Project: **1372**
Project title: **ExaOcean**
Principal investigator: **Maximilian Witte**
Report period: **2023-05-01 to 2024-04-30**
Maximum of 2 pages including figures. 9 pt minimum font size.

Allocations 7/2023 to 6/2024

Resource	Granted	Utilization	Remaining	Files
Levante CPU nodes (Node hours)	3000	2259	741	
Levante GPU nodes (Node hours)	10000	8196	1804	
Levante storage (TiB)	100	55	45	
Archive project (TiB)	150		150	
Swift Object Storage (GiB)	10240		10240	

Figure 1: Allocated and used ressources.

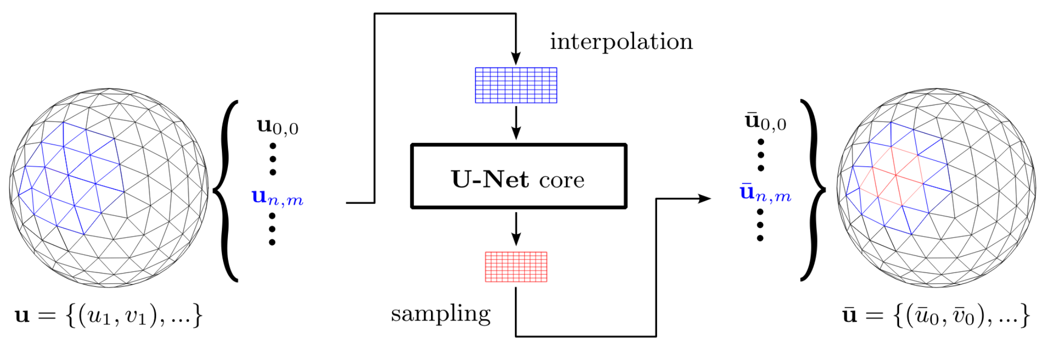

Figure 2: Taken from [1] Implemented approach for the combination of super-resolution with ICON. The ML model is locally inferred so that its global application can be parallelized. The input data on the ICON grid is interpolated to enable efficient operations on a regular grid (e.g. convolutions). A sampling module was developed to map the data back onto the ICON grid. [1]

Justification of expired ressources (Figure 1)

An ML model has been successfully developed that demonstrates scalability by dynamically adapting the size of the spatial domain to the specific problem (Figure 2). Although the allocated resources were not fully utilised, several factors contribute to this result. Firstly, during the development phase, considerable effort was invested in establishing efficient interfaces between grids and adapting specific modules within the U-Net core. Second, the use of a lightweight local approach for both training and inference significantly reduces computational requirements. The locality of the model not only helps to conserve resources, but also enhances its ability to generalise effectively. In addition, the choice of a 2D Galewsky test case for method validation, characterised by fewer variables and dimensions compared to a full 3D model, contributes to the efficient use of resources. The training data were obtained from MPI-M resources.

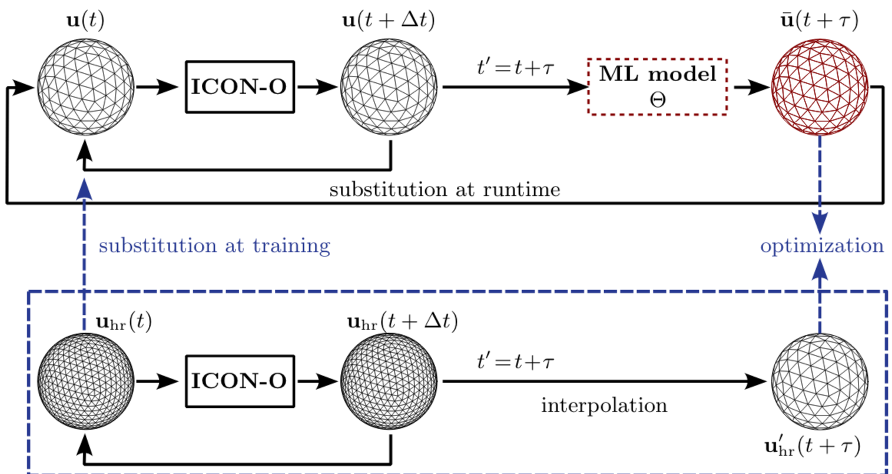

Figure 3: Taken from [1]. Hybrid approach combining numerical simulation using the ICON-O model with machine learning based super-resolution. The data flow during runtime is indicated by the solid arrows (top panel), while the dashed blue lines indicate the data flow during training. Δt is the time step used in the numerical simulation (ICON-O). At t = 0 and τ ≫ Δt, we use a distance-weighted interpolation method (CDO) to map the high-resolution ground truth of the simulation onto the low-resolution grid for the input and the ground truth of the ML model Θ training [1].

Results of ExaOcean

The approach used is shown in Figure 2.
The ML model is trained to learn the difference between the low- and high-resolution simulations at each τ. For the chosen test-case (Galewsky, shallow water), τ was chosen as $\tau = 12h$. At runtime, the ML model corrects the low-resolution data. See the preprint for details. The results have been submitted to the SIAM Journal of Scientific Computing (SISC). For results refer to the published preprint [1].

[1] Witte et al. Dynamic Deep Learning Based Super-Resolution For The Shallow Water Equations (2024). arXiv:2404.06400