SPECIAL PROJECT INTERIM REPORT

Interim Reports should be 2 to 10 pages in length, depending on importance of the project. All the following mandatory information needs to be provided.

Reporting year	2009
Project Title:	Variational data assimilation with the OPA/NEMO OGCM
Computer Project Account:	SPFRVODA
Principal Investigator(s):	Anthony WEAVER Isabelle MIROUZE Maria VALDIVIESO DA COSTA (left in January 2009) Andrea PIACENTINI Thomas PANGAUD (from December 2009)
Affiliation:	CERFACS, Toulouse
Name of ECMWF scientist(s) collaborating to the project (if applicable)	Magdalena BALMASEDA Kristian MOGENSEN
Start date of the project:	2003
Expected end date:	2011

Computer resources allocated/used for the current year and the previous one

(if applicable)

Please answer for all project resources

		Previous year		**Current year**	
		Allocated	Used	Allocated	Used
High Performance Computing Facility	(units)	136000	35100	150000	11655
Data storage capacity	(Gbytes)	1500		2000	

Summary of project objectives

The aim of this project is to further the scientific and technical development of the NEMOVAR system. NEMOVAR is an incremental variational data assimilation system developed specifically for the ocean model component (OPA) of NEMO (Nucleus for European Modelling of the Ocean). The development of NEMOVAR is being pursued in collaboration with ECMWF, the Met Office, and INRIA/LJK. The structure of NEMOVAR is closely related to that of OPAVAR, the variational assimilation system developed at CERFACS for a previous version of OPA. The OPAVAR system has been used for several studies and in the European research projects DEMETER, ENACT and ENSEMBLES. The NEMOVAR system is destined for operational applications at ECMWF and the Met Office. Ocean analyses from NEMOVAR will be used by several groups (including CERFACS) for decadal forecast initialization in the European COMBINE project.

Summary of problems encountered (if any)

Progress has been delayed due to difficulties in finding an appropriate candidate for a postdoctoral position at CERFACS for developing an ensemble method for NEMOVAR. This work will be based on that of Daget *et al*. (2008, 2009) who developed an ensemble method for OPAVAR. The computing resources requested in this project are mainly intended for the ensemble experiments. The postdoctoral position has recently been filled and work on the ensemble NEMOVAR system will start in late 2009.

Summary of results of the current year (from July of previous year to June of current year)

See Appendix.

List of publications/reports from the project with complete references

Cummings, J., Bertino, L., Brasseur, P., Fukumori, I., Kamachi, M., Martin, M., Mogensen, K., Oke, P., Testut, C. E., Verron J. and A. T. Weaver, 2009: Ocean data assimilation systems for GODAE. *Oceanography Magazine*, In press.

Daget, 2008: "Estimation d'ensemble des paramètres des covariances d'erreur d'ébauche dans un système d'assimilation variationnelle de données océaniques". PhD thesis, Université Paul Sabatier, Toulouse, France.

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. Meteorol. Soc*., **135**, 1071–1094.

Daget, N., Weaver, A. T. and M. A. Balamaseda, 2008: An ensemble three-dimensional variational ocean data assimilation for the global ocean: sensitivity to the observation- and background-error formulation. ECMWF Technical Memorandum No. 562. See http://www.ecmwf.int/publications/library/do/references/list/14.

Davey, M. and co-authors, 2005: Enhanced Ocean Data Assimilation and Climate Prediction. Final report for the FP5 ENACT project, Contract No. EVK2-CT2001-00117. Available at http://www.hadobs.org

Mirouze, I. and A. T. Weaver, 2009: Representation of correlation functions using a one-dimensional implicit diffusion equation. In preparation.

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

Ricci, S., Weaver, A. T., Vialard, J. and P. Rogel, 2005: Incorporating state-dependent temperature-salinity relationships in the background error covariance of variational ocean data assimilation. *Mon. Wea. Rev*., **133**, 317–338. Also available as an ECMWF Technical Memorandum No. 441. See http://www.ecmwf.int/publications/library/do/references/list/14.

Tshimanga, J., Gratton, S., Weaver, A. T. and A. Sartenaer, 2008: Limited-memory preconditioners, with application to incremental four-dimensional variational ocean data assimilation. *Q. J. R. Meteorol. Soc*., **134**, 753–771.

Vialard, J., Weaver, A. T., Anderson, D. L. T. and P. Delecluse, 2003: Three- and four-dimensional variational assimilation with an ocean general circulation model of the tropical Pacific Ocean. Part 2: physical validation. *Mon. Wea. Rev*., **131**, 1379–1395.

Weaver, A. T., Deltel, C., Machu, E., Ricci, S. and N. Daget, 2005: A multivariate balance operator for variational ocean data assimilation. *Q. J. R. Meteorol. Soc*. **131**, 3605–3626. Also available as an ECMWF Technical Memorandum No. 491. See http://www.ecmwf.int/publications/library/do/references/list/14.

Weaver, A. T., Vialard, J. and D. L. T. Anderson, 2003: Three- and four-dimensional variational assimilation with an ocean general circulation model of the tropical Pacific Ocean. Part 1: formulation, internal diagnostics and consistency checks. *Mon. Wea. Rev*., **131**, 1360–1378.

Weaver, A. T., Vialard, J., Anderson, D. L. T. and P. Delecluse, 2002: Three- and four-dimensional variational assimilation with an ocean general circulation model of the tropical Pacific Ocean. ECMWF Technical Memorandum No. 365. See http://www.ecmwf.int/publications/library/do/references/list/14.

Summary of plans for the continuation of the project

Work in the next period will focus on the following developments to NEMOVAR:

1. implementation of an ensemble 3D-Var in NEMOVAR;
2. development of improved models for representing correlated background and observation error in NEMOVAR; and
3. development of improved minimization and preconditioning algorithms for NEMOVAR.

Appendix: Summary of results of the current year

Performance and validation of the NEMOVAR system

The NEMOVAR system is based on an incremental formulation and currently supports 3D-Var (FGAT). The NEMOVAR 3D-Var system has reached a fairly mature state over the last year (Mogensen *et al.* 2009). It has a multivariate background-error formulation that includes a local relationship between temperature and salinity obtained by imposing approximate preservation of water masses, an equation of state relating temperature and salinity to density, geostrophic adjustment of horizontal components of velocity, and a local relationship between sea level and vertical density. It is able to assimilate observations from sub-surface profiles of temperature and salinity, along-track sea-level anomaly data from satellite altimeters, and sea-surface temperature data, and employs an online, automatic system for quality control of real-time observations. An on-line model bias correction algorithm has also been implemented. CERFACS has contributed to various aspects of the NEMOVAR system development.

A cycled 3D-Var experiment using a global 1^o version of NEMOVAR has been conducted by ECMWF for the 20-year period 1987-2006. The assimilated data consist of temperature and salinity profiles from EN3v1c. The surface forcing fluxes are derived from ERA40 and include a strong relaxation to SST analyses from the Reynolds OIv2 in the surface heat flux. Results from the experiment are summarized in Fig. 1 which shows the global mean and root-mean-square (rms) of the model fit to the temperature (left panel) and salinity data (right panel). The first year of the experiment has been excluded from the statistics. The control analysis, produced without assimilating data, is too warm and salty compared to observations (thin red curves in both panels in Fig. 1). The assimilation experiment improves the mean and rms fit to the data in both the analysis and the background, the fit for the former being somewhat better as expected. Note that the fit to the data achieved by the 3D-Var analysis is degraded with the IAU procedure, with the model-data fit after IAU lying in between the 3D-Var analysis residual and the background-minus-observation difference. This is an unavoidable property of IAU which achieves temporal smoothness in the analyses at the expense of degrading the fit to the data near the beginning of the assimilation window. Overall, these results from the 3D-Var experiment are encouraging as this was the first multi-year experiment conducted with NEMOVAR.

Results (not shown) from data retention experiments (i.e., experiments in which the model is integrated freely after assimilating data for a certain period) indicate that, on the global average, the assimilated temperature and salinity information is retained for up to 6 months in the thermocline. Near the surface, the information is lost more rapidly, but still retained for 3 to 4 months for temperature and even longer for salinity. These results are encouraging for the initialization of seasonal forecasts.

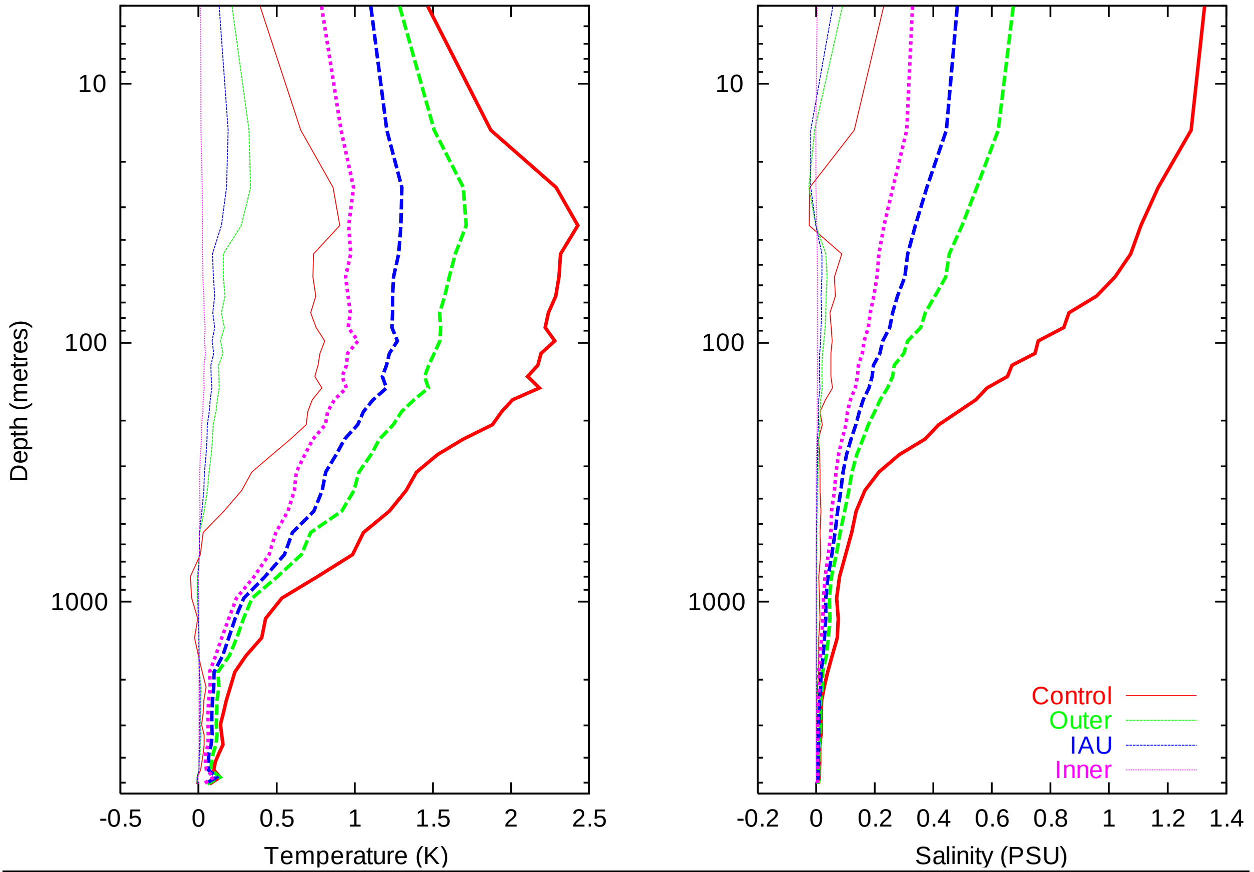


Figure 1: Vertical profiles of the 1988-2006 time-mean (thin curves) and rms (thick curves) of the globally-averaged model-minus-observation differences for temperature (left panel) and salinity (right panel) in a 1^o x 1^o global NEMO configuration. Displayed are the statistics from a control experiment in which no profile data are assimilated (red solid curve) and from a 3D-Var (NEMOVAR) experiment. For the 3D-Var experiment, the statistics are shown for the background-minus-observation (green dashed curve), the analysis-minus-observations after minimization (pink dotted curve), and the analysis-minus-observation after IAU (blue dashed curve).

References:

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

Towards the assimilation of velocity data in NEMOVAR

Velocity is an analysis variable in NEMOVAR, but velocity observations are not yet assimilated. As a first step towards the assimilation of velocity data in NEMOVAR, an observation operator has been developed for measuring differences between the model velocity field and velocity measurements from current-meters and ADCPs from the TAO array. Currentmeter and ADCP data were obtained from the TAO-NOAA website (http://tao.noaa.gov/proj_overview) and combined into a more convenient NetCDF (ENACT-like) data format for using with NEMOVAR (one file for the entire period for both instrument types). About 100,000 daily-averaged velocity measurements are available in the upper 500 metres, with an interval of 2 to 10 metres, from 1980 to the present. Currentmeter data consist of 19 sites across the Tropical Pacific from 1980 to the present. ADCP data consist of 6 sites along the Equator from 1988 to the present (Fig. 2). F90 routines were added to the outer loop of NEMOVAR to read these data.

The velocity observation operator is a 3D interpolation operator similar to the one developed for temperature and salinity, although is technically more complicated since the u and v components of the model velocity field are located at different points on the model C-grid and are aligned with the directions of the curvilinear model grid. Additional F90 routines were developed to write out the data and their model equivalent in a standard feedback file format. The code was successfully tested (on single and multiple processors) in a single cycle (5-day) experiment of a global 1^{o} NEMO configuration. The next step will be to use the velocity observation operators in a multi-cycle control and data assimilation experiment to perform online model-TAO velocity data comparisons. The effective assimilation of these data will require careful tuning of the background-error statistics, particularly for the unbalanced components of the velocity variables.

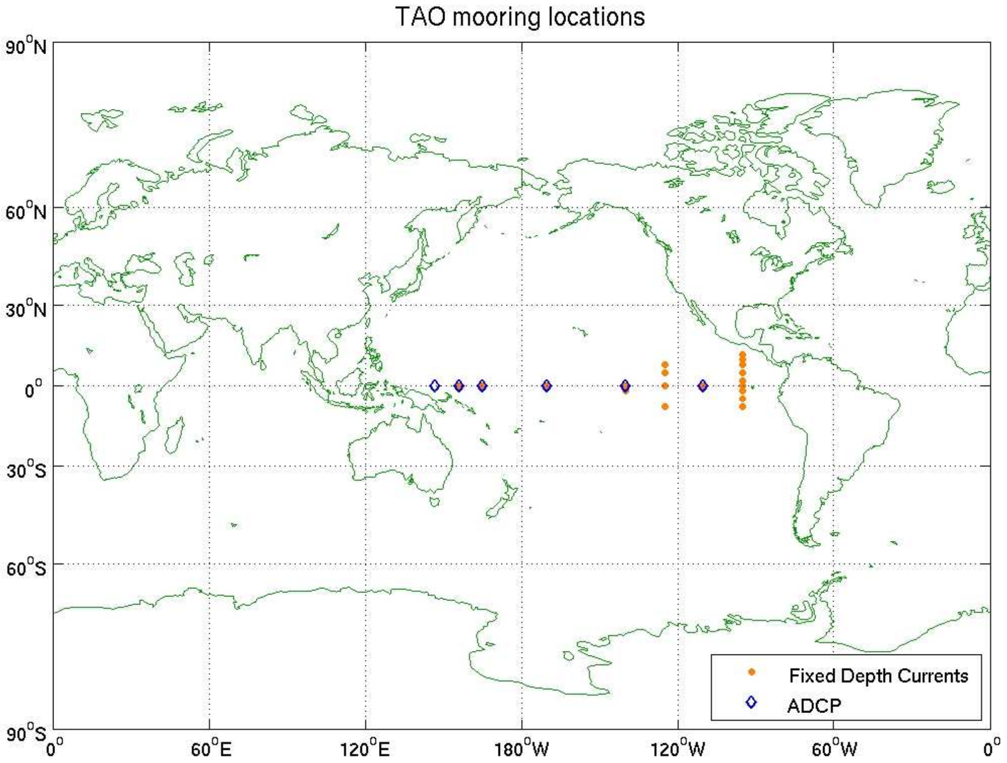
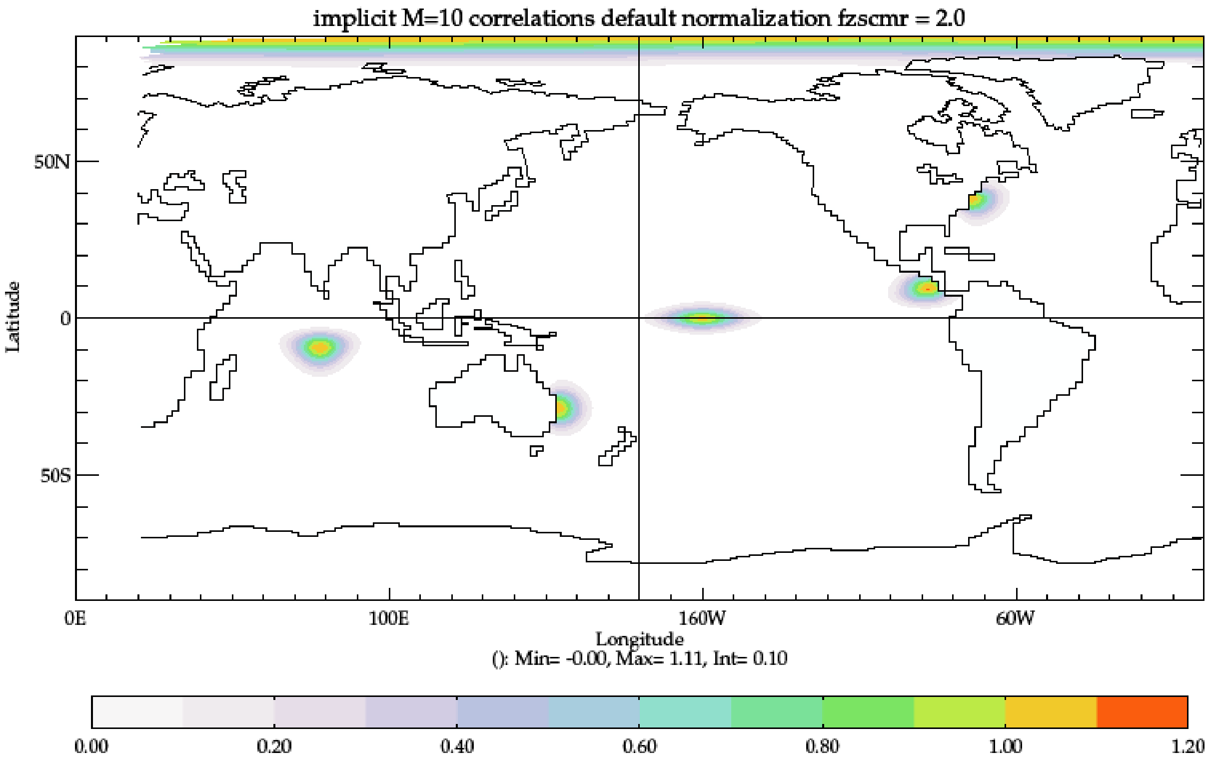


Figure 2: The velocity observation sites on TAO moorings which have been included in NEMOVAR. Data include measurements from fixed-depth currentmeters (orange dots) and high-vertical resolution ADCPs (blue triangles).

Correlation modelling using implicit diffusion operators

The theoretical basis for employing a diffusion equation to represent the action of a correlation operator is described in detail in Weaver and Courtier (2001), Weaver and Ricci (2004) and more recently Mirouze and Weaver (2009). A new three-dimensional correlation operator has been developed for NEMOVAR. It is formulated as a symmetric product of one-dimensional (1D) implicit diffusion operators acting in each of the model's curvilinear coordinate directions. Application of an M-step 1D implicit diffusion operator to a given field is shown to be equivalent to convolving that field with an M-th order auto-regressive function. The new formulation has several advantages over the previous formulation based on an explicit diffusion algorithm, the most important being its ability to reduce significantly the number of diffusion iterations for applications with high resolution models and inhomogeneous length scales. This is important for our future plans which will employ inhomogeneous background-error length scales estimated from analysis ensembles (Daget, 2008). Normalization coefficients are required to compute unit-amplitude (correlation) functions with the diffusion algorithm. An important aspect of the inhomogeneous problem is the estimation of the normalization factors which are not constant. A novel technique has been developed to provide accurate estimates of these factors at a much lower cost than can be achieved using the traditional technique based on randomization.

Figure 3 illustrates the correlations generated by the new 3D implicit diffusion model at selected points in the global domain. Ten iterations have been used for each 1D implicit diffusion operator and the new technique has been used to estimate the normalization factors. The selected points are at the equator, near certain continental boundaries, at 80^{o}E where lateral periodicity is enforced, and near the pole where the tri-polar model grid is severely distorted. No spurious numerical artifacts are produced at these "difficult" points and the amplitudes are close to one as desired.



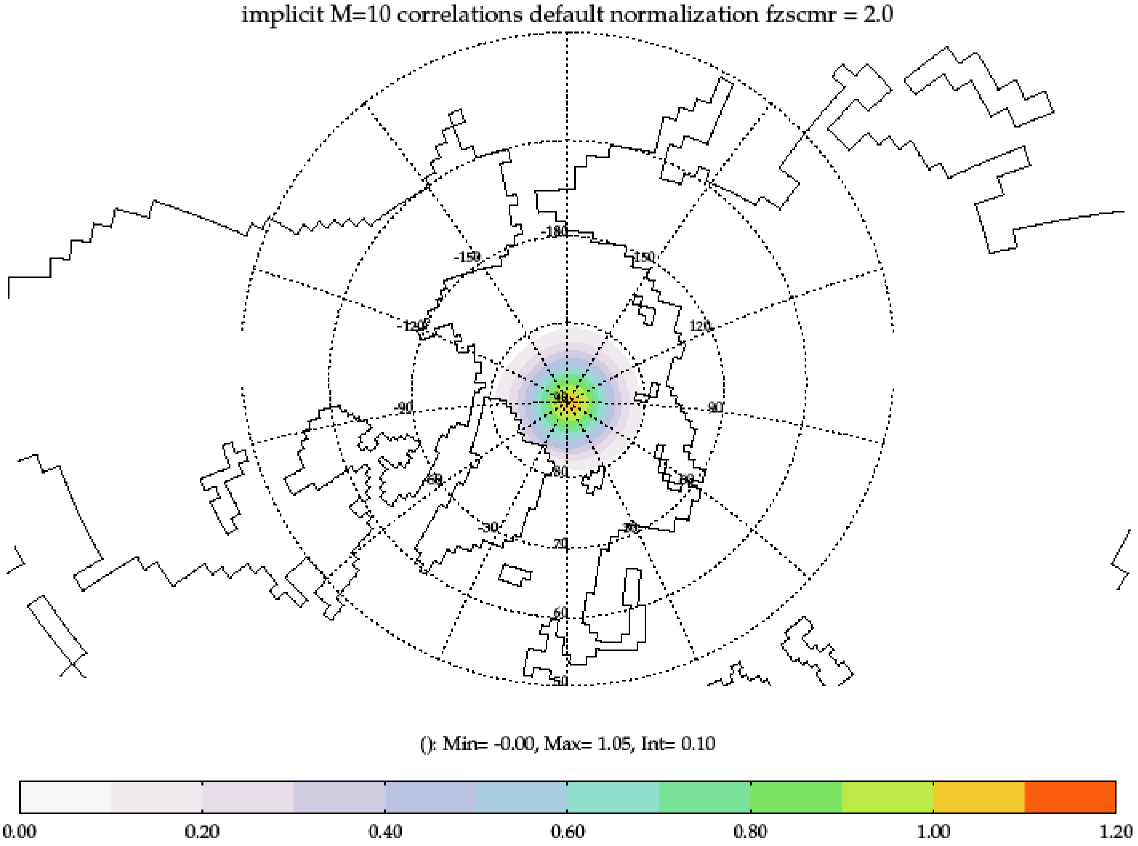


Figure 3: Correlations generated by the 3D implicit diffusion model at selected points in the global domain.

Implicit diffusion operators require the inversion of a matrix. For the 1D correlation operators, this matrix is small enough that it can be inverted using standard direct methods such as Cholesky factorization combined with a forward elimination and backward substitution algorithm. This algorithm has been implemented in NEMOVAR and used for the examples shown in Fig. 3. This method is not convenient, however, for massively parallel processor applications using the domain decomposition approach available with NEMO. Iterative techniques, based on conjugate gradient or Gauss-Seidel (successive-over-relaxation) for example, are better suited for parallel implementations of the implicit diffusion operator and are currently being explored.

References:

Daget, 2008: “Estimation d’ensemble des paramètres des covariances d’erreur d’ébauche dans un système d’assimilation variationnelle de données océaniques”. PhD thesis, Université Paul Sabatier, Toulouse, France.

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.

Mirouze, I. and A. T. Weaver, 2009: Representation of correlation functions using a one-dimensional implicit diffusion equation. *In preparation*.

Weaver, A. T. and S. Ricci, 2004: Constructing a background-error correlation model using generalized diffusion operators. In “Recent developments in data assimilation for atmosphere and ocean”, ECMWF Seminar Proceedings, 8–12 September 2003, pp. 327–340.