REQUEST FOR A SPECIAL PROJECT 2015–2017

MEMBER STATE: ITALY

Principal Investigator[1]: ANDREA STORTO

Affiliation: CMCC

Address: VIALE ALDO MORO, 44, 40127, BOLOGNA, ITALY

……………………………………………………………………………

……………………………………………………………………………

E-mail: **andrea.storto@cmcc.it**

Other researchers: Andrea Storto, Raffaele Farina(1), Andrea Cipollone, Salvatore Cuomo (2), Ardelio Galletti, Livia Marcellino, Giulio Giunta(3)

(1) Istituto di Calcolo e Rete ad Alte Prestazioni – CNR- ITALY

(2) Dipartimento di Matematica e Applicazioni "R. Caccioppoli" Università degli Studi di Napoli Federico II

(3) Laboratorio di Modellistica Numerica e Calcolo Parallelo, Università degli Studi di Napoli Parthenope

……………………………………………………………………………

Project Title:

Optimization of the OceanVar oceanographic data assimilation system for high-resolution applications

If this is a continuation of an existing project, please state the computer project account assigned previously.	**SP** ________________	
Starting year: (Each project will have a well defined duration, up to a maximum of 3 years, agreed at the beginning of the project.)	2015	
Would you accept support for 1 year only, if necessary?	YES X	NO ☐

Computer resources required for 2015-2017: (The maximum project duration is 3 years, therefore a continuation project cannot request resources for 2017.)		**2015**	**2016**	**2017**
High Performance Computing Facility	(units)	150000	150000	150000
Data storage capacity (total archive volume)	(gigabytes)	500	500	500

An electronic copy of this form ***must be sent*** *via e-mail to:* *special_projects@ecmwf.int*

1 The Principal Investigator will act as contact person for this Special Project and, in particular, will be asked to register the project, provide an annual progress report of the project's activities, etc.

Electronic copy of the form sent on (please specify date):

01/04/15

Continue overleaf

Principal Investigator: **Andrea Storto**

Project Title: Optimization of the OceanVar oceanographic data assimilation system for high-resolution applications

Extended abstract

OceanVar is the variational ocean data assimilation system developed at CMCC to provide initial conditions for ocean reanalyses (Storto et al., 2011) and operational forecasts (Dobricic and Pinardi, 2008). It is a three-dimensional variational (3DVAR) system, formulated in its classical incremental variant. It supports the assimilation of all ocean in-situ measurements of temperature and salinity and the assimilation of remotely sensed data (altimetry, sea surface temperature and salinity), from a number of different datasets depending on the specific application. The background-error covariance matrix is composed by two linear terms accounting, respectively, for vertical covariances and horizontal correlations. The former are modeled through the use of multivariate EOFs, while the latter through the application of recursive filters (RF). Multivariate balances can be specified as purely statistical or through the application of steady state balances or simplified models. Pre-processing of the observations includes climatology and background quality checks and thinning of dense observations. OceanVar is hybrid MPI-OpenMP parallel and runs on a variety of different architectures.

CMCC has recently developed a very high resolution Ocean General Circulation Model (OGCM) based on NEMO ocean engine, at 1/16 horizontal resolution (Iovino et al., 2014). The use of OceanVar for this global model is clearly demanding. It is therefore necessary to improve the code in terms of optimization and scalability. From previous profiling reports, the horizontal operator, responsible for the horizontal spread of the observational information, is by far the most demanding part of the data assimilation system. We will therefore concentrate our optimization efforts on the horizontal operator, with the aim of improving its performance in a parallel architecture.

At the moment, the horizontal operator consists of recursive filter. Recently, the recursive filter was upgraded to a third-order formulation (Farina et al., 2015), which allowed improved performance and accuracy since it is able to mimic a Gaussian shape in only 1 iteration of the filter, while the previous first-order recursive filter needed at least 4 iterations to provide a Gaussian-like shape of the horizontal correlation. On the other hand, the formulation was modified to allow non-uniform horizontal correlations (Storto et al., 2014), which turned out to have crucial impact especially in areas of strong mesoscale activity.

However, the current system is characterized by some weaknesses, regarding its boundary (land-sea) conditions and its parallelization strategy, which compromise the overall performances. Concerning the former, to overcome the land discontinuity, the filter projects the geographical domain onto an imaginary irregular grid, with "ghost points" at the sea boundary in order to neglect the effects of the recursive filter boundary conditions. This turns out to i) increase the memory consumption of OceanVar (the recursive filter imaginary grid size is typically much larger than the geographical grid) and ii) make the parallelization non-trivial as it requires a separate domain decomposition with respect to that of the geographical grid. The second weakness regards the parallelization strategy: at the moment OceanVar has a hybrid parallelism. OpenMP serves the parallelization along the vertical levels, while MPI serves the parallelization on macro domain on the horizontal. The macro-domains are blind to each other, i.e. the OceanVar solution is not globally optimal, and to limit the discontinuities of the solutions we use a halo region that typically extends to a few tenths of gridpoints. This intrinsically limits the scalability of the system, further to not providing a solution that is globally optimal.

In this project, we aim at overcoming the two weaknesses. We will begin with determining suitable analytic boundary condition for the RF as described in the work of Triggs et al (2006). The analytic boundary conditions want to replace the existing ghost points, in order to remove the edge artifacts, introduced by the RF for the computing of the background-error-horizontal-covariance, as described by Cuomo et al (2014).

From a numerical point of view, this novel solution can accelerate the convergence of the L-BFGS algorithm, used for minimizing the cost function in OceanVar. From a parallel point of view it removes the imaginary sea points on the lands providing a better balance between computational system loads in the horizontal domain decomposition and hence eliminating possible bottlenecks represented by the sub-domains with many coastlines.

In OceanVar, RF is a loop that has loop-carried dependencies along x and y directions. The existing horizontal parallel strategy of OceanVar is the domain decomposition where the RF has been localized on the macro-areas but they are blind and hence the RF is stopped without transferring of the information to the neighbor sub-domains. This kind of solution replaces the use of pipeline method that needs to transfer date to the near sub-domains but introduces significant discontinuities in the global solution if the observations are too close at the borders of the sub-domain. The proposed improvement is to determine new fitting conditions on the borders to remove these discontinuities and preserve the quasi-Gaussian shape of the horizontal correlation. This represents a novel solution in the parallel strategy of the 3DVAR method, in order to obtain local solutions such whose union satisfactorily approximates the global solution.

References

Cuomo, Salvatore, et al. "An error estimate of Gaussian Recursive Filter in 3Dvar problem." Computer Science and Information Systems (FedCSIS), 2014 Federated Conference on. IEEE, 2014.

Dobricic, S., and N. Pinardi, 2008: An oceanographic three-dimensional variational data assimilation scheme, Ocean Modelling, 22, 89-105.

Farina, R., Dobricic, S., Cuomo, S., Some numerical enhancements in a data assimilation scheme (2013) AIP Conference Proceedings, 1558, pp. 2369-2372. Cited 1 time. DOI: 10.1063/1.4826017

Farina, R; S Dobricic; A Storto; S Masina; S Cuomo; A Revised Scheme to Compute Horizontal Covariances in an Oceanographic 3D-VAR Assimilation System. J. Comput. Phys. (2015), http://dx.doi.org/10.1016/j.jcp.2015.01.003

Iovino D, Storto A, Masina S, Cipollone A, Stepanov V, 2014, GLOB16, the CMCC global mesoscale-eddying ocean, CMCC Research Paper RP0247, December 2014, available at http://www.cmcc.it/wp-content/uploads/2015/02/rp0247-ans-12-2014.pdf

Storto, A., S. Dobricic, S. Masina, and P. Di Pietro, 2011: Assimilating along-track altimetric observations through local hydrostatic adjustments in a global ocean reanalysis system. Mon. Wea. Rev. 139, 738-754.

Storto, Simona Masina, and Srdjan Dobricic, 2014: Estimation and Impact of Nonuniform Horizontal Correlation Length Scales for Global Ocean Physical Analyses. J. Atmos. Oceanic Technol., 31, 2330–2349.

Triggs, Bill, and Michaël Sdika. 2006 "Boundary conditions for Young-van Vliet recursive filtering." Signal Processing, IEEE Transactions on 54.6: 2365-2367.