

Diagnostics of linear and  
incremental approximations in  
4D-Var revisited for higher  
resolution analysis

G. Radnóti,  
Y. Trémolet, E. Andersson,  
L. Isaksen, E. Hólm and  
M. Janisková

Research Department

June 2005

This paper has not been published and should be regarded as an Internal Report from ECMWF.  
Permission to quote from it should be obtained from the ECMWF.



**Series: ECMWF Technical Memoranda**

A full list of ECMWF Publications can be found on our web site under:

<http://www.ecmwf.int/publications/>

Contact: library@ecmwf.int

**© Copyright 2005**

European Centre for Medium Range Weather Forecasts  
Shinfield Park, Reading, Berkshire RG2 9AX, England

Literary and scientific copyrights belong to ECMWF and are reserved in all countries. This publication is not to be reprinted or translated in whole or in part without the written permission of the Director. Appropriate non-commercial use will normally be granted under the condition that reference is made to ECMWF.

The information within this publication is given in good faith and considered to be true, but ECMWF accepts no liability for error, omission and for loss or damage arising from its use.

## Abstract

Diagnostic experiments on the validity of the tangent linear approximation, essential in the context of ECMWF 4D-Var, are examined using the most recent versions of the assimilation system. The results are compared with those obtained with a three years earlier model version. The tests are extended to T799 horizontal resolution for the full non-linear model and to T319 for the tangent-linear, in view of planned resolution upgrades to the ECMWF forecast system. The role of various components of the physics is investigated, as is the impact of a new conserving interpolation scheme, and a newly developed grid-point representation for humidity in the tangent-linear model.

Our results show that improved accuracy can be obtained from increased analysis resolution, and in particular that the T799/T255 system is viable, and should be pursued for pre-operational testing. It has been affirmed that a rapid initial growth of error near the surface is present in ECMWF 4D-Var. The tests show that the discrepancies between the inner- and outer-loop resolution, non-linearities in the vertical diffusion and in other physical processes equally contribute to this. Tests with the new conserving interpolation being under development and to be used in the near future within 4D-Var indicate its potential positive impact. The grid-point representation for humidity in the tangent-linear model, according to first tests, seems to better represent humidity increments associated with smaller scale phenomena, such as organized convection.

## 1. Introduction

The accuracy and efficiency of the currently adopted 4D-Var solution algorithm (Courtier et al., 1994) depend on the validity of the tangent-linear assumption. This has been previously examined by Trémolet (2004) (hereafter YT2004) using the January 2002 version<sup>1</sup> of ECMWF's operational 4D-Var data assimilation system (Rabier et al. 2000). Here we revisit the diagnostics of YT2004, using current (spring 2005), improved versions<sup>2</sup> of the 4D-Var system with a view to forthcoming analysis resolution increases, in particular.

A tangent-linear model of atmospheric dynamics and physics, linearized around the current best estimate of the atmospheric state, is used to evolve the analysis increments over the 4D-Var assimilation window (currently 12 hours). Similarly, tangent-linear observation operators are used for the comparison between model and observations. Linearity is a prerequisite for the most efficient iterative minimisation algorithms, such as conjugate gradient (Fisher, 1998). Non-linearities are accounted for by re-linearizing around the state incremented by the output of the linear iterations. This re-linearization can be repeated, thus obtaining the combined 'inner/outer' solution algorithm (Laroche and Gauthier, 1998; Lawless et al., 2005) comprising linear inner loops nested within non-linear outer-loops; see YT2004 and Andersson et al. (2004) for further information on the ECMWF implementation.

In the ECMWF operational system, because of computational cost, the tangent-linear (and adjoint) models are run at a lower resolution than the non-linear model. Furthermore, the physics is simpler in the tangent-linear and adjoint than in the non-linear model. The YT2004 diagnostics measure the errors introduced by these approximations, by comparing the output of the linear model (i.e. an evolved analysis increment) with the finite difference obtained by running the non-linear model twice, with and without adding the analysis increment. YT2004 found that the linearization errors (at that time) were larger than expected, and that large errors appeared very early in the assimilation window. An important conclusion of YT2004 was that higher-resolution 4D-Var will require more accurate linear physics.

In this paper we present results obtained by repeating the YT2004 diagnostic tests with the current versions of the Integrated Forecasting System (IFS), and extending it to experiments at higher resolutions: up to T799 outer-loop and T319 inner-loop resolutions. Our experiments were performed to assess the impact, in this

<sup>1</sup> Labelled IFS cycle 24r3

<sup>2</sup> Labelled IFS cycle 28r4 and 29r1

context, of the model enhancements of the last three years, which include: changes in the non-linear and tangent linear physics (Janisková et al. 2002) observation usage (Thépaut and Andersson, 2003), formulation of the assimilation background term ( $\mathbf{J}_b$ ) (Fisher, 2003), and a new humidity analysis (Hólm et al., 2002). Additionally, we seek further insight into some specific issues that are relevant in preparation for the planned resolution increases to both outer- and inner loops of the operational data assimilation system:

- How do the tangent-linear approximation errors depend on resolution in a range of configurations up to T799 outer loops and T319 inner loops?
- Which inner-loop resolution is required with T799 outer loops to obtain similar relative accuracy as with the currently operational T511/T159 system?
- What are the dominant sources of the fast-growing tangent linear error, near the surface, during the very early stages of forecast (as first identified in YT2004)?
- To what extent do interpolation errors contribute to inner/outer loop inaccuracies?
- What is the impact of the newly developed tangent-linear grid-point representation for humidity, consistent with the non-linear model?

The methodology, similar to YT2004, is outlined in Section 2. In Section 3, we revisit some of the YT2004 tests, using the current version of the IFS. The fast-growing near-surface error is investigated in Section 4, and resolution issues are studied in Section 5. Grid-point humidity and interpolation issues are also investigated in Section 5. Conclusions are given in Section 6.

## 2. Methodology

In this section we recall the methodology of YT2004, which is adopted for the diagnostics in the present study. We evaluate the linear assumption in the conditions that are most relevant for the current incremental 4D-Var: the perturbations we consider are analysis increments, and their low-resolution linear evolution will be evaluated with respect to the operational high-resolution forecast model.

4D-Var seeks to minimize the discrepancy between observations of the atmosphere and a forecast over a period of time called the assimilation window (currently 12 hours, Bouttier 2001). The control variable of the problem is the initial condition of the model, i.e. the atmospheric state at the initial time of the assimilation window. The cost function, which is minimized includes three terms and can be written:

$$J(\mathbf{x}) = (\mathbf{x} - \mathbf{x}_b)^T \mathbf{B}^{-1} (\mathbf{x} - \mathbf{x}_b) + (H(\mathbf{x}) - \mathbf{y})^T \mathbf{R}^{-1} (H(\mathbf{x}) - \mathbf{y}) + J_c$$

where  $\mathbf{x}$  is the control variable,  $\mathbf{x}_b$  is the background state,  $\mathbf{y}$  is the vector of observations,  $\mathbf{B}$  is the background error covariance matrix,  $\mathbf{R}$  is the observation error covariance matrix,  $H$  is the nonlinear observation operator and  $J_c$  is an initialization term used to control gravity waves. The latter term will be omitted in the remaining of this paper for simplicity.  $H$  computes the observation equivalent at the correct location and time and includes the forecast model in our notation. In its incremental formulation (Courtier et al. 1994), the minimization problem is written as a function of the departure from the background:  $\delta\mathbf{x} = \mathbf{x} - \mathbf{x}_b$ . At the minimum,  $\delta\mathbf{x}$  will be the analysis increment. A first-order approximation of the cost function is given by:

$$J(\delta\mathbf{x}) = \delta\mathbf{x}^T \mathbf{B}^{-1} \delta\mathbf{x} + (\mathbf{H}\delta\mathbf{x} - \mathbf{d})^T \mathbf{R}^{-1} (\mathbf{H}\delta\mathbf{x} - \mathbf{d})$$

where  $\mathbf{H}$  is the linearized observation operator and  $\mathbf{d} = \mathbf{y} - H(\mathbf{x}_b)$  is the departure from observations. In this notation, the tangent-linear model is considered part of the linearized observation operator. The incremental 4D-Var algorithm is shown schematically in Figure 1. The minimization problem is solved using an iterative algorithm (conjugate gradient or quasi-Newton algorithms). This is the inner loop of 4D-Var. In order to reduce the computational cost of the assimilation, the inner loop is run at lower resolution than the forecast. However, in order to retain the maximum information from the observations, the departures are computed at high resolution. The starting point for the minimization (background) is interpolated to the inner-loop resolution using an operator  $S$ . After the minimization, the analysis increments at low resolution are projected onto the high-resolution space, and added to the reference state. Then the high-resolution trajectory and the departures are recomputed. This is the outer loop of incremental 4D-Var. The high-resolution nonlinear runs also define the state around which the observation operator is linearized, after interpolation to the inner-loop resolution. Then the cost function to be minimized in a given inner loop of a 4D-Var with several outer loop iterations can be given by:

$$J(\delta \mathbf{x}_i) = (\delta \mathbf{x}_i + S \mathbf{x}_{i-1} - S \mathbf{x}_b)^T \mathbf{B}^{-1} (\delta \mathbf{x}_i + S \mathbf{x}_{i-1} - S \mathbf{x}_b) + (\mathbf{H}_{i-1} \delta \mathbf{x}_i - \mathbf{d}_{i-1})^T \mathbf{R}^{-1} (\mathbf{H}_{i-1} \delta \mathbf{x}_i - \mathbf{d}_{i-1})$$

$$\mathbf{x}_i = \mathbf{x}_{i-1} + S^{-1} \delta \mathbf{x}_i$$

$$\mathbf{d}_i = \mathbf{y} - H \mathbf{x}_i$$

where  $i$  is the outer loop index,  $S$  is an interpolation operator from high to low resolution, its “pseudo-inverse” is an interpolator from low to high resolution, the states  $\mathbf{x}_i, \mathbf{x}_b$  are defined on the resolution of the

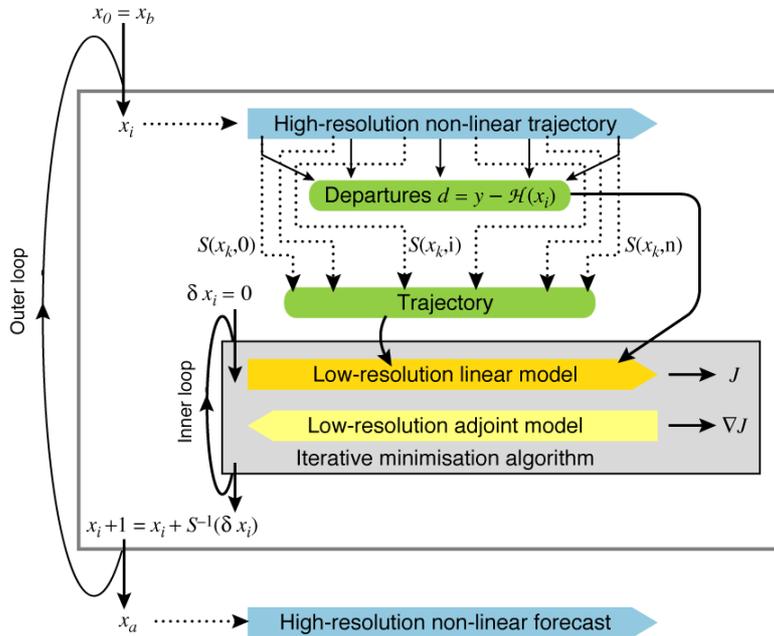


Figure 1: Incremental 4D-Var algorithm: The non-linear trajectory and departures ( $d$ ) from the observations ( $y$ ) are computed at high resolution. The first guess ( $x_0$ ) and the nonlinear trajectory are interpolated to low resolution using the operator  $S$ . The cost function is minimized at low resolution using an iterative algorithm (inner loop). The resulting increment  $\delta \mathbf{x}$  is interpolated back to high resolution (symbolized by  $S^{-1}$ ) and added to the current state  $x_i$ . The process is repeated (outer loop, subscript  $i$ , currently two iterations at ECMWF) until the analysis  $x_a$  is obtained.

outer loop, while the increment  $\delta\mathbf{x}_i$  is given on the resolution of the inner loop. Currently at ECMWF, two iterations of the outer loop are run. For further reduction of the computational cost, the linear physics is omitted in the first inner-loop minimization. The first inner-loop minimisation is run at T95 resolution, and the second at T159, following the multi-resolution incremental approach described by Veersé and Thépaut (1998).

Tangent linear models are usually validated by comparing the output of a linear run with the finite difference between two nonlinear runs of the corresponding model using the ratio:

$$r_k = \frac{\mathbf{M}_k \delta\mathbf{x}}{M_k(\mathbf{x} + \delta\mathbf{x}) - M_k(\mathbf{x})}$$

Where  $M$  is the full nonlinear model,  $\mathbf{M}$  is the tangent linear model linearized around the trajectory of model states and  $r$  is evaluated for each component ( $k=1;\dots;N$ ,  $N$  being the dimension of the model state vector) of the model output. According to the Taylor formula, when the size of the perturbation tends to zero, the finite difference and the linear model should behave similarly. In practice, this is limited by machine precision. The ratio between the output of the linear model and the value of the finite difference linearly approaches 1 and then diverges when machine precision is reached. This test should be true for any perturbation and in practice a set of random perturbations are used. In order for this test to be valid, both models are run at the same resolution, with the same physical processes included and the same values for all parameters.

However, in operational data assimilation, these conditions are not satisfied. We have seen that for computational cost reasons the minimization is run at lower resolution than the forecast, and the linear model does not contain all the physical processes that are represented in the full forecast model. Furthermore, in the current ECMWF system, humidity is a spectral variable in the linear model, whereas the nonlinear model uses a grid-point representation. The option to run the tangent-linear model with grid-point humidity has been recently developed, and we assess its impact in Section 5. When adding the low-resolution increment to the first guess, the operator  $S^{-1}$  is not in practice the inverse of  $S$  but only a pseudo-inverse:  $S$  being a truncation operator it is not invertible. Furthermore, a super-saturation check is applied to the updated state vector after the increment has been added. Finally, the perturbation is not arbitrary in size or direction: it is an analysis increment.

In order to diagnose the resulting errors in data assimilation, the output of the linear model used in ECMWF 4D-Var assimilation system and the difference between two runs of the forecast model are compared in this paper. An analysis increment will be used as initial perturbation. The typical maximum amplitude of the perturbation will be of the order of 3 K and 12 m/s.

4D-Var is an iterative process in which several integrations of both the linear and nonlinear models are performed in the inner and outer loops respectively. Consequently, all the necessary information to perform a test of the linearization is naturally available. Using the approach of YT2004, the output of the last integration of the linear model in a given minimization is saved as well as the output of the two high-resolution forecasts surrounding it. The difference between the two high-resolution runs is then compared with the output of the linear run. This approach has two advantages: the added computational cost is negligible when data assimilation is running and most importantly, the linear and non-linear models are run exactly as used in data assimilation which could be difficult to ensure in any other way.

In the following sections, the relative error

$$r = \frac{\|M(\mathbf{x}_i + S^{-1}\delta\mathbf{x}_i) - M(\mathbf{x}_i) - \mathbf{M}_i(\delta\mathbf{x}_i)\|}{\|M(\mathbf{x}_i + S^{-1}\delta\mathbf{x}_i) - M(\mathbf{x}_i)\|}$$

will be presented where  $\mathbf{x}_i$  is the first guess,  $\delta\mathbf{x}_i$  the analysis increment,  $M$  the nonlinear forecast model,  $\mathbf{M}_i$  the tangent linear model linearized around  $\mathbf{x}_i$  and  $S^{-1}$  is the pseudo-inverse of the truncation operator.  $r$  is computed for each field and each level. In this study the nonlinear model was run at resolutions up to T799, which is the horizontal resolution of the next proposed upgrade of the ECMWF operational forecasting system. The linear model is run at a range of resolutions up to T319. The diagnostics presented here are globally averaged RMS errors computed in grid-point space on the model's reduced Gaussian grid. Both models were run with the current operational 60-level vertical resolution.

In most of the experiments presented below, we use a single outer-loop 4D-Var. This setting is different from the operational one that presently uses two outer-loop iterations. However, this change is appropriate as it provides a single increment, its linear evolution and the corresponding evolution of non-linear model finite differences, from which absolute and relative errors that are characteristic of the 4D-Var procedure can easily be computed.

### 3. Reproduction of the Trémolet 2004 diagnostics using a current version of the forecast system

At first, it is important to make sure that two successive nonlinear trajectory runs produce identically the same forecast when the analysis increment is zero. YT2004 had shown that to obtain this trivial result requires deactivating saturation check and conversion to virtual temperature in the final of the two trajectory runs. We show the first 3 hours only, as from 3 hours of trajectory integration onward an additional discrepancy between the initial and final trajectory is due to interaction through coupling with the ocean wave-model. The differences are shown here in Figure 2, which can be compared to Figure 2 of YT2004.

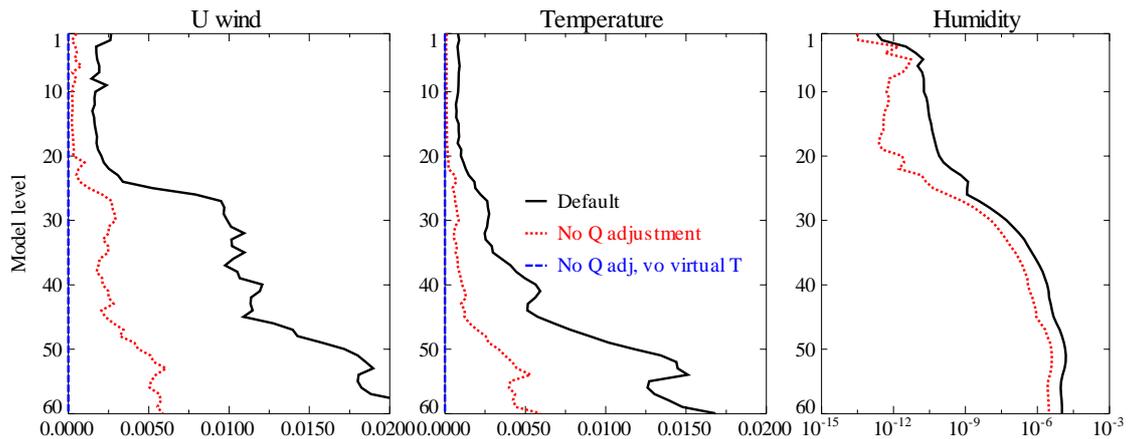


Figure 2: Average difference between successive T159 nonlinear trajectories after 3h when increments are set to zero.

When the two identified items deactivated (blue dashed curves) the two trajectory runs are identical. However, comparison with YT2004 shows that when these two discrepancies are present (black and red curves, respectively), the resulting errors are orders of magnitude larger than in YT2004. This suggests that

small initial perturbations (in the present case initial temperature and humidity differences due to saturation check and virtual temperature conversion) are growing much more rapidly in the current version of the non-linear model.

If the resolutions in the tangent linear and non-linear integrations were the same and the abovementioned discrepancies were not present in the settings of the subsequent non-linear integrations, initial linearity errors would be zero for any increment. However, mostly due to the resolution differences, these errors occur already at the beginning of the 4D-Var window. Figure 3 shows these initial relative errors for different inner-loop resolutions (as Figure 3 in YT2004). One would expect that when the inner-loop resolution is closer to that of the outer-loop these initial errors are smaller. However, except for humidity, these figures show the opposite behaviour. This is due to the way the diagnostics were computed: the comparison of tangent linear increments and finite differences of non-linear trajectories in each case was done on the inner-loop resolution of the given experiment. This can make the comparison between different inner-loop resolution experiments misleading. When the diagnostics are computed on a fixed resolution (we chose T255 in the present tests) independently of the resolution of the given experiment, this discrepancy disappears and the expected results (Figure 4) are obtained.

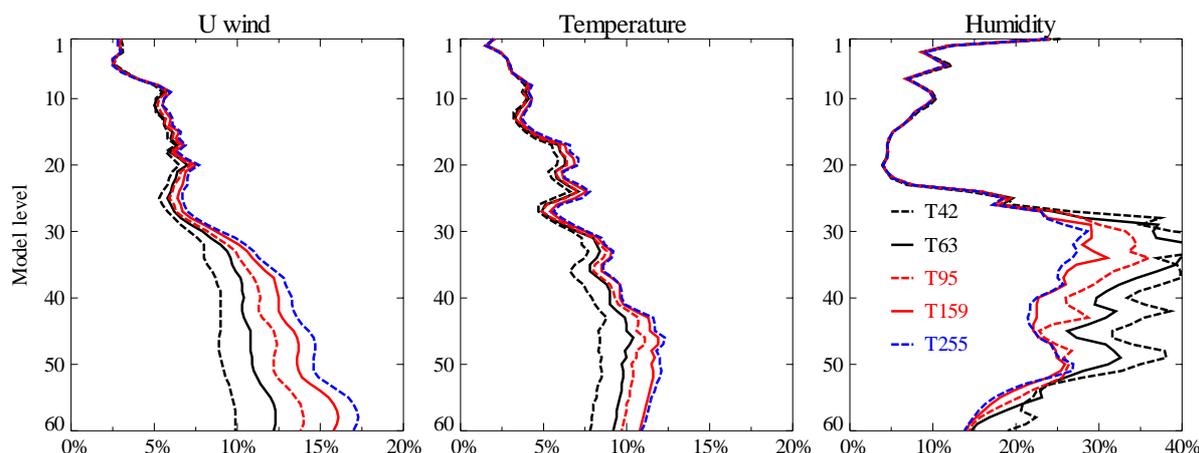


Figure 3: Relative error in the initial condition for several inner loop resolutions as indicated by the legend compared to T511 outer loop

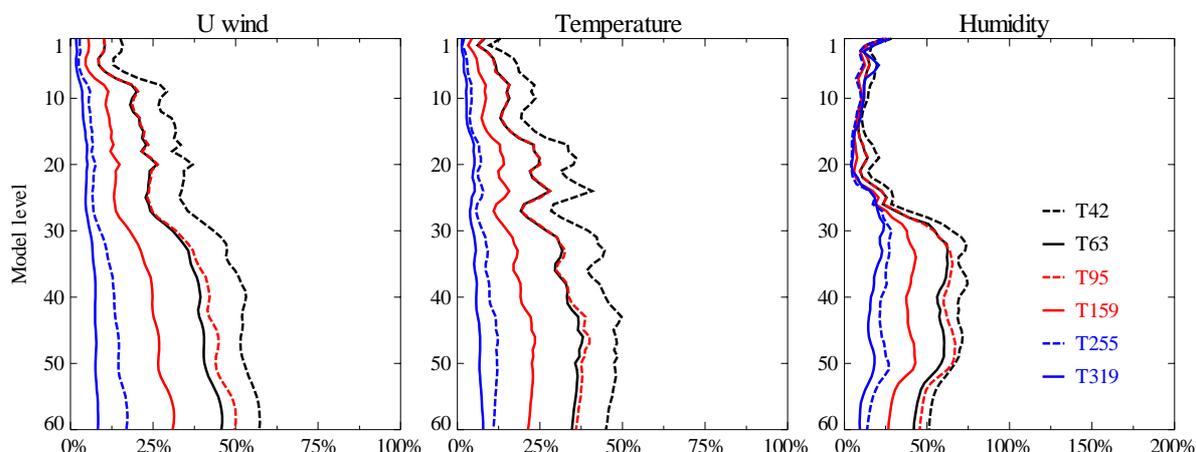


Figure 4: Same as Figure 3, but diagnostics are for all experiments evaluated on the T255 resolution grid.

Figure 5 shows the evolution of the relative errors within the 12-hour assimilation window (counterpart of Figure 4 of YT2004), for the T511/T159 experiment. Curves are shown for 0, 1, 3, 6, 9 and 12 hours.

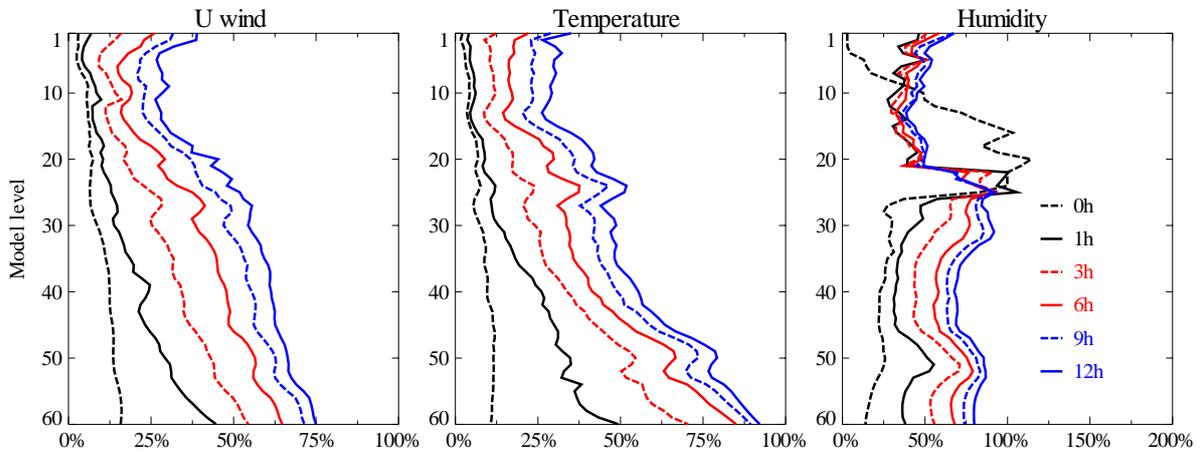


Figure 5: Evolution of the relative error in the T159 tangent linear model with respect to the T511 forecast model over the length of the assimilation window

Comparison shows that the relative errors evolve in the same way as in YT2004, though lower-level wind errors are somewhat larger now. As in YT2004, very rapid error growth occurs already in the first hour of integration, particularly for temperature and wind near the surface. This feature will be the subject of the next Section. Here we note that the initial-time (black dashed) stratospheric humidity errors seem to be erroneous. In the present version of humidity analysis the stratosphere is practically not analyzed: stratospheric humidity background error variances are artificially reduced by two orders of magnitude to avoid spreading humidity increments from the humidity data rich troposphere across the tropopause to the humidity data sparse stratosphere. Therefore, what we see as initial humidity relative errors on the upper levels is purely noise. To verify this the above experiment was rerun without reducing the humidity background-error variances (left panel of Figure 6). Indeed, most of the noise disappeared, but a secondary peak approximately at the tropopause level remained.

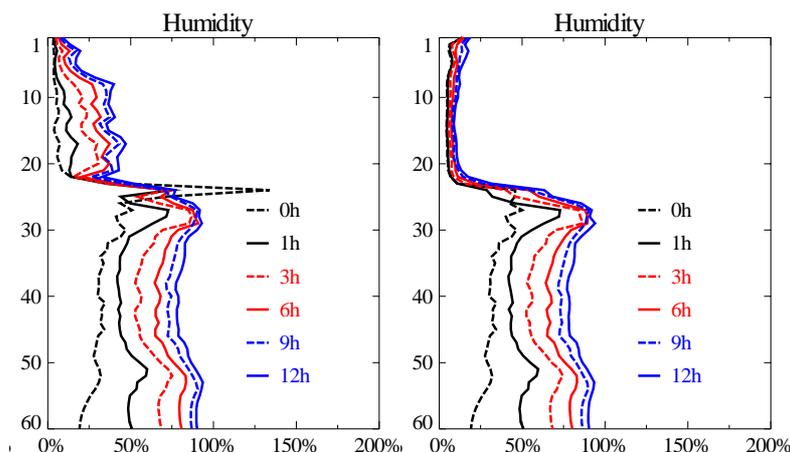


Figure 6: Evolution of humidity relative error in the T159 tangent linear model with respect to the T511 forecast mode. Left panel: non-reduced stratospheric humidity background error variances. Right panel: non-reduced stratospheric humidity background error variances + RH-based control variable in the stratosphere.

Further investigations showed that this secondary peak is the consequence of the abrupt change from relative humidity (RH) based control variable to a specific humidity ( $q$ ) based one at the tropopause level, which is used in the Hólm et al. (2002) formulation of the humidity analysis. To demonstrate that this is indeed the explanation, we switched to using the RH based control variable also in the stratosphere (right panel of Figure 6), and we can see the secondary peak then disappears. This result suggests that the introduction of a smooth vertical transition between RH- and  $q$ -based control variables would be preferable to an abrupt one, even if humidity analysis increments were no longer totally suppressed in the stratosphere. To test this, the humidity error variances and the temperature part of the humidity control variable were tapered off smoothly by multiplying by (0.1, 0.2, 0.5, 0.8, 0.9) for the first five levels below the diagnosed tropopause, and in the stratosphere, the term was set to a very small number ( $1E-9$ ). This configuration was later implemented in CY29R2 of the IFS. To eliminate this noise from humidity most of the experiments presented in the following use this modified humidity analysis setting.

#### 4. On the rapid initial growth of error near the surface

In Figure 5, we showed a very rapid initial error growth in the lowest 10 to 15 levels of the model, especially for temperature, but also for winds. This was observed in YT2004 too. It is the dominant feature of Figure 5 and requires closer examination. First, we show maps of the initial-time increment (Figure 7, top panel), and the increment evolved for one hour by the non-linear (middle) and tangent-linear model (bottom panel), respectively, for lowest-level temperature.

In this case, we have run a 4D-Var experiment with two iterations of the outer-loop. The increment fields are taken from the second minimization while the finite differences are correspondingly computed from the final and the penultimate non-linear trajectory runs. We do this because in this case the rapid initial growth of temperature errors is even stronger than in a single iteration 4D-Var. The first impression from Figure 7 is that the evolution of the increments is quite similar; though some increment patterns (e.g. positive increments appearing as red spots over South-America, or Greenland) develop in the finite difference fields (middle) without any trace in the tangent linear model results (bottom panel). The fields shown here have been interpolated to the T255 grid for comparison. Even so, the increments obtained from the high resolution (T511) non-linear trajectories show much smaller scale details of the increment patterns already after 1 hour of integration. These small scale patterns will typically show up as large relative errors in the RMSE like metric we use in the diagnostics of this paper. This should be kept in mind when these results are evaluated. It is also to be noted, that the non-linear differences seem to evolve faster than the increments in the tangent linear model: the 1-hour increment field of the tangent linear model is at several geographical locations more similar to the initial finite difference than to the 1-hour one.

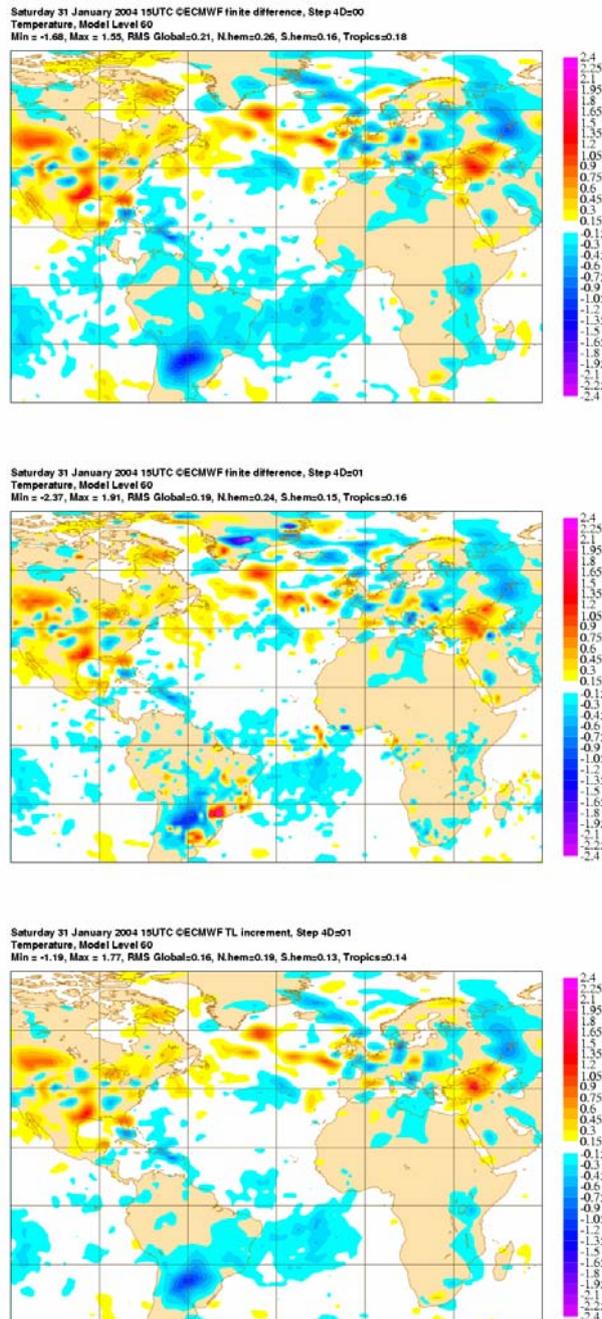


Figure 7: Initial (top) and 1 hour (middle) temperature increment fields on the lowest model level (at level 60) as evolving in the non-linear model and 1 hour increment of the tangent linear model (bottom). All computed during the second iteration of T511/T159 4D-Var

As shown in Figure 5 of YT2004, and confirmed with the present model version (not shown), the rapid initial error growth is not present in an adiabatic experiment where both the non-linear trajectory runs and the minimization are performed without physics. Thus the question may naturally arise: what parts of the physical parameterization, i.e. pairs of non-linear and corresponding tangent linear physical parameterizations, are mostly responsible for this lack of linearity already in the early parts of model integration? To answer this question we have performed “partially diabatic” experiments. However, it is not

always entirely straightforward or meaningful because of the interactions between different physical processes. Here we present some results of our attempts within the framework of 4D-Var to separate different physical processes. Figure 8 shows the relative errors after 1 hour in 4D-Var experiments for different settings of the physics.

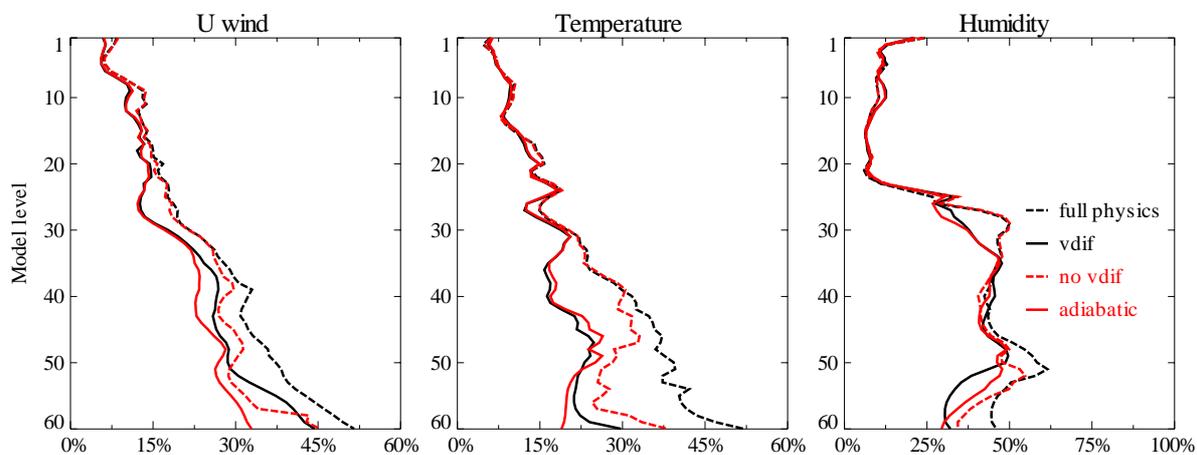


Figure 8: Relative error of the tangent linear model for various settings of physics with respect to T511 nonlinear, diabatic model after 1h

In the experiment denoted “vdif” only vertical diffusion was activated, while in the “no vdif” experiment all the physical parameterization was activated like in the operational model, except for vertical diffusion, which was switched off. The non-linear model physics was in each experiment set consistent with the corresponding tangent-linear physics. The results broadly indicate that both vertical diffusion and the rest of physics contribute nearly equally to the rapid error growth of wind and temperature and to the large vertical gradient of errors in the lowermost model levels. Based on these results it will be important to test the tangent linear vertical diffusion with respect to its own simplified non-linear scheme, but in the present version of the model code it is hardly possible. The work that is presently undergoing to separate the tangent linear and adjoint vertical diffusion code from the full non-linear version will make these tests available in the near future. In addition to the above tests, the experiments were also performed (not shown) with and without surface parameterization within the non-linear trajectory run, but the presence or absence of evolving surface forcing had no impact on the above results, probably because within the time range of an hour the change of surface fields can be neglected.

To see if interpolation errors, due to the resolution difference, might contribute to the rapid initial growth of error we investigate how the errors evolve in time when inner- and outer-loop resolutions are identical. In this case, no interpolation is performed between inner- and outer-loop. Thus, Figure 9 shows the evolution of relative errors in a T159/159 4D-Var.

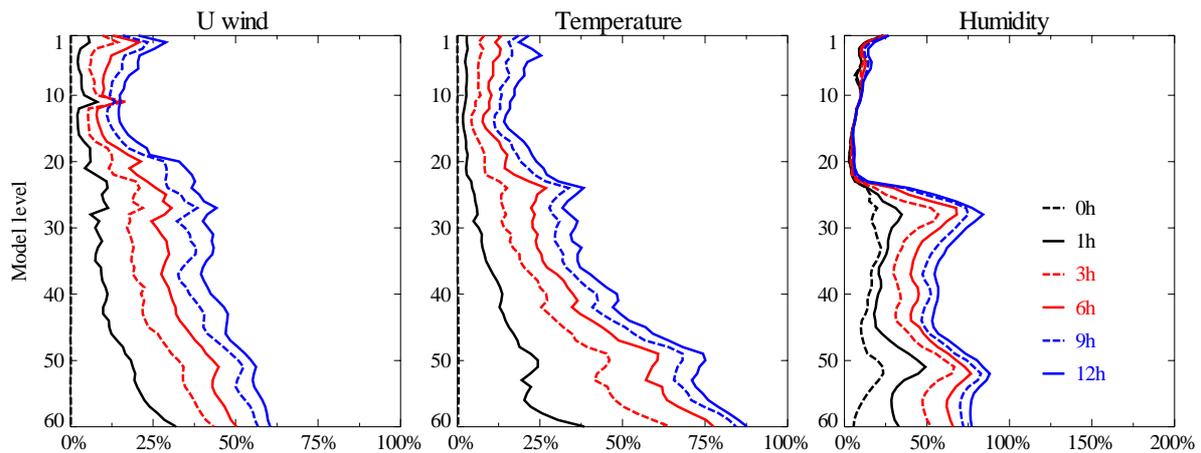


Figure 9: Evolution of the relative error in the T159 tangent linear model with respect to the T159 forecast model over the length of the assimilation window

In this case initial-time errors (black dashed lines) for temperature and wind are very small, but they are non-negligible for humidity, due the saturation check applied at the beginning of the final trajectory run (as mentioned in Section 3). Comparing the error evolutions on Figure 5 and Figure 9 it can be seen that the relative errors for the early hours of the 4D-Var window are significantly smaller when the two resolutions are the same. However, after 12 hours temperature relative errors are almost the same in the two experiments. Consequently, the resolution difference of the inner- and outer-loop is an important factor in the rapid initial temperature error growth near the surface.

In the present operational 4D-Var the tangent linear model is linearized around the non-linear trajectory, interpolated to the lower resolution at one-hourly frequency. It is worth verifying whether or not this update frequency has a significant impact on the accuracy of the tangent linear model, especially in the first hour where the rapid error growth has been seen. At the same time, we wanted to see how the rapid initial error growth is distributed within the first hour of the assimilation window and if it appears in a same way at a Eulerian and at a semi-lagrangian 4D-Var. To check all this a Eulerian 4D-Var experiment has been performed with a 90 sec time-step in the nonlinear model, 3 minutes time-step in the minimization and 15 minutes trajectory update frequency. In this experiment, quadratic truncation T341 has been used in the outer-loop and T106 in the inner-loop (the quadratic counterpart of the present operational T511/159 linear truncations). Figure 10 shows the error evolution within the first hour of the 4D-Var window. Wind, temperature and surface pressure errors grow gradually, reaching at one hour the same level of error as in the standard semi-lagrangian 4D-Var with hourly trajectory update (1-hour curve of Figure 5). The lowermost level wind errors grow suddenly already in the very first time-step. All this suggests that neither the trajectory update frequency nor the choice of the advection scheme within the 4D-Var algorithm play an important role in the error evolution and that the rapid initial error growth is not related only to the very first time-step, but it is rather a gradual process during the whole one-hour initial integration period.

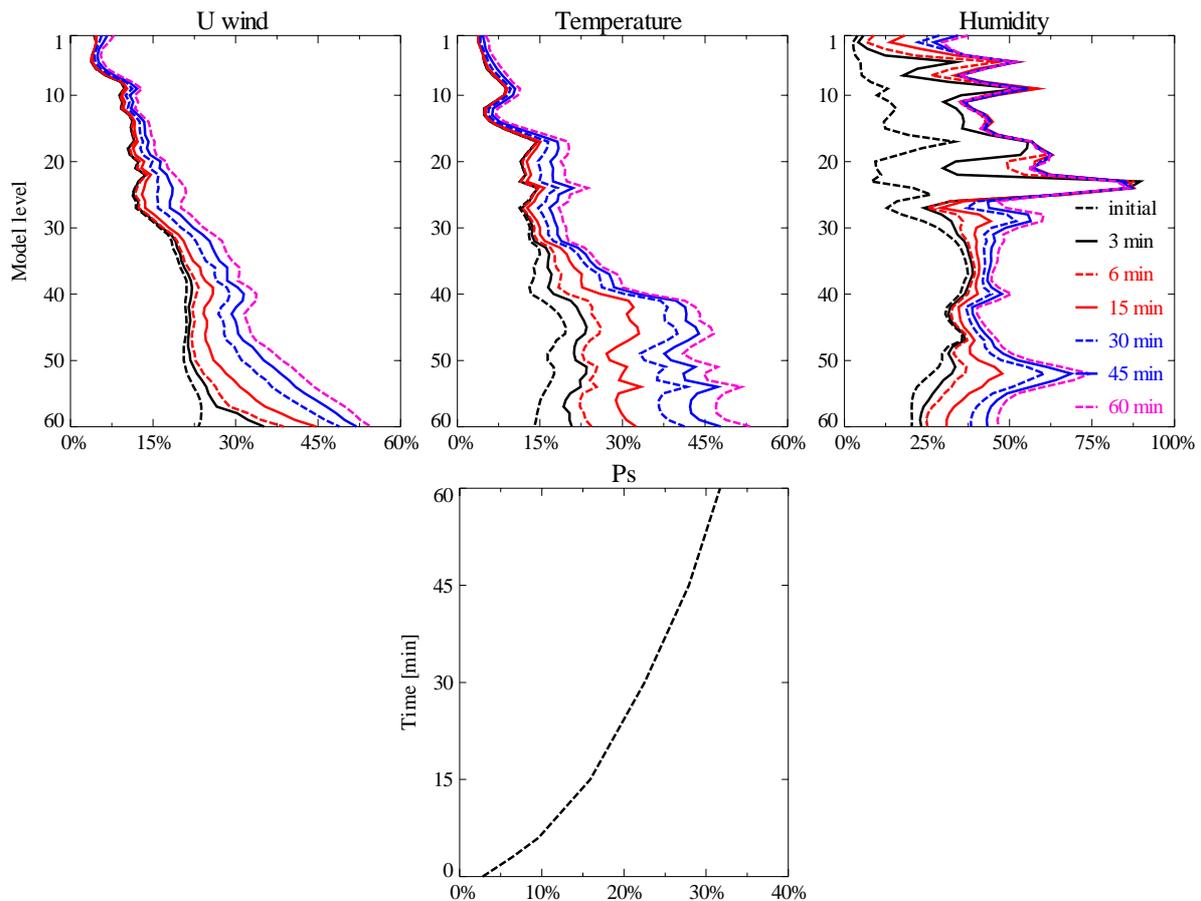


Figure 10: Evolution of the relative error in the T106 tangent linear model with respect to the T341 forecast model over the first hour of the assimilation window within a eulerian 4D-Var experiment

Based on these investigations we conclude that the rapid initial error growth is mostly related to the discrepancy between inner- and outer-loop resolutions and to the non-linearities in the full model. These discrepancies have a significant impact already within the time range of an hour of integration.

## 5. Towards higher resolutions

### 5.1 Sensitivity to inner-loop resolution

Another subject to be studied is the relative error in 12-hour evolved increments for inner-loop resolutions ranging from T42 to T319. The results are shown in Figure 11 similarly to Figure 3 above in Section 3, and should be compared to Figure 7 in YT2004. Additionally, the fourth panel in Figure 11 shows the evolution of surface pressure errors as a function of time within the 4D-Var window. A significant improvement with respect to YT2004 is to be emphasized. In YT2004 the inner-loop resolution increase from T159 to T255 did not bring any further benefit in terms of the accuracy of the tangent linear integration. In our present experiments, however, a gradual improvement with the resolution increase continues across the whole range of inner-loop resolutions, even up to T319. This very promising result is the positive outcome of a combination of the developments that took place in the last three years, notably in the tangent-linear physics. It is conceivable that changes in the structure and reduced amplitude of the increments themselves, following changes in the background-error formulation and observation data usage, may have contributed to the result. That is why it is important to do these tests inside the 4D-Var configuration.

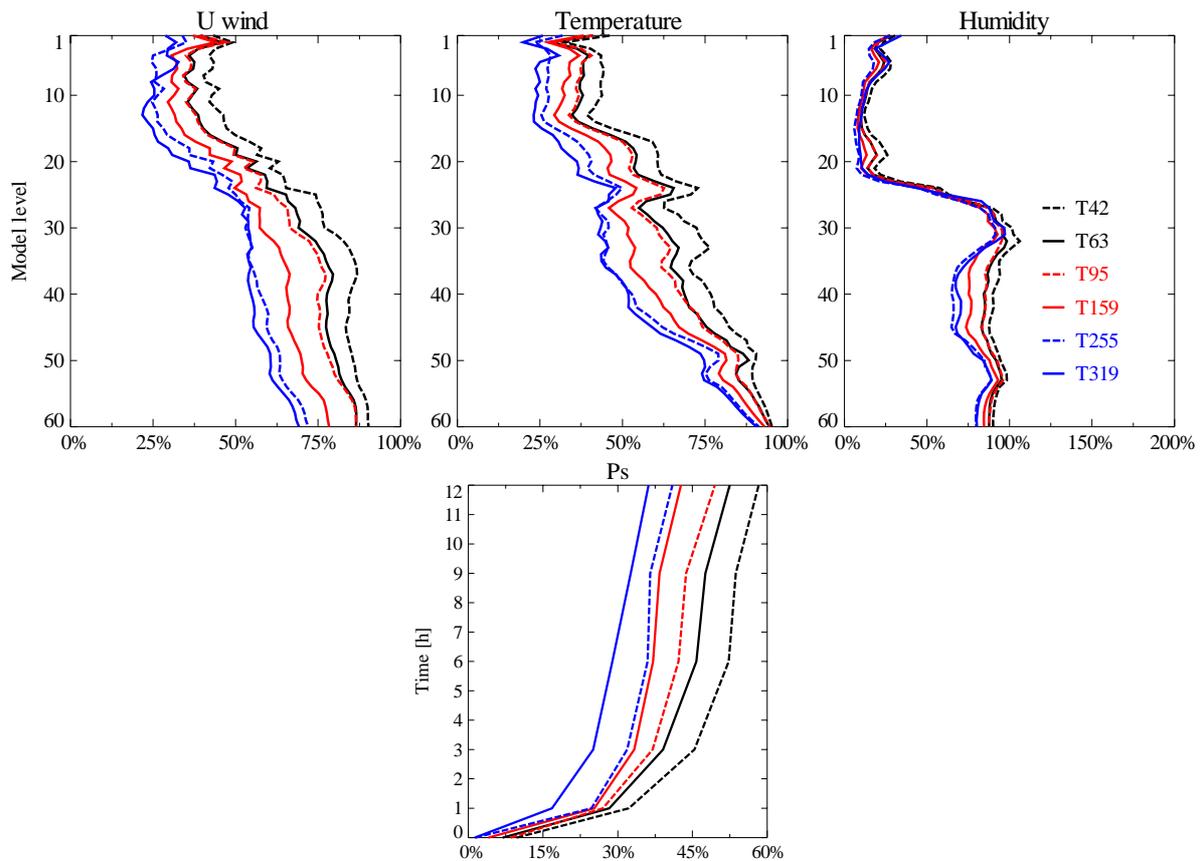


Figure 11: Relative error of the tangent linear model for various resolutions with respect to T511 nonlinear, diabatic model after 12h for the 3 dimensional variables and for the whole 4D-Var window for surface pressure. Diagnostics are computed on the T255 resolution grid.

## 5.2 Higher outer-loop resolution

It is planned that the next version of the ECMWF forecasting system will run at T799 resolution. We have therefore performed some tests with T799 outer-loop resolution. Figure 12 shows the evolution of the relative error (like Figure 5) for a T799/T159 experiment, in CY29R1.

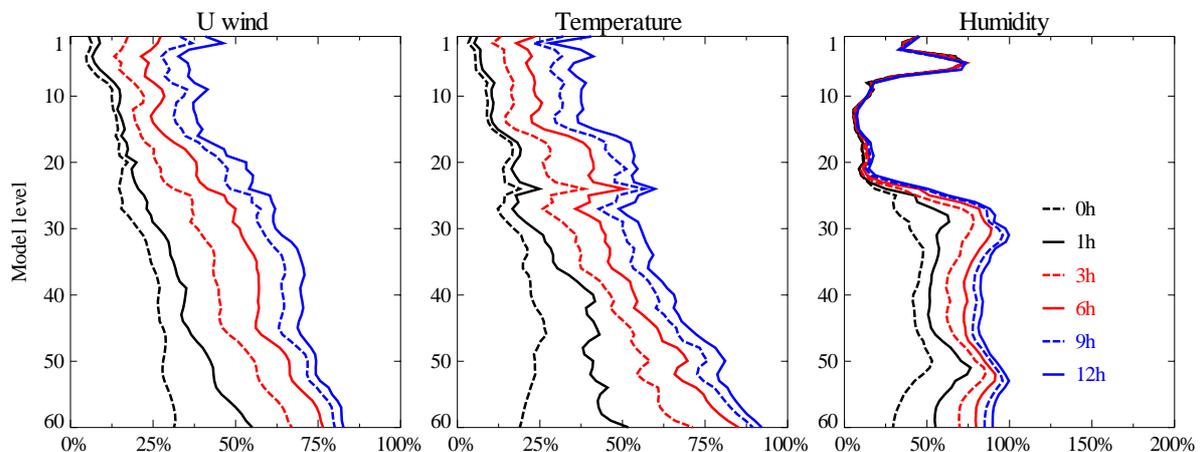


Figure 12: Evolution of the relative error in the T159 tangent linear model with respect to the T799 forecast model over the length of the assimilation window, cy29r1.

The relative errors evolve similarly to the T511/T159 case, but the corresponding relative errors are slightly larger here due to the larger difference between the inner- and outer-loop resolutions (the difference in stratospheric humidity error behaviour is due to different humidity analysis settings, as explained in Section 3 of this paper). As previously noted by Bouttier (2001) it appears that the inner-loop resolution ought to be increased at the same time as the outer-loop resolution, in order to retain a similar level of relative errors.

The question arises: What inner-loop resolution is required in the T799 system, to retain or improve the relative error with respect to the current T511/T159 system? To answer this question we have compared 12-hour relative errors (Figure 13) for fixed T799 outer-loop resolution, but varying the inner-loop resolution from T95, to T159 and T255 (see legend). As a reference, we show the presently operational T511/159 system (black dashed). The corresponding time-steps are 720 sec for the T799, 900 sec for the T511 non-linear model integration and 1800 sec for all the tangent linear model integrations.

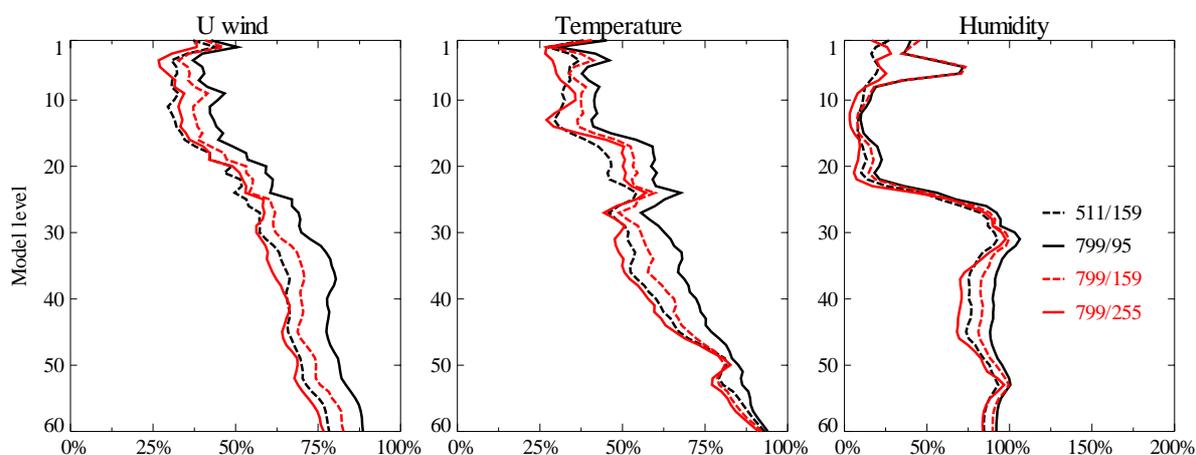


Figure 13: Relative error of the tangent linear model with respect to the T799 nonlinear, diabatic model after 12 hours for various inner-loop resolutions (see legend). The T511/T159 system is shown as reference (black dashed lines). Diagnostics are computed on the T255 resolution grid.

As the results confirm, the magnitude of relative errors is largely determined by the ratio of inner- and outer-loop resolution. The errors of the T799/T255 and the presently operational T511/159 resolution combinations are quite comparable, for humidity and for tropospheric wind and temperature the T799/255 relative errors are even smaller than those of T511/159.

### 5.3 Interpolation

As displayed on Figure 11, the tangent-linear model is more accurate at higher resolution. From that point of view, there is thus scope to increase the analysis resolution beyond the current T159. This is affordable within the adopted multi-resolution incremental approach discussed above. We have also seen that considerable error is introduced from the interpolations between inner and outer loop resolutions. It is crucial that these errors are reduced when going to higher and higher resolution in the data assimilation. A new conserving higher-order interpolation scheme has therefore recently been developed which, when fully tested, will replace the current interpolation techniques: simple spectral truncation, its pseudo-inverse or polynomial interpolation. The conserving interpolation does a spatial integration of the original field over each grid cell at the target resolution. The fields that are integrated need to be in mass units, and this is achieved by multiplying mixing ratios, wind, and temperature by the pressure thickness of each layer. The final step of the interpolation is to divide by the pressure thickness of the output grid. This interpolation is also applied to surface pressure and all surface fields. The accuracy of the interpolation is determined by the

assumed order of the polynomial describing the sub-cell distribution of the field within each grid-cell. For higher order interpolations, a nonlinear limiter is under development to avoid over and undershoots. The chosen interpolation plays an important role within the 4D-Var algorithm due to the multi-resolution incremental approach, and it has a visible impact on the diagnostic calculations presented in this paper as well. We have tested the sensitivity of the resulting relative errors by applying different interpolation options in the diagnostic computations (keeping for now the present interpolation method within 4D-Var). Figure 14 shows the evolution of temperature relative errors with different interpolation options within the diagnostics.

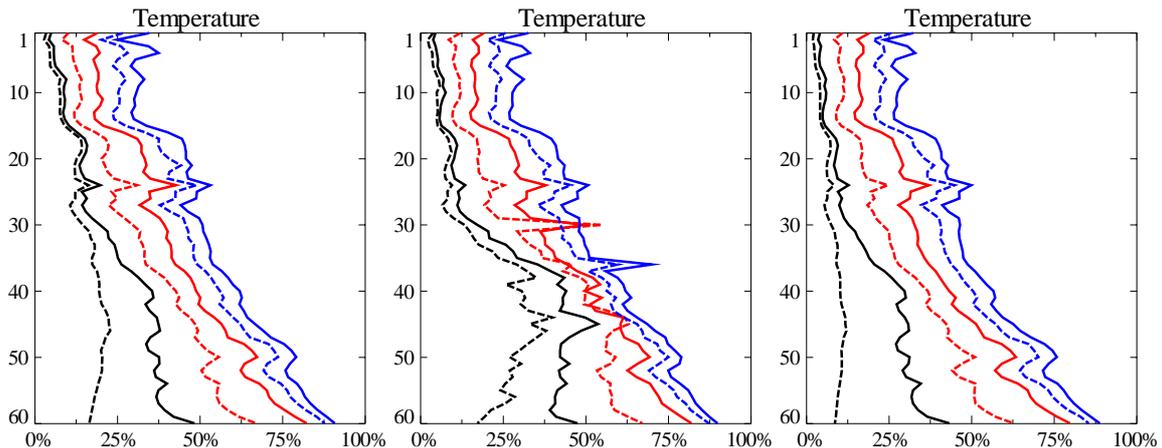


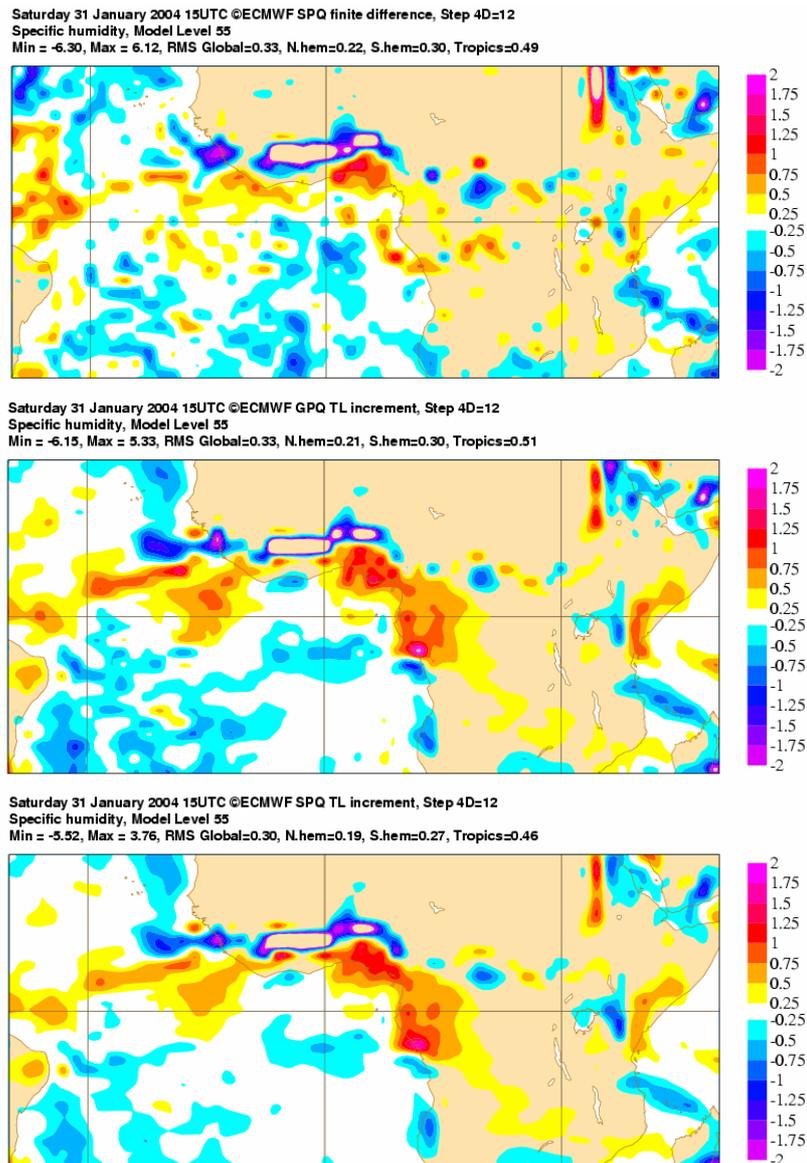
Figure 14: Evolution of the temperature relative error in the T159 tangent linear model with respect to the T511 forecast model over the length of the assimilation window with different interpolation methods for the diagnostics (colour of curves as on Figure 12). Left plot: cubic interpolation. Middle plot: conserving interpolation with non-linear limiter. Right plot: Conserving interpolation without limiter.

The present version of the conserving interpolation with non-linear limiter produces unrealistic error peaks at some fields (the algorithm is under development at the time of writing this paper). However, when the limiter is deactivated, the errors of the conserving interpolation become smaller than in the cubic interpolation (this is especially clear for the initial errors, that mostly come from interpolation). The linear version of the limiter has also been tested and its results are nearly identical to those when no limiter is applied. However, lower order interpolation is only recommended when going from higher to lower resolution. The larger the difference between the resolutions, the less the difference is between higher order and linear interpolation because high-resolution grid-cells entirely within the low-resolution grid-cell integrate exactly to the mean value within the cell, whatever the sub-cell distribution. Going from lower to higher resolution, the order of the interpolation is the deciding factor for accuracy. Slight differences between the left panel of Figure 14 and the temperature error evolution on Figure 5 are seen because the experiment for Figure 14 has run with CY29R1 (instead of CY28R4 of Figure 5) and it uses the wavelet formulation of Jb.

#### 5.4 Grid-point representation of humidity within the inner-loop

We have performed some experiments with the recently developed version of minimization that uses grid-point representation of specific humidity and ozone. Figure 15 shows the so obtained 12-hour humidity increment fields of a single outer loop T799/255 4D-Var on model level 55 (zoomed over a tropical area) as derived from the non-linear model (top), from the tangent linear model in a 4D-Var minimization with grid-point representation of specific humidity (middle, denoted by GPQ) and from the tangent linear model in a 4DVAR minimization with spectral representation of specific humidity (bottom, denoted by SPQ). The increment patterns developed by the GPQ minimization are very similar to the reference SPQ increments.

However, in the GPQ increment field we can identify more small scale structures presumably associated with tropical deep-convection. These structures are, at least qualitatively, more similar to the ones seen on the first map generated by the non-linear model (see e.g. on the Southern tropical Atlantic area) even if the finite difference map is derived from the SPQ experiment. This is promising, though not surprisingly, the quantitative analysis shows that the relative errors of the GPQ and SPQ 4D-Var experiments are almost exactly the same for all vertical levels and variables (not shown), including specific humidity.



*Figure 15: 12-hour evolved specific humidity increments on model level 55 (zoomed over a tropical area) as derived from the T799 nonlinear model (top), from the tangent linear model in a 4D-Var minimization with grid-point representation of specific humidity (middle, denoted by GPQ) and from the tangent linear model in a 4D-Var minimization with spectral representation of specific humidity (bottom, denoted by SPQ). Both GPQ and SPQ minimizations use T255 resolution. Units on the maps are in g/kg.*

## 6. Conclusions

The validity of the tangent linear approximation within the incremental 4D-Var has been re-tested with the current version of the IFS model and some specific problems have been examined.

It has been shown that the gradual resolution increase in the inner-loop brings the tangent linear model closer and closer to the finite differences of the non-linear model in the whole range from T42 to T319 inner-loop resolutions. This is an important improvement with respect to the earlier model versions, and is expected to be of benefit for the forthcoming resolution increases to T799/T255, and makes a higher inner-loop resolution conceivable. As far as the ratio of inner and outer-loop resolutions is concerned, it has been found that the planned T799/255 configuration shows slightly better linearization error characteristics than the presently operational T511/159 configuration.

The reasons for the rapid initial growth of error near the surface have been examined. The results of the experiments suggest that discrepancies between the inner- and outer-loop resolution, non-linearities in the vertical diffusion and in other physical processes equally contribute to this. On the other hand it has been shown that the update frequency of trajectory used for linearization and the discrepancy in the use of surface parameterization do not play an important role in this phenomenon, neither does the choice of advection scheme, semi-Lagrangian versus Eulerian, used in the 4D-Var.

The first tests with the new conserving interpolation scheme to be used in the near future within 4D-Var and with the grid-point representation of humidity and ozone variables within the minimization indicate their potential positive impact. We have shown that humidity analysis increments evolved by the tangent linear model can more realistically describe smaller scale structures when the minimization is based on their grid-point representation even if the inner-loop resolution is the same.

## References

- Andersson, E., C. Cardinali, M. Fisher, E. Hólm, L. Isaksen, Y. Trémolet and A. Hollingsworth, 2004: Developments in ECMWF's 4D-Var system. Proc. AMS Symposium on "Forecasting the Weather and Climate of the Atmosphere and Ocean". Seattle, Washington State, 11-15 Jan. 2004. Available from the American Meteorological Society, <http://ams.confex.com/ams/84Annual/20WAF16NW/program.htm>
- Bouttier, F., 2001: The development of 12-hourly 4D-Var. ECMWF Tech. Memo. 348.
- Courtier, P., J.-N. Thépaut and A. Hollingsworth, 1994: A strategy for operational implementation of 4D-Var, using an incremental approach. *Q. J. R. Meteorol. Soc.* **120**, 1367-1388.
- Fisher, M., 1998: Minimization algorithms for variational data assimilation. Proc. ECMWF Seminar on "Recent developments in numerical methods for atmospheric modelling", Reading, U.K. 7-11 September, 1998. 364-385.
- Fisher, M., 2003: Background error covariance modelling. Proc. *ECMWF Seminar on "Recent Developments in Data Assimilation for Atmosphere and Ocean"*, 8-12 Sept 2003, Reading, UK., 45-63.
- Hólm, E., E. Andersson, A. Beljaars, P. Lopez, J.-F. Mahfouf, A. J. Simmons and J.-N. Thépaut, 2002: Assimilation and Modelling of the Hydrological Cycle: ECMWF's Status and Plans. *ECMWF Tech. Memo.* **383**, pp 55.
- Janisková, M., J.-F. Mahfouf, J.-J. Morcrette and F. Chevallier, 2002: Linearized radiation and cloud schemes in the ECMWF model: Development and Evaluation. *Q. J. R. Meteorol. Soc.*, **128**, 1505-1527
- Laroche, S. and P. Gauthier, 1998: A validation of the incremental formulation of 4D variational data assimilation in a nonlinear barotropic flow. *Tellus*, **50A**, 557-572.

Lawless, A., S. Gratton and N. Nichols, 2005: An investigation of incremental 4D-Var using non-tangent linear models. *Q. J. R. Meteorol. Soc.*, **131**, 459-476.

Rabier, F., H. Järvinen, E. Klinker, J.F. Mahfouf and A. Simmons, 2000: The ECMWF operational implementation of four-dimensional variational assimilation. Part I: experimental results with simplified physics. *Q. J. R. Meteorol. Soc.*, **126**, 1143-1170.

Thépaut, J.-N., and E. Andersson, 2003: Assimilation of high-resolution satellite data. *ECMWF Newsletter*, **97**, 6-12.

Trémolet, Y., 2004: Diagnostics of linear and incremental approximations in 4d-Var. *Q. J. R. Meteor. Soc.*, **130**, 2233-2251.

Veersé, F. and J.-N. Thépaut, 1998: Multiple-truncation incremental approach for four-dimensional variational data assimilation. *Q. J. R. Meteorol. Soc.* **124**, 1889—1908.