

Project: **1088**

Project title: **DKRZ_MIP_Pool_Analysis**

Principal investigator: **Stephan Kindermann**

Report period: **2020-11-01 to 2021-08-31**

The DKRZ_MIP_Pool_Analysis project wants to support data analysis activities requiring access to existing DKRZ data collections (accessible via /pool/data) with specific focus on the CMIP6 (and CORDEX) collections stored as part of the DKRZ data project ik1017.

This project supported in the reporting period individual users and research teams (having no existing project allocation at DKRZ) and is coordinated with similar efforts at IPSL (France), STFC (UK) and CMCC (Italy) as part of the IS-ENES3 project to provide coordinated European support for CMIP related data analysis activities.

Individual users were supported with a light weight (short term, small compute profile) access to analysis capabilities mostly via the DKRZ jupyter-hub installation (see <https://portal.enes.org/data/data-metadata-service/climate-analytics-service>). These users only need to pass a light weight registration process and are reviewed and selected locally at DKRZ.

Research groups with larger computational demands have to send proposals responding to specific calls which are then reviewed by the established IS-ENES review boards (separate technical and scientific board) based on well defined selection criteria. (<https://portal.enes.org/data/data-metadata-service/analysis-platforms>).

In the reporting period about 20 people used the service and spent around 5200 node hours from the available 8000 node hours until now, so all 8000 node hours will be spent in the allocation period (only 98 node hours expired in the reporting period). 6 research groups were granted access to DKRZ by the IS-ENES review committee. Summary reports on their activities will be collected and are currently pending.

A large online webinar/workshop with associated online live tutorial on how to exploit CMIP data via the provided DKRZ service was held in March with more than 100 participants (and more than 50 users registering and using the DKRZ jupyter-hub service). The event was co-organized by DKRZ and CMCC as part of the IS-ENES3 project as well as the EOSC-hub project ¹.

As part of this a collection of support material (e.g. jupyter notebooks) was made available which will also provide a good starting point for future users ².

In summary the mitigation steps to increase end user uptake of this service like outreach activities and extending the service to provide light weight access to jupyter-hub etc. (see an overview in the allocation report for 2020) were quite successful and in this reporting period and a diverse user community was supported:

- IPCC working group members producing figures for the IPCC report
- ESMValTool users (partly overlapping with IPCC working group members)
- Climate service centers and commercial companies wanting to have a short look at the data and the associated jupyter-hub service to support their activities.
- Research groups involved in different types of CMIP data analysis
- Individual researchers (e.g. from the climate impact community)

The IS-ENES3 project and the European Consortium of CMIP data providers plan to continue the service during 2022 and thus an extension of this project is requested. Additionally the demand for having short term grants for data analysis resources via jupyterhub grows, as the CMIP data pool at DKRZ currently provides the largest and most complete CMIP6 data collection in Europe. This type of resource request is currently not very well fitting into the long(er) term project application procedure at DKRZ.

¹ <https://www.dkrz.de/de/kommunikation/aktuelles/training-datenanalytik>

² ENES github repository: <https://github.com/IS-ENES-Data/Climate-data-analysis-service>