Project: 1344
Project title: Horizon EU-project EERIE [**E**uropean **E**ddy-**RI**ch **E**arth System Models]
Principal investigator: Jin-Song von Storch (m221050, MPIM)
Project leads: Thomas Jung (AWI), Fabian Wachmann (DKRZ), Nikolay Koldunov (AWI), Dian Putrasahan (MPIM)
Report period: 2023-01-01 to 2023-10-31

Achievements in the previous allocation period:

MPI-M

New ICON configuration for EERIE

A change in the configuration for ICON in EERIE from R2B8L128-oce/R2B8L90-atm to R2B9L72-oce/R2B8L90-atm was made to better address the role of ocean mesoscale eddies on the climate. R2B8L128 for the ocean gives a nominal horizontal resolution of 10km, which technically does not resolve mesoscale eddies in the high latitudes like the Southern Ocean and subpolar/polar oceans. Since the project is about the role of ocean mesoscale eddies, we found it much more appropriate to go to R2B9L72, which has a horizontal resolution of 5km. We reduced the vertical levels and have it inline with EPOC, where vertical resolution of depths at which we are scientifically concerned about are still resolved, but we gain back computational throughput that would allow us to complete the long-integrations set forth in EERIE project goals.

Increased throughput

Continued collaboration with DKRZ and CIMD from MPIM, and project joint efforts (MPI-M Ruby/Sapphire, EPOC, CliCCS A6, NextGEMS, DestinE), helped us to get a technically-ready and an almost climate-ready configuration. We have also been able to increase the throughput of simulations, from ~1 SYPD to ~1.33SYPD on 400 nodes, which translates to 7200 node hours per simulation year. Note that this improvement, while on absolute terms is only ~33%, it is actually much more given that we have doubled the horizontal resolution and reduced vertical resolution by 40%. However, there's still much progress to be made if we are to reach our goals in EERIE, with multiple centennial simulations using coupled R2B9/R2B8 ICON-ESM. The coupled R2B9/R2B8 version is named ICON-ESM-ER.

Heterogeneous setup

We (together with DKRZ and CIMD) made plans to test out a heterogeneous setup with EERIE configuration in order to increase throughput. It was not until about a month ago that a heterogenous setup for ICON (with a lower resolution configuration) showed preliminary signs that it is technically working (work done by DKRZ). We await for this test setup to functionally run and will apply it to the EERIE setup. In parallel, we did get a heterogeneous setup with EERIE configuration working on JSC, and would therefore use the knowledge gained from there to help get it to work on Levante.

Alignment of model output with data request

Much effort had been placed in aligning the model output to the data request for the project. For example, ssh had to be corrected to account for sea-ice. Mixed layer depth needed to be referenced to 10m depth, not the surface. Several quantities could not be directly outputted and backtracking of required variables to compute said quantities were made. Changes in ICON atmospheric variable names had to be included. Balancing data storage and the desire for high temporal resolution data were made. There was a strong request for 3D daily ocean output as the project is meant to understand how eddies can impact circulation, air-sea interactions and the climate. Due to the enormity of the data storage, we opted for performing time slices with such output.

Synergy with other projects

EERIE wants to adopt the NextGEMS setup to keep the number of configurations and settings between the various projects to its minimum, as well as to enable cleaner comparisons/analysis and more synergy between various projects (EERIE, NextGEMS, EPOC, DestinE, etc.) Also, an energetically consistent and energy conserving climate system would be desirable. Yet, EERIE has an added challenge to contend with, which is getting the simulations climate-ready.

Climate readiness of EERIE ICON simulation (spin-up)

A climatically-stable (climate-ready) control simulation is crucial for having long-integration climate runs, and assessing changes under a warming world. We technically got it running and spun up the coupled system for 46 years first, then later extended the spin-up by 6 years with changes to sea-ice model parameters (explained further below in “Arctic holes”). Thus, after 52 years of spin up, we then branched off a control 1950-forcing run. A brief description of the climate state during the spin-up phase is given here.

In the atmosphere, top of the atmosphere (TOA) short wave (SW) and long wave net fluxes are out of balance and both are too low (230 and -224 W/m^2, respectively) The imbalance points at a leak in our system, the overall too low fluxes come with too many (liquid) clouds, too much reflected SW, and too cold climate, which is expected when using a 10-km ICON-A without parameterized convection.

In the ocean, the global mean surface temperature cools down within the first 20 years by about 1^{o}C and seems to further cool down at a lower rate (0.1^{o}C per decade) (Fig. 1). Accordingly, the total heat flux at the ocean surface is negative with an initial magnitude of 5 W/m^2 that went down to below 1 W/m^2 within about a decade (Fig. 2).

The strength of the Atlantic meridional overturning circulation (AMOC) has quasi leveled between 20 and 22 Sv after speeding up for about 25 years (Fig. 3, black line) but is far from stationary as the extended spin-up and the first 15 years of the control run reveal (Fig. 3, red and blue lines). Here, the negative trend may already be indicative of a long term variability of the AMOC in the control run.

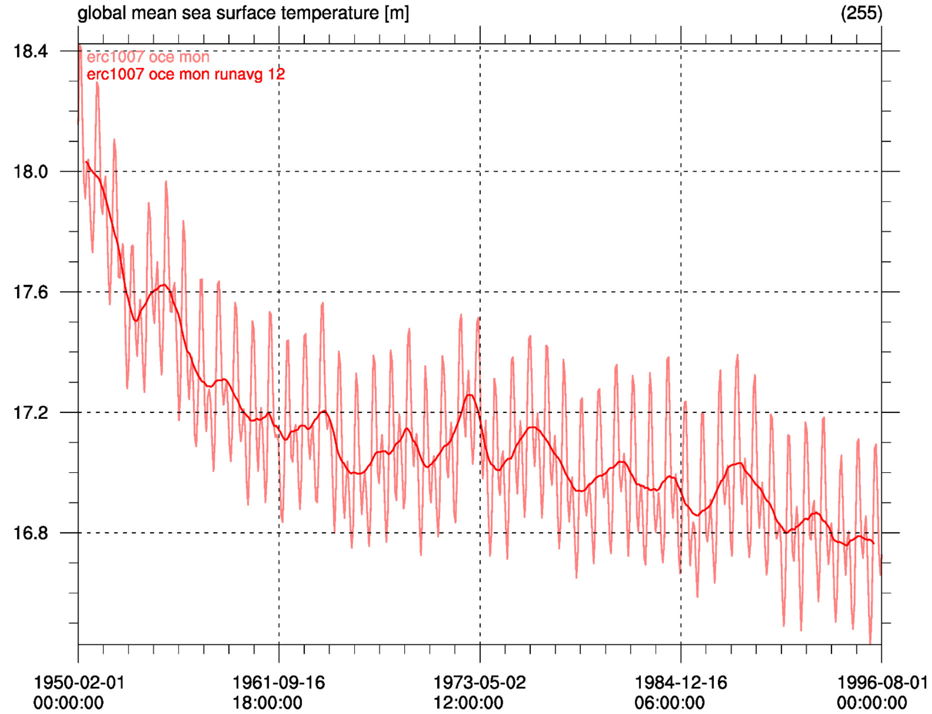


Figure 1: Evolution of global mean sea surface temperature in °C, time series of monthly means with (bold) and without (thin) smoothing by a running mean.

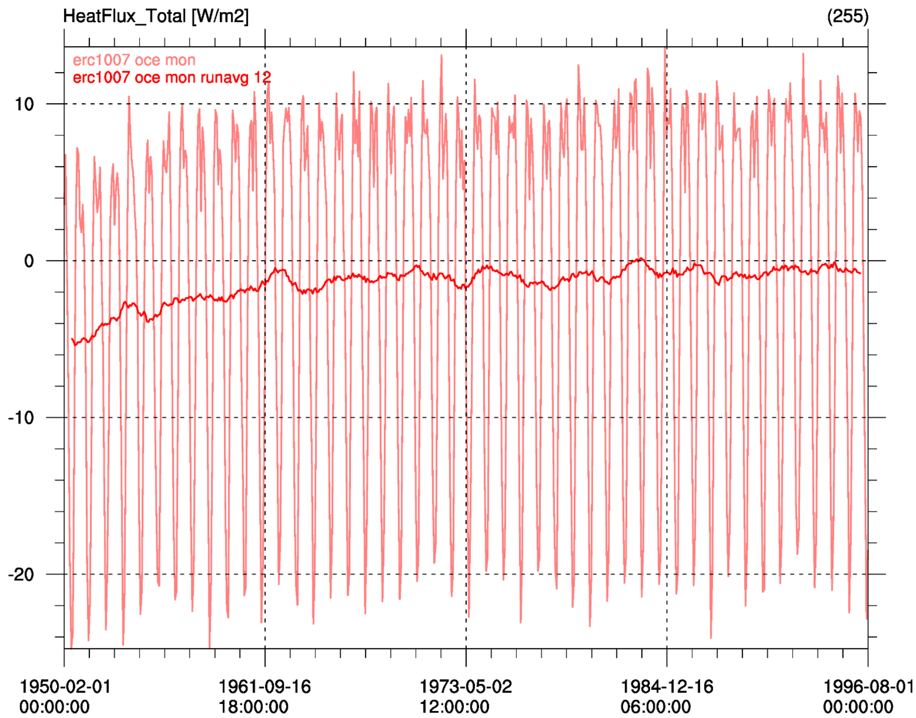


Figure 2: Evolution of total heat flux at the sea surface in W/m^2, time series of monthly means with (bold) and without (thin) smoothing by a running mean.

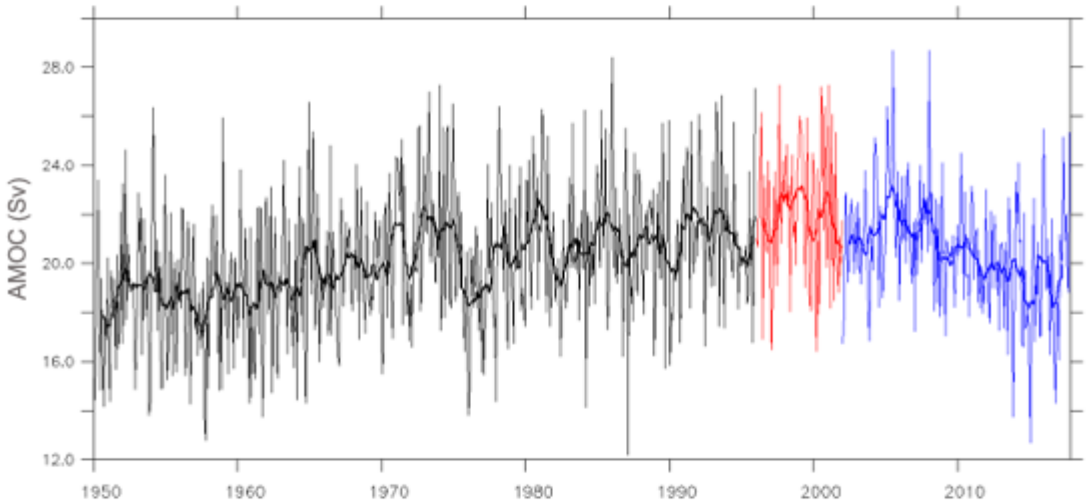


Figure 3: Evolution of the AMOC (in Sv) at 26N and 1000m depth in the spin-up phase plus the extended spin-up (black and red lines) and in the control run (blue line). Time series of monthly means with (bold) and without (thin) smoothing by a running mean.

During the spin-up phase, we noticed that sea-ice holes are intermittently forming in the Arctic Ocean (Fig. 4). As indicated by one case study, such regions with low sea ice concentration coincide with high sea ice thickness, start to form at the end of summer, move around the Arctic Ocean, change size during the course of the successive fall and winter seasons and may last until at least the end of April (not shown). This can have implications for deep water formation. We are currently tackling this feature/bug, which resulted in 6 years of additional spin-up ('extended spin-up') in which first changes in parameters in the ice model were tested. These parameters describe changes in ice-thickness distribution during both freezing and melting, plus a reduced response to wind stress in the vertical mixing scheme under ice was introduced. More years of runs to further test and address the issue of sea-ice holes are to follow. Nevertheless, a control run had to be branched off from the extended spin-up as explained in the next paragraph.

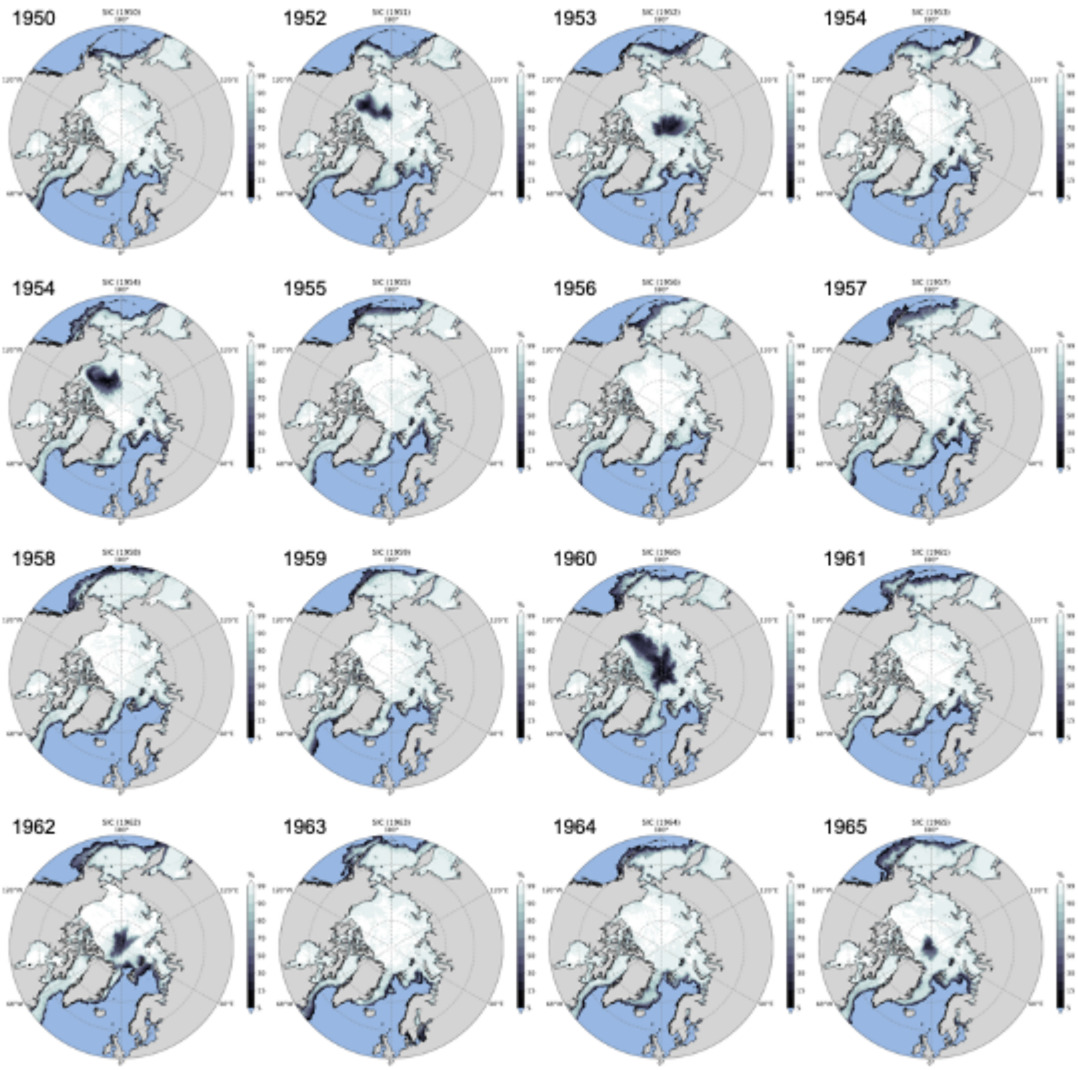


Figure 4: February Arctic sea ice concentration in the spin-up run for the first 16 years. Regions with low coverage (dark shading) are considered ‘holes’ in the ice.

EERIE ICON control-1950 simulation

The EERIE hackathon is coming up on Nov. 9-10, 2023, and we need to have 20-30 years of simulation prepared before then. As such, we branched off after 52 years of spin-up run, and began a control-1950 run with currently 15 years of high resolution data available. We conduct preliminary analysis to assess mesoscale air-sea coupling from a Lagrangian perspective by tracking Agulhas rings in the Southeast Atlantic and building composites from these tracked eddies. Fig. 5a shows the long-lived tracked eddies associated with Agulhas leakage, and their averaged composites of 700km high-pass filtered SST, 10m wind speed, precipitation, downwind SST gradients, and wind stress divergence (Fig. 5b-f respectively). SST anomaly peaks in the southwest quadrant of the composite, and coincides with faster surface wind speed anomalies and larger latent heat flux out of the ocean (turbulent heat flux composite not shown but easily inferred as warm SST anomalies and faster winds act in the same direction to flux more heat out of the ocean). This suggests warm SSTs could induce faster wind speeds through the vertical mixing mechanism and together enhance heat flux into the atmosphere. On the other hand, precipitation anomaly peaks on the southeast quadrant, which coincides with the downwind SST convergence and windstress convergence, again suggesting the vertical mixing mechanism at work on the mesoscales.

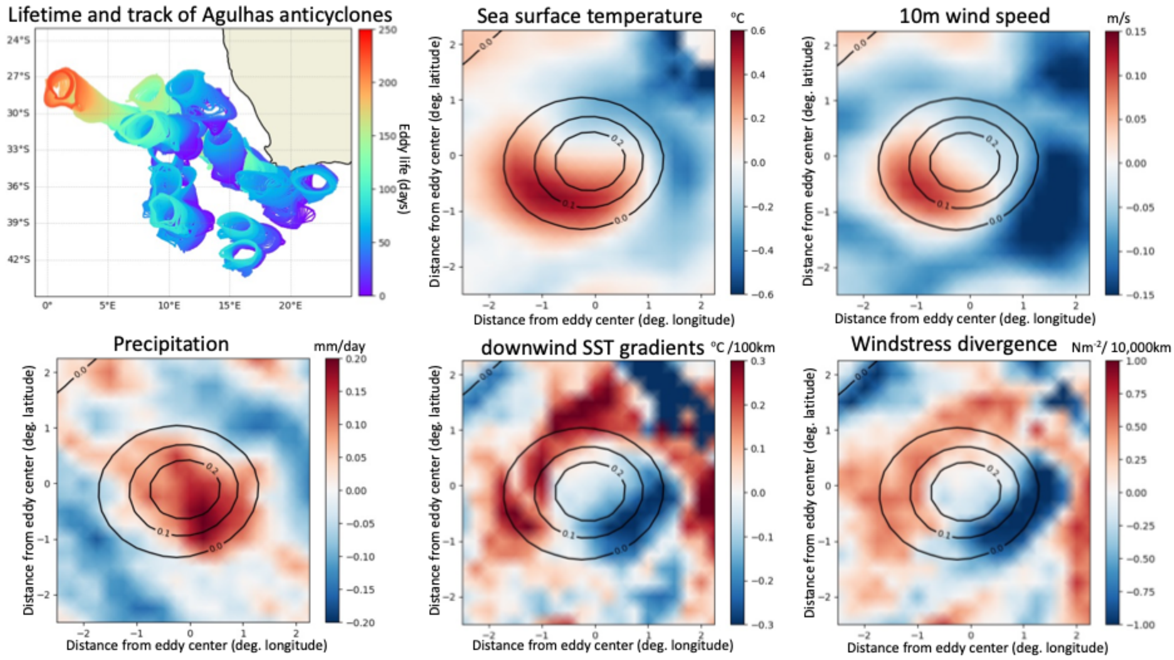


Figure 5: a) Lifetime (days) and track of Agulhas anticyclones that live longer than 60 days from a 7-year period of EERIE ICON simulation output.; averaged composites of tracked Agulhas anticyclones for 700-km high-pass filtered b) SST,; c) 10m wind speed,; d) precipitation,; e) downwind SST gradients,; f) windstress divergence.

Age tracer as a proxy for carbon uptake

We have been working towards the EERIE project deliverable of implementing an online ventilation age diagnostic in ICON. Such diagnostics can be used to produce a “Green’s function for the ocean which allows us to calculate distributions of any number of passive tracers in the ocean — from anthropogenic carbon to CFCs. This will allow us to explore the role of eddies in both transporting and storing anthropogenic heat and carbon. This approach has not been pursued before in the context of an eddying ocean model; however, when successfully implemented will enable the offline analysis of quantities too computationally intensive to compute otherwise.

To obtain the Green’s function we must first calculate the mean age and the Peclet number of the flow in an ocean model. Figure 6 shows both these quantities at a depth of 50 m in a coarse R2B4 prototype model. Regions marked in red in Fig. 6b mark areas where advective tracer transport is dominant and in blue where diffusive fluxes dominate. We are currently working towards getting the age tracer working in a 5 km resolution model. In such a model we would expect to see advection to play a much more important role with advective tracer pathways better resolved.

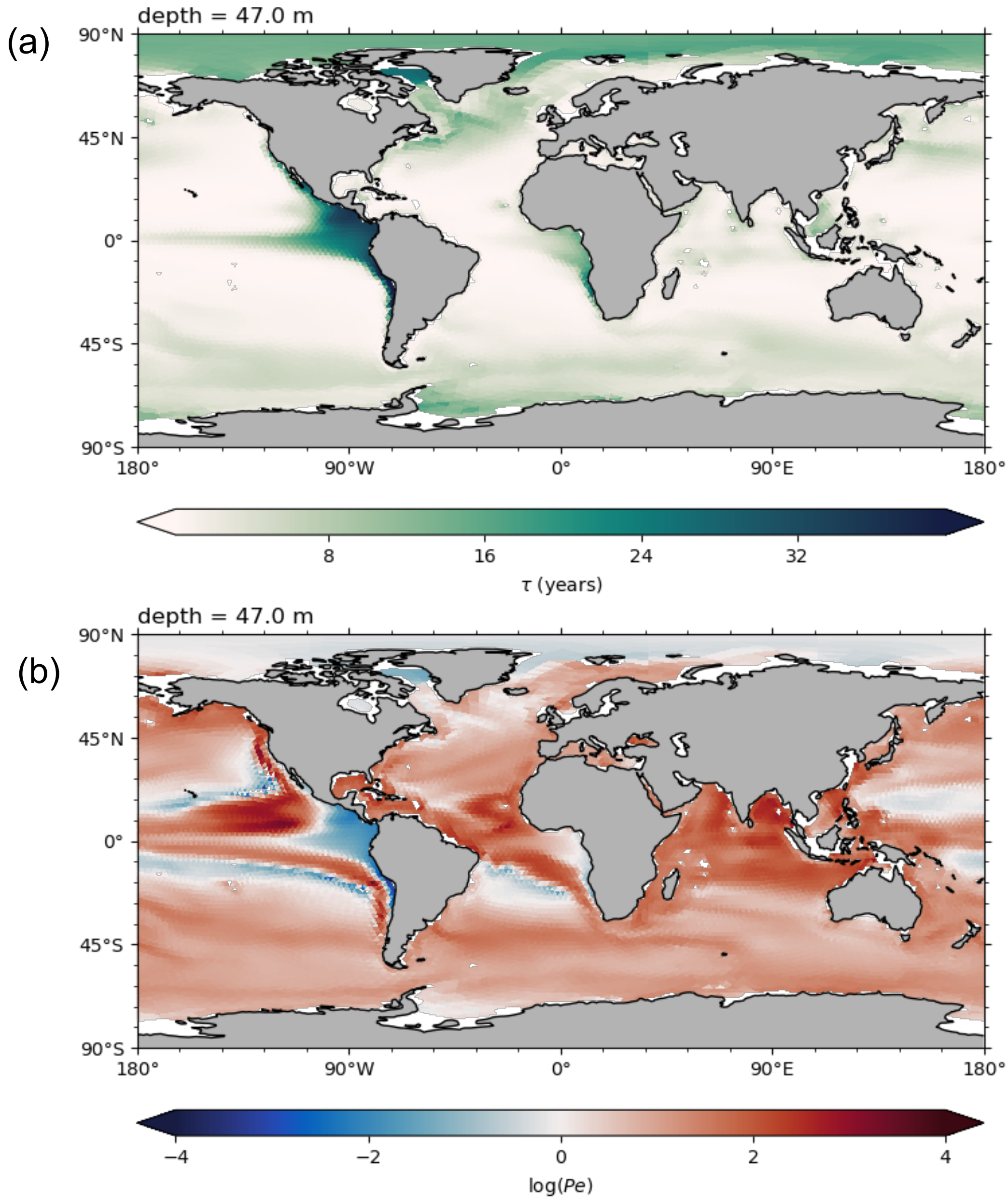


Figure 6: a) Ventilation age and (b) logarithm of the Peclet number in a coarse (R2B4) ICON-O model. Areas in red show regions where the flow is predominantly advective, and in blue regions tracer transport is predominantly diffusive.

AWI

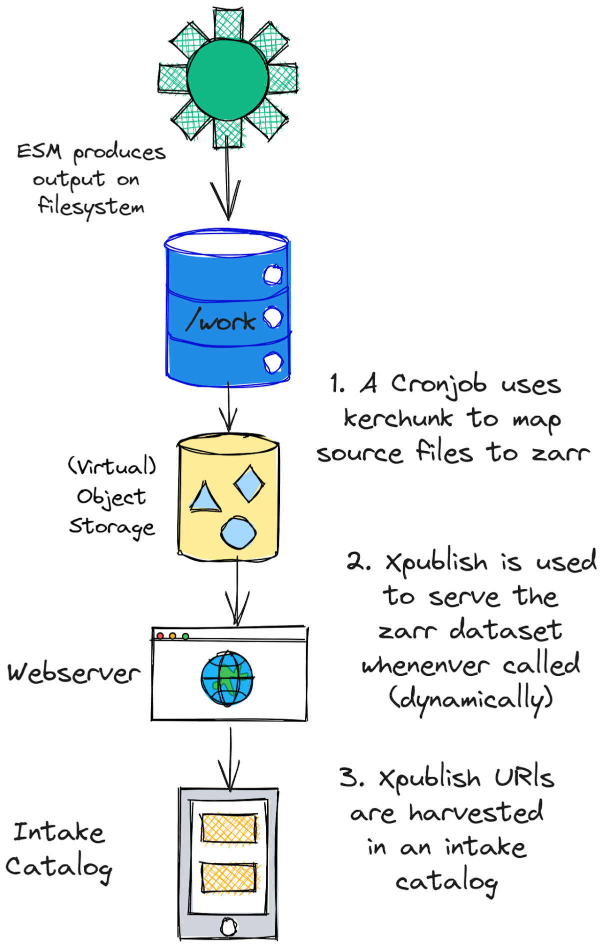
During the reporting period, the focus for AWI was primarily on the utilization of disk space and management of simulation data, particularly in light of the project's broader objectives and the challenges encountered in the development phases. We experienced delays in the coupled system's development that prevented us from running both low resolution simulations (planned for DKRZ) and high resolution simulations planned for JSC (JUWELS).

While the coupled simulations were delayed, we conducted an evaluation of high-resolution ocean-only NG5 configurations on JUWELS. This data was transferred to the DKRZ using uftp. The transfer process uncovered a bug in the Uftp client. Eventually, the general uftp transfer could be optimized by changing the Uftp server configuration and fixing Uftp client bugs at DKRZ.

We embraced the opportunity to gain experience in managing and post-processing of a similar dataset, generated by nextGEMS project. The nextGEMS cycle3 data have the same NG5 resolution, that will be used in high resolution IFS-FESOM EERIE runs, and two times higher resolution in the atmosphere. We have tested efficient ways of data interpolation to regular grid, creation of intake catalogs, converting data to zarr format. Resulted datasets were tested by users during nextGEMS Cycle 3 hackathon in Madrid. We gain important experience on working and distributing the data that will be further used during the EERIE project for more efficient and user-friendly data management.

DKRZ

Using the ICON-ESM-ER spin-up for EERIE, we set-up a post processing workflow for providing efficient data access to EERIE production NetCDF output by using a python tool set consisting of kerchunk, intake and xpbulish. For each output pattern (e.g. “atm_2d_1mth_mean” in ICON), we generate one so called kerchunk json file which contains all necessary metadata to map the entire time series to the zarr format for simple, high granular and parallel data access. During this step, we modify coordinates and metadata of the virtual dataset without writing additional data to disk: We add a grid description, attributes and improve the time axis. We collect the kerchunked files in an intake catalog which is provided to Levante users for open access via github and in /pool/data/Catalogs.



Based on this intake catalog, we set up an archival workflow which lossy compresses the output with a blosc compressor and conservation of 99.9% information to ensure further kerchunk application and analysis and to reduce the required storage space. It also targets to write netcdf files with a disk size of 1-100GB to optimize transfers to and from the archive. For the ICON-ESM-ER spin-up experiment, we already archived 46 years of daily and monthly output (not the full EERIE namelist) which correspond to 30TB. We generate an intake catalog for archived files (right now available under /work/bm1344/DKRZ/intake/dkrz_eerie_archive.yaml) which is enriched with all metadata by the HSM meta database. This in turn has previously harvested the attributes from the NetCDF files.

Fig. 7:
Data workflow for hosting EERIE production output

We also implemented a data server which hosts the raw output of the data live on eerie.cloud.dkrz.de based on the kerchunk intake catalog. We use xpublish and nginx for this set-up. A sketch of the entire workflow is provided in Fig. 7.

Another post-processing of EERIE simulation we have started is the remapping to pressure levels and on regular grids. Remapped data of the control run is available in intake catalogs for disk and cloud. This data is of high value for the community as standard tools like ESMVal can be used for the analysis. It is also the first step for the data standardization which is part of the ESGF pipeline and eventually results in the data publication.

All intake catalogs are combined in one main eerie catalog at https://github.com/eerie-project/intake_catalogues .