CMORization of CMIP6 data: workflow from the data request to the final ESGF-publishable data

Arthur Ramos (*arthur.amaral@bsc.es*)[1], Pierre-Antoine Bretonnière (*pierre-antoine.bretonniere@bsc.es*)[1]

[1]Earth Sciences Department, Barcelona Supercomputing Center (BSC)

Introduction

The process of implementing a Model Intercomparison Project (**MIP**) goes through several steps to assure the confiability of the data being produced by a research institute. In order to facilitate this process, many tools are being developed by the **EC-Earth** community that allow an easier approach to pre and post-process the model's inputs/outputs. Some of these tools, and how they fit in the workflow, are shown here.

1 - MIPs with participation of BSC

DECK	Diagnostic, Evaluation and Characterization of Klima
DCPP	Decadal Climate Prediction Project
VolMIP	Volcanic Forcings Model Intercomparison Project
HiResMIP	High Resolution Model Intercomparison Project
ScenarioMIP	Scenario Model Intercomparison Project

2 - Data Request (*drq*) (https://pypi.org/project/dreqPy/)

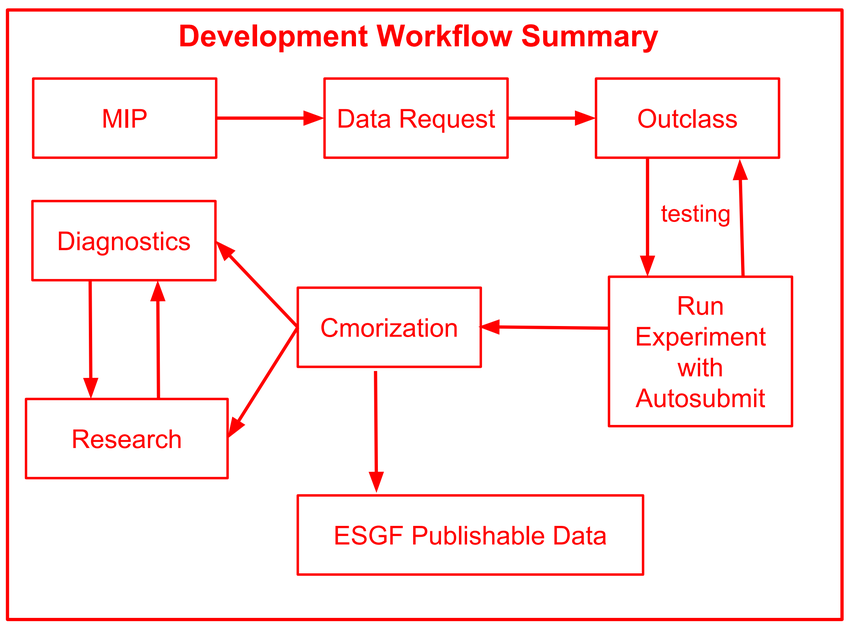
By using a combination of the packages **drq** and **ece2cmor3**, one can run:

```
./generate-ec-earth-namelists.sh DCPP DCPP 1 1
```

It generates the xml files (for the ocean, ice and biogeochemical) to be used by the ocean/ice model NEMO, and the ppt files to be used by the atmospheric/land model IFS. Together with the input files, a series of excel tables are also generated, in which they carry the right names and metadata to be used later for the CMORization. In this case, the reference MIP is the DCPP MIP, Tier 1, Priority 1.

3 - Outclass Development

As part of the workflow, the **Outclass** contains all the files and information regarding that specific MIP requested in the Data Request. In this step we can add variables, vertical levels, increase/decrease the frequency of the output, all depending on the needs of each group or project. Each MIP has its own **Outclass** files.



4 - Autosubmit (https://pypi.org/project/autosubmit/)

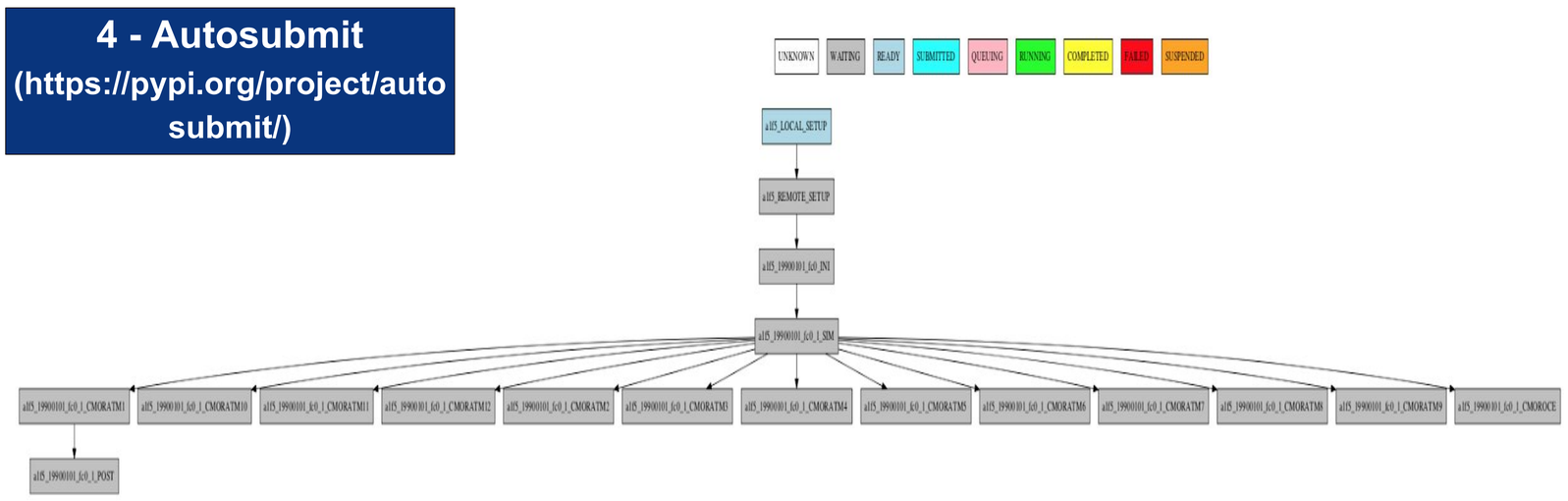
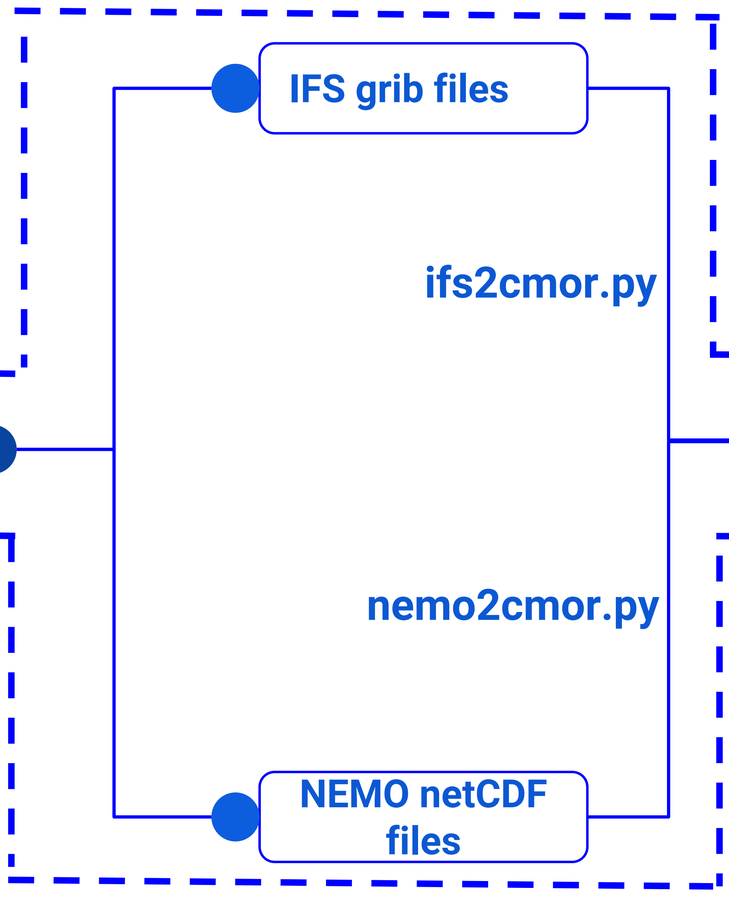


Fig. 1: Autosubmit workflow.

A python-based tool to create, manage and monitor experiments by using Computing Clusters, HPC's and Supercomputers remotely via ssh. The experiment manager is able to interact with git and svn repositories and controls the entire process, from setting up local and remote machines, going through the simulations and finally the storage.

5 - Climate Model Output Rewriter (CMOR) (github.com/PCMDI/cmor) and ece2cmor3 (github.com/EC-Earth/ece2cmor3)

CMOR library is used to standardize the model output with the DRS required by the project conventions (CMIP6, CORDEX, SPECS, etc.).



Reads the excel table generated by the **drq** to fill the metadata required by each project with the right names and frequencies.

ece2cmor3 reads and converts the output from different models, including tm5, LPJ-GUESS, NEMO and IFS, into netCDF files. Can be used online or offline.

Can be processed by different post-processing or diagnostic tools, such as Barakuda, ECE3-POSTPROC, Earth diagnostics/ESMValTool (https://earth.bsc.es/gitlab/es/earthdiagnostics).

Get the poster here

6 - Diagnostics

Example of diagnostics using a customized version of Barakuda that allows CMORized data and will soon be implemented in the Autosubmit workflow. .

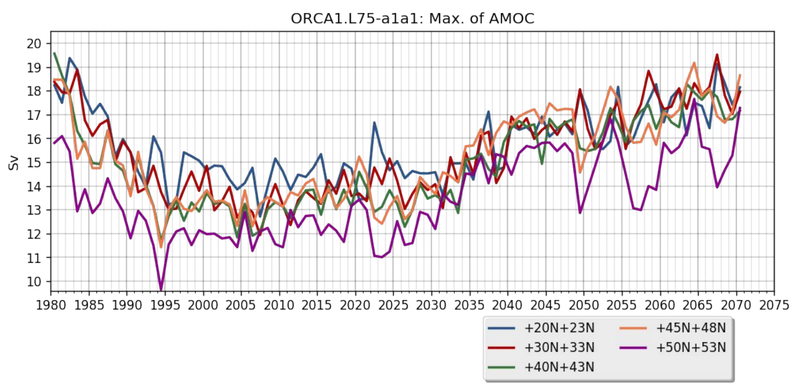


Fig. 2: Maximum values of AMOC between 1980 and 2070 using the Spinup Outclass. .

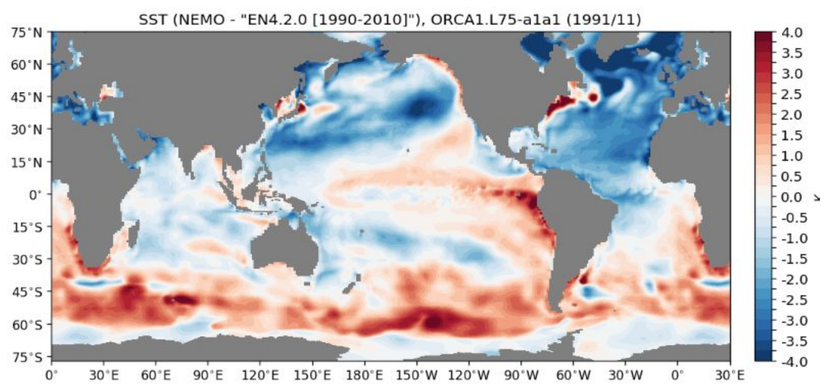


Fig. 3: Monthly Sea Surface Temperature (SST) bias from 1991 using the Spinup Outclass. .

Acknowledgments

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 641727.