Multigrid methods and data assimilation

applied to a linear advection equation

Neveu Emilie *, Debreu Laurent **, Le Dimet François-Xavier ***

INRIA MOISE Project-Team and Laboratoire Jean Kuntzmann
Grenoble, France
* Emilie.Neveu@imag.fr

** Laurent.Debreu@imag.fr,

*** Francois-Xavier.Le-Dimet@imag.fr

ABSTRACT. In order to limit the computational cost of the variational data assimilation process, we investigate the use of multigrid methods to solve the associated optimal control system. On a linear advection equation, we study the impact of the regularization term of the optimal control and also the impact of the discretization errors on the efficiency of the coarse grid correction step introduced by the multigrid method. We show that even if for a perfect numeric model, the optimal control problem leads to the solution of an elliptic system, discretization errors introduce implicit diffusion that can alter the success of the multigrid methods.

RÉSUMÉ. Afin de limiter le coût de calcul lié aux méthodes variationnelles d'assimilation de données, nous nous intéressons ici à l'utilisation de méthodes multigrilles pour la solution de problème de contrôle optimal. Sur un modèle simple d'advection linéaire, nous étudions l'impact du terme de régularisation du contrôle optimal ainsi que l'impact des erreurs de discrétisation sur l'efficacité de la correction grille grossière introduite par cette méthode. En particulier, nous montrons que pour un modèle numérique parfait, le problème de contrôle optimal est elliptique mais que les erreurs de discrétisation introduisant une diffusion implicite peuvent altérer les performances de la méthode multigrille.

KEYWORDS : variational data assimilation, multigrid methods, optimal control

MOTS-CLÉS : assimilation variationnelle de données, méthodes multigrilles, contrôle optimal

1. Introduction

Data assimilation methods are a way of combining different sources of information: observations and numerical models according to error statistics on these sources. These methods have lead to strong improvements in the operational context of weather or ocean forecast. Data assimilation methods can be divided into two groups [1]. First, sequential methods are based on the Kalman filtering approach and make the state vector evolve in time along with its error statistics. Then, variational methods are based on optimal control techniques and minimize the distance between the model trajectory and observations according to a cost function J. We will focus on the 4D-variational data assimilation (4D-var), introduced by LeDimet and Talagrand in 1986 [4].
Both methods have huge computational costs and have to be simplified for operational purposes. For this reason, multigrid methods are attractive. In variational data assimilation, assuming $\mathbf{x}_0$ is the control vector, the necessary condition of optimality is given by the Euler-Lagrange equation $\nabla_{\mathbf{x}_0} J = 0$. In this paper, the multigrid methods are used for solving this resulting system.
In the optimal control framework, several attempts have been made to apply multigrid methods, either for linear or non linear optimization ([8], [7], [6]). [6] focuses on the control of the initial condition for a linear advection equation with a specific cost function and discretization scheme that makes the problem really well suited for multigrid methods.
In this paper, we present multigrid methods in the general case of linear systems in section 2. In section 3, we apply multigrids on a data assimilation problem characterized by a linear advection equation and a cost function associated with a typical regularization term. Additionally, in section 4, using Fourier analysis, we study how discretization errors in the numerical model can alter the efficiency of the coarse grid correction step.

2. Multigrid methods for the solution of linear systems

Let $\mathbf{x}_0^*$ be the solution of the following linear system:

$$A\mathbf{x}_0^* = g \quad \text{on } \Omega \qquad [1]$$

Let $\mathbf{x}_0$ be an approximation of $\mathbf{x}_0^*$ and $\delta\mathbf{x}_0 = \mathbf{x}_0^* - \mathbf{x}_0$ the error.
The residual r, given by $r = g - A\mathbf{x}_0$, satisfies the residual equation

$$A\delta\mathbf{x}_0 = r \qquad [2]$$

Multigrid methods use different levels of grid resolution to find an efficient solution of equation [1]. In the following, for sake of clarity, we restrict ourselves to a two-level multigrid method. The superscript c (resp. f) stands for the coarse (resp. fine) grid.
To exchange informations between the two grids, we use a restriction operator I_f^c and an interpolation or prolongation operator I_c^f. We denote by A^c and A^f the discretizations of A and by g^c and g^f the right hand side.
Knowing this, the algorithm can be introduced:

Loop over k until convergence, $\mathbf{x}_b^f$ being the first estimate of $\mathbf{x}_0^f$

1) Apply ν_1 times a relaxation method (e.g. weighted-Jacobi (M being the diagonal of A^f)):

$$\mathbf{x}_0^f = \left[\mathbf{I} - \omega M^{-1} A^f\right] \mathbf{x}_0^f + \omega M^{-1} g^f \quad \omega \in [0,1] \qquad [3]$$

2) Coarse grid correction step:

a) Solve exactly

$$A^c \delta \mathbf{x}_0^c = r^c, \quad \text{on } \Omega^c \qquad [4]$$

where r^c is the fine grid residual transfered to the coarse grid: $r^c = I_f^c r^f$.

b) Interpolate the correction and deduce the approximate solution on Ω^f, $\mathbf{x}_0^{f,\,\text{after CGC}} = \mathbf{x}_0^f + I_c^f \delta \mathbf{x}_0^c$, which may be expressed by:

$$\mathbf{x}_0^{f,\,\text{after CGC}} = \left(\mathbf{I} - I_c^f (A^c)^{-1} I_f^c A^f\right) \mathbf{x}_0^f + I_c^f (A^c)^{-1} I_f^c g^f \qquad [5]$$

3) Apply ν_2 times a relaxation method to $\mathbf{x}_0^{f,\,\text{after CGC}}$ as in iteration [3] to find the new $\mathbf{x}_0^f$

End Loop

Knowing the fine grid solution, $\mathbf{x}_0^{f^*}$, which is also a fixed point of the iteration [5], the evolution of the error during the coarse grid correction (step 2) can be written:

$$\mathbf{x}_0^{f,\,\text{after CGC}} - \mathbf{x}_0^{f^*} = K_{\text{mgrid}}(\mathbf{x}_0^f - \mathbf{x}_0^{f^*}) \qquad [6]$$

where

$$K_{\text{mgrid}} = \mathbf{I} - I_c^f (A^c)^{-1} I_f^c A^f \qquad [7]$$

If the operator A is elliptic, relaxation methods applied on the fine grid, like Jacobi or Gauss-Seidel methods, will efficiently remove high frequencies in the error $\delta \mathbf{x}_0$ (see [10]). After some iterations of relaxation methods on the fine grid, the remaining low frequencies of the error can be passed to a coarser grid on which they appear as higher frequencies.

3. Multigrid methods for variational data assimilation

In a variational data assimilation method, the aim is to minimize a given cost function $J(\mathbf{x}_0)$ playing with the control vector $\mathbf{x}_0$. The cost function measures the distance between a set of observations $\mathbf{y}^o$ and the model as follows

$$J(\mathbf{x}_0) = \frac{1}{2} \sum_i \|\mathbf{H}_i(\mathbf{x}_0) - \mathbf{y}_i^o\|^2_{R_i^{-1}} + \frac{1}{2} \|\mathbf{x}_0 - \mathbf{x}_b\|^2_{B^{-1}} \qquad [8]$$

The first term measures the misfit to data while the last term is a regularization one, $\mathbf{x}_b$ being the background or current estimate of the initial state.
$\mathbf{H}_i$ is an operator that includes both the model trajectory $\mathbf{M}_{0,i}$ from time 0 to time i and the observation operator that goes from model space to the observations space. B is the associated background error covariance matrix and R_i the observational error covariance matrix.
In many cases, as operational data assimilation for ocean or meteorological models, the resulting system is non-linear. But in this preliminary study, we will restrict us to linear system to understand the behavior on elementary equations.

Let $\mathbf{x}_0^*$ be the minimum of the cost function : $\mathbf{x}_0^* = \min_{\mathbf{x}_0} J(\mathbf{x}_0)$.
Then, a necessary condition for $\mathbf{x}_0^*$ to be a minimum of $J(\mathbf{x}_0)$ is the Euler equation :

$$\nabla_{\mathbf{x}_0} J(\mathbf{x}_0^*) = 0 \qquad [9]$$

Variational data assimilation ([4] [5], [1]) is based on a minimization procedure, that requires the knowledge of the gradient of J which itself requires integration of both direct and adjoint models. As one integration of an operational model is already of a high cost of resolution, the solution of the data assimilation problem can then become very expensive. Our objective is to alleviate this cost by solving the Euler equation [9] using a multigrid method. We base our study on the work of Nash [7], and Ta'asan [9] who have studied some optimal control problems solved by multigrid methods.
In order to relate this problem to the one introduced in the previous section, we write the Euler equation under the form $A(\mathbf{x}_0) = g$ where $A = \nabla J$ and $g = 0$.
Then the convergence factor of coarse grid correction can be expressed as:

$$K_{\text{mgrid}} = \mathbf{I} - I_c^f (\mathcal{H}^c)^{-1} I_f^c \mathcal{H}^f \qquad [10]$$

where $\mathcal{H}$ is the hessian of J: $\mathcal{H}.\delta\mathbf{x}_0^* = \nabla J(\delta\mathbf{x}_0^*)$.

4. Convergence study on an advection equation

In this section, we study the application of the previously described method on a linear advection equation. First, the continuous and discrete problems are introduced, then we make a Fourier analysis of the convergence properties of the coarse grid correction step.

4.1. Model and cost function

We use a elementary advection equation on a one dimensional domain Ω

$$\left.\begin{array}{l} \partial_t \mathbf{x} + c\,\partial_x \mathbf{x} = 0 \quad \text{with } c > 0 \\ \mathbf{x}(x, t = 0) = \mathbf{x}_0(x) \end{array}\right\} \qquad [11]$$

with periodic boundary conditions .
We suppose that the observation data set $\mathbf{y}^o$ is available continuously (the observation operator is identity).
We suppose that the observational error covariance matrix R is a diagonal matrix with a constant variance equal to σ_o^2 on fine and coarse grids.
The cost function [8] makes use of the background error covariance matrix B. In typical applications, B is representative of errors correlated with a Gaussian shape function. In that case, the regularization term can be approximated using the spatial derivatives of the initial state (see [1]), σ_b^2 being the error variance and l the correlation length.
The continuous cost function is given by

$$J(\mathbf{x}_0) = \frac{1}{2\sigma_o^2} \int_0^T \|\mathbf{x}(x,t) - \mathbf{y}^o\|^2 + \frac{\beta}{2\sqrt{2}\sigma_b^2 l} \left(\|\mathbf{x}_0 - \mathbf{x}_b\|^2 + \|l^2 \triangle (\mathbf{x}_0 - \mathbf{x}_b)\|^2 \right) \qquad [12]$$

β is a positive constant that puts more or less weight on the regularization term.
Using the continuous solution of equation [11], $\mathbf{x}(x,t) = \mathbf{x}_0(x - ct)$, the hessian of J can be derived :

$$\mathcal{H} = \frac{T}{\sigma_o^2} \left(1 + \frac{\gamma}{\sqrt{2}l} \left(\mathbf{I} + l^4 \frac{d^4}{dx^4} \right) \right), \quad \text{where } \gamma = \frac{\beta\sigma_o^2}{T\sigma_b^2} \qquad [13]$$

$\widetilde{\mathcal{H}}(k)$ the symbol of the Hessian defined by $\mathcal{H}(e^{ikx}) = \widetilde{\mathcal{H}}(k) e^{ikx}$ is given by the expression:

$$\widetilde{\mathcal{H}}(k) = \frac{T}{\sigma_o^2} \left(1 + \frac{\gamma}{\sqrt{2}l} \left(1 + l^4 k^4 \right) \right) \qquad [14]$$

With $\gamma = 0$, that is to say if the cost function does not include the regularization term, then the hessian does not depend on k. The optimization method will have a similar behavior at all frequencies so that the multigrid method will loose its theoretical efficiency. However for practical uses, the regularization term is needed and so γ is stricly positive. In that case, an elliptic operator is obtained.

Equation [11] is discretized using a finite difference method based on a uniform grid with time step Δt and space step Δx, using an Euler upwind scheme.
We note $\mathbf{x}_n^j$ the approximation of the value of $\mathbf{x}(x,t)$ at $x = j\Delta x$ and $t = n\Delta t$.
Defining $\lambda = c\Delta t/\Delta x$ the Courant number, the numerical scheme writes:

$$\mathbf{x}_{n+1}^j - \mathbf{x}_n^j = -\lambda\left(\mathbf{x}_n^j - \mathbf{x}_n^{j-1}\right) \qquad [15]$$

The laplacian $\triangle$ of the cost function is discretized using a second order centered scheme. We use two different grids, Ω^f, a fine grid of resolution $\Delta x^f = 0.1$ in space and $\Delta t^f = \lambda\Delta x^f/c$ in time, and a coarse one Ω^c of resolution Δx^c in space and Δt^c in time.
The properties of the Euler upwind scheme are well known. It is first order accurate in space and time and is conditionally stable under the constraint $\lambda \leq 1$. By a Taylor expansion, it can be proved that the numerical solution produced by scheme [15] is a first order approximation in time and space of the advection equation [11] but is a second order approximation of the following advection-diffusion equation (modified equation):

$$\partial_t \mathbf{x} + c\,\partial_x \mathbf{x} = \epsilon \partial_{xx}\mathbf{x}, \text{ with } \epsilon = \frac{c}{2}\left(\Delta x - c\,\Delta t\right) \qquad [16]$$

Note that for the particular case of $\lambda = 1$, $\epsilon = 0$ and the numerical model actually leads to the true solution.
If $\lambda \neq 1$, ϵ is strictly positive due to the stability constraint $0 \leq \lambda \leq 1$ and the additional term on the right hand side corresponds to a diffusion term.
Using the same cost function with the advection-diffusion equation gives us a new expression of the Hessian, which is in Fourier space:

$$\widetilde{\mathcal{H}}(k) = \frac{1 - e^{-\epsilon k^2 T}}{\sigma_o^2 \epsilon k^2} + \frac{\beta}{\sqrt{2}\sigma_b^2 l}\left(1 + l^4 k^4\right) \qquad [17]$$

We give its Taylor expansion at order 4 :

$$\widetilde{\mathcal{H}}(k) = \frac{T}{\sigma_o^2}\left(1 - \frac{\epsilon k^2 T}{2} + \frac{\epsilon^2 T^2 k^4}{3!} + \frac{\gamma}{\sqrt{2}l}\left(1 + l^4 k^4\right)\right) \qquad [18]$$

This expression shows that at large scales there is a balance between artificial numerical diffusion and the regularization term.

4.2. Fourier analysis

Brandt, 1984 [2] shows that h-ellipticity for the system is required for the multigrid methods to be efficient. Elliptic problems have the properties that high frequencies are local. An elliptic operator would react to a high-frequency change by a local high frequency. In our case, the Hessian needs to be elliptic so the following inequality must be satisfied in the discrete case:

$$|\widetilde{\mathcal{H}}(k)| \geq \sum_{j>0} C_j \sin(k\Delta x/2)^{2j}, \quad C_j \geq 0.$$

In the following, we study the convergence factor of one coarse grid correction step.

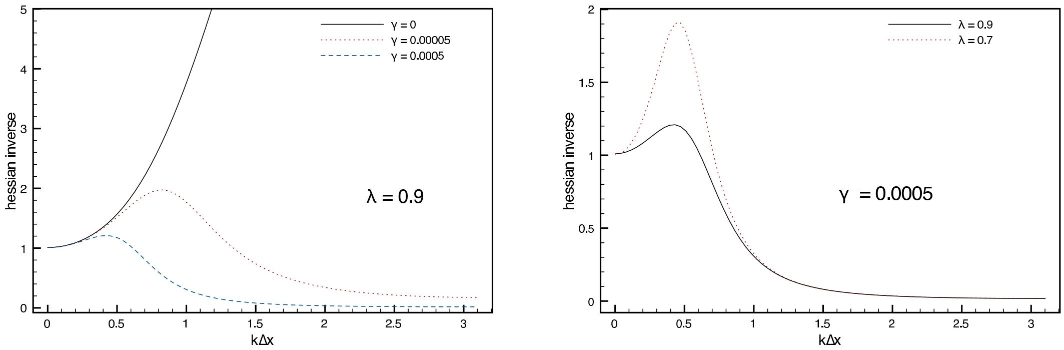

Figure 1. *The inverse of the hessian, $|\widetilde{\mathcal{H}^{-1}}(k)|$, according to $k\Delta x$ for different values of γ to the left and different values of λ to the right.*

4.2.1. Hessian ellipticity

The hessian of data assimilation system depends on the numerical model we use, on the equation we study, and on the cost function we choose. We have briefly studied the behavior of the Hessian for low frequencies (k near of 0). To complete this study, we numerically compute the inverse of the discrete hessian according to the mode k, $|\widetilde{\mathcal{H}^{-1}}(k)]$. The two parameters T and L are given by $T = 1$ and $L = 4$. Values of $|\widetilde{\mathcal{H}^{-1}}(k)]$ shown here are to the true discrete values, the expression [17] being only valid at large scales $k\Delta x \ll 1$.

Figure 1 shows on the left the influence of γ on the modulus of the inverse of the fine grid Hessian for $\lambda = 0.9$ and $l = 1$ and on the right the influence of λ for $\gamma = 0.0005$ and $l = 1$. When the regularization term is omitted ($\gamma = 0$), the increase of numerical diffusion makes the model less controllable. Even if we wouldn't have used multigrid methods, a regularization term $\gamma > 0$ makes the fine grid optimization easier and helps the solution $\mathbf{x}_0$ to have physical meaning. With $\gamma > 0$, at medium scales there is a competition between the artificial diffusion induced by the discretization and the regularization term. Even if the Hessian is not truly elliptic, the regularization term also improves its properties with regard to the convergence of the multigrid process.

4.2.2. The coarse grid correction

Assuming we have solved exactly the coarse grid system [4] $\mathcal{H}^c.\delta\mathbf{x}_0^c = -I_f^c\nabla\mathcal{J}^f(\mathbf{x}_0^f)$, the impact of the coarse grid correction can be studied using the convergence factor [10]. To be efficient for the multigrid process, it should be small at large scales.

We assume here that the spatial and temporal refinement factors are equal to 2. Thus the value of the Courant number λ is the same for the two grids, but numerical diffusion is twice on coarse grid. The other parameters (γ and l) have identical values on fine and coarse grids.

The chosen restriction operator corresponds to a full-weighting operator while the interpolation operator is linear. Their stencils are given by :

$$I_f^c = \frac{1}{2}\left[\begin{array}{ccc} \frac{1}{2} & 1 & \frac{1}{2} \end{array}\right]_f \quad I_c^f = \left[\begin{array}{c} \frac{1}{2} \\ 1 \\ \frac{1}{2} \end{array}\right]_c$$

In Figure 2, on the left, we plot the convergence factor for different values of γ and λ. The local maximum we find for low frequencies between $[0.2; 0.8]$ is the same as found on

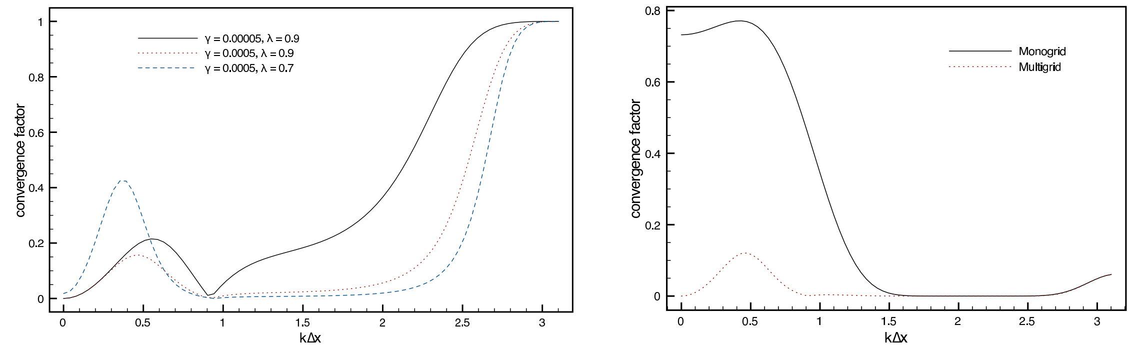

Figure 2. *To the left, the convergence factor for one coarse grid correction according to $k\Delta x$, for different values of γ and λ. To the right, the convergence factor for the weighted-Jacobi on Ω^f and for the multigrids using the same relaxation method on fine grid.*

the inverse of Hessian and is due to the fact that the regularization term, which damps the nonellipticity effect, isn't effective for very low frequencies. Anyway, the low frequencies are quite better resolved than the high ones and a coarse grid correction step efficiently reduces the errors at large scales.

In Figure 2, to the right, using $\gamma = 0,005$, $\lambda = 0.9$ and $l = 1$, we plot the convergence factor of ten high resolution weighted-Jacobi relaxation (see equation [3] with $\omega = 2/3$). We compare it to the convergence factor of the all-multigrid algorithm combining the convergence factor of the relaxation method on fine grid ($\nu_1 = \nu_2 = 5$) with K_{mgrid}. The multigrid method removes much better the low frequencies than monogrid does, which emphasizes the coarse grid correction efficiency.

5. Conclusion

We have introduced a multigrid algorithm for solving an idealized 4D-var problem. The 4D-var resolution has a high computational cost that prevents its use in an efficient way for operational context. The multigrid methods are a way of solving a system solver by using grids at coarser resolutions. Following Nash [6], the idea was to adapt the multigrid methods to data assimilation to low its cost of resolution. We have studied the efficiency of the coarse grid correction according to the parameters of the cost function and show the importance of defining a well suited cost function.

More precisely, we have shown that the regularization term is important for two reasons, the ellipticity of the hessian and the regularization of the solution $\mathbf{x}_0$ to be found. Discretization errors can also have an impact on the coarse grid correction step.

Obviously, much work still has to be done. Operational models in ocean or weather forecasts can have strong non-linearities. Fortunately, multigrid methods can be adapted to non-linear systems using the Full-Approximation Scheme (FAS) [10]. Further efforts will be made to adapt it on a non-linear 4D-var.

Moreover one variant of 4D-var method, used operationally, is the incremental 4D-Var given by Courtier [3]. In this approach, the model is linearized around the current estimate, like in a Gauss Newton optimization method. In the multi-incremental method [11], additionally, the increment is searched in a lower dimensional space, typically corresponding to a grid with coarser resolution. Because the multi-incremental approach makes use of grids of different resolutions, it is linked in some way to multigrid methods. The idea is to experiment the multigrid approach on a non-linear 4D-var and compared its results to the incremental and multi-incremental 4D-var method.

6. Biography

Emilie Neveu *is a PhD student since 2007, she's working on the multigrid methods and its applications to data assimilation for geophysical fluids, in particular in the context of assimilation of images.*

Laurent Debreu *is an INRIA Research Scientist. His main fields of interest are numerical methods such as grid refinement and data assimilation for geophysical fluids*

Francois-Xavier Le Dimet *is Professor at University Joseph- Fourier, Grenoble, France. His main field of interest is the theory of variational data assimilation for geophysical fluids.*

7. References

[1] ANDREW F. BENNETT. "Inverse Modeling of the Ocean and Atmosphere", *Cambridge University Press*, 2002.

[2] A. BRANDT. " Multigrid techniques 1984 Guide with applications to fluid dynamics", *Gesellschaft fuer Mathematik und Datenverarbeitung m.b.H. Bonn (GMD), St. Augustin (Germany)*, 1984.

[3] PHILIPPE COURTIER, J. N. THÉPAUT, AND A. HOLLINGSWORTH. " A strategy for operational implementation of 4d-var, using an incremental approach." *Q. J. R. Meteorol. Soc.*, vol. 123:1367–1387, 1994.

[4] FRANÇOIS-XAVIER LE DIMET AND O. TALAGRAND. " Variational algorithms for analysis and assimilation of meteorological observations: Therotical aspects." *Tellus*, vol. 38(A):97–110, 1986.

[5] IDE, KAYO AND COURTIER, PHILIPPE AND GHIL, MICHAEL AND LORENC, ANDREW C. "Unified notation for data assimilation: Operational, and Variational." *J. Met. Soc. of Japan*, vol. 75(1B):181-189, 1997.

[6] ROBERT MICHAEL LEWIS AND STEPHEN G. NASH. " Model problems for the multigrid optimization of systems governed by differential equations." *SIAM J. SCI. COMPUT.*, vol. 26(6):1811–1837, 2005.

[7] STEPHEN G. NASH. " A multigrid approach to discretized optimization problems." *Journal of Optimization Methods and Software*, vol. 14:99–116, 2000.

[8] SHLOMO. TA'ASAN. " One shot methods for optimal control of distributed parameters systems i: Finite dimensional control." *Technical Report No 91-2, ICASE Report*, 1991.

[9] SHLOMO. TA'ASAN. " Multigrid one-shot methods and Design Strategy. " *Carnegie Mellon University*, 1997.

[10] ULRICH TROTTENBERG, CORNELIS W. OOSTERLEE, AND ANTON SCHÜLLER. " Multigrid " *Academic Press*, 2001.

[11] F. VEERSÉ AND J. N. THÉPAUT. "Multiple-truncation incremental approach of four-dimensional variational data assimilation. " *Q. J. R. Meteorol. Soc.*, vol. 124:1889–1908, 1998.