# SPECIAL PROJECT PROGRESS REPORT

All the following mandatory information needs to be provided. The length should *reflect the complexity and duration* of the project.

| | |
|---|---|
| **Reporting year** | 2024 |
| **Project Title:** | Speeding up ocean spin-up using stochastic parametrisations |
| **Computer Project Account:** | spgbchri |
| **Principal Investigator(s):** | Dr Hannah Christensen<br><br>Dr Kristian Strommen<br>Dr Robert (Bobby) Antonio |
| **Affiliation:** | University of Oxford |
| **Name of ECMWF scientist(s) collaborating to the project** (if applicable) | Chris Roberts………………………………………..……<br>……………………………………………………….…… |
| **Start date of the project:** | September 2023 |
| **Expected end date:** | March 2026 |

## Computer resources allocated/used for the current year and the previous one
(if applicable)
Please answer for all project resources

| | | Previous year | | Current year | |
|---|---|---|---|---|---|
| | | Allocated | Used | Allocated | Used |
| **High Performance Computing Facility** | (units) | 3500000 | 3500000 | 50000000 | 7161 |
| **Data storage capacity** | (Gbytes) | | 85 | 15500 | 85 |

## Summary of project objectives (10 lines max)

The primary aim of this project is to use recent advances in machine learning (ML) to speed up the ocean spin-up in a coupled earth system model, to improve the quality of climate simulations as part of the EERIE (European Eddie-Rich ESMs) project. Our approach to do this is by coupling a machine learning emulator of the atmosphere (Graphcast) to a standard ocean model (NEMO), which will significantly reduce the compute time required per year of spin up. Once we have successfully created a scheme to couple these models and stabilised the machine learning model, we will produce one or more spun-up ocean states, and assess the quality of these states by running a conventional climate model (EC-Earth) using the spun-up state and assessing the drift and biases. We also hope to use this novel setup to explore scientific questions about how the atmospheric dynamics affect the ocean.

## Summary of problems encountered (10 lines max)

The machine learning emulator we are using becomes unstable after around 20 days of simulation, and so we have spent some time exploring ways to make it operate up to climate timescales. Coupling the machine learning model to the ocean model is also non-trivial, and we are at the early stages of having something working for this. Some time was also spent setting up and performing initial runs of EC-Earth; whilst the model appears to compile and run successfully, there seem to be some problems occurring after running for a few years, which we have not yet resolved. Our eventual use of compute resources will be unconventional, requiring both a GPU and substantial numbers of CPUS, so we anticipate this will be a problem to be addressed with the help of expertise from the HPC team.

## Summary of plans for the continuation of the project (10 lines max)

Once we have arrived at a robust scheme to stabilise the ML model and couple it to NEMO, we will run both models together for many decades to produce one or more spun-up ocean states. We will then run EC-Earth initialised with these spun-up states to assess the quality of the spin up. If time permits, we also hope to use the coupled setup to explore scientific questions about the impact of the atmosphere on ocean variability.

## List of publications/reports from the project with complete references

N/A

## Summary of results (September 2023 – June 2024)

Over the last 8 months we have experimented with running a Machine Learning emulator of the atmosphere (Graphcast) out beyond the time range at which it was trained, to investigate its suitability to be used to spin-up an ocean model for climate simulations. Our results indicate the model becomes unstable at around 20 days (Figure 1). However, before the model becomes unstable, evidence for a seasonal cycle is visible (Figure 1), important for using GraphCast on longer timescales than the ten days for which it was trained.

We have then been assessing methods to stabilise Graphcast for climate timescales, and assessing how well this stabilised model captures the long-term dynamics of the atmosphere. We find that this stability is particularly sensitive to near-surface temperature fields, indicating the possibility that coupling to an ocean may have a stabilising effect. However, we anticipate this will not be sufficient to fully stabilise the model. We have therefore developed a re-initialisation approach. Every T days, we encode the state of GraphCast into a reduced order latent space. In this latent space, we compare GraphCast to an ERA5 database, to find the closest historical match. ERA5 fields from this day are used to reinitialise GraphCast. In this way, we are guaranteed that predictions will never go outside of the range of ERA5. Figure 2 shows this approach in practice. Here we have matched using a PCA approach on 2m Temperature only, every T = 10 days.

We are beginning to investigate methods to couple it to a conventional ocean model (NEMO). Once we are happy with these components, we will start running coupled simulations and explore whether or not the resultant ocean state is a good initial state for EC-Earth.

We have successfully compiled and performed several initial runs of EC-Earth (with the help of Uwe Fladrich), and set up the NEMO model. We have run into some undiagnosed failures occurring while running EC-Earth, which may be because this particular version is still at an early stage of development.

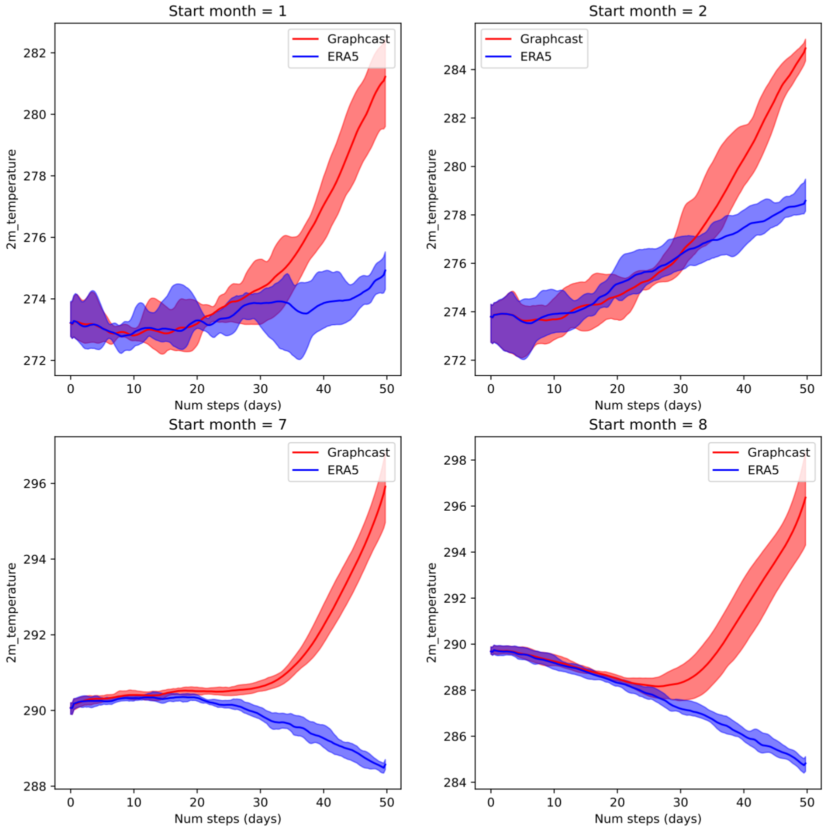


Figure 1: 50-day GraphCast simulations (red) initialised from dates in a) January b) February c) July d) August, for 5 different years. These are compared to the variability observed in the same five years of ERA5 data. All panels show 2m Temperature averaged over the North Hemisphere.

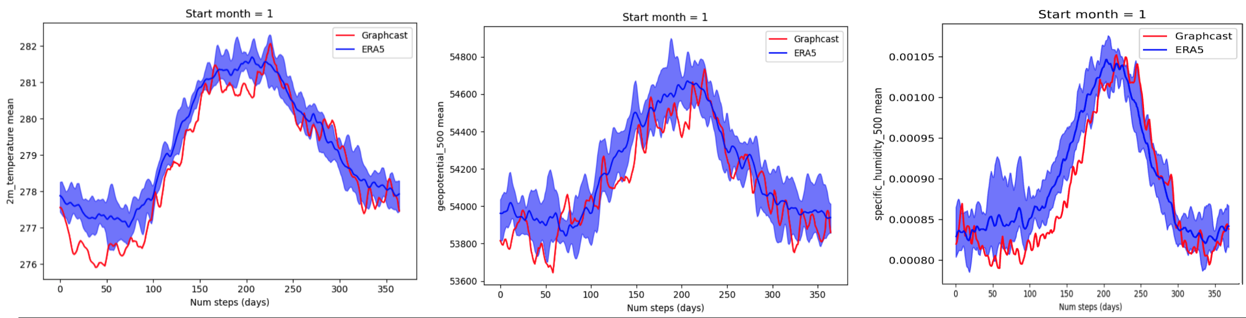


Figure 2: One year-long GraphCast simulation (red) starting in January, using the reinitialization approach described. Global mean values of a) T2m b) Z500 and c) q500 are shown. Each variable shows a realistic annual cycle when compared to the previous 5 years of ERA5 (blue).