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: (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: (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	2,5	2,5
Accumulated data storage (total archive volume)[2]	(TB)	2,5	5	7,5

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 three joint projects within the CMEMS Service Evolution (SE) named SCRUM[5], SCRUM2[6] and MULTICAST[7]. The latter SE research project is currently ongoing and will end in 2024. The resources requested in this Special Project are designed to support the R&D activities of the University of Athens and LEGOS/CNRS, performing Ensemble simulations for MULTICAST, as part also for the Project PredictOnTime (PIs: G. Coppini et al.) under the Coast Predict[8] 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"[9]. 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 is 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

An Ensemble forecasting system based on two-way coupled regional to coastal high-resolution ocean models is 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

[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://marine.copernicus.eu/about/research-development-projects/2016-2018/scrum
[6] https://marine.copernicus.eu/about/research-development-projects/2018-2020/scrum2
[7] https://marine.copernicus.eu/about/research-development-projects/2022-2024/MULTICAST
[8] https://www.coastpredict.org/coastpredict-observing-and-predicting-the-global-coastal-ocean/
[9] CMEMS service evolution strategy: R&D priorities

IBI36 MFC[10] 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)[11] [12] [13] data and arrays, and considering also simulated SWOT[14] data from the ocean SWOT simulator[15].
- Perform Ensemble-based nested simulations and incremental analysis incorporating an augmented state vector spanning both "parent-child" domains.

Methodology

In this Special Project, we setup a two-way coupled regional to coastal high-resolution ocean model domains based on NEMOv4.2[16] (Nucleus for European Modelling of the Ocean; Madec and NEMO System Team, 2019). We make use of the AGRIF[17] 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 Galician coast as a 1:3 subset of BISCAY36. In **Figure 1**, the prototype two-way coupled ocean modelling system is designed to perform Ensemble nested simulations.

A series of sensitivity twin-experiments are performed to assess the added value of downscaling model errors, using Ensemble-based empirical consistency metrics. The parent-child ocean modelling system is interfaced with the SEQUOIA ocean Data Assimilation Platform - SDAP[18] which we have been using in the previous ECMWF Special Projects, incorporating an augmented state vector spanning both domains. We use the stochastic EnKF (Ensemble Kalman Filter) kernel of SDAP for that purpose. The model performance of the parent-child system is assessed against an Ensemble forecast system without any zooms and a quasi-reliable set of pseudo-observations from one or several discarded members. We also run Ensemble simulations where we take into account local errors in addition to Ensemble downscaling errors from the parent (Ghantous et al., 2020). In particular, we perturb the surface atmospheric forcing, so that we can test various options in the state and control vectors being augmented with those perturbations.

Technical Requirements and Computational Resources

The proposed two-way coupled ocean modelling system (physics only) is installed at the ECMWF HPCF and is assessed against the resources allocated for the three previous SPs. First Ensemble simulation tests at ATOS show that, the cost of the proposed two-way coupled configuration i.e., only physics for the NEMOv4.2 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 has approximately the same number of grid points as BISCAY36.

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

Let us recapitulate briefly a few technical requirements, computational resources and the most recent setup compiling and running the code for the dual-grid modelling system, in order to make a valid estimate for the resources required for the years 2022-2024. The test simulations of the model Ensembles are performed at ATOS and the code is compiled with the following loaded modules: prgenv/intel, intel-mkl/19.0.5, intel/2021.4.0, hpcx-openmpi/2.9.0, netcdf4-parallel/4.9.1 and hdf5-parallel/1.12.2. The same environment is used for the compilation of the latest versions of XIOS and SDAP platforms. We use -O3 optimization in the FCFLAGS of the compilation architecture file. The physics model output consists of daily files of the ocean state vector for both the parent and the child grids.

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 are carried out by just one call of the executable. Both domains BISCAY36+zoom scale out using 128 processor cores of domain decomposition per Ensemble member, excluding land processors. The configuration uses the NEMO I/O and is 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 are designed to fit a scalability problem approximately of $O(\mathbf{10^3})$ cores. Taking into account the ECMWF's hardware/software specifications, we tested the following resources geometry i.e., for 7 members we used 896 NEMO processors and 128 XIOS servers using a total of 9 nodes (instead of 8 nodes that would normally fit at ATOS, to handle memory allocation). 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 330 minutes, with runtime standard deviation of approximately 10 minutes, including the first and last reading and writing time-steps, and yielding a total of 150.000 SBUs.

In line with the above technical requirements and considering that we perform several nesting and stochastic protocols e.g., 1-way vs. 2-way, perturbing the parent and/or the child, and generating at least 21 members for the parent and child grids per model Ensemble experiment, we ask resources and data storage at about 2,5 MSBUs and 2,5 TB per year i.e., for the whole SP period 2022-2024 a total of 7,5 MSBUs and 7,5 TB. Depending on whether data assimilation experiments are performed, we may ask more resources for the last year 2024.

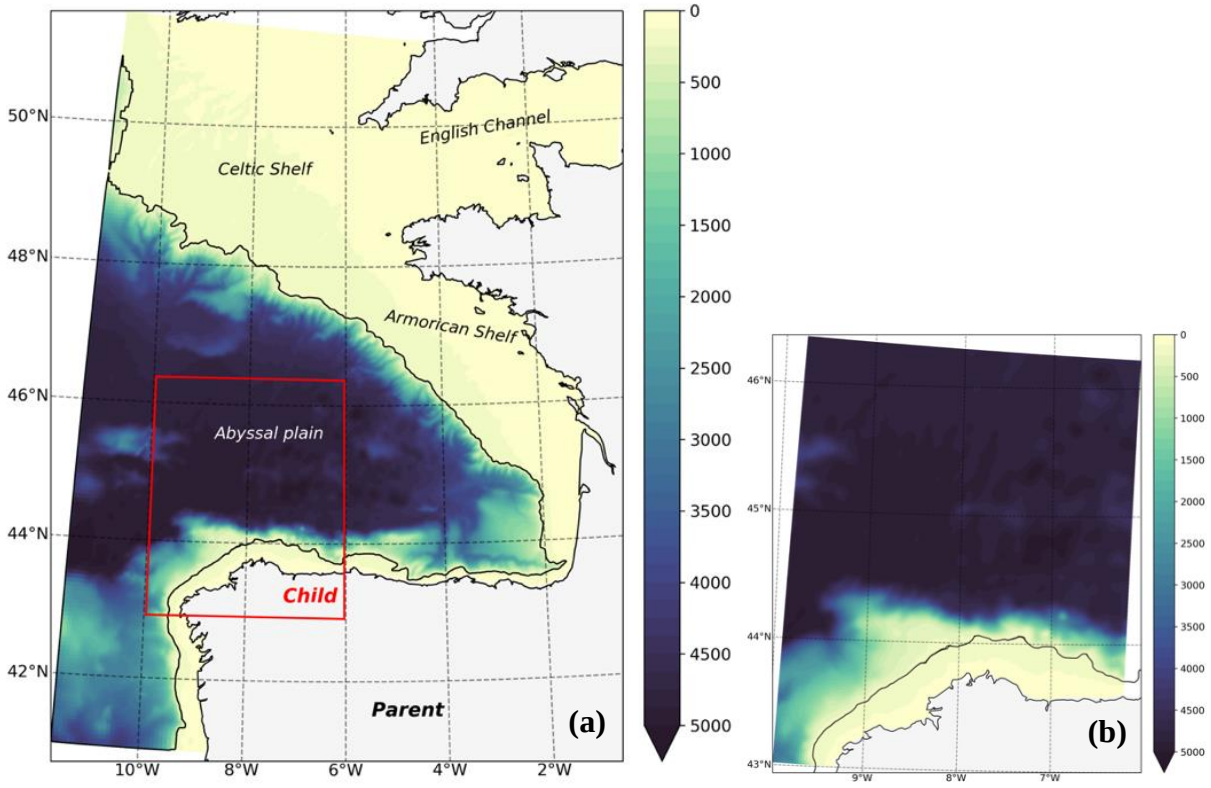

Figure 1: Ocean model "parent-child" dual-grid NEMOv4.2 configuration using the AGRIF library: (a) "parent" BISCAY36 regional domain (1/36°), (b) "child" high-resolution 1:3 refinement (1/108°).

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.