Final Report for Project 1372

ExaOcean

Principal investigator: Maximilian Witte

Report period: July 1, 2023 – June 30, 2026

Project title:	ExaOcean
Grant identifier:	16ME0679K / bk1372
Institutional context:	Deutsches Klimarechenzentrum (DKRZ)
Project runtime:	October 2022 – September 2025

Abstract

ExaOcean pursued a combined numerical and machine-learning approach to improve ICON-O simulations on modern high-performance computing systems. On the machine-learning side, the primary research focus was the use of correction models to recover fine-scale information during runtime, allowing coarse ICON-O simulations to approach the accuracy of finer-grid references at lower cost. Within the project this dynamic super-resolution strategy was established for a shallow-water benchmark, extended to three-dimensional baroclinic turbulence with a depth-aware neural network, and prepared for HPC workflows through parallel inference strategies. In parallel, the Field-Space Transformer was developed as a structure-preserving multiscale architecture motivated by the long-term goal of learning on variable-resolution Earth-system data such as the submesoscale telescope configuration. Although deployment on the targeted submesoscale telescope workflow was not completed within ExaOcean, the resulting architecture has clear relevance beyond the project itself.

1 Overview

The overall scientific goal of ExaOcean was to improve ICON-O simulations through a combination of more efficient numerical algorithms and machine-learning-based correction. A central project objective was to reduce the effective cost of high-resolution ocean simulation by learning corrections that can be injected during runtime, so that a coarser model trajectory stays close to a finer-resolution reference. In practical terms, the project started from conventional convolutional architectures, in particular U-Net-style models, because they provide a robust and computationally efficient baseline for super-resolution and hybrid correction tasks.

Alongside this ML line, ExaOcean also developed parallel spectral deferred correction (PSDC) time-stepping methods to improve small-scale parallelism in time on multicore CPU systems, with algorithmic advances and shared-memory performance studies carried out on Jülich resources [3, 4].

This report therefore follows the scientific progression of the project:

1. First, ExaOcean examined whether a conventional U-Net, applied periodically during runtime, can correct coarse ICON-O simulations sufficiently well to recover the evolution of a finer-grid solution.

2. Second, it assessed whether this idea can be transferred from a two-dimensional benchmark to a three-dimensional baroclinic turbulence setting that is substantially closer to realistic ocean dynamics.
3. Third, it addressed whether the resulting ML components can be embedded efficiently enough into HPC workflows to support hybrid simulation.
4. Finally, motivated by the longer-term target of working with submesoscale telescope data on variable-resolution grids, the project also explored more structure-preserving architectures beyond the original U-Net line.

2 U-Net based hybrid correction for ICON-O

2.1 Dynamic super-resolution for shallow-water dynamics

The first major result of the project was the demonstration that dynamic super-resolution works for ICON-O in a controlled shallow-water benchmark. In the published study by Witte et al. [1], a U-Net-type neural network was trained to predict the correction between low- and high-resolution ICON-O solutions and was then inserted into the time-stepping loop as a periodic runtime correction. The benchmark used the Galewsky barotropic instability test case, which is demanding because it requires the model to maintain balance while also capturing the transition to turbulence.

Scientifically, this result established the feasibility of the hybrid approach. A 20 km simulation corrected by the neural network achieved discretization errors comparable to a 10 km simulation, while retaining the balanced large-scale flow and the onset of instability seen in the higher-resolution reference. After eight simulated days, the corrected run had an L_2 error similar to the finer-grid simulation. Mass conservation was retained, although some spurious kinetic-energy generation remained. These results demonstrated that ML corrections can be integrated into a prognostic ocean model in a dynamically meaningful way rather than being limited to offline post-processing.

2.2 Extension to three-dimensional baroclinic instability

The second major result was the extension of the DSR idea to a much more demanding three-dimensional baroclinic channel configuration. This work is documented in the manuscript *A CNN-Ocean Model Hybrid for Improving the Representation of Baroclinic Instability in Coarse-Resolution Simulations*, led by MPI-M collaborators and based on the ExaOcean hybrid workflow [2]. Here, the coarse ICON-O simulation at 5 km resolution was corrected using a depth-aware U-Net with Feature-wise Linear Modulation and sinusoidal depth embeddings, allowing the network to learn corrections for three-dimensional temperature and velocity structure without the full cost of 3D convolutions.

This step was scientifically important because baroclinic turbulence contains the multiscale eddies, jets, vertical structure, and short predictability horizons that make hybrid correction a stringent test of both dynamical fidelity and short-term predictive skill. The hybrid simulation reduced the L_2 error growth rate to a level comparable to an unassimilated 2 km simulation, even though the underlying numerical model ran at 5 km resolution. Spatial decorrelation was delayed relative to the free-running 5 km run, and the corrected simulation improved the representation of key physical diagnostics including the mean temperature structure, zonal jets, eddy kinetic energy, and vertical buoyancy flux. The work therefore showed that learned corrections can bridge the gap between coarse and intermediate resolution not only for stylized 2D dynamics but also for fully three-dimensional ocean turbulence.

The study also clarified the limitations of the present approach. Temperature correction at the smallest scales remained harder than velocity correction, likely because tracer errors depend strongly on unresolved vertical advection. The vertical buoyancy flux improved but remained below the 2 km reference. These limitations indicate that future hybrid architectures need stronger vertical coupling and more explicit physical constraints.

2.3 Workflow and HPC relevance

The ML work in ExaOcean was not limited to model design. A hybrid simulation is only useful if the inference step can be embedded efficiently into the larger HPC workflow. To support this, the project implemented a spatially parallel inference strategy for the U-Net-based correction model on CPUs. The field can be partitioned in space and processed without loss of accuracy, which is important for running the ML component within the same parallel environment as the numerical model.

Within the broader ExaOcean effort, these ML advances complemented the development of more efficient time integration through PSDC methods [3, 4] and the porting effort toward heterogeneous HPC systems including JUPITER. The project did not reach the final fully integrated production stage in which all acceleration routes were combined for the targeted submesoscale application. Still, it established several of the required components for that integration, including a published U-Net-based online correction method, a 3D baroclinic extension, CPU-parallel inference, and initial porting efforts toward JUPITER, while leaving full end-to-end coupling of these components to future work.

3 Field-Space Transformer as a parallel development

The Field-Space Transformer did not emerge as the primary architecture used in implemented ExaOcean hybrid ICON-O workflow. Instead, it emerged from the project's longer-term scientific target of learning from submesoscale telescope (SMT) data and related variable-resolution ocean configurations. The SMT experiment used ICON-O on a grid with strongly varying resolution, reaching about 600 m in the North Atlantic and about 11 km near Australia [6]. This kind of multiscale, non-uniform setting motivated the search for an architecture that is more naturally aligned with variable-resolution and structure-preserving field representations than a standard latent-space transformer.

This led to the development of *Field-Space Attention*, documented in the project preprint by Witte et al. [5]. The central idea is to replace purely latent token processing by a structure-preserving multiscale representation that remains in physical field space throughout the network. On global temperature super-resolution experiments on HEALPix grids, the resulting Field-Space Transformer trained more stably than conventional vision-transformer and transformer-U-Net baselines and achieved better accuracy with substantially fewer parameters. In the reported scaling experiment, a 15.5M-parameter Field-Space Transformer reached a lower RMSE than much larger baseline models while preserving interpretable multiscale intermediate states.

The project did not reach the stage where this architecture could be deployed on the intended SMT-style ICON-O workflow. It should therefore be interpreted as an architecture motivated by ExaOcean's longer-term application target rather than as a component of the implemented hybrid ICON-O workflow. Field-Space Transformer should therefore be understood as a parallel scientific development that grew out of ExaOcean's final target but was not itself an operational ExaOcean deliverable inside ICON-O. Even so, its significance is broader than the original project scope. It provides a promising architecture for future work on multiresolution Earth-system data, including data assimilation, prediction, generative modeling, and hybrid correction on geometrically complex grids.

4 Assessment against the original objectives

Objective	Status	Outcome
Develop online ML correction for ICON-O	Achieved	Dynamic super-resolution was demonstrated in ICON-O and published for the shallow-water benchmark.
Transfer DSR to more realistic ocean turbulence	Achieved	A depth-aware hybrid setup for three-dimensional baroclinic instability was developed and shown to recover key dynamics of finer-resolution runs.
Optimize ML execution for HPC workflows	Partially achieved	CPU-parallel inference was implemented and JUPITER porting began, but full heterogeneous deployment was not completed during the project.
Combine algorithmic acceleration routes into one production workflow	Partially achieved	Individual speedups were demonstrated, but the full integrated throughput target was not realized within the funded period.
Enable online learning during simulation	Not achieved	This part of the work plan was deferred because effort was prioritized toward robust offline training, hybrid coupling, and inference optimization.
Develop next-generation structure-preserving ML models	Achieved in parallel	Field-Space Attention was developed as a new multiscale transformer concept motivated by the project's longer-term multiresolution target, but it was not integrated into the final ICON-O workflow.

Taken together, the project met several of its core scientific objectives, while only partially reaching the originally envisioned integrated endpoint. The core scientific line was clearly established: ExaOcean showed that conventional U-Net-based runtime correction can improve coarse ICON-O simulations, that the idea transfers from idealized 2D flow to 3D baroclinic turbulence, and that the resulting workflow can be aligned with HPC constraints. In parallel, the project generated a more forward-looking architectural development in the Field-Space Transformer, which extends beyond the immediate ExaOcean scope.

5 Publications, impact, and outlook

The project produced one completed DSR publication directly tied to ExaOcean, one advanced manuscript on three-dimensional baroclinic instability, and one major methodological preprint on Field-Space Attention. The resulting methods are not limited to ICON-O. The U-Net-based correction line is relevant wherever coarse PDE-based solvers can benefit from learned multiscale correction. The Field-Space Transformer line is potentially even broader, because it addresses a general representation problem for continuous multiscale fields and is not tied to one ocean-model workflow.

Several next steps follow directly from these results. First, the 3D hybrid U-Net framework should be pushed from benchmark channels toward more realistic ocean configurations. Second, the inference path should be optimized further for heterogeneous systems, including GPU-resident workflows where possible. Third, the structure-preserving ideas behind Field-Space Attention should be tested directly on variable-resolution Earth-system data and then evaluated for hybrid correction and data assimilation settings. These directions define a concrete continuation path toward more accurate and more efficient Earth-system simulation.

References

[1] M. Witte, F. R. Lapolli, P. Freese, S. Götschel, D. Ruprecht, P. Korn, and C. Kadow. Dynamic deep learning based super-resolution for the shallow water equations. *Machine Learning: Science and Technology*, 6(1):015060, 2025. https://doi.org/10.1088/2632-2153/ada19f

[2] F. R. Lapolli, M. Witte, P. Korn, and C. Kadow. A CNN-Ocean Model Hybrid for Improving the Representation of baroclinic instability in coarse-resolution simulations. Manuscript in preparation, based on ExaOcean project results, 2026.

[3] G. Čaklović, T. Lunet, S. Götschel, and D. Ruprecht. Improving efficiency of parallel-across-the-method spectral deferred corrections. *SIAM Journal on Scientific Computing*, 47(1):A430–A453, 2025. https://doi.org/10.1137/24M1649800

[4] P. Freese, S. Götschel, T. Lunet, D. Ruprecht, and M. Schreiber. Parallel performance of shared memory parallel spectral deferred corrections. *The International Journal of High Performance Computing Applications*, 2025. In press. http://arxiv.org/abs/2403.20135v1

[5] M. Witte, J. Meuer, E. Plésiat, J. Baehr, and C. Kadow. Field-Space Attention for structure-preserving Earth system transformers. arXiv:2512.20350, 2025. Under review.

[6] M. Epke, L. Linardakis, P. Korn, and N. Brüggemann. Overturning of mixed layer eddies in a submesoscale resolving simulation of the North Atlantic. *Journal of Physical Oceanography*, 56(6), 2026. https://doi.org/10.1175/JPO-D-25-0015.1