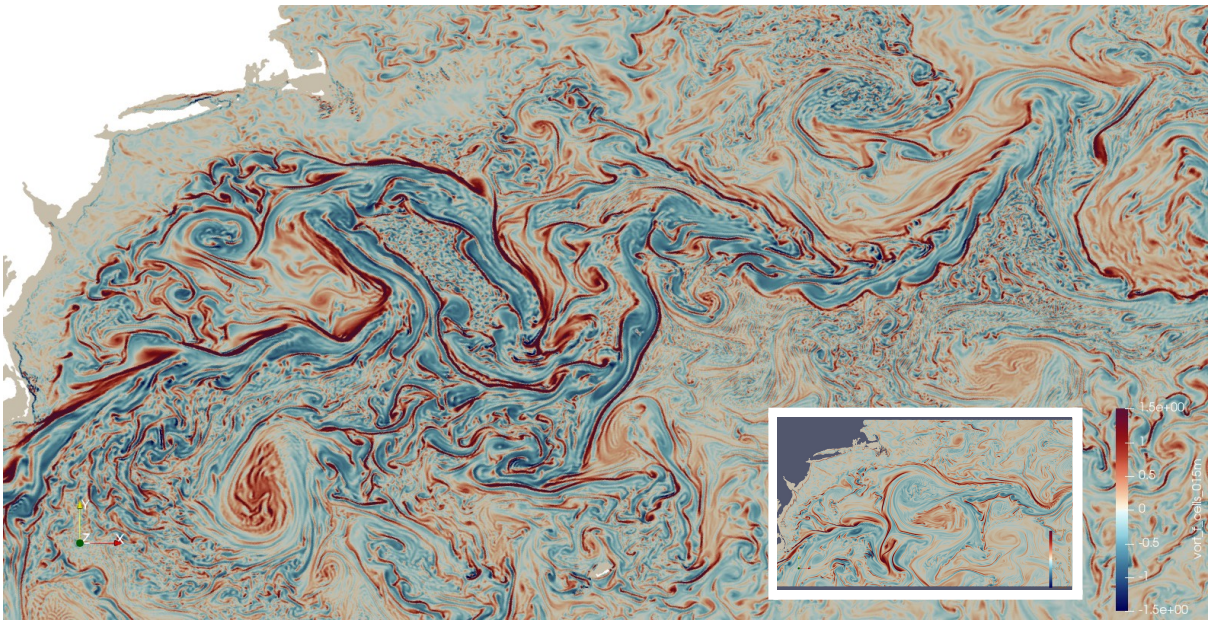
Report: ESiWACE: Scalability of Earth System Models

Project: 1040
Project title: ESiWACE: Scalability of Earth System Models
Principal Investigators: Florian Ziemen (DKRZ), Daniel Klocke (DWD, MPI-M), Bjorn Stevens (MPI-M), Joachim Biercamp (DKRZ)
Report period: **2020-11-01 to 2021-08-31**

ICON SIMULATIONS

In late 2020, we finished the first coupled one-year simulation at 5 km resolution. In the meantime, MPI-M has performed another almost two years on institutional resources.
In the context of bk1040, we mainly performed the planned development simulations at 1.25 km resolution. Most of the computational resources used here were spent on the ocean component, and contributed to the setup for a two-month 1.25 km resolution simulation that was performed on MPI-M resources.



Near-surface Rossby-Number in the North Atlantic after one month of 1.25 km resolution ocean simulations with ICON-O. The color scale covers ±1.5. The inset shows the same field 12h after the initialization from the 5 km resolution ocean. The comparison with the eddy field that has developed after a month of simulation clearly demonstrates the massive increase in resolved eddy activity associated with the quadrupling of the model resolution.

For the atmospheric setup, we are still facing challenges during the initialization phase of the model and with the lower boundary conditions. We therefore were not able to use much of our allocated resources.

DYAMOND Data library

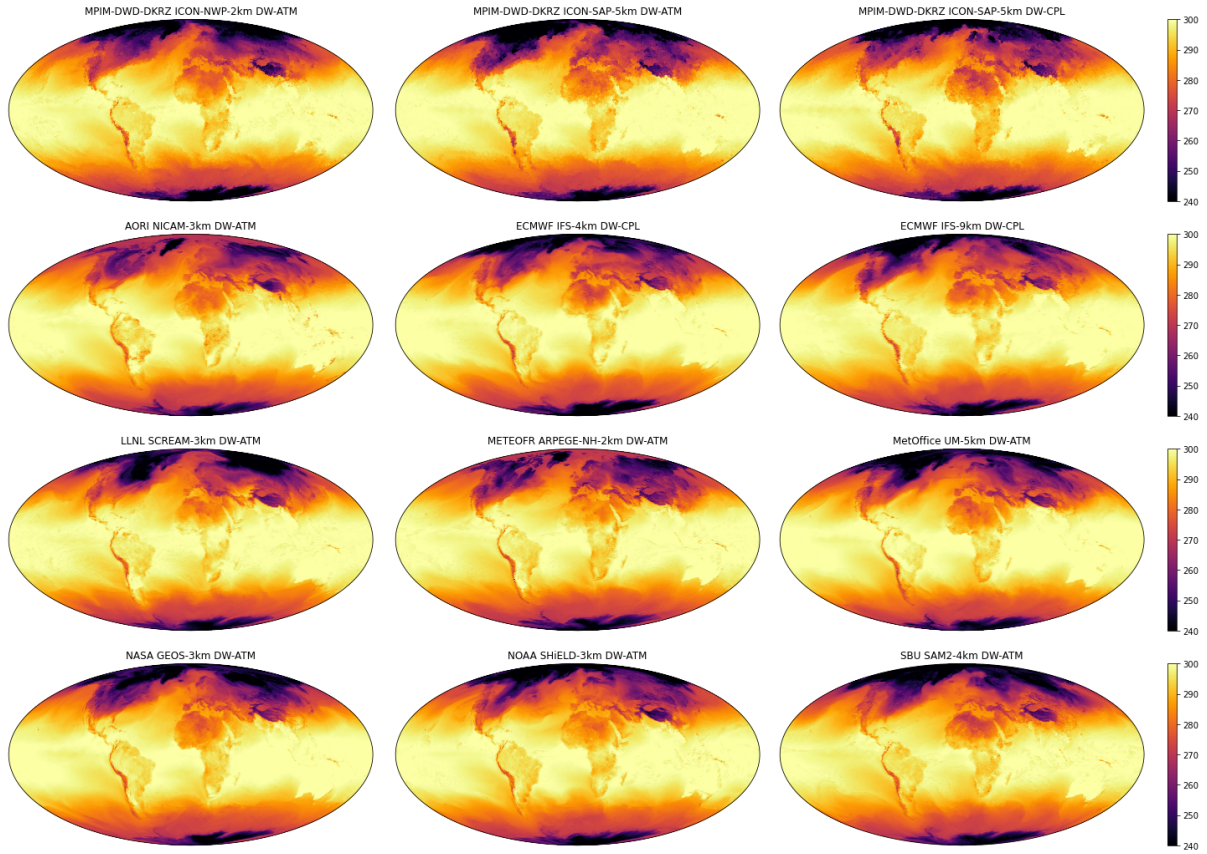
We are using the disk space for the DYAMOND model intercomparisons.
All data sets are available via the *DYAMOND Data Library*[1] as a centralized repository, where users can access data on disk and create requests for downloading data sets from the tape archive.
While we initially planned to simply store the incoming data in the formats provided by the other modeling teams, we decided to put in additional effort to standardize the incoming data. This process is still ongoing, so not all contributions we have received are in the standardized library yet.
We are standardizing the data for several reasons:

1. By doing the standardization once, we spare all other users from having to do this and repeating the invention of the wheel over and over. We expect a net increase in scientific productivity from this burden sharing, although it comes at our expense.
2. By providing standardized files, we reduce the incentive for users to keep "standardized" copies of the files on disk, and thus reduce the load on the file system.
3. Moving the data to compressed netcdf4 reduces file sizes in several cases, again reducing the load on the file system.
4. The standardized files can be directly accessed via freva and the python intake-esm library, thus providing more user-friendly interfaces, and increasing the productivity of the scientists analyzing the data (see below for a set of example plots).
5. The semantic access with freva and intake-esm allows us to prepare for the move to semantic db-access that we expect to result from activities planned in ACROSS (H2020) and WarmWorld (BMBF, expected to start in 2022).

The intake-esm access will be a foundation for the NextGEMS hackathon in October, and thus directly contributes to NextGEMS.

To further ease the user access we have partnered with the *Climate Informatics and Technologies* group of DKRZ and have set up https://gems.dkrz.de/, where users can search the standardized datasets using a convenient web interface at https://gems.dkrz.de/solr/data-browser/ (use the guest login at the top right). Freva also can be used with command-line tools, e.g. in combination with CDO[2]. Furthermore, together with NextGEMS, we are developing a documentation platform at https://easy.gems.dkrz.de/. We are aiming at making these platforms coss-project activities and are applying for funding via WarmWorld for further development of the freva interface.

[1] https://easy.gems.dkrz.de/DYAMOND/dyamond-library
[2] https://easy.gems.dkrz.de/DYAMOND/Winter/index.html#using-freva-and-gems-dkrz-de-to-browse-the-data



Surface air temperatures on Feb 5, 2020 at 6AM UTC. With the standardization and intake-esm we were able to generate these plots in one easily customizable script[3]. P.S.: Spot the two outliers we found with these plots. ;)

[3] https://easy.gems.dkrz.de/Processing/Intake/grid-guessing-test.html