REQUEST FOR A SPECIAL PROJECT 2013–2015

MEMBER STATE: France

Principal Investigator[1]**:** Anthony Weaver

Affiliation: CERFACS

Address: 42, avenue Gaspard Coriolis

31057 Toulouse Cedex 1

France

E-mail: weaver@cerfacs.fr

Other researchers: Selime Gurol, Andrea Piacentini

Project Title: Ensemble variational data assimilation with NEMOVAR

If this is a continuation of an existing project, please state the computer project account assigned previously.	**SP**FRVODA	
Starting year: (Each project will have a well defined duration, up to a maximum of 3 years, agreed at the beginning of the project.)	2012	
Would you accept support for 1 year only, if necessary?	YES	NO X

Computer resources required for 2013-2015: (The maximum project duration is 3 years, therefore a continuation project cannot request resources for 2015.)		**2013**	**2014**	**2015**
High Performance Computing Facility	(units)	1,000,000	1,000,000	
Data storage capacity (total archive volume)	(gigabytes)	10,000	10,000	

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

Electronic copy of the form sent on (please specify date): April 30, 2012

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 an annual progress report of the project's activities, etc.

Principal Investigator: Anthony Weaver

Project Title: Ensemble variational data assimilation with NEMOVAR

Extended abstract

It is expected that Special Projects requesting large amounts of computing resources (500,000 SBU or more) should provide a more detailed abstract/project description (3-5 pages) including a scientific plan, a justification of the computer resources requested and the technical characteristics of the code to be used. The Scientific Advisory Committee and the Technical Advisory Committee review the scientific and technical aspects of each Special Project application. The review process takes into account the resources available, the quality of the scientific and technical proposals, the use of ECMWF software and data infrastructure, and their relevance to ECMWF's objectives. - Descriptions of all accepted projects will be published on the ECMWF website.

General context and objectives

The aim of this project is to further the development of the variational ocean data assimilation system NEMOVAR (Mogensen *et al.* 2012) for the NEMO model. NEMOVAR is developed as part of a collaboration between CERFACS, ECMWF, Met Office and INRIA/LJK (Mogensen *et al.* 2009), and more recently the University of Reading. NEMOVAR is used with the operational monthly/seasonal forecasting system at ECMWF (OceanS4) and for ocean renanalysis (ORAS4; Balmaseda *et al*. 2012). The Met Office is developing NEMOVAR for their ocean assimilation activities, covering short-range ocean forecasting applications using the FOAM system, and seasonal forecasting. NEMOVAR analyses are used by several other groups for initializing decadal predictions. This Special Project will help to advance research and development with NEMOVAR.

The main focus of this project will be to develop and evaluate a hybrid data assimilation method for NEMOVAR, which incorporates statistical information from an ensemble into the background-error covariance formulation of the variational analysis. There are three main objectives to this project: 1) to develop numerical methods for solving a 2D or 3D implicit diffusion equation for representing anisotropic background-error correlation functions; 2) to use an ensemble version of NEMOVAR to improve the estimates of both the background-error variances and the parameters of the (local) anisotropic tensor of the background-error correlation functions implied by the diffusion operator; and 3) to evaluate the impact of the improved estimates of the background-error covariances in global ocean reanalysis experiments.

Representation of anisotropic and inhomogeneous background-error correlations via an implicit diffusion equation

Differential operators derived from the explicit or implicit solution of a diffusion equation are widely used for modelling background-error correlations in variational ocean data assimilation. The theoretical basis for employing a diffusion equation to represent the action of a correlation operator is described in Weaver and Courtier (2001), Mirouze and Weaver (2010) and more recently Weaver and Mirouze (2012). The latter article lays out the framework for defining both isotropic and anisotropic correlation functions. In the isotropic case, the correlation functions implied by explicit diffusion are approximately Gaussian, whereas those implied by implicit diffusion belong to the larger class of Matérn functions which contains the Gaussian function as a limiting case. Anisotropic Gaussian and Matérn correlation functions can then be constructed from their isotropic counterparts by replacing the standard normalized Euclidean distance measure with a Mahalanobis distance measure that accounts for directionality via the inverse of an aspect tensor. Analytical expressions relate the aspect tensor to the so-called Daley tensor, which can be viewed as a generalization of the Daley length-scale in the isotropic case. These tensors can in turn be related to the parameters of the anisotropic formulations of the explicit and implicit diffusion operators.

Explicit diffusion methods are subject to a stability criterion that can make them expensive. Implicit diffusion methods have the advantage of being unconditionally stable (as well as having the capacity to represent a larger class of correlation functions) but require the solution of a linear system. A 3D correlation operator formulated as a symmetric product of 1D implicit diffusion operators has been developed recently for NEMOVAR (Mirouze 2010). In this approach the 1D implicit operators are applied independently in each orthogonal direction of the computational coordinates. The new formulation has resulted in a significant reduction in computational cost compared to the explicit formulation. Implicit diffusion operators require the inversion of a matrix. For each 1D correlation operator, this matrix is small enough that it can be inverted using Cholesky factorization combined with a forward elimination and backward substitution algorithm. Within this framework, the ensemble data assimilation system will be used to estimate the directional scale parameters of the 1D implicit diffusion operators.

The 1D approach is not practical for accounting for anisotropic correlations and has limitations on massively parallel machines. As an alternative, a 3D correlation model will be constructed directly from a 3D or 2D x 1D anisotropic formulation of the implicit diffusion equation. Different iterative methods (conjugate gradient, multi-grid) will be explored for solving the diffusion equation, paying particular attention to their suitability to be applied with NEMOVAR in a massively parallel framework.

Estimation of the background-error covariance parameters using an ensemble

An ensemble version of the NEMOVAR 3D-Var system will be used to obtain estimates of the background-error variances and elements of the anisotropic tensor. The procedure for computing ensembles of 3D-Var analyses and forecasts is similar to the one developed for NEMOVAR's predecessor OPAVAR (Daget *et al.* 2009). At ECMWF a multi-annual 5-member ensemble has been used to provide multiple ocean analyses for initializing probabilistic coupled forecasts. Here we wish to exploit this system for extracting information about the background-error covariances. The ensemble-generation procedure currently employs perturbed data-sets for surface windstress, different initial states sampled from an existing long reanalysis, and a random rejection criterion for the assimilated observations. Possible extensions to this procedure will be investigated such as including perturbations to the precipitation and sea surface temperature (SST) data-sets, and to the actual values of the assimilated observations.

The estimation of the variances is straightforward but care will be needed to determine an adequate temporal and spatial filtering strategy to minimize sampling errors. The estimation of the Daley tensor elements will be done using the ensemble-based method of Belo Pereira and Berre (2006) which is described in the context of diffusion modelling by Weaver and Mirouze (2012). The key point is that since the number of independent parameters needed to specify the local Daley tensor is of the order of the total number of grid points N, sampling errors are inherently much smaller than those involved in the order N^2 estimation problem of the full correlation matrix. Conventional background-error covariance formulations constructed from small ensembles require somewhat *ad hoc* methods to localize the covariances. It is important to emphasize that *no* such localization is required with the method proposed here. A thorough diagnostic study will be done using reanalysis results from the existing ECMWF NEMOVAR ensemble system before attempting to apply ensemble covariance estimates in an assimilation experiment. Special attention will be given to assess the importance of the off-diagonal elements of the tensor as these elements indicate the importance of anisotropy in the xy, xz and yz directions.

Global ocean reanalysis with ensemble-estimated background-error covariances

The third aspect of this project is to evaluate the impact of the ensemble-estimated covariances in global ocean reanalysis experiments. The resources requested for this project are primarily needed for experimentation with the ensemble system. Experiments will employ the same global model configuration (ORCA1) and assimilation environment as in System 4. In a first step, the assimilation diagnostics of Desroziers *et al.* (2005) will be used to provide improved estimates of the observation-error *variances* for the assimilated temperature (T) and salinity (S) profiles. The current procedure for specifying the observation-error variances for T and S is very simple, being based on analytical profiles that are independent of horizontal position, except near coastlines where the variances are inflated. The absence of an adequate spatial variation in these variances can be problematic in non-eddy resolving configurations such as ORCA1 where representativeness error is large and in combination with an adaptive background-error variance formulation (Daget *et al.* 2009). Preliminary work has already been done to estimate the observation-error variances using the Desroziers *et al.* (2005) method. The procedure has involved binning time-averaged estimates of the variances onto a 2 deg. x 2 deg. regular grid and then interpolating these variances to the observation points. As expected, large variances appear in boundary current regions such as the Gulf Stream which are dominated by (unresolved) eddy-variability.

A first set of experiments will be performed using the existing quasi-isotropic background-error correlation model in combination with variance and length-scale scale estimates from the ensemble. A second set of experiments will then be performed with an anisotropic version of the correlation model where the local anisotropic tensor is estimated from the ensemble. Physical and assimilation diagnostics will be used to evaluate the quality of the analyses compared to those produced with the existing background-error covariance formulation.

References:

Balmaseda, M. A., Mogensen, K. and A. T. Weaver, 2012. Evaluation of the ECMWF Ocean Reanalysis ORAS4. In preparation for *Climate Dynamics*.

Belo Pereira M. and L. Berre, 2006: The use of an ensemble approach to study the background error covariances in a global NWP model. *Mon. Wea. Rev*., **134**, 2466-2489.

Daget, N., Weaver, A. T. and M. A. Balmaseda, 2009: Ensemble estimation of background-error variances in a three-dimensional variational data assimilation system for the global ocean. *Q. J. R. Meteor. Soc*., **135**, 1071-1094.

Desroziers, G., Berre, L., Chapnik, B. and P. Poli, 2005: Diagnosis of observation, background and analysis error statistics in observation space. *Q. J. R. Meteor. Soc*., **131**, 3385-3396.

Mirouze, I, 2010: Régularisation de problèmes inverses à l'aide de l'équation de diffusion généralisée. PhD thesis, Université Paul Sabatier, Toulouse.

Mirouze, I. and A. T. Weaver, 2010: Representation of correlation functions in variational assimilation using an implicit diffusion operator. *Q. J. R. Meteorol. Soc.*, **136**, 1421-1443.

Mogensen, K., Balmaseda, M. A., Weaver, A. T., M. Martin and A. Vidard, 2009: NEMOVAR: A variational data assimilation system for the NEMO ocean model. ECMWF Newsletter, Number 120, Summer 2009.

Mogensen, K., M. Balmaseda and A. T. Weaver, 2012: The NEMOVAR ocean data assimilation system as implemented in the ECMWF ocean analysis for System 4. ECMWF Technical Memorandum, No. 668.

Weaver, A. T. and P. Courtier, 2001: Correlation modelling on the sphere using a generalized diffusion equation. *Q. J. R. Meteorol. Soc*., **127**, 1815-1842.

Weaver, A. T. and I. Mirouze, 2012: On the diffusion equation and its application to isotropic and anisotropic correlation modelling in variational assimilaton. *Q. J. R. Meteorol. Soc., In press.*