EMI R&D PROJECT FINAL REPORT

All the following mandatory information needs to be provided.

Project Title:	Machine learning techniques for ocean spin-up acceleration
Computer Project Account:	spitsozz
Start Year - End Year :	2025 - 2026
Principal Investigator(s)	Alessandro Sozza
Affiliation/Address:	Institute of Atmospheric Sciences and Climate, National Research Council of Italy (CNR-ISAC)
Other Researchers (Name/Affiliation):	Paolo Davini, Susanna Corti, Matteo Nurisso, Renata Coppo

The following should cover the entire project duration.

Summary of project objectives

(10 lines max)

The MARTINI project aims to accelerate the spin-up of the EC-Earth4 climate model using both deterministic data-driven methods and machine learning (ML) techniques. By reducing the time and computational cost needed to reach climate equilibrium, the project seeks to improve the efficiency of long-term climate simulations. Key objectives include: (1) testing low-resolution configurations of EC-Earth4, (2) developing and validating spin-up acceleration strategies based on trend extrapolation, and (3) extending these approaches to ML methods such as ARX and GANs. The ultimate goal is to establish a robust framework for efficient and physically consistent model initialization.

Summary of problems encountered

(If you encountered any problems of a more technical nature, please describe them here.)

The main technical challenge encountered during this period concerns the implementation of the full stop-and-go acceleration tool, particularly with respect to convective adjustment and the preservation of vertical stability in the ocean restart fields. When trend extrapolation modifies temperature and salinity profiles, the resulting fields may violate the conditions required for static stability, triggering spurious convective mixing upon model restart. Ensuring that the corrected fields remain physically consistent — without introducing instabilities that could alter the model trajectory — has proven non-trivial and remains an active area of work.

Another issue relates to OpenIFS crashes caused by hardcoded array accumulation. Because of this, we were forced to run segmented simulations of 350 years, restarting each subsequent run from the previous ocean restart. Since OpenIFS restarts can be difficult to build and initialize from scratch, we used the 1990 initial conditions. Although the atmosphere evolves rapidly, this approach introduces a minor initialization error. Furthermore, this workaround requires careful management of restart files across simulation chunks. To address this, we successfully implemented a systematic workaround: we designed a procedure that bypasses the issue by manipulating the time-lag indices, effectively preventing OpenIFS from creating arrays of continuously increasing size. While this custom solution resolved our immediate bottleneck, we hope this issue will be addressed in future official releases of OpenIFS through a more robust and native implementation.

Experience with the EMI R&D Project framework

(Please let us know about your experience with administrative aspects like the application procedure, progress reporting etc.)

The experience with the EMI R&D Project framework has been highly positive. Both the application procedure and the progress reporting requirements are streamlined, clear, and straightforward. The low administrative overhead is particularly appreciated, as it allows researchers to focus fully on the scientific and computational aspects of the project on the Atos platform. Additionally, the accounting and system status pages are exceptionally well-designed, making it very easy to monitor resource consumption and machine availability in real-time. The management of system issues and maintenance communication has also been highly efficient, allowing us to focus fully on our computational work.

Summary of results

(This section should comprise up to 10 pages, reflecting the complexity and duration of the project, and can be replaced by a short summary plus an existing scientific report on the project.)

Over its entire duration, the MARTINI project has successfully met its core objectives, bridging advanced high-performance computing (HPC) engineering on the ECMWF Atos platform with theoretical geophysical fluid dynamics. The project has established a validated, robust, and highly accelerated framework for the spin-up of the coupled climate model EC-Earth4.

1. HPC Performance, Scaling, and Load Balancing on Atos

The initial phase of the project focused on porting and optimizing the low-resolution configuration of EC-Earth4 (**TL63L31-ORCA2Z31**) on the ECMWF Atos HPC system. To ensure the highly intensive multi-century simulations were computationally efficient, we conducted extensive strong and weak scaling tests. A primary engineering challenge was optimizing the load balancing between the atmospheric (**OpenIFS**) and the ocean components (NEMO). We performed a detailed analysis of domain decomposition, carefully considering the geometric constraints of the ORCA2 grid and the high-frequency coupling interfaces. The highest computational throughput and lowest idle times were achieved by maintaining an MPI task ratio between NEMO and OIFS of approximately 1:1. This minimized coupling-related delays and neutralized I/O bottlenecks. To automate this optimization process, we developed a Python-based diagnostic tool to collect, parse, and visualize timing and memory-use metrics across runs. This utility was crucial for tuning the hardware configurations of our production runs.

2. Overcoming OpenIFS memory overflows

As the project transitioned to multi-century integrations, a critical technical bottleneck emerged. OpenIFS systematically experienced fatal memory overflows and crashed after several centuries of continuous simulated time. Deep investigation traced the issue to improper memory handling in OIFS, specifically a hardcoded, continuous accumulation of diagnostic arrays over simulated time. On the Atos nodes, this accumulation eventually exhausted the physical memory allocated to the tasks.

To bypass this software-inherent limitation without modifying the core source code of the official OIFS release, we first developed a segmented simulation protocol that splits the long-term integration into manageable segments of approximately 350 years. In this initial approach, NEMO restart files were used to carry the ocean state across segments, while the atmosphere was re-initialized each time from a standard 1990 state. While resetting the atmosphere introduces a minor initialization error, its memory is incredibly short (on the order of weeks) compared to the millennial timescales of the deep ocean; hence, any initialization anomaly decays rapidly and has a negligible impact on the overall trajectory of the ocean spin-up. However, this strategy required careful management of restart files across simulation chunks and posed significant extra work in the post-processing pipeline due to the handling of multiple fragmented runs.

To avoid segmented runs and eliminate even these minor initialization errors, we designed and implemented a more sophisticated, alternative workaround that preserves full continuity in both the atmosphere and the ocean. This procedure directly targets and resets the internal time-stepping and accumulation counters of OpenIFS by manipulating the time-lag indices within the restart metadata. By effectively "faking" these internal indices, the procedure prevents OpenIFS from creating and allocating diagnostic arrays of continuously increasing size as the simulation years advance. This elegant, handcrafted engineering solution allows the coupled model to run continuously across centuries with zero atmospheric drift and no initialization shocks, keeping the memory footprint on Atos completely flat.

This ultimate workaround was successfully validated and deployed, bypassing the OIFS system limitations and enabling the successful, uninterrupted completion of a 5000-year pre-industrial control run—providing the project with an unprecedented, high-quality reference dataset.

3. Multi-millenial pre-industrial control runs

The completion of the uninterrupted 5000-year pre-industrial control run on the Atos platform represents a cornerstone achievement of the MARTINI project, providing an unprecedented baseline for assessing deep-ocean equilibration. A fundamental scientific question that had to be addressed before deploying any trajectory-acceleration strategy (such as the *stop-and-go* protocol) was the uniqueness and stability of the model's equilibrium state. Under identical pre-industrial forcing, does the coupled climate system converge to a single, unique climate equilibrium, or is it highly sensitive to initial conditions?

If the configuration allowed for multiple stable equilibria, any artificial perturbation introduced to accelerate the spin-up (such as trend extrapolation in the restart fields) could permanently drift the simulation into an alternative, physically inconsistent climate state. To investigate this, we analyzed long-duration simulations initialized from two vastly different initial states: (1) a homogeneous, idealized **cold start**; (2) An observation-based initialization using **WOA13 Levitus** climatology.

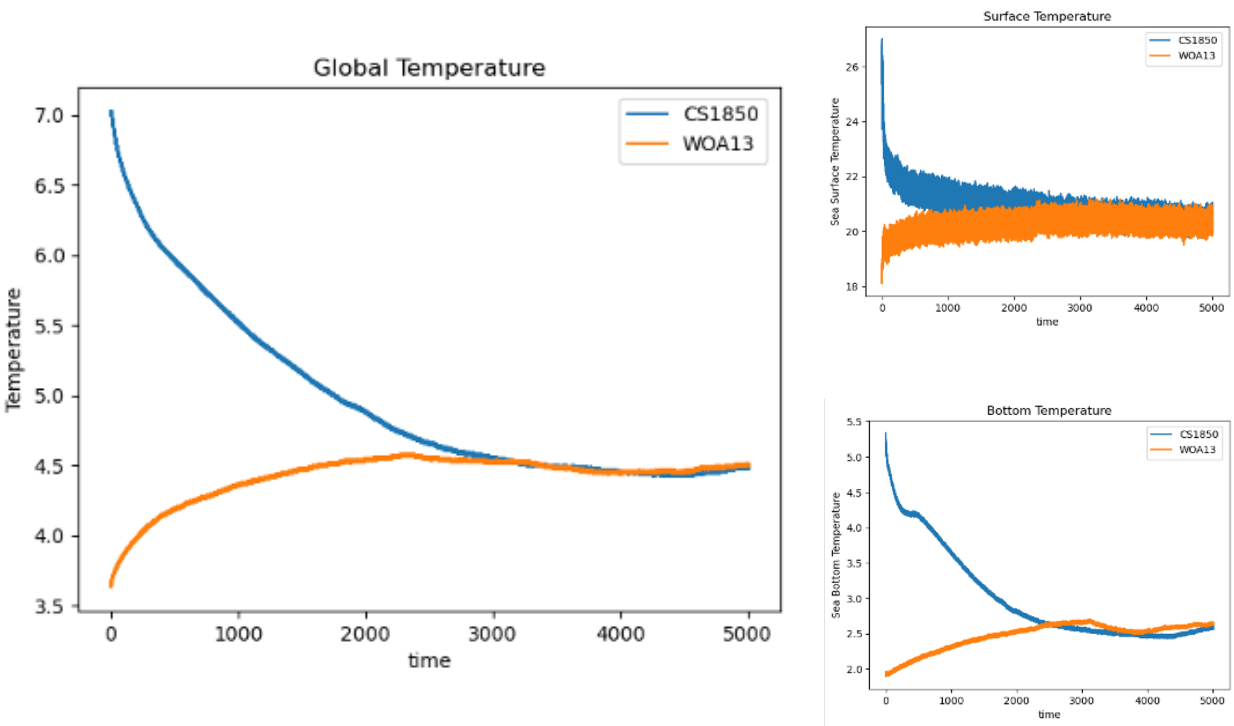

Figure 1: *Timeseries of the control runs starting from two initial conditions, namely cold start (blue), WOA13 Levitus (orange), representing different temperature averages: global ocean temperature (left), ocean surface temperature (upper right), ocean bottom temperature (lower right).*

Over multi-century timescales, both independent trajectories converged asymptotically toward the exact same dynamic equilibrium. This behavior rigorously demonstrates that this low-resolution EC-Earth4 configuration (**TL63L31-ORCA2Z31**) possesses a **single, unique equilibrium state** under pre-industrial forcing. For the MARTINI framework, this result provides the ultimate scientific justification for our *stop-and-go* acceleration protocol. Because the model's final equilibrium is unique, any temporary out-of-equilibrium perturbation introduced during the restart manipulation phase is guaranteed to ultimately decay and relax back toward the correct, physical pre-industrial state. The acceleration method does not risk stabilizing non-physical climate states, and its convergence rate and fidelity can now be evaluated and benchmarked unambiguously against our multi-millennial control run.

4. Run-and-forecast strategy for spin-up acceleration

To drastically reduce the wall-clock time and computational resources required to reach deep-ocean thermodynamic equilibrium, we developed and implemented a deterministic *"run-and-forecast"* acceleration protocol. This framework systematically bypasses the slow, diffusive physical adjustment of the deep ocean by using data-driven extrapolations directly on the model restart states.

The protocol operates as an iterative, semi-automated external supervisor integrated into the EC-Earth4 workflow. The core mechanism is structured as follows:

1. **Run:** The coupled model is integrated for a relatively short temporal window (e.g., a 10-year or 20-year "run" chunk), allowing the fast-varying components of the system to adjust and establish physical trends.
2. **Stop & Analysis:** The simulation is automatically paused. The tool parses the historical outputs to extract spatial and temporal trends of the ocean temperature and salinity fields, typically using Empirical Orthogonal Function (EOF) reconstructions.
3. **Forecast:** Using the calculated trends, the tool performs a mathematical projection (a "forecast jump") into the future. It extrapolates the temperature and salinity anomalies forward in time (representing, for instance, a 10-year or 30-year leap toward equilibrium) and applies this correction directly to the 3D fields of the NEMO restart file.
4. **Convective Adjustment:** The modified fields are checked for physical consistency. Any density inversion is resolved before relaunching.
5. **Restart:** The model is initialized with the newly extrapolated, physically stabilized restart files, and a new cycle begins.

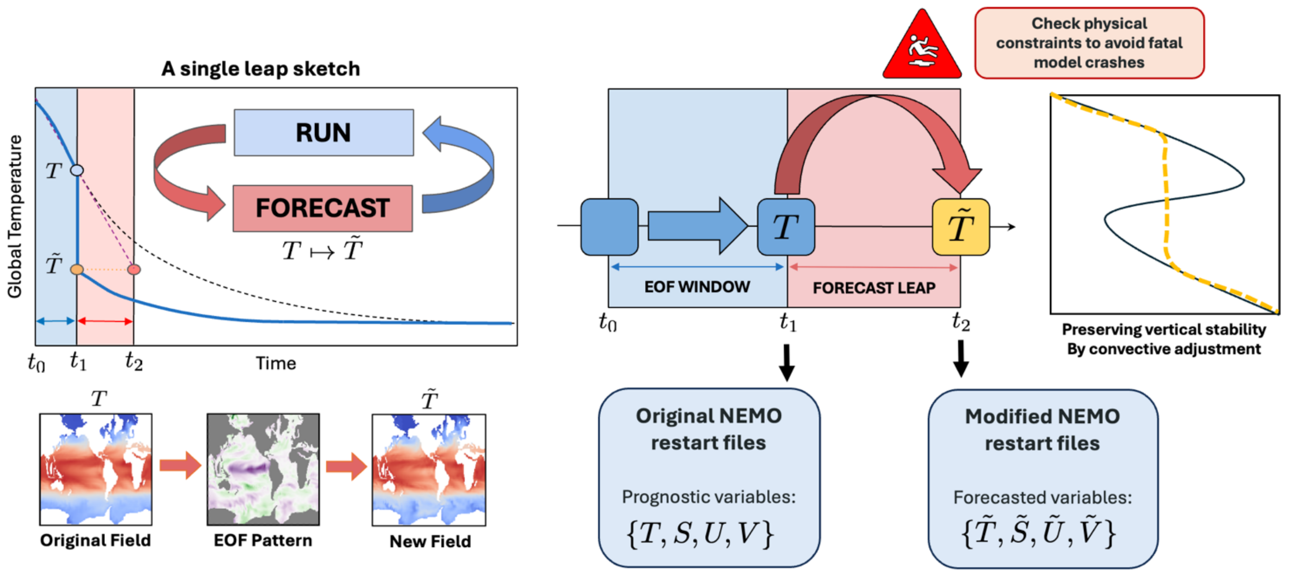

Figure 2: *Sketch of the run-and-forecast method.*

A key challenge of the run-and-forecast method is that the forecast output is intrinsically unconstrained by physical and thermodynamic laws. Because the trend projection is purely data-driven, the extrapolated temperature and salinity fields can easily become physically inconsistent. If these unconstrained fields are written directly into the restart files, they can trigger severe numerical and physical instabilities upon model restart.

Because the EOF trend projection is applied layer-by-layer across the 3D grid, the resulting temperature and salinity anomalies can locally alter the vertical density gradient in a non-physical manner. If a modified profile violates the static stability condition: $\partial\rho/\partial z < 0$, where ρ is the potential density and z is the vertical coordinate, pointing upwards), it creates a localized density inversion (heavier water sitting on top of lighter water). To prevent this, we integrated a convective adjustment module directly into the restart manipulation tool. Before saving the extrapolated fields back to the NEMO binary restart file, the tool performs a 1D column-by-column diagnostic that scans the entire 3D grid to detect any points of static instability. Wherever an instability is flagged, it applies a localized vertical mixing algorithm. This algorithm smooths the temperature and salinity profiles until static stability is restored, while strictly conserving the column's integrated heat and salt content. This ensures that when the coupled simulation resumes, NEMO experiences a smooth transition with zero numerical shock or spurious convection, allowing the accelerated trajectory to target the long-term equilibrium cleanly.

5. Preliminary Acceleration Performance and Convergence Analysis

To evaluate the efficiency and physical fidelity of the run-and-forecast protocol, we performed an initial accelerated simulation spanning a physical model time of approximately 250 years.

A preliminary assessment of the accelerated integration reveals a speed-up of approximately 5 times compared to the standard control run over the simulated period. It is crucial to highlight that this speed-up factor is not static but increases with each iteration of the run-and-forecast cycle. This cumulative performance gain is due to the nominal forecast jump (e.g., 10 years), gained at each leap, plus a "residual" temporal because the trend extrapolation projects the slow-varying components deeper into their transient adjustment phase. Consequently, the effective acceleration rate is non-linear and grows as the simulation progresses. Extending this run to its full completion is therefore necessary to obtain a definitive, quantitative estimate of the method's ultimate computational efficiency.

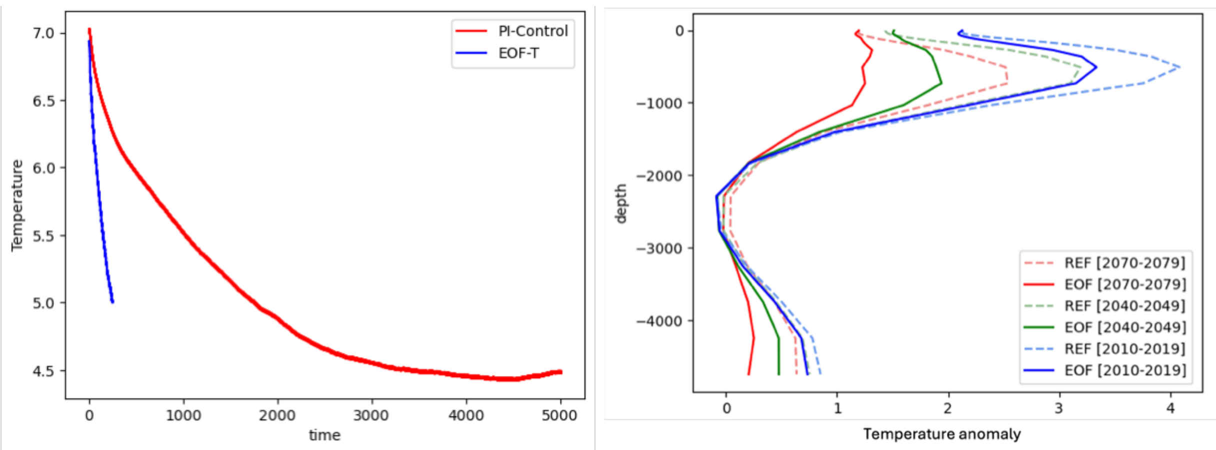

Figure 3. *Left panel: Time series of global mean ocean temperature for the control run (red) and the accelerated run (blue), the latter configured with a 10-year EOF window and a 10-year forecast jump. Right panel: Vertical profiles of ocean temperature anomaly, measuring the distance from the equilibrium state (defined as the 10-year average centered at model year 5000). The control run (dashed lines) and the accelerated run (solid lines) are evaluated over different 10-year averaging windows to show convergence: years 2010–2019 (blue), 2040–2049 (green), and 2070–2079 (red).*

In these initial benchmarks, we evaluated the distance to equilibrium using a straightforward metric based on the 3D field anomaly relative to the end of the 5000-year control run. For the current phase of the project, this simple approach has proven entirely sufficient and highly effective in tracking the general convergence behavior of the model.

Looking ahead, however, as we aim to fully characterize the fine-scale equilibrium state, more sophisticated and robust metrics may be explored. These include evaluating radiative and heat budget imbalances, applying statistical distance measures (such as the Kullback-Leibler divergence), or defining multi-variate cost functions.

When comparing the accelerated simulation to a point in the control run that shares the exact same global mean ocean temperature, we observe that the spatial patterns of Sea Surface Temperature (SST) are distinct. This indicates that the accelerated run and the control run follow different physical trajectories through the model's phase space. This divergence, however, does not imply a failure of the method. On the contrary, looking at the vertical profiles of the temperature anomaly relative to the 5000-year equilibrium (Figure 2, right panel), the accelerated run is systematically and rapidly converging toward the deep-ocean equilibrium.

While the standard control run slowly diffuses heat through the water column over millennia, the accelerated run is forced to bypass the slow, intermediate transient phases of the control trajectory, taking a direct, optimized path toward the final 5000-year equilibrium state.

To identify the optimal configuration for the acceleration protocol, we also conducted a systematic sensitivity analysis across the key parameters of the *run-and-forecast* method. Specifically, we explored the parameter space by varying the run window (simulation chunk, e.g. 10 to 20 years), EOF windows and forecast jump (e.g. 10 to 30 years). The results of this sensitivity sweep reveal a fundamental trade-off. While larger forecast jumps mathematically yield higher nominal speed-ups, the use of linear regression to project the EOF principal components imposes a strict physical limit on the length of the jump. Because the real climate system approaches equilibrium through non-linear radiative and convective feedback, a linear trend extrapolation over too long a horizon (e.g., beyond 20–30 years per jump) overshoots the physical target, leading to severe thermodynamic inconsistencies. This analysis confirms that for larger forecast jumps, simple linear projection is insufficient, highlighting the clear necessity of transitioning to non-linear extrapolation methods—such as the Machine Learning approaches planned for the next phase—to safely achieve larger jumps without violating physical boundaries.

In summary, the project has successfully achieved a complete reference simulation on the Atos platform and established a robust physical and theoretical basis for the "Run-and-Forecast" acceleration strategy. The focus now shifts to finalizing the automated restart-manipulation tool, preparing it to be tested under different climate forcing conditions and adapting it to accelerate high-resolution model configurations. This will allow us to perform a systematic benchmarking sweep of the data-driven method to optimize the balance between computational speed-up and physical consistency. Finally, these baselines will pave the way for initiating the machine learning phase of the project, training non-linear models like ARX and GANs to handle physically consistent forecast jumps.

List of publications/reports from the project with complete references

Alessandro Sozza, Paolo Davini, and Susanna Corti, Data-driven approaches for accelerating ocean spin-up in coupled climate simulations, EGU25-4573, https://doi.org/10.5194/egusphere-egu25-4573, EGU General Assembly 2025, Session ITS1.2/OS4.8 "Machine Learning for Ocean Science".

Future plans

(Please let us know of any imminent plans regarding a continuation of this research activity, in particular if they are linked to another/new EMI R&D Project.)