Climate4Impact (C4I) and icclim
Tools to ease access to climate data

Christian Pagé
CERFACS, Toulouse, France

Abel Aoun
CERFACS Toulouse, France

Alessandro Spinuso
KNMI, De Bilt, Netherlands

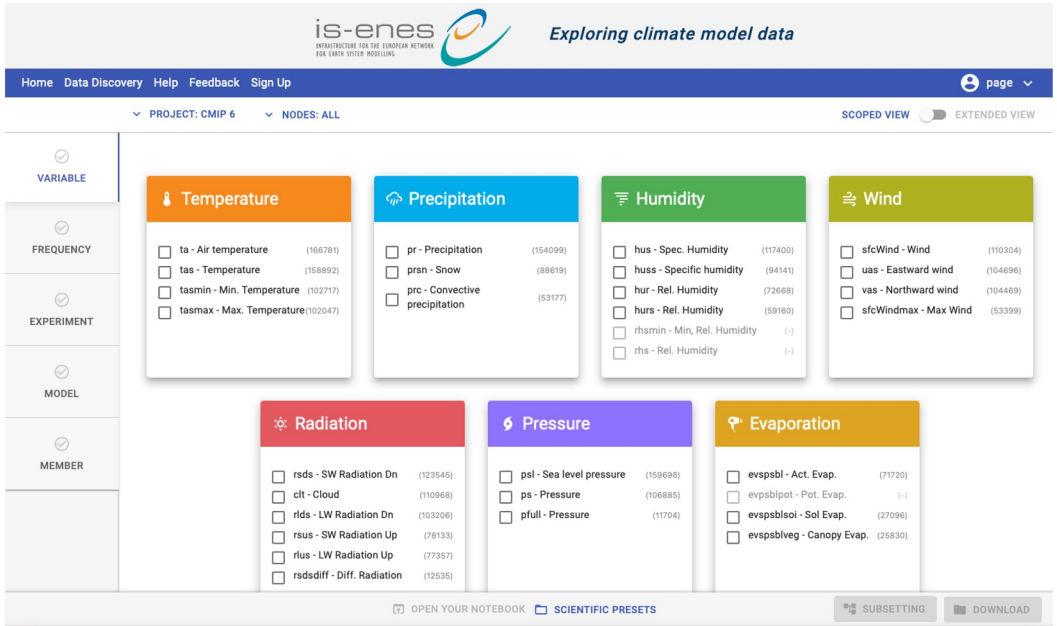
IS-ENES3 Central&Eastern Europe Autumn School
From Climate Projections to Climate impacts via Regional Downscaling
Prague, 28 November-2 December 2022

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

What is the climate4impact (C4I) portal? (1/2)

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

Findable Accessible Interoperable Reusable

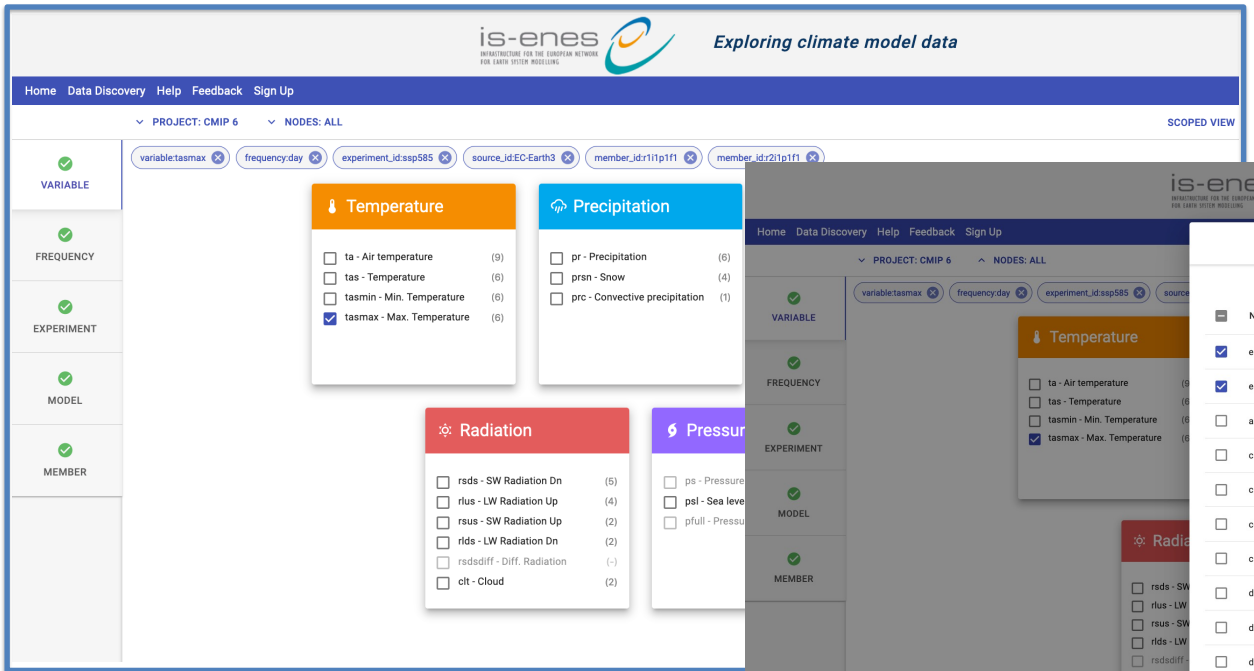
IS-ENES3 has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824084

Climate4Impact (v2) Workflows & Workspaces

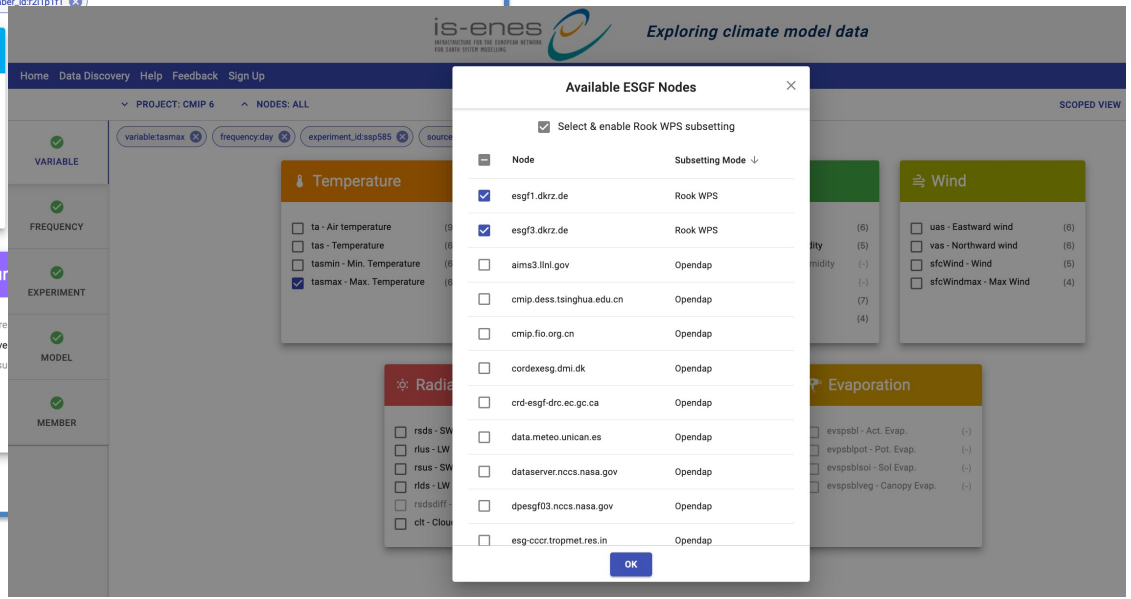
IS-ENES3 has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824084

Climate4Impact (v2) Search & Nodes

Search Parametrisation made easier



Nodes Selection by Service

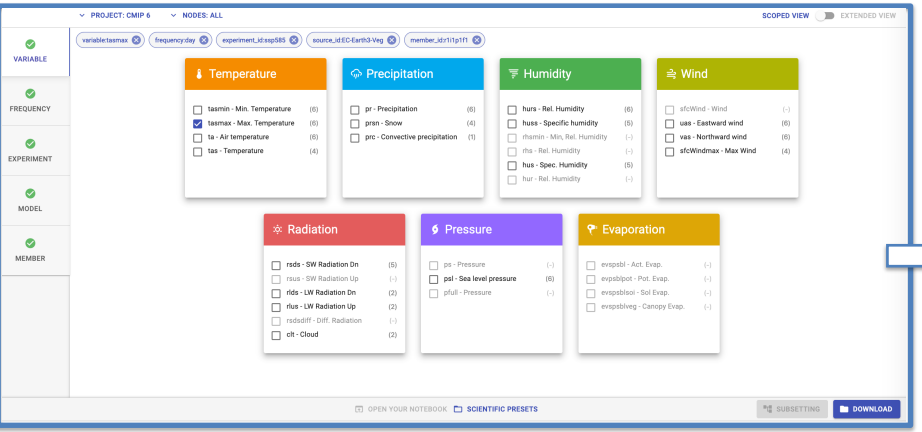


IS-ENES3 has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824084

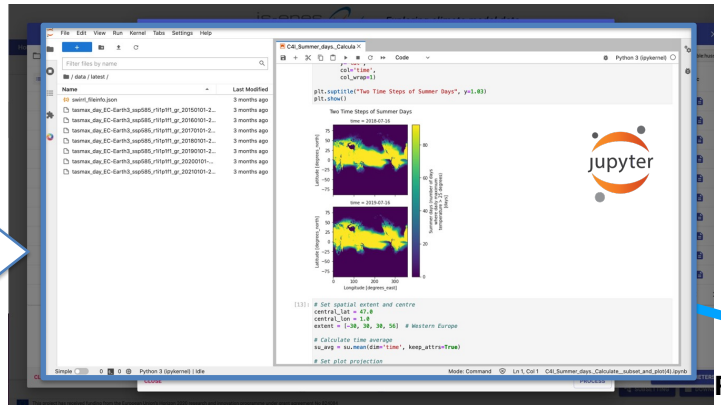
Climate4Impact (v2) Workflows & Workspaces

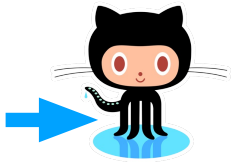
Climate4Impact Search for CMIP5/6 CORDEX Data

https://dev.climate4impact.eu



Workflows for data staging & reduction onto icclim powered Notebooks





Save/Share Progress to Git

Reduced Data MyBinder Reproduce

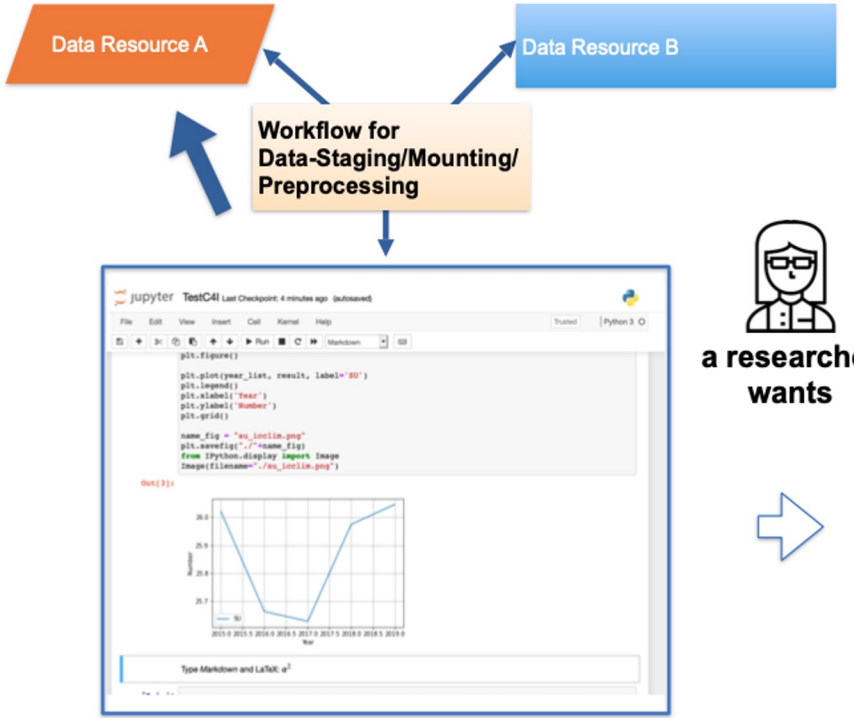
- **Trace Changes to Restore, Recover Software and/or Data**

W3C PROV

CERFACS

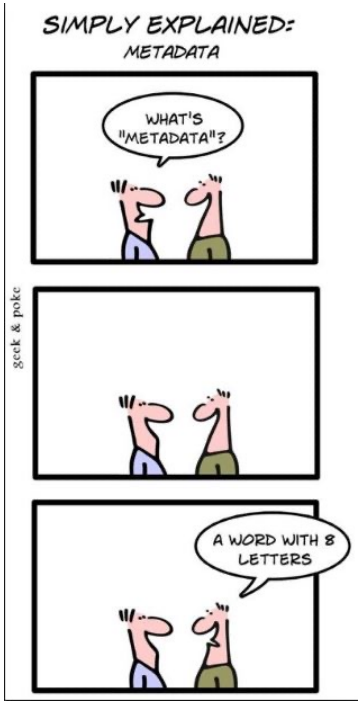
IS-ENES3 has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824084

C4I Workspace Use Case



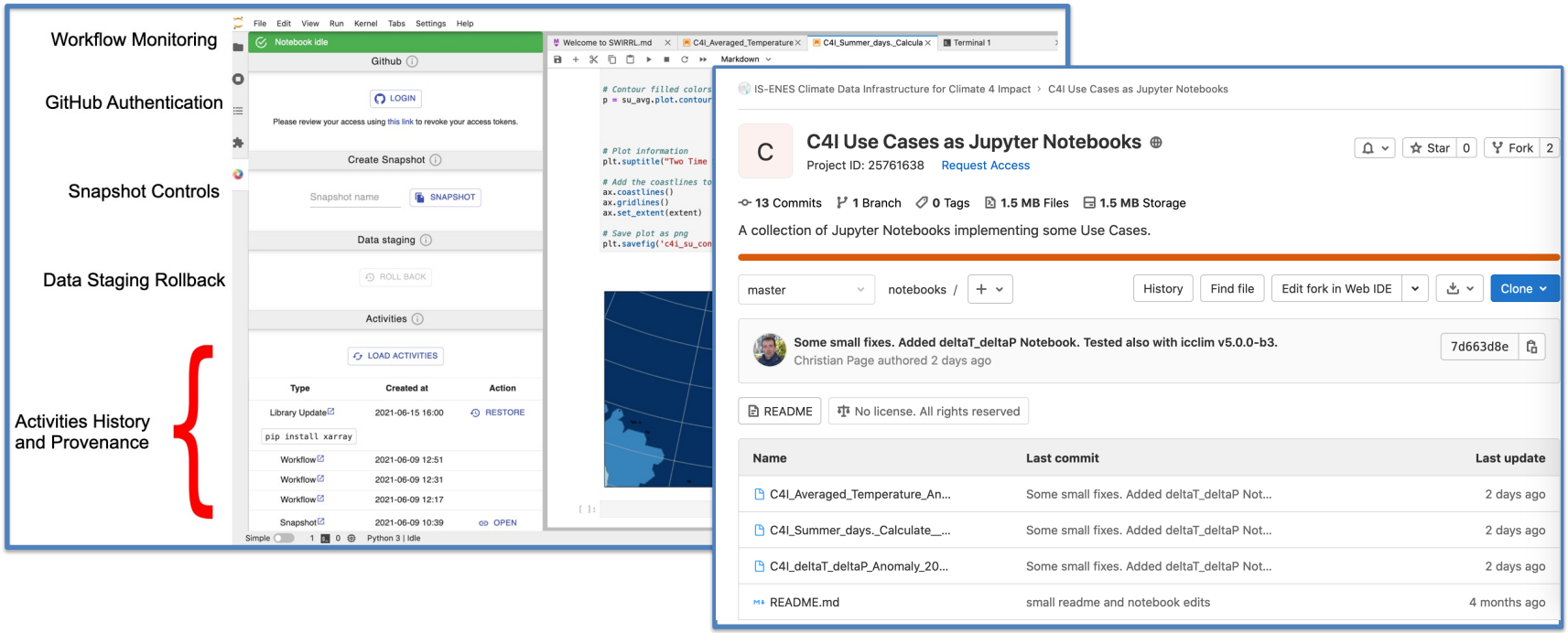
a researcher wants

- **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)



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

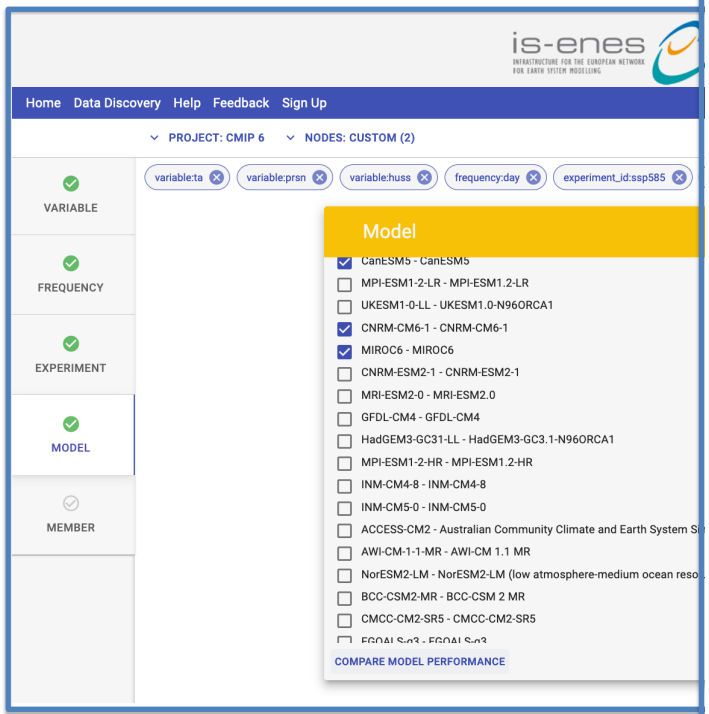
SWIRRL JupyterLab Extension and Sample Notebooks based on icclim (climate indices)



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

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

ESMValTool in C4I
Models' Performance Comparison

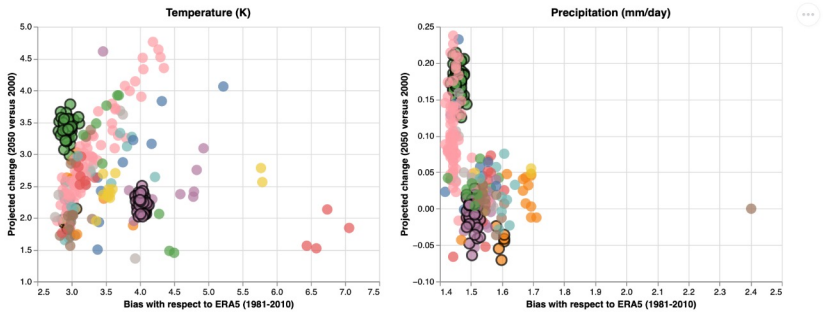


Climate impact result viewer

This applications shows a 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 calcated 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 CMIP6

Selected datasets:

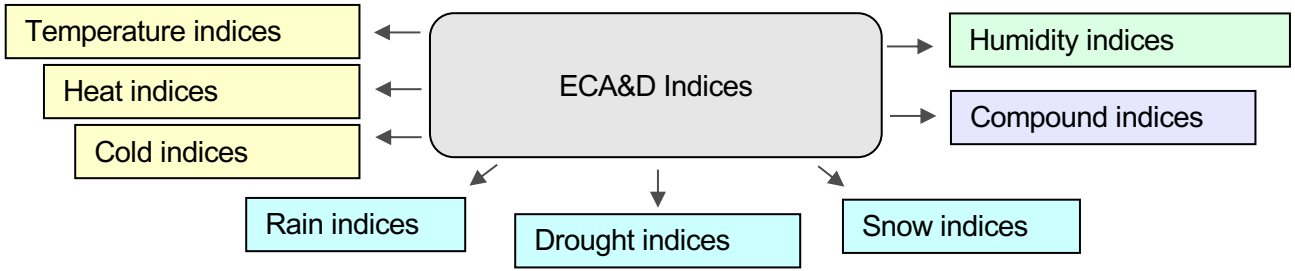
- CNRM-CM6-1
- MIROC6
- CanESM5

View static recipe output

Netherlands eScience Center

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

icclim: Index Calculation for Climate



- 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 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)
  - Available indices: https://icclim.readthedocs.io/en/latest/explanation/climate_indices.html#icclim-capabilities

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

icclim: climate indices

Documentation: https://icclim.readthedocs.io/en/latest/python_api.html
Source code: https://github.com/cerfacs-globc/icclim
Current Version 6.1.2: https://github.com/cerfacs-globc/icclim/releases/tag/v6.1.2

`icclim.index(**kwargs)`

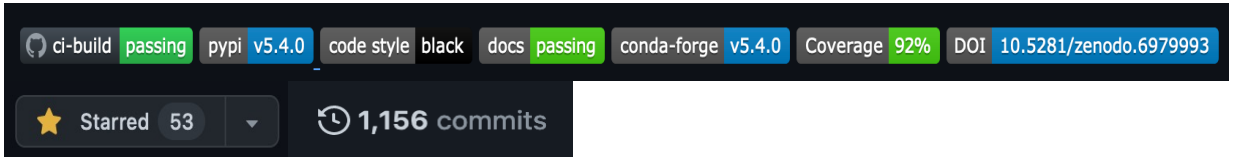
Parameters:

- **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.

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

icclim v6.1.2

https://github.com/cerfacs-globc/icclim




```
import icclim
summer_days = icclim.su("netcdf_files/tasmax_1990-2100.nc", out_file="summer_days.nc")
```

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

A **F.A.I.R** icclim

- Findable:
  - Open source code, licensed under the permissive Apache 2 license.
  - Try to follow OpenSSF principles
  - Documentation freely available on readthedocs
  - Dissemination by Christian in conferences (AGU, AMS, EGU...).
  - Integrated in C4I and C3S.
- Accessible
  - Documentation on readthedocs.
  - Github issues and pull requests.
  - CERFACS support the development.
  - Can be installed with **pip** or **conda-forge**.
- Interoperable
  - Based on well knowns libraries: xarray, xclim, numpy.
  - Support NetCDF and xarray.Dataset input formats.
  - Indices are deriviated from ECA&D and CF conventions standards.
  - icclim custom parsing for non CF complying inputs.
- Reusable
  - NetCDF output, with enriched metadata.
  - Easy to integrate within existing infrastructures.
  - Easy to write small scripts with it.

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

icclim: examples

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

```
## KNMI TX
tx_files = glob.glob(f"netcdf_files/knmi/clean/*tx*.nc")
bp = [datetime.datetime(1901, 1, 1), datetime.datetime(1921, 12, 31)]
icclim.index(index_name='tx90p',
             in_files=tx_files,
             base_period_time_range=bp,
             slice_mode='YS',
             out_file="netcdf_files/output/out.nc")
```

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

[new in 5.1.0] Rechunk how data is stored

Documentation: https://icclim.readthedocs.io/en/latest/how_to/dask.html#create-an-optimized-chunking-on-disk

```python
import icclim

ref_period = [datetime.datetime(1980, 1, 1), datetime.datetime(2009, 12, 31)]
with icclim.create_optimized_zarr_store(
        in_files="netcdf_files/tas.nc",
        var_names="tas",
        target_zarr_store_name="opti.zarr",
        keep_target_store=False,
        chunking={"time": -1, "lat": "auto", "lon": "auto"},
) as opti_tas:
    icclim.index(
        index_name="TG90p",
        in_files=opti_tas,
        slice_mode="YS",
        base_period_time_range=ref_period,
        out_file="netcdf_files/output/tg90p.nc",
    )
```

ement No 824084

[new in 5.1.0] Compute multiple indices at once

Documentation https://icclim.readthedocs.io/en/latest/how_to/recipes_ecad.html#multi-index-computation

```python
bp = [datetime.datetime(1991, 1, 1), datetime.datetime(1999, 12, 31)]
tr = [datetime.datetime(1991, 1, 1), datetime.datetime(2010, 12, 31)]
res = icclim.indices(
    index_group=IndexGroup.HEAT,
    in_files="./netcdf_files/climpact.sampledata.gridded.1991-2010.nc",
    base_period_time_range=bp,
    time_range=tr,
    out_file="pouetpouet.nc",
)
```

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

[new in 5.2.0] one function per index

Documentation: https://icclim.readthedocs.io/en/latest/references/ecad_functions_api.html

```python
import glob
import icclim
summer_days = icclim.su(
    in_files=glob.glob("netcdf_files/tasmax*.nc"),
    out_file="summer_days.nc",
)
```

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

[new in 5.3.0] custom season between exact dates

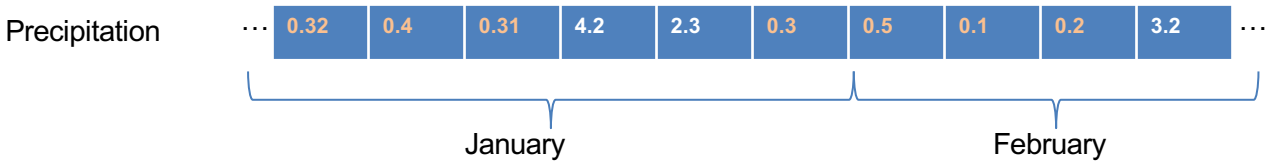
```
result = icclim.su(
    in_files=xr.open_dataset(climp_file).tmax,
    slice_mode=["season", ["19-07", "8 Aout"]],
).compute()
```

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

icclim v6
Generic climate indices made easy

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

[new in 6.0] Spells starting before the season bounds are properly counted



Say we want to compute cdd on a monthly basis…

CDD: number of consecutive where pr < 1 mm/day

```
cdd = icclim.cdd(in_files=precipitation, slice_mode="month").CDD.compute()
```

Should February cdd be 3 days or 4 ?

With icclim 6 we count spells length before resampling into monthly values,
So CDD in February would be 4 here

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

[new in 6.0] Generic indices – What ?

What's the difference between SU, FD and Tx90p ?

Number of days where tx > 25 degC

Number of days where tn < 0 degC

Number of days where tx > 90th (daily) percentile

An operator

A studied variable

A threshold

For our output, we want the right computation to executed and to have specific metadata to clearly identify :

- What was the computation configuration
- What was the initial data used

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

[new in 6.0] Generic indices – How ?

```
icclim.index(in_files=climp_file,
    Operator    index_name="count_occurrences",
Studied variable var_name="tmax",
   Threshold    threshold="≥ 20 deg_C",
                ).compute()
```

Not CF proof!

standard_name: number_of_days_with_maximum_air_temperature_above_threshold

long_name: Number of days when maximum air temperature is greater or equal to 293.15 K for each year.

units: d

cell_methods: time: sum over days

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

R75p (ECAD)

```
thresh = Threshold("≥ 75 period_per",
                   threshold_min_value="1 mm/day",
                   reference_period=['1991-01-01', "31 décembre 2000" ])
r75p = icclim.index(climp_file,
```

CW (ECAD)

```
  pr_thresh = Threshold("≥ 75 doy_per", threshold_min_value="1 mm/day")
  temp_thresh = Threshold("≥ 75 doy_per")
  c_w = icclim.index(climp_file,
```

Cold, wet and windy (Bretagne?)

```
    pr_thresh = Threshold("≥ 75 doy_per", threshold_min_value="1 mm/day")
    temp_thresh = Threshold("≥ 75 doy_per")
    wind_thresh = Threshold("≥ 95 doy_per")
    c_w_w = icclim.index(ds,
                         index_name="count_occurrences",
                         var_name=["precip", "tmax", "sfcWind"],
                         threshold=[pr_thresh, temp_thresh, wind_thresh])
```


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

[new in 6.0] Generic indices – Operators

count_occurrences	SU, TR, TG90p, CD, R95p SD50cm …
max_consecutive_occurrence	CSU, CFD, CDD, CWD
sum_of_spell_lengths	WSDI, CSDI
excess	GD4
deficit	HD17
fraction_of_total	R75pTOT, R95pTOT, r99pTOT
maximum	RX1day, TXx, TNx, custom__maximum
minimum	TXn, TNn, user_index - minimum
average	TG, TN, TX, SDII, SD, custom__mean
sum	PRCPTOT, custom__sum
standard_deviation	**new**
max_of_rolling_sum	RX5day, user_indexrolling_sum
min_of_rolling_sum	custom__rolling_sum
min_of_rolling_average	custom__rolling_mean
max_of_rolling_average	custom__rolling_mean
mean_of_difference	DTR
difference_of_extremes	ETR
mean_of_absolute_one_time_step_difference	vDTR
difference_of_means	custom_anomaly

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

[new in 6.0] Generic indices – Thresholds

Scalar	Threshold(query=">=", value=25, unit="degC")
Day of year percentiles	Threshold(query=">=", value=99, unit="doy_per")
Period percentiles	Threshold(query=">=", value=99, unit="period_per")
Bounded period percentiles.	Threshold(">= 75 period_per”. , threshold_min_value="1 mm/day” , reference_period=['1991-01-01’ , "31 décembre 2000"])
Per-grid cell threshold	Threshold(query=">=", value=“data.nc”, threshold_var_name="tmin", unit="K")
Sequence of scalars	Threshold(query=">=", value=[25, 29], unit="degC")

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

Thanks !

On behalf of the climate4impact and icclim teams
https://dev.climate4impact.eu
https://icclim.readthedocs.io/en/stable/

Christian Pagé christian.page@cerfacs.fr
Alessandro Spinuso alessandro.spinuso@knmi.nl
Abel Aoun 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

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

Follow us on Twitter !
@ISENES_RI

Contact us at
is-enes@ipsl.fr

YouTube
Follow our channel
IS-ENES3 H2020

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

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