

EGU24-17111, updated on 17 Mar 2025
<https://doi.org/10.5194/egusphere-egu24-17111>
EGU General Assembly 2024
© Author(s) 2025. This work is distributed under
the Creative Commons Attribution 4.0 License.



Cloudifying Earth System Model Output

Fabian Wachsmann

Deutsches Klimarechenzentrum, Datenmanagement, Hamburg, Germany (wachsmann@dkrz.de)

We introduce `eerie.cloud` (`eerie.cloud.dkrz.de`), a data server for efficient access to prominent climate data sets stored on disk at the German Climate Computing Center (DKRZ). We show how we “cloudify” data from two projects, EERIE and ERA5, and how one can benefit from it.

The European Eddy-rich Earth System Model (EERIE) project aims to develop state-of-the-art high-resolution Earth System Models (ESM) that are able to resolve ocean mesoscale processes. These models are then used to perform simulations over centennial scales and make their output available for the global community. At present, the total volume of the EERIE data set exceeds 0.5PB and is rapidly growing, posing challenges for data management.

ERA5 is the fifth generation ECMWF global atmospheric reanalysis. It is widely used as forcing data for climate model simulations, for model evaluation or for the analysis of climate trends. DKRZ maintains a 1.6 PB subset of ERA5 data at its native resolution.

We use `Xpublish` to set up the data server. `Xpublish` is a python package and a plugin for Pangeo's central analysis package `Xarray`. Its main feature is to provide ESM output by mapping any input data to virtual `zarr` data sets. Users can retrieve these data sets as if they were cloud-native and cloud-optimized.

`eerie.cloud` features

- Parallel access to data subsets on chunk-level
- Interfaces to make the data more FAIR
 - User friendly content overviews with displays of `xarray`-like dataset representations
 - Simple browsing and loading data with an intake catalog
- On-the-fly server-side computation
 - Register simple `xarray` routines for generating customized variables
 - Compression for speeding up downloads
- Generation of interactive geographical plots, including animations

`Eerie.cloud` is a solution to make EERIE data more usable by a wider community.