## COMMENTS ON THE PAPER

Nonlinear effects in 4D-Var

by Massimo Bonavita, Peter Lean and Elias Holm

Date: July 16, 2018

## 1 Recommendation

This paper revisits some important issues studied previously with simpler models several years ago. The objective is to assess the extent to which nonlinearities are correctly dealt with by using an incremental approach in which successive linearizations are used to minimize a non-quadratic objective function. The source of nonlinearities are first and foremost the model itself but also the treatment of humidity and even new formulation of the objective function that include non-Gaussian observation error (e.g., the Huber norm). This is an interesting paper that raises these issues and presents results obtained with the state-of-the-art 4D-Var of ECMWF. Although one could think of many other experiments that could be done, what is presented here is very interesting. The paper is well written and the results presented clearly. My recommendation is then that the paper be accepted with minor corrections. However, the authors should consider the specific comments that I think need to be addressed. I consider those to be minor except for comment p8, Fig10: the different color bars used for the three figures make it difficult to draw conclusions.

## 2 Specific comments

p1L15: Rabier and Courtier (QJRMS 1992) presented a good study to measure the accuracy of the tangent-linear with a "realistic" global baroclinic model, the IFS of the time but without the physical parameterizations. Lacarra and Talagrand (1988) did also conclude that the TL model was a reasonable approximation of the evolution of a perturbation for ~48-h. This is the basis of ECMWF's EPS.

p1L19: the reference to Pires *et al.* (1996) is missing in the list of references.

p1L24: "the use of a linear model".

p2L22: although this is mentioned in the legend, the notation TL511, for example, should be explained: eg., 40 km (TL511, spectral truncation 511 with a linear grid). The same for TCo1279, which was a new one for me.

p4 eq.(5): the Hessian should be written as $\mathbf{B}^{-1}+\sum_k \mathbf{G}_k^T\mathbf{R}_k^{-1}\mathbf{G}_k$ . The perturbation $\delta\mathbf{x}_0$ is not part it.

p3L23: the notation $\mathbf{x}^t$ is usually used for the true state in the literature while here it refers to the trajectory. I suggest that the notation $\mathbf{x}^{tr}$ should be used instead.

p5L4:I can understnd that interpolation of a high resolution field with high variability would create a "noisy" low resolution field. However, such cases require special treatment of the interpolaitonto be representative (e.g, aggregation instead of geometrical interpolation). Can the authors comment on this?

p5L10: Rabier and Courtier (QJ 1992) have looked into this with a similar approach. This should be referred to and discussed in relation with the results of Fig.1. The difference

$$\Delta \mathbf{X} = M\left(\mathbf{x}^{n-1} + \delta \mathbf{x}^{n}\right) - M\left(\mathbf{x}^{n-1} - \delta \mathbf{x}^{n}\right)$$

shows how long the linearity assumption holds regardless of how the TLM has been formulated. If $\Delta \mathbf{X}$ is not small, it implies that it is hopeless to be able to find a linear model that would provide a reasonable evolution of the perturbation in the initial conditions.

p5L15: reusing the observations implies that the accuracy of the background state has changed and the B-matrix should reflect that. The first-guess trajectory is therefore not the background state or is it? My interpretation is that the first-guess is just a starting point of the minimization as are each state used in an outer loop.

p5L23: this argument holds if the errors are Gaussian but non-Gaussian errors could lead to a non-zero mean ensemble.

p7L5: convergence needs to consider that the objective function represents the fit of a given realisation of the observations. Convergence to numerical accuracy is meaningless and it is justified to reduce the requested accuracy of the minimization.

p8, Fig10: it is difficult to say that the departures with respect to wind observations is smaller when more outer loops are used since the scales (color bars) are not the same. In fact, that of the first-guess has a value of 9.97, while for one outer iteration it is 11.91, with three we get 13.52 and finally with 5, 11.91. It is not so "visually" apparent that 5 outer loops is better.

p8L15: this particular situation involves the physical parameterizations (convection). Looking at the initial physical tendencies (see Rodwell and Palmer, QJ2007) may reveal interesting information for this particular case.

p9L10 (Fig.11): a reduction of O-A is not a good measure of the quality of the analysis. This can be obtained by reducing the observation error but result in even unphysical forecasts. The O-B is a better measure in that sense and indicates that increasing the outer loop from 3 to 4 or 5 leads to a more modest gain.

With one outer loop, is the resolution of the analysis increment TL95 or does it correspond to that used in the third outer loop (TL399). If it is TL95, the degradation may be attributed to the degraded resolution and in that case, it would be better to redo it using the same higher resolution as used for the 3, 4 and 5 outer loops. Even in 3D-Var, TL95 would be considered too low.

p9L25: with significantly different observation errors, the minimization would focus first on those with small errors and it is only when convergence is reached for those that it would take care of others. This can happen when artificially large observation error are assigned to some satellite observations which were then incapable to have a significant influence on the analysis. In this particular case, it may be that satellite observations have now impacted the analysis significantly more than before.

p10L10: an experiment assimilating only satellite measurements sensitive to humidity and precipitation could show more about the nonlinearity associated with those.

p10L14: "seen that the impact".

p10 Conclusion: increasing the resolution implies a reduction of the nonlinear timescale. The assimilation window would have to be shortened. It is important to evaluate what this timescale is. If the TLM of the full model cannot be achieved, Tanguay et al. (1995) have shown with a simple model that even in the best of cases convergence cannot be reached. To what extent can we say that a weak constraint 4D-Var would be needed? Being at ECMWF hwere the weak constraint 4D-Var has been extensively studied, the authors are in a position to comment on this.