Report on the joint project "WarmWorld" bk1341

Project: **1341**
Project title: **WarmWorld**
Principal investigator: **Julia Duras**
Allocation period: **2024-01-01 to 2024-12-31**

1 Calibration of liquid-cloud microphysics in ICON *(Leipzig University, AWI)*

Consumption by end of October: 21.000 CPU node hours

In the WarmWorld-Better project, working group 2, we have so far used in total 21.000 CPU node hours for

- running global test simulations with ICON-AES,
- pre- and postprocessing of input and output data,
- and testing new developments within the ICON 1-moment microphysics scheme.

To create a base simulation, we have tested ICON using different configurations (e.g. with and without *tmx*), resolutions (R2B4 and R2B9) and microphysical schemes (1-moment and 2-moment) to get acquainted with the details of the model performance. These test simulations were run for 5 days each to include adjustment time and to ensure performance stability.

Based on this we have decided on one base configuration to start with for further developments. These new developments which are being tested at the moment target specifically at

- the use of external data (CAMS reanalysis) to read in fields of cloud condensation nuclei (*ccn*) within the 1-moment cloud microphysical scheme (*mo_cloud_mig*),
- scaling and conversion from *ccn* to cloud condensation nuclei (*cdnc*),
- and the coupling of cloud microphysics and radiation (using *rte-rrtmgp*) via a common 3-D field of *cdnc*.

This work is ongoing.

2 Development of ICON turbulence parameterization *(University Hamburg)*

Consumption by end of October: 4.300 CPU node hours

For the development of the software package turbulent mixing (tmx), 4.283 CPU node hours were used for validation tests. The tmx package was refactored to improve the modularity of the software package, and new diagnostics are provided for the velocity gradient tensor to simplify the implementation of turbulence parameterizations. Validation tests were successfully performed for the warm-bubble test case on a torus grid, as well as for an AMIP-R2B8 configuration. For the AMIP runs with a 3-D Smagorinsky scheme, single and double precision runs were performed for 3 simulated months on 300 nodes. The comparison of the different AMIP test cases shows that the refactoring leads to deviations in global quantities after approximately 6-10 days, which depends on the numerical precision (single precision leads to earlier deviations). This behaviour is to be expected, as round-off errors arise due to the changed computation order by the refactoring and are thus amplified by the internal variability of the climate simulations.
The project is currently working on the implementation of a new turbulence parameterization based on the Leonard term for the subgrid-scale stress tensor. In addition, the coordination and integration of further turbulence parameterizations within the turbulent mixing package is being pursued in collaboration with the WarmWorld groups led by Reiner Schnur (MPI-M) and Linda Schlemmer (DWD).

3 Frozen microphysics and limited-area model testings *(MPI-M)*

3.1 Fine-tuning of ice/snow microphysical processes

Consumption by end of October: 10.000 GPU node hours

Over the past year, the granted GPU-allocation and part of the CPU's has allowed us to perform a comprehensive parameteric study on the first-moment microphysics scheme. The goals were two-fold: 1), to identify a compact set of parameters allowing to fine-tuning the mass fraction ratio between snow and ice, and 2) assess the radiative impact of snow, recently accounted for in the flux calculations. This is part of an ongoing effort to broadly investigate the role of frozen condensates in tropical meteorology, and adress the overpopulation of snow measured in ICON. These are:

- The ice fall velocity v_i, which determines the depletion rate of cloud ice.
- The snow fall velocity v_s, which determines the depletion rate of snow *and* the collection rate of falling snow with cloud water.
- A prefactor coefficient in front of all snow producing processes, in an effort to mitigate its creation.

The allocation allowed us to run one full year on global R2B8 mesh for all of the following six configurations: 1) a case using the original microphysics scheme (labelled *reference*), 2) a case doubling the snow fall velocity v_s (labelled *2* v_s), 3) a case halving the ice fall velocity v_i (labelled *0.5* v_i), 4) a case halving all snow production rates (labelled *0.5 x2S*), a case doubling the snow fall velocity and halving that of ice (labelled *2* v_s*-0.5* v_i). Initially, only four cases were expected, as we anticipated the fall velocities to have a sufficient effect. As is explained below, the *0.5 x2S* had the greater impact and were added to the list of test case, resulting in a slightly higher consumption of GPU nodes. The CPU node were instead used in the development of the limited-area model, and for its use in studying the interplay between turbulence diffusivity and resolution.

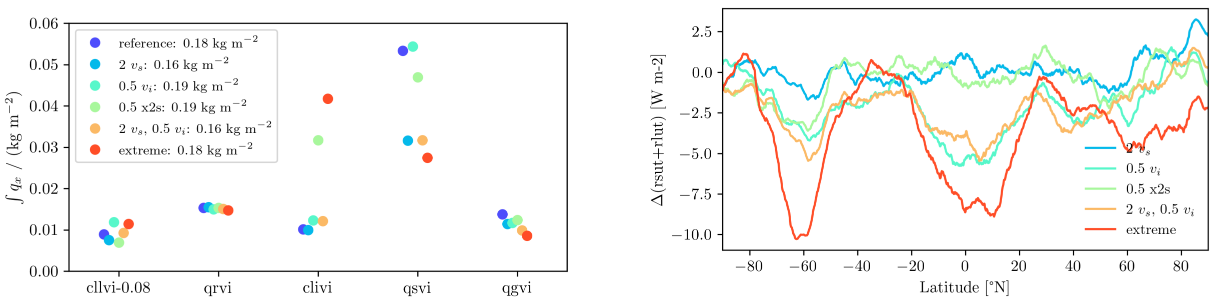

Figure 1: Left) Averaged values of column-integrated water mass for all hydrometers, for all six microphysics configuration. Right) Perturbation of the top-of-the-atmosphere radiative budget for all microphysics configurations.

As seen in Figure 1 (left), the modifications can have drastic impacts on the frozen hydrometeors populations, and provide a few surprises. First, doubling the snow fall velocity nearly halves the amount of atmospheric snow, logically, while keeping the other population virtually unchanged. More surprising, halving the ice fall velocity only marginally increase the mass of ice and snow, suggesting the microphysical conversion processes dominate over the sedimentation rate. That observation is comforted by the strong increase in ice water content for the case halving all snow production rates (incl. ice-to-snow conversion). In fact, the mass of ice triples while the amount of snow merely decrease by about 15%. Further, combining both changes to ice and snow fall speed results in a similar state as the one only doubling snow fall velocity. Finally, the combination of all changes reverses the ice/snow mass ratio from the reference case, changing from a 1-to-5 to a 4-to-3 ratio. To conclude, it is more efficient to tweak microphysical processes acting on the snow population rather than ice, as any increase/decrease of the latter will be repercussioned on the former.

The impacts of the microphysics tweaking on the top-of-the-atmospheric radiative budget is presented in Figure 1 (right). Overall, as the ice/snow mass ratio increases, we observe a decrease in outgoing longwave radiative flux across all latitudes. This is coherent with the larger effective radius the new model computes for snow, as compared to that of ice. Conversely, the solar scattering of ice particles being more important than that of snow, the microphysics rebalancing tends to increase the upwelling shortwave radiative flux. However, this increase is insufficient to counter-balance the aforementioned decrease of LW fluxes over the tropics. As a result, we observe a 7.5 W/m^2 increase in greenhouse effect at the equator. Interestingly, the mid-latitudes are impervious to the reversing of the ice/snow mass ratio, which could indicate that warm clouds have a greater influence on the radiative budget over these regions. Finally, we were surprised to see no noticeable impact on tropical precipitation rates, which encourages us to look towards tweaking the warm cloud microphysics in the future.

3.2 Limited-area modelling, resolution and turbulent diffusivity

Consumption by end of October: 60.000 CPU node hours

In the original proposal, the requested CPU node hours were split between the global and regional simulations. As the global runs could be mostly carried on the GPU instead, the majority of the CPU allocation was used to 1) to re-enable the limited-area modelling capabilities of the latest ICON "Sapphire version" and 2) use such configuration to perform regional simulations. Initially, these regional runs were supposed to expand the microphysics tests into the realm of higher-resolution simulations. However, recent internal work has shown that ICON has been suffering from a "deaggregating" bias preventing the growth and sustainability of mesoscale convective systems. Unfortunately, the aforementioned microphysics tuning had little impact on cloud organisation (not shown). This convinced us to redirect the scope of these runs towards studying the impact of the new TMIX turbulence scheme on convection.

The configuration that was chosen is presented in Figure 2. The parent domain uses a global R2B9 mesh (Δx = 5 km), with a first-level nested domain using an R2B10-like resolution (2.5 km), and two second-level nested meshes using an R2B11-like resolution (1.25 km). The two highest-resolution meshes allows us to simultaneously scrutinize regions with different cloud regimes, that is deep convection in the tropical Atlantic and shallow convection off the Namibian coast. Figure 2 shows for the latter region how an improvement of resolution has a quantifyably similar effect than increasing the stability factor b, which itself neutralizes turbulent diffusivity in convective region. This mecanism helps sustain smaller systems, which can in time organise into mesoscale systems, and is therefore crucial to resolving tropical convection.

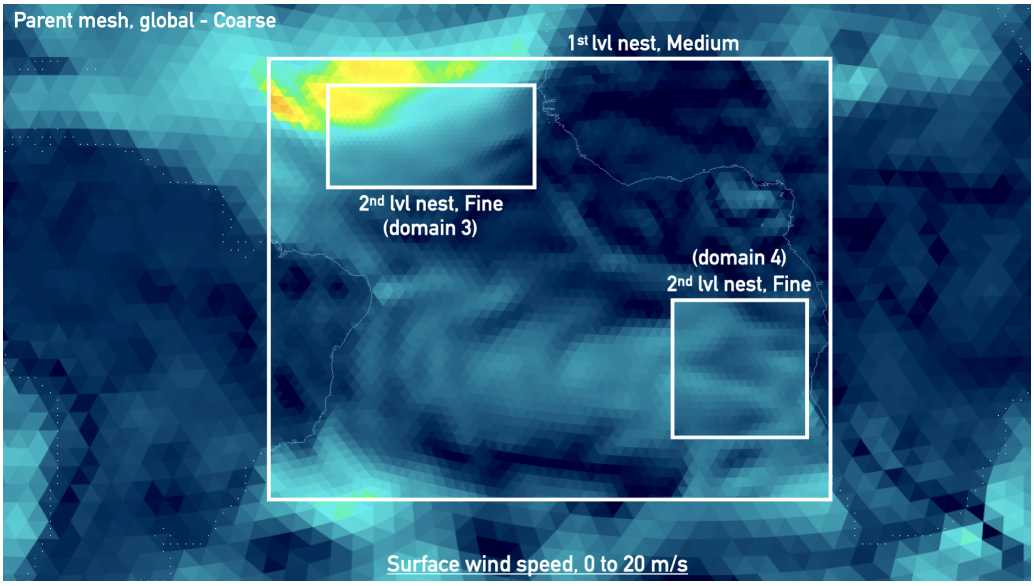

Figure 2: Overview of the global+nested configuration, with a parent domain using a global R2B9 mesh (Δx = 5 km).

Finally, a part of the allocation was spent in the development and test-run for the regional simulations in support of the ongoing field campaign ORCESTRA, though the runs themselves were priced-in another project (mh4092).

4 Turbulent mixing simulations *(DWD, MPI-M)*

Consumption by end of October: 25.000 CPU node hours

About 25.000 CPU node hours were used doing a global coupled storm-resolving simulation for the nextGEMS km scale hackathon in March 2024. This was the first application of the newly developed software package for turbulent mixing (tmx) with the 3-D Smagorinsky scheme on production scale and allowed us to quantify the impact of the scheme both computationally and scientifically. We changed the atmospheric model time step to 60 seconds for R2B8 atmosphere and R2B9 ocean horizontal resolution. With this, and parallel HEALPix Zarr output (hiopy), we performed 26 simulated months using 462 nodes. This simulation was extended to 30 simulated years using CPU resources from the nextGEMS project (bm1235). A publication of all nextGEMS simulations, including the here reported run is in preparation (Segura et al.).

Consumption by end of October: 260 CPU node hours and 65 GPU node hours

Besides the computationally intense global simulations, several smaller verification runs were made. For this purpose, a limited area simulation (LAM) covering Germany (ICON-D2 mesh) with a grid resolution of about

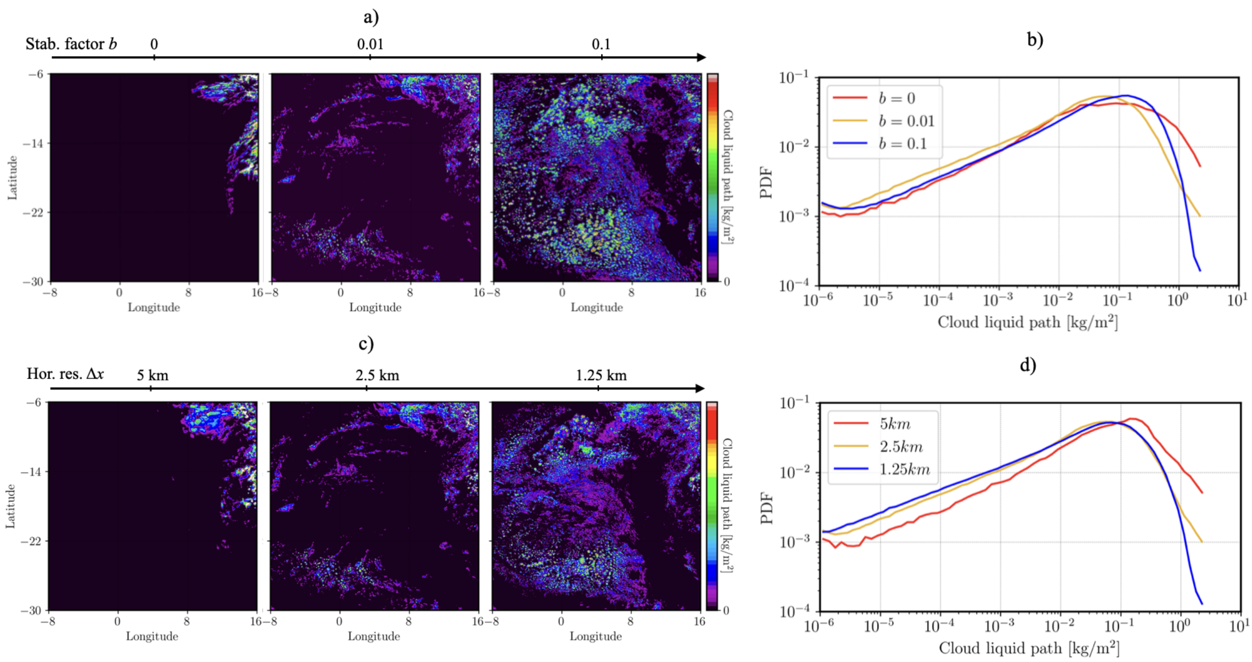

Figure 3: a) Snapshots of liquid cloud water path for increasing stability factor b and b) corresponding probability density functions. c) a) Snapshots of liquid cloud water path for increasing spatial resolution Δx and d) corresponding probability density functions.

$\Delta x = 2.1\,km$ and 90 vertical layers was set up. In contrast to generic benchmark tests and global simulations, LAM simulations allow for testing and validating newly implemented functionalities in a realistic environment on a highly resolved mesh at comparatively low computational costs. New features integrated into the turbulent mixing package (*tmx*) include:

- Implementation of an additional turbulence model based on the turbulent total energy besides the already existing Smagorinsky model
- An enhanced modularization of *tmx*
- Porting of *tmx* routines to GPUs

In total, about 260 CPU node-hours were utilized to set up the LAM simulation, verify the new features of *tmx*, and visualize the simulation results. Additionally, the development progress was rigorously tested on GPU nodes, consuming about 65 GPU node-hours.

5 Evaluate and tune mixed-phase and ice cloud microphysics *(KIT)*

Consumption by end of October: 2.000 CPU node hours

The overarching aim is to evaluate the effect of mixed-phase and ice-cloud microphysical processes on cloud glaciation, both with 1 and 2 moment cloud physics schemes. For this purpose, four aircraft-observed (ACAPEX, MC3E, DCMEX, HALO-AC³-CAO), and one satellite-observed (ORCESTRA) cloudy cases are planned to simulate. These cases were consisted of a range of cloud types, including deep convection (in ORCESTRA, MC3E, DCMEX), orographic convection (in APCAPEX) and supercooled stratiform arctic clouds (CAO).

Initially, about 200 CPU node hours were used from project bb1163 (C2Phase) to set-up MC3E and ACAPEX simulations and validation of MC3E case.

Furthermore, with both ACAPEX and MC3E simulations, the following tasks are ongoing, which consumed about 2.000 CPU node hours so far;

- validation of microphysical and macrophysical properties,
- evaluation of microphysical responses from 1 and 2 moment physics schemes,
- implementation and testing of CAMS ccn data,
- implementation and testing of time dependence of heterogeneous INP freezing,

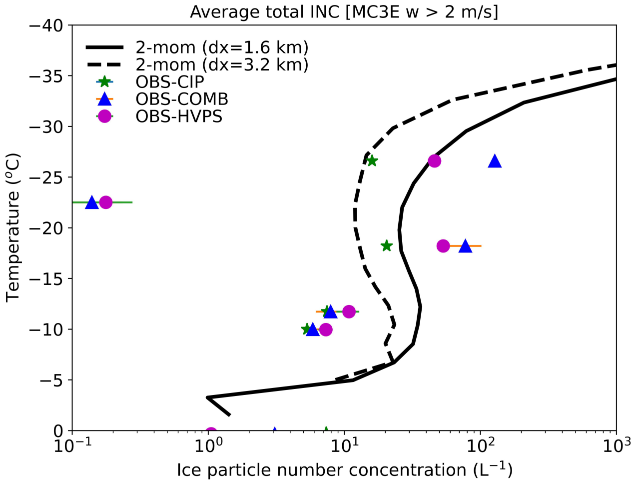

Figure 4: Comparison of the ICON predicted total ice particle number concentrations, conditionally averaged over updrafts $w > 2m/s$ with coincident aircraft observations from Cloud Imaging Probe (CIP), High Volume Precipitation Spectrometer (HVPS), and with a combined spectrum (COMB) of both CIP and HVPS, for the simulated MC3E case.

6 Latitude-belt simulations with ParFlow *(FZ Jülich)*

Consumption by end of October: 45.000 CPU node hours and 3.300 GPU node hours

The ParFlow hydrological model has been coupled to ICON-Land/JSBACH via the YAC coupler. 45.000 CPU node hours have been used for:

- Development and testing the coupling including new interface modules both in ParFlow and ICON.
- Proof-of-concept simulations with ICON running globally based on the QUBICC setup (R2B7) and a regional ParFlow setup covering a Pan-European domain with 12.5 km horizontal resolution. Outside the ParFlow coverage, ICON relies on the standard JSBACH hydrological scheme.
- A first study on the impact of the ParFlow coupling on the hydrological cycle involving reference simulations without ParFlow for comparison and simulation periods of several months.

As these first simulations have shown promising results in terms of realistic effects on soil-water variability (see Figure 5) and land-atmosphere coupling, the number of simulations has been increased in an ensemble-like way in order to demonstrate that the observations made in the model output are robust. For that reason, considerably more CPU node hours than projected have been consumed. Apart from that, the ICON simulations were run globally instead of regionally as planned in the first place.
Lastly, the GPU resources were started to be used computing the ”ICON-only” reference simulations, each covering several months in the course of a few consecutive years. The remaining GPU node hours will be consumed until end of year as a working GPU setup for the coupled simulations had to be developed first.
A paper is in preparation (Weinkaemmerer et al.).

7 Data handling workflow development *(DKRZ)*

Consumption by end of October: 400 CPU node hours and 60 GPU node hours

Within the framework of the project ”WarmWorld Easier”, our principal focus lies in developing refined and efficient data handling workflows to manage the extensive volumes of data integral to our operations. The compute resources have been employed strategically and purposefully, with the primary intent being the development, refinement, and rigorous testing of these workflows and supplementary tools.

Moreover, essential to our project progression has been the development of in-situ workflows, which necessitate the availability of a fully operational setup for rigorous testing. The necessity of running a full-scale operational setup for in-situ workflow testing cannot be overstressed, as it enables an accurate simulation and examination of how the workflows cope in a live environment.

This rigorous testing approach has facilitated the progress from nascent conceptual stages to developing first prototypes, providing an experimental sandbox for attempting and evaluating various innovative ideas. Through

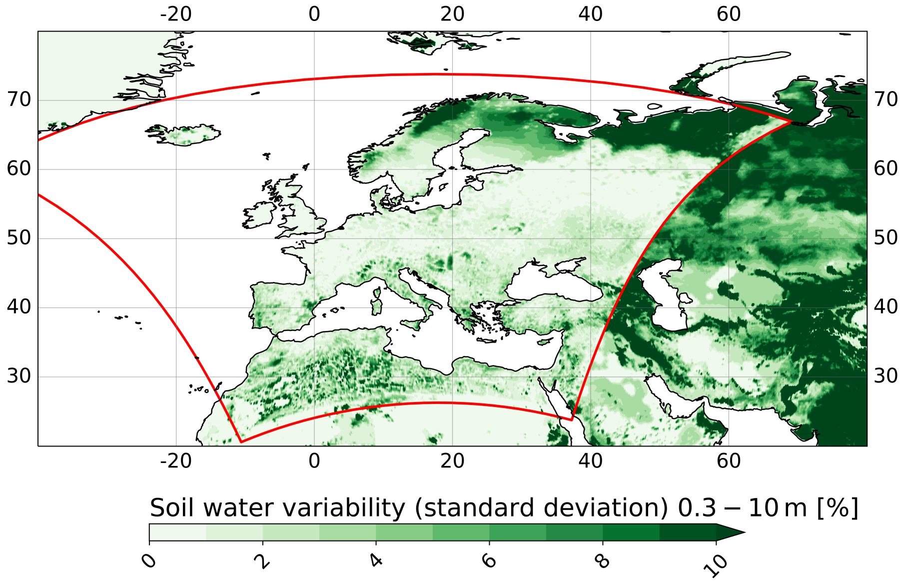

Figure 5: Inside the ParFlow coverage (red frame), the temporal soil-water variability in the deeper soil layers (based on weekly averages from April to November 2004) is effectively reduced as lateral groundwater flow is enabled.

this iterative process, not only have we been able to test the stamina and integrity of our tools and workflows, but also constantly improve them and extract actionable insights that feed back positively into our development process.

The overview of the data and metadata workflow from simulation to the end user is illustrated in Figure 6.

The computing resources granted for 2024 as part of the Warm World Easier project were also utilized in the following activities:

- Development of STAC catalogs for the Km scale simulations.
- Deployment of the catalog on various backends, storage and retrieval of the metadata to address the user queries to facilitate the discoverability of Km scale ESM simulations.
- Development of software to reorganize the data as appropriate to the storage tier, for example to optimally fit into the tape archives, to keep in view the size, data transfer rate constraints etc. imposed by the Hierarchical storage management (HSM) system, see Figure 7.

8 ICON performance benchmarks *(DKRZ)*

Consumption by end of October: 3.400 CPU node hours and 1.300 GPU node hours

As part of "WarmWorld Faster" project module, compute resources have been used for benchmarks and scaling runs of global high-resolution ICON simulation with experiment set-ups consistent with the current flagship science projects, nextGEMS and DestinE. The performance and the scalabilty of the coupled experiment (R2B8-L90 for the atmosphere and R2B9-L72 ocean) as well as it corresponding atmosphere-land experiment (R2B8-L90) has been evaluated with respect to the total run but also with respect to different part of the code. To gain better performance, the runs has been setup as heterogeneous jobs, where the atmosphere runs on the *gpu* partition and the ocean on the *cluster* partition. Within the scope of the WarmWorld project but also in cooperation with Destination Earth, these scaling runs have been performed also on the super-computer JUWELS-BOOSTER and LUMI.

Within the reporting period, these benchmarks showed several bugs, e.g. in the coupling or the aerosols when the atmosphere is runs on GPUs. This helped to improve the ICON atmosphere performance at GPUs significantly.

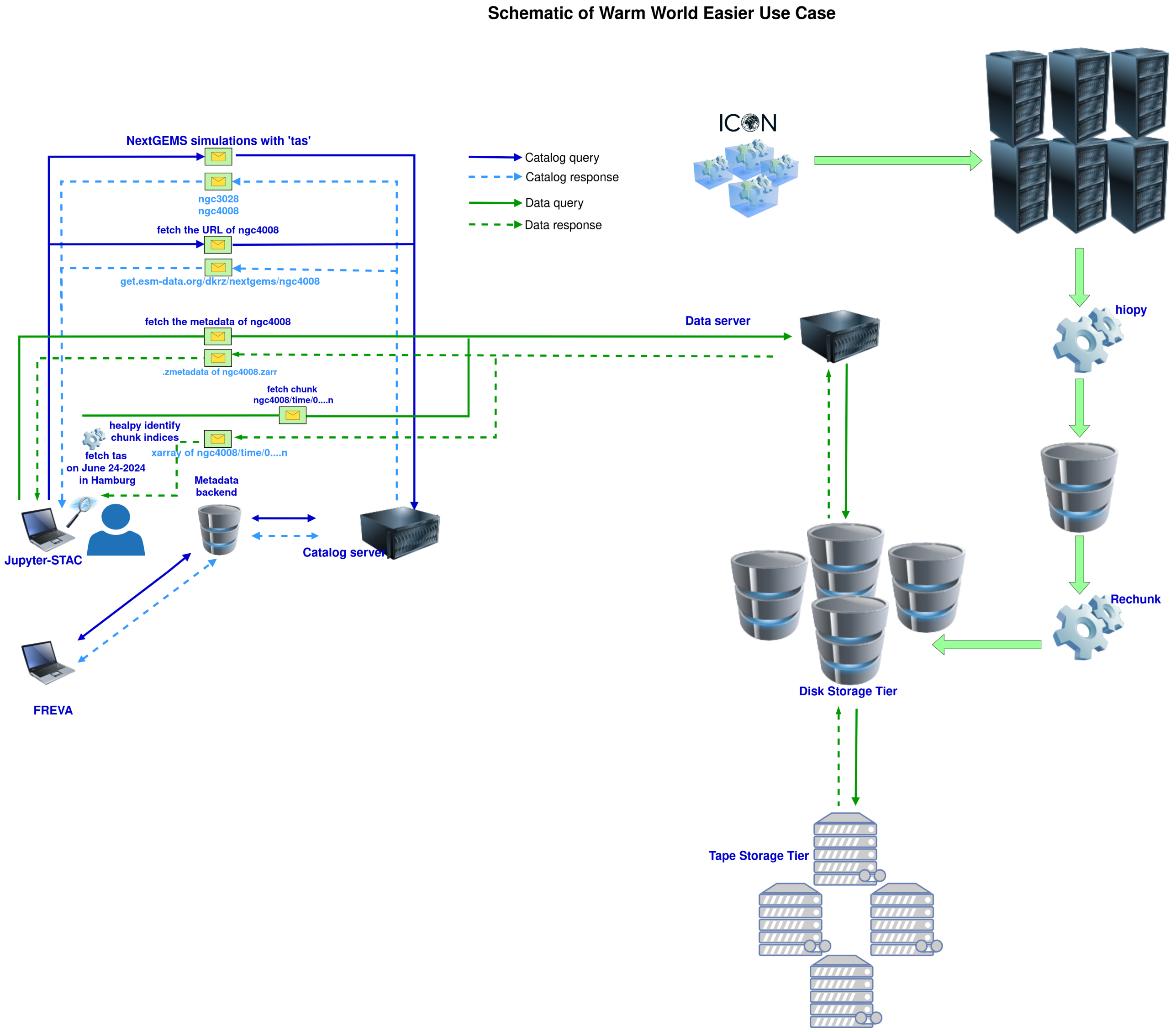

Figure 6: Schematic of the HPC resources for simulation, storage and discovery of earth system model output.

9 Archive 1.25km coupled baseline run of 2023 *(MPIM)*

As analysis and writing publication of the "ICON SR-ESM baseline 1.25 km coupled" run (see report of 2023) is still in progress the runs has not yet been archived. But the data of 220 TiB was reduced by 40 TiB by deleting non necessary files.

10 ICON model development on GPUs *(DKRZ)*

Consumption by end of October: 200 CPU node hours and 300 GPU node hours

As part of "WarmWorld Faster" WP12 D2, we used the compute hours (170 CPU and 220 GPU) to evaluate several heterogeneous programming models which could be an alternative to Fortran + OpenACC (this includes gitlab continuous integration support). The prototype repository has five C++ backends: Serial, Kokkos, OpenMP, OpenACC and SYCL. Preliminary measurements of the microphysics standalone show that the Kokkos implementation could be 1.5-3× faster than Fortran+OpenACC on recent NVIDIA and AMD GPU, while the same code compiled for `x86-64` provides a multi-threaded implementation which is 15-20× faster than the (sequential) Fortran version.
Final performance evaluation of full ICON simulations that use the C++ microphysics has just started. This is expected to be the most compute intensive part. The 4500 GPU node hours estimated initially included this step, but the implementation was delayed and the developers reached this point only in September 2024; therefore only a fraction of the allocation was actually used.

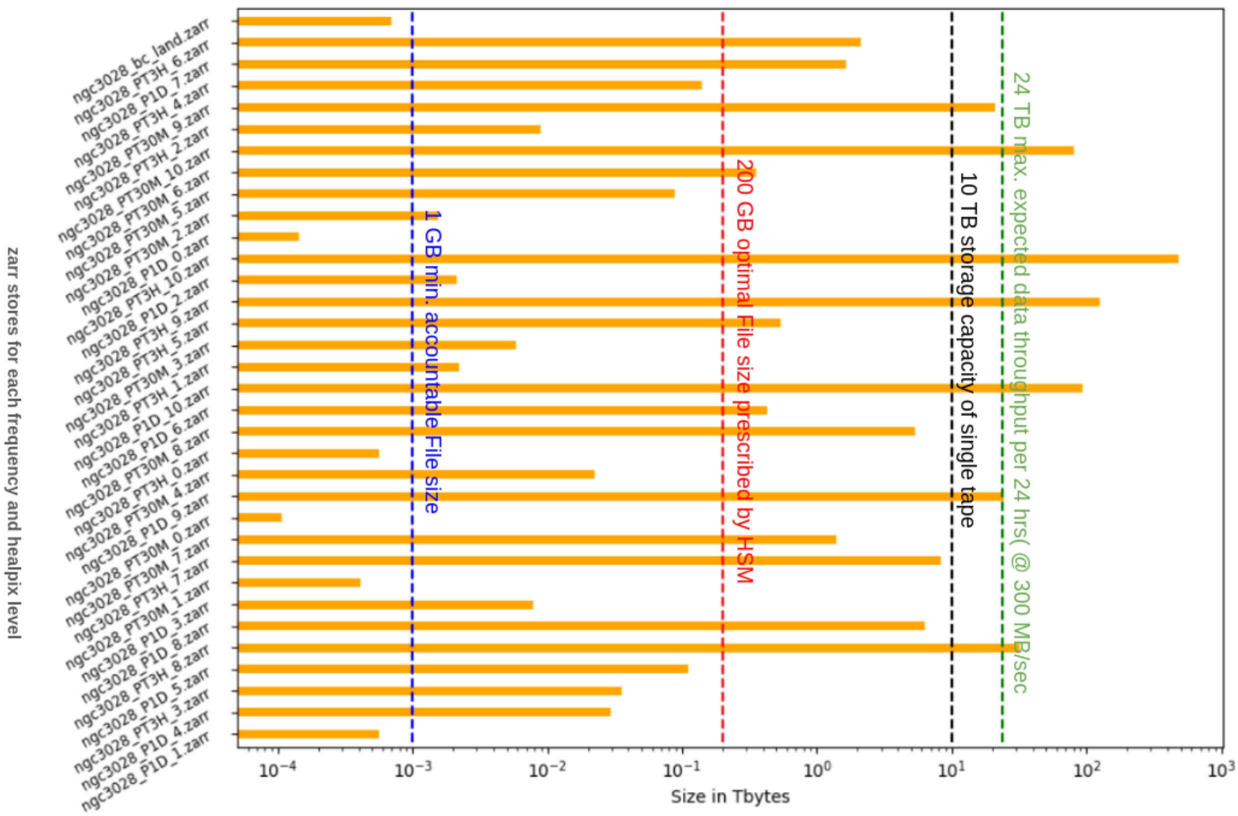

Figure 7: Typical near Km scale simulation output plotted against the constraints of the Hierarchical Storage Management system at DKRZ.

11 IFS-FESOM configurations and simulations *(AWI)*

Along the course of Warmworld project, it is intended to develop FESOM ocean model to be able to perform km-scale simulations (initially with IFS). Along the way, simulations with new test configurations are used to ghost the simulations performed with ICON model to provide another realization of climate for assesment of uncertainiy. In the first phase the project, the simulations with ICON are mostly with Atmosphere-only configurations to be able to compare with observational campaigns such as ORCHESTRA. To be able to compare a coupled simulation of FESOM, that has more degrees of freedom, with an atmosphere only simulations and observations, it is intended to test and develop spectral nudging as a framework to bring large scales of the atmospheric model to focus on a time-bound state.

11.1 Nudging methodology and sensitivity tests

Consumption by end of October: 11.400 CPU node hours

Nudging methodology was initially applied with AWICM3 (FESOM with openIFS) at a relatively coarse resolution, various sensitivity tests were performed to find the nudging timescale and wave numbers for a high resolution (≈10km atmosphere and ≈5km ocean scale) IFS-FESOM model configuration. The understanding from this was used to perform a extended simulation on LUMI as heavy-nudging datasets were readily available as part of DestinE project. The manuscript from the results of these experiments is currently underway. Initially allocated resources to do simulations over MOSAIC campaign (2019-2020) were exchanged with DestinE and NextGEMS projects, as described below.

Consumption by end of October: 13.700 CPU node hours

For the nextGEMS project, 13.700 CPU node hours were used to run three years of an IFS-FESOM historical at TOC1279 resolution for IFS and NG5 for FESOM. This simulation was run for 30 years in total (1990-2020) and compliments the IFS-FESOM scenario simulation performed for nextGEMS. The IFS-FESOM historical simulation manages to capture global mean temperature evolution over this 30-year period well (see Figure 8). This increases the credibility of the IFS-FESOM nextGEMS scenario. Lessons learnt from finding the best IFS-FESOM configuration for nextGEMS also feed back to WarmWorld and Destination Earth model development.

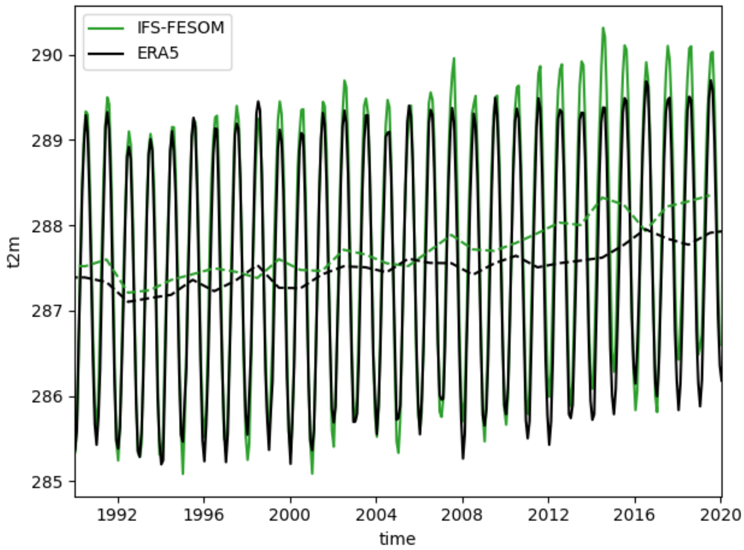

Figure 8: Time series of monthly (solid) and yearly (dashed) global mean 2 meter temperature (in Kelvin) in the IFS-FESOM nextGEMS historical simulation with a TOC1279-NG5 resolution.

11.2 Spin-up of IFS-FESOM model

Consumption by end of October: 45.000 CPU node hours

To use a coupled model with nudging it is essential to start with a well-balanced spin-up state of atmosphere and ocean. A steady-control state of the coupled model is essential and would also benefit various projects utilising the same km-scale IFS-FESOM simulation configuration. Initially, a coupled spin-up simulation was proposed to achieve this steady state. The model exhibited strong drift immediately after the start of the coupled spin-up, necessitating a longer simulation period of approximately 50 years to reach a stable control state (see Figure 9). Of this extended spin-up, about 10 years (with an approximately 45.000 CPU node hours per year) were conducted using resources from the WarmWorld project. The climate state at the conclusion of this spin-up simulation provides appropriate initial conditions for subsequent nudged simulations, ensuring a more accurate representation of climate conditions.

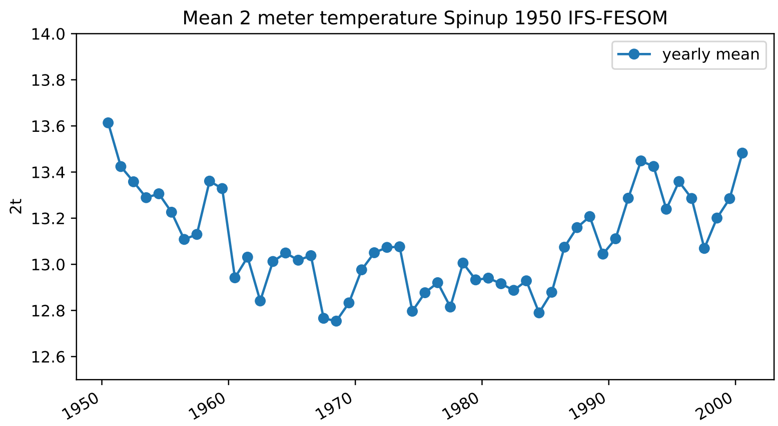

Figure 9: Time series of global yearly mean 2 meter temperature (in degree Celsius) in the coupled spin-up simulation of the IFS-FESOM at 9km atmosphere (TOC1279) and 5km Ocean (NG5) resolution under 1950 CMIP6 radiative forcing.

11.3 New IFS-FESOM coupled configurations

Consumption by end of October: 22.200k CPU node hours

The initial coupling of IFS and FESOM configurations as reported above were developed by ECMWF. A hypothesis for unforseen spin-up time reported above is due to non conservative remapping from a very unstructured ocean model and atmospheric model using different land-sea masks. To test this hypothesis and to develop flexible IFS-FESOM configurations the compute resources were used. The simulations are relatively new and are yet to be analysed. It is anticipated to lead a development that is more consistent and conserved coupled configuration.

11.4 ICON-FESOM development

Consumption by end of October: 1.800 CPU node hours

Using the compute resources, a 2 year-old version of ICON-FESOM port done by Nils Dreier from DKRZ using YAC, was tested. An effort to upgrade the FESOM model to a recent, majorly refactored version is underway.

11.5 FESOM-GPU development for efficient km-scale simulations

While it was initially intended to further develop GPU-port of FESOM2 model for tracer advection, it was difficult to timely obtain resources (and insufficient) on Levante to be able to perform km-scale simulations. Most GPU-development hence focused on LUMI HPC. The tests at lower resolutions were performed on Levante by the HPC support group of AWI with resources accounted on project ab0995. For the next year, more attention will be paid for accounting of the GPU-resources.

References

Hans Segura, Xabier Pedruzo-Bagazgoitia, Philipp Weiss, Sebastian K. Müller, Thomas Rackow, Junhong Lee, Edgar Dolores-Tesillos, Imme Benedict, Razvan Aguridan, Suvarchal K. Cheedela, Ioan Hadade, Rene Redler, Karl-Hermann Wieners, Florian Ziemen, and the nextGEMS community. NextGEMS: entering the era of kilometre-scale earth system. *In preparation.*

Jan Weinkaemmerer, Reiner Schnur, Klaus Görgen, and Stefan Kollet. Coupling the parflow hydrological model to ICON-Land. *In preparation.*