# REQUEST FOR A SPECIAL PROJECT 2022–2024


**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:** CASCADE (Coupled regional coAStal oCeAn moDel Ensembles)

| | | |
|---|---|---|
| If this is a continuation of an existing project, please state the computer project account assigned previously. | **SPGRVER2** | |
| Starting year:<br>(A project can have a duration of up to 3 years, agreed at the beginning of the project.) | 2022 | |
| Would you accept support for 1 year only, if necessary? | YES ☒ | NO ☐ |

| **Computer resources required for 2022-2024:**<br>(To make changes to an existing project please submit an amended version of the original form.) | | **2022** | **2023** | **2024** |
|---|---|---|---|---|
| High Performance Computing Facility | (MSBU) | 2,5 | - | - |
| Accumulated data storage (total archive volume)[2] | (TB) | 2,5 | - | - |

*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:** CASCADE (Coupled regional coAStal oCeAn moDel Ensembles)

# 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, 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 3,000,000 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, and expands, the three previous ECMWF Special Projects, with account ids SPGRVERV[2] and SPGRVER2[3] [4], and the two joint projects within the CMEMS Service Evolution named SCRUM[5] and SCRUM2[6]. A new research project will be submitted in March 2022, within the CMEMS Service Evolution framework[7]. 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 in the preparation of the proposed CMEMS Service Evolution project, and in the project itself if selected. That project is a part of the FA-1 Core Project (PIs: G. Coppini, P. De Mey-Frémaux and G. Liu) submitted for labelling under the United Nations Ocean Decade[8]. The FA-1 Core Project is itself a part of the Coast Predict[9] Programme, which is a labelled Ocean Decade programme.

*"The Copernicus Marine Environment Monitoring Service (CMEMS) provides regular and systematic reference information on the physical state, variability and dynamics of the ocean and marine ecosystems for the global ocean and the European regional seas"*[10]. Data assimilation approaches bridge ocean models and observations to provide an optimal estimate of the ocean state. The CMEMS operational services are constantly evolving by making use of high-resolution ocean models and observations, providing accurate information to the users. In this context, ocean data assimilation forecasting systems need to be improved as a response to this demand.

This project chiefly aims at strengthening CMEMS in the areas of regional and coastal ocean uncertainty modelling, empirical ensemble consistency verification and ensemble data assimilation. Our work will be based on the development of an ensemble ocean data assimilation system, using two-way coupled regional to coastal high-resolution ocean domains, as a case study for the CMEMS Service Evolution and the future capabilities of MFCs.

**Objectives**

[2] https://www.ecmwf.int/en/research/special-projects/spgrverv-2016
[3] https://www.ecmwf.int/en/research/special-projects/spgrver2-2018
[4] https://www.ecmwf.int/en/research/special-projects/spgrver2-2021
[5] https://www.mercator-ocean.fr/portfolio/scrum-2/
[6] https://www.mercator-ocean.fr/en/portfolio/scrum2/
[7] https://www.mercator-ocean.eu/en/tender51/
[8] https://www.oceandecade.org
[9] https://www.coastpredict.org/coastpredict-observing-and-predicting-the-global-coastal-ocean/
[10] CMEMS service evolution strategy: R&D priorities

An Ensemble forecasting system based on two-way coupled regional to coastal high-resolution ocean models will be installed at ECMWF HPCF by the University of Athens and LEGOS/CNRS research teams. The area of interest is the Iberian-Biscay-Irish (IBI) coastal and shelf region, relevant to the CMEMS IBI36 MFC[11] operational system. The objectives of the Special Project can be summarized to the following topics:

- Consolidate the stochastic modelling work which has already been undertaken in past projects using ECMWF HPC Facilities, by using up-to-date perturbation and analysis techniques.
- Expand the ensemble-based work to dual-scale model uncertainties, using two-way coupled techniques for a regional and coastal "parent-child" forecasting system.
- Introduce and test methodologies aimed at checking the consistency of ensembles with respect to CMEMS Thematic Assembly Centres (TAC)[12] [13] [14] data and arrays, and considering also simulated SWOT[15] data from the ocean SWOT simulator[16].
- Perform ensemble-based nested simulations and incremental analysis incorporating an augmented state vector spanning both "parent-child" domains.

## Methodology

In this Special Project, we will setup two-way coupled regional to coastal high-resolution ocean model domains based on NEMOv4.+[17] (Nucleus for European Modelling of the Ocean; Madec and NEMO System Team, 2019). We will make use of the AGRIF[18] library (Adaptive Grid Refinement In Fortran), with the parent configuration based on the regional model BISCAY36 at 1/36º resolution (Vervatis et al., 2016; 2021a; 2021b), which is a sub-domain of the IBI36 MFC, and the child domain extending North from the northern coast of the Iberian Peninsula as a 1:3 subset of BISCAY36 (or a 1:2 subset depending on the computational demand). In **Figure 1**, the prototype two-way coupled ocean modelling system will be designed to perform ensemble nested simulations.

A series of sensitivity twin-experiments will be performed to assess the added value of downscaling model errors (if any), using ensemble-based empirical consistency metrics. To that end, we note that the parent-child ocean modelling system will be interfaced with the SEQUOIA ocean Data Assimilation Platform - SDAP[19] which we have been using in the previous ECMWF Special Projects, incorporating an augmented state vector spanning both domains. We will use the stochastic EnKF (Ensemble Kalman Filter) kernel of SDAP for that purpose. The model performance of the parent-child system will be assessed against an ensemble forecast system without any zooms and/or pseudo-observations from a discarded member. We also intend to run ensemble simulations where we take into account local errors in addition to ensemble downscaling errors (Ghantous et al., 2020). In particular, we intend to perturb surface atmospheric variables, so that we can test various options in the state and control vectors being augmented with those variables.

## Technical Requirements and Computational Resources

The proposed two-way coupled ocean modelling system (physics only) will be installed at ECMWF HPCF, upon selection of the project. Therefore, here, we provide only a rough estimate of the resources required to perform ensemble simulations based on the three aforementioned SPs (the one ended just the previous year). This new prototype modelling system will help us better refine HPC

[11] https://marine.copernicus.eu/about-us/about-producers/ibi-mfc/
[12] https://marine.copernicus.eu/about-us/about-producers/sl-tac/
[13] https://marine.copernicus.eu/about-us/about-producers/sst-tac/
[14] https://marine.copernicus.eu/about-us/about-producers/insitu-tac/
[15] https://earth.esa.int/web/eoportal/satellite-missions/s/swot
[16] https://github.com/SWOTsimulator/swotsimulator
[17] http://www.nemo-ocean.eu
[18] http://agrif.imag.fr
[19] https://sourceforge.net/projects/sequoia-dap/

resources and needs. A rough estimate of computational resources is based on one important assumption: the cost of the proposed two-way coupled configuration i.e., only physics for the NEMO BISCAY36+zoom, is similar to the cost (including memory allocation) of the physics-biogeochemistry online coupled system already installed at ECMWF premises for the previous SPs (i.e., NEMO-PISCES: BISCAY36), pertaining also to the fact that the high-resolution zoom area will have approximately the same grid points as the BISCAY36.

Let us recapitulate briefly a few technical requirements, computational resources and the most recent setup compiling and running the code for the previous NEMO-PISCES BISCAY36 modelling system, in order to make valid assumptions for the resources required for the new NEMO BISCAY36+zoom modelling system. The model ensembles for the previous SPs were generated 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 code was 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. The physics model output consisted of daily files of the ocean state vector.

We 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. BISCAY36 scaled out using 96 processor cores of domain decomposition per ensemble member, excluding land processors. The configuration used the NEMO I/O and was connected to an external server (i.e. XIOS controlled by an XML file), thus increasing the total number of processors for the free ensemble simulations including those handling the I/O specifications (e.g. model variables, domains, grid, output frequencies etc.). Our free ensemble experiments were designed to fit a scalability problem approximately of $O(\mathbf{10^3})$ cores. Taking into account the ECMWF's hardware/software specifications, we have tested the following resources geometry: (a) for 10 members, we have used 960 NEMO processors and 48 XIOS servers filling a total of 28 nodes, (b) for 20 members, we have used 1920 NEMO processors and 24 XIOS servers filling a total of 54 nodes. The free ensemble simulations were submitted as batch jobs for a 30-day run. For these examples, the ECMWF's job epilogue during production indicated a runtime average of about 489 minutes, with runtime standard deviation of approximately 29 minutes, including the first and last reading and writing time-steps.

In line with the above technical requirements, we estimate that the computational and data storage resources comparing the previous and the new modelling system will be approximately the same. We also note that by activating the AGRIF library we do not necessarily need more processors. The simulation experiments planned for the new modelling configuration are similar in terms of computational resources with the previous SPs and the main difference is for the data storage, since we won't archive biogeochemical variables in the new modelling system (but we will archive the ocean physics state for the child domain). Overall, in this SP, we ask resources and data storage at about 2,5 MSBUs and 2,5 TB only for one year i.e., the year 2022, which will be mostly for model development and to perform ensemble test runs. Depending on whether our Service Evolution project is selected by CMEMS, we may ask for more resources the following years 2023-2024.

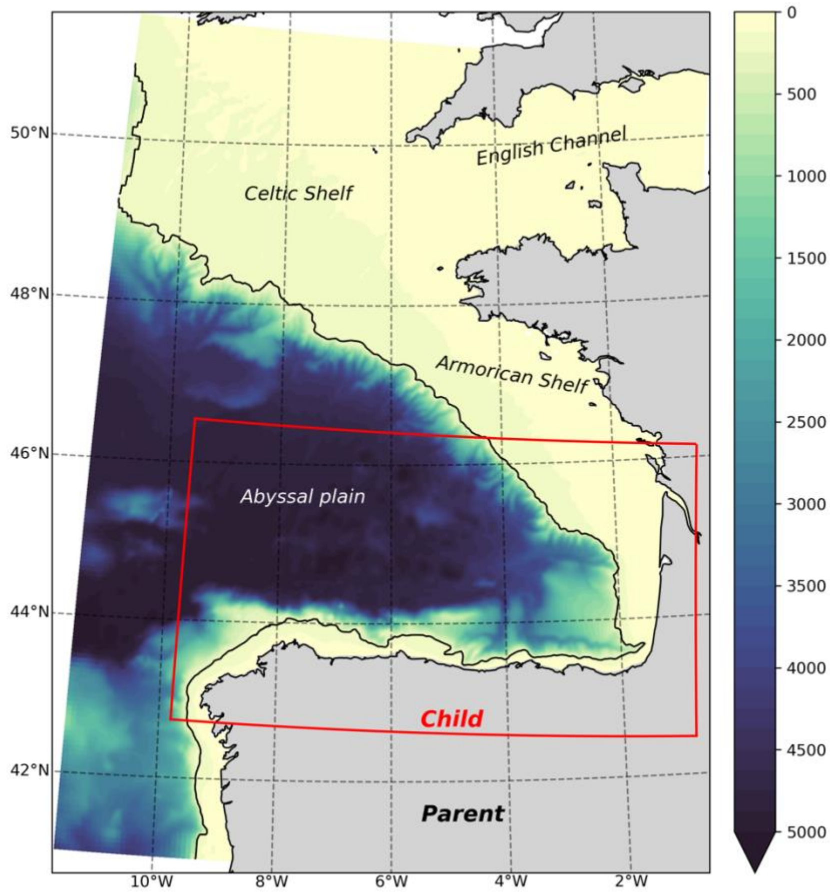


**Figure 1:** Ocean model domains for one possible implementation of the two-way coupled "parent-child" configuration based on NEMO and the AGRIF library. The "parent" configuration is the BISCAY36 regional domain (1/36°) and the "child" configuration will be a high-resolution coastal 1:3 (or 1:2) refinement of BISCAY36.

**References**


Ghantous, M., Ayoub, N., De Mey-Frémaux, P., Vervatis, V., and Marsaleix, P., 2020: Ensemble downscaling of a regional ocean model. Ocean Modelling, 145, https://doi.org/10.1016/j.ocemod.2019.101511.

Madec, G., and NEMO System Team, 2019: NEMO ocean engine. In Notes du Pôle de modélisation de l'Institut Pierre-Simon Laplace (IPSL) (v4.0, Number 27). Zenodo. https://doi.org/10.5281/zenodo.3878122.

Vervatis, V., C.E. Testut, P. De Mey, N. Ayoub, J. Chanut, and G. Quattrocchi, 2016: Data assimilative twin-experiment in a high-resolution Bay of Biscay configuration: 4D EnOI based on stochastic modelling of the wind forcing. Ocean Modelling, 100, 1-19, http://dx.doi.org/10.1016/j.ocemod.2016.01.003.

Vervatis, D. V., P. De Mey-Frémaux, N. Ayoub, J. Karagiorgos, M. Ghantous, M. Kailas, C.-E. Testut and S. Sofianos, 2021: Assessment of a regional physical-biogeochemical stochastic ocean model. Part 1: Ensemble generation, Ocean Modelling, 160, 101781, https://doi.org/10.1016/j.ocemod.2021.101781.

Vervatis, D. V., P. De Mey-Frémaux, N. Ayoub, J. Karagiorgos, S. Ciavatta, R.J.W. Brewin and S. Sofianos, 2021: Assessment of a regional physical-biogeochemical stochastic ocean model. Part 2: Empirical consistency, Ocean Modelling, 160, 101770, https://doi.org/10.1016/j.ocemod.2021.101770.