

EGU21-4805

<https://doi.org/10.5194/egusphere-egu21-4805>

EGU General Assembly 2021

© Author(s) 2021. This work is distributed under the Creative Commons Attribution 4.0 License.



Bringing ESMValTool to the Jupyter Lab

Peter C. Kalverla, Stef Smeets, Niels Drost, Bouwe Andela, Fakhreh Alidoost, and Jaro Camphuijsen

Netherlands eScience Center, Amsterdam , Netherlands (p.kalverla@esciencecenter.nl)

Ease of use can easily become a limiting factor to scientific quality and progress. In order to verify and build upon previous results, the ability to effortlessly access and process increasing data volumes is crucial.

To level the playing field for all researchers, a shared infrastructure had to be developed. In Europe, this effort is coordinated mainly through the IS-ENES projects. The current infrastructure provides access to the data as well as compute resources. This leaves the tools to easily work with the data as the main obstacle for a smooth scientific process. Interestingly, not the scarcity of tools, but rather their abundance can lead to diverging workflows that hamper reproducibility.

The Earth System Model eValuation Tool (ESMValTool) was originally developed as a command line tool for routine evaluation of important analytics workflows. This tool encourages some degree of standardization by factoring out common operations, while allowing for custom analytics of the pre-processed data. All scripts are bundled with the tool. Over time this has grown into a library of so-called 'recipes'.

In the EUCP project, we are now developing a Python API for the ESMValTool. This allows for interactive exploration, modification, and execution of existing recipes, as well as creation of new analytics. Concomitantly, partners in IS-ENES3 are making their infrastructure accessible through JupyterLab. Through the combination of these technologies, researchers can easily access the data and compute, but also the workflows or methods used by their colleagues - all through the web browser. During the vEGU, we will show how this extended infrastructure can be used to easily reproduce, and build upon, previous results.