Project: **1149**

Project title: **CCLM2ESGF (COSMO-CLM/ICON-CLM to be published on ESGF)**

Principal investigator: **Christian Steger**

Report period: **2025-11-01 to 2026-10-31**

This project provides storage for EURO-CORDEX data produced by the CLM Community (www.clm-community.eu) and published via the Earth System Grid Federation (ESGF). These datasets form the basis for climate services, climate adaptation and policy consultancy in Germany and many other European countries, so the EURO-CORDEX simulations and their provision via the ESGF are very important.

Several years ago, DKRZ recommended that the CLM Community establish this project to provide the necessary storage for data published via the ESGF. The project was launched in 2019 and is renewed annually as required.

From our point of view, this procedure is not ideal because there is a risk that insufficient storage will be granted, or that previously granted storage will not be extended. This would mean that datasets already published would have to be withdrawn from the ESGF. Ideally, DKRZ would provide storage directly from its own resources for data published in the ESGF (e.g. /pool/data storage, as for CMIP and CORDEX-CMIP5). Unfortunately, this is not currently possible.

Therefore, it is of great importance that sufficient storage is allocated to this project and that it is extended for as long as the data must remain in the ESGF (several years).

The downscaling of the CMIP6 simulations for EURO-CORDEX using ICON-CLM and COSMO-CLM is at an advanced stage. Most of the ICON-CLM simulations produced within the UDAG project ('Updating the data basis for adaptation to climate change in Germany', bb1364) have already been published. All the simulations for the EUR-12 domain have been published. Publication of the high-resolution (3 km) data for the Central European domain (CEUX-3) is pending due to issues with the ESGF.

Other institutions in the CLM Community from across Europe also contribute to EURO-CORDEX, using COSMO-CLM and ICON-CLM. Simulation production is coordinated within the CLM Community. Some groups will use model configurations that include ocean models or other components of the Earth system. Some of this data has already been published.

Most of the data for the planned EUR-12 simulations with ICON-CLM has already been transferred to project storage and published in the Earth System Grid Federation (ESGF). These include simulations for the SSP1.2-6, SSP2-4.5, SSP3-7.0 and SSP5-8.5 scenarios, with boundary conditions from the global climate models MPI-ESM1-2-HR, EC-Earth3-Veg, CNRM-ESM2-1, MIROC6 and NorESM2-MM.

Furthermore, COSMO-CLM simulation data and COSMO-CLM simulation data with an ocean model for the North and Baltic Seas are already stored in /work/bb1149. The total amount of data mentioned above is currently 710 TB (06 October 2026).

Last year, colleagues from DKRZ who participated in UDAG also used the project storage as intermediate storage for UDAG 3 km data (additional variables to the ones that will be published in the ESGF), which will be published in the WDCC (currently ~500 TB). This data will be deleted once it has been transferred to WDCC. Additionally, there are around 200 TB of model output that have now been CMORised and require deletion.

Over the next months, we expect to receive several hundred terabytes of data from partner institutions in the CLM Community who are running additional EUR-12 COSMO-CLM simulations, in some cases coupled with ocean and hydrology models (Parflow). Additionally, due to an issue with the land-sea mask in the Mediterranean, we must reproduce ICON-CLM 12 km simulations with boundary conditions from MIROC6. Colleagues from CMCC in Italy may be able to publish their data on an ESGF node in Italy. If this happens, the storage space allocated for these simulations would not be necessary. However, this has not yet been decided, so we are including the data in our calculations.

In summary, 710 TB of the project's 1,915 TB of storage is already in use for data published in the ESGF. Most of the remaining storage is currently being used by DKRZ for data that will be published in WDCC. This data can be deleted afterwards. According to our estimate, the remaining storage of ~900 TB should be sufficient for the data currently being produced by the CLM Community for EURO-CORDEX. Therefore, we are not requesting additional storage for this project, merely an extension to the 1,915 TB of storage that have already been allocated. If CMCC can publish their data in Italy, we will reduce the request for the next allocation period accordingly.