REQUEST FOR ADDITIONAL RESOURCES IN THE CURRENT YEAR FOR AN EXISTING SPECIAL PROJECT

MEMBER STATE: Greece, France …….……...…….……...…

Principal Investigator[1]: Vassilios D. Vervatis (1), Pierre De Mey (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/CNES 18, av. Edouard Belin, 31401 Toulouse cédex 9, France …….

E-mail: (1) vervatis@oc.phys.uoa.gr …….....
(2) pierre.de-mey@legos.obs-mip.fr …….

Other researchers: Sarantis Sofianos (1), Nadia Ayoub (2), …….……
Charles-Emmanuel Testut (Mercator Ocean, Ramonville St. Agne, France) ……

Project title: Stochastic Coastal/Regional Uncertainty Modelling: sensitivity, consistency and potential contribution to CMEMS ensemble data assimilation …….

Project account: **SPGRVERV**

Additional computer resources requested for	**2016**
High Performance Computing Facility (units)	7 MSBU
Data storage capacity (total) (Gbytes)	11.500 GB

Continue overleaf

[1] The Principal Investigator is the contact person for this Special Project

Technical reasons and scientific justifications why additional resources are needed

This is a technical report requesting additional resources for the year 2016, following the progress report of the project SPGRVERV. The ECMWF-SP initial resources were based on an estimate from a previous study, implementing a different ensemble approach and no ecosystem model. For this, we have performed test runs to reassess the computational resources and the disk storage of the ensemble data. In the following paragraph we give an analytical description of the scheduled sensitivity experiments within 2016 (cf. also ECMWF-SP progress report). The ECMWF-SP resources are used in a joint project within the CMEMS Service Evolution open tender under Lot 3: links with coastal environment. The CMEMS project is a two year project with a tight calendar, especially for the first year. An important objective of our runs so far has been to precisely determine the resources that would be needed for the work planned in the CMEMS project, another one being the single-source sensitivity experiments, both of which explaining why we haven't consumed much yet.

The ECMWF-SP is on its first year, where we have installed an ocean-biogeochemical system for the Bay of Biscay (i.e. BISCAY36). The modelling platform is based on the latest stable version of NEMOv3.6-PISCESv2 with ensemble capabilities. Test runs have been carried out on CCB using the new architecture of the cluster (Intel Broadwell node: 36 physical cores and memory limit 120 MB). In Table 1, we present approximately the resources needed per ensemble member and simulation period of one day for two perturbation approaches: a) perturbing only 2D atmospheric-ocean variables focusing on upper-ocean properties; b) perturbing only 3D BGC variables. The latter perturbation approach consists of 24 prognostic variables and therefore is more expensive compared to the former experiment. In addition, the model outputs are organized in daily, hourly and 6-hrs files for the ocean state, the ecosystem model and the stochastic fields (Table 1). In order to free some disk space the following files will not be permanently saved: a) the hourly files in all experiments i.e. file2.nc; b) the stochastic outputs almost in all experiments i.e. file5.nc and file6.nc; c) the ecosystem variables, except chlorophyll, in all medium range experiments i.e. file3.nc; d) the ocean state variables when only BGC variables are perturbed i.e. file1.nc are identical files with the free run, since the ocean is one-way coupled to the ecosystem model; e) all restart files at the end of each simulation. The optimal BISCAY36 domain decomposition is 96 CPU cores (excluding land points). The modelling system is compiled on Intel environment using the XIOS library to handle I/O processes. In Table 2, we illustrate in more details the ensemble sensitivity experiments and their resources. The experiments comprise of ensembles of 20 and 40 members, for one and seven months respectively. In order to use our resources with caution we prepare batch jobs consisting of 10 ensemble members for the following reasons: a) due to XIOS memory constraints; b) avoid spending resources on a false experiment or a late cancelled job; c) optimal use of PBS options e.g. nodes, memory, walltime, wrappers etc. A batch job of 10 members on CCB Broadwell nodes consists of 96 * 10 = 960 CPU cores (hyperthreading = 1). The specific resources geometry fills 26 nodes (i.e. 26 * 36 = 936 CPU cores), plus 24 additional CPU cores on the 27^{th} node. In order to fill properly the last node we use 12 CPU cores for XIOS, dedicated only to I/O processes for the model's diagnostics (using the option writing in one netcdf file and not in domain decomposition of several files). Batch jobs are designed to be resubmitted every month of simulation, not exceeding a 16 hours walltime. In Table 2, we present the summary of all modelling systems and sensitivity experiments. The estimated resources/storage for the first year of the project is approximately (in total after rounding numbers) 9 MSBU computational resources and 15,5 TB disk storage. In the initial ECMWF-SP proposal we have underestimated the first year resources requesting 2 MSBU and 4 TB disk storage. Therefore, the additional resources requested for 2016 is 9 – 2 = 7 MSBU and 15,5 – 4 = 11,5 TB. Finally, we estimate that the level of resources that will be needed for the second year, are close to what we have requested in the initial proposal i.e. 4 MSBU and 8 TB for the year 2017.

Table 1. Estimate of computational resources and data storage per ensemble member and simulation period of one day.

perturbation approach	comp. resources (SBU) per member/day	netcdf files and data storage (MB) per member/day
2D variables of atmospheric-ocean models (experiments S1-7 and Ens1 in Table 2)	~350	file1.nc (daily state variables): 228 file2.nc (hourly state variables): 140 file3.nc (daily ecosystem variables): 592 file4.nc (daily chlorophyll variable): 123 file5.nc (6-hrs 2D stochastic fields): 22 file6.nc (6-hrs 3D stochastic field): 162 restart files will not be saved
3D variables of ecosystem model (experiments S8 and Ens2 in Table 2)	~500	

Table 2. Scheduled ensemble sensitivity experiments and estimated resources/storage for all modelling systems (cf. Table 1 and ECMWF-SP progress report).

experiment	variable	*simulation period/ members	computational resources (MSBU)	files and data storage (TB)
	Atmospheric model			
S1	Wind	1-month/ 20-members	0,21	file1.nc; file4.nc; 0,2106
S2	Tair	1-month/ 20-members	0,21	file1.nc; file4.nc; 0,2106
S3	SLP	1-month/ 20-members	0,21	file1.nc; file4.nc; 0,2106
	Ocean model			
S4	Cd, Ch, Ce	1-month/ 20-members	0,21	file1.nc; file4.nc; 0,2106
S5	K_{PAR}	1-month/ 20-members	0,21	file1.nc; file4.nc; 0,2106
S6	Cd_b	1-month/ 20-members	0,21	file1.nc; file4.nc; 0,2106
S7	S1-6	1-month/ 20-members	0,21	file1.nc; file4.nc; file5.nc; 0,2238
Ens1	S1-6	7-months/ 40-members	2,94	file1.nc; file3.nc; file4.nc; 7,9212
	Ecosystem model			
S8	all BGC	1-month/ 20-members	0,3	file4.nc; file6.nc; 0,171
Ens2	all BGC	7-months/ 40-members	4,2	file3.nc; file4.nc; 6,006
	Summary for all modelling systems and experiments			
			8,91 (~9)	15,5856 (~15,5)

*1-month: medium range ensemble on April, 2012; 7-months: seasonal range ensemble on 01 Dec, 2011 - 30 June, 2012.