LATE REQUEST FOR A EMI R&D PROJECT 2026–2028

Country/Organisation: Netherlands

Principal Investigator[1]: Dr Isma Abdelkader Di Carlo

Affiliation: Institute for Marine and Atmospheric research Utrecht (IMAU), Utrecht University

Address: Princetonplein 5, 3584 CC Utrecht, The Netherlands

Other researchers:

Project Title:

A physically constrained machine-learned emulator of a high-resolution climate model for past and future climates

To make changes to an existing project please submit an amended version of the original form.

If this is a continuation of an existing project, please state the computer project account assigned previously.	**SP (new project)**	
Starting year: (A project can have a duration of up to 3 years, agreed at the beginning of the project.)	2026	
Would you accept support for 1 year only, if necessary?	YES [X]	NO [ ]

Late EMI R&D Projects are normally approved for 1 year, but exceptionally can be approved for 3 years.

Computer resources required for project year:		**2026**	**2027**	**2028**
High Performance Computing Facility	[SBU]	200,000	500,000	500,000
Graphics Processing Unit Cluster-A	[GBU]	540,000	1,050,000	1,050,000
Graphics Processing Unit Cluster-B	[GBU]	0	0	0
Accumulated data storage (total archive volume)[2]	[GB]	15,000	20,000	25,000

[1] The Principal Investigator will act as contact person for this EMI R&D Project and, in particular, will be asked to register the project, provide annual progress reports of the project's activities, etc.

[2] These figures refer to data archived in ECFS and MARS. If e.g. you archive x GB in year one and y GB in year two and don't delete anything you need to request x + y GB for the second project year etc.

[3] The number of vGPU is referred to the equivalent number of virtualized vGPUs with 8GB memory.

EWC resources required for project year:		2026	2027	2028
Number of vCPUs	[#]	0	0	0
Total memory	[GB]	0	0	0
Storage	[GB]	0	0	0
Number of vGPUs[3]	[#]	0	0	0

Continue overleaf.

Principal Investigator: Dr Isma Abdelkader Di Carlo

Project Title: A physically constrained machine-learned emulator of a high-resolution climate model for past and future climates

Extended abstract

All EMI R&D Project requests should provide an abstract/project description including a scientific plan, a justification of the computer resources requested and the technical characteristics of the code to be used. The completed form should be submitted/uploaded at https://www.ecmwf.int/en/research/special-projects/special-project-application/special-project-request-submission.

Following submission by the relevant Member State the EMI R&D Project requests will be published on the ECMWF website and evaluated by ECMWF and its Scientific Advisory Committee. The requests are evaluated based on their scientific and technical quality, and the justification of the resources requested. Previous EMI R&D Project reports and the use of ECMWF software and data infrastructure will also be considered in the evaluation process.

Requests exceeding 10,000,000 SBU should be more detailed (3-5 pages).

Aim. High-resolution climate models resolve storms, ocean eddies and extremes, but they are too expensive to run for the many climates and scenarios that questions about past and future climate need. Machine-learned emulators of the atmosphere can run for decades at a small fraction of that cost, but they are trained on the recent past, and it is not known whether they remain stable and physically consistent in very different climates. This project trains an emulator of the atmosphere of a high-resolution coupled climate model (CESM, with an atmosphere at about 25 km and an eddying ocean) and tests it in climates it never saw, simulated by the same model: the pre-industrial climate, the Last Glacial Maximum, the mid-Pliocene, the early Eocene and an abrupt quadrupling of CO2, as well as the twenty-first century scenarios. The emulator is constrained by conservation laws taken from the parent model's own equations (global dry air mass, water and energy), and its surface fluxes over the ocean are tied to the prescribed sea surface temperature. The scientific question is how far a data-driven atmosphere trained on the recent past can be trusted in other climates, and which physical constraints make it trustworthy.

Data. The training data are the 6-hourly output of nine historical members of the high-resolution simulations (1920 to 1999), coarse-grained to a 1 degree grid; validation uses 2000 to 2005, and the out-of-sample tests use the pre-industrial, glacial, Pliocene, Eocene and 4xCO2 simulations of the same model. All are public in the NCAR data archive. The reduced training set is about 13 GB per member-year, about 9.5 TB for the 720 member-years, plus about 4 TB of validation and test data. For the historical period, the emulator's climatology will also be compared with ERA5 from MARS, to place the parent model's and the emulator's biases side by side.

Code and technical characteristics. The emulator is built in PyTorch on NCAR's open-source CREDIT framework (Apache 2.0), with a vision transformer (CrossFormer) that maps the atmospheric state at three past times and the forcing (sea surface temperature, sea ice, insolation, greenhouse gases) to the state 6 hours later, and is then run forward for years. The conservation constraints are applied after every step in double precision; they were tested on the model output before use. Training is staged (one-step, then two- and four-step rollouts), with checkpoints chosen on multi-day validation rollouts. Jobs run on a single node with one to four GPUs; data are read from NetCDF or zarr by several loader processes per GPU. The code has been run on NVIDIA A100 GPUs at NCAR: one training step takes 0.13 s for a network of 120 million parameters and 0.45 s for 0.9 billion parameters (25 GB of GPU memory) at batch size 1, and short test trainings of a small network (9,000 updates and a 30-day rollout) take about one hour.

Work plan. Late 2026: (1) transfer the reduced training and test data to ECMWF and port the environment to the A100 cluster, reproducing the NCAR test results; (2) first trainings of the mid-size emulator (about 120 million parameters) on part of the training set. 2027: (3) train the mid-size emulator on the full historical training set; (4) run multi-year rollouts in all test climates and evaluate stability, energy and water budgets, mean climate and variability against the parent model and, for the historical period, ERA5. 2028: (5) repeat with the constraints and settings chosen in 2027, and report.

Justification of resources. GPU Cluster-A: one pass over the training set (about one million samples) takes about 38 A100 hours for the mid-size network, so about five passes with multi-step fine-tuning take about 220 hours; multi-year rollouts in five climates and their evaluation take about 40 hours; tests and porting about 30 hours. In total about 290 A100 hours (1,050,000 GBU) in 2027. In late 2026, porting, data checks and the first trainings take about 150 A100 hours (540,000 GBU). HPC Facility: data conversion to zarr, normalisation statistics and the evaluation of the rollouts on CPUs, about 25,000 core-hours (500,000 SBU) a year, 200,000 SBU in late 2026. Storage: the training and test data (about 13.5 TB) are archived in late 2026 (15 TB), with checkpoints and reduced rollout output growing the total to 20 TB in 2027 and 25 TB in 2028. If the approach works, a regular request with larger resources will follow by the 30 June deadline.