons. Bias-adjusted CORDEX simulations should be used carefully with full understanding of all potential limitations of the bias adjustment approach.

- Bias may not be removed completely
- When using only a bias-adjustment based on averages, **the extremes may still be over/underestimated**
- When an adjustment is performed per climate variable **this may reduce the consistencies between climate variables**
- The **bias may be different for different regions**, so adjustment should take this into account.
- The reference dataset determines strongly the bias and how it is adjusted. **If this reference dataset also contains biases, then the bias in the climate model runs can not be removed.**

IS-ENES climate4impact (C4I)

https://climate4impact.eu (also called C4I)

- Access all ESGF distributed datasets
  - CMIP5
  - CMIP6
  - CORDEX
  - CORDEX-Adjust
  - ...
- Exhaustive Documentation/Guidance
  - Climate System
  - Climate Numerical Modelling
  - Downscaling techniques
  - Case studies
- On-demand Processing Capabilities

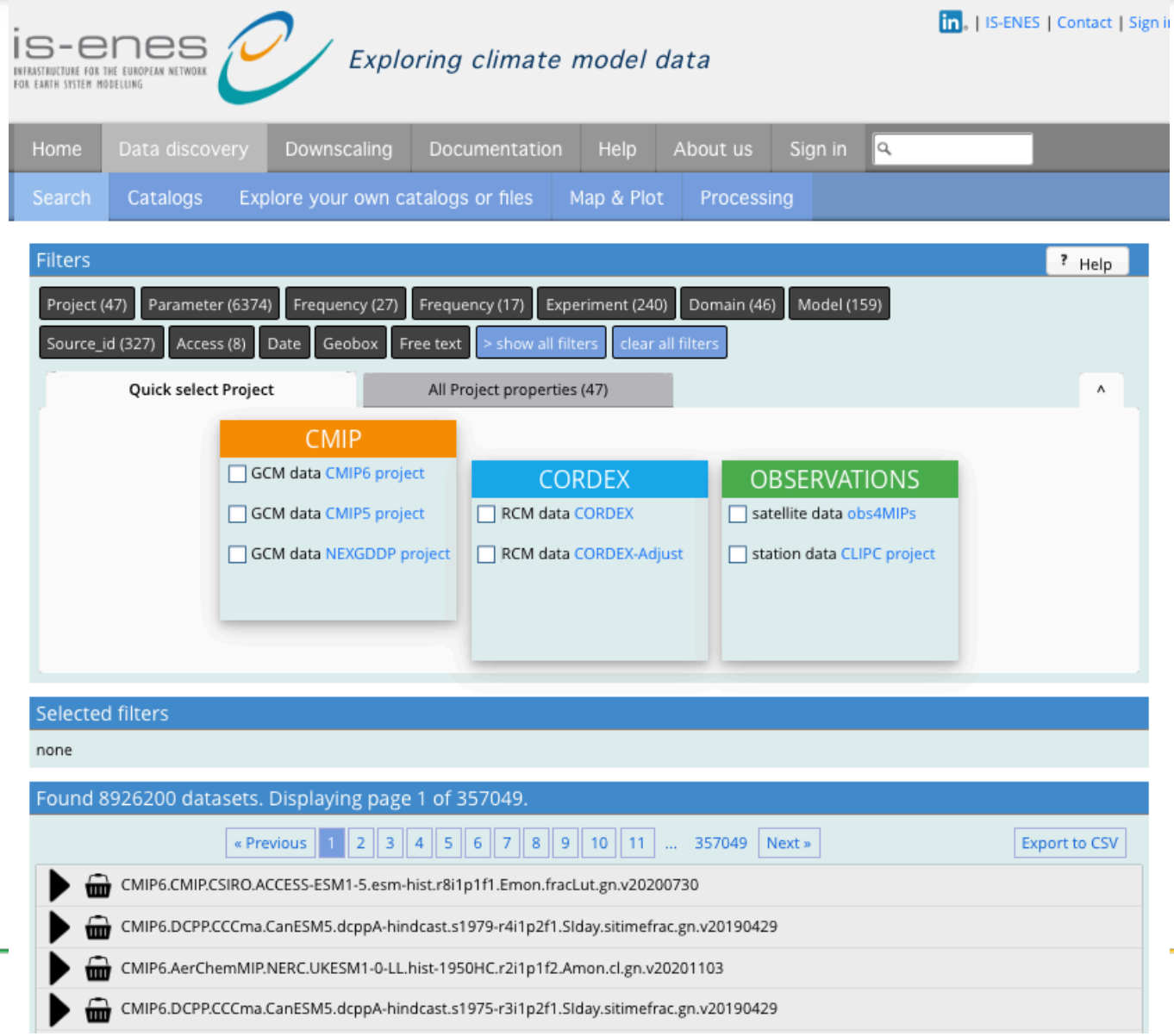


IS-ENES climate4impact: Search Interface

https://climate4impact.eu

- Search Interface with facets with Quick Selects
- Explore and use your own datasets
- On-demand Map & Plots
- On-demand Processing

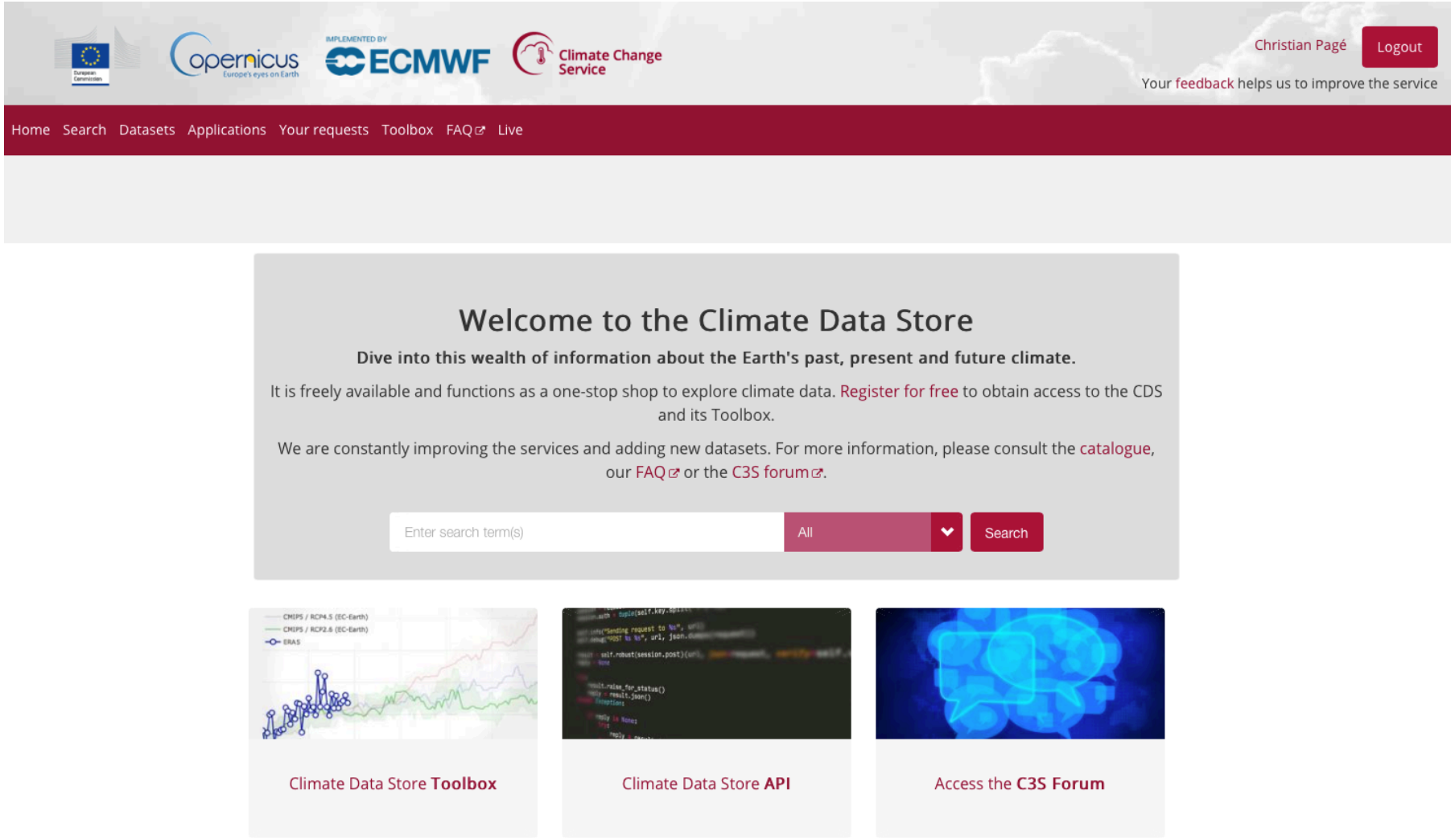
A completely updated C4I portal is being developed and more will be explained next week!



Copernicus Climate Data Store

Copernicus CDS
https://cds.climate.copernicus.eu

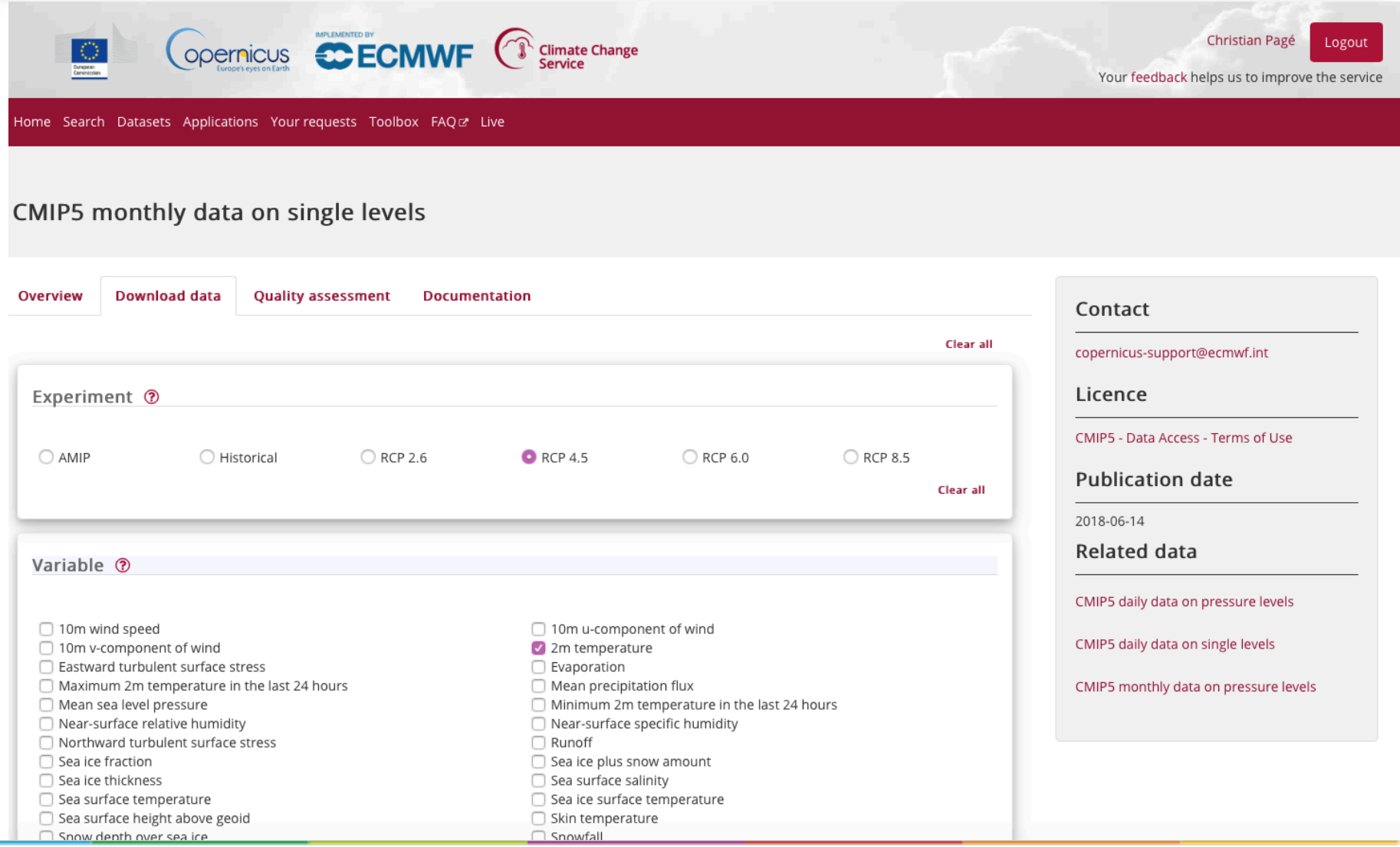
- Operational Service
- Toolbox, API and Applications
- Climate-Service like approach
- CMIP5, CORDEX, EOBS, ... most used datasets are available (not all experiments)
- CMIP6 datasets not yet available
- Not all experiments/models are available, a selection is made
- There is a quality assurance
- Support is available



Copernicus Climate Data Store: Download data

Copernicus CDS
https://cds.climate.copernicus.eu

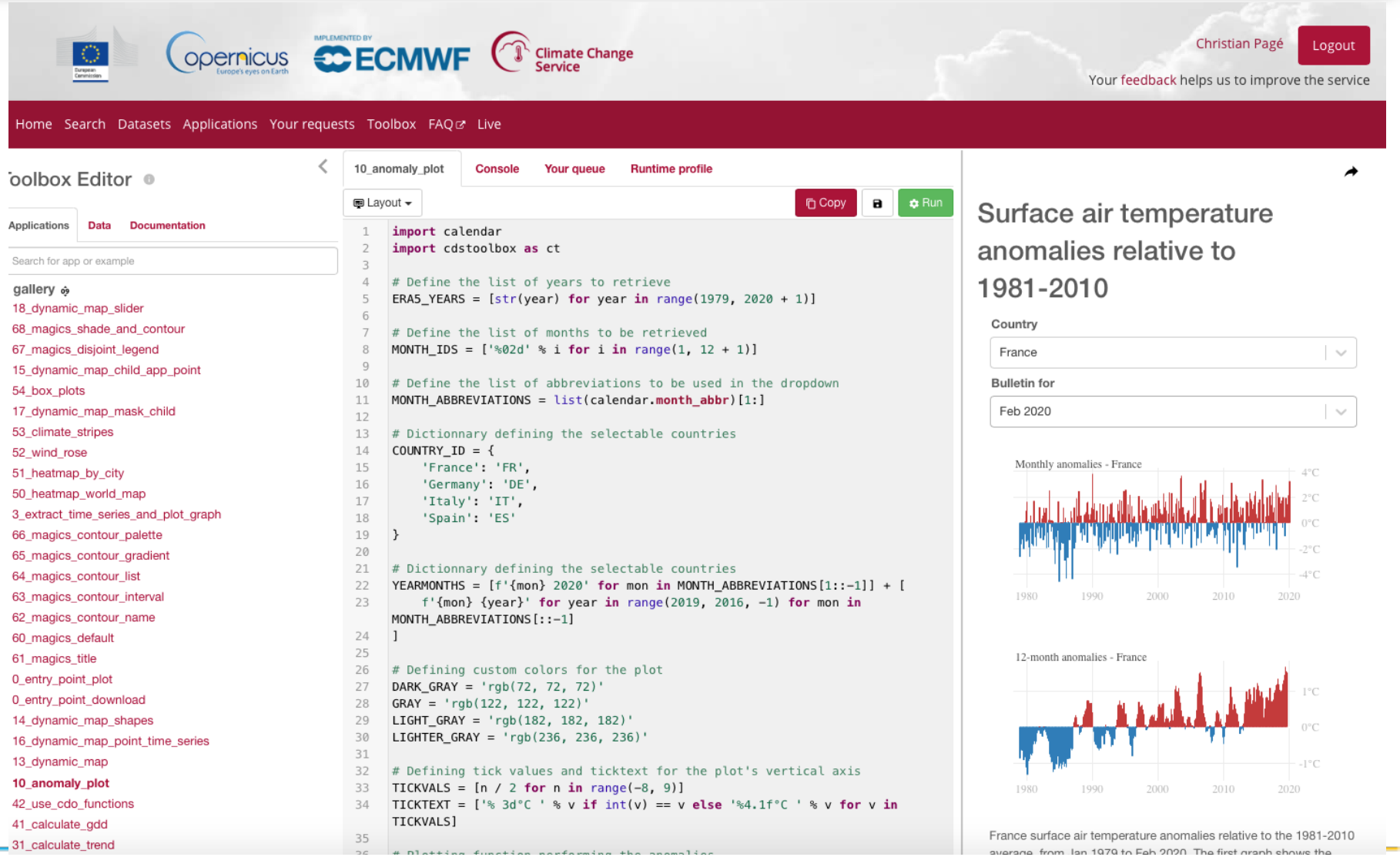
- Example for CMIP5 data selection
- Tutorials are available in the User Learning Service



Copernicus Climate Data Store: Toolbox

Copernicus CDS
https://cds.climate.copernicus.eu

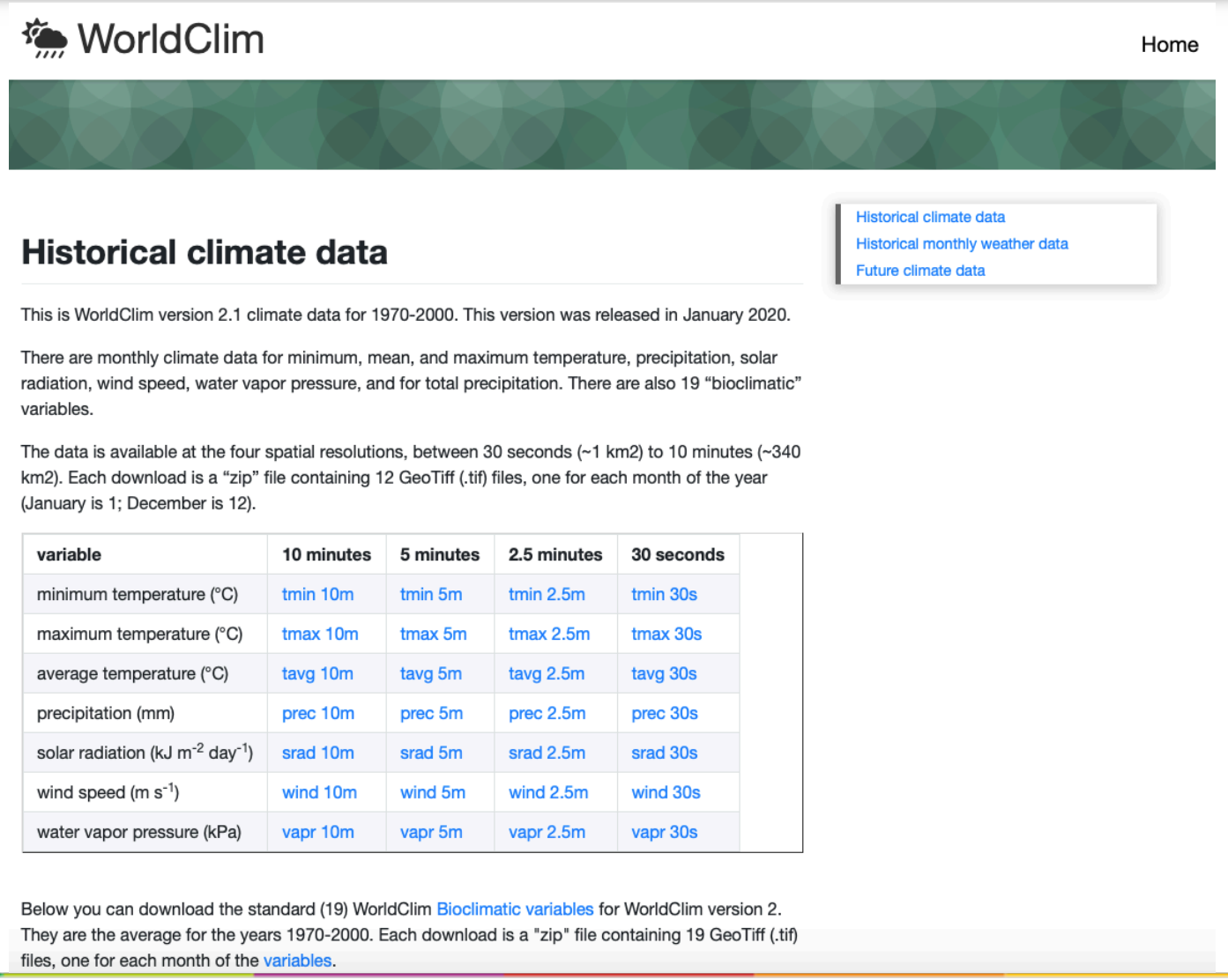
- Toolbox using python code
- Processing done remotely
- Applications Community Sharing
- Can do the same using API for programming applications



WorldClim: basic methodology

WorldClim
https://www.worldclim.org

- Fine resolution observation and future projections
- Very basic scientific downscaling methodology
  - No corrections for inhomogeneities
  - Interpolation using elevation and distance
- Difficult to assess pros and cons of datasets
- Very sparse documentation and information on the methods used
- Be very cautious ⚠⚠⚠⚠ know what you are using.... because data may not be appropriate and usable for your application



WorldClim

Home

Historical climate data
Historical monthly weather data
Future climate data

Historical climate data

This is WorldClim version 2.1 climate data for 1970-2000. This version was released in January 2020.

There are monthly climate data for minimum, mean, and maximum temperature, precipitation, solar radiation, wind speed, water vapor pressure, and for total precipitation. There are also 19 "bioclimatic" variables.

The data is available at the four spatial resolutions, between 30 seconds (~1 km2) to 10 minutes (~340 km2). Each download is a "zip" file containing 12 GeoTiff (.tif) files, one for each month of the year (January is 1; December is 12).

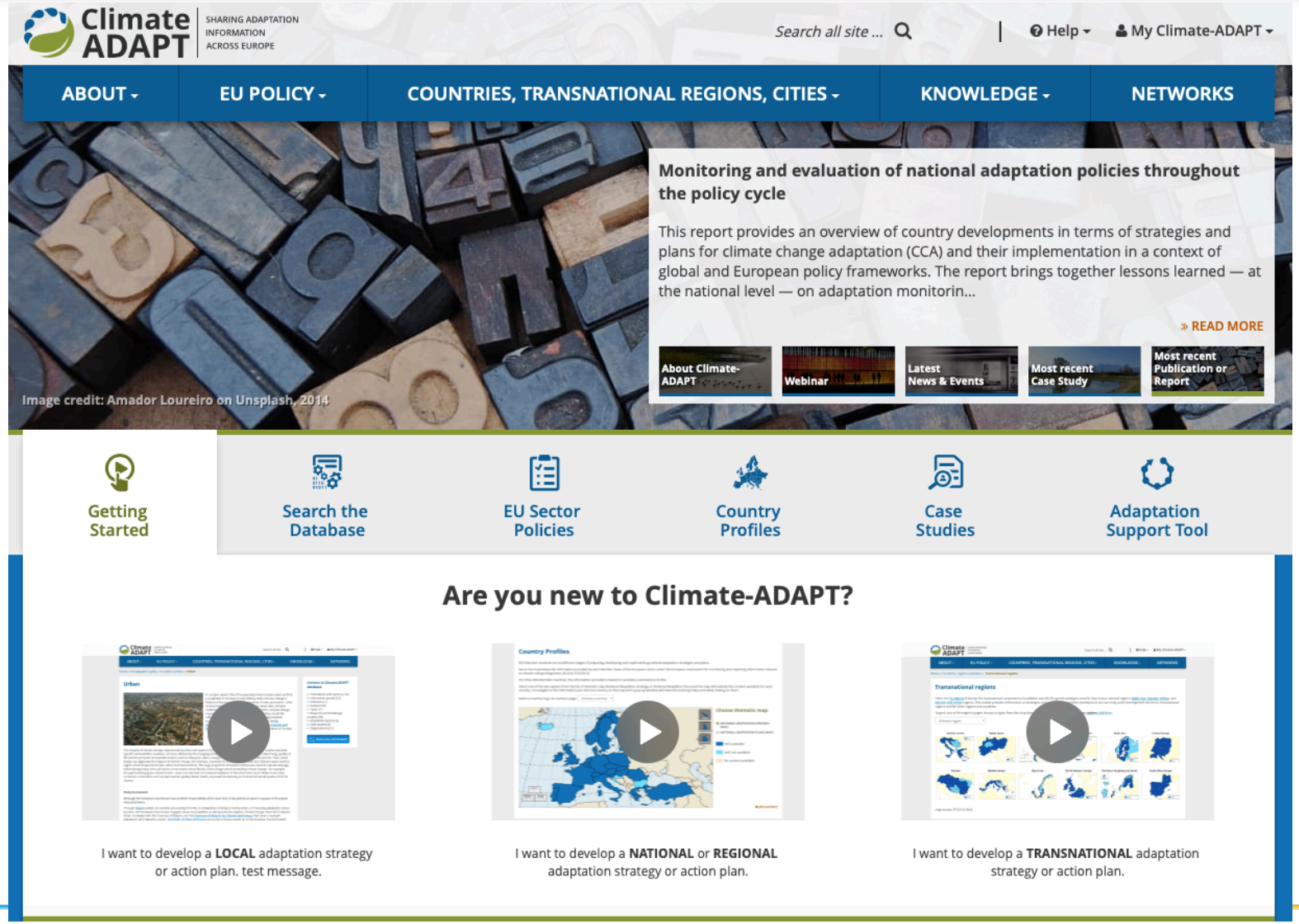
variable	10 minutes	5 minutes	2.5 minutes	30 seconds
minimum temperature (°C)	tmin 10m	tmin 5m	tmin 2.5m	tmin 30s
maximum temperature (°C)	tmax 10m	tmax 5m	tmax 2.5m	tmax 30s
average temperature (°C)	tavg 10m	tavg 5m	tavg 2.5m	tavg 30s
precipitation (mm)	prec 10m	prec 5m	prec 2.5m	prec 30s
solar radiation (kJ m^{-2} day^{-1})	srad 10m	srad 5m	srad 2.5m	srad 30s
wind speed (m s^{-1})	wind 10m	wind 5m	wind 2.5m	wind 30s
water vapor pressure (kPa)	vapr 10m	vapr 5m	vapr 2.5m	vapr 30s

Below you can download the standard (19) WorldClim Bioclimatic variables for WorldClim version 2. They are the average for the years 1970-2000. Each download is a "zip" file containing 19 GeoTiff (.tif) files, one for each month of the variables.

Climate ADAPT: Resources on how to do adaptation plans

Climate ADAPT (EEA)
https://climate-adapt.eea.europa.eu

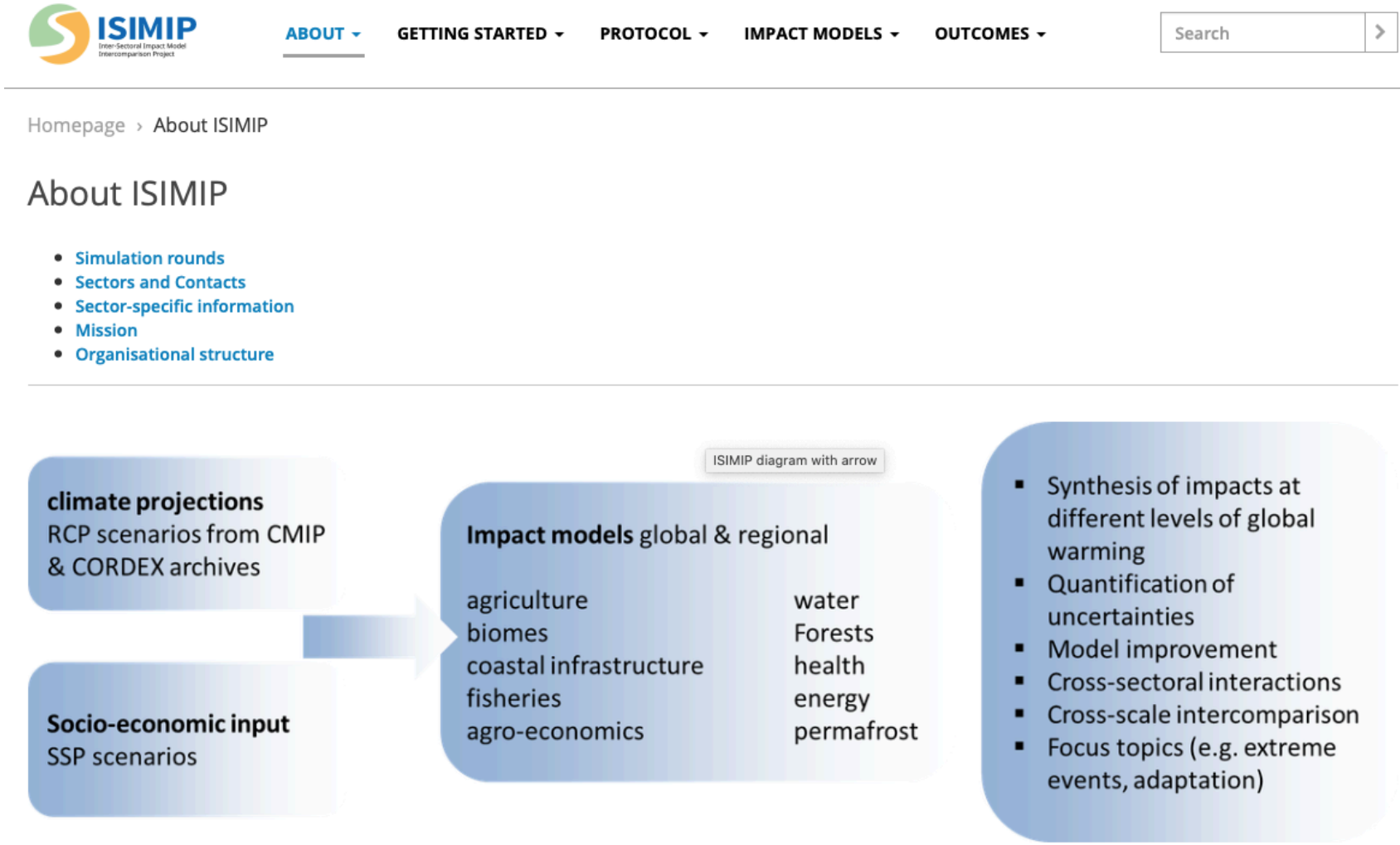
- Large documentation and guidance
- Explore other adaptation studies
- Guidance on how to develop adaptation strategies
- Not a data portal but give links to external resources



ISIMIP: Impact Model Intercomparison

ISIMIP
https://www.isimip.org

- Inter-Sectoral Impact Model Intercomparison Project
- Large number of impact modelling datasets
- Guidance on how to run impact models
- Access to outputs of the impact models through the ESGF data nodes



ISIMIP: Impact Model Intercomparison

ISIMIP
https://www.isimip.org

- Inter-Sectoral Impact Model Intercomparison Project

GETTING STARTED

Contact information, information for impacts modellers, including simulation protocol, input data, and ISIMIP newsletter.

IMPACT MODELS

Searchable and downloadable database of all climate-impact models participating in ISIMIP, with technical summaries, key characteristics and contact person.

OUTPUT DATA

See which climate-impact data have been generated within ISIMIP on a sector-by-sector basis, inluding a log of changes to uploaded data.

OUTCOMES

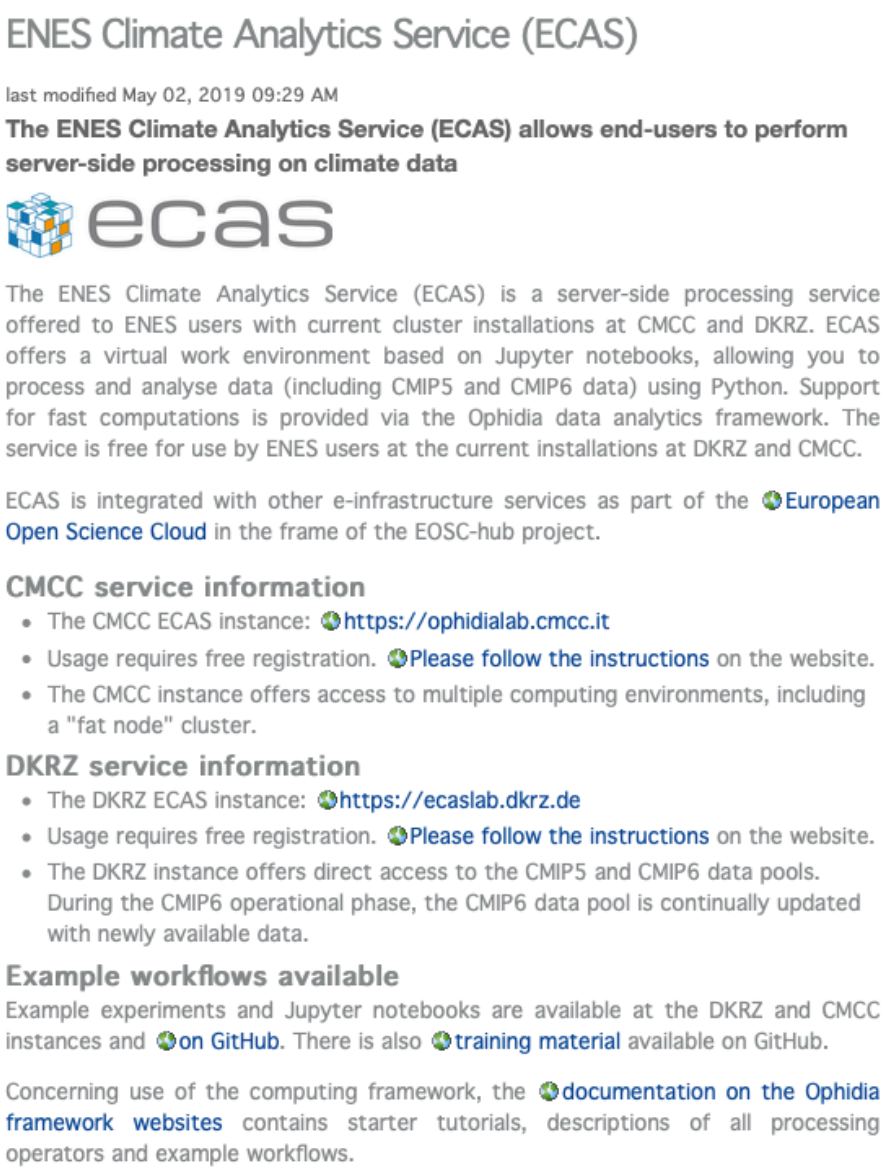
Papers, reports and other products based on ISIMIP climate-impacts simulations. Links to past ISIMIP workshops and events.

ENES Climate Analytics Service (ECAS)

https://portal.enes.org/data/data-metadata-service/processing/ecas

IS-ENES ECAS

- Remote Calculations
- Two endpoints: CMCC (Italy) and DKRZ (Germany)
- Significant computing resources
- Free registration
- Example workflows, python-based using Jupyter notebooks
- Training material available



ENES Climate Analytics Service (ECAS)

last modified May 02, 2019 09:29 AM

The ENES Climate Analytics Service (ECAS) allows end-users to perform server-side processing on climate data

ecas

The ENES Climate Analytics Service (ECAS) is a server-side processing service offered to ENES users with current cluster installations at CMCC and DKRZ. ECAS offers a virtual work environment based on Jupyter notebooks, allowing you to process and analyse data (including CMIP5 and CMIP6 data) using Python. Support for fast computations is provided via the Ophidia data analytics framework. The service is free for use by ENES users at the current installations at DKRZ and CMCC.

ECAS is integrated with other e-infrastructure services as part of the European Open Science Cloud in the frame of the EOSC-hub project.

CMCC service information

- The CMCC ECAS instance: https://ophidialab.cmcc.it
- Usage requires free registration. Please follow the instructions on the website.
- The CMCC instance offers access to multiple computing environments, including a "fat node" cluster.

DKRZ service information

- The DKRZ ECAS instance: https://ecaslab.dkrz.de
- Usage requires free registration. Please follow the instructions on the website.
- The DKRZ instance offers direct access to the CMIP5 and CMIP6 data pools. During the CMIP6 operational phase, the CMIP6 data pool is continually updated with newly available data.

Example workflows available

Example experiments and Jupyter notebooks are available at the DKRZ and CMCC instances and on GitHub. There is also training material available on GitHub.

Concerning use of the computing framework, the documentation on the Ophidia framework websites contains starter tutorials, descriptions of all processing operators and example workflows.

Some other portals

Portals with Data and Tools

- Climate Explorer: https://climexp.knmi.nl/start.cgi (new version under flag of WMO)
- ECA&D and ICA&D (European/international Climate Assessment Database): https://www.ecad.eu/, https://www.ecad.eu/icad.php (E-OBS through C3S CDS)
- Climate Data Guide: https://climatedataguide.ucar.edu/

Portal with Tools

- Climate Data Tool: https://iri.columbia.edu/our-expertise/climate/tools/cdt/

Impact Data

- Impact2C: https://www.atlas.impact2c.eu/en/

National Portals

- Example for France, the DRIAS portal: http://www.drias-climat.fr

Some thoughts

- A large variety of portals is available
  - Non-homogeneity in available datasets
  - Very different capabilities and interfaces
- May be difficult to know which one is suitable for specific needs and knowledge
- Guidance is usually not always sufficient
- Support may not be available
- Applicability and cautions about datasets and their possible use may not always be obvious
- On-demand remote calculations and subsetting is not always possible
- Downloading results in digital format is not always possible
- Also be attentive to FAIRness compliance, especially reproducibility and lineage

- For national, regional and local studies, always seek first for National portals, but be careful about uncertainties aspects
- The Climate Adapt portal is a good resource for guidance on how to perform an impact study properly

Some more thoughts ...

- There has never been so many accessible climate datasets!
- On-demand remote calculations is rapidly developing
- There is an increasing number of higher spatial and time resolution datasets
  - This is also true for gridded observation and reanalysis datasets
- Uncertainties can be better estimated thanks to the higher number of ensembles of simulations
- European-based portals integrating several infrastructures are being developed and launched
- Cloud-based solutions are emerging

CERFACS

CENTRE EUROPÉEN DE RECHERCHE ET DE FORMATION AVANCÉE EN CALCUL SCIENTIFIQUE

is-enes

INFRASTRUCTURE FOR THE EUROPEAN NETWORK FOR EARTH SYSTEM MODELLING

www.cerfacs.fr