# Interactive and Flexible Environment for on-demand Climate Data Analysis

Christian Pagé, Abel Aoun
CECI, Université de Toulouse, CNRS, CERFACS
Toulouse, France
christian.page@cerfacs.fr

Alessandro Spinuso, Mats Veldhuizen, Ian van der Neut
R&D Observations and Data Technology
KNMI
De Bilt, Netherlands

Klaus Zimmermann, Lars Bärring
SMHI Rossby Centre
Norrköping, Sweden

Phillip Kershaw
NCEO / Centre for Environmental Data Analysis,
STFC, Rutherford Appleton Laboratory
Didcot, United Kingdom

## Abstract

Many end users of climate change information often need to rely on external data analysis environments, having proper tools to produce meaningful information for their specific applications. Data volumes as well as the number of datasets are increasing very rapidly. The ability to select, process and download all needed data is getting complex, technical, and very time-consuming, especially when datasets are distributed among several data centres and into a large quantity of files. This is the case for the climate data distribution infrastructure ESGF (Earth System Grid Federation).

This means that science gateways are very important for end users to ease the access to climate data analysis. There are several platforms available to hide this complexity from users and provide a seamless access to climate analysis on large pre-generated archives of datasets' copies. We can cite the Copernicus Data Store[1], Pangeo[2] and ESGF Amazon Web Services[3], as they provide a (limited) subset of the most used climate datasets. Another important platform is developed within the H2020 IS-ENES3 project: climate4impact 2.0[4], providing access to all ESGF datasets. It is using an enhanced Jupyter-Lab environment orchestrated via the SWIRRL[5] API [1] along with a collection of Jupyter notebooks[6] as a useful set of examples on how to use and process data. Similar solution to EODAG Jupyter extension can be envisaged once the backend of the ESGF will be migrated. The selected datasets, using an intuitive and simplified interface (that can also be set in "Expert Mode"), are asynchronously staged from the ESGF, via SWIRRL workflows, onto the user space, obtaining custom and limited sized data-spaces. Users can choose to reduce the size of the transferred data by configuring a subsetting workflow. This invokes the remote WPSs available at data nodes of the ESGF, which will slice the data before sending. The portal also provides interactive pages for the evaluation of climate models, using ESMValTool [7][2], to guide users on selecting appropriate climate datasets.

The notebooks that can be executed in climate4impact, are developed using a very convenient software library to calculate climate indices and indicators called icclim[8] [3]. This library is a flexible python software package to calculate climate indices and indicators. This tool adheres as much as possible to metadata conventions such as the Climate & Forecasting Conventions (CF-1.x) as well as the clix-meta[9] work that is being done in IS-ENES3[10]. It provides the implementation of most of the international standard climate indices. It has been validated against Climpact[11] as well.

This project (IS-ENES3) has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N°824084.

***Keywords— climate; analysis; compute; toolbox; notebook; climate indices***

## References

[1] A. Spinuso, M. Veldhuizen, D. Bailo, V. Vinciarelli, T. Langeland; SWIRRL. Managing Provenance-aware and Reproducible Workspaces. *Data Intelligence* 2022; doi: https://doi.org/10.1162/dint_a_00129

[2] Righi, M., Andela, B., Eyring, V., Lauer, A., Predoi, V., Schlund, M., ... & Zimmermann, K. (2020). Earth System Model Evaluation Tool (ESMValTool) v2. 0–technical overview. *Geoscientific Model Development*, *13*(3), 1179-1199

[3] Abel Aoun, Christian, Natalia, Krahshtrea, marev711, Lars Bärring, bekozi, Petros Gasteratos, & bascrezee. (2022). cerfacs-globc/icclim: v5.1.0 (v5.1.0). Zenodo. https://doi.org/10.5281/zenodo.6380604

[1] https://cds.climate.copernicus.eu/cdsapp#!/toolbox
[2] https://pangeo.io
[3] https://registry.opendata.aws/cmip6/
[4] https://dev.climate4impact.eu
[5] https://gitlab.com/KNMI-OSS/swirrl
[6] https://gitlab.com/is-enes-cdi-c4i/notebooks
[7] https://www.esmvaltool.org/
[8] v5.1.0 https://github.com/cerfacs-globc/icclim
[9] https://github.com/clix-meta
[10] https://is.enes.org/
[11] https://github.com/ARCCSS-extremes/climpact