4D variational data assimilation for locally nested models: complementary theoretical aspects and application to a 2D shallow water model.

Ehouarn Simon\textsuperscript{1}, Laurent Debreu\textsuperscript{2,4,*} and Eric Blayo\textsuperscript{3,2,4}

\textsuperscript{1} Nansen Environmental and Remote Sensing Center, Bergen, Norway.
\textsuperscript{2} INRIA, MOISE project team, Saint Martin d’Hères, France.
\textsuperscript{3} Joseph Fourier University, Saint Martin d’Hères, France.
\textsuperscript{4} Laboratoire Jean Kuntzmann, Saint Martin d’Hères, France.

SUMMARY

We consider the application of a four dimensional variational data assimilation method to a numerical model which employs local mesh refinement to improve its solution. We focus on structured meshes where a high resolution grid is embedded in a coarser resolution one which covers the entire domain. The formulation of the nested variational data assimilation algorithm was derived in a preliminary work [14]. We are interested here in complementary theoretical aspects. We present first a model for the multigrid background error covariance matrix. Then we propose a variant of our algorithms based on the addition of control variables in the inter-grid transfers in order to allow for a reduction of the errors linked to the interactions between the grids. These formulations are illustrated and discussed in the test case experiment of a 2D shallow water model. Copyright © 2000 John Wiley & Sons, Ltd.

KEY WORDS: Variational Data Assimilation, Nesting, Shallow Water, Structured grid

1. INTRODUCTION

Nested models are commonly used in meteorology and oceanography. Such systems, developed within the framework of structured grids, allow a local increase of the mesh resolution in areas where it seems to be necessary, by running the same model on a hierarchy of grids. In the case of one-way interaction, coarse grid solution provides boundary conditions for the high resolution grid. In the case of two-way interaction, a feedback from the fine grid to the coarse grid is added. These methods have been applied successfully in atmosphere and ocean modelling [34, 28, 24, 17, 12].

\textsuperscript{*}Correspondence to: Laurent Debreu, Laboratoire Jean Kuntzmann, 51 rue des Mathématiques, 38400 Saint Martin d’Hères, France, Tel +33-4-76-51-48-60 Fax +33-4-76-63-12-63, Email: Laurent.Debreu@imag.fr
The development and implementation of efficient data assimilation algorithms in such models are presently becoming important challenges, particularly in the context of operational meteorology and oceanography. Some developments and applications have been performed using a stochastic approach (Kalman filtering) \[4\]. However the problem of variational data assimilation in such systems has almost not yet been studied to our knowledge. This approach is based on the optimal control theory, and aims at minimizing a cost function measuring the gaps between the model solution and the observations and between the model solution and a model first guess. The minimization can be performed using a descent method, the computation of the gradient being performed with an adjoint model (e.g. \[22\]).

In the first part of this study \[14\], we have proposed a formulation of the variational data assimilation for configurations where a high resolution grid is nested in a coarser resolution one which covers the entire domain. The cost function is defined as the sum of the misfits between the model solutions and the observations over both domains. The control variable is composed of the initial conditions on both grids. Both discrete and continuous versions of the derivation of the adjoint of the nested model are presented. Numerical results obtained with a 1D shallow water model are promising: the approach performs well and leads to a fast convergence rate. However the aim of this first part was only to introduce the formulation, and we did not deal with several important features of variational data assimilation. In particular the problem of the multi-grid modelling of the background error covariance matrix was not addressed. This is however a major issue in order to obtain an efficient data assimilation system, due to the important role that plays this matrix in the spreading of the information throughout the state vector. Furthermore the question of the model error and its control has to be investigated in the framework of nested models. The control of the model error has already been studied in a mono-grid framework. This problem appears to be more complex in a multi-grid framework, the non perfect interactions present between the grids leading to complicated interactions between the model error of the nested models. For example the negative effects of the coarse resolution model error on the high resolution model, due to the interactions between the two models, could be important.

The aim of this paper is to suggest a way to deal with the multi-grid modelling of the background error covariance matrix, in order to make possible the use of 4D-Var algorithms in a realistic nested context. Furthermore, we investigate a way to introduce the model error in the nested formulation of the 4D-Var algorithm. In section 2 we recall the formulation of variational data assimilation for locally nested models in both cases of one-way and two-way interactions. We also propose a multi-grid modelling of the background error covariance matrix. We describe our experimental framework in section 3 and we evaluate the nested data assimilation algorithms and the benefits of the multi-grid modelling of the background error covariance matrix in section 4. In section 5, we propose a variant of the nested data assimilation algorithms in which new control variables are added in order to model and reduce the errors linked to the interactions between the grids. Corresponding numerical results are also shown.
2. THE NESTED VARIATIONAL DATA ASSIMILATION FORMULATION

2.1. Formulation of the nested models

We consider the general case of a high resolution model, covering the local domain \( \omega \), embedded in a coarser resolution model covering the larger domain \( \Omega \). With obvious notations, the local high resolution grid and the global coarse resolution grid are denoted respectively \( \omega_h \) and \( \Omega_H \). The corresponding state vectors are denoted respectively \( x_h \) and \( x_H \). We also denote \( \omega_H \) the part of the grid \( \Omega_H \) corresponding to the local domain \( \omega \) (see Figure 1).

For both one-way and two way interactions, we use the same notation \( F \) for the coarse and fine models in order to simplify the notations. It can of course be different, at least at the discrete level.

In the case of one-way interaction, the coarse grid model provides boundary conditions to the high resolution model using an interpolation operator \( I_h^H \). Semi-discretized equations of the nested system can be written as follows:

\[
\begin{align*}
\frac{\partial x_H}{\partial t} &= F(x_H) \quad \text{on} \quad \Omega_H \times [0,T] \\
x_H(t=0) &= x_H^0 \\
x_{\partial \omega} &= I_h^H(x_H) \quad \text{on} \quad \partial \omega_H \times [0,T]
\end{align*}
\]

where \( x_{\partial \omega} \) represents the information coming from the coarse grid onto \( \partial \omega_H \), the boundary of the fine grid. The one-way interaction is said passive since there is no retroaction from the local model onto the global model. From a practical point of view, this also means that both models do not have to be run simultaneously (the global model can be run first, and its solution can then be used off line by the local model).

In the case of two-way interaction, a feedback term from the fine grid onto the coarse grid is added. The coarse solution is updated locally (in \( \hat{\omega}_H \), the interior of \( \omega_H \)) by the high resolution solution using a restriction operator \( G_h^H \). Semi-discretized equations of the nested system can be written as follows:

\[
\begin{align*}
\frac{\partial x_H}{\partial t} &= F(x_H, x_\omega) \quad \text{on} \quad \Omega_H \times [0,T] \\
x_H(t=0) &= x_H^0 \\
x_\omega &= G_h^H(x_h) \quad \text{on} \quad \hat{\omega}_H \times [0,T] \\
x_{\partial \omega} &= I_h^H(x_H) \quad \text{on} \quad \partial \omega_H \times [0,T]
\end{align*}
\]

where \( x_\omega \) represents the information coming from the coarse grid onto the fine grid in \( \hat{\omega}_H \). The two-way interaction is said active. In that case both models must be run simultaneously since they permanently exchange information.

After discretization, the problems have to be integrated in time in a specific order. The model is first integrated on the coarse grid \( \Omega_H \) and then on the high resolution \( \omega_h \) grid with
boundary conditions given by a spatial and temporal interpolation of the coarse values. Finally a feedback can be applied. An example of this integration order with a time refinement factor of 2 is shown on figure (2).

[Figure 2 about here.]

For more details about the specific design of the two-way retroaction in the coarse resolution model, we refer to [13].

2.2. Variational data assimilation

In order to obtain a formulation of variational data assimilation in the specific context of a two-grid system, we have first to re-define several fundamental notions.

**Observations and observation operator** The observations used at time $t_i$ on $\Omega_H$ and $\omega_h$ are denoted respectively $y_i^H$ and $y_i^h$. These observations belong to the same observation space. The subscripts $H$ and $h$ highlight on which grid the observations are used to compute the cost function. Discrete observation operators and observation error covariance matrices are defined on the coarse grid ($H_i^H$ and $R_i^H$) and on the fine grid ($H_i^h$ and $R_i^h$) in the same way as in the context of a single grid. Note that an observation located in $\omega$ can possibly be assimilated on both grids.

The question of the optimal distribution of the observations on the grids is important in the context of multiresolution systems. However it is out of the scope of the present paper, and is not addressed in the following.

**State vector** The state vector $x$ of the two-grid model is composed of the state vectors of both grids:

$$x = \begin{bmatrix} x_H \\ x_h \end{bmatrix}$$

It corresponds to the "multigrid state vector" proposed in [4] for the application of a SEEK filter in a two-way nested model of the Ligurian Sea. Therefore the initial condition $x^0$, which will be the control variable for the minimization problem, and its background value $x^b$ (i.e. the first guess at the beginning of the optimization process) read with obvious notations:

$$x^0 = \begin{bmatrix} x_0^H \\ x_0^h \end{bmatrix} \quad x^b = \begin{bmatrix} x_b^H \\ x_b^h \end{bmatrix}$$

The dimension of such vectors can be large, as it is the sum of the dimensions of the state vectors of the different grids. This size can be typically 2 or 3 times the size of the coarsest initial condition in actual applications. It can be a practical issue when using high resolution GCM-type models. For the specific case of the 2D shallow water model used in the present paper (sections 3 to 5), the dimension of vectors $x^0$ and $x^b$ is equal to 27300.
Cost function

The cost function is defined as the sum of the misfit to the first guess and the misfits to the observations on both grids:

\[
J(x^0) = J_b(x^0) + J_{obs}^H(x^0) + J_{obs}^h(x^0)
\]

\[
= \frac{1}{2}(x^0 - x^b)^T B^{-1}(x^0 - x^b)
\]

\[+ \frac{1}{2} \sum_{i=0}^{N} \left[H_i^H(x^0) - y_i^H \right]^T R_i^{-1} \left[H_i^H(x^0) - y_i^H \right]
\]

\[+ \frac{1}{2} \sum_{i=0}^{N} \left[H_i^h(x^0) - y_i^h \right]^T R_i^{-1} \left[H_i^h(x^0) - y_i^h \right]
\]

An important aspect is the expression of the background error covariance matrix \(B\). Its modelling is described in §2.4.

2.3. The multigrid optimality systems

An usual way to optimize the cost function \(J\) is to make use of descent gradient methods. Such methods require to assess the gradient of the cost function. The gradient of the background term \(J_b\) reads \(\nabla J_b(x^0) = B^{-1}(x^0 - x^b)\). The gradients of the observation terms \(J_{obs}^H\) and \(J_{obs}^h\) can be obtained using the adjoint method (e.g. [22]) applied to the nested system.

The derivation of the adjoint model (semi-discretized and discretized versions) for the cases of one-way and two-way interactions is detailed in the first part of this study [14] using, as usual, a \(L_2\) inner-product. We note \(P\) the coarse resolution component of the adjoint vector (\(P\) is the adjoint variable of \(x_H\)) and \(Q\) the high resolution one (\(Q\) is the adjoint variable of \(x_h\)).

In the case of one-way interaction, the optimality system reads:
One-way optimality system

\[
\begin{align*}
\text{Domain } & \Omega_H \quad \left\{ \begin{array}{ll}
\frac{\partial x_H}{\partial t} &= F(x_H) \quad \text{on } \Omega_H \times [0, T] \\
x_H(x,0) &= x_H^0
\end{array} \right. \\
\frac{\partial P}{\partial t} + \left[ \frac{\partial F}{\partial x_H} \right]^* \cdot P + I_H^* \left[ \frac{\partial F}{\partial x_{\partial \omega}} \right]^* \cdot Q &= H_H^{-1} (H_H x_H(t) - y_H(t)) \\
P(T) &= 0 \\
\nabla x_H^0 J_{\text{obs}} &= -P(0)
\end{align*}
\]

\[
\begin{align*}
\text{Domain } & \omega_h \quad \left\{ \begin{array}{ll}
\frac{\partial x_h}{\partial t} &= F(x_h, x_{\partial \omega}) \quad \text{on } \omega_h \times [0, T] \\
x_h(x,0) &= x_h^0 \\
x_{\partial \omega} &= I_H^h(x_H) \quad \text{on } \partial \omega_h \times [0, T]
\end{array} \right. \\
\frac{\partial Q}{\partial t} + \left[ \frac{\partial F}{\partial x_h} \right]^* \cdot Q &= H_h^{-1} (H_h x_h(t) - y_h(t)) \\
Q(T) &= 0 \\
\nabla x_h^0 J_{\text{obs}} &= -Q(0)
\end{align*}
\]

where the operator \( I_h^H = I_h^H(x_H) \) is the linearization of \( I_h^H \) in \( x_H \) and the operator \( I_H^h \) is the adjoint of \( I_h^H \). We observe a feedback from the high resolution adjoint model onto the coarse resolution one, in the opposite direction of the interactions existing in the direct formulation. Due to this interaction, the high resolution adjoint model must be integrated first. However, at each iteration of the minimization process, the integration of the fine resolution direct model (resp. coarse resolution adjoint model) can still be performed offline, once the coarse resolution direct model (resp. fine resolution adjoint model) has been run.

In the case of two-way interaction, the adjoint system reads:
Two-way optimality system

\[ \begin{align*}
\frac{\partial \mathbf{x}_H}{\partial t} &= F(\mathbf{x}_H, \mathbf{x}_\omega) \quad \text{on } \Omega_H \times [0,T] \\
\mathbf{x}_H(x,0) &= \mathbf{x}_0^H \\
\mathbf{x}_\omega &= G_h^H(\mathbf{x}_h) \quad \text{on } \omega_H \times [0,T]
\end{align*} \]

Domain \( \Omega_H \)

\[ \begin{align*}
\frac{\partial \mathbf{P}}{\partial t} + \left[ \frac{\partial F}{\partial \mathbf{x}_H} \right]^* \cdot \mathbf{P} + \mathbf{H}_h^H \left[ \frac{\partial F}{\partial \mathbf{x}_\omega} \right]^* \cdot \mathbf{Q} &= \mathbf{H}_h^H \mathbf{R}_h^{-1} (\mathbf{H}_H \mathbf{x}_H(t) - \mathbf{y}_H(t)) \\
\nabla_{x_h}^H J_{obs} &= -\mathbf{P}(0)
\end{align*} \]

Domain \( \omega_h \)

\[ \begin{align*}
\frac{\partial \mathbf{x}_h}{\partial t} &= F(\mathbf{x}_h, \mathbf{x}_\partial \omega) \quad \text{on } \omega_h \times [0,T] \\
\mathbf{x}_h(x,0) &= \mathbf{x}_0^h \\
\mathbf{x}_\partial \omega &= I_H^h(\mathbf{x}_H) \quad \text{on } \partial \omega_h \times [0,T]
\end{align*} \]

\[ \begin{align*}
\frac{\partial \mathbf{Q}}{\partial t} + \left[ \frac{\partial F}{\partial \mathbf{x}_h} \right]^* \cdot \mathbf{Q} + \mathbf{G}_h^H \left[ \frac{\partial F}{\partial \mathbf{x}_\omega} \right]^* \cdot \mathbf{P} &= \mathbf{H}_h^H \mathbf{R}_h^{-1} (\mathbf{H}_h \mathbf{x}_h(t) - \mathbf{y}_h(t)) \\
\nabla_{x_h}^H J_{obs} &= -\mathbf{Q}(0)
\end{align*} \]

where the operator \( \mathbf{G}_h^H = \mathbf{G}_h^H(\mathbf{x}_h) \) is the linearization of \( \mathbf{G}_h^H \) in \( \mathbf{x}_h \) and the operator \( \mathbf{G}_h^H \) is the adjoint of \( \mathbf{G}_h^H \). Once again, the interactions between the grids in the adjoint formulation are in the opposite direction of the interactions existing in the direct formulation. The addition of a feedback from the high resolution solution to the coarse resolution one leads to an interaction in the adjoint model from the coarse resolution adjoint solution to the high resolution one. Like for the direct system, the two grids must be integrated simultaneously in this adjoint system.

2.4. Multigrid background error covariance matrix

The background error covariance matrix is a key point of any data assimilation system, because of its role in spreading the information brought by the data throughout the entire state vector. Numerous approaches exist, based on more or less empirical modelling and on statistical considerations, that lead to the construction of such matrices for single grid applications (e.g. [25, 16, 32, 26, 33, 23]). Given these approaches, the question here is to provide a way to build a background error covariance matrix for a multigrid system.

The simplest way to construct such a matrix is to assume that the background errors on different grids are independent. One then obtains a block diagonal matrix, each block being a single grid background error covariance matrix obtained by an appropriate single grid algorithm, for example algorithm based on the resolution of a diffusion equation [5, 32].

However the provision of better correlations between the corrections to the initial conditions on both grids requires a full modelling (non-diagonal blocks) of the background error covariance matrix.
matrix. We suggest to address the multigrid character of the $B$ matrix by reproducing the structure of interactions between the grids present in two-way nested models. Since the derivation of such a matrix is rather technical, we only give the main results in the following, the calculations being detailed in Appendix.

Let assume that background vectors are available on each grid, denoted by $\hat{x}^b_H$ and $\hat{x}^b_h$. Starting from these vectors, we can define a two-grid background vector $x^b = (x^b_H, x^b_h)^T$, consistent with the two-way interactions, as follows:

- $x^b_H$ is the restriction of $\hat{x}^b_h$ in the interior of $\omega_H$ and is equal to $\hat{x}^b_H$ elsewhere.
- $x^b_h$ is the interpolation of $\hat{x}^b_h$ on the boundary of $\omega_h$ and is equal to $\hat{x}^b_h$ in the interior.

Let $x^t$ be the true value of the solution. We can then define:

- $\hat{e}_h^H = \hat{x}_h^b - x^t_H$ the “initial” background error on the coarse grid
- $\hat{e}_h^H = \hat{x}_h^b - x^t_h$ the “initial” background error on the fine grid
- $e^G = G^H_h (x^b_H) - x^t_H$, defined in $\omega_H$, the error due to the restriction operator $G^H_h$
- $e^I = I^H_h (x^b_H) - x^t_h$, defined on $\partial\omega_h$, the error due to the interpolation operator $I^H_h$

We assume now that, for each following couple, the two random variables are independent: $(e^G; \hat{e}_h^H) , (e^G; \hat{e}_h^H), (e^I; \hat{e}_h^H)$ and $(e^I; \hat{e}_h^H)$. This simply means that the choice of the restriction operator $G^H_h$ and of the interpolation operator $I^H_h$ is independent of the choice of the initial first guess on each grid. This assumption seems reasonable and appears in a sense close to the assumption of independence between the model error and the background error usually done in data assimilation.

Furthermore we assume (see Appendix) that $\hat{e}_h^H$ and $\hat{e}_h^H$ are independent, which means that $\hat{x}_h^b$ and $\hat{x}_h^b$ were obtained independently. This hypothesis is more questionable, and probably not satisfied in many practical cases where the first guesses are obtained after an integration of a nested model. It can thus lead to some loss of information. However, even with this hypothesis, this approach allows a real spreading of the information on the entire multigrid vector in agreement with the interactions which are present in the two-way nested models. It is therefore already an improvement in comparison with the simplest modelling of the background covariance error matrix, namely the block diagonal matrix. Future work will probably have to study a way to relax this hypothesis.

Given these hypotheses we demonstrate that $B$ can be approximated by

$$B = SS^T + Q$$

where the exact expressions of $S$ and $Q$ are given in Appendix. $SS^T$ is a two-grid background error covariance matrix, depending solely of tilted background errors and their images by the interpolation/restriction operators. $Q$ models the covariances corresponding to the errors $e^I$ and $e^G$ of the interpolation/restriction operators.

We do not claim that such modelling is the only way, neither the best way, to couple initial errors between grids. Nevertheless we expect that the modelling based on two-way interaction will improve the spreading of the high resolution information issued from the backward integration of the adjoint model on the coarse resolution correction for both one-way and two-way nested models. Even if this modelling appears to be better adapted for two-way interaction, we expect also benefits in one-way interaction by obtaining boundary conditions from the coarse model in agreement with the dynamics of the high resolution solution.
3. DESCRIPTION OF THE EXPERIMENTAL FRAMEWORK

3.1. The 2-D shallow water model

We consider here a reduced gravity 2-D shallow water system in a flat bottom rectangular
domain $\Omega$. Using a cartesian coordinate system, $\Omega$ reads $\Omega = [0, L_x] \times [y_0 - \frac{L_y}{2}, y_0 + \frac{L_y}{2}]$, where $y_0$ is the value of $y$ at the center of the basin. The governing equations are:

$$\begin{align*}
\frac{\partial \phi}{\partial t} + \frac{\partial (\phi u)}{\partial x} + \frac{\partial (\phi v)}{\partial y} &= 0 \\
\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} + v \frac{\partial u}{\partial y} - f v + g \frac{\partial \phi}{\partial x} + ru - \nu \Delta u &= \frac{\tau_x}{\rho_0} \\
\frac{\partial v}{\partial t} + u \frac{\partial v}{\partial x} + v \frac{\partial v}{\partial y} + f u + g \frac{\partial \phi}{\partial y} + rv - \nu \Delta v &= \frac{\tau_y}{\rho_0}
\end{align*}$$

where $\phi$ denotes the water height and $(u, v)$ is the horizontal velocity. Since we use a reduced
gravity model, $\phi$ must be interpreted as the depth of an idealized model of the upper ocean
layer. The system is closed with noslip boundary conditions: $u = v = 0$ on $\partial \Omega$. $g$ is a
reduced gravity and $\rho_0$ is the water density. $f$ is the Coriolis parameter, given by the $\beta$-plane
approximation: $f(y) \approx f(y_0) + \beta(y - y_0)$, with $\beta = \frac{\partial f}{\partial y}(y_0)$. $\nu$ and $r$ are the diffusivity and
bottom friction coefficients respectively. $\tau = (\tau_x, \tau_y)$ is the wind stress. It is chosen constant
in time, as follows: $\tau_x = \tau_0 \cos \left( \frac{2 \pi (y - y_0)}{L_y} \right)$ and $\tau_y = 0$, which leads to a classical double gyre
circulation (see figure 3).

Numerical values of the parameters are chosen as follows: $L_x = L_y = 2000 km$, $f(y_0) = 0.7 \times 10^{-4}s^{-1}$, $\beta = 2 \times 10^{-11} m^{-1}s^{-1}$, $r = 10^{-7}s^{-1}$, $\rho_0 = 1000 kg m^{-3}$, $g = 0.02 ms^{-2}$, $\tau_0 = 0.015 Nm^{-2}$. These values correspond to those used in the simulations reported in [2, 30].

The numerical value for $\nu$ will depend on the horizontal resolution (see Table I).

This model is discretized on an Arakawa C grid [1], the water height $\phi$ being located at the
center of the cells and the velocity components at the center of the cell edges. A third order
predictor-corrector scheme is used. It corresponds to the scheme used for the barotropic flow
in the numerical model ROMS-AGRIF developed by the Center for Earth Systems Research
(CESR) at UCLA and the French Institut de Recherche pour le Développement (IRD). More
information on this scheme can be found in [27]. Furthermore this model includes the mesh
refinement library AGRIF† (Adaptive Grid Refinement in Fortran) [15].

3.2. Configuration of the numerical experiments

High resolution reference experiment A reference high resolution simulation was run on a grid
$\Omega_h$ with a resolution of $\frac{25}{3} km$ in the two horizontal directions. The time step is equal to 1800s.

†http://www-ljk.imag.fr/MOISE/AGRIF/index.html
Starting from rest \( \phi = \phi_0 = 500 \, m, u = v = 0 \, m.s^{-1} \), a five-year spin-up was performed. At the end of this simulation, the system has reached a stabilized level of energy and the nested system can be set up (see below). Then an additional seven months run was performed, which trajectory will be denoted \( x^{ref} \), and will be considered as the true state in the following data assimilation experiments. These latter experiments, which are detailed below, correspond to the seventh month of \( x^{ref} \). The temporal unfoldment of our experiments is summarized on figure 4.

The nested system. A nested system was set up, with both one-way and two-way interactions. A coarse grid, with 25 km horizontal resolution, covers the entire domain. A local fine grid, with \( \frac{25}{3} \, km \) horizontal resolution, is embedded in the coarse grid in the region of the central jet (see figure 3). The diffusivity coefficients are the same for the local high resolution simulation (on \( \omega_h \)) and the reference simulation (on \( \Omega_h \)). The diffusivity coefficient is 9 times smaller for the fine grid than for the coarse grid, which makes the high resolution dynamics quite different from the coarse resolution one. The coarse model is not able to generate meso-scale eddies, and diffuses the existing ones if initially present, while the local high resolution model has the same dynamics as the reference model. The values of the resolution dependent parameters are summarized in Table I.

In order to avoid numerical instabilities which may appear due to a large discontinuity in dissipation, we introduce a transition layer (5 grid points wide) near the boundaries on the fine grid. In this layer, the diffusion coefficient varies progressively from its fine resolution value to its coarse resolution value. This transition layer differs from a sponge layer in the fact that the increase in diffusivity is only guided by the ratio between the coarse and fine grid resolution, and not by any need of damping the dynamics (see [8]).

The coarse-to-fine operator \( I^h_H \) is a simple linear interpolation, while the fine-to-coarse operator \( G^H_h \) is a linear "Full Weighting" restriction (see [29]). This operator uses a larger stencil than an average operator (5 grid points rather than 3 for a mesh refinement factor of 3) but has better spectral properties. Furthermore, it is proportional to the adjoint of \( I^h_H \), which is desirable property in multigrid systems. In order to avoid additional interpolation errors, we decided to run both the coarse model and the fine model with the same time step of 1800 s, as for the reference simulation.

A Flather condition (1976) is used on \( \partial \omega \) for the one-way interaction, in order to improve the behavior of the open boundaries. This condition imposes the value of the incoming characteristic of the shallow water equations. See [7] for details on this open boundary condition. For the case of two-way interaction, the Flather condition is not useful longer, and may even lead to numerical instabilities, particularly near boundaries mostly outflowing, the coarse resolution solution being strongly dependent of the local high resolution solution. That’s why Dirichlet (or "clamped") conditions are used. The feedback from the high resolution solution to the coarse resolution one consists simply in replacing the coarse resolution solution at grid points located in \( \omega_H \) by a local average of the high resolution solution (\( x_\omega \)). Potential improvements can probably be obtained by adding flux correction procedures in order to ensure the balance of mass and tracers through the boundary, as recommended in [13].
Data assimilation experiments

Four data assimilation experiments were performed. They differ by the nesting algorithm (one-way vs two-way) and by the structure of the background error covariance matrix (see §3.3). The length of the assimilation window is one month and corresponds, as mentioned previously, to the seventh month of \( x^{\text{ref}} \). All the experiments share the following features:

Observations

Observations of the layer thickness \( \phi \) only are available, and only on the fine grid \( \omega_h \). They are obtained by sampling spatially and temporally the water height component of the reference true state \( x^{\text{ref}} \): one observation field per day with a spatial resolution of \( (125 \text{ km})^2 \) (every five grid points in each direction). We do not add noise to the reference true state to obtain the observations. Observation errors are supposed to be spatially uncorrelated, with a standard deviation of \( 1 \text{ m} \) (which represents a relative error of the order of 1\%). Therefore \( R_i = \text{Id} \) at each observation time \( t_i \). So we assimilate 30 observation fields over an assimilation window of one month. Not that experiments with noised data (e.g. with observation errors around 10\% of the solution) were also performed (not shown), and confirmed the results presented in the following.

This observation network can be seen as some idealized satellite data.

Background state

The background state \( x^b \) is obtained in the following way. At the end of the 5 years spin up simulation, the nested system is set up: the initial condition of the coarse model is obtained by a restriction (with a Full Weighted scheme) of the global high resolution solution \( x^{\text{ref}} \). Then an interpolation of this coarse initial condition gives a local high resolution initial condition. As the composition of a Full Weighting restriction and a linear interpolation is almost an unity operator, the initial condition of the local high resolution model is close to the reference solution. Then a six-month simulation of this nested system in two-way interaction is performed, which solution at the final time will be \( x^b \). The differences between the reference and background solutions are mostly a result of the six-month nested simulation, and not a result of important differences in initial conditions. The maps of the background state (figure 3 a-) and the reference initial condition (figure 3 b-) reveal strong differences between these two solutions.

Cost function

Using the usual 4D-Var preconditionning proposed in [10] \( x^0 - x^b = B^{1/2} v_0 \), the cost function is defined in the preconditionned space as follows

\[
J(v^h_0, v^H_0) = \frac{1}{2} \| v^h_0 \|_2^2 + \frac{1}{2} \| v^H_0 \|_2^2 + \sum_{i=0}^{N} (H_i[x^i_h(v^h_0, v^H_0)] - y_i)^T R_i^{-1}(H_i[x^i_h(v^h_0, v^H_0)] - y_i)
\]

(5)

3.3. Background error covariance matrix

Modelling \( B \)

We saw in §2.4 how the background error covariance matrix can be approximated in the context of a two-grid system. In order to illustrate the impact of the non-diagonal blocks of \( B \), which correspond to the inter-grid covariances, let us consider two versions of \( B \), with and without these blocks.
Block diagonal $B$  The first guesses on both grids are assumed to be independent. This leads to the block diagonal matrix:

$$B_{\text{indep}} = \begin{bmatrix} B_H & 0 \\ 0 & B_h \end{bmatrix}$$

This matrix is clearly a simplistic approximation of the true $B$. However its implementation is relatively easy. We do not build the entire matrix $B_{\text{indep}}$, but the operator realizing the product of its square root with any vector $v$ (which allows for the use of the preconditioning mentioned previously $x^0 - x_b = B_{\text{indep}}^{1/2} v_0$). $B_{\text{indep}}^{1/2}$ reads

$$B_{\text{indep}}^{1/2} = \begin{bmatrix} B_H^{1/2} & 0 \\ 0 & B_h^{1/2} \end{bmatrix}$$

On each grid, the square root of the multivariate error covariance matrix $B_{\alpha}^{1/2}$ ($\alpha \in \{ h, H \}$) takes the form:

$$B_{\alpha}^{1/2} = K_{\alpha} \Sigma_{\alpha} C_{\alpha}^{1/2}$$

(6)

$C_{\alpha}^{1/2}$ is the square root of a Gaussian-shaped monovariate covariance operator, obtained by the resolution of a generalized diffusion equation as proposed by [32]. This operator includes a normalization matrix to ensure that diagonal terms are equal to one. $\Sigma_{\alpha}$ is a diagonal matrix of background-error standard deviations. $\Sigma_{\alpha} C_{\alpha}^{1/2}$ is therefore the square root of a monovariate background error covariance matrix. Then a balance operator [16] $K_{\alpha}$, based on the geostrophic equilibrium, is applied to get a multivariate error covariance matrix on each grid.

At each grid point, we use the same spatio-temporal variances in the modelling of $\Sigma_{\alpha}$. The values of the variances have been estimated using a one-month model simulation and are summarized in Table II. These values and their spatial distributions are of course not adapted for realistic applications, but are sufficient in the present idealized context.

[Table 2 about here.]

“Full” $B$  Following the approach developed in §2.4 and Appendix, a more complex approximation of $B$ can be built, denoted $B_{\text{multi}}$. We have $B_{\text{multi}} = SS^T + Q$ where $S$ is obtained from $B_{\text{indep}}^{1/2}$ and the interpolation/restriction operators. In the present case, we suppose that the error covariances of the interpolation/restriction operators (matrix $Q$) are negligible compared to the multigrid background error covariances (matrix $SS^T$), which leads to

$$B_{\text{multi}}^{1/2} \approx S = \begin{bmatrix} \Psi_{\omega_n} B_H^{1/2} \\ \Psi_{\omega_n} G_{\omega_n}^H B_h^{1/2} \end{bmatrix}$$

where the exact definitions of the indicator functions $\Psi$, $\bar{\Psi}$ and $\Phi$ are given in Appendix.
Illustration of \( \mathbf{B} \)  

In order to visualize the structure of the background error covariance and to better understand the information transfers between the grids, we have performed numerical experiments assimilating one single observation available at the end of the integration window. For such experiments, the analysis increment \( \mathbf{x}_a - \mathbf{x}_b \) corresponds, by ignoring the second order and higher derivatives, to a column of the multigrid matrix \( \mathbf{BM}^T \), with \( \mathbf{M}^T \) the multigrid adjoint of the nested model. Even if it does not exactly visualize \( \mathbf{B} \), such an experiment is frequently performed in data assimilation studies because it gives some insight into the structure of \( \mathbf{B} \) and into the way the information is spread by the adjoint model.

The single observation of the water height was set at the center of the fine grid. Its location is denoted with a red cross on figure 5 (resp. figure 6) representing the maps of the high (resp. coarse) resolution analysis increments. This experiment was performed with the two different matrices \( \mathbf{B}_{\text{indep}} \) and \( \mathbf{B}_{\text{multi}} \), both for one-way and two-way interaction. For each case, starting from the background described in §3.2, the integration of the model led to a misfit about 33 m between the high resolution solution and the single observation.

A maximum number of iterations was introduced in the minimization process. The four analysis increments shown on figures 5 and 6 correspond to the optimized correction given at the end of the fortieth iteration. In the four cases, the cost functions were reduced by a factor 6000 during the minimization, and the error to the single observation was oscillating during the last 30 iterations between values from \( 5 \times 10^{-4} \) m to 0.3 m.

[Figure 5 about here.]

[Figure 6 about here.]

On the fine grid, the differences in the structure of the analysis increment are quite weak when comparing the four cases (figure 5). The corrections are very similar for the case of one-way interaction with \( \mathbf{B}_{\text{multi}} \) (figure 5 c-) and for the cases of two-way interaction with both modelling of the covariance matrices (figures 5 b- and d-). The correction is mainly localized on the western boundary, where the jet enters the domain. The case of one-way interaction with \( \mathbf{B}_{\text{indep}} \) (figure 5 a-) is slightly different: the amplitude of this western correction is weaker and the correction in the interior of the domain is slightly stronger than the ones for the three other cases. The high resolution correction highlights the signature of the adjoint model. For that specific configuration of an eastward jet, the integration backward in time by the adjoint model of the misfit between the high resolution solution and the single observation leads for the initial condition to a significant correction of the position of the jet close to the western boundary. Finally the signature on the high resolution increment of the non diagonal blocks in \( \mathbf{B}_{\text{multi}} \) is more significant in one-way interaction than in two-way interaction.

On the contrary, the differences between the corrections on the coarse grid are very significant (figure 6). The amplitude of the correction for both one-way and two-way interactions, which is of the order of 5 m for the water height and 1 cm s\(^{-1}\) for the velocity with the matrix \( \mathbf{B}_{\text{indep}} \) (figures 6 a- and b-), is increased by a factor of almost 10 when using the matrix \( \mathbf{B}_{\text{multi}} \) (figures 6 c- and d-). Moreover, the spatial structure of the correction changes also drastically. The addition of the intergrid covariances in \( \mathbf{B}_{\text{multi}} \) leads to a much better consistency between the corrections on both grids, the restriction to \( \omega_H \) of the correction on the coarse grid (figures 6 c- and d-) being roughly the same as the correction on the fine grid.

Copyright © 2000 John Wiley & Sons, Ltd.  

Int. J. Numer. Meth. Fluids 2000; 00:1–6

Prepared using fldauth.cls
We note that the use of $B_{\text{multi}}$ in one-way interaction (figure 6 c- and d-) leads to an analysis increment very close to the two-way analysis increment (figure 6 d-).

The use of the matrix $B_{\text{indep}}$ leads to a slight correction, mostly located just west of $\omega_H$ both for one-way (figure 6 a-) and two-way interactions (figure 6 b-). It results in a correction of the western boundary condition on the high resolution grid. Moreover, unlike what we noted with the use of the matrix $B_{\text{multi}}$, strong differences can be observed between the one-way and two-way interactions (figures 6 a- and b-). In the latter the correction is weaker, because the update procedure (starting at the first time step) is able to transfer, at least partly, on the coarse grid the influence of the correction identified on the fine grid.

The signature of the modelling of the non-diagonal block in $B_{\text{multi}}$ appears clearly in the coarse resolution analysis increment and may dominate the one of the adjoint in the coarse resolution correction.

4. PERFORMANCES OF THE NESTED DATA ASSIMILATION ALGORITHMS

We now compare in this section the results of data assimilation experiments performed in the configuration described in §3.2. Like in the preceding section, four experiments have been performed: with one-way or two-way interaction, and using the matrix $B_{\text{indep}}$ or the matrix $B_{\text{multi}}$. As it was done in the single observation experiment, the number of iterations in the optimization is limited to 40.

4.1. Cost function

There is no significant difference, depending on whether one uses the matrix $B_{\text{multi}}$ or the matrix $B_{\text{indep}}$, in the evolution of the cost function defined in Eq. (5) during 40 iterations of the minimization (not shown). In the four cases, the cost functions were divided by a factor being between 1600 and 2000. In such a configuration the use of the matrix $B_{\text{indep}}$ allows a very efficient minimization of the cost function, and so a reduction of the misfit between the high resolution solution and the observations. That is why the potential improvements of the two grid solution due to the use of the matrix $B_{\text{multi}}$ seem to have a weak impact in this diagnostic.

4.2. Error on the fine grid $\omega_h$

The normalized root mean square (RMS) error on the high resolution grid $\omega_h$ globally compares the solution of the high resolution grid and the true state. It is defined separately for the layer thickness $\phi$ and the velocity $u$ by:

$$\text{RMS}_h = \left( \frac{\frac{1}{T} \int_T^0 \|x_h(t) - \hat{x}_{e,f}(t)\|_2^2 dt}{\frac{1}{T} \int_T^0 \|\hat{x}_{e,f}(t)\|_2^2 dt} \right)^{1/2}$$

where $\| \cdot \|_2$ denotes the Euclidian norm on $\omega_h$ and $x_h$ denotes either $\phi$ or $u$.

We can observe (figure 7 a-) that all four experiments lead to a significant reduction of the error on the water height component of the high resolution solution, which is the observed variable. After 40 minimization iterations, the error level varies from 4% with the use of the matrix $B_{\text{multi}}$ in two-way interaction to 4.7% with the use of the matrix $B_{\text{indep}}$ in one-way
interaction. The differences on the errors on the velocity (the unobserved variable) are more significant (figure 7 b-). The two experiments with two-way interaction lead to improved results with regard to the two experiments with one-way interaction.

When looking at the evolution with time of the instantaneous RMS error of the solutions obtained at iteration 40 (called optimized solutions in the following):

\[
\text{RMS}_h(t) = \left( \frac{1}{T} \int_0^T \frac{1}{2} \left( \frac{1}{2} ||x_{\text{ref}}(\tau)||_{\Omega_H,2}^2 \right) d\tau \right)^{1/2}
\]

the differences between the \(\phi\) component of the high resolution solutions are highlighted (figure 8 a-). First, the use of two-way interaction leads to better optimized solutions over a larger part of the time interval. This illustrates the impact and the benefits of the addition of interactions between the grids in both direct and adjoint models in the two-way nested 4D-Var formulation. Nevertheless one notes that the two-way optimized solutions present more errors than one-way optimized solutions at the beginning of the assimilation window (first 100 hours). Secondly, in the case of one-way interaction, the use of the matrix \(B_{\text{multi}}\) clearly improves the solution compared to the use of the matrix \(B_{\text{indep}}\). Even if the errors of the high resolution solutions are similar at the initial step, the differences between the coarse resolution initial conditions lead to differences in the boundary conditions of the high resolution model and thus a significant divergence of RMS errors during the second half of the month. This result is coherent with those that we noted in the "single observation" experiments: this matrix compensates for a part of the weakness of the one-way interaction.

However this is no longer true in the case of two-way interaction for which the high resolution initial condition has a strong influence on the coarse model dynamics due to the regular feedback from the high resolution solution to the coarse resolution one. In our configuration the coarse resolution solution is indeed updated by the high resolution solution at each coarse resolution time step. Therefore a solution very close to the initial condition on the high resolution model is transferred on the coarse grid at the end of the first coarse resolution time step, thus limiting the impact of the lack of intergrid information in the formulation of the matrix \(B_{\text{indep}}\). For configurations where the update of the coarse grid would be less frequent, we can assume however that the matrix \(B_{\text{multi}}\) would be more influential.

Concerning the velocity (figure 8 b-), the use of the matrix \(B_{\text{multi}}\) appears mainly at the end of the assimilation window in the case of one-way interaction. Once again this is no longer true in the case of two-way interaction.

4.3. Error on the coarse grid \(\Omega_H\)

The performances of the data assimilation methods on the coarse grid \(\Omega_H\) are evaluated by defining the global RMS error on this grid by:

\[
\text{RMS}_H = \left( \frac{1}{T} \int_0^T \left( \frac{1}{2} \left( \frac{1}{2} ||x_{H}(t) - \Pi_{\text{ref}}^H(x_{\text{ref}}(t))||_{\Omega_H,2}^2 \right) dt \right)^{1/2}
\]

Copyright © 2000 John Wiley & Sons, Ltd.

Int. J. Numer. Meth. Fluids 2000; 00:1–6

Prepared using fidauth.cls
where $\Pi^H_h$ denotes the canonical injection from $\Omega_h$ onto $\Omega_H$. Defining the coarse resolution true state with $\Pi^H_h$ rather than $G^H_h$ leads to take into account the non-perfectness of the restriction operator $G^H_h$ in our idealized framework (the restriction error $\epsilon^G$ is non null).

This diagnostic is plotted on figure 9. We notice first that the algorithms are able to decrease the errors on the coarse grid without assimilating any observation on this grid. However the decrease is less important than on the fine grid (we typically obtain a reduction of 10% to 20% of the relative error).

For one-way interaction, the use of the matrix $B_{\text{multi}}$ allows a significant improvement of the coarse solution, compared to the use of the matrix $B_{\text{indep}}$, which highlights the interest of the non-diagonal blocks in the matrix $B_{\text{multi}}$. For the two-way interaction, as already observed on the fine grid, the behavior of the errors are identical for the two matrices $B_{\text{multi}}$ and $B_{\text{indep}}$, due to the feedback term in the algorithm. The final level of the error on the water height (figure 9 a-) and velocity (figure 9 b-) components are slightly greater than for the one-way case. The update of the coarse resolution solution by the high resolution solution present in two-way interaction is simple in this testcase. It corresponds only to a restriction of the high resolution solution. It leads to instabilities in the vicinity of the western boundary (inflow boundary) due to inconsistencies between the coarse resolution which has been updated in $\omega_H$ and the coarse resolution which has not been updated outside $\omega_H$. Improvements could be obtain by including a mass conservation correction on the coarse resolution grid. As one-way interaction does not present a feedback from the high resolution solution to the coarse resolution one, we do not observe numerical artifacts.

Finally, we are interested in the spatial differences between the optimal coarse resolution solutions. We define the average of the difference between the true state and the optimized solution by:

$$\text{DIFF} = \frac{1}{N} \sum_{i=1}^{N} (x_H - \Pi^H_h(x_{\text{ref}}))(t_i)$$

where $(t_i)_{i=1,N}$ corresponds to a temporal sample of the assimilation window. The maps of this diagnostic applied to the water height are plotted on figure 10. Sample are extracted every 50 hours to compute the temporal average ($N = 14$). In the case of two-way interaction, these results are obtained with the matrix $B_{\text{multi}}$. Quite identical results are obtained with the matrix $B_{\text{indep}}$.

We notice first that the use of the matrix $B_{\text{multi}}$, in the case of one-way interaction (figure 10 d-), leads to an important reduction of the error in the area $\omega_H$ (where the mesh refinement takes place), compared to the use of the matrix $B_{\text{indep}}$ (figure 10 c-). Even if the patterns of the errors appear to be similar right on the eastern boundary between the two one-way solutions, these one present strong differences (not shown). The solution obtained with the matrix $B_{\text{multi}}$ represents fairly well the structures (the two gyres and the first meander of the jet) present in the true state, while the solution obtained with the matrix $B_{\text{indep}}$ let appear an erroneous important eddy already detached form the northern gyre. Furthermore the amplitude of the
errors of the one-way solution obtained with the matrix $B_{\text{multi}}$ is close to those of the two-way solution (figure 10 b-) in the subdomain $\omega_H$. Nevertheless the use of two-way interaction, due to the feedback from the high resolution solution onto the coarse resolution solution, leads to a better solution inside and downstream from the refined area.

Finally we notice in all cases that important errors remain upstream from $\omega_H$ (in the subdomain left to the western boundary). Even if the use of an adjoint model allows an impact of the circulation in this area, the non-eddy resolving properties of the model do not allow the introduction in the coarse resolution solution of the eddies present in the true state (figure 10 a-). It explains the errors of all the optimized coarse resolution solutions in the South part of this area.

5. CONTROL OF THE INTER-GRID INTERACTION ERRORS

5.1. The inter-grid interaction errors

The formulation introduced in the previous section implicitly assumes that the models on both grids are perfect, as it is generally the case in strong-constraint 4D-Var data assimilation (except in some particular approaches like 4D-PSAS [10], the representer method [9] or like methods that explicitly control some model error [31, 20, 3]). However this assumption is of course not satisfied in actual simulations. Moreover, in the framework of nested systems, the error of the coarse model can be particularly large, and is propagated into the fine resolution model through the specification of its boundary conditions. Furthermore the operators of interpolation are not perfect and are potential sources of errors in the high resolution model. These interpolation errors can be defined as the difference on the boundary between an hypothetic global high resolution solution and the interpolation of the coarse resolution solution (in a similar way of the error $\epsilon_I$ defined in §2.4). They include the high scales that are not resolved on the coarse resolution grid. We propose here to model and control these errors by adding a new control term on $\partial\omega$. We expect by this way to be able to control the effects on the fine model of both the interpolation error and the coarse model error. In the case of two-way interaction, we can similarly add a control term on $\Omega_H$ for the restriction errors, in order to control the effects on the coarse model of both the fine model error and the restriction error.

In the general case of two-way interaction, we obtain the new formulation:

\[
\begin{align*}
\text{Domain } \Omega_H: & \quad \frac{\partial \mathbf{x}_H}{\partial t} = F_H(x_H, x_\omega) \\
x_H(x, 0) &= x'_H \\
x_\omega &= G_H(x_h) + \epsilon_\omega
\end{align*}
\]

\[
\begin{align*}
\text{Domain } \omega_h: & \quad \frac{\partial \mathbf{x}_h}{\partial t} = F_h(x_h, x_{\partial \omega}) \\
x_h(x, 0) &= x'_h \\
x_{\partial \omega} &= I_H(x_H) + \epsilon_{\partial \omega}
\end{align*}
\]

These inter-grid interaction errors will be denoted in the following by the vector $\epsilon^{ig} = \begin{bmatrix} \epsilon_\omega \\ \epsilon_{\partial \omega} \end{bmatrix}$.

In this formulation, the variable $\epsilon_\omega$ is controlled at each time step of the coarse resolution grid and the variable $\epsilon_{\partial \omega}$ is controlled at each high resolution time step. It means that the size of the vector $\epsilon^{ig}$ can be very important in the case of long assimilation time windows or for very small time steps on both grids. This can be problematic for the application of this algorithm.
in such situations. Nevertheless an easy way to reduce the size of $\varepsilon^g$ if necessary is to control the inter-grid interaction errors only $k$ time steps where $k > 1$.

Unfortunately almost no prior information about these errors is available, since these terms model the combination of several error sources which cannot be accurately modeled.

5.2. A new data assimilation problem

Given this new formulation of the nested system, we can define a corresponding new data assimilation problem in which we control not only the initial condition on both grids but also the inter-grid interaction error. The new control vector is therefore $\begin{bmatrix} x^0 \\ \varepsilon^g \end{bmatrix}$. In the following, we will denote wic algorithms (for "with interaction control") the algorithms using this extended control variable.

In order to regularize the problem, we add to the cost function defined in the previous formulation (§2.2) a new term $J_{\varepsilon^g}(\varepsilon^g) = \frac{1}{2}\|K\varepsilon^g\|^2$ which appears as a weak constraint $K$ on $\varepsilon^g$. A traditional way is to add a digital filter to the cost function [18]. Nevertheless the specific case of nested model requires to prevent the appearance of discontinuities between the grids due to the introduction of $\varepsilon^g$. The expression of $K$ will be discussed later.

We then obtain the following optimization problem:

$$\min_{(x^0, \varepsilon^g)} J(x^0, \varepsilon^g) = J_b(x^0) + J_{\varepsilon^g}(\varepsilon^g) + J_{\text{obs}}^H(x^0, \varepsilon^g) + J_{\text{obs}}^h(x^0, \varepsilon^g)$$

$$= \frac{1}{2}(x^0 - x^b)^T B^{-1} (x^0 - x^b) + \frac{1}{2}\|K\varepsilon^g\|^2$$

$$+ \frac{1}{2} \sum_{i=0}^{N} [H^i_H(x^i_H(x^0, \varepsilon^g)) - y^i_H]^{T} R^{-1}_H [H^i_H(x^i_H(x^0, \varepsilon^g)) - y^i_H]$$

$$+ \frac{1}{2} \sum_{i=0}^{N} [H^i_h(x^i_h(x^0, \varepsilon^g)) - y^i_h]^{T} R^{-1}_h [H^i_h(x^i_h(x^0, \varepsilon^g)) - y^i_h]$$

The adjoint models are identical to the ones defined in the previous formulation (§2.3). For that reason, the software developments required for adding this new control variable are relatively light.

In the case of two-way interaction, the gradients of the term measuring the misfit to the observations on both grids are given by the following system:

For the domain $\Omega_H$:

$$\nabla_{x^0_H} J_{\text{obs}} = -P(0)$$

$$\nabla_{\varepsilon^g_H} J_{\text{obs}} = - \left[ \frac{\partial F_H}{\partial x^0} \right]^* \cdot P$$

For the domain $\omega_h$:

$$\nabla_{x^0_h} J_{\text{obs}} = -Q(0)$$

$$\nabla_{\varepsilon^g_h} J_{\text{obs}} = - \left[ \frac{\partial F_h}{\partial x^0} \right]^* \cdot Q$$

where $P$ and $Q$ are solutions of the two-way nested adjoint model defined in §2.3.

Note that, formally, the components of the gradients related to $\varepsilon_\omega$ and $\varepsilon_{\partial\omega}$ have the same expression as gradients related to a control of the boundary conditions.

5.3. Numerical experiments

In the following numerical experiments we compare the performances of the wic algorithms in the one-way and two-way cases with the corresponding nested algorithms with no additional
control terms. The configuration of the experiments is the same as in the previous section. The background error covariance matrix is $B_{\text{multi}}$, which construction was detailed in §3.2. The same experiments realized with the matrix $B_{\text{indep}}$ led to similar conclusions.

Choice of the regularization term $K$ The introduction of the inter-grid interaction errors $\epsilon^{ig}$ to model and control the effects on a given grid of both the errors due to the non-perfectness of the inter-grid operators (restriction/interpolation) and the model error of the other grid can lead to a weakening of the interactions between the grids, the control of $\epsilon^{ig}$ substituting to the control of the initial conditions. The weak constraint $K$ has to be defined in order to prevent this potential substitution and the consequent generation of discontinuities between the grids. A natural idea is to constraint the component of the inter-grid interaction errors on a given grid to be invisible for the other grid.

Such a filtering can be obtained using a scale decomposition which is well-known in multigrid theory (e.g. [29]). Let $F_{\Omega_H}$ and $F_{\Omega_h}$ be the spaces of grid functions defined on $\Omega_H$ and $\Omega_h$ respectively. Let $A^h_H$ be a linear operator from $F_{\Omega_H}$ into $F_{\Omega_h}$ (interpolation). Therefore the following relation stands:

$$F_{\Omega_h} = \text{Ker}(A^H_h) \oplus \text{Im}(A^h_H)$$

where $A^H_h$ is the adjoint of $A^h_H$. This means that any vector $x_h$ of the high resolution grid $\Omega_h$ can be decomposed in a unique way under the form $x_h = u_h + A^H_h v_H$ with $u_h \in \text{Ker}(A^H_h)$ and $v_H$ a vector of the coarse resolution grid $\Omega_H$. By definition $u_h$ is not “visible” on the coarse grid (it is composed only of high frequencies), while $A^H_h v_H$ contains only low frequencies.

Therefore the properties we want for $\epsilon_\omega$ and $\epsilon_{\partial\omega}$ can be obtained by imposing in a weak sense that they lie in the kernels of the adjoints of the inter-grid transfer operators, i.e. by minimizing

$$J^{\epsilon^{ig}}(\epsilon^{ig}) = \|G^h_H \epsilon_\omega\|^2 + \|I^H_h \epsilon_{\partial\omega}\|^2$$

where $G^H_h$ and $I^H_h$ are respectively the adjoints of the restriction and the interpolation operators used in the nested system. This corresponds, in the case of two-way interaction, to defining $K$ as:

$$K = \begin{bmatrix} G^h_H & 0 \\ 0 & I^H_h \end{bmatrix}$$

The adjoint of the Full Weighting restriction operator being an injective interpolation, one has $\text{Ker}(G^H_h) = 0$. One imposes in a weak sense that the corrections of the inter-grid interaction errors in the coarse resolution grid are null, the only control term which is not visible from the high resolution grid. $I^H_h$ being an injective interpolation, $\text{Ker}(I^H_h) \neq 0$ and there is at least one non null vector which is not visible on the coarse resolution grid.

This choice for $K$ is not perfect however. It leads to some problems of non physical corrections due to the lack of physics in the definition of the operator $K$. A solution to improve the correction of the interaction errors could be to use balance operators between the model variables, but defining such operators for variables defined only on the boundaries is challenging. This is still an open issue at the present time.

Cost function The evolution of the term $J^{obs}_h$ (misfit between the high resolution solution and the observations) of the cost function $J$ during the 40 iterations of the minimization is plotted on figure 11. We normalized $J^{obs}_h$ by its initial value (value obtained with the free run) for all the cases. One observes a real impact of the addition of the control of the inter-grid
transfer errors. The minimization of $J_{obs}$ is more significant for the wic algorithms (factor 2 in one-way interaction, factor 5 in two-way interaction). The addition of the control of inter-grid interaction errors leads to optimized high resolution solutions (solutions obtained after 40 iterations) which are closer to the observations.

[Figure 11 about here.]

Finally the combination of two-way interaction and control of the inter-grid interaction errors leads after 40 iterations to a more significant decrease of the cost function (not shown).

Error on the fine grid $\omega_h$. The evolution of the high resolution normalized errors during the 40 iterations of the minimization is indicated on figure 12.

[Figure 12 about here.]

Contrary to what we have noted for the minimization of the cost function, the correction brought by the wic algorithms of the error on both the observed layer thickness and unobserved velocity variables is less significant at the end of the minimization than the nested algorithm with the only control of the initial condition. Nevertheless the differences are very small: less than 1 percentage point on the layer thickness and about 5 percentage points on the velocity. These results are confirmed when looking to the evolution on the assimilation window of the RMS error of the optimized solution (figure 13). So the solutions achieved from the algorithm without the control of the inter-grid transfer errors are quite better.

[Figure 13 about here.]

We think that this is due to the fact that the layer thickness (what we observe) inside the domain is strongly sensitive to the variations of the velocity on the boundaries, as it is suggested by [19]. The inter-grid transfer errors control makes it possible to bring potentially non physically balanced corrections (due to the lack of balance operator on the control variable) during the model integration. The misfits to the observations (and so the cost function) could be reduced without improving the solution in the area (in time and space) where there is no observation. So the performances of the wic algorithms in term of error correction could be improved by adding relevant balance operators on the inter-grid transfer error control.

Furthermore in the case of two-way interaction, we note figure 13 an acceleration of the decrease of the RMS error (mainly on the velocity, figure 13 b-) during the first iterations of the minimization with the wic formulation. This property can be very valuable given the cost of variational data assimilation algorithms. Indeed, in the context of operational data assimilation or in the perspective of an incremental approach, it will be essential to reduce the number of iterations of the minimization.
Error on the coarse grid $\Omega_H$ The evolution of the coarse resolution normalized errors during the 40 iterations of the minimization is indicated on figure 14.

First we note that the RMS error of the coarse resolution solution decreases during the minimization for both one-way and two-way interactions when adding the control of the inter-grid interaction errors. Nevertheless, as noted on the high resolution grid, the optimized solution achieved with the \textit{wic} algorithms presents a larger error than the solution achieved by controlling only the multigrid initial condition. This is mainly due to the configuration of the experiments: observations are present only on $\omega_h$. By adding degrees of freedom in the data assimilation system, the algorithm does not have to strongly constrain the coarse resolution solution in order to minimize the misfits of observations localized only on the fine grid. The control of the inter-grid interaction errors tends to substitute to the control of the initial condition of the coarse resolution grid. In preliminary experiments (not shown) with observations on both grids, the addition of the inter-grid transfer error control allowed a slightly more significant decrease of the coarse resolution error. Such results highlight some potential of the \textit{wic} algorithms. Nevertheless the confirmation of these results requires investigations of the question of the choice of the appropriate density of observations as a function of the grid resolution.

6. SUMMARY

The aim of this study was to analyze the 4D variational data assimilation in the context of locally nested models. Starting from the formulation of the 4D-Var algorithm that we introduced previously in [14], this study focused mainly on two points: the introduction and the modelling of the multigrid background error covariance matrix $B$, and the improvement of the nesting method by introducing and controlling what we have called the inter-grid interaction errors.

Numerical experiments in a 2D shallow water model with a two-grid configuration have confirmed the good results that we obtained previously with this approach in a simpler case ([14]). Furthermore it is important to mention that the results obtained on the high resolution grid were always better than those obtained with a classical 4D-Var algorithm with control of initial and boundary conditions in the local high resolution model with open boundaries. These results are not shown here, since the present study does not focus on this topic, but they clearly bring an additional value to the nested system approach.

In the first part of this study, we have suggested a multigrid modelling of the matrix $B$ that takes into account the interactions between the grids. In the case of one-way interaction, this modelling allows a real improvement of the coarse and fine resolution optimal solutions in comparison with solutions obtained with a simple block diagonal matrix. For the case of two-way interaction, this modelling leads only to a slight improvement of the solutions. We noted that the algorithms in both types of interactions were not able to improve the coarse resolution solution upstream from the refinement area. Since the assumption of independence of the "initial" background errors is not satisfied in practice, the data assimilation system needs more information on the correlation between the errors on both grids. A possible solution to
better catch the flow dependencies between the grids could be to apply the reduced order strategy suggested in [26] to the space of the multigrid state vector.

In the second part of this study, we have introduced a new control term in the formulation of the inter-grid interactions in order to reduce the effects on a given model of both the model error of the other model and the inter-grid transfer operator error. This new formulation allows a more significant decrease of the misfit between the high resolution solution and the observations \( J_{h}^{obs} \) in comparison to the formulation without this additional control term. However the corrections brought by the \( wic \) algorithms on the RMS errors are less important at the end of the minimization. We think that it is mainly due to the lack of balance operators on the inter-grid transfer errors, allowing non physically balanced corrections. The definition of such operators for variables defined only on the boundaries is still an open issue and has to be investigated.

The problem of the relevance of the observation density as a function of the grid resolution was out of the scope of this paper. However, an observation which is relevant at some given scale can not be relevant any more at another scale, and can even deteriorate the performance of the data assimilation system [21]. This problem will have to be addressed before implementing data assimilation in operational multigrid systems. The concept of adaptive observations [6, 11] could be an interesting tool for tackling this problem.

ACKNOWLEDGEMENT

This research was supported by a DGA fellowship. The authors thank the three anonymous referees as well as Dr Arthur Vidard for their relevant and constructive comments.
APPENDIX

Background error covariance matrix

We develop here the proof of the approximation of the matrix $B$ by equation (3):

$$B = SS^T + Q$$

The exact expressions of $S$ and $Q$ are given in the following.

Let $x^b = (x_H^b, x_h^b)^T$ be the two-grid background vector, with $x_H^b$ and $x_h^b$ the background vectors on both grids. The background error $\epsilon^b$ is defined as:

$$\epsilon^b = \begin{bmatrix} \epsilon_H^b \\ \epsilon_h^b \end{bmatrix} = \begin{bmatrix} x_H^b - x_H^t \\ x_h^b - x_h^t \end{bmatrix}$$

where $x_H^t$ and $x_h^t$ are the true states on the coarse grid and the fine grid. Let assume that one has at his disposal background vectors on each grid, built independently. We note them $\tilde{x}_H^b$ and $\tilde{x}_h^b$.

Inter-grid transfer operators

We itemize here several operators playing a role in the definition of the inter-grid transfers.

**Coarse grid**

Let $\Psi_{\omega_H}$ a bounded linear operator from $F_{\omega_H}$ to $F_{\Omega_H}$ (indicator function of $\Omega_H \setminus \hat{\omega}_H$):

$$\Psi_{\omega_H} : F_{\omega_H} \to F_{\Omega_H}$$

$$x_H \mapsto \begin{cases} x_H & \text{in } \Omega_H \setminus \hat{\omega}_H \\ 0 & \text{in } \hat{\omega}_H \end{cases}$$

(7)

Let $\bar{\Psi}_{\omega_H}$ a bounded linear operator from $F_{\omega_H}$ to $F_{\Omega_H}$ (indicator function of $\hat{\omega}_H$ and extension by zero):

$$\bar{\Psi}_{\omega_H} : F_{\omega_H} \to F_{\Omega_H}$$

$$x_H \mapsto \begin{cases} 0 & \text{in } \Omega_H \setminus \hat{\omega}_H \\ x_H & \text{in } \hat{\omega}_H \end{cases}$$

(8)

Let $\Phi_{\omega_H}$ the restriction of $x_H$ to the interior of $\omega_H$:

$$\Phi_{\omega_H} : F_{\Omega_H} \to F_{\omega_H}$$

$$x_H \mapsto \begin{cases} 0 & \text{in } \partial \omega_H \\ x_H & \text{in } \hat{\omega}_H \end{cases}$$

(9)

One has the relation:

$$\Psi_{\omega_H} + \bar{\Psi}_{\omega_H} \Phi_{\omega_H} = I_{F_{\Omega_H}}$$

(10)

with $I_{F_{\Omega_H}}$ the identity operator on $F_{\Omega_H}$.

**Fine grid**

Let $\Psi_{\omega_h}$ a bounded linear operator from $F_{\omega_h}$ to $F_{\omega_h}$ (indicator function of $\omega_h$):

$$\Psi_{\omega_h} : F_{\omega_h} \to F_{\omega_h}$$

$$x_h \mapsto \begin{cases} x_h & \text{in } \omega_h \\ 0 & \text{in } \partial \omega_h \end{cases}$$

(11)

Let $\bar{\Psi}_{\omega_h}$ a bounded linear operator from $F_{\partial \omega_h}$ to $F_{\omega_h}$ (injection):

$$\bar{\Psi}_{\omega_h} : F_{\partial \omega_h} \to F_{\omega_h}$$

$$x_h \mapsto \begin{cases} 0 & \text{in } \omega_h \\ x_h & \text{on } \partial \omega_h \end{cases}$$

(12)
Let $\Phi_{\partial\omega_H}$ the restriction of $x_H$ to the boundary of $\omega_H$:

$$\Phi_{\partial\omega_H} : F_{\omega_H} \mapsto F_{\partial\omega_H}, \quad x_H \mapsto x_H$$

(13)

Building of the background vector $x^b$

Coarse grid: We define the coarse background vector as the restriction of $\tilde{x}^b_h$ to the interior of $\omega_H$ and $\tilde{x}^b_H$ elsewhere:

$$x^b_H = \Psi_{\omega_H} \tilde{x}^b_H + \bar{\Psi}_{\omega_H} G^H_h(\tilde{x}^b_h)$$

Fine grid: We define the fine background term as the interpolation of $\tilde{x}^b_H$ on the boundaries of $\omega_h$ and $\tilde{x}^b_h$ in the interior:

$$x^b_h = \Psi_{\omega_h} \tilde{x}^b_h + \bar{\Psi}_{\omega_h} I^b_H(\Phi_{\partial\omega_h} \tilde{x}^b_H)$$

The background error $\epsilon^b$

Coarse grid

The background error on the coarse grid $\epsilon^b_H$ becomes:

$$\epsilon^b_H = x^b_H - x^b_H$$

$$\epsilon^b_H = \Psi_{\omega_H} \tilde{x}^b_H + \bar{\Psi}_{\omega_H} G^H_h(\tilde{x}^b_h) - x^b_H$$

(14)

The last line was obtained using equation 10. The first term $\bar{\epsilon}^b_H = \tilde{x}^b_H - x^b_H$ corresponds to the initial background error on the coarse grid. One can define $\epsilon^C_H = G^H_h(x^b_h) - \Phi_{\partial\omega_h} \tilde{x}^b_H$ the error due to the restriction operator $G^H_h$. Furthermore a linearization of $G^H_h(x^b_h)$ in the vicinity of $\tilde{x}^b_h$ gives:

$$G^H_h(x^b_h) = G^H_h(\tilde{x}^b_h) - G^H_h \epsilon^b_h + \alpha_H$$

where $G^H_h = G^H_h(\tilde{x}^b_h)$.

Finally, one obtains:

$$\epsilon^b_H = \Psi_{\omega_H} \bar{\epsilon}^b_H + \bar{\Psi}_{\omega_H} G^H_h \epsilon^b_h + \bar{\Psi}_{\omega_H} \epsilon^C_H + \alpha_H$$

with $\|\alpha_H\| = o(\|\epsilon^b_h\|)$

Fine grid

The background error on the fine grid $\epsilon^b_h$ becomes:

$$\epsilon^b_h = x^b_h - x^b_h$$

$$\epsilon^b_h = \Psi_{\omega_h} \tilde{x}^b_h + \bar{\Psi}_{\omega_h} I^b_H(\Phi_{\partial\omega_h} \tilde{x}^b_H) - x^b_h$$

(15)

The first term $\bar{\epsilon}^b_h = \tilde{x}^b_h - x^b_h$ corresponds to the initial background error on the fine grid. One can define $\epsilon^I = I^b_H(\Phi_{\partial\omega_h} \tilde{x}^b_H) - x^b_h$ the error due to the interpolation operator $I^b_H$. Furthermore a linearization of $I^b_H(\Phi_{\partial\omega_h} \tilde{x}^b_H)$ in the vicinity of $\Phi_{\partial\omega_h} \tilde{x}^b_H$ gives:

$$I^b_H(\Phi_{\partial\omega_h} x^b_H) = I^b_H(\Phi_{\partial\omega_h} \tilde{x}^b_H) - I^b_H(\Phi_{\partial\omega_h} x^b_h) + \beta_h$$

where $I^b_H = I^b_H(\Phi_{\partial\omega_h} \tilde{x}^b_H)$.

Finally one obtains:

$$\epsilon^b_h = \Psi_{\omega_h} \bar{\epsilon}^b_h + \bar{\Psi}_{\omega_h} I^b_H(\Phi_{\partial\omega_h} \tilde{x}^b_H) + \bar{\Psi}_{\omega_h} \epsilon^I + \beta_h$$

with $\|\beta_h\| = o(||\epsilon^b_H||)$

Copyright © 2000 John Wiley & Sons, Ltd.

Int. J. Numer. Meth. Fluids 2000; 00:1–6

Prepared using fldauth.cls
Approximation of $B$  
Let assume that $\alpha_h$ and $\beta_h$ are small (the initial first guesses are good on each grid). We assume now that, for each following couple, the two random variables are independent: $(\epsilon^G; \epsilon^H_h)$, $(\epsilon^G; \epsilon^h)$, $(\epsilon^I; \epsilon^H_h)$ and $(\epsilon^I; \epsilon^h)$. We also assume that $\tilde{\epsilon}_H$ and $\tilde{\epsilon}_h$ are independent. We refer to §2.4 for more information about these assumptions. Since the covariance operator $\text{Cov}(\ldots)$ is bilinear, one obtains an approximation of $B$: $B \approx B^0$, with

$$B^0 = \begin{bmatrix} B^0_{HH} & B^0_{Hh} \\ B^0_{Hh} & B^0_{hh} \end{bmatrix}$$  \hfill (16)

and

$$\begin{aligned}
B^0_{HH} &= \Psi_{\omega_h} \tilde{B}_{HH} \Psi_{\omega_h}^T + \Psi_{\omega_h} G_h^H B_{h_h} (G_h^H)^T \Psi_{\omega_h}^T + \Psi_{\omega_h} B_{h_h} \Psi_{\omega_h}^T \\
B^0_{Hh} &= \Psi_{\omega_h} \tilde{B}_{HH} \Phi_{\omega_h} (I_h^H)^T \Psi_{\omega_h} + \Psi_{\omega_h} G_h^H B_{h_h} \Psi_{\omega_h}^T + \Psi_{\omega_h} \text{Cov}(\epsilon^G, \epsilon^H) \Psi_{\omega_h}^T \\
B^0_{hh} &= \Psi_{\omega_h} \tilde{B}_{hh} \Psi_{\omega_h} + \Psi_{\omega_h} \Phi_{\omega_h} B_{h_h} \Phi_{\omega_h}^T (I_h^H)^T \Psi_{\omega_h} + \Psi_{\omega_h} \text{Cov}(\epsilon^G, \epsilon^I) \Psi_{\omega_h}^T
\end{aligned}$$  \hfill (17)

$B_{HH}$ and $B_{hh}$ are symmetric positive, so they admit some Choleski decomposition:

$$\exists \ (S_H, S_h), \quad \tilde{B}_{HH} = S_H S_H^T \quad \text{and} \quad \tilde{B}_{hh} = S_h S_h^T$$

Then, if one defines

$$S = \begin{bmatrix} \Psi_{\omega_h} S_H \\ \Psi_{\omega_h} \Phi_{\omega_h} S_H \end{bmatrix} \quad \text{and} \quad Q = \begin{bmatrix} \Psi_{\omega_h} B_{h_h} \Phi_{\omega_h} \Psi_{\omega_h}^T \\ \Psi_{\omega_h} \Phi_{\omega_h} \text{Cov}(\epsilon^G, \epsilon^I) \Psi_{\omega_h}^T \end{bmatrix}$$

it comes:

$$B = SS^T + Q$$  \hfill (18)
REFERENCES

28. Skamarock WC, Klemp JB. Adaptive grid refinement for two-dimensional and three-dimensional


List of Figures

1 Notations used in the definition of the nested models .......................... 29
2 Integration of the two grids over one coarse grid time step in the case of a time refinement ratio equal to 2 ................................................. 30
3 a- Coarse resolution background solution of the nested model. b- True state. In red: location of the high resolution grid. .......................... 31
4 Framework of the data assimilation experiments. ................................. 32
5 Single observation experiments: map of the high resolution analysis increment for the case of : a- one-way interaction and the matrix $B_{indep}$, b- two-way interaction and the matrix $B_{indep}$, c- one-way interaction and the matrix $B_{multi}$, d- two-way interaction and the matrix $B_{multi}$. The observation is localized by a red cross in the centre of the maps. .......................... 33
6 Single observation experiments: map of the coarse resolution analysis increment for the case of : a- one-way interaction and the matrix $B_{indep}$, b- two-way interaction and the matrix $B_{indep}$, c- one-way interaction and the matrix $B_{multi}$, d- two-way interaction and the matrix $B_{multi}$. The high resolution grid is localized by the red box. The observation is localized by a red cross in the centre of fine grid. .......................... 34
7 Fine grid $\omega_h$: normalized RMS error as a function of the iteration number. a- Water height. b- Velocity. .................................................. 35
8 Fine grid $\omega_h$: normalized RMS error of the optimized solutions as a function of the integration time. a- Water height. b- Velocity. .................................................. 36
9 Coarse grid $\Omega_H$: normalized RMS error as a function of the iteration number. a- Water height. b- Velocity. .................................................. 37
10 Coarse grid $\Omega_H$: temporal average of the difference between the water height components of the optimized solution and the true state. a- Temporal average of the true state. b- Two-way interaction and the matrix $B_{multi}$, c- One-way interaction and the matrix $B_{indep}$, d- One-way interaction and the matrix $B_{multi}$. .......................... 38
11 Evolution of the normalized $J_{obs}^{\text{free}}$ during the minimization (logarithmic scale). $J_{obs}^{\text{free}}$ is divided by its initial value (free run). ................. 39
12 Fine grid $\omega_h$: normalized RMS error as a function of the iteration number. a- Water height. b- Velocity. .................................................. 40
13 Fine grid $\omega_h$: normalized RMS error of the optimized solutions as a function of the integration time. a- Water height. b- Velocity. .................................................. 41
14 Coarse grid $\Omega_H$: normalized RMS error as a function of the iteration number. a- Water height. b- Velocity. .................................................. 42
Figure 1. Notations used in the definition of the nested models
Figure 2. Integration of the two grids over one coarse grid time step in the case of a time refinement ratio equal to 2.
Figure 3. a- Coarse resolution background solution of the nested model. b- True state. In red: location of the high resolution grid.
Figure 4. Framework of the data assimilation experiments.
Figure 5. Single observation experiments: map of the high resolution analysis increment for the case of: a- one-way interaction and the matrix $B_{indep}$, b- two-way interaction and the matrix $B_{indep}$, c- one-way interaction and the matrix $B_{multi}$, d- two-way interaction and the matrix $B_{multi}$. The observation is localized by a red cross in the centre of the maps.
Figure 6. Single observation experiments: map of the coarse resolution analysis increment for the case of:

a- one-way interaction and the matrix $B_{\text{indep}}$, b- two-way interaction and the matrix $B_{\text{indep}}$, c- one-way interaction and the matrix $B_{\text{multi}}$, d- two-way interaction and the matrix $B_{\text{multi}}$. The high resolution grid is localized by the red box. The observation is localized by a red cross in the centre of fine grid.
Figure 7. Fine grid $\omega_h$: normalized RMS error as a function of the iteration number. a- Water height. b- Velocity.
Figure 8. Fine grid $\omega_h$: normalized RMS error of the optimized solutions as a function of the integration time. a- Water height. b- Velocity.
Figure 9. Coarse grid $\Omega_H$: normalized RMS error as a function of the iteration number. a- Water height. b- Velocity.
a- Temporal average of the true state

b- Two-way $B_{\text{multi}}$

c- One-way $B_{\text{indep}}$

d- One-way $B_{\text{multi}}$

Figure 10. Coarse grid $\Omega_H$: temporal average of the difference between the water height components of the optimized solution and the true state. a- Temporal average of the true state. b- Two-way interaction and the matrix $B_{\text{multi}}$. c- One-way interaction and the matrix $B_{\text{indep}}$. d- One-way interaction and the matrix $B_{\text{multi}}$. 

Copyright © 2000 John Wiley & Sons, Ltd.

Int. J. Numer. Meth. Fluids 2000; 00:1–6

Prepared using fdlauth.cls
Figure 11. Evolution of the normalized $J_{h}^{obs}$ during the minimization (logarithmic scale). $J_{h}^{obs}$ is divided by its initial value (free run).
Figure 12. Fine grid $\omega_h$: normalized RMS error as a function of the iteration number. a- Water height. 
b- Velocity.
Figure 13. Fine grid $\omega_h$: normalized RMS error of the optimized solutions as a function of the integration time. a- Water height. b- Velocity.
Figure 14. Coarse grid $\Omega_H$: normalized RMS error as a function of the iteration number. a- Water height. b- Velocity.
List of Tables

I  Resolution dependent parameters ................................................. 44
II  Multigrid background error covariance matrix: temporal variances used in $\Sigma$. . 45
<table>
<thead>
<tr>
<th>Grid # meshes</th>
<th>Resolution (km)</th>
<th>Time step (s)</th>
<th>Diffusivity coefficient ($m^2.s^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global high resolution model (&quot;true&quot; state)</td>
<td>240 × 240</td>
<td>$h = \frac{25}{3}$</td>
<td>1800</td>
</tr>
<tr>
<td>Global coarse resolution model</td>
<td>80 × 80</td>
<td>$H = 25$</td>
<td>1800</td>
</tr>
<tr>
<td>Local high resolution model</td>
<td>45 × 60</td>
<td>$h = \frac{25}{3}$</td>
<td>1800</td>
</tr>
</tbody>
</table>

Table I. Resolution dependent parameters
Table II. Multigrid background error covariance matrix: temporal variances used in $\Sigma$.

<table>
<thead>
<tr>
<th></th>
<th>$\phi$ (m)</th>
<th>$u$ (m.s$^{-1}$)</th>
<th>$v$ (m.s$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global coarse resolution model</td>
<td>35.8</td>
<td>0.0744</td>
<td>0.055</td>
</tr>
<tr>
<td>Local high resolution model</td>
<td>40.1</td>
<td>0.0832</td>
<td>0.0622</td>
</tr>
</tbody>
</table>