

## ***Interactive comment on “EnKF and 4D-Var Data Assimilation with a Chemistry Transport Model” by S. Skachko et al.***

**S. Skachko et al.**

sergey.skachko@aeronomie.be

Received and published: 27 May 2016

[11pt,a4paper]article color

C1

### **Response to the Anonymous Referee 2**

May 27, 2016

The author's responses are marked in blue. We first would like to thank the anonymous referee 2 to the useful remarks.

This paper presents a comparison of the EnKF and 4D-Var data assimilation methods applied to a chemistry transport model. These two methods have been used extensively in applications to numerical weather prediction (NWP), a field with which I am more familiar. It is interesting to see such an intercomparison for an atmospheric chemistry model for simultaneously assimilating observations of the concentrations of several chemical species. General Comments:

#### 1 Set up of assimilation windows

One major concern I have is related to the basic setup of the experiments. Whereas intercomparisons in the context of NWP have generally used the same data assimilation window length for the two approaches (typically 6h), this study uses a different window length: 4D-Var is applied in its strong constraint formulation with a 24h window, whereas the EnKF is used to sequentially assimilate observations every 30 minutes (the time step length of the forecast model). In addition, model error perturbations are added every 30 minutes in the EnKF experiment, in contrast with the assumption of no model error over the 24h assimilation window in the 4D-Var setup. This difference in

C2

the configurations of the two methods seems to affect the conclusions of the study, as acknowledged by the authors near the end of the paper.

Alternative configurations could have possibly been chosen that would have reduced these differences. For example, the 4D-Var could have been implemented in its weak-constraint formulation, employing a model error term every 30 minutes based on the same covariance matrix as used for model error in the EnKF experiment. Alternatively, a 4D-EnKF approach could have been used with the EnKF in which ensembles of model solutions over a longer assimilation window are used to construct a 4D background error covariance matrix (as is done in most current implementations of both EnVar and the EnKF for NWP). In this case a window shorter than 24h (say 6h) would likely be preferable (and used for both EnKF and 4D-Var) to avoid problems with horizontal localization when a chemical species is advected within the time window by a distance comparable to the localization length scale. The authors should mention these two alternative approaches and discuss how they could have reduced the differences seen in some of their results.

As a consequence, the difference in window length also likely has a direct impact on the quality of the forecasts, since the 4D-Var forecasts are started from a state that may not be in as good agreement with the most recent observations as in the case of the EnKF. It would be useful to show the fit of the state used to initialize the 24h forecasts with respect to the observations valid at the same time for the EnKF and 4D-Var experiments, or at least mention in the text how this fit differs between the two. I wonder if this explains some of the differences in results for HCl.

We appreciate the concern raised by the referee to mention alternative 4D-EnVar approach (and it was done), however the purpose of this paper is not to try to develop a hybrid method, but first to compare two data assimilation approaches in the set-up that is usually found in chemical applications. It is current practice to use a 24 h window in stratospheric 4D-Var chemistry because of diurnal cycles with many chemical species and because of the polar orbiter limb sounder whose spatial coverage is far too limited

C3

in 6h, whereas a 24 h does provide an acceptable complete spatial coverage. The BASCOE 4D-Var system has been running in near-real time since the early 2000's using the 24 h assimilation window. On the other hand, since the early 2000's sequential methods, such as simplified KF and later with the EnKF, have been running with model error and with a time window of 1 hour. In our previous work, we did compare 4D-Var and EnKF but for chemical transport with no chemistry (Skachko et al 2014), referred as S14, and found nearly identical results, despite the fact that BASCOE EnKF used 0.5 h ensemble model forecasts and 4D-Var, the 24 h assimilation window. From a dynamical system point of view, chemical transport and NWP are quite different. We do not find growth of error in chemical transport. And adding chemistry produces a system with a rather dissipative behaviour. This is corroborated by the fact that chemical transport model simulation (without chemical data assimilation) shows stable differences with chemical observations over periods of months and even years. So we disagree that the difference in window length has such an impact in the context of chemical transport. What we did here is to add chemistry to this problem and see what are the implications due to chemistry. In particular we acknowledge that the window length and having model error or not does have an impact on certain species. However, in the context where the chemistry would be directly coupled with an NWP model, a 6hr time window would need to be considered. The cost of operational implementation of such an approach would be nevertheless prohibitive. Lower resolution, one way coupling or simplified chemistry (which displays very different characteristics in terms of model error) are most likely to be considered for such set-up.

2 Inter-species background error covariances The discussion regarding the estimation and treatment of inter-species background error covariances is somewhat unclear. The authors state that because it seems suitable to use the same model error covariances for each species it follows that the main source of error is from the transport. However, then the authors claim that the cross-species background error covariances are weak and not sufficiently well estimated with the EnKF ensemble and therefore should be set to zero when assimilating the observations. It seems to me that if the errors of

C4

the different species are dominantly affected by the same common source (i.e. the transport), then their errors should be correlated. Or maybe this is not the case due to the dependence of the concentration errors from transport on the concentration spatial gradients, which may be very different for each chemical species. Please provide some comment on this. Also, I think the authors need to stress how the relatively small size of the ensemble (20 members versus the usual size for NWP of  $O(100)$ ) could affect their conclusions regarding the utility of the cross-species covariances. Presumably with a much larger ensemble they would be better estimated and therefore could be more useful.

Although it is the same transport that acts on all species, the resulting distribution of species are not necessarily correlated. Long-lived species for instance which are the best candidates for such correlations show in some cases complicated correlations patterns that depend on latitude and height and vary in time (e.g. Sankey and Shepherd 2003). Besides, preliminary experiments with BASCOE-4D-Var system showed that the cross-covariance between innovations of long-lived species is rather noisy and assimilation experiments that accounts for cross-covariance between long-lived species in the background error term do not show in practice significant improvements in the analysis quality. This additional information is added in the article text.

An example of weakly related chemical species is considered in the article: the ozone and  $N_2O$  are assimilated taking into account the cross-correlation terms in the background error covariance matrix. The spurious errors, that arise from the use of a rather small ensemble, can be reduced by using larger ensemble, however the NWP ensemble of  $O(100)$  would not be sufficient to solve the problem. And the computation cost of such ensemble data assimilation becomes unreasonably expensive. To our knowledge, all groups working on chemical data assimilation simply put the cross-correlations between weakly chemically related species to zero. We don't try to solve this problem, but only provide an illustration. The solving of the automatic localization procedure between species would be a subject of another study which is out of scope of the current

C5

paper.

Specific comments:

In several places the term "observation errors" is used where I believe "observation error variances" or covariance or statistics is really what is meant.

Done.

Similarly, on page 3, line 2 I am guessing that "the covariance inflation" refers to the inflation of the "background error covariance", please be more precise.

The sentence is rewritten as: "In the same line of thought, the Desroziers' method (Desroziers, 2005) was also used to simultaneously estimate the covariance inflation factor and the observation error variance(Li et al. 2009, Gaubert et al. 2014)."

page 1, line 10 "Æùhere we keep both estimates:" This is not clear. Is it meant that you keep both estimates fixed in time? Please clarify.

We wanted to say that in the 4D-Var experiments, both background and observation error covariance matrices were estimated using the Desroziers' method. To be more clear on this, I reformulated the sentence as follows: "For comparison purposes, we apply the same estimate procedure in the 4D-Var data assimilation, where both, the background and observation error covariance matrices are estimated using the Desroziers' method."

page 1, line 12 "Æà single model error basedÅġ' It is difficult to decipher what is meant. I suppose this refers to using the same specified model error covariance matrix for each chemical species. Please make this more clear.

This part is reformulated as: "However in EnKF, the background error covariance is modelled using the full chemistry model and a model error term which is tuned using an adjustable parameter. We found that it is adequate to have the same value of this parameter based on the chemical tracer formulation that is applied for all observed

C6

species.”

page 1, line 15 “Æsampling noise errors” and “These errors need to be filtered out”. This is awkward wording. Better to talk about “sampling error due to the use of a small ensemble leading to spurious covariances” and “setting these spurious covariance to zero”.

This is rewritten in one sentence as: “The second issue in EnKF with comprehensive atmospheric chemistry models is the spurious error, that occurs when species are weakly chemically related at the same location.”

page 1, line 20 “Æot too small chemical life times” This does not sound sufficiently quantitative. Could you instead say something like “chemical lifetimes longer than” where you compare the lifetimes to some relevant time scale, e.g. the model time-step or assimilation window length.

This is rewritten as: “If the erroneous chemical modelling is associated with moderately fast chemical processes, but whose life-times are longer than the model time step, then EnKF performs better,…”

page 2, line 6-7: It would make sense to also refer here to one of the first EnKF/4D-Var comparisons performed with real NWP systems done at the same center as one of the coauthors: Buehner et al. 2010 (2-part paper in MWR).

Done.

page 2, line 29-34: It is not clear why the issue of estimating error statistics for many chemical species is specific to the EnKF, as the text now seems to imply. Clearly this is equally important and challenging for the application of 4D-Var? Please clarify.

Our previous work, S14, based on a chemical tracer version of the CTM showed that the EnKF is much more sensitive to the parametrization of the error statistics than the 4D-Var. In order to make it more clear, I put the following text at the end of the next paragraph: “Contrary to the EnKF, the 4D-Var is much more tolerant to the parametriza-

C7

tion of the error statistics, as it was shown in S14. Hence, the online estimation of the error statistics is of great importance only for EnKF.”

page 3, line 4 “background error” should be “background error covariance”.

Done.

page 4, line 13, 15: Define the acronyms PSC and MACC.

Done.

page 5, line 15: It would help to state here that the same background error correlations are used for each chemical species and that the between-species covariances are assumed to be zero.

The sentence of line 24-25: “... $\Lambda^{1/2}$  is the spatial correlation matrix defined on a spherical harmonic basis hence diagonal;” is modified as follows: “... $\Lambda^{1/2}$  is the spatial correlation matrix, identical for each chemical species, defined on a spherical harmonic basis hence diagonal;” And a new sentence is added at the end of the same paragraph: “The between-species covariances are assumed to be zero in the background error covariance matrix  $B_0$ .”

page 6, line 8: gamma is really set to 5? According to equation 5 this means an observation is rejected if its innovation is larger than 2.2 stddev (i.e.  $\sqrt{5} = 2.2$ ). Maybe equation 5 should have  $\gamma^2$  instead of just  $\gamma$ ?

Yes,  $\gamma^2$  is written now in the equation.

page 6, line 20: the parameter N (ensemble size) needs to be defined here, since this is where it first appears.

The parameter N is now defined here.

page 6, line 23: “Æare normally distributed random numbers” should probably be “Æare vectors of independent normally distributed random numbers” otherwise

C8

equations 6 and 7 do not make sense.

Done.

page 6, lines 26-29: The two sentences about the application of the Desroziers method seems out of place here, since the method has not yet been introduced. Considering moving this to section 2.5.

The text is modified: "...where the adjustable scaling factor  $s_o(i)$  is estimated using the method described in Sect.2.5". And the second sentence is written as: "Besides, the current version of EnKF allows for more accurate observation error covariance estimation with respect to S14 because it computes  $s_o(i)$  as a function of observation vertical pressure level."

page 7, lines 8-9: "To this end, the EnKF algorithm accounts now for a new effective procedure to find a current local sub domain in the model space." It is not at all clear what this sentence means. It must be better explained. Is the algorithm similar to that described by Houtekamer et al. (2014, MWR - "Parallel implementation of an EnKF")?

The sentence is now written as: "To this end, the EnKF algorithm accounts now for a procedure to find a current local sub domain in the model space using the K-D tree, which is a binary search tree where the comparison key is cycled between K components (K = 3 in our case, because the observation location is a 3-dimensional vector). More information on the method can be found on the Web or in any textbook on data structures (e.g. Gonnet and Baeza-Yates 1991)."

page 8, line 3: I believe "variance" should be "standard deviation". Please also check the entire paper to ensure standard deviation and variance are used correctly. Also, as already mentioned, ensure the word "error" is not used in places where "error standard deviation" or "error statistics" or "error covariance" is actually what is meant. This is a common mistake that can be very confusing for some readers.

Done.

C9

page 8, lines 32: "and thus goes along in arguing that they represent some true error statistics." This does not sound very solid as a logical argument. Consider improving. Similarly for the following 2 sentences. It is probably not necessary to make this assertion at this point in the text.

The text is modified as: "The time evolution of the error variance scaling factors at individual levels is, as in Figure 3, (result not shown) generally consistent over time, especially for the EnKF estimates. Furthermore, we note that for the EnKF results, the scaling factors of any species show no drift in time. We argue from this result that there is apparently no need to have a different model error  $\alpha$  for different species. Thus we conclude that for a chemical transport models, the main source of model error can be attributed to transport errors primarily."

page 9, line 4: "10 iterations" Is this enough to obtain a substantial reduction in the amplitude of the cost function gradient (i.e. at least a factor of 10)?

The reduction in the gradient of the cost function depends on the assimilated observations. In the case of EOS Aura-MLS data, the reduction varies between the factor of 6 to 10. The reduction is more important when the number of iteration is doubled, however it does not lead to any significant improvement in the OmF statistics. For the practical purposes, the near-operational BASCOE 4D-Var system used in the MACC project is run with 10 iterations, because the chemical modelling is essentially time-consuming task. In a similar data assimilation case, Elbern et al. (2010) utilize between 12 and 16 iterations.

page 9, section 3: Please provide some additional general information about the observations assimilated: how much of the globe is observed during 24h? what is the horizontal and vertical spatial resolution?

The 24h of the 4D-Var assimilation window is chosen because within this period, Aura-MLS data provide near-global coverage with gaps that are dispersed regularly within the correlation length of the data assimilation, i.e. 800km. The following text is added:

C10

“The EOS Aura-MLS data cover the latitude range between 82°S and 82°N with an along-track separation of around 165 km between consecutive scans. Around 3500 vertical scans are performed every day. The vertical resolution varies for different species.”

page 10, section 4.1 and verification in general: What observations are used for verification? For example, are all observations with a valid time within 1h or 3h or ?h of the valid time of the forecast used? Since you are always verifying 24h forecasts valid at 0000UTC each day (if I understood correctly), does this tend to focus the verification in only certain geographical regions due to the orbit of the satellite?

At every 0.5 h model time step of the 24 h forecast, the BASCOE system computes the OmF statistics. Thus, figures 4-10 show the statistics where all observations between 0 h and 24 h UTC, distributed within a given latitude band and chosen period, are taken into account.

page 11, line 16-17: related to the first general comment above, I think the difference seen here between the 4D-Var and EnKF results may be due to the difference in assimilation window lengths, please add a comment here.

In the context of chemistry, the difference in data assimilation window lengths really has implications, as pointed out by the referee. Our conclusions, page 14, line 32-32 and page 15 line 1: “Two main reasons are responsible for this better performance. First, EnKF has a short-time forecast followed by frequent observational updates that is possibly more adequate for moderately fast chemical processes (but not for processes of life-time smaller than the model time step). Second, the ensemble of CTM's provides better representation of the model variance.”

page 12, line 32: “This is due to the automatic rejection by the 4D-Var of most observations..” What does this mean? Is it because the 4D-Var cannot make the forecast model solution fit the observations over the 24h assimilation window (due to model error) or is it referring to some quality control procedure (which I though was deactivated

C11

for this chemical species at the pressures considered here)?

Yes, the background quality check (BgQC) procedure rejects the observations. BgQC is active for both EnKF and 4D-Var data assimilation systems, though it works differently for each systems when their background states differ.

page 13, line 23: “..model error covariance..” should be “..background error covariance in observation space..”

Done.

page 13, line 26: “..localization of the error variance..” should be “..localization of the error covariance..”

Done.

page 15, last paragraph: Another limitation is that much fewer ensemble members were used as compared with typical NWP applications. This should be mentioned.

The ensemble used in this study is typical for chemical data assimilation. Here, we don't aim to compare BASCOE system with NWP applications.

## References

- Elbern, H., J. Schwinger, and R. Botchorishvili (2010), Chemical state estimation for the middle atmosphere by four-dimensional variational data assimilation: System configuration, *J. Geophys. Res.*, 115, D06302, doi:10.1029/2009JD011953.
- Skachko, S., Errera, Q., Ménard, R., Christophe, Y., and Chabrilat, S.: Comparison of the ensemble Kalman filter and 4D-Var assimilation methods using a stratospheric tracer transport model, *Geosci. Model Dev.*, 7, 1451-1465, doi:10.5194/gmd-7-1451-2014, 2014.
- Correlations of long-lived chemical species in a middle atmosphere general circulation model, *J. Geophys. Res.*, 108, 4494, doi:10.1029/2002JD002799, D16.

C12