EGU24-8146, updated on 25 Jul 2024
https://doi.org/10.5194/egusphere-egu24-8146
EGU General Assembly 2024

Xarray-regrid: regridding with ease

Bart Schilperoort[1], Claire Donnelly[1], Yang Liu[1], and Gerbrand Koren[2]
[1]Netherlands eScience Center, Delft, Netherlands (b.schilperoort@esciencecenter.nl)
[2]Copernicus Institute of Sustainable Development, Utrecht University, Utrecht, Netherlands

In geosciences different sources of data are often on different grids. These can be at different resolutions, but also have the grid centers at different locations. To be able to use these different sources of data in models or analyses, they have to be re-projected to a common grid. Popular tools for this are the command-line tool 'Climate Data Operators' (CDO) and the Earth System Modeling Framework (ESMF).

These tools work well but have some downsides: CDO is a command-line tool and as such the regridded data has to be written to disk. ESMPy, the Python package for ESMF, is only available on Linux and Mac OSX, and does not support out-of-core computing. Both tools rely on binary dependencies, which can make them more difficult to install. Additionally, many geoscientists already use xarray for analyzing and processing (netCDF) data.

For this use case we developed xarray-regrid, a lightweight xarray plugin which can regrid (rectilinear) data using the linear, nearest-neighbor, cubic, and conservative methods. The code is open source and modularly designed to facilitate the addition of alternative methods. Xarray-regrid is fully implemented in Python and therefore can be used on any platform. Using Dask, the computation is fully parallelized and can be performed out-of-core. This allows for fast processing of large datasets without running out of memory.

Xarray-regrid is available on the Python Package Index (*pip install xarray-regrid*), and its source code is available on GitHub at https://github.com/EXCITED-CO2/xarray-regrid