

# EOSC Compute Platform

## *Status and Way Forward*

**Smitesh Jain**, Innovation Management Specialist, EGI Foundation

**Giuseppe La Rocca**, Community Support Lead at the EGI Foundation

**Christian Pagé**, Research Engineer and Project Management, CERFACS

**József Kovács**, Senior Research Fellow, SZTAKI

**Hakan Bayındır**, Senior Researcher, TUBITAK ULAKBIM

EOSC Symposium - 14-17 November 2022, Prague



EGI-ACE receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no. 101017567.

# Outline of this session



- **Project overview**
  - Smitesh Jain, Innovation Management Specialist, EGI Foundation
- **User engagement and impact**
  - Giuseppe La Rocca, Community Support Lead, EGI Foundation
- **User experience stories:**
  - *Building a Climate indices dataset for climate change impacts assessment*
    - Chrisitan Pagé Research Engineer and Project Management, CERFACS
  - *Bring-your-own-resources: How the NEANIAS project became compatible with EGI computing services and introduced a new resource utilization approach*
    - Jozsef Kovacs, Senior Research Fellow, SZTAKI
- **HPC services in the EOSC Compute Platform**
  - Hakan Bayındır, Senior Researcher, TUBITAK ULAKBIM

# Building a Climate indices dataset for climate change impacts assessment

**Christian Pagé**, [christian.page@cerfacs.fr](mailto:christian.page@cerfacs.fr)

Research Engineer and Project Management, CERFACS



EOSC Symposium - 14-17 November 2022, Prague

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

EGI-ACE receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no. 101017567.



# Climate data distribution

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



- 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.

Powered 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/LLNL node [Technical Support](#)

Home | Contact Us | Data Nodes Status

WARNING: Not all models include a variant "1112111", 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 -r should never be combined.

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

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

The search returned 6 results.

ESGF sponsors and partners: DOE Office of Science | ES-ENES | NASA | NOAA | NCI | RDP

ESGF version v4.0.002

ESGF WCRP Version v4.0.0

Earth System CoG sponsors and partners: NOAA | NASA | NSF | DOE Office of Science | ES-ENES

Privacy & Legal Notice

# Gap between Users needs and available data

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



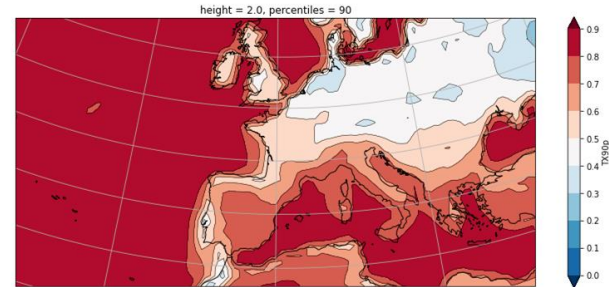
- Often significant gaps between distributed datasets and users' needs:
  - Assessing climate change anomalies
  - Evaluating climate extremes
  - Understanding climate change impacts
  - ...
- Users' Stories examples
  - Will there be more droughts in northeast Spain?
  - How likely landslides will 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  $\geq 25$  °C*) – **SU**
- Most of Climate Indices are standardized within the international community
  - ETCCDI, ECA&D, ET-SCI, ...

Percentage of days when Tmax > 90th percentile Period 2081-2100 Reference 1981-2000 TX90P





EUROPEAN  
STATE OF THE  
CLIMATE

## Mediterranean summer extremes 2021



Mid  
JULY

Many parts of the Mediterranean were hit by an **intense and long-lasting heatwave** in July and August 2021.

A provisional temperature record for Europe, of **48.8°C**, was set in **Sicily**. A provisional national record was set in **southern Spain**.

In parts of Italy, Greece and Turkey, the **heatwave** lasted for as long as **two to three weeks**.

Italy, Greece and the Balkans experienced **significant droughts** throughout the summer.

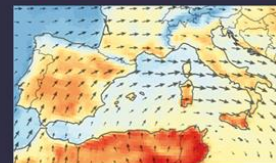
Parts of Italy, Spain, Greece and the Balkans experienced '**very strong heat stress**' during the summer months.

The hot and dry conditions were conducive to **numerous large wildfires**, particularly in Italy, Greece and Turkey.

Mid  
AUGUST



The total area burnt during July and August exceeded **800,000** hectares.



# icclim: a flexible tool, but still

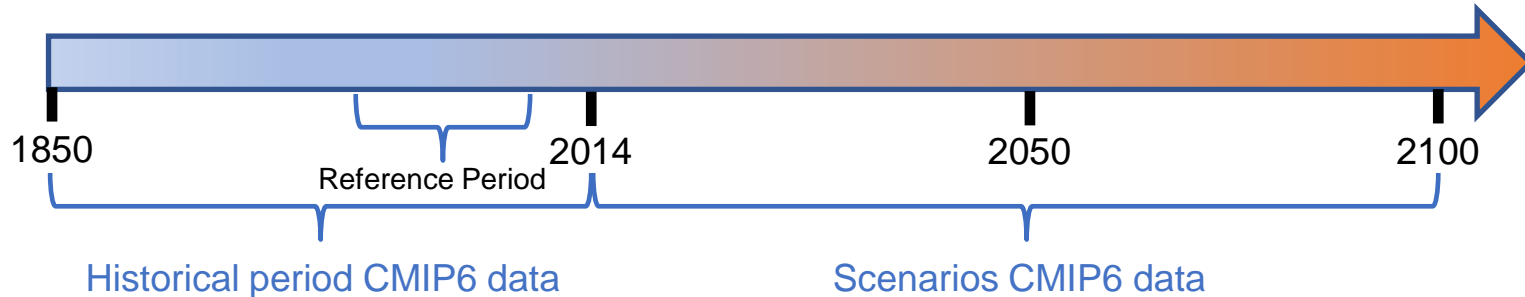


- Tool: **icclim**, an open source python software package to calculate climate indices
- Simple and flexible API and interface, fast processing
- Difficult for users to process a sufficient numbers of climate projections to calculate those climate indices
  - Assess Uncertainties
  - Explore several 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



- Pre-generate 50 standard climate indices
  - **CMIP6 (most common experiments used)**
  - +ERA5
  - +CORDEX
  - +CMIP5...
- Core set of simulations
  - **All:** climate models, greenhouse gas scenarios (aka SSPs...), ensemble members, versions
  - Daily time frequency

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



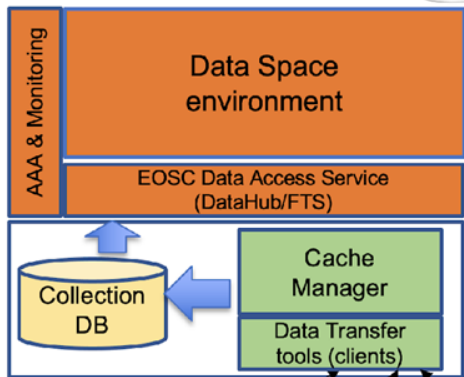
- Standard thresholds of standard indices
  - **Example:** Summer day is a day with maximum temperature  $\geq 25^{\circ}\text{C}$

# Computations



**cmcc**  
Centro Euro-Mediterraneo  
sui Cambiamenti Climatici

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




<http://is.enes.org/>




# Running on EGI-ACE resources



File Edit View Run Kernel Tabs Settings Help

Name	Last Modified
dask-worker-space	3 months ago
data	3 months ago
C4I_Summer_days_Calculat...	a month ago
C4I_Summer_days_Calculat...	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_licfcm.py	4 days ago
tas.txt	4 months ago
tasmax.txt	4 months ago
tasmin.txt	4 months ago

```
jovyan@jupyter-cpage: ~/w X | jovyan@jupyter-cpage: ~/w X | C4I_Summer_days_Calculat... X +
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"

cmip6_dir = HOME + '/data/CHIP6'

dirs = {}
dirs["historical"] = cmip6_dir + '/' + historical
dirs["ssp"] = cmip6_dir + '/' + ssp

indices = {"tas": ["TG", "GD4", "HD17", "TG10p", "TG90p"],
           "tasmin": ["TN", "TNx", "TNn", "TR", "FD", "CFD", "TN10p", "TN90p", "CSDI"],
           "tasmax": ["TX", "TXx", "TXn", "SU", "CSU", "ID", "TX10p", "TX90p", "WSDI"],
           "tasminmx": ["DTR", "ETR", "vDTR"],
           "pr": ["PRCPTOT", "RRI", "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"],
               "tasminmx": ["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: ~/w X +
Process == false
Index: HD17
/home/jovyan/work/data/CHIP6/CHIP/BCC/BCC-ESM1/historical/r3ilipfl/HD17/gn/v20181220/HD17_day_BCC-ESM1_historical_r3ilipfl_gn_18500101-20141231.nc
Process == false
Index: TG10p
/home/jovyan/work/data/CHIP6/CHIP/BCC/BCC-ESM1/historical/r3ilipfl/TG10p/gn/v20181220/TG10p_day_BCC-ESM1_historical_r3ilipfl_gn_18500101-20141231.nc
Processing TG10p and creating /home/jovyan/work/data/CHIP6/CHIP/BCC/BCC-ESM1/historical/r3ilipfl/TG10p/gn/v20181220/TG10p_day_BCC-ESM1_historical_r3ilipfl_gn_18500101-20141231
.nc
Key: tasmin
tasmin
First var test: /home/jovyan/data/CHIP6/CHIP/BCC/BCC-ESM1/historical/r3ilipfl/day/tasmin
Key: tasmax
tasmax
First var test: /home/jovyan/data/CHIP6/CHIP/BCC/BCC-ESM1/historical/r3ilipfl/day/tasmax
Key: tasminmx
tasmin mx
First var test: /home/jovyan/data/CHIP6/CHIP/BCC/BCC-ESM1/historical/r3ilipfl/day/tasmin
Key: pr
pr
First var test: /home/jovyan/data/CHIP6/CHIP/BCC/BCC-ESM1/historical/r3ilipfl/day/pr
Key: prsn
prsn
First var test: /home/jovyan/data/CHIP6/CHIP/BCC/BCC-ESM1/historical/r3ilipfl/day/prsn
Key: taspr
tas pr
First var test: /home/jovyan/data/CHIP6/CHIP/BCC/BCC-ESM1/historical/r3ilipfl/day/tas
Version: v20181220
Period: 18500101-20141231
Members: r3ilipfl
Key: tas
tas
First var test: /home/jovyan/data/CHIP6/CHIP/BCC/BCC-ESM1/historical/r3ilipfl/day/tas
Version: v20181220
Period: 18500101-20141231
Index: TG
/home/jovyan/work/data/CHIP6/CHIP/BCC/BCC-ESM1/historical/r3ilipfl/TG/gn/v20181220/TG_day_BCC-ESM1_historical_r3ilipfl_gn_18500101-20141231.nc
Process == false
Index: GD4
/home/jovyan/work/data/CHIP6/CHIP/BCC/BCC-ESM1/historical/r3ilipfl/GD4/gn/v20181220/GD4_day_BCC-ESM1_historical_r3ilipfl_gn_18500101-20141231.nc
Process == false
Index: HD17
/home/jovyan/work/data/CHIP6/CHIP/BCC/BCC-ESM1/historical/r3ilipfl/HD17/gn/v20181220/HD17_day_BCC-ESM1_historical_r3ilipfl_gn_18500101-20141231.nc
Processing TG10p and creating /home/jovyan/work/data/CHIP6/CHIP/BCC/BCC-ESM1/historical/r3ilipfl/TG10p/gn/v20181220/TG10p_day_BCC-ESM1_historical_r3ilipfl_gn_18500101-20141231
.nc
Key: tasmin
tasmin
First var test: /home/jovyan/data/CHIP6/CHIP/BCC/BCC-ESM1/historical/r3ilipfl/day/tasmin
Version: v20181220
Period: 18500101-20141231
```

Simple 9 1 Python 3 (ipykernel) | Idle Mem: 863.93 / 15360.00 MB

Mode: Command Ln 1, C

- 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, New CORDEX...



# Experiences and the Future of the EOSC Compute Platform

**Smitesh Jain**, Innovation Management Specialist, EGI Foundation

**Tiziana Ferrari**, Director, EGI Foundation

**Christian Pagé**, Research Engineer and Project Management, CERFACS

**József Kovács**, Senior Research Fellow, SZTAKI

**Hakan Bayındır**, Senior Researcher, TUBITAK ULAKBIM

EOSC Symposium - 14-17 November 2022, Prague



EGI-ACE receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no. 101017567.



# Thank you!

Contact: [egi-ace-po@mailman.egi.eu](mailto:egi-ace-po@mailman.egi.eu)  
Website: [www.egi.eu/projects/egi-ace](http://www.egi.eu/projects/egi-ace)



[EGI Foundation](#)



[@EGI\\_einfra](#)



EGI-ACE receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no. 101017567.