Use of a digital filter as weak constraint in variational data assimilation

Nils Gustafsson

Swedish Meteorological and Hydrological Institute

Abstract

A simple experiment with a 1-dimensional spectral shallow-water model has been carried out to test the possible use of a digital filter for control of gravity wave noise in variational data assimilation based on the adjoint model technique. The digital filter is applied as a weak constraint by adding of an extra quadratic term to the cost function that is minimized through the assimilation procedure. The applied technique with the digital filter as a weak constraint is compared with other techniques for gravity noise control e.g. the non-linear normal mode initialization applied at the start of the data assimilation period, minimization of the gravity mode tendencies in a quadratic norm and a time-stepping scheme based on non-linear normal mode initialization. The comparisons are carried by the aid of simple identical twin observing system simulation experiments. The digital filter approach proves to have similar characteristics as gravity noise control based on minimization of the gravity mode tendencies in a quadratic norm.

1. Introduction

For any full scale meteorological data assimilation problem, the degrees of freedom of the model is at least one order of magnitude larger than the degrees of freedom of the available observations. To circumvent this problem, it is necessary to add various constraints in order to reduce the degrees of freedom of the model system. In the traditional forward intermittent data assimilation, constraints are applied through some analysis procedure, e.g. statistical interpolation (Gandin 1963), that filters out small scale noise and forces the assimilation increments to fulfill some simple diagnostic relation like the geostrophic wind and also through some initialization procedure like the non-linear normal mode initialization (Machenhauer 1977). In 4-dimensional data assimilation, when addition of the time dimension increases the degrees of freedom of the observational data, the assimilation problem is still under-determined. Within the framework of 4-dimensional data assimilation based on variational principles (Le Demit and Talagrand 1986), it is possible add constraints on the assimilation increments similar to that of statistical interpolation (Lorenc 1986) and non-linear normal mode initialization. Constraints based on normal mode initialization can either be added as strong constraints, by inclusion of the initialization procedure into the data assimilation cycle, or as a weak constraint by minimization of a cost function that includes gravity wave tendencies in a quadratic term.

One problem with the constraints based on non-linear mode initialization is that the normal modes of the model linearized around a simplified basic state need to be determined. This could be cumbersome and costly, in particular if we move towards more complete, e.g. non-hydrostatic, forecast model equations. Recently Lynch and Huang (1992) successfully applied the idea of using digital filters to initialize a full scale high resolution forecast model. Initialization by this technique consists of an integration of the forecast model from un-initialized data over a certain time-period, e.g. 6 hours, followed by a filtering of the series of forecast model states by a time-filter (grid-point by grid-point or spectral component by spectral component). The aim of the present paper is to study the use of a digital filter in a weak constraint to control gravity wave noise in variational data assimilation based on the adjoint model technique.

2. A digital filter

Consider a time-series of "noisy" forecast values ...... f_{-2} f_{-1} f_0 f_1 f_2 ...... with the subscripts referring to the time-step number of the model centered around some reference time 0. The basic idea of using digital filters for initialization is to (1) Calculate the Fourier transform in the time dimension, (2) Set coefficients of high frequencies to zero, (3) Calculate the inverse Fourier transform to obtain filtered forecast values and, finally, (4) Use the filtered values as initial values for a new forecast starting at time 0. The filter could be applied grid-point by grid-point or spectral component by spectral component. For any practical application, the three first steps are combined into a filter that is applied in analogy with a simple weighted averaging for each model component in the time-domain. The same non-recursive filter as applied by Lynch and Huang (1992) will be utilized

$$f_n = \sum_{k=-N}^{N} a_k f_{n-k}$$

For derivation of the filter coefficients a_k we will consider the time history of the model variables as Fourier coefficients in the following expansion

$$F(\theta) = \sum_{n=-\infty}^{\infty} f_n e^{-in\theta}$$

where θ is a digital frequency. To filter in the time-domain we introduce a cut-off function

$$H_d(\theta) = \sum_{m=-\infty}^{\infty} H_m e^{-im\theta}$$

with $H_d(\theta) = 0$ for ABS(θ) $\geq \theta_c$ and $H_d(\theta) = 1$ for ABS(θ) $< \theta_c$ where θ_c is the cut-off frequency. The coefficients h_m are given by

$$h_m = \frac{1}{2\pi} \int_{-\pi}^{\pi} H_d(\theta)\, e^{im\theta} d\theta$$

Combining the two expansions and substituting m+n with l, we will have:

$$F(\theta) H_d(\theta) = \sum_{l=-\infty}^{\infty} \left[\sum_{m=-\infty}^{\infty} h_m f_{l-m} \right] e^{-il\theta}$$

The expression in square brackets obviously consists of our filtered forecast values. For any real application we need to introduce a truncated series:

$$\tilde{f}_l = \sum_{m=-M}^{M} h_m f_{l-m}$$

For the cut-off function defined above, we will have the following filtering weights:

$$h_m = \frac{\sin(m\theta_c)}{m\pi}$$

By using a truncated Fourier series, Gibbs-waves are introduced, and the effect of the filter will be less sharp. To partly avoid the effect of the Gibbs-waves, the filtering weights will be multiplied by the same window function as was used by Lynch and Huang (the Lanczos window). The final expression for the filtering weights, excluding normalization factors, is

$$h_m = \left[\frac{\sin(m\pi/(N+1))}{m\pi/(N+1)}\right]\left[\frac{\sin(n\theta_c)}{n\pi}\right]$$

In the application described below, we will use a cut-off frequency θ_c = 6 hours.

3. A one-dimensional spectral shallow water model and its adjoint.

To test the idea of using digital filters as weak constraints in variational data assimilation including the time dimension, a simple spectral 1-dimensional shallow water model on a beta-plane will be utilized. The basic model equations, before elimination of the y-dependency, are:

$$\frac{du}{dt} - fv + \frac{\partial\phi}{\partial x} = 0$$

$$\frac{dv}{dt} + fu + \frac{\partial\phi}{\partial y} = 0$$

$$\frac{d\phi}{dt} + \phi\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right) = 0$$

$$\phi = gh$$

$$f=f_0+\beta y$$

For elimination of the y-dependency, we introduce a mean zonal wind $\overline{u}$, a mean zonal geopotential gradient $\frac{\partial\overline{\phi}}{\partial y}$ and a variation of deviations from this mean state in the x-direction only:

$$u=\overline{u}+u'(x,t)$$

$$v=v'(x,t)$$

$$\phi=\phi_0+\frac{\partial\overline{\phi}}{\partial y}y+\phi'(x,t)$$

Introducing a spectral representation with cyclic boundary conditions in the x-direction, the non-linear model equations for the spectral coefficients, with non-linear terms computed by the transform method, are given by:

$$\frac{d\hat{u}_k}{dt}=-ik'\overline{u}\hat{u}_k-F_k[F^{-1}(\hat{u})F^{-1}(il'\hat{u})]+f\hat{v}_k+i\frac{\beta}{k'}\hat{u}_k-ik'\hat{\phi}_k$$

$$\frac{d\hat{v}_k}{dt}=-ik'\overline{u}\hat{v}_k-F_k[F^{-1}(\hat{u})F^{-1}(il'\hat{v})]-f\hat{u}_k+i\frac{\beta}{k'}\hat{v}_k$$

$$\frac{d\hat{\phi}_k}{dt}=-ik'\overline{u}\hat{\phi}_k-ik'F_k[F^{-1}(\hat{u})F^{-1}(\hat{\phi})]+f\overline{u}\hat{v}_k-\overline{\phi}ik'\hat{u}_k$$

where F denotes Fourier transform , F^{1} inverse Fourier transform and k' scaled wave-number:

$$\hat{a}_k=F_k(a)=\frac{1}{N}\sum_{j=1}^{N}a(x_j)\,e^{-ik'x_j}$$

$$a(x_j)=F^{-1}(\hat{a})=\sum_{k=-M}^{M}\hat{a}_k e^{ik'x_j}$$

with $k' = \frac{2\pi k}{L}$, L the length of area in the x-direction and with N=3*M+1 to permit aliasing-free computation of the quadratic terms.

Starting from these spectral equations it is a straight-forward algebraic task to introduce time-stepping schemes, normal-mode initialization, tangent linear equations, the adjoint equations and the adjoints of the time-stepping and initialization schemes. Space does not permit us to present all these equations in detail. We will present some of these equations, that we need for the derivation of the digital filter constraint, in symbolic form. Thus the tangent linear equations are derived by linearization of the full non-linear spectral equations around a particular solution. Let us denote this particular solution by the vector ψ_t containing all spectral coefficients of the wind components and geopotential at time t. The tangent linear equation for small perturbations $\delta\psi_t$ around the solution ψ_t can then symbolically be written:

$$\frac{d\delta\psi_t}{dt} = A_t \delta\psi_t$$

where A_t is a matrix, the elements of which are depending of the non-linear solution at time t.

For derivation of the adjoint of the tangent linear model, it is needed to define an appropriate scalar product. Since the model is formulated in complex Fourier coefficients, the following scalar product is utilized:

$$\langle a,b \rangle = \frac{1}{2}(a^{*T}b + a^{T}b^{*})$$

Here superscript T denotes vector transpose and * denotes complex conjugate. Utilizing this definition of the scalar product, the adjoint of the tangent linear model can be shown to be:

$$\frac{d\delta\varphi_t}{dt} = -A_t^{*T}\delta\varphi_t$$

For comparison with gravity wave control by means of a digital filter as a weak constraint in variational data assimilation, a non-linear normal mode initialization scheme is also available. Time stepping is carried out by a leap-frog scheme and in addition also a time-stepping scheme based on normal mode initialization each time-step is available (Daley 1980). The latter time-stepping scheme allows for a longer time-step and also for an efficient gravity wave control. These features of the model will not be described in further details here.

4. Digital filters and variational data assimilation

There are several possible ways in which digital filters could be applied as weak constraints, as well as hard constraints, in variational data assimilation including the time dimension. For this preliminary study, we will simple require the deviations between filtered and unfiltered forecast states to be as small as possible. This is achieved by adding a quadratic term J_{DF} to the cost function that is being minimized. In addition to this "Digital Filter" term, the cost function includes another quadratic J_{OBS} term describing the deviations between the forecast state and the observations.

$$J = J_{OBS} + J_{DF}$$

where

$$J_{DF} = \alpha_{DF}^2 \sum_{t=t_1}^{t_2} \langle \psi_t - \sum_{k=-N}^{N} h_k \psi_{t+k}, \psi_t - \sum_{k=-N}^{N} h_k \psi_{t+k} \rangle$$

Here α_{DF} is a scaling factor for the digital filter constraint, t_1 to t_2 the time-period over which the digital filter constraint is applied and 2*N+1 the number of time-steps over which the digital filter is computed (the time-span of the filter). Introduce $g_k=h_k$ for $k \neq 0$ and $g_k=h_k-1$ for $k=0$ and we will have

$$J = J_{OBS} + \alpha_{DF}^2 \sum_{t=t_1}^{t_2} \sum_{k=-N}^{N} \sum_{l=-N}^{N} g_k g_l \langle \psi_{t+k}, \psi_{t+l} \rangle$$

Taking the first order variation of J with respect to small perturbations $\delta\psi_t$ we will have

$$\delta J = \delta J_{obs} + \alpha_{DF}^2 \sum_{t=t_1}^{t_2} \sum_{k=-N}^{N} \sum_{l=-N}^{N} g_k g_l [\langle \delta\psi_{t+k}, \psi_{t+l} \rangle + \langle \psi_{t+k}, \delta\psi_{t+l} \rangle]$$

This expression can be re-arranged to the following form

$$\delta J = \delta J_{obs} + \sum_{t=0}^{T} \langle 2\alpha_{DF}^2 \sum_{t'=tmin(t)}^{tmax(t)} g_{t-t'} [\sum_{l=-N}^{N} g_l \psi_{t'+l}], \delta\psi_t \rangle$$

Here T denotes the final time-step of the data-assimilation period, tmin(t) = max(t_1,t-N) and tmax(t) = min(t_2,t+N). From this we can derive the forcing b_t at time t for the adjoint model that is used to calculate the gradient of the norm with respect to the initial conditions ψ_0. This consist of one part for the forcing from observations, b_t^{obs}, and one part for the forcing from the digital filter, b_t^{df} :

$$b_t = b_t^{obs} + b_t^{df} - b_t^{obs} + 2\alpha_{DF}^2 \sum_{t'=tmin(t)}^{tmax(t)} g_{t-t'} \sum_{l=-N}^{N} g_l \psi_{t'+l}$$

For comparison with the weak constraint based on a digital filter, data assimilation experiments will also be carried out based on the following (weak) quadratic constraint on the gravity mode tendencies:

$$J = J_{OBS} + J_{GW} = J_{OBS} + \sum_{t=1}^{T} \langle G(\psi_t - \psi_{t-1}), G(\psi_t - \psi_{t-1}) \rangle$$

where G denotes the matrix for transformation from the spectral model coefficients to the gravity mode coefficients.

The cost function weighing factors α_{DF} and α_{GW} need to be determined. For this preliminary study, coefficients were determined empirically to give a reasonable balance between the aim to fit the observations and the aim to reduce the level of "gravity wave noise". Thus the following values were used: $(\alpha_{DF})^2 = 10.0$ and $(\alpha_{GW})^2 = 0.1$. In addition, it was necessary to scale the variables entering the digital filter constraint to give an equal contribution from mass and wind field information; the geopotentials were scaled by a factor 10^{-4}.

5. Experimental comparison of different gravity wave control methods.

Identical twin observing system simulation experiments (OSSE) were carried out to test the idea of using digital filters as a weak constraint for gravity wave control in variational data assimilation. A low resolution version of the spectral shallow water model, with the following characteristics, was used for the experiments: L_x = length of the model area in the x-direction = 5000 km, M = maximum wave-number = 3, N= number of transform grid-points = 10, $\bar{u}$ = mean zonal wind = 20 ms^{-1}, f_0 = Coriolis parameter = 10^{-4}, $\beta = 10^{-11}$, Δt = time-step for model integrations = 720 s, T = number of time-step for data-assimilation = 100.

The observing system simulation experiments were carried out through the following three steps:
1) Generation of a reference solution by a non-linear normal-mode initialization and a model integration with the leap-frog time-stepping scheme starting from an exponentially shaped initial height-field, with amplitude = 200 m, and winds in geostrophic balance.
2) Generation of simulated observations from the reference solution at random positions and with random observational errors added. 3 wind vector observations and 3 height observations were generated at each of the time-steps 0, 25, 50, 75 and 100. Two different observation sets were produced; a "good" data set with standard-deviations of observational errors = 1.5 ms^{-1} for the wind components and = 15 m for the height; a "poor" data set with the error standard-deviations = 5 ms^{-1} and = 30 m respectively.
3) These simulated observations were assimilated in data assimilation experiments, using different methods for gravity wave control. Each data assimilation experiment was started from a

simplified first guess model state (amplitude of exponential initial wave = 50 m).

A quasi-Newton optimization package (M1GC2) was used (as a black box by the author) to minimize the cost-function during the data assimilation. We will start by presenting results from the "good data" simulations. The decrease of the cost function as a function of the number of the cost function evaluation is plotted in Fig. 1 for data assimilation without gravity wave control (full curve), with non-linear normal mode initialization at the start of the data assimilation period (dashed curve) and with a quadratic constraint on gravity mode tendencies (dotted curve).The observation part of the cost function was normalized by the standard deviations of the observational errors and by the total number of observations. This means that when the fit to the observations is within the standard-deviations of the observational errors, the observation part of the cost function should have a value below 1.0.

In the case of the leap-frog scheme, without any specific gravity mode control applied, the convergence is quite slow. After about 40 iteration the cost function goes below 1.0 and levels off at a value of about 0.65. This means an overfit by the data assimilation, adjusting the solution also to observational errors. Since there is no gravity wave control applied, also gravity waves are free to adjust the solution to the observations. This can be seen in Fig. 2 (a) which shows the reference solution (full line) together with the assimilation solution without gravity wave control (dotted line) at time-step 10. Due to the basic model assumptions, the deviations of the u-component from the mean zonal wind is purely divergent. For the reference solution, this divergent u-component perturbation is very small, as should be expected since the reference run was initialized by a non-linear normal-mode initialization.

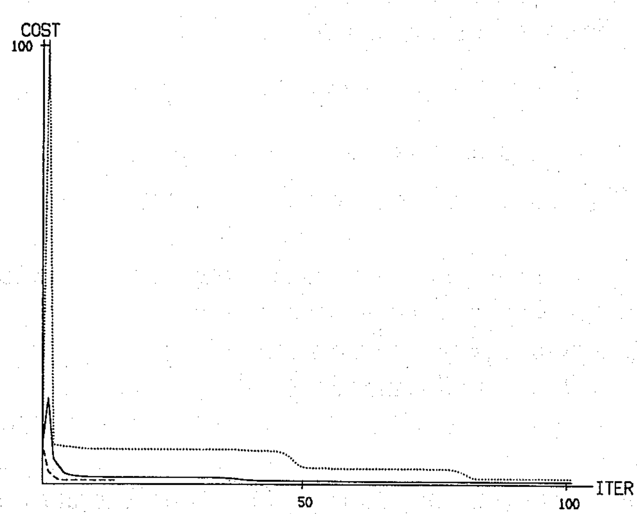

Figure 1: Cost function values as a function of the number of the cost function evaluation. Without gravity wave control (full line), with NLNMI at the start of the data assimilation period (dashed line) and with a weak gravity wave tendency constraint (dotted line).

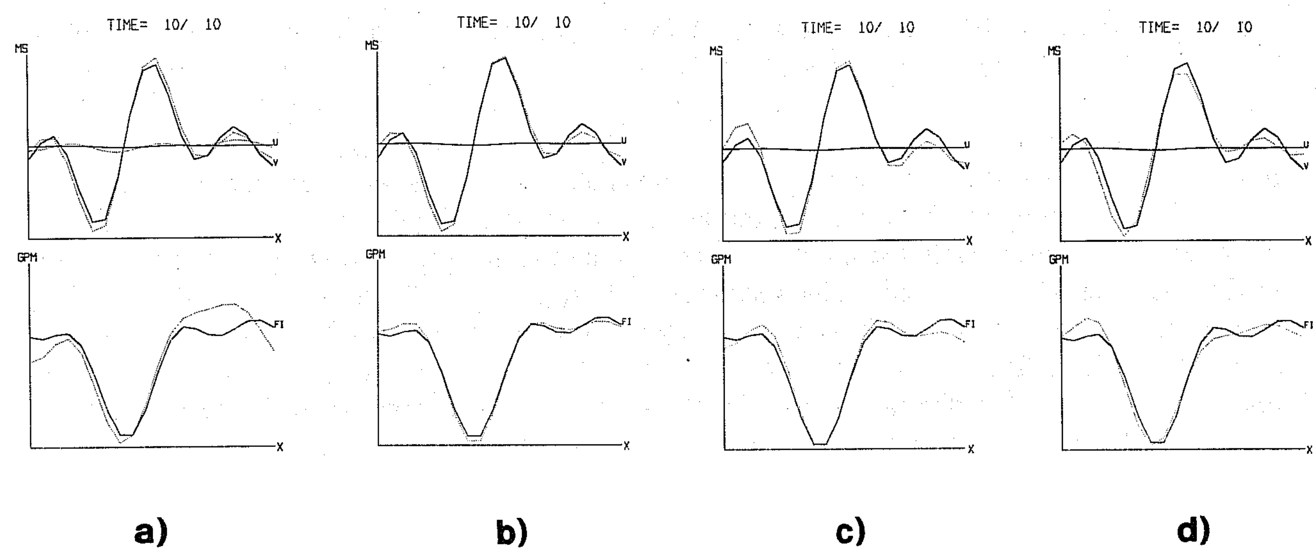

Figure 2: Reference solution (full line in all figures) and data assimilation solutions (dotted lines) using the "good" data set. All solutions given for time-step 10 as deviations of the u-component and the geopotential from the mean state and the full v-component. (a) Without gravity wave control, (b) NLNMI at the start of the data assimilation period, (c) Weak gravity wave tendency constraint and (d) Weak digital filter constraint.

When non-linear normal mode initialization (NLNMI) is applied (as a hard constraint) in the data assimilation, the convergence of the optimization process is fast (dashed line in Fig. 1). After 15 cost function evaluations (11 iterations), the optimization routine stopped at a cost function minimum of approximately 0.85. The level of divergence in the solution is as low as in the reference solution (Figure 2 (b)) and, in general, the assimilation solution is much closer to the reference solution as compared to the case when NLNMI is not applied (figure 2 (a)). The data assimilation solution for the experiment when the normal mode time-stepping is applied (not shown) is very similar to the solution when NLNMI is applied only at the start of the data assimilation period.

When the quadratic constraint on gravity mode tendencies is applied, again the convergence of the optimization process is very slow. After about 85 function evaluations, the cost function levels off at a value of approximately 1.35. It should be noted, however, that a larger final value of the cost function is expected in this experiment, since the cost function also contains a term for control of the gravity wave noise. The final value, at which the cost function levels off, of course depends on the weight α_{DF}^2 given to the new term. As can be seen in Fig. 2 (c), the final solution for the case the gravity wave (GW) constraint is applied, lies in between the no constraint solution and the NLNMI solution.

It is more complicated to obtain convergence of the data assimilation process in the case when the digital filter (DF) constraint is applied. In fact, when the digital filter constraint is applied directly with "full" weight, the cost function levels off at an un-realistically large value with a solution (not shown) that is quite far from the reference solution. However, by a stepwise introduction of the digital filter constraint, a reasonable solution and convergence is obtained. In a fist optimization cycle, the weight of the DF constraint is multiplied by a factor 0.001 and in a second optimization cycle the DF constraint is given full weight. Of course, this is equivalent to the use of an almost un-constrained solution as first guess for the digital filter solution. The convergence curves for the 2 DF optimization cycles are shown in Fig. 3 ("almost unconstrained" solution full line; final DF solution dashed line). The DF solution (Fig. 1 (d)) certainly has managed to reduce the divergence to a level that is about the same as in the GW constraint solution and the solution is much closer to the reference solution than the un-constrained solution (Fig. 2(a)).

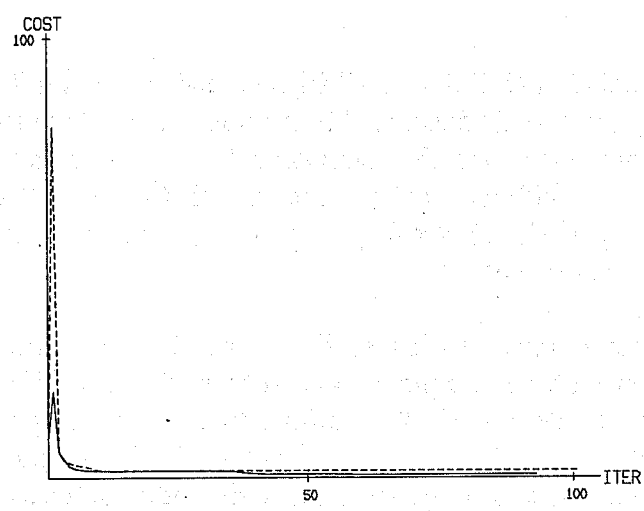

Figure 3: Cost function values as a function of the number of the cost function evaluation. First optimization cycle of the assimilation using the weak digital filter constraint (full line), second optimization cycle of the assimilation using the weak digital filter constraint (dashed line).

As a measure of the relative level of gravity wave noise in the reference run and in the 4 different data assimilation experiments, we present a table below containing the gravity mode u-component tendencies (averaged over the first 10 time-steps of forecasts based on the final results of the data assimilation).

Experiment	du_{grav}/dt (ms^{-2} per 3h)
Reference run	$3*10^{-4}$
No GW constraint	5
Hard NLNMI constraint	$6*10^{-3}$
Weak GW tendency constraint	0.2
Weak DF constraint	1.0

The results from experiments with assimilation of the "poor" data set are qualitatively very similar to the results from the "good" data assimilation experiments. The main difference is a less good fit to the reference solution, due to the larger simulated observational errors. We will therefore restrict ourselves to the presentation in Fig. 4 of the final solutions at time-step 10 of the different experiments.

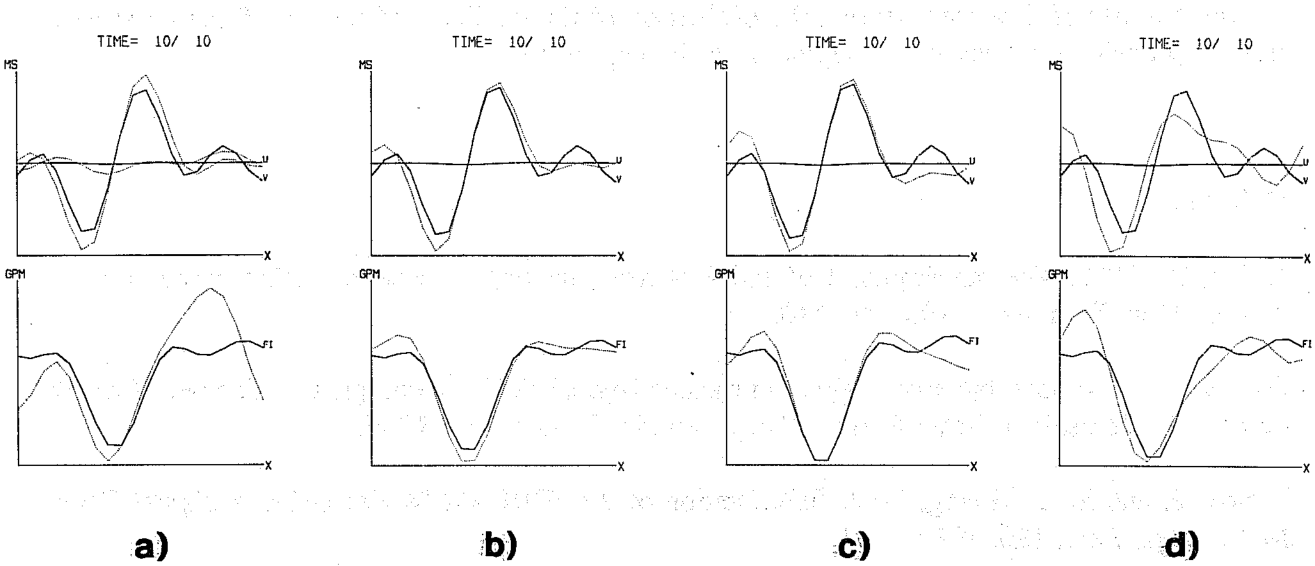

Figure 4: Reference solution (full line in all figures) and data assimilation solutions (dotted lines) using the "poor" data set. All solutions given for time-step 10 as deviations of the u-component and the geopotential from the mean state and the full v-component. (a) Without gravity wave control, (b) NLNMI at the start of the data assimilation period, (c) Weak gravity wave tendency constraint and (d) Weak digital filter constraint.

6. Summary and concluding remarks

A preliminary test of using a digital filter as a weak constraint in variational data assimilation including the time dimension has been carried out. A simple low order one-dimensional spectral shallow water model and its adjoint were used for the experiments. It is certainly too early to conclude from these preliminary experiments, but at least it could be stated that the addition of the digital filter constraint reduces the level of gravity wave noise significantly as compared to data assimilation without any gravity wave control. The convergence of the optimization process turns out to be rather slow, similar to the case when a weak constraint on gravity mode tendencies is used (for comparison reasons). Several aspects of the proposed technique for control of gravity wave noise certainly have to be re-considered, e.g. tuning of the optimization procedure, choice of digital filter, time-span of the filter, cut-off frequency and the time-period for the application of the filter constraint. Also application of digital filters as hard constraints, built into the data-assimilation cycle, need to be considered.

After this preliminary study, it seems appropriate to turn to test the idea in a more complicated model system, e.g. in a 2-dimensional shallow water model using real atmospheric data. The practical application of the weak digital filter constraint in a full scale atmospheric model should be quite straight-forward - the main problem being the need of fast access to model states from a large number of time-steps during the evaluation of the gradients of the cost function at each time-step during the backward integration of the adjoint model.

References

Daley, R., 1980: The development of efficient time integration schemes using model normal modes. Mon. Wea. Rev., 108, 100-110.

Gandin, L. S., 1963: Objective analysis of meteorological fields. (Leningrad: Gidromet). English translation (Jerusalem: Israel Program for Scientific Translation), 1965.

Lynch, P. and X. Y. Huang, 1992: Initialization of the HIRLAM Model Using a Digital Filter. Moth. Wea. Rev., 120, 1019-1034.

Le Dimet, F. and O. Talagrand, 1986: Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects. Tellus, 38A, 97-110.

Lorenc, A., 1986: Analysis methods for numerical weather prediction. Quart. J. Roy. Meteor. Soc., 112, 701-721.

Machenhauer, B., 1977: On the dynamics of gravity oscillations in a shallow water model with applications to normal mode initialization. Beitr. Atmos. Phys., 50, 253-271.