

Project: **1363**

Project title: **CORDEX-FPSCONV2ESGF**

Principal investigator: **Klaus Görgen**

Report period: **2024-05-01 to 2025-04-30**

Project report

Project 1363 is a pure data project, whose entire storage volume on Levante's WORK filesystem is used to publish data from the CORDEX-FPSCONV project through the DKRZ ESGF data node.

The reporting period is the first allocation period, which saw the publication of the first batch of CORDEX-FPSCONV data.

The following main activities took place to date in the previous allocation periods since 2023-07:

- CMORized and QA-checked netCDF files were uploaded as parallel data streams from the jsc-cordex data exchange server, which is hosted at Jülich Supercomputing Centre Helmholtz on the JSC Cloud OpenStack VM cloud infrastructure, using specific, newly developed UFTP upload scripts.
- To ensure data consistency and check once more for completeness after data transfer, sha256 checksums as generated on jsc-cordex from the reference data repositories were checked on Levante; this was in addition to the sha1-based checks during UFTP transfers.
- The DKRZ data publication agreement was organized within the CORDEX-FPSCONV community (support contact persons), finalized, and handed in.
- Minor metadata and Data Reference Syntax issues were fixed directly on Levante after data checking by DKRZ Data Management staff.
- Before data publication (April 2024) through ESGF by DKRZ Data Management, all published data were backed up on tape archive, using dedicated scripts with `slk archive`.

These procedures were preceded by a test-publication of a single daily variable in August 2023 to check for the work-flow and the functioning of the ESGF ini-file.

Data of the first batch of CORDEX-FPSCONV data is now available under the "CORDEX-FPSCONV" project ID.

We split the CORDEX-FPSCONV simulation results on jsc-cordex into batches, each containing several complete simulation experiments, based on the availability of the final data (CMORized and QA-checked). We originally assumed that there would be a second batch during 2024/25 and a third, final batch in 2025/26 until all data will have been made available through ESGF data nodes; this planning is still in place albeit shifted by one year as some groups are still preparing data for ESGF dissemination in parallel to analysis work.

A data set publication, based on the CORDEX-FPSCONV publication of data via DKRZ ESGF and other ESGF data nodes by Goergen et al. is close to submission for Scientific Data.