Dear reviewer#1,

Thank you for your carefully reading the new version of the article and for the references you have suggested, they were valuable for producing the final version of the article.

We have addressed all your remarks. Special attention was given to remarks 1 and 2, which refer to the latest version of the article, as well as to remark 3 that we judge very important due to its philosophical character.

In the following we explain how we dealt with your remarks and where, in the text, we have made changes.

1) The perturbed initial condition is a model field 20 days apart from the considered unperturbed initial conditions. We have modified lines 603-604 in the new version of the article.

2) We agree that the errors in the initial conditions identified by the 4Dvar should also be projected on the stable manifold. So, we have removed this phrase from the text.

3) Concerning the 'nudging' point, the introduction has been modified in order to explain clearly what it is and when it appeared in the literature (lines 20-41 and 73-87). Nudging is nothing else than Luenberger observer which was invented by David Luenberger in 1966 in a paper where the link with Kalman-Bucy filter was explained and it was presented in a very clear paper in 1971 (the two Luenberger's papers are given in the references). In the linear case for an observable system (i.e. satisfying Kalman's observability criterion), it was proved , thanks to the pole shifting theorem, that there exists a matrix so that the Luenberger observer is asymptotically exponentially stable. This Luenberger observer is still nowadays a fundamental tool of control theory (automatics) and estimation and appears in all the text books in the paragraph 'observers'(the book of Gelb et al. in English or the one of Bonnans and Rouchon in French for instance). Kalman-Bucy appears generally in the chapter concerned with filtering.

This observer appeared, as it is explained in the introduction, in the geophysical literature under the word 'nudging', but it is exactly the same as Luenberger observer. This is not an exception in the world of geophysics. The 4D-VAR method is nothing else than optimal control, EOF technique is nothing else than POD, and for our present matter, nudging is nothing else than the Luenberger's observer.

Its advantage compared to Kalman filtering is that it does not require any information on the various covariance matrices, but as it was pointed out by Luenberger, the Kalman-Bucy filter appears as a particular Luenberger observer which corresponds to the minimum of the quadratic cost-function, with its weights corresponding to the covariances of the various errors. The stochastic observer unifies the concepts of deterministic Luenberger observer theory and stochastic Kalman filtering theory as it is explained in Gelb's book for instance. However, both are useful in practice and have advantages and drawbacks.

Some authors present the Luenberger observer as the deterministic observer, whereas Kalman-Bucy is the stochastic equivalent (see for instance on Wikipedia the paper on observateur d'état in French : http://fr.wikipedia.org/wiki/Observateur_d'état or state observers in English : http://en.wikipedia.org/wiki/State_observer ).

Concerning the remark that nudging is already present in Morel's paper, it is not the nudging concept corresponding to Luenberger's observer and a remark in this sense has been added in the introduction, where it is mentioned that back and forth algorithms have already been used for reversible systems, but it is not back and forth nudging except if one takes the nudging coefficient infinite as it was already mentioned in Auroux-Blum 2008 in the same journal (it is called 'assimilation par aller et retour' in Talagrand's thesis).

Finally concerning this question, it is worth noting that this remark was an interesting point of discussion raised with a few colleagues during the INDAM Workshop on Mathematical Paradigms of Climate Science held in Roma, June 2013.

4) We have changed "The feedback term changes the dynamical equations and forces the state variables to fit the observations as well as possible." to "The feedback term changes the dynamical equations and is a penalty term that forces the state variables to get closer to the observations". (lines 142-143).

5) We have changed "The BFN novelty with respect to conventional nudging methods is the model integration backward in time." to "The difference of the BFN with respect to conventional nudging

methods is the model integration backward in time.” For the references to Morel and Talagrand, these papers have been mentioned in the modified introduction (see also the answer to question 3).

6) We are sorry about the bad citation of the Courtier's work. We have modified it in the lines 231-248.

7) We have corrected the model description to say that it solves five prognostic equations. (lines 257-261).

8) We would say that the iterations compensate for the lack of a priori information about the background errors. What we mean is that the model dynamics propagates the increments on space and time during the iterative process given a dynamical character to the final increment. We have changed the text accordingly (lines 532-534).

9) We do believe that the bad performance of the 4Dvar for the assimilation window of five days is due to spurious increments, which may be related to bad convergence. The fact is that for a five days window inertia-gravity waves may still dominate the next first guess for the new assimilation cycle. To avoid entering into this topic we prefer to remove this phrase from the text as you have suggested.

10) With the figure 5 we only want to show that for our implementation both methods work properly.

11) We have changed the caption of figure 6.

12) The ordinary nudging has been cited in the lines 388-389.

13) We changed it in the second paragraph of the introduction.