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	2025……………………
Project Title:	Towards a land/atmosphere/ocean/ice coupled data assimilation system: 20-year coupled reanalysis to initialise longer time-range forecasts, specifically seasonal forecasts (towards C3S multi-model project)
Computer Project Account:	Spitcard
Principal Investigator(s):	Carla Cardinali and Marcelo Guatura
Affiliation:	CMCC………………………….……
Name of ECMWF scientist(s) collaborating to the project (if applicable)	NA…
Start date of the project:	April 2025……
Expected end date:	December 2027……………………

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)			300,000,000	3,322,583
Data storage capacity	(Gbytes)			100.000	14.000

Summary of project objectives (10 lines max)

The goal of this project is to build and deploy a state-of-the-art, fully coupled data-assimilation system, bringing together land, atmosphere, ocean, and sea-ice components to deliver a physically consistent picture of the entire Earth system. By tightly integrating observations across all four domains, we will constrain model fields simultaneously, capturing their complex interactions and feedbacks. Once in place, this framework will produce a 20-year coupled reanalysis, a seamless historical record to furnish high-quality, balanced initial conditions. These improved initial states will reduce spin-up shock, refine key climate modes, and boost the skill and reliability of seasonal-to-subseasonal forecasts. Ultimately, our system and its datasets will offer major added value to international multi-model forecasting efforts and directly support the objectives of the Copernicus Climate Change Service (C3S).

Summary of problems encountered (10 lines max)

Migrating the DA system to the new HPC platform revealed major software-engineering and data-management hurdles. We had to untangle complex dependencies among version-specific libraries (NetCDF, MPI) and compilers, which led to a thorough overhaul of the build system, adapting compilation flags and relinking every component to guarantee numerical reproducibility on the new architecture. At the same time, securely moving and byte-verifying terabytes of initial-condition data exposed weaknesses in our I/O workflows, so we completely rewrote data-access scripts and I/O routines to fit the platform's filesystem. By resolving these dependency conflicts, build errors, and I/O mismatches, we've now laid a rock-solid foundation for the ongoing testing phase.

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

Over the next few months, we will first complete a one-month campaign of standalone land and atmosphere experiments, then move into a fully coupled land–atmosphere configuration for the traditional DA system. In parallel, driven by the recent machine-learning revolution and its promise of faster, more flexible modelling, we will kick off an ML extension that mirrors this workflow. That means retraining our graph-neural-network module on standalone land runs and standalone atmosphere runs, using both observational and model data to capture non-linear interactions and feedbacks. Once both the conventional coupled system and the AI-augmented system are up and running, we'll perform a head-to-head comparison of analysis accuracy, forecast skill, and computational efficiency. This dual path will let us quantify the operational and scientific benefits of each approach and chart the best course for a next-generation, fully coupled Earth-system DA.

List of publications/reports from the project with complete references

……………………………………………………………………………………………………
……………………………………………………………………………………………………
……………………………………………………………………………………………………
………………………………………………

Summary of results

If submitted **during the first project year**, please summarise the results achieved during the period from the project start to June of the current year. A few paragraphs might be sufficient. If submitted **during the second project year**, this summary should be more detailed and cover the period from the project start. The length, at most 8 pages, should reflect the complexity of the project. Alternatively, it could be replaced by a short summary plus an existing scientific report on the project attached to this document. If submitted **during the third project year**, please summarise the results achieved during the period from July of the previous year to June of the current year. A few paragraphs might be sufficient.

………

We now have fully functional standalone configurations for both land-surface and atmospheric components. In the land-surface setup, the model and its assimilation module use prescribed atmospheric forcing from established reanalyses to isolate and validate the land-surface analysis. These experiments quantify the model's response to land-based observations, such as satellite-retrieved soil moisture, leaf area index (LAI), and snow cover, and provide the foundation for tuning land-surface error covariances. Simultaneously, the atmospheric model and its assimilation module are run independently under fixed surface boundary conditions (e.g., climatological soil properties and sea-surface temperatures). By decoupling surface feedbacks, these tests rigorously evaluate atmospheric analysis and forecast performance in response to both conventional and satellite observations. In each 6-hour cycle, approximately 1.5 million observations are assimilated.