REQUEST FOR A SPECIAL PROJECT 2021–2023

MEMBER STATE: Greece, France

Principal Investigator[1]: Vassilios D. Vervatis (1), Pierre De Mey-Frémaux (2) …………

Affiliation: (1) National Kapodistrian University of Athens (UoA) ………….
(2) Laboratoire d'Etudes en Géophysique et Océanographie Spatiales (LEGOS) ……………………………………………

Address: (1) Ocean Physics And Modelling group (OPAM), Faculty of Physics, Division of Environmental Physics and Meteorology, Campus Building PHYS-V, 15784 Athens, Greece………………
(2) LEGOS/CNRS 18, av. Edouard Belin, 31401 Toulouse cédex 9, France ………………………………………………………….

Other researchers: Sarantis Sofianos (1), Nadia Ayoub (2)
…………………………………………………………………….…

Project Title: SDAP/NEMO system
…………………………………………………………………..……

If this is a continuation of an existing project, please state the computer project account assigned previously.	**SPGRVER2**	
Starting year: (A project can have a duration of up to 3 years, agreed at the beginning of the project.)	2021	
Would you accept support for 1 year only, if necessary?	YES ☒	NO ☐

Computer resources required for 2021-2023: (To make changes to an existing project please submit an amended version of the original form.)		**2021**	**2022**	**2023**
High Performance Computing Facility	(MSBU)	1	-	-
Accumulated data storage (total archive volume)[2]	(TB)	1	-	-

Continue overleaf

[1] The Principal Investigator will act as contact person for this Special 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.

Principal Investigator: Vassilios D. Vervatis and Pierre De Mey-Frémaux …………

Project Title: SDAP/NEMO system

Extended abstract

The completed form should be submitted/uploaded at https://www.ecmwf.int/en/research/special-projects/special-project-application/special-project-request-submission.

All Special 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.

Following submission by the relevant Member State the Special Project requests will be published on the ECMWF website and evaluated by ECMWF as well as the Scientific Advisory Committee. The evaluation of the requests is based on the following criteria: Relevance to ECMWF's objectives, scientific and technical quality, disciplinary relevance, and justification of the resources requested. Previous Special Project reports and the use of ECMWF software and data infrastructure will also be considered in the evaluation process.

Requests asking for 1,000,000 SBUs or more should be more detailed (3-5 pages). Large requests asking for 10,000,000 SBUs or more might receive a detailed review by members of the Scientific Advisory Committee.

The work proposed here builds upon the two previous ECMWF Special Projects, with account ids SPGRVERV[3] and SPGRVER2[4], and the two joint projects within the CMEMS Service Evolution named SCRUM[5] and SCRUM2[6]. The resources requested in this new Special Project will be used to support the R&D activities of the University of Athens and LEGOS/CNRS, interfacing the SEQUOIA ocean Data Assimilation Platform - SDAP[7] with the ocean model NEMO[8] (Nucleus for European Modelling of the Ocean; Madec, 2008).

The SDAP system is a complete and modular data assimilation code, specialized in ensemble assimilation with integrated array design and probabilistic forecasting capabilities. SDAP has been used on the ECMWF HPC facilities throughout the whole duration of the aforementioned Special Projects. Recently, its interface with NEMO started to undergo extensive upgrades and changes in the above-mentioned projects and because the changes were extensive, a new branch (1.6) of the code was created in the SDAP repository[9].

In this project, we aim at modifying further the structure of the SDAP code to permit NEMO to fully control the double MPI parallelization, including domain decomposition, and ensemble operation. The R&D activities, codenamed here "Spruce", consists in using the SDAP libraries in a different way than before, i.e. the "Sequoia" flavor, with their differences highlighted in Fig 1.

The following two objectives will be envisaged in the course of 2021:

- A working prototype of "Spruce" using a pilot implementation with a Lorenz-63 model.
- A prototype based on the benchmark NEMO-GYRE configuration to illustrate first steps towards a correction and restart strategy, using the NEMO double parallelization capabilities.

Depending on the results we may ask for a more substantial allowance for production runs in 2022 and perhaps beyond.

---

[3] https://www.ecmwf.int/en/research/special-projects/spgrverv-2016
[4] https://www.ecmwf.int/en/research/special-projects/spgrver2-2018
[5] https://www.mercator-ocean.fr/portfolio/scrum-2/
[6] https://www.mercator-ocean.fr/en/portfolio/scrum2/
[7] https://sourceforge.net/projects/sequoia-dap/
[8] http://www.nemo-ocean.eu
[9] https://sourceforge.net/p/sequoia-dap/code/HEAD/tree/branches/1.6/

Technical Requirements and Computational Resources

Our estimate on the computational resources requested in this work is based on the model ensembles performed during the previous Special Projects. Let us recapitulate briefly a few technical requirements. In the previous projects, we generated ocean model ensembles on CCA and CCB clusters, which are Cray XC40 systems integrating Intel Broadwell nodes, with 36 cores per node and 128 GB (2400 MHz DDR4) memory per node. The NEMO and SDAP codes were compiled under the Intel Broadwell software environment using the Cray Development Toolkit (CDT) cdt/17.03, with intel/17.0.3.053 compiler, and the following libraries: cray-netcdf-hdf5parallel/4.4.1.1 and cray-hdf5-parallel/1.10.0.1. The same environment was used for the compilation of XIOS version 2.0. We used -O3 optimization in the FCFLAGS of the compilation architecture file.

We also made use of NEMO's enhanced MPI strategy whose features allow for parallelization in both the spatial domain and across ensemble members. Free ensemble simulations were then carried out by just one call of the executable. In case of ensemble-based data assimilation experiments we roughly estimate that the computational resources will be double of those of a free ensemble. The prototype Lorenz-63 model is not expensive in terms of computational resources and most of the resources requested here will be used for the benchmark NEMO-GYRE configuration. We anticipate that 1MSBUs and 1TB will suffice to perform ensemble simulations.

Sequoia vs. Spruce

	Sequoia (all-in-1)	Spruce (script-based)
Status	Current	In development
Base	SDAP libraries v1.6	SDAP libraries v1.6 & later
Social behavior	SDAP calls model	Model calls SDAP
Run-time	One unique executable does everything	Several executables launched by a script
Ensemble synchronization	• Ensemble is run in parallel via MPI • Instance 0: Master run • Uses ASEYA library and MPI_Barrier()	• Ensemble is run in parallel via MPI, in sequence, or mixed • Script-based sequencing of tasks (bash or py2f)
Compatibility with double parallelized model and associated diagnostics	• NOT compatible • Domain decomposition NOT allowed	• Compatible • Domain decomposition and // diags allowed
Scalar analysis output	Master analysis (as proxy of Ensemble average)	Ensemble average

Fig 1: Differences between the Sequoia and Spruce flavors of SDAP.

References

Madec, G., 2008: NEMO ocean engine. Note du Pôle de modélisation, Institut Pierre-Simon Laplace (IPSL), France, No 27 ISSN No 1288-1619.