14B.1
Variational assimilation in a prototype limited area model

Pierre Bernardet*
Meteo-France, Toulouse, France

July 7, 2005

1 Introduction

Most implementations of the adjoint method employ global models and attention now turns to limited area models where the use of variational methods is more complicated . This problem has already been given attention; Lu and Browning (2000) follow the theory of Gustaffson et al (1972) for systems of hyperbolic equations, and show the well posedness of the adjoint equations; however they do not address the discretization problems and the determination of the incoming boundary forcing in their numerical experiment. On the practical side, Zou and Kuo (1996) used the grid point model MM5 and made a first attempt to determine a boundary forcing along with the interior solution : the trend at the boundary was optimized, and they concluded that this trend had a major impact on the quality of the retrieved fields.

Trying to address the boundary control, we will encounter problems with the convergence of the adjoint; use of the discretization of the continuous adjoint model is often advocated. We will meet this problem and instead we will bring the adequate modification to the direct model so that the problem disappears.

The difficulties encountered even with the simplest models lead us to address first the problem with the simplest prototype model, advection in one dimension, discussed at length in Bennett (2002) and we consider the determination of the boundary forcing along with the initial state of the model.

**author address :* Pierre Bernardet, MétéoFrance, CNRM/GMME, 57 Ave G. Coriolis, 31057 Toulouse, France. Phone : (33)561079378. Email : Pierre.Bernardet@meteo.fr

Different space discretizations have been studied: centered differences, grid-point methods, Legendre polynomial finite element method, Ikawa spectral method; we have retained only centered differences in this presentation.

The questions we will address are :

- Does the adjoint of the discretized direct equations approximate the continuous adjoint?
- does the optimal control of the discretized problem converge to the physical one, even in the case the adjoint of the model equations has a pathological behavior?
- An iterative minimization method is practical only when the discretized problem is well conditioned; is it the case?

2 Adding boundary forcing in the control :

We consider the advection equation on the segment $[a, b] = [0, 1]$. The direct equation is :

$$\begin{aligned}
\frac{\partial u}{\partial t} &= Lu = -U\frac{\partial u}{\partial x} \\
u(a,t) &= f(t) \\
u(x,t_i) &= u_0
\end{aligned}$$

and the resolvent is :

$$u(x,t_f) = \mathcal{L}(u_0, f)$$

We assume observations are complete and located at final time $t_f = 0$, that is u is given at final time :

$$u(x, t=0) = \widetilde{u}$$

Variational assimilation considers a functional of the initial state u_0 and boundary forcing $f(t)$ designed to minimize the discrepancy at final time between model state and observations :

$$\mathcal{J}(u_0, f) = 1/2 \int_a^b \left(u(x,t_f) - \widetilde{u}\right)^2 dx$$

Notice that we omit at first a penalty on the smoothness of the control (u_0, f). $\mathcal{J}$ will be at a minimum if its first variation is null; let $v(x,t_f) = u(x,t_f) - \widetilde{u}$; then

$$\delta\mathcal{J} = \int_a^b v(x,t_f)\delta u(x,t_f)dx$$

If the adjoint variable v satisfies :

$$\begin{aligned} v(x,t_f) &= u(x,t_f) - \widetilde{u} \\ \frac{\partial v}{\partial t} &= -L^* v \\ v(b,t) &= 0 \end{aligned}$$

the variation of $\mathcal{J}$ is expressed under the required form :

$$\delta\mathcal{J} = \int_a^b v(x,0)\delta u(x,0)dx + U\int_{t_i}^0 v(a,t)\delta f dt$$

where the variation of the controls appears solely.

We see that two scalar products are present : one for $u(x,t_f)$ on $[a,b]$:

$$\langle u; v\rangle = \int_a^b uvdx$$

and the other one on the control space $\{u(x,t_i), f(t)\}$:

$$\langle\langle (u,f);(v,g)\rangle\rangle = \int_a^b uvdx + U\int_{t_i}^0 fgdt$$

and $\mathcal{L}^*$ is defined by :

$$\langle \mathcal{L}(u,f); v\rangle = \langle\langle (u,f); \mathcal{L}^*(v)\rangle\rangle$$

We see that the adequate scalar product for the control considers the advecting speed U. So, as it was obvious, the adjoint is the same advection backwards, and boundary control should be given as $f(t) = \widetilde{u}(a - Ut)$ for $t_f - t < (b-a)/U$

3 Pathology of the adjoint integrations :

There are two steps in the forward model: we replace the left value of u by the boundary forcing after advection on the other points. We leave to the reader the detail of the adjoint; the result is :

$$\begin{aligned} g^n &= \frac{v_1^n \Delta x}{U\Delta t}\left(= v_2^{n+1}\right) \\ v_1^{n-1} &= \frac{U\Delta t}{\Delta x} v_2^n \\ v_j^{n-1} &= v_j^{n+1} + \frac{U\Delta t}{\Delta x}(v_{j+1}^n - v_{j-1}^n) \\ v_J^{n-1} &= v_J^{n+1} + \frac{U\Delta t}{\Delta x}(0 - v_J^n - v_{J-1}^n) \end{aligned}$$

The boundary gradient takes the value $g(t_f) = \frac{v_1^N \Delta x}{U\Delta t}$ at time step N for t_f, thus the shock especially for small time steps; null values enter by relaxation to the right and here the refection of physical modes to computational modes at the left boundary is obvious (fig.1)

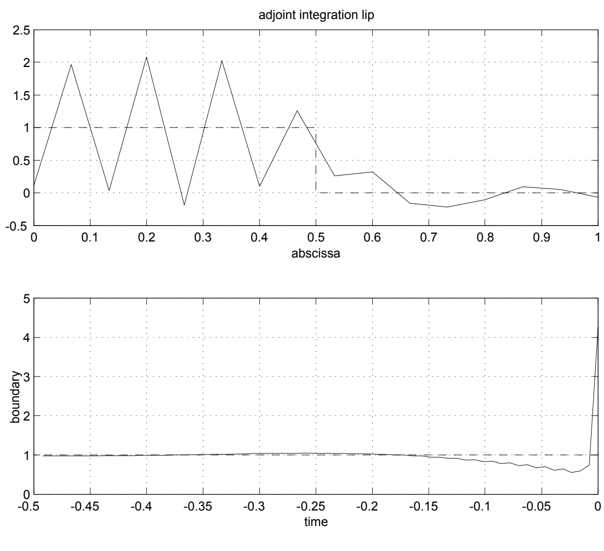

Figure 1: Adjoint integration with a uniform final state and courant number .125; a) (top) : adjoint v at $t = -.5$ b) (bottom) : boundary values g from $t_f = 0$ to $t_i = -.5$. Dashed lines show theoretical solution.

4 Modification of the forward model :

4.1 The scheme

Given the form of the boundary forcing at the right in the adjoint model, we guess what should be the forcing in the direct model by adjusting the weight of the scalar product at the boundary and the coefficient of relaxation upon u_1, u_2; the outcome is :

$$\frac{\partial u_1}{\partial t} = \frac{U}{\Delta x}\left(2f(t) - u_1 - u_2\right) \quad (1)$$
$$\frac{\partial u_j}{\partial t} = \frac{U}{2\Delta x}\left(u_{j-1} - u_{j+1}\right)$$
$$\frac{\partial u_J}{\partial t} = \frac{U}{\Delta x}\left(u_{J-1} - u_J\right) \quad (2)$$

with scalar product

$$\langle u; v\rangle = \Delta x\left(\frac{1}{2}u_1 v_1 + \sum_2^{J-1} u_j v_j + \frac{1}{2}u_J v_J\right)$$

. The adjoint is identical to the direct model. We notice that the forcing in a implies the sum of u_1 and u_2; this is precisely a radiative condition for the computational modes that travel to the left in the direct model.

So the key to a convergent adjoint is to enter information in the direct model through a Newtonian relaxation. This way of dealing with the boundary forcing is not standard; we will call it weak forcing, and the ordinary forcing with imposition of the boundary value will be called Tau method following the terminology of finite elements. We see now (fig.2) that the adjoint integration behaves normally.

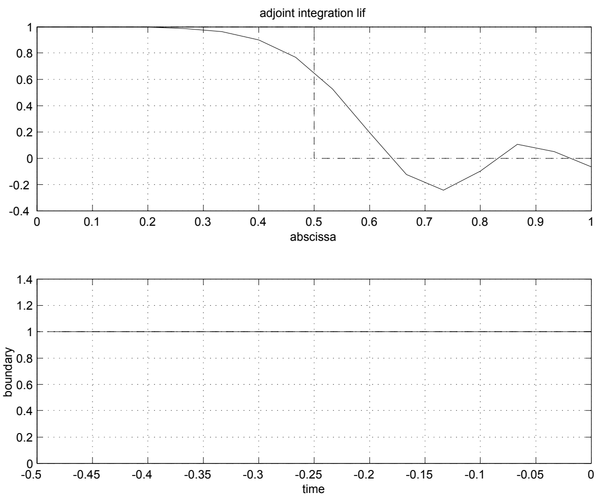

Figure 2: Same as fig.1 but with weak forcing;

4.2 Accuracy of weak forcing :

Again we consider a semi-infinite domain $[a, \infty[$ and forcing by a periodic boundary value $f = \hat{f}\exp(i\omega t)$; for centered differences the steady response with strong forcing is equal to the ideal response $u = \hat{u}\exp i(kx - \omega t)$ with $\hat{u} = \hat{f}$ and with k determined by the dispersion relation; we want to estimate the response with weak forcing. As the computational mode propagates to the left, only the physical wave is present. Replacing in Eq.(1) we get :

$$\hat{u}\left(1 + \exp(ik\Delta x) + 2i\sin(k\Delta x)\right) = 2\hat{f}$$

which means that the response is correct for small wave-numbers, but that the forcing accuracy matches the first order accuracy of the radiating condition Eq.(2). As during the minimization process we determine the forcing $f(t)$ from boundary gradients, we are sure that high frequencies are absent (see below).

5 Solution space :

We ask what control $\theta = \left(u^0(x), f(t)\right)$ such that at final time $\mathbf{L}\theta = u$ is chosen by a descent method, with $\mathbf{L}$ the resolvent for the N time steps of the discretized model with J modes, minimization starting from the guess θ_G. Gradients are $\nabla\mathcal{J} = \mathbf{L}^*(\mathbf{L}\theta - \widetilde{u})$, so the minimization runs in the subspace $\theta_G + \mathbf{L}^*(\mathcal{B}_J)$ of the $K + N$ space of controls, where $\mathcal{B}_J$ is the space of grid-point values. Let v be the departure from the guess : $v = \tilde{u} - \mathbf{L}(\theta_G)$; the condition for the minimum is $\nabla\mathcal{J} = 0$:

$$\mathbf{L}^*\mathbf{L}\Delta\theta = \mathbf{L}^* v$$

$\mathbf{L}$ is rectangular of dimension $J \times (N+J)$, so $\mathbf{L}^*\mathbf{L}$ is not invertible; however $\mathbf{LL}^*$ is, its eigenvalues are those non-null eigenvalues of $\mathbf{L}^*\mathbf{L}$. One verifies that the solution in $\theta_G + \mathbf{L}^*(\mathcal{B}_J)$ is :

$$\theta = \theta_G + \mathbf{L}^* \left(\mathbf{L}\mathbf{L}^*\right)^{-1} v$$

6 Conditioning of the system

The solution above will easily be reached by a gradient method provided that κ, condition number of $\mathbf{L}\mathbf{L}^*$ is small; the condition number κ has been evaluated for our two discretization methods with appropriate time steps and $t_i = -3$ and is displayed below.

cond($\mathbf{L}\mathbf{L}^*$)	$J = 8$	$J = 16$
Tau forcing	$1.E4$	$2.2E5$
weak forcing	$4.E2$	$9.E3$

Condition number , number of points J.

For all the methods, κ increases rapidly with K; conditioning is better with weak method than with Tau method; the high condition number of finite differences reflects the fact that these methods have a plentiful of computational modes with a reverse group velocity that, forced by the wrong side, decay exponentially to the right : if the observation $\widetilde{u}$ presents computational noise, it will be difficult to attain.

Thus another approach has to be taken to show controllability of physical modes only; one should consider a subspace of eigenvectors of the advection equation describing long modes, as is suggested in Infante and Zuazua, 1998, and show that the solution can be approached using that subspace; however here the eigenvectors are more complex and their computational part remains dominant near $x = a$ even for small eigenvalues ω; we have not pursued along this line.

Instead, we have tried to detect when the minimization makes its way through completely unphysical states by displaying the quantity $\hat{u} = \left(\mathbf{L}\mathbf{L}^*\right)^{-1} \widetilde{u}$; we will call $\hat{u}$ the precursor of the solution θ as $\theta = \mathbf{L}^*\hat{u}$.

7 Convergence of the solution

Solution $\theta = (u_s, f_s)$ of the minimization problem with, again, $\widetilde{u}(x) = 1, U = 1, t_f = 0, t_i = -1/2 = -(b-a)/2, \Delta x = (b-a)/16$ and using traditional centered differences is displayed in fig.3; the solution departs from the theoretical one significantly. The boundary part of the solution f_s is roughly equal to $1/2$; u_s in the left part of the domain is a superposition of a constant unity value, which will be advected to the right part, and of numerical noise of wave-length $2\Delta x$ and amplitude $1/2$, which will be reflected in a and add to values forced by f_s to give the unit amplitude of $\widetilde{u}$. Examination of the precursor u_p shows its left part has a reasonable amplitude, but that its right part shows large amplitude noise, that will propagate to the right by $\mathbf{L}^*$ to give the right part of u_s; as we have shown, the reflection coefficient is small, which explains the size of the noise. One might wonder why the internal solution u_s needs to have unit values to the right instead of zero values as one would expect : in this latter case, we would have also have short waves propagating from the discontinuity to the left, spoiling the solution.

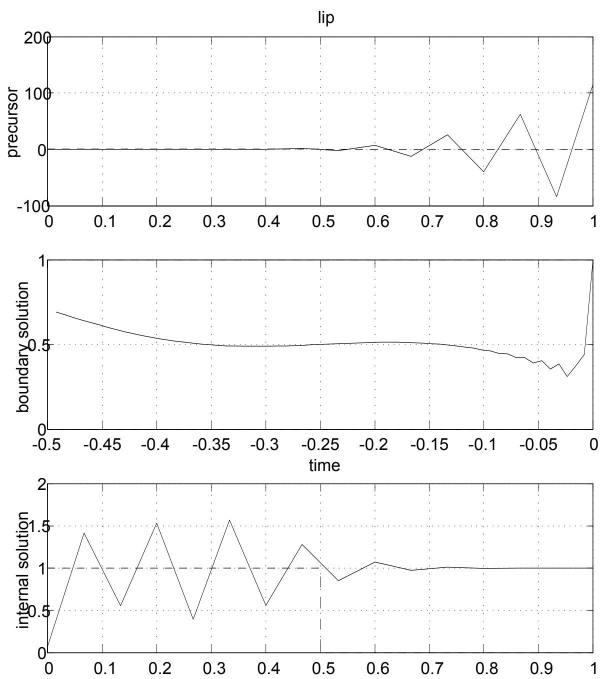

Figure 3: Solution with normal forcing : a) precursor $\hat{u} = \left(\mathbf{L}\mathbf{L}^*\right)^{-1} \widetilde{u}$; b) boundary solution; c) Internal solution

Fig..4 shows the solution for centered differences with weak forcing; the solution is nearly perfect, with, again, unit values of u_s extending in the right half of the domain and a similar precursor.

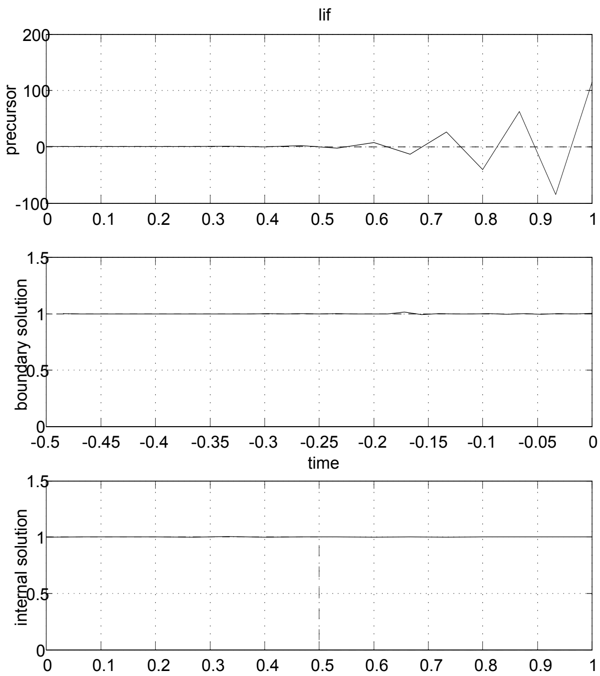

Figure 4: Same as fig.3 but with weak forcing

Apparently when we solve the minimization problem up to the exact solution we generate unphysical states. So we would like to penalize the precursor. First let us notice that it is equivalent to minimize :

$$\begin{aligned}\mathcal{J}_\theta(\theta) &= \frac{1}{2}(\mathbf{L}\theta - \widetilde{u})^t(\mathbf{L}\theta - \widetilde{u}) \\ \nabla\mathcal{J}_\theta &= \mathbf{L}^*\mathbf{L}\theta - \mathbf{L}^*\widetilde{u}\end{aligned}$$

or :

$$\begin{aligned}\mathcal{J}_u(\hat{u}) &= \frac{1}{2}\hat{u}^t\mathbf{L}\mathbf{L}^*\hat{u} - \hat{u}^t\widetilde{u} \\ \nabla\mathcal{J}_u &= \mathbf{L}\mathbf{L}^*\hat{u} - \widetilde{u}\end{aligned}$$

as their solution are related by $\theta = \mathbf{L}^*\hat{u}$ and gradients by :

$$\nabla\mathcal{J}_\theta = \mathbf{L}^*\nabla\mathcal{J}_u$$

So we can penalize $\mathcal{J}_u$ by considering :

$$\mathcal{J}_u(\hat{u}) = \frac{1}{2}\hat{u}^t\mathbf{L}\mathbf{L}^*\hat{u} - \hat{u}^t\widetilde{u} + \frac{\varepsilon}{2}\hat{u}^t\hat{u}$$

which is equivalent to the problem for θ :

$$\mathcal{J}_\theta(\theta) == \frac{1}{2}(\mathbf{L}\theta - \widetilde{u})^*(\mathbf{L}\theta - \widetilde{u}) + \frac{\varepsilon}{2}\theta^*\theta$$

as we can check from the gradients. This formulation has been tried for the grid-point method with weak forcing and $\varepsilon = 1$ (fig.5); the scheme limits the value of the precursor without significant damping of the solution.

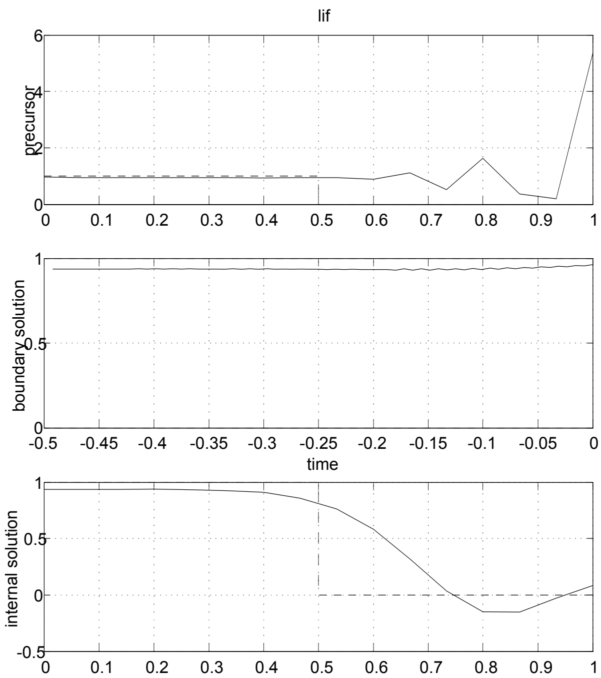

Figure 5: Same as fig.4 but with penalization

When computational modes are not allowed to interfere, it is either that they are not present or that the scheme suppresses them (Lax-Wendroff scheme damps them, weak forcing allows no reflection), there is no other choice than the exact solution. To illustrate this, we take $t_i = -2$ and the traditional centered differences; as the model is damping, we expect that the interior solution will play no part, so the computational modes; this is what we observe in fig.6 : the boundary solution has this time the right amplitude.

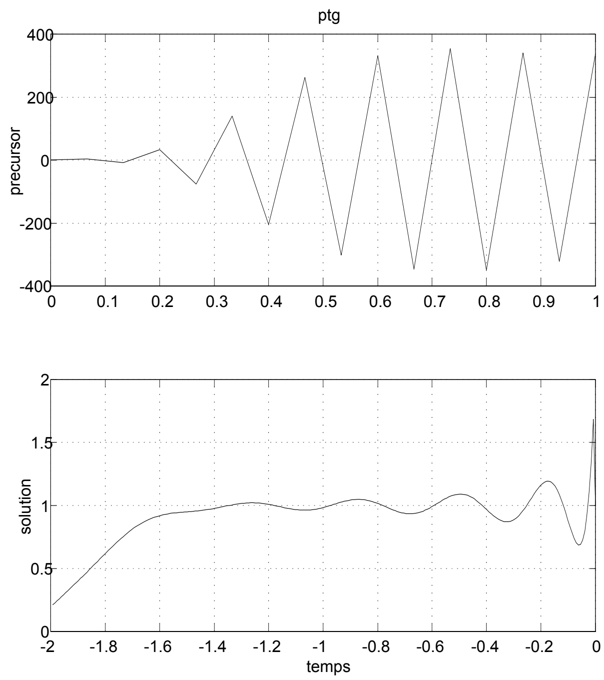

Figure 6: normal forcing, no penalization, $t_i = -2$; solution : top :Precursor; bottom: boundary forcing

8 Discussion :

We have evidenced that the adjoint of the discretized forward model should converge towards the continuous adjoint in order to get the right solution when computational modes are allowed to influence the minimization process. With a cautious discretization of the forward model, the adjoint of the discretized converges; that means that, given a state $\widetilde{u}$ to be reached with an appropriate boundary forcing θ, and an error target, it suffices to choose the discretization J to attain the solution at the first step of a descent method.

We have shown that, even in the case of a weak forcing of our advection equation, the condition number of the Hessian of the cost function is only indicative; its high value shows it does not convey the appropriate information about the difficulty of the minimization process. We have proposed another criterion, amplitude of the precursor, to give additional information.

We have employed for didactical purposes the most simple cost function; caution should be employed in its formulation : a penalty upon the control should be incorporated so that to prevent erroneous solutions due to forcing in a short period of time; this point has been advocated already by Bennett et al (1990); computational modes should be eradicated by an appropriate representation error penalty or by a damping scheme; we should seek a smooth control to handle the problem of the shock in the adjoint integrations when observations depart from the guess at the boundaries of the domain.

In a more complete study, we have shown some superiority of the spectral method. The only spectral methods practically used in limited area modeling are Fourier methods with periodisation (Haugen and Machenhauer, 1993), or a sin expansion null at the boundary, completed by a cosine to match the boundary values (Ikawa, 1987); Tau and weak method have been tested on the latter with the same conclusions.

Ongoing studies pertaining to the forcing of limited area models tend to get rid of the traditional Davies (1976) relaxation zones, and to discriminate between the different types of waves; weak forcing as introduced here should be considered, as it is more natural and eliminates reflections under the form of computational modes at the boundary.

I thank L. Amodei for fruitful discussions during the development of this work.

REF-
ERENCES

Bennett, A. F.and R. N. Miller, 1990: Weighting initial conditions in variational assimilation schemes. *J. Atmos. Sci.,* **119**, 1098-1102.

Bennett,A. F., 2002: Inverse Modeling of the Ocean and Atmosphere.Cambridge University Press.234pp.

Davies, H. C.,1976: A lateral loundary formulation for multi-level prediction models. *Quart. J. Roy. Met. Soc.* **102**, 405-418.

Haugen, J and B. Machenhauer, 1993 : A spectral limited-area model formulation with time-dependent boundary conditions applied to the shallow-water equations. *Mon. Wea. Rev.* **121**, 2618-2630.

Ikawa, M., 1987 : A consideration of Spectral methods for Limited area models. *J. Met. Soc. Japan*, **65**, 439-452.

Infante, J.-A. and E. Zuazua, 1998 : Boundary observability for the space-discretizations of the 1-D wave equation. *Comptes Rendus de l'Academie des Sciences Series I Mathematics,* March 1998.

Gustafsson, B., H.-O. Kreiss, and A. Sundström, 1972 : Stability theory of difference approximations for mixed initial boundary value problems. II. *Math. Comp.*, **26**, 649–686.

Laurent P-J., 1972 : Approximation et optimisation. *Hermann.*

Lu C. and G. L. Browning, 2000 : Four-dimensional variational data assimilation for limited-area models : lateral boundary conditions, solution uniqueness and numerical convergence. *J. Atmos. Sci*. **57**, 1341-1353.

P. Lax and B. Wendroff, 1960: Systems of conservation laws. *Comm. Pure Appl. Math.* **13**, 267-293.

Orlanski, I.,1976 : A simple boundary condition for unbounded hyperbolic flows. *J. Comput. Phys.* **21**, 251-269.

Vukicevic T., M. Steyskal and M. Hecht, 2001 : Properties of Advection Algorithms in the Context of Variational Data Assimilation. *Mon. Wea. Rev.* **129**, 1221–1231.

Zou, X. and Y.-H. Kuo, 1996 : Rainfall assimilation through an optimal control of initial and boundary conditions in a limited-area mesoscale model. *Mon. Wea. Rev.* **124**, 2859-2882.