

Building actionable climate products for end users using EGI-ACE resources

Christian Pagé
christian.page@cerfacs.fr
Abel Aoun

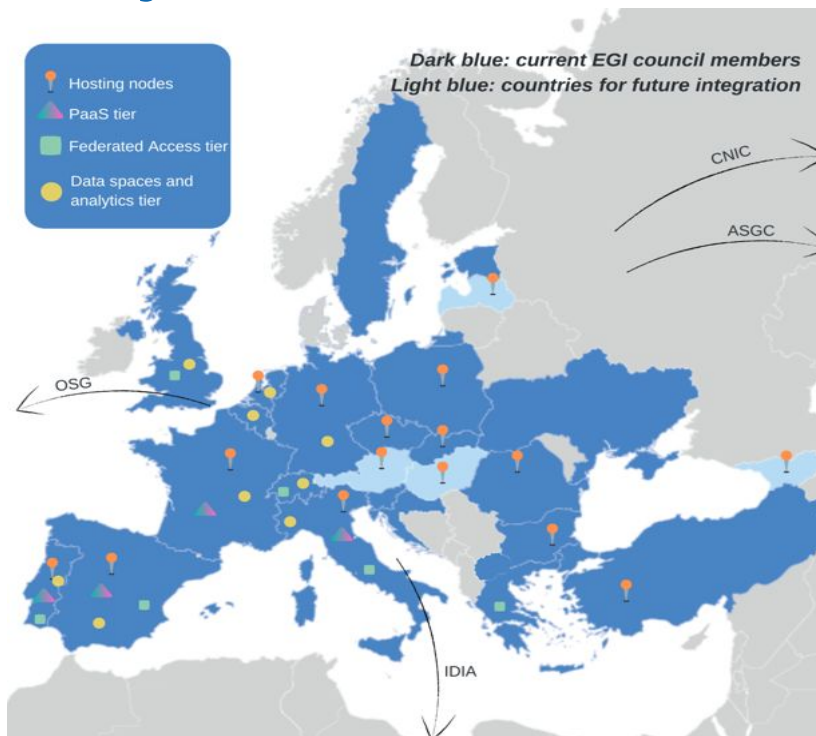
CERFACS, Toulouse, France

Pan-European digital assets
supporting research communities -
Benefits & opportunities
05-06 December 2022
Online

EGI-ACE Mission

Implement the **Compute Platform of the EOSC** and contribute to the **EOSC Data Commons** by delivering integrated computing, platforms, data spaces and tools as an integrated solution that is **aligned with** major European cloud federation projects and HPC initiatives.

Project Overview



EGI Advanced Computing for EOSC Grant agreement ID: 101017567

Budget

- Total budget: € 12,009,988
- EC budget: € 8,000,000

Consortium

- Coordinator - Stichting EGI
- 33 Partners, 23 third parties

Effort

- 1472 PMs, 48 FTEs
- **49% Virtual Access** (35 services, 38 providers)

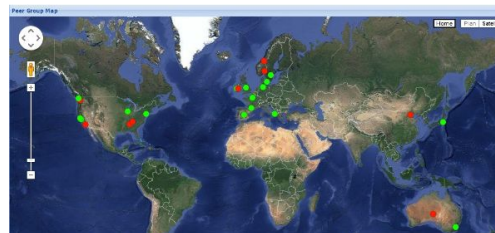
Duration

- Jan 2021 - June 2023 (30 months)

Climate data distribution

- Climate data is distributed using the Earth System Grid Federation (ESGF)
- Data Nodes interface is not straightforward to use for non-expert users
- Available variables are “raw” output from climate models: temperature, humidity, precipitation, ...
- Daily, monthly, ... frequencies

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



Hosted by:  Department of Energy
Lawrence Livermore National Laboratory

Powered by:  ESGF
Welcome, Guest. | Login | Create Account

WCRP CMIP6
World Climate Research Programme

You are at the ESGF@DOE/LNL node
Technical Support

Home Contact Us Data Nodes Status

MP Era +
Activity +
Model Cohort +
Product +
Source ID +
Institution ID +
Source Type +
Nominal Resolution +
Experiment ID +
Sub-Experiment +
Variant Label +
Grid Label +
Table ID +
Frequency +
Realm +
Variable +
CF Standard Name +
Data Node +

WARNING: Not all models include a variant "r11p111", and across models, identical values of variant_label do not imply identical variants! To learn which forcing datasets were used in each variant, please check modeling group publications and documentation provided through ES-DOC.

CMIP6 project data downloads 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.

Enter Text: [\[More Search Options \]](#)

☐ Show All Replicas ☐ Show All Versions ☐ Search Local Node Only (including All Replicas)

The search returned 0 results.

ESGF sponsors and partners
DOE Office of Science | IS-ENES | NASA | NOAA | NC | NSF

ESGF version v4.0.0
ESGF PDP version v4.0.4

Earth System Grid sponsors and partners
NOAA | NASA | NSF | DOE Office of Science | IS-ENES

<http://esgf-node.llnl.gov> [Privacy & Legal Notice](#)

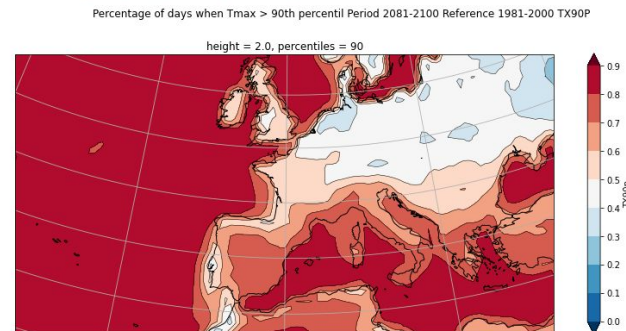
Gap between Users needs and available data

- Often significant gaps between distributed datasets and users' needs:
 - Assessing climate change anomalies
 - Evaluating climate extremes
 - Understanding climate change impacts
 - ...
- Users' Stories examples
 - Will there be more droughts in northeast Spain?
 - How likely landslides will occur in this mountainous valley?
 - Which region in my Europe will see the greatest change in heatwave intensity and occurrence?

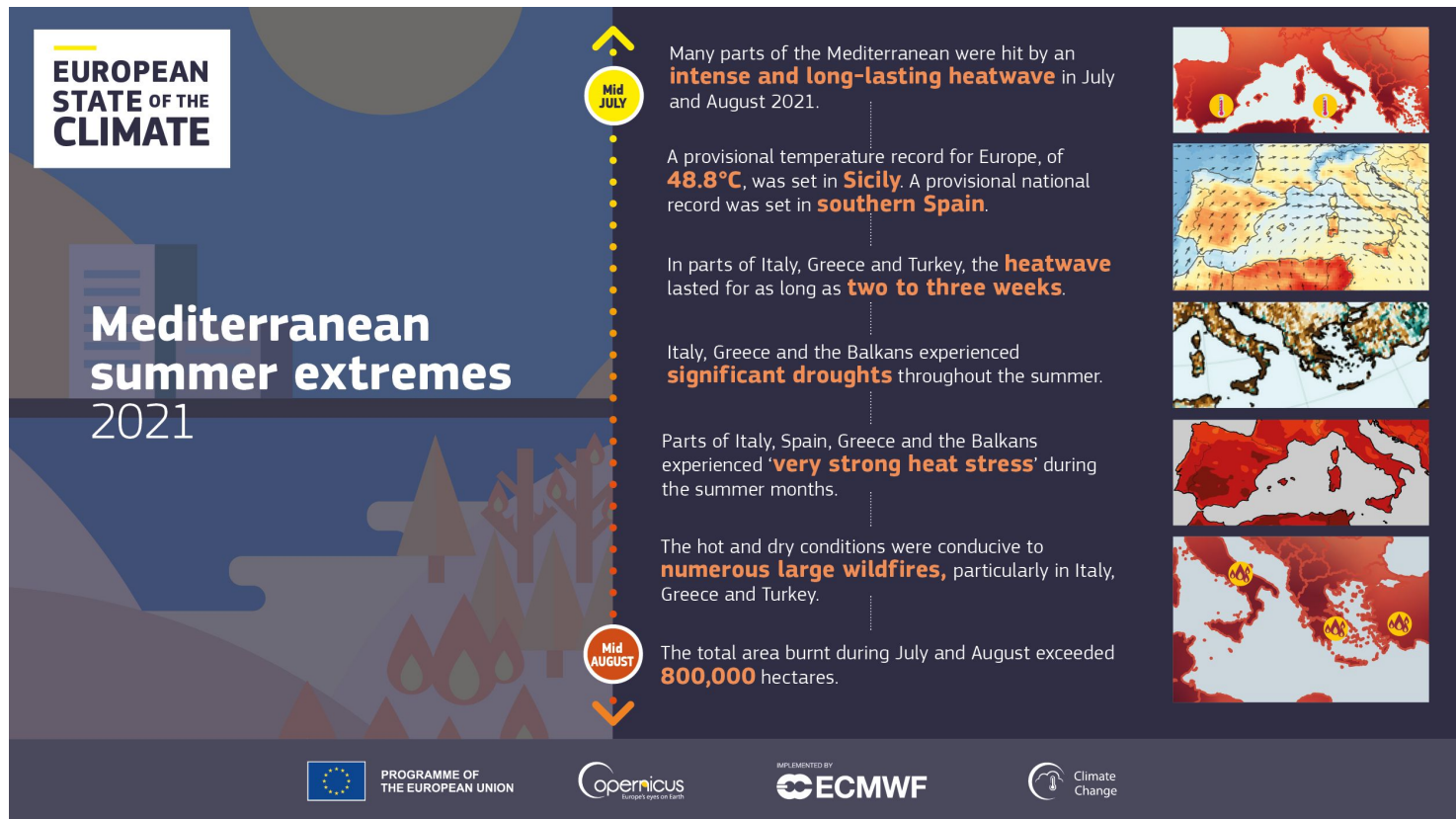
In the future climate compared to now

What is a climate index

- A Climate Index is derived from basic climate variables such as temperature, humidity, precipitation, wind, ...
 - Warm days (*days with mean temperature > 90th percentile of daily mean temperature*) – **TG90p**
 - Summer days (*days with max temperature ≥ 25 °C*) – **SU**
- Most of Climate Indices are standardized within the international community
 - ETCCDI, ECA&D, ET-SCI, ...



What is a climate index



icclim: a flexible tool, but still



- Tool: **icclim**, *an open source python software package to calculate climate indices*
- Simple and flexible API and interface, fast processing
- Difficult for users to process a sufficient numbers of climate projections to calculate those climate indices
 - Assess Uncertainties
 - Explore several Greenhouse Gas Emission Scenarios
 - Impossibility to download all required input data
 - Even with all data available, very time consuming and complex to calculate all what's needed

Project

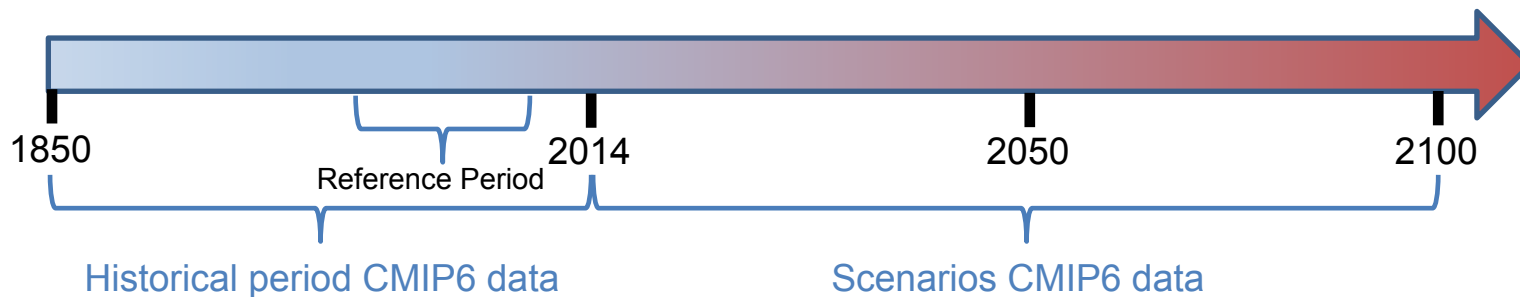
- Pre-generate 50 standard climate indices

CMIP6 (most common experiments used)

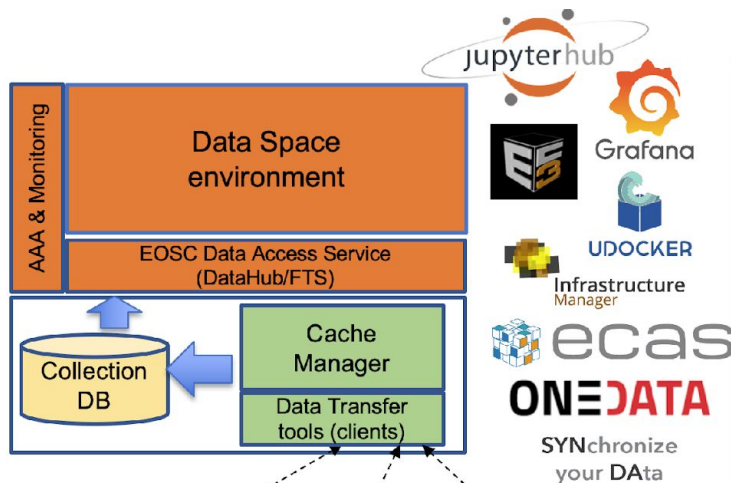
- Core set of simulations
 - **All**: climate models, greenhouse gas scenarios (aka SSPs...), ensemble members, versions
 - Daily time frequency

Choices to be made

- Reference period for percentiles
 - 1981–2010 (within historical period of climate simulations 1850–2014)



- Standard thresholds of standard indices
 - **Example:** Summer day is a day with maximum temperature $\geq 25^{\circ}\text{C}$



https://enesdataspace.vm.feccloud.eu

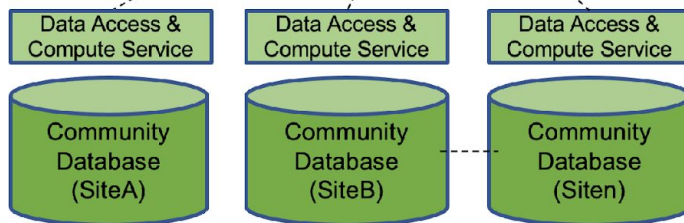
ENES Data Space

Home Notebooks User guide Login

The ENES Data Space delivers an open, scalable and cloud-enabled data science environment for climate data analysis on top of the EOSC Compute Platform. It provides both storage and computational capabilities.

It consists of a JupyterHub instance jointly with a large set of pre-installed Python libraries for running data manipulation, analysis, and visualization, and a data publication service enabling file browsing and data access for scientific datasets.

The ENES Data Space hosts (open) data from the ESGF federated data archive on compute cloud to support meteorological and industrial researchers in realistic climate model analysis experiments.



EUROPEAN OPEN SCIENCE CLOUD

Find resource... All resour... Q My EOSC Marketplace

Resources Processing & Analysis Data Analysis Image/Data Analysis ENES Data Space

cmcc
IPSL

ENES Data Space

Data science environment for climate data analysis on top of the EOSC Compute Platform

Organization: Euro-Mediterranean Center on Climate Change

★★★★★ (0.0/5) 0 reviews ☐ Add to comparison ☐ Add to favourites

Webpage Helodesk e-mail

Access the resource

OPEN ACCESS

Ask a question about this resource?

File Edit View Run Kernel Tabs Settings Help

Name	Last Modified
dask-worker-space	3 months ago
data	3 months ago
C4I_Summer_days_Calcul...	a month ago
C4I_Summer_days_Calcul...	a day ago
file_pr.txt	4 months ago
file_tas.txt	4 months ago
file_tasmax.txt	4 months ago
file_tasmin.txt	4 months ago
filelist.txt	4 months ago
nohup.out	2 days ago
pr.txt	4 months ago
su_icclim.py	4 days ago
tas.txt	4 months ago
tasmax.txt	4 months ago
tasmin.txt	4 months ago

jovyan@jupyter-cpage: ~/wX

jovyan@jupyter-cpage: ~/wX

C4I_Summer_days_Calcul...

Python 3 (ipykernel)

Code

```
x1, x2 = range1.start, range1.stop
y1, y2 = range2.start, range2.stop
return x1 <= y2 and y1 <= x2
```

```
[ ]: HOME = os.getenv('HOME')
```

```
historical = "CHIP"
```

```
ssp = "ScenarioMIP"
```

```
frequency = "day"
```

```
cmp6_dir = HOME + '/data/CHIP6'
```

```
dirs = {}
```

```
dirs["historical"] = cmp6_dir + '/' + historical
```

```
dirs["ssp"] = cmp6_dir + '/' + ssp
```

```
indices = {"tas": ["TG", "GD4", "HD17", "TG10p", "TG90p"],
```

```
"tasmin": ["TN", "TNx", "TNn", "TR", "CFD", "TN10p", "TN90p", "CSDI"],
```

```
"tasmax": ["TX", "TXx", "TXn", "SU", "CSU", "ID", "TX10p", "TX90p", "WSDI"],
```

```
"tasminmax": ["DTR", "ETR", "vDTR"],
```

```
"pr": ["PRCPTOT", "RR1", "SDII", "CWD", "CDD", "R10mm", "R20mm", "RX1day", "RX5day", "R75p",
```

```
"prsn": ["SD", "SD1", "SD5cm", "SD58cm"],
```

```
"taspr": ["CD", "CW", "WD", "WW"]
```

```
}
```

```
indices_percentiles = ["CD", "CW",
```

```
"R75p", "R75pTOT", "R95p", "R95pTOT", "R99p", "R99pTOT",
```

```
"TG10p", "TG90p", "TN10p", "TN90p", "TX10p", "TX90p",
```

```
"WD", "WW"]
```

```
indices_vars = {"tas": ["tas"],
```

```
"tasmin": ["tasmin"],
```

```
"tasmax": ["tasmax"],
```

```
"tasminmax": ["tasmin", "tasmax"],
```

```
"pr": ["pr"],
```

```
"prsn": ["prsn"],
```

```
"taspr": ["tas", "pr"]
```

```
}
```

```
reference_period = [1981, 2010]
```

```
# base period
```

```
base_dt1 = datetime.datetime(1981,1,1)
```

```
base_dt2 = datetime.datetime(2010,12,31)
```

```
institutes = {}
```

```
institutes["historical"] = os.listdir(dirs["historical"])
```

```
institutes["ssp"] = os.listdir(dirs["ssp"])
```

jovyan@jupyter-cpage: ~/wX

Process == false

Index: HD17

/home/jovyan/work/data/CHIP6/CHIP/BCC-ESM1/historical/r3ilp1f/HD17/gn/v20181220/HD17_day_BCC-ESM1_historical_r3ilp1f_gn_18500101-20141231.nc

Process == false

Index: TG10p

/home/jovyan/work/data/CHIP6/CHIP/BCC-ESM1/historical/r3ilp1f/TG10p/gn/v20181220/TG10p_day_BCC-ESM1_historical_r3ilp1f_gn_18500101-20141231.nc

Processing TG10p and creating /home/jovyan/work/data/CHIP6/CHIP/BCC-ESM1/historical/r3ilp1f/TG10p/gn/v20181220/TG10p_day_BCC-ESM1_historical_r3ilp1f_gn_18500101-20141231.nc

Key: tasmin

tasmin

First var test: /home/jovyan/data/CHIP6/CHIP/BCC-ESM1/historical/r3ilp1f/day/tasmin

Key: tasmax

tasmax

First var test: /home/jovyan/data/CHIP6/CHIP/BCC-ESM1/historical/r3ilp1f/day/tasmax

Key: tasminmax

tasminmax

First var test: /home/jovyan/data/CHIP6/CHIP/BCC-ESM1/historical/r3ilp1f/day/tasmin

Key: pr

pr

First var test: /home/jovyan/data/CHIP6/CHIP/BCC-ESM1/historical/r3ilp1f/day/pr

Key: prsn

prsn

First var test: /home/jovyan/data/CHIP6/CHIP/BCC-ESM1/historical/r3ilp1f/day/prsn

Key: taspr

taspr

First var test: /home/jovyan/data/CHIP6/CHIP/BCC-ESM1/historical/r3ilp1f/day/tas

Version: v20181220

Period: 18500101-20141231

Members: r3ilp1f

Key: tas

tas

First var test: /home/jovyan/data/CHIP6/CHIP/BCC-ESM1/historical/r3ilp1f/day/tas

Version: v20181220

Index: TG

/home/jovyan/work/data/CHIP6/CHIP/BCC-ESM1/historical/r3ilp1f/TG/gn/v20181220/TG_day_BCC-ESM1_historical_r3ilp1f_gn_18500101-20141231.nc

Process == false

Index: GD4

/home/jovyan/work/data/CHIP6/CHIP/BCC-ESM1/historical/r3ilp1f/GD4/gn/v20181220/GD4_day_BCC-ESM1_historical_r3ilp1f_gn_18500101-20141231.nc

Process == false

Index: HD17

/home/jovyan/work/data/CHIP6/CHIP/BCC-ESM1/historical/r3ilp1f/HD17/gn/v20181220/HD17_day_BCC-ESM1_historical_r3ilp1f_gn_18500101-20141231.nc

Processing TG10p and creating /home/jovyan/work/data/CHIP6/CHIP/BCC-ESM1/historical/r3ilp1f/TG10p/gn/v20181220/TG10p_day_BCC-ESM1_historical_r3ilp1f_gn_18500101-20141231.nc

Key: tasmin

tasmin

First var test: /home/jovyan/data/CHIP6/CHIP/BCC-ESM1/historical/r3ilp1f/day/tasmin

Version: v20181220

Period: 18500101-20141231

- Delays in initial planning
 - Delay in starting the action
 - Several Technical adjustments and Support actions in August (thanks CMCC!)
 - Complex processing script (parsing proper datafiles)
 - September extremely busy (project on hold)
- Current actions
 - Small adjustments to script
 - Not optimized: significant time to aggregate input files as xarray datasets and some pre-processing
 - Calculations in progress

Significant step toward
more actionable climate
data information



- Future actions
 - Validate calculations (end of 2022 – beginning of 2023)
 - Decide on where to store database permanently
 - NetCDF, zarr, Commercial and Public Clouds, ...
 - Make it accessible within the IS-ENES C4I platform
 - Use database to support Horizon Europe interTwin project
 - Disseminate information about this climate indices database
- Possible extensions
 - ERA5, and other re-analyses
 - CORDEX
 - CMIP5
 - CMIP7, Future CORDEX...

Thanks !



Christian Pagé christian.page@cerfacs.fr
Abel Aoun



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



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



Follow us on Twitter !
@ISENES_RI



Contact us at
is-enes@ipsl.fr



Follow our channel
IS-ENES3 H2020

