Final Response to the Anonymous Referees

July 11, 2016

We thank all the referees for the detailed revision of the paper. The author's responses are marked in blue.

1 Response to the Referee 1

General

This paper represents a unique contribution on the comparison of EnKF and 4D-Var approaches for the assimilation of chemical species. It provides much insight, notably on issues related to inter-species error correlations and localization. I congratulate the authors. I provide here only minor suggested corrections.

Minor corrections:

L16: Change to: "one issue is the large number of "

Done.

L19: Change "comparison reason " to "comparison purposes "

Done.

L182: Change " will be a subject to " to "will be subject to "

Done.

L208: Define PSC

Done.

L211 Do not refer to MACC or define it

The MACC acronym has been defined.

L303 Change "has the background quality" to "includes a background quality "

Done.

L 501 -504 Apparent contradiction where precision level 14 ppbv seems first to correspond to 38% error at L 501 while at line 503 it corresponds to 250%. This is likely because in the first case it corresponds to 4.6 hPa level while in the other it corresponds to 1 hPa level. Sentence could be clearer.

We have modified the sentence as follows: "...the Aura MLS N_2O precision is 24-14 ppbv (9-38%, relative to the observation mean at given altitude) and its accuracy is 70-3 ppbv (9-25%) in the pressure range 100-4.6 hPa but its precision drops to 14 ppbv (250%) at 1 hPa..."

L 698: why qualify as “tricky ” the Aura-MLS dataset?

We modified the phrase with “the Aura-MLS N_2O dataset containing a bias”

L707 End of sentence point needed after 5 hPa.

Done.

L709 “ approaching to them”, define what “them” represents. Not clear as is.

Replaced with “keeping the model closer to the assimilated observations”

L727 Drop the reference to paper in preparation!

Done.

L733: Suggest to change “accidentally ” by “incidentally ”

Done.

L826 Life time of ozone is that small, lower than a model time step presumably of order 15 min?

Indeed.

Conclusion: Perhaps add as comment that an hybrid approach such as popular 4D-EnVar approach could emerge as a good way forward for chemical data assimilation.

The following sentence is added to the last paragraph of the conclusions: “A future development of the BASCOE chemical data assimilation system would be a hybrid 4D-EnKF approach using the ensemble of models to construct a 4D background error covariance matrix.”

2 Responses to the Referee 2

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 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 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 considred. 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 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 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 "where 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 "a single model error based" 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 species."

page 1, line 15 "sampling 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 oc-

curs when species are weakly chemically related at the same location."

page 1, line 20 "not 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 parametrization 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: "...$\mathbf{\Lambda}^{1/2}$ is the spatial correlation matrix defined on a spherical harmonic basis hence diagonal;" is modified as follows: "...$\mathbf{\Lambda}^{1/2}$ 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 $\mathbf{B}_0$."

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 γ^2 instead of just γ?

Yes, γ^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: "are normally distributed random numbers" should probably be "are vectors of independent normally distributed random numbers" otherwise 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.

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 α 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: "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 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.

General comments:

I am confused by the authors’ response to my main concern, related to the difference in the window length used for the 4D-Var and EnKF experiments. In response to my first general comment, the authors’ response is: ”So we disagree that the difference in window length has such an impact in the context of chemical transport.” Then, when I later bring up the same point again in relation to the discussion of the results, the authors’ response is: ”In the context of chemistry, the difference in data assimilation window lengths really has implications, as pointed out by the referee.”

We should be more clear on this. The first mentioned sentence means that in the context of chemical tracer transport only (without chemistry system), there is no difference in using an EnKF with 30 min ensemble model forecasts and a model error term or a 4D-Var with 24 h assimilation window without model error term. This was shown in our previous article (Skachko et al 2014). The purpose of the present work is to reveal the role of the chemistry system (including interactions between chemical species) in the context of our two data assimilations that are configured as they are normally used in chemical data assimilation applications: one model time step ensemble model forecasts within EnKF, and 12 - 24 h of 4D-Var assimilation window.

Also, I believe the authors’ misinterpreted part of my first general comment. I made no suggestion that a hybrid 4D-EnVar experiment be performed, or even mentioned. What I did suggest was that a 4D-EnKF approach (with model error perturbations only applied at the beginning of each

window to be equivalent with strong-constraint 4D-Var) should be considered and mentioned, since this would allow a longer window to be used for the EnKF. In this case, the analysis would be forced to simultaneously fit all of the observations distributed over a longer window, while still satisfying the model equations, as in 4D-Var.

I appreciate that the authors have tested two data assimilation methods in configurations as they are usually used for chemical applications. This point should be emphasized in the paper to justify the choice. However, it would be helpful to inform the reader that other configurations are possible that would reduce the differences between the two approaches (i.e. including model error in weakconstraint 4D-Var and using 4D covariances with a longer window in the EnKF). Otherwise, readers will conclude that one approach (i.e. EnKF or 4D-Var) is fundamentally better or worse than the other in some respects, whereas it is more likely the choice of how each approach was implemented that is more important.

The fourth paragraph of the introduction is modified as follows: "But how do the EnKF and the 4DVar methods compare when photochemical reactions are taken into account? Do the results depend on the assimilated chemical species? Using actual satellite datasets and operational configurations, what are their respective performances in terms of precision, accuracy and computational efficiency? What is the role of the practical implementation of each method, when the full description of the stratospheric chemistry is taken into account in the CTM. These are the main questions addressed in this paper."

The conclusions start with: "We have conducted a comparison of an EnKF and 4DVar data assimilation system using a comprehensive stratospheric chemical transport model. We considered 4D-Var and EnKF configurations that are normally used for chemical data assimilation applications. Both data assimilation systems have online estimation of error variances based on the Desroziers' method and share the same correlation model for all prescribed error correlations (i.e. the background error covariance for 4D-Var, initial error and model error for EnKF) so that each data assimilation system is nearly optimal and can also be compared to each other. A previous comparison study by (Skachko et al. 2014) showed that for chemical tracer transport only both assimilation system provide results of essentially similar quality despite the difference in practical implementation of each method: the 4D-Var was applied in its strong constraint formulation with a 24 h assimilation window with the assumption of no model error over this period, whereas the EnKF was used to sequentially assimilate observations every 30 minutes with model error perturbations added every 30 minutes."

Then the following text is added at the end of our conclusions: "Another possibilities may be considered to properly compare two essentially different data assimilation systems. First, a 4D-EnKF approach, where model error perturbations only applied at the beginning of each 4D-Var assimilation

window to be equivalent with a strong-constraint 4D-Var, may be considered. This would allow a longer assimilation window to be used for the EnKF. In this case, the analysis would be forced to simultaneously fit all of the observations distributed over a longer window, while still satisfying the model equations, as in 4D-Var. Second, the use of a weak-constraint 4D-Var including model error would also reduce the differences between two considered approaches. "

Specific comments:

In response to the third specific comment, your revised sentence seems imprecise: "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." I presume it is only the scale factors for both covariance matrices that are estimated and not the full matrices? Please improve the wording.

The sentence is now written as: "For comparison purposes, we apply the same estimate procedure in the 4D-Var data assimilation, where both scale factors of the background and observation error covariance matrices are estimated using the Desroziers method."

In response to the fifth specific comment, your revised sentence does not clear up my concern: "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." The term "spurious error" is very ambiguous how can error be spurious? I believe this is again where "error" is used in place of "error covariance". Only the "estimated error covariance" is spurious. [The word "error" on its own really should be reserved for the difference between an estimate and the truth and I don't think this is what is meant in this case. I realize that some published papers have used "error" to mean "error standard deviation" or "error covariance", but I believe this has needlessly caused confusion for some people in the DA community.]

The sentence is rewritten as follows: "The second issue in EnKF with comprehensive atmospheric chemistry models is the noise in the cross-covariance between species, that occurs when species are weakly chemically related at the same location."

I am satisfied with the authors responses. I only suggest the following text be used instead of the added text on page C4 of the authors latest response (which incorrectly referred to the "4D-Var" window when describing the 4DEnKF approach): "Other possibilities may be considered to properly compare two essentially different data assimilation systems. For example, the 4DEnKF (Hunt et al., 2004) approach could be used that computes 4D error covariances from the ensemble of forecasts at several times within the assimilation window. This would allow a longer assimilation window to be used in the EnKF experiment, making it more comparable to the configuration of 4D-Var. In this case, the EnKF analysis would be forced to

simultaneously fit all of the observations distributed over a longer window, while still satisfying the model equations, as in 4D-Var. Applying the model error in this 4DEnKF only at the beginning of each assimilation window would make it similar to the strong-constraint version of the 4D-Var that was used.

We thank the referee for the fruitful discussion of the paper. The proposed text has been taken into account as is.

3 Response to the Referee 3

This study compares EnKF and 4D-Var data assimilation methods applied to a chemistry transport model. The purpose is to compare relative merits of the two methods on long time (short windows) atmospheric chemistry data assimilation with prescribed flow fields.

Major comments:

1 EnKF Experimental setup: Page 6: "the model error term is added to observed species only." What is the rationale for this?

Perturbing all 58 species of the model state vector results in the noisy cross-species error covariances. A simplified example of such set-up (where the cross-covariances between the ozone and N_2O only) is shown in the experiment EnKF-CC. When non-observed species are not (or weakly) chemically related with the observed species, the noise introduced to the EnKF error covariances essentially

The same L operator seems to be used both for 4D-Var and EnKF, but at lest in the definition of η in (7), and (1) or (2), L lives in different spaces.

The operator $\mathbf{L}$ is defined by Eq. (3) for both, 4D-Var and EnKF systems in the spectral space. The algorithm to generate EnKF state perturbations is then identical to the algorithm of the 4D-Var background error covariance generation. However, the operator $\mathbf{L}$ is applied to the normally distributed random deviate ζ_i (Eq. (6)) rather than to the control variable ξ (Eq. (1))

The authors claim that the same error covariances are used in both cases [page 14: "the same correlation model for all prescribed error correlations (i.e. the background error for 4D-Var, initial error and model error for EnKF)"]; however, on page 8 around line 10, they seem to indicate different localization operators that come in to build B. This should be clarified.

We have given in the manuscript a reference to our previous study, where it had been explained in more details: "The EnKF uses as localization method a Schur product with a compact support correlation function. The use of Schur product reduces the resulting correlation length scales. In order to maintain the correlations of the EnKF analysis comparable to those of the 4D-Var system, a different setting of the correlation length scales is adopted to generate the model error. Let $\mathbf{C}$ be a matrix resulting from the Schur product of two matrices $\mathbf{A}$ and $\mathbf{B}$: $\mathbf{C} = \mathbf{A} \circ \mathbf{B}$. If the correlation length

scales of **A** and **B** are, respectively L_A and L_B, the correlation length scale of **C** is given by Gaspari and Cohn, (1999):

$$\frac{1}{L_C^2} = \frac{1}{L_A^2} + \frac{1}{L_B^2}. \quad (1)$$

In our case, L_A corresponds to the correlation length scale L_{loc} of the compact support correlation function ρ and L_B corresponds to the correlation length scale of the forecast ensemble covariance matrix $\mathbf{B}_e$, denoted in the following by L_e. Similarly, L_C corresponds to the correlation length scale of the analysis ensemble covariance matrix, denoted in the following by the effective correlation length scale L_{eff}. As we would like to maintain the L_{eff} equal to the Gaussian correlation length scales used in the 4D-Var (i.e. L_0^h=800 km and L_0^v=1 level), we need to set L_{loc} and L_e such that $L_{eff} = L_0$. "

Cross species localization: In Section 5 the authors discuss the effects of interspecies localization. It is unclear to me what is done here. Is ENFK-CC the same as EnKF except that in EnKF-CC the O3 and NO2 are localized? If that is the case, then this is problematic because one cannot choose to localize some species and not localize the others because it introduces transients that may lead to spurious bias oscillations. This should be clarified as well.

The experiment, denoted as EnKF-CC, involves the cross-species error covariance between two weakly chemically related species, the ozone and N_2O. In other words, we consider the assimilation test, where the observational updates of ozone are obtained using both the ozone and N_2O measurements, and the updates of N_2O are also computed using both datasets. This is done in addition to the procedure of spatial localization that has been applied to all observed species.

Minor comment:

Page 4 line 2: "cross-covariance between species are taken into account automatically using the 4D-Var adjoint mode" is not clear to me. How is this achieved?

In the 4D-Var approach, we have a direct and an adjoint chemistry model. The chemistry model includes all possible chemical interactions between species. Hence, the 4D-Var computes the observational updates of all model state variables (observed and non-observed) from all available observations.

References

[Elbern et al (2010)] Elbern, H., J. Schwinger, and R. Botchorishvili (2010), Chemical state estimation for the middle atmosphere by fourdimensional variational data assimilation: System configuration, J. Geophys. Res., 115, D06302, doi:10.1029/2009JD011953.

[Gaspari and Cohn (1999)] Gaspari, G. and Cohn, S. E.: Construction of correlation functions in two and three dimensions, Q. J. R. Meteorol. Soc., 125, 723757, 1999.

[Skachko et al. (2014)] Skachko, S., Errera, Q., Mnard, R., Christophe, Y., and Chabrillat, 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.

[Sankey, David and Shepherd, Theodore G. (2003)] Correlations of long-lived chemical species in a middle atmosphere general circulation model, J. Geophys. Res., 108, 4494, doi:10.1029/2002JD002799, D16.

EnKF and 4D-Var Data Assimilation with ~~a~~ the Chemistry Transport Model BASCOE (version 05.06)

Sergey Skachko[1], Richard Ménard[2], Quentin Errera[1], Yves Christophe[1], and Simon Chabrillat[1]

[1]Royal Belgian Institute for Space Aeronomy, BIRA-IASB, Brussels, Belgium
[2]Air Quality Research Division, Environment Canada, Dorval, Canada

Correspondence to: Sergey Skachko
(Sergey.Skachko@aeronomie.be)

Abstract. We compare two optimized chemical data assimilation systems, one based on the ensemble Kalman filter (EnKF) and the other based on four-dimensional variational (4D-Var), using a comprehensive stratospheric chemistry transport model (CTM). The work is an extension of the Belgian Assimilation System for Chemical ObsErvations (BASCOE), initially designed to work with a 4D-Var data assimilation. A strict comparison of both methods in the case of chemical tracer transport was done in a previous study and indicated that both methods provide essentially similar results. In the present work, we assimilate observations of ozone, HCl, HNO_3, H_2O and N_2O from EOS Aura-MLS data into the BASCOE CTM with a full description of stratospheric chemistry. Two new issues related to the use of full chemistry model with EnKF are taken into account. One issue ~~concerns to~~ is a large number of error variance parameters that need to be optimized. We estimate an observation error variance parameter as function of pressure level for each observed species using the Desroziers' method. For comparison ~~reasons~~purposes, we apply the same estimate procedure in the 4D-Var data assimilation, where ~~we keep both estimates:~~ both scale factors of the background and observation error ~~variances~~covariance matrices are estimated using the Desroziers' method. 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 ~~a single model error~~ the same value of this parameter based on the chemical tracer formulation that is applied for all observed species. This is an indication that the main source of model error in chemical transport model is due to the transport. The second issue in EnKF with comprehensive atmospheric chemistry models is ~~the sampling errors between species. When~~ is the noise in the cross-covariance between species, that occurs when species are weakly chemically related ~~, cross-species sampling noise errors occur~~ at the same location. These errors need to be filtered out, in addition to a localization based on distance. The performance of two data assimilation methods was assessed through an eight-month long assimilation of limb sounding observations from EOS Aura-MLS. The paper discusses the differences in results and their relation to stratospheric chemical processes. Generally speaking, EnKF and 4D-Var provide results of comparable quality but differ substantially in presence of model error or observation biases. If the erroneous chemical modelling is associated with ~~not too small chemical~~ moderately fast chemical processes, but whose life-times are longer than the model time step, then EnKF performs better, while 4D-Var develops spurious increments in the chemically related species. If, on the other hand, the observation biases are significant, then 4D-Var is more robust and is able to reject erroneous observations, while EnKF does not.

1 Introduction

The Ensemble Kalman Filter (EnKF) and the four-Dimensional Variational algorithm (4D-Var) are widely used data assimilation methods that utilize the model to propagate observational information in time and space into an estimate of the state. Each method is built around different assumptions and has its own merits. But to some extent, the relative merits are application dependent. In the context of meteorological data assimilation, the relative advantages of these two methods have been discussed by ~~Lorenc (2003) and Kalnay et al. (2007)~~ Lorenc (2003); Kalnay et al. (2007); Buehner et al. (2010a, b) to name a few, and it has promoted the development of new hybrid methods such as 4DEnVar (Lorenc et al., 2015) and En4DVar (Liu et al., 2008; Poterjoy and Zhang, 2015). In atmospheric chemistry there are, however, very few comparison studies. The purpose of this paper is to compare carefully optimized EnKF and 4D-Var chemical data assimilation systems for an extended time period using the same Chemistry-Transport Model (CTM) and same observations.

A short literature review discussing the Chemical Data Assimilation (CDA) problems related to EnKF and 4D-Var, their inter-comparison and application to the atmospheric chemistry modelling is already given in Skachko et al. (2014, hereafter denoted S14). More recent review including future prospects for coupled chemistry-meteorology models is given by Bocquet et al. (2015).

As in S14, here we use BASCOE (Belgian Assimilation System for Chemical ObsErvations) environment. BASCOE was designed to assimilate satellite observations of chemical composition into a stratospheric CTM originally using the 4D-Var assimilation method (Errera et al., 2008; Errera and Ménard, 2012). S14 described the implementation of the EnKF as an alternative assimilation method in BASCOE and compared it with the original 4D-Var approach, using carefully calibrated error variances for both methods and applying them to observations of ozone which was considered as a passive tracer. Indeed this preliminary paper performed the comparison in a chemical tracer transport framework, i.e. taking only transport into account while neglecting chemical reactions. Our results showed that in this framework the two methods give nearly identical performance. This outcome can be interpreted as a consequence of the dynamics of tracer error covariances: as noted early on by Cohn (1993) and Ménard and Daley (1996), such error covariances follow the characteristics of the flow. Hence in the absence of model error, there is thus no distinction between a filtering (EnKF) and a smoothing (4D-Var) algorithm.

But how do the EnKF and the 4DVar methods compare when photochemical reactions are taken into account? Do the results depend on the assimilated chemical species? Using actual satellite datasets and operational configurations, what are their respective performances in terms of precision, accuracy and computational efficiency? What is the role of the practical implementation of each method, when the full description of the stratospheric chemistry is taken into account in the CTM. These are the main questions addressed in this paper.

The application of the multi-variate EnKF method to an assimilation system with the full chemistry should in principle address two important issues: the estimation of a large number of input error statistics; and the problem of localization between chemical species.

The first issue is the large number of input error statistics that is needed (e.g. the observation error variances for each species at each vertical levels). Clearly, an online estimation of error statistics is desirable to accomplish this task. In an

idealized framework, Mitchell and Houtekamer (2000) proposed an adaptive EnKF where the model error parameters were estimated using innovation statistics within a maximum likelihood method. In the same line of thought, the Desroziers' method (Desroziers et al., 2005) was also used to simultaneously estimate the covariance inflation factor and the observation ~~errors~~ error variance (Li et al., 2009; Gaubert et al., 2014). Ménard (2016) also showed that the Desroziers' observation variance estimates converge to the truth if the background error covariance is close to the truth, which seems to be a reasonable assumption for EnKF background error covariances, when the χ^2 condition (Ménard and Chang, 2000) is respected. Contrary to the EnKF, the 4D-Var is much more tolerant to the parametrization 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.

The second issue related to the implementation of a multi-species EnKF is the localization between species. It is well known in EnKF applications that a tapering of the sampling error correlations is needed when the true error correlation is not close to +1 or -1 (e.g. Anderson, 2012). For correlations that depend on distance, a widely used sampling error correction is provided by the Schur product of a compact support correlation function (Gaspari and Cohn, 1999) to the sample covariance. However, in comprehensive atmospheric chemistry models that have many prognostic chemical species, sampling errors between species at the same location are also expected to occur. 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. The approach to ~~this~~ the cross-species sampling correlation noise within a sequential data assimilation has not been fully explored yet. Several studies in EnKF chemical data assimilation use a brute force species localization that consists in zeroing-out cross-species covariances. This is the case for example for Tang et al. (2011) and Gaubert et al. (2014) where only O_3 is observed and where all cross-covariances between ozone and other species are zeroed-out in order to reduce the noise in the analysis. In an ozone assimilation study, Curier et al. (2012) kept the cross-covariances between O_3 and some other strongly coupled species, in particular NO, NO_2 and VOC's, as well as the error covariances with the boundary conditions (O_3 dry deposition and model top boundary condition). They showed that each of these kept cross-covariances give rather similar impact on the ozone analysis. Eben et al. (2005) in an multi-species air quality EnKF assimilation of surface O_3, NO and NO_2 measurements indicated that in order to reduce the sampling noise they kept the cross-covariances between these species only at the surface. In another study, Miyazaki et al. (2012) assimilated simultaneously NO_2, O_3 CO, and HNO_3 tropospheric chemical species along with the estimation of surface emissions. Using verification against satellite observations, such as innovation variance, they found that cross-covariance between chemical species need to be set to zero unless they are strongly chemically related. Examples of strongly chemically related species are members of the NOy family, or CO with VOC's. Miyazaki et al. (2012) also allowed the coupling between NO_2 and emissions of NO_2, or the CO with the emissions of CO, but set to zero the cross-covariance between emissions of NO_x and CO. Keeping the cross-covariance with the boundary conditions (surface emissions and lateral boundary conditions) was also argued in Constantinescu et al. (2007). Overall, these

studies indicate that when there is a believed strong correlation between observed and modelled species (or boundary condition) then these can be kept in the EnKF, but otherwise to reduce noise, all other cross-covariances are better be zeroed-out.

In this paper we perform an assimilation with EnKF and 4D-Var of several species in the stratosphere that are not necessarily directly chemically linked and with real-life constraints. The lifetimes of the assimilated species are quite diversified and vary with altitude. We use a state-of-the-art CTM that is in fact in constant improvement, but also has some deficiencies. We use limb sounding observations that give vertically resolved measurements, and thus there is a need to have vertically resolved error statistics. As it was shown in S14, the EnKF is more sensitive to the observation error statistics than 4D-Var assimilation. Yet, to provide a consistency between the two assimilation systems, the observation error statistics of 4D-Var will be ~~a~~ subject to the same Desroziers' estimation procedure. Localization between species, that is needed in EnKF, is in fact not applied to 4D-Var, since the cross-covariance between species are taken into account automatically using the 4D-Var adjoint model.

The paper is organized as follows. The next section describes the main components of the BASCOE Data Assimilation System (version 5.8): the common CTM, the 4D-Var system and the EnKF system. It also describes the implementation of Desroziers' method and the tuning of the error covariances in each system. The assimilated observations and independent data used to validate the results are given in Sect. 3. Section 4 describes the results of our assimilation and model experiments. And Sect. 5 discusses a separate EnKF experiment where the cross-species correlations are taken into account. Finally, some conclusions are given in Sect. 6.

2 The BASCOE Data Assimilation System

2.1 The Chemistry-Transport Model

In this study, all numerical experiments are performed with the Belgian Assimilation System for Chemical ObsErvations (BASCOE) and its underlying Chemistry Transport Model (CTM). The BASCOE CTM computes the temporal evolution of 58 stratospheric chemical species accounting for the advection, photochemical reactions and a parametrization of PSC (Polar Stratospheric Clouds) microphysics. We used a CTM configuration nearly identical to the one described by (Lefever et al., 2015) for Near Real-Time production of 4D-Var analyses as part of the MACC (Monitoring Atmospheric Composition and Climate) project. Here we provide a brief reminder of its most salient features.

All species are advected by the Flux-Form Semi-Lagrangian scheme (Lin and Rood, 1996), here driven by ERA-Interim wind fields (Dee et al., 2011). The horizontal resolution is set at 3.75° longitude by 2.5° latitude. The model considers 37 levels from the surface to 0.1 hPa, which is a subset of the 60 levels of ERA-Interim that excludes most tropospheric levels. Hence The CTM state is described by the vector $\mathbf{x} \in \Re^n$ of length $n = 96 \times 73 \times 37 \times 58 \approx 1.5 \times 10^7$. The model time step is set to 30 minutes.

The photochemical scheme of BASCOE account for 208 stratospheric chemical reactions: 146 gas-phase, 53 photolysis, 9 heterogeneous. Photolysis rates are provided by the Jet Propulsion Laboratory (JPL) recommendations (Sander et al., 2006). The computation of the photolysis rates is based on the Tropospheric Ultraviolet and Visible (TUV) radiative transfer package (Madronich and Flocke, 1999).

2.2 Setting up the time windows

In order to describe the practical implementations of the 4D-Var and EnKF algorithms in BASCOE, we must first explain the different set-up of their assimilation windows with respect to time. This is schematically shown by Fig. 1. The 4D-Var assimilation window is set to 24 h, i.e. this is the duration of the forward and backward integrations of the CTM and its adjoint. Each 4D-Var iteration is followed by a minimizing procedure (see Sect. 2.3 for more details). In this 4D-Var implementation, the 24 h forecast is defined as the first forward model simulation starting from the analysis of the previous assimilation window. All 4D-Var assimilation cycles save the model state in observation space during these forecasts, in order to compute Observation-minus-Forecast (OmF) statistics discussed below.

The EnKF initializes its ensemble of model states from one given state using a procedure described in Sect. 2.4. The EnKF assimilation is then based on ensembles of short model forecasts which have the same duration as the CTM time step, i.e. 30 minutes, followed by the observational update of each ensemble member. The updated ensemble states (analyses) are used then as initial states for the next ensemble forecast. Hence, there is no practical need to compute the 24 h forecast (green line) as in the 4D-Var approach. However we have introduced this option in the EnKF in order to allow a consistent comparison with the 4D-Var forecasts. Hence in the EnKF approach, the 24 h forecast is defined as a model simulation started from the ensemble mean analysis at 0 h UTC. As in the 4D-Var system, the 24 h forecast of the EnKF stores the OmF statistics.

2.3 The 4D-Var system

The BASCOE 4D-Var of this study was already used by S14 and is described in detail by Errera and Ménard (2012). The 4D-Var data assimilation is carried out by minimizing the so-called cost function which measures the discrepancy between the model state and observations (Talagrand and Courtier, 1987). Here, the model state vector contains 58 prognostic variables, where only 7 chemical species are observed among them, see Sect. 3.

The background error covariance matrix $\mathbf{B}_0$ is parametrized using a control variable transform

$$\mathbf{L}\xi = \mathbf{x}_0 - \mathbf{x}_0^b \equiv \delta\mathbf{x}_0, \tag{1}$$

where ξ is a new control variable, $\mathbf{x}_0$ the first guess field, $\delta\mathbf{x}_0$ is the analysis increment and $\mathbf{L}$ is the square root of $\mathbf{B}_0$:

$$\mathbf{B}_0 = \mathbf{L}^T\mathbf{L}. \tag{2}$$

As in S14, the error covariance of the first guess field expresses spatial correlations on a spherical harmonic basis (Courtier et al., 1998), allowing a representation of homogeneous and isotropic horizontal correlations by a diagonal matrix with diagonal values repeated for the same zonal wave number. The operator $\mathbf{L}$ is defined by

$$\mathbf{L} = \mathbf{\Sigma}\mathbf{S}\mathbf{\Lambda}^{1/2}, \tag{3}$$

where $\mathbf{\Sigma}$ is the (diagonal) background error standard deviation matrix with $s_b(l)\sigma_b(l)$ values on its diagonal, $s_b(l)$ is an *adjustable background error scaling factor* on the level l; $\mathbf{\Lambda}^{1/2}$ is the spatial correlation matrix, identical for each chemical

species, defined on a spherical harmonic basis hence diagonal; and $\mathbf{S}$ is the spectral transform operator from the spectral space to the model space. The spatial correlation matrix considers Gaussian correlations in the horizontal and in the vertical directions with length scales L_0^h and L_0^v in horizontal and vertical directions, respectively. The between-species covariances are assumed to be zero in the background error covariance matrix $\mathbf{B}_0$.

The observation errors are assumed to be uncorrelated both horizontally and vertically. The observation error covariance matrix $\mathbf{R}_k$ is thus defined diagonal:

$$\mathbf{R}_k(i,j) \quad = \quad \begin{cases} (s_o(i)\,\sigma_y(i)|_{t_k})^2, & \text{if } i = j \\ 0, & \text{if } i \neq j, \end{cases} \tag{4}$$

where $s_o(i)$ is an *adjustable observation error variance scaling factor* and $\sigma_y(i)|_{t_k}$ is the measurement error at level i and time t_k. The observations and their errors are described in Sect. 3. The adjustment of s_b and s_o scaling factors is performed in observation space for every observed species separately, where they are functions of vertical pressure level (see Sect. 2.5).

Finally, the BASCOE 4D-Var implementation ~~has~~ includes the background quality control procedure (BgQC, Anderson and Järvinen, 19 This procedure rejects observations when:

$$(\mathbf{y}_{i,l} - H_{i,l}(\mathbf{x}_b))^2 > \gamma^2(\mathrm{diag}(\mathbf{R})_{i,l} + H_{i,l}(\mathrm{diag}(\mathbf{B}))) \tag{5}$$

where the operator $\mathrm{diag}(\mathbf{A})$ is a diagonal matrix of $\mathbf{A}$ and i, l are the data indices of profile and level, respectively. The value of γ is set to 5 in BASCOE, so that BgQC rejects only obviously wrong observations.

2.4 The EnKF system

The BASCOE EnKF of this study is similar to the system used in S14. An ensemble of initial states $\tilde{\mathbf{x}}_i(t_0)$ is generated by adding to the model state $\mathbf{x}_0$ a set of spatially correlated perturbations according to the prescribed initial error covariance. This procedure is schematically represented on Fig. 1 on the left-hand side. The ensemble of model states is propagated forward in time using the same CTM as used in the 4D-Var (see Sect. 2.1). The ~~model~~ background error covariance is represented by the addition of a stochastic noise $\boldsymbol{\eta}_i$ to each ensemble member at each model time step. In the current implementation, the model error term is added to observed species only. The non-observed model species evolve with ensemble and are influenced by the analysis increments only implicitly through the chemistry scheme of BASCOE CTM.

The operator $\mathbf{L}$ described in Sect. 2.3 is used to generate the initial deviation $\tilde{\mathbf{x}}_i(t_0)$ and the model error $\boldsymbol{\eta}_i(t_k)$ of the EnKF system. This ensures that at the initial time, both EnKF and 4D-Var systems have identical error statistics. Initial deviation is defined as

$$\tilde{\mathbf{x}}_i(t_0) = \mathbf{L}\boldsymbol{\zeta}_i(t_0), \quad i \in [1, N], \tag{6}$$

whereas the model error term is written as

$$\boldsymbol{\eta}_i(t_k) = \alpha\mathbf{L}\boldsymbol{\psi}_i(t_k), \quad i \in [1, N], \tag{7}$$

where $\boldsymbol{\zeta}_i(t_0)$ and $\boldsymbol{\psi}_i(t_k)$ are vectors of independent normally distributed random numbers with zero mean and variance equal to 1 ~~,~~ defined in the spectral space, N is the ensemble size; and where α is an empirical model error parameter. The definition of this parameter is explained in Sect. 2.6.

The observation error covariance matrix $\mathbf{R}$ is defined by Eq. 4, where the adjustable scaling factor $s_o(i)$ is ~~also~~ estimated using the ~~Desroziers'~~ method described in Sect. 2.5. The fact that the matrix $\mathbf{R}$ is calibrated automatically without using a trial and error procedure for every observed species makes EnKF essentially easier to parametrize than in our previous study. Besides, the ~~Desroziers' methods~~ current version of EnKF allows for more accurate observation error ~~estimation~~ variance estimation with respect to S14 because it computes $s_o(i)$ as a function of observation vertical pressure level. It should be noted that EnKF uses the same background quality control procedure described in Sect. 2.3.

As in our previous study with a chemical tracer model, BASCOE EnKF uses the Schur (element-wise) product of the ensemble covariance matrix with a compact support correlation function. This function is the 5th-order piecewise rational function of Gaspari and Cohn (1999) which is isotropic and decreases monotonically with distance depending on the correlation length scale L_{loc}. The function is positive only for distances that are less than $2L_{loc}$ and zero otherwise. We applied this procedure to both horizontal and vertical correlations, using the compact support correlation functions with correlation length scales L^h_{loc} and L^v_{loc}, respectively. The choice of these parameters is discussed in Sect. 2.6. In order to make feasible the computation of much more expensive EnKF in the framework of full-chemistry model, the analyses are computed locally, around the area where current satellite observations are situated. To this end, the EnKF algorithm accounts now for a ~~new effective~~ 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) .

The EnKF analyses of this study are performed in parallel for every observed species in its own space. Thus, such analysis increments of every species do not account for cross-correlations between different chemical observations, which is not the case for the 4D-Var system. However, it is technically possible to keep all observations from multiple species in one observation space, introducing thus the cross correlations between species. An example of such EnKF data assimilation is discussed in Sect. 5.

2.5 The Desroziers' method

We use the Desroziers et al. (2005) method to estimate error variance scaling factors for each observed species and each vertical levels. The diagnosis relies on linear estimation theory where the statistics is computed using observation-minus-background, observations-minus-analysis, and analysis-minus-background differences. The estimation of the background error variance is written as

$$s_b(l)^2\sigma_b(l)^2 = \langle(\mathbf{d}_b^a)^T\mathbf{d}_b^o\rangle, \tag{8}$$

and the observation error variance is then

$$s_o(l)^2\sigma_o(l)^2 = \langle(\mathbf{d}_a^o)^T\mathbf{d}_b^o\rangle, \tag{9}$$

where the vector $\mathbf{d}_b^a$ is the difference between analysis and background, $\mathbf{d}_b^o$, the difference between observations and background, $\mathbf{d}_a^o$, the difference between observations and analysis in observation space, and $\langle\rangle$ denote the mathematical expectation. Note that in practical implementation, the expectation is replaced by a horizontal mean and time mean of a day.

The BASCOE data assimilation is initialized using $s_b(l) = 1$ and $s_b(l) = 1$ for both, EnKF and 4D-Var. These initial values are kept in the system for the first 24h of system integration. The analysis increment and model innovation statistics are accumulated during this time. Then the estimation of the scaling factors is performed using expressions 8 and 9. The following 24h analyses (on day 2), both EnKF and 4D-Var use the day 1 estimated error variance scaling factors. The procedure is sequentially updated every 24h assimilation cycle using the statistics accumulated during the previous cycle. Note that we could have used an estimated $s_b(l)$ in EnKF to tune the model error, but we decided not to apply it and use the χ^2 tendency to this end (see next section, Sect. 2.6).

The Desroziers' estimates appear to be asymptotically stable after only one day. That means that changing the initial parameter value has little to no effect on the resulting time series of estimated parameter values. Figure 2, shows examples of the observation error ~~variance~~ standard deviation for O_3, H_2O and HCl (at each vertical levels) when we perturb the initial parameter value by a factor 1.5. The dashed curves represent the initial values and the solid curves the values estimated after one day using the Desroziers' method.

2.6 Tuning of error covariances in the two systems

Each assimilation system (i.e. EnKF and 4D-Var) has its own optimized error variances but shares a common error correlation for the prescribed $\mathbf{B}_0$ in 4D-Var and the prescribed model and initial condition error correlations in EnKF. As in S14, our starting point is the calibration of the error covariance matrix $\mathbf{B}_0$ used by the 4D-Var system. This is realized through a calibration of the spatial correlation associated with the operator $\mathbf{L}$ described in Sect. 2.3. The operator $\mathbf{L}$ has similar parameters as in S14: $L_0^h = 800$ km and $L_0^v = 1$ model level, $\sigma_b(l) = 0.2$ (with scaling factor $s_b(l) = 1$) at all levels. Then, we take into account the fact that the use of the Schur product results in shorter correlation length scales (See S14 for more details). Similarly to our previous work, EnKF uses an effective correlation length scales of $L_e^h = 872$ km and $L_e^v = 1.3$ in model level coordinates, given that $L_{loc}^h = 2000$ km and $L_{loc}^v = 1.5$ are chosen a priori. The calibrated operator $\mathbf{L}$ is then used in the EnKF system.

The observation error variance scaling factor s_o (see Sect. 2.4) is estimated for both systems using the Desroziers' method described in Sect. 2.5. The background error ~~scaling factor s_o~~ covariance scaling factor s_b used in 4D-Var is also estimated using the Desroziers' method. In EnKF, the background error covariance is evolved using CTM where we add a model error term that uses a calibration parameter α. The value of α equal to 0.025 was found in S14 in the case study of O_3 tracer. This value is based on the property that the time tendency (over periods of weeks and months) of the χ^2 diagnostic should be nearly zero as argued in Ménard and Chang (2000). In general, we have found (in S14) that the value of α changes the slope of the O_3

χ^2 distribution, whereas the observation error variance scaling factor s_o is responsible for the mean value of it. In the absence of better knowledge, we use the value $\alpha = 0.025$ for all observed species described in Sect. 3.

The performance of each data assimilation system of BASCOE can be monitored by the χ^2 diagnostic. During the whole period of our experiments, $\langle \chi^2_k \rangle / m_k$ values remain close to 1 (result not shown). This is achieved by using the error variances estimated by the Desroziers' method. In the case of 4D-Var where both observation and the background error variances are estimated, the Desroziers' method gives estimates that achieve the innovation variance consistency (Ménard, 2016). For EnKF where only the observation error variance is estimated, the fact that $\langle \chi^2_k \rangle / m_k$ values remain close to 1 is an indirect confirmation that the model error is tuned appropriately. Figure 3 shows the evolution of the adjustable parameters for both systems. The solid lines show the vertically mean values of the observation variance parameter s_o and the dashed lines, the vertically mean values of the 4D-Var background error variances. 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~~, and thus goes along in arguing that they represent some true error statistics.~~. 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 α 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.

Finally, we wish to remark that to keep comparable CPU costs in both data assimilation systems, and that can be carried out in a reasonable time, 4D-Var is run with 10 iterations (including 10 adjoint iterations) and the EnKF uses 20 ensemble members. As in S14, the computation of the EnKF Kalman gain is performed using Cholesky decomposition where the full observation vector is considered at a given time step for a given species. No simplification is used to compute the inversion of the innovation matrix $[\mathbf{H}\mathbf{B}_e\mathbf{H}^T + \mathbf{R}]$ or the matrix $\mathbf{B}_e\mathbf{H}^T$ (see S14 for more details). The actual use of local domain decomposition and integration of the ensemble members on different processors in parallel decreases essentially the CPU costs as compared to the previous version of BASCOE EnKF.

3 Observations

The dataset assimilated in this study is the version 4.2 of the retrievals from the Microwave Limb Sounder (MLS) on-board the EOS (Earth Observing System) Aura satellite (Livesey et al., 2015). 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. Here we assimilate the retrievals of five species which are listed in Table 1 along with some key parameters of the dataset and the validation reference for each species.

Some results of the data assimilation will be validated against independent observations. This will be the case for N_2O because the Aura MLS N_2O precision is 24-14 ppbv (9-38%, relative to the observation mean at given altitude) and its accuracy is 70-3 ppbv (9-25%) in the pressure range 100-4.6 hPa but its precision drops to 14 ppbv (250%) at 1 hPa, where the accuracy is estimated to 16% (Livesey et al., 2015). Hence we will validate the BASCOE N_2O. with observations retrieved from the Atmospheric Chemistry Experiment - Fourier Transform Spectrometer (ACE-FTS) satellite instrument (Bernath et al., 2005)

Table 1. List of species retrieved in Aura-MLS v4.2 and assimilated for this paper.

Name	Resolution (km) Vertical x Horizontal	Vertical range of assimilation (hPa)	Accuracy	Precision	Validation paper (for Aura-MLS v2.2)
O_3	3 - 6 x 200 - 300	0.1 - 261	3 - 20%	2 - 40%	Froidevaux et al. (2008a)
N_2O	4 - 6 x 300 - 620	4.64 (1) - 100	9 - 25% (12%)	7 - 38% (250%)	Lambert et al. (2007)
H_2O	3 - 4 x 220 - 440	0.1 - 316	4 - 11%	4 - 9%	Lambert et al. (2007)
HNO_3	3 - 5 x 300 - 500	1.5 - 215	$\pm$0.5 - 2 ppbv	$\pm$0.7 ppbv	Santee et al. (2007)
HCl	3 - 6 x 200 - 400	0.32 - 100	5 -50%	10 - 50%	Froidevaux et al. (2008b)

which uses solar occultation to provide around 28 profiles per day. Strong et al. (2008) validated N_2O retrievals from ACE-FTS (version 2.2) and found a bias of $\pm$15% between 6-30 km and a bias of $\pm$4 ppbv between 30-60 km . Here we use the retrieval version 3.5.

We will also use the Michelson Interferometer for Passive Atmospheric Sounding (MIPAS) retrievals by the IMK/IAA (Institut für Meteorologie und Klimaforschung, Karlsruhe/Instituto de Astrofisica de Andalucia, Grenada) to validate the unconstrained distributions of CH_4 and NO_x (NO_x= NO + NO_2). The MIPAS IMK/IAA retrievals of CH_4were validated by Laeng et al. (2015) and the retrievals of NO_x were described by Funke et al. (2005).

4 Numerical experiments

This section reports the numerical experiments performed in this study: the control run, i.e. an unconstrained simulation by the BASCOE CTM including photochemistry; the "EnKF" and "EnKF tracer" experiments, the first one including photochemistry and the second one neglecting it (i.e. assimilation in chemical tracer mode as done in S14); and the two corresponding "4D-Var" and "4D-Var tracer" experiments. All experiments start on 1 April 2008 from the same initial condition, i.e. a 4D-Var analysis of Aura-MLS retrievals (Lefever et al., 2015), and end on 1 November 2008 i.e. after 7 months.

The results of our model and data assimilation experiments will be assessed using Observations-minus-Forecast (OmF) statistics, relative bias and standard deviation, computed in the observation space. In the case of N_2O the relative bias and standard deviation are not good diagnostics because its volume mixing ratio decreases by two orders of magnitude between 100 and 1 hPa. Hence we will simply compare the mean profiles of N_2O by the five numerical experiments with assimilated and independent measurements. The statistics are computed in three different latitude bands covering the globe: South Pole (90°S-60°S), middle latitudes and Tropics (60°S-60°N) and North Pole (60°N-90°N). The analyses of the assimilated species are verified by comparison with the assimilated observations (sections 4.1–4.5). Section 4.6 evaluates the results for methane and nitrogen oxides which are not assimilated.

4.1 Verification of ozone

Figure 4 shows the OmF statistics for ozone over the period September 2008-October 2008, i.e. during the period of the Antarctic ozone hole. The results of the tracer experiments are not shown above 1 hPa because the tracer approximation is not valid in this region. The CTM experiment delivers rather large biases (10 to 30%) in the lower and upper stratosphere and at all levels above the South Pole region.

All data assimilation experiments succeed in eliminating these biases nearly completely in the lower and middle stratosphere. The resulting biases are smaller than 2%, except for the 4D-Var experiment which overestimates ozone depletion in the Antarctic ozone hole region (around 50 hPa) by up to 5%. Compared with the CTM results, the 4D-Var and EnKF experiments also reduce significantly the OmF standard deviation in the lowest levels. The smallest OmF standard deviations are delivered by the 4D-Var experiment, with results about 1% smaller than those delivered by the EnKF in pressure range 30-2 hPa.

The experiments 4D-Var tracer and EnKF tracer allow us to assess the impact of stratospheric chemistry. Neglecting this process results in larger biases and OmF standard deviations above the South Pole in the region 10-2 hPa, where both tracer data assimilation systems overestimate ozone by $\sim$5% and deliver OmF standard deviations reaching 10%.

The photochemical lifetime of ozone decreases rapidly in the upper stratosphere and reaches values as short as a few minutes (Brasseur and Solomon, 2005) in the lower mesosphere. In these regions, our CTM experiment has a significant ozone deficit reaching about 20% at the stratopause (1 hPa). The sources of this model bias are out of scope of the present paper. However, its presence helps to assess the behaviour of our assimilation algorithms. It is found that both data assimilation algorithms fail to correct this model bias: ozone is still underestimated by $\sim$15% at the Stratopause. In the upper stratosphere and mesosphere, data assimilation does not improve OmF standard deviations either: these remain nearly identical to those obtained by the CTM. These results indicate that when the photochemical lifetime is short (e.g. smaller than the time step of the model) and the model error is important, both data assimilation systems fail to improve the representation of the model state. Since this issue also involves species that have strong chemical interactions with ozone, it will be further discussed in Sections 4.2, 4.4 and 4.6.

4.2 Verification of HCl

During the largest part of the CTM simulation, the HCl distribution is in agreement with the Aura-MLS observations. And the EnKF and 4D-Var experiments deliver nearly identical results where the small CTM biases are completely corrected (not shown). The only exception is in the South Pole latitude band, during the period May-June 2008 which is shown on figure 5. During this period the chemical lifetime of HCl is much shorter than at other latitudes, because the heterogeneous removal due to the formation of Polar Stratospheric Clouds has already started. This loss process is currently overestimated in the BASCOE CTM, due to a crude cold-point temperature parametrisation (section 2.2.2 in Lefever et al., 2015). As a result, the CTM experiment underestimates HCl by up to 45% at 30 hPa in the Antarctic polar vortex region and its OmF standard deviation also reaches $\sim$ 45% . While the 4D-Var approach essentially fails to correct this large disagreement, the bias is nearly halved in the EnKF experiment and the OmF standard deviation is significantly reduced as well.

Staying in the lower stratosphere (100-10 hPa), the outcome of the experiments is different than above the South Pole. Northward of 60°S, the CTM biases do not exceed 15% and they are nearly eliminated by both data assimilation experiments. The OmF standard deviations of both data assimilations are also quite similar in these regions.

In the middle stratosphere, the chemical lifetime of HCl decreases from about one week at 10 hPa to about one day at 1 hPa (Brasseur and Solomon, 2005). The CTM experiment delivers quite accurate results in this region: the OmF biases do not exceed 3% and the standard deviations are less than 10%, in every latitude band for the pressure range 10-0.46 hPa. The EnKF and 4D-Var experiments both succeed in correcting these small CTM biases and reducing the OmF standard deviations, except at the 1 hPa level where the 4D-Var does not correct the CTM deficit of 3% for HCl.

4.3 Verification of HNO_3

Figure 6 shows the HNO_3 OmF statistics between the assimilated Aura MLS data and the CTM, EnKF and 4D-Var experiments for the period September-October 2008. For all three latitude bands, the CTM shows a significant underestimation reaching 20-25% around 30 hPa. This model bias nearly disappears at 10 hPa but grows again above this level. In the lower stratosphere, the OmF standard deviations of the CTM experiment reach minimum values of 10-15% but at the lowermost levels the standard deviation is much larger in the Antarctic polar vortex region than at other latitudes.

Both data assimilation experiments correct the OmF model bias at all latitudes and at all pressure levels between 100 and 10 hPa. Above that level, the quickly increasing model OmF bias is not corrected by either assimilation algorithm. The explanation for this different behaviour in the upper stratosphere is twofold. First, the observation error grows quickly with altitude, reducing the weight of observations in the assimilation experiments. Second, a large discrepancy between the model and the observed data leads to rejection of most measurements above 10 hPa by the background quality control procedure (see Sect. 2.3 for more details).

The 4D-Var OmF bias is generally less than 3% in the pressure range 100-7 hPa, except for an 8% OmF bias at 70 hPa in the Tropics. The EnKF delivers even smaller OmF biases in the whole pressure range and at all latitudes. Both data assimilations results in almost identical OmF standard deviations, except in the Antarctic polar vortex region where the EnKF errors are slightly larger below 20 hPa.

4.4 Verification of water vapour

Water vapour is a long-lived tracer in the whole stratosphere, with a photochemical lifetime still longer than one month at the stratopause (Brasseur and Solomon, 2005). The OmF statistics for H_2O are shown on Fig. 7. The CTM provides OmF biases smaller than 10% in the whole pressure range and at all latitudes, except in the Antarctic polar vortex between 100 and 10 hPa, where H_2Ounderestimation reaches 30%. The OmF standard deviation by the CTM is also largest in this region, reaching 23% while it does not exceed 15% elsewhere.

Both data assimilations mostly correct the OmF bias and standard deviation errors with respect to the CTM. Their OmF biases do not exceed 2%, except for the OmF bias by the 4D-Var which reaches 3% at 1 hPa, i.e. the level where the ozone

deficit described in Sect. 4.1 is maximum. The OmF standard deviation errors resulting from the two assimilation experiments are also quite similar, with slightly larger EnKF errors in the Antarctic polar vortex below 10 hPa.

4.5 Verification of N_2O

The relative error statistics shown for other species are difficult to interpret in the case of N_2O because its volume mixing ratio decreases by two orders of magnitude between 100 and 1 hPa. Hence figure 8 simply compares mean profiles of forecasts and observations. We display the assimilated Aura MLS observations along with their validation uncertainties (grey filled region as reported by Lambert et al., 2007). Since these uncertainties are very large in the upper stratosphere, we also compare with independent observations by the ACE-FTS solar occultation instrument (see Sect. 3).

In the lower stratosphere the two satellite datasets and the CTM experiment are in good agreement. Above 10 hPa the mixing ratios retrieved from Aura-MLS are much larger than those from ACE-FTS, and above 5 hPa they become pressure-independent which is not realistic. As expected, the CTM experiment agrees much better with the ACE-FTS N_2O retrievals since they are much more precise in the upper stratosphere.

How do the 4D-Var and EnKF treat the ~~tricky~~ Aura-MLS ~~dataset~~N_2O dataset containing a bias? To answer this question we inhibited any a priori filtering of the Aura MLS observations of N_2O above 5 hPa, and we used both the full chemistry CTM and its transport-only version. Figure 8 shows that both EnKF experiments follow the assimilated Aura MLS data in the upper stratosphere, whereas the mean profile delivered by the 4D-Var experiment remains closer to the CTM. This is due to the automatic rejection by the 4D-Var of most Aura-MLS observations of N_2O above 5 hPa. However, 4D-Var assimilation with a chemical tracer transport model (cyan dashed curve) assimilates more Aura MLS data ~~approaching closer to them~~keeping the model closer to the assimilated observations. This episode reveals the role of chemistry in the multivariate assimilation: it acts as a strong constraint within 4D-Var, preventing it from assimilating erroneous observations.

4.6 Evaluation of non-observed species

Finally, the forecasts of two non-observed species issued from both data assimilation systems will be validated, CH_4 and NO_x, the sum of NO_2 and NO, (Fig. 9). CH_4 CTM forecasts agree well with the MIPAS IMK/IAA data. And both data assimilation system keep generally this agreement, except the region around 2-1 hPa where 4D-Var develops an artificial bias related to the presence of O_3 model bias and the fact that O_3 data were assimilated in the upper stratosphere. As we saw this before, 4D-Var tends to develop such biases in many assimilated and non-assimilated species to compensate the O_3 bias. The problem of model O_3 bias is out of scope of the present article. We should only note that it can not be solved directly by data assimilation without an improved version of CTM~~(see Skachko et al, the paper in preparation to ACP for more details)~~. NO_x CTM forecasts are essentially different with data due to absent NO_x sources in the model. As for CH_4, both data assimilation keep the model state unperturbed, except the region around 2-1 hPa where 4D-Var develops a bias for the same reason as stated above. ~~Accidentally~~Incidentally, this bias provides better agreement between the model and data in this region.

5 EnKF with cross correlations between species

All the EnKF experiments done so far used a brute force species localization, in other words, the sample covariance between species is set to zero. This type of localization should not be confused with the localization based on distance for the same species, which we keep. Now let us see what happens when we keep the sample cross-covariance intact.

To this end, we conducted an experiment where we assimilate O_3 and N_2O, two species that are not strongly related via the chemistry system. We will call this experiment the EnKF-CC, standing for EnKF with Cross-Covariances. In principle, we would expect that an observation of O_3 does not change significantly N_2O and vice versa. In EnKF-CC, O_3 and N_2O are put into a common observation space, defined by the observation vector $\mathbf{y}$ and the observation error covariance matrix $\mathbf{R}$. The ensemble of model vectors in observation space $H\mathbf{x}_i$ contain thus two blocks of O_3 and N_2O. This provides the cross-correlation terms in the ~~model~~ background error covariance matrix in observation space $H\mathbf{B}H^T$ computed as

$$H\mathbf{B}H^T = \sum_i^N H(\mathbf{x}_i - \bar{\mathbf{x}})H(\mathbf{x}_i - \bar{\mathbf{x}})^T, \tag{10}$$

where $i \in [1, N]$ is the number of ensemble member, N is the ensemble size, $\bar{\mathbf{x}}$ is the ensemble mean (see S14 for more details). The cross correlation terms between species remain after the localization of the error ~~variance~~ covariance via the Schur product, because it filters out only the spurious spatial correlations.

Figure 10 shows an example of such EnKF-CC data assimilation comparing its results with EnKF discussed in the previous sections where the sample covariance between species was set to zero. The figure shows the O_3 OmF bias and standard deviation for EnKF (red) and EnKF-CC (yellow) analyses during 24h of September 15 2008. We observe that the EnKF-CC has noisy bias and an increased and noisy error standard deviations in the OmF correlations compared with the EnKF experiment. A similar kind of impact is also obtained when we assimilate only one species and examine the OmF of the other non-observed species (results not shown). We thus conclude, as other studies have indicated, that the sample cross-covariance between weakly chemically related species, give rise to spurious analysis increments with a deterioration of the overall performance of the assimilation system.

6 Conclusions

We have conducted a comparison of an EnKF and 4DVar data assimilation system using a comprehensive stratospheric chemical transport model. We considered 4D-Var and EnKF configurations that are normally used for chemical data assimilation applications. Both data assimilation systems have online estimation of error variances based on the Desroziers' method and share the same correlation model for all prescribed error correlations (i.e. the background error covariance for 4D-Var, initial error and model error for EnKF) so that each data assimilation system is nearly optimal and can also be compared to each other. A previous comparison study by Skachko et al. (2014) showed that for chemical tracer transport only both assimilation system provide results of essentially similar quality despite the difference in practical implementation of each method: the 4D-Var was applied in its strong constraint formulation with a 24 h assimilation window with the assumption of no model

error over this period, whereas the EnKF was used to sequentially assimilate observations every 30 minutes with model error perturbations added every 30 minutes. This study examines in what way the inclusion of chemistry changes the performance of the assimilation ~~system~~systems, but perhaps more importantly how an EnKF and a 4D-Var chemical data assimilations can be implemented in a real-life situation with several modelled and assimilated species. In this study we assimilate ozone, HCl, HNO_3, H_2O and N_2O observations from EOS Aura-MLS.

In the context of atmospheric chemistry, EnKF and 4D-Var differ in a number of ways. While 4D-Var, built on the assumption of a perfect model, tries to find a strong constraint solution that fits observations over a 24h window, EnKF on the other hand provides estimates at each model time step but allows for modelling error (mainly as a ~~model~~ background error covariance). Furthermore, while 4D-Var infers information based on error correlation between observed and non-observed species, EnKF, on the other hand, introduces noise between weakly chemically-related species, and so far in practice, these cross-species error covariances are set to zero. So the question is: to which extent the chemical modelling is an important component of the analysis? The implementation of a multi-species sequential chemical data assimilation is challenging by the need to properly tune and automate the estimation of a large number of input error parameters.

The comparison done in this paper shows that, in general, there is not a significant improvement in the OmF statistics of the system when the cross-correlation between species is kept (4D-Var) versus the EnKF system where the cross-species error correlation has been filtered out. Differences do occur, however, when there is an important chemical modelling error or when there are large biases between model and observed values.

For example, the BASCOE CTM has an important model O_3 deficit near or above 1 hPa. The source of this model bias is unclear and is not discussed in this paper. The experiments show however that assimilating O_3 at these altitudes gives a poor agreement with observations. At these altitudes the chemical life-time of O_3 is smaller than the time step of the model and consequently, any correction on the O_3 concentrations by the assimilation of O_3 measurements simply cannot correct for the model error. For the other species, such as HCl and HNO_3, the OmF statistics for EnKF are always better than for the 4D-Var. 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. On the other hand, the cross-species covariances, implicit to a 4D-Var assimilation system, have a negative effect in the presence of strong model O_3 bias. The 4D-Var system tries to compensate the bias and thus develops small artificial biases in many chemically related with O_3 species, observed and non-observed. This is shown using OmF statistics for two observed species, HCl and H_2O, and two non-observed, CH_4 and the NO_x family.

The effect of large observation biases has a very different impact. For example, the EOS Aura MLS N_2O has significant biases above 4 hPa. In this case, EnKF reaches the state close to observations from the first observation updates during the spin up phase, and keeps model close to observations afterwards because of short ensemble model forecasts and frequent observational updates. On the other hand, 4D-Var appears to be robust to erroneous observations. A significant number of observations are rejected by the quality control, and in the end, 4D-Var provides analyses with more weight given to the model forecast rather than to the observations.

We have also examined the need to have cross-species localisation in an EnKF. Our study shows that the simultaneous assimilation of O_3 and N_2O, two species that are only weakly chemically related, gives rise to spurious cross-species error correlations that deteriorates the performance of EnKF, and it is then better to simply ignore those error correlations. To have a more sensible approach to species localization could be the object of future work.

An important aspect of this study is the implementation of an online estimation of error variance parameters. The estimation of observation error variance and, in addition the background error variance for 4D-Var is done at each observation vertical level, using the Desroziers' method. The variance parameters being estimated are in fact very robust over time, showing little variability one day to the next.

Finally all the experiments were done with comparable wall clock time for EnKF and 4D-Var settings.

The study has also some limitations. An acknowledged difficulty often encountered in chemical data assimilation is the situation where both the model and the observations suffer from significant biases. This is the case for example with the BASCOE CTM CO and ClO when using the Aura MLS datasets. Solving this problem represents a challenging task that we have not conducted here, and would necessitate a dedicated study. Another limiting factor is the correlation length used in this study. We have not attempted to estimate it, but rather have used what appears to be a reasonable value from past 4D-Var experiments. The estimated error variances and thus the weight given to the observations are also linked to the correctness of the error correlation, and this issue could also be investigated further. A future development of the BASCOE chemical data assimilation system would be a hybrid 4D-EnKF approach using the ensemble of models to construct a 4D background error covariance matrix.

Other possibilities may be considered to properly compare two essentially different data assimilation systems. For example, the 4DEnKF (Hunt et al., 2004) approach could be used that computes 4D error covariances from the ensemble of forecasts at several times within the assimilation window. This would allow a longer assimilation window to be used in the EnKF experiment, making it more comparable to the configuration of 4D-Var. In this case, the EnKF analysis would be forced to simultaneously fit all of the observations distributed over a longer window, while still satisfying the model equations, as in 4D-Var. Applying the model error in this 4DEnKF only at the beginning of each assimilation window would make it similar to the strong-constraint version of the 4D-Var that was used.

7 Code availability

~~The numerical code of BASCOE CTM as well as two data assimilation methods, 4D-Var and EnKF , are provided upon an email request to the authors~~ Readers interested in the BASCOE code can contact the developers through http://bascoe.oma.be

Author contributions. S. Skachko and R. Ménard designed the experiments and S. Skachko carried them out. Q. Errera developed the codes of 4D-Var and the Desroziers' method. S. Skachko and Y. Christophe developed the EnKF code. S. Chabrillat and Y. Christophe worked on the CTM code. S. Skachko prepared the manuscript with contributions from all co-authors.

Acknowledgements. This research was financially supported at BIRA-IASB by the Belspo/ESA/PRODEX programme. ~~S.Skachko acknowledges fruitful discussions~~ The authors thank the ISSI (International Space Science Institute) for the funding of two work meetings in Bern to prepare the present article within the international "Study group on the added-value of chemical data assimilation in the stratosphere and upper-troposphere"~~at ISSI (International Space Science Institute) in Bern.~~ . The authors also thank three anonymous referees for their useful comments that have essentially improved the article.

References

Anderson, E. and Järvinen, H.: Variational quality control, Q. J. R. Meteorol. Soc., 125, 697–722, doi:10.1002/qj.49712555416, 1999.

Anderson, J. L.: Localization and Sampling Error Correction in Ensemble Kalman Filter Data Assimilation, Monthly Weather Review, 140, 2359–2371, doi:10.1175/MWR-D-11-00013.1, http://dx.doi.org/10.1175/MWR-D-11-00013.1, 2012.

Bernath, P. F., McElroy, C. T., Abrams, M. C., Boone, C. D., Butler, M., Camy-Peyret, C., Carleer, M., Clerbaux, C., Coheur, P.-F., Colin, R., DeCola, P., DeMazière, M., Drummond, J. R., Dufour, D., Evans, W. F. J., Fast, H., Fussen, D., Gilbert, K., Jennings, D. E., Llewellyn, E. J., Lowe, R. P., Mahieu, E., McConnell, J. C., McHugh, M., McLeod, S. D., Michaud, R., Midwinter, C., Nassar, R., Nichitiu, F., Nowlan, C., Rinsland, C. P., Rochon, Y. J., Rowlands, N., Semeniuk, K., Simon, P., Skelton, R., Sloan, J. J., Soucy, M.-A., Strong, K., Tremblay, P., Turnbull, D., Walker, K. A., Walkty, I., Wardle, D. A., Wehrle, V., Zander, R., and Zou, J.: Atmospheric Chemistry Experiment (ACE): Mission overview, Geophysical Research Letters, 32, n/a–n/a, doi:10.1029/2005GL022386, http://dx.doi.org/10.1029/2005GL022386, l15S01, 2005.

Bocquet, M., Elbern, H., Eskes, H., Hirtl, M., Žabkar, R., Carmichael, G. R., Flemming, J., Inness, A., Pagowski, M., Pérez Camaño, J. L., Saide, P. E., San Jose, R., Sofiev, M., Vira, J., Baklanov, A., Carnevale, C., Grell, G., and Seigneur, C.: Data assimilation in atmospheric chemistry models: current status and future prospects for coupled chemistry meteorology models, Atmospheric Chemistry and Physics, 15, 5325–5358, doi:10.5194/acp-15-5325-2015, http://www.atmos-chem-phys.net/15/5325/2015/, 2015.

Brasseur, G. and Solomon, S.: Aeronomy of the middle atmosphere : chemistry and physics of the stratosphere and mesosphere, Springer Netherlands, doi:10.1007/1-4020-3824-0, previous ed.: Dordrecht : Reidel, 1986, 2005.

Buehner, M., Houtekamer, P. L., Charette, C., Mitchell, H. L., and He, B.: Intercomparison of Variational Data Assimilation and the Ensemble Kalman Filter for Global Deterministic NWP. Part I: Description and Single-Observation Experiments, Mon. Weather Rev., 138, 1550–1566, 2010a.

Buehner, M., Houtekamer, P. L., Charette, C., Mitchell, H. L., and He, B.: Intercomparison of Variational Data Assimilation and the Ensemble Kalman Filter for Global Deterministic NWP. Part II: One-Month Experiments with Real Observations, Mon. Weather Rev., 138, 1567–1586, 2010b.

Cohn, S. E.: Dynamics of Short-Term Univariate Forecast Error Covariances, Mon. Weather Rev., 121, 3123–3148, 1993.

Constantinescu, E. M., Chai, T., Sandu, A., and Carmichael, G. R.: Autoregressive models of background errors for chemical data assimilation, Journal of Geophysical Research: Atmospheres, 112, n/a–n/a, doi:10.1029/2006JD008103, http://dx.doi.org/10.1029/2006JD008103, d12309, 2007.

Courtier, P., Andersson, E., Heckley, W., Pailleux, J., Vasiljevic, D., Hamrud, M., Hollingsworth, A., Rabier, F., and Fisher, M.: The ECMWF implementation of three-dimensional variational assimilation (3D-Var). I: Formulation, Q. J. R. Meteorol. Soc., 124, 1783–1807, 1998.

Curier, R., Timmermans, R., Calabretta-Jongen, S., Eskes, H., Segers, A., Swart, D., and Schaap, M.: Improving ozone forecasts over Europe by synergistic use of the LOTOS-EUROS chemical transport model and in-situ measurements, Atmospheric Environment, 60, 217 – 226, doi:http://dx.doi.org/10.1016/j.atmosenv.2012.06.017, http://www.sciencedirect.com/science/article/pii/S1352231012005687, 2012.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Q. J. R. Meteorol. Soc., 137, 553–597, 2011.

Desroziers, G., Berre, L., Chapnik, B., and Poli, P.: Diagnosis of observation, background and analysis-error statistics in observation space, Q. J. R. Meteorol. Soc., 131, 3385–3396, doi:10.1256/qj.05.108, 2005.

Eben, K., Juruş, P., Resler, J., Belda, M., Pelikán, E., Krüger, B. C., and Keder, J.: An ensemble Kalman filter for short-term forecasting of tropospheric ozone concentrations, Quarterly Journal of the Royal Meteorological Society, 131, 3313–3322, doi:10.1256/qj.05.110, http://dx.doi.org/10.1256/qj.05.110, 2005.

Errera, Q. and Ménard, R.: Technical Note: Spectral representation of spatial correlations in variational assimilation with grid point models and application to the Belgian Assimilation System for Chemical Observations (BASCOE), Atmospheric Chemistry and Physics, 12, 10 015–10 031, doi:10.5194/acp-12-10015-2012, http://www.atmos-chem-phys.net/12/10015/2012/, 2012.

Errera, Q., Daerden, F., Chabrillat, S., Lambert, J. C., Lahoz, W. A., Viscardy, S., Bonjean, S., and Fonteyn, D.: 4D-Var assimilation of MIPAS chemical observations: ozone and nitrogen dioxide analyses, Atmos. Chem. Phys., 8, 6169–6187, 2008.

Froidevaux, L., Jiang, Y. B., Lambert, A., Livesey, N. J., Read, W. G., Waters, J. W., Browell, E. V., Hair, J. W., Avery, M. A., McGee, T. J., Twigg, L. W., Sumnicht, G. K., Jucks, K. W., Margitan, J. J., Sen, B., Stachnik, R. A., Toon, G. C., Bernath, P. F., Boone, C. D., Walker, K. A., Filipiak, M. J., Harwood, R. S., Fuller, R. A., Manney, G. L., Schwartz, M. J., Daffer, W. H., Drouin, B. J., Cofield, R. E., Cuddy, D. T., Jarnot, R. F., Knosp, B. W., Perun, V. S., Snyder, W. V., Stek, P. C., Thurstans, R. P., and Wagner, P. A.: Validation of Aura Microwave Limb Sounder stratospheric ozone measurements, Journal of Geophysical Research: Atmospheres, 113, doi:10.1029/2007JD008771, http://dx.doi.org/10.1029/2007JD008771, 2008a.

Froidevaux, L., Jiang, Y. B., Lambert, A., Livesey, N. J., Read, W. G., Waters, J. W., Fuller, R. A., Marcy, T. P., Popp, P. J., Gao, R. S., Fahey, D. W., Jucks, K. W., Stachnik, R. A., Toon, G. C., Christensen, L. E., Webster, C. R., Bernath, P. F., Boone, C. D., Walker, K. A., Pumphrey, H. C., Harwood, R. S., Manney, G. L., Schwartz, M. J., Daffer, W. H., Drouin, B. J., Cofield, R. E., Cuddy, D. T., Jarnot, R. F., Knosp, B. W., Perun, V. S., Snyder, W. V., Stek, P. C., Thurstans, R. P., and Wagner, P. A.: Validation of Aura Microwave Limb Sounder HCl measurements, Journal of Geophysical Research: Atmospheres, 113, n/a–n/a, doi:10.1029/2007JD009025, http://dx.doi.org/10.1029/2007JD009025, d15S25, 2008b.

Funke, B., López-Puertas, M., von Clarmann, T., Stiller, G. P., Fischer, H., Glatthor, N., Grabowski, U., Höpfner, M., Kellmann, S., Kiefer, M., Linden, A., Mengistu Tsidu, G., Milz, M., Steck, T., and Wang, D. Y.: Retrieval of stratospheric NOx from 5.3 and 6.2 nm nonlocal thermodynamic equilibrium emissions measured by Michelson Interferometer for Passive Atmospheric Sounding (MIPAS) on Envisat, Journal of Geophysical Research: Atmospheres, 110, n/a–n/a, doi:10.1029/2004JD005225, http://dx.doi.org/10.1029/2004JD005225, d09302, 2005.

Gaspari, G. and Cohn, S. E.: Construction of correlation functions in two and three dimensions, Q. J. R. Meteorol. Soc., 125, 723–757, 1999.

Gaubert, B., Coman, A., Foret, G., Meleux, F., Ung, A., Rouil, L., Ionescu, A., Candau, Y., and Beekmann, M.: Regional scale ozone data assimilation using an ensemble Kalman filter and the CHIMERE chemical transport model, Geoscientific Model Development, 7, 283–302, doi:10.5194/gmd-7-283-2014, http://www.geosci-model-dev.net/7/283/2014/, 2014.

Gonnet, G. H. and Baeza-Yates, R.: Handbook of Algorithms and Data Structures: In Pascal and C (2Nd Ed.), Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1991.

Hunt, B. R., Kalnay, E., Kostelich, E. J., Ott, E., Patil, D. J., Sauer, T., Szunyogh, I., Yorke, J. A., and Zimin, A. V.: Four-dimensional ensemble Kalman filtering, Tellus A, 56, 273–277, doi:10.1111/j.1600-0870.2004.00066.x, http://dx.doi.org/10.1111/j.1600-0870.2004.00066.x, 2004.

Kalnay, E., Li, H., Miyoshi, T., Yang, S.-C., and Ballabrera-Poy, J.: 4-D-Var or ensemble Kalman filter?, Tellus A, 59, 758–773, 2007.

Laeng, A., Plieninger, J., von Clarmann, T., Grabowski, U., Stiller, G., Eckert, E., Glatthor, N., Haenel, F., Kellmann, S., Kiefer, M., Linden, A., Lossow, S., Deaver, L., Engel, A., Hervig, M., Levin, I., McHugh, M., Noël, S., Toon, G., and Walker, K.: Val-

idation of MIPAS IMK/IAA methane profiles, Atmospheric Measurement Techniques, 8, 5251–5261, doi:10.5194/amt-8-5251-2015, http://www.atmos-meas-tech.net/8/5251/2015/, 2015.

Lambert, A., Read, W. G., Livesey, N. J., Santee, M. L., Manney, G. L., Froidevaux, L., Wu, D. L., Schwartz, M. J., Pumphrey, H. C., Jimenez, C., Nedoluha, G. E., Cofield, R. E., Cuddy, D. T., Daffer, W. H., Drouin, B. J., Fuller, R. A., Jarnot, R. F., Knosp, B. W., Pickett, H. M., Perun, V. S., Snyder, W. V., Stek, P. C., Thurstans, R. P., Wagner, P. A., Waters, J. W., Jucks, K. W., Toon, G. C., Stachnik, R. A., Bernath, P. F., Boone, C. D., Walker, K. A., Urban, J., Murtagh, D., Elkins, J. W., and Atlas, E.: Validation of the Aura Microwave Limb Sounder middle atmosphere water vapor and nitrous oxide measurements, J. Geophys. Res., 112, D24S36, doi:10.1029/2007JD008724, 2007.

Lefever, K., van der A, R., Baier, F., Christophe, Y., Errera, Q., Eskes, H., Flemming, J., Inness, A., Jones, L., Lambert, J.-C., Langerock, B., Schultz, M. G., Stein, O., Wagner, A., and Chabrillat, S.: Copernicus stratospheric ozone service, 2009–2012: validation, system intercomparison and roles of input data sets, Atmospheric Chemistry and Physics, 15, 2269–2293, doi:10.5194/acp-15-2269-2015, http://www.atmos-chem-phys.net/15/2269/2015/, 2015.

Li, H., Kalnay, E., and Miyoshi, T.: Simultaneous estimation of covariance inflation and observation errors within an ensemble Kalman filter, Quarterly Journal of the Royal Meteorological Society, 135, 523–533, doi:10.1002/qj.371, http://dx.doi.org/10.1002/qj.371, 2009.

Lin, S.-J. and Rood, R. B.: Multidimensional Flux-Form Semi-Lagrangian Transport Schemes, Mon. Weather Rev., 124, 2046–2070, 1996.

Liu, C., Xiao, Q., and Wang, B.: An Ensemble-Based Four-Dimensional Variational Data Assimilation Scheme. Part I: Technical Formulation and Preliminary Test, Monthly Weather Review, 136, 3363–3373, doi:10.1175/2008MWR2312.1, http://dx.doi.org/10.1175/2008MWR2312.1, 2008.

Livesey, N. J., Read, W. G., Wagner, P. A., Froidevaux, L., Lambert, A., Manney, G. L., Pumphrey, H. C., Santee, M. L., Schwartz, M. J., Wang, S., Fuller, R. A., Jarnot, R. F., Knosp, B. W., and Martinez, E.: Earth Observing System (EOS) Aura Microwave Limb Sounder (MLS) Version 4.2x Level 2 data quality and description document, Jpl d-33509 rev.a, Jet Propulsion Laboratory, http://mls.jpl.nasa.gov/data/v4-2_data_quality_document.pdf, 2015.

Lorenc, A. C.: The potential of the ensemble Kalman filter for NWP – a comparison with 4D-Var, Q. J. R. Meteorol. Soc., 129, 3183–3203, 2003.

Lorenc, A. C., Bowler, N. E., Clayton, A. M., Pring, S. R., and Fairbairn, D.: Comparison of Hybrid-4DEnVar and Hybrid-4DVar Data Assimilation Methods for Global NWP, Monthly Weather Review, 143, 212–229, doi:10.1175/MWR-D-14-00195.1, http://dx.doi.org/10.1175/MWR-D-14-00195.1, 2015.

Madronich, S. and Flocke, S.: The Role of Solar Radiation in Atmospheric Chemistry, in: Environmental Photochemistry, edited by Boule, P., vol. 2 / 2L of *The Handbook of Environmental Chemistry*, pp. 1–26, Springer Berlin Heidelberg, doi:10.1007/978-3-540-69044-3_1, http://dx.doi.org/10.1007/978-3-540-69044-3_1, 1999.

Ménard, R.: Error covariance estimation methods based on analysis residuals: theoretical foundation and convergence properties derived from simplified observation networks, Q. J. R. Meteorol. Soc., 142, 257–273, doi:10.1002/qj.2650, http://dx.doi.org/10.1002/qj.2650, 2016.

Ménard, R. and Chang, L.-P.: Assimilation of Stratospheric Chemical Tracer Observations Using a Kalman Filter. Part II: χ2-Validated Results and Analysis of Variance and Correlation Dynamics, Mon. Weather Rev., 128, 2672–2686, 2000.

Ménard, R. and Daley, R.: The application of Kalman smoother theory to the estimation of 4DVAR error statistics, Tellus A, 48, 221–237, 1996.

Mitchell, H. L. and Houtekamer, P. L.: An Adaptive Ensemble Kalman Filter, Monthly Weather Review, 128, 416–433, doi:10.1175/1520-0493(2000)128<0416:AAEKF>2.0.CO;2, http://journals.ametsoc.org/doi/abs/10.1175/1520-0493%282000%29128%3C0416%3AAAEKF%3E2.0.CO%3B2, 2000.

Miyazaki, K., Eskes, H. J., Sudo, K., Takigawa, M., van Weele, M., and Boersma, K. F.: Simultaneous assimilation of satellite NO_2, O_3, CO, and HNO_3 data for the analysis of tropospheric chemical composition and emissions, Atmos. Chem. Phys., 12, 9545–9579, 2012.

Poterjoy, J. and Zhang, F.: Systematic Comparison of Four-Dimensional Data Assimilation Methods With and Without the Tangent Linear Model Using Hybrid Background Error Covariance: E4DVar versus 4DEnVar, Monthly Weather Review, 143, 1601–1621, doi:10.1175/MWR-D-14-00224.1, http://dx.doi.org/10.1175/MWR-D-14-00224.1, 2015.

Sander, S., Friedl, R., Golden, D., Kurylo, M., Moortgat, G., Keller-Rudek, H., Wine, P., Ravishankara, A., Kolb, C., Molina, M., Finlayson-Pitts, B., Huie, R., and Orkin, V.: Chemical Kinetics and Photochemical Data for Use in Atmospheric Studies. Evaluation Number 15, JPL Publication 06-2, Jet Propulsion Laboratory, Pasadena, http://jpldataeval.jpl.nasa.gov, 2006.

Sankey, D. and Shepherd, T. G.: Correlations of long-lived chemical species in a middle atmosphere general circulation model, Journal of Geophysical Research: Atmospheres, 108, n/a–n/a, doi:10.1029/2002JD002799, http://dx.doi.org/10.1029/2002JD002799, 4494, 2003.

Santee, M. L., Lambert, A., Read, W. G., Livesey, N. J., Cofield, R. E., Cuddy, D. T., Daffer, W. H., Drouin, B. J., Froidevaux, L., Fuller, R. A., Jarnot, R. F., Knosp, B. W., Manney, G. L., Perun, V. S., Snyder, W. V., Stek, P. C., Thurstans, R. P., Wagner, P. A., Waters, J. W., Muscari, G., de Zafra, R. L., Dibb, J. E., Fahey, D. W., Popp, P. J., Marcy, T. P., Jucks, K. W., Toon, G. C., Stachnik, R. A., Bernath, P. F., Boone, C. D., Walker, K. A., Urban, J., and Murtagh, D.: Validation of the Aura Microwave Limb Sounder HNO3 measurements, Journal of Geophysical Research: Atmospheres, 112, n/a–n/a, doi:10.1029/2007JD008721, http://dx.doi.org/10.1029/2007JD008721, d24S40, 2007.

Skachko, S., Errera, Q., Ménard, R., Christophe, Y., and Chabrillat, S.: Comparison of the ensemble Kalman filter and 4D-Var assimilation methods using a stratospheric tracer transport model, Geoscientific Model Development, 7, 1451–1465, doi:10.5194/gmd-7-1451-2014, http://www.geosci-model-dev.net/7/1451/2014/, 2014.

Strong, K., Wolff, M. A., Kerzenmacher, T. E., Walker, K. A., and et al.: Validation of ACE-FTS N2O measurements, Atmos. Chem. Phys., 8, 4759–4786, 2008.

Talagrand, O. and Courtier, P.: Variational assimilation of meteorological observations with the adjoint vorticity equation. I: Theory, Q. J. R. Meteorol. Soc., 113, 1311–1328, 1987.

Tang, X., Zhu, J., Wang, Z. F., and Gbaguidi, A.: Improvement of ozone forecast over Beijing based on ensemble Kalman filter with simultaneous adjustment of initial conditions and emissions, Atmospheric Chemistry and Physics, 11, 12 901–12 916, doi:10.5194/acp-11-12901-2011, http://www.atmos-chem-phys.net/11/12901/2011/, 2011.

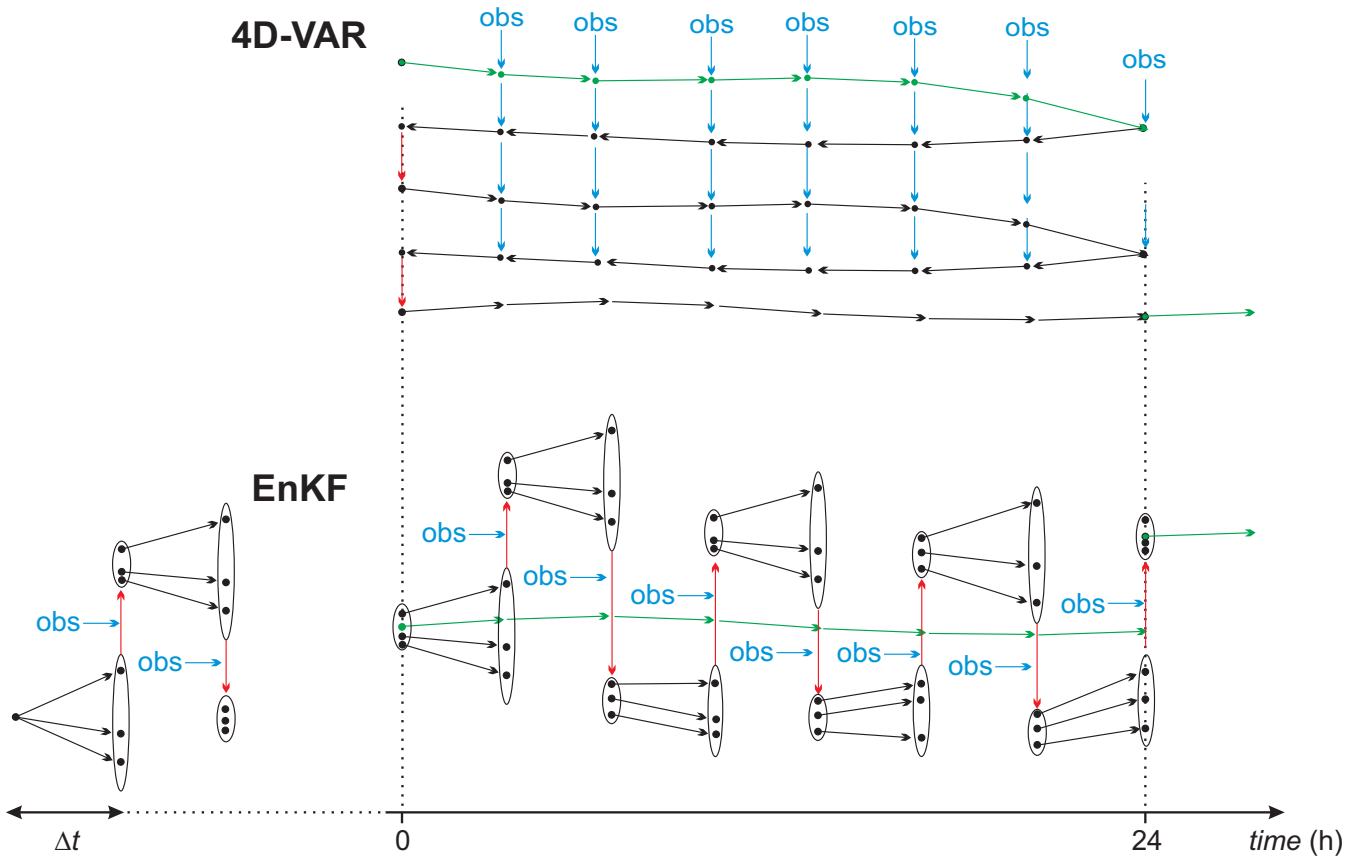

Figure 1. Schematic representation of the practical implementation of the 4D-VAR (top) and EnKF (bottom) assimilation methods in BASCOE. Black dots represent model state and observational information is depicted in blue. The black arrows represent model integrations by one time step, vertical red arrows represent model state optimization (4D-VAR) or Kalman filter (EnKF). Green dots represent the analyses at 0h which are used as initial conditions for the diagnostic 24-h forecasts (green arrows). For clarity, the number of 4D-VAR iterations has been limited to 2 and the number of EnKF members has been limited to 3.

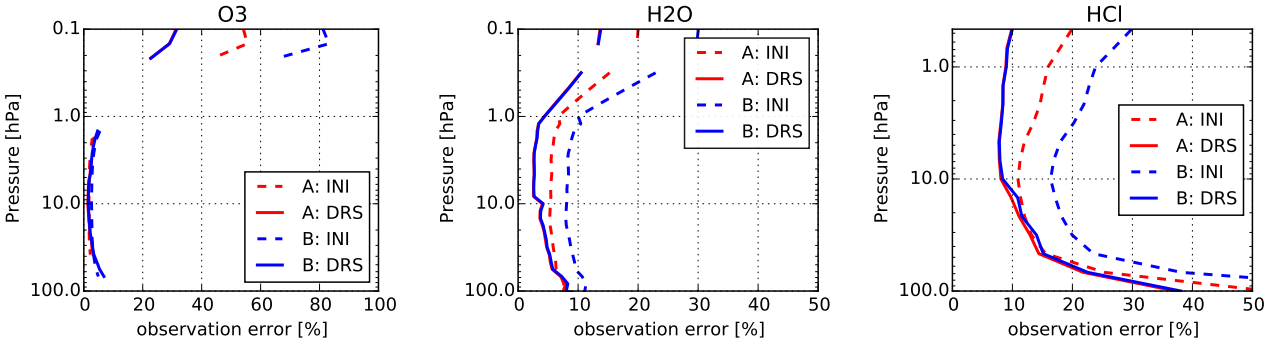

Figure 2. Initial (INI) observation error covariance matrix of experiment A (dashed red), starting with $\mathbf{R} = (\sigma_o)^2$, and B (dashed blue), starting with $\mathbf{R} = (1.5\ \sigma_o)^2$ and their Desroziers' (DRS) estimations (solid red and blue lines, respectively) using statistics of the first 24h. The statistics is shown for O_3, H_2O and HCl.

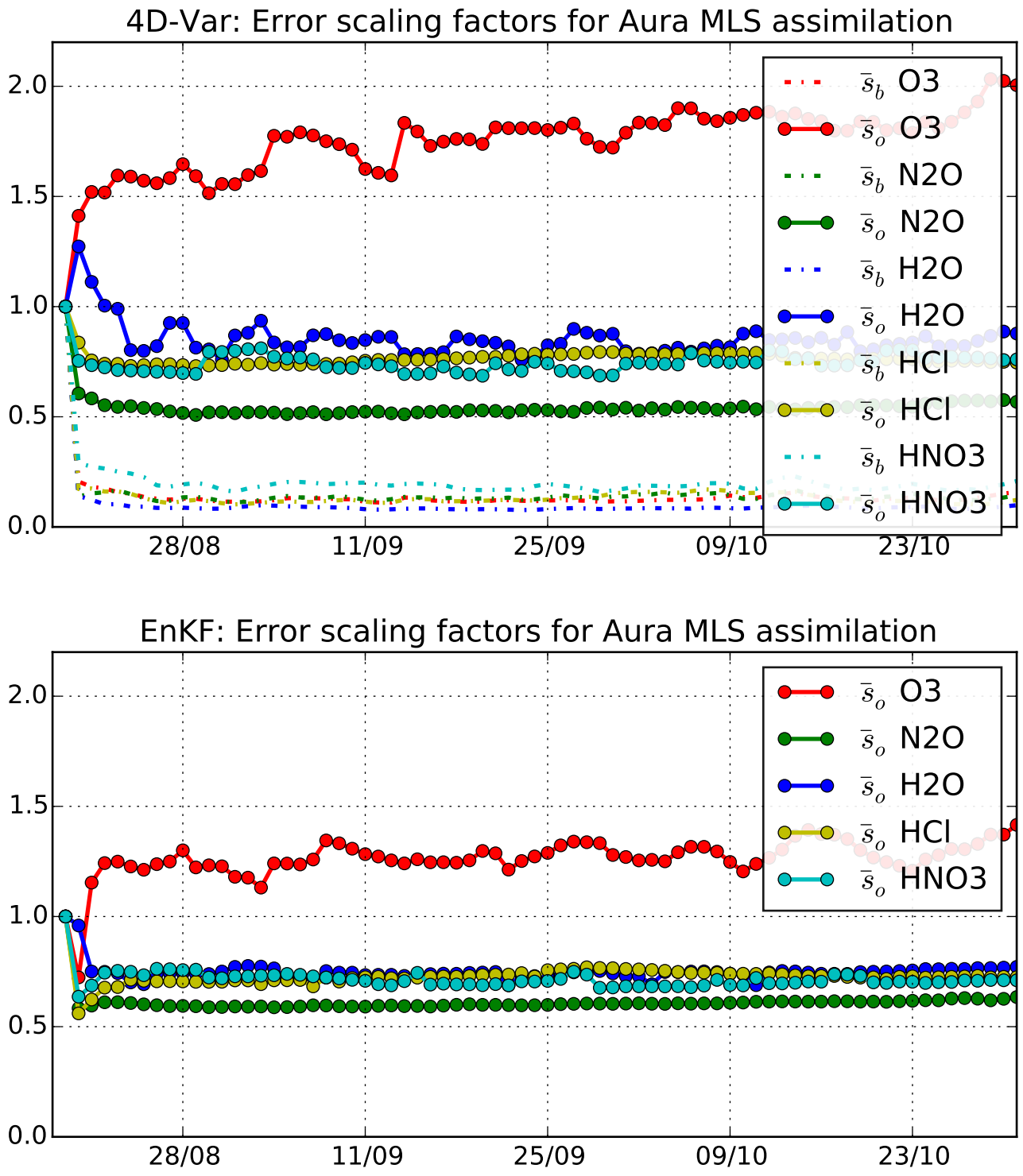

Figure 3. Estimated error scale factors within 4D-Var (left) and EnKF (right) for the period April-November 2008.

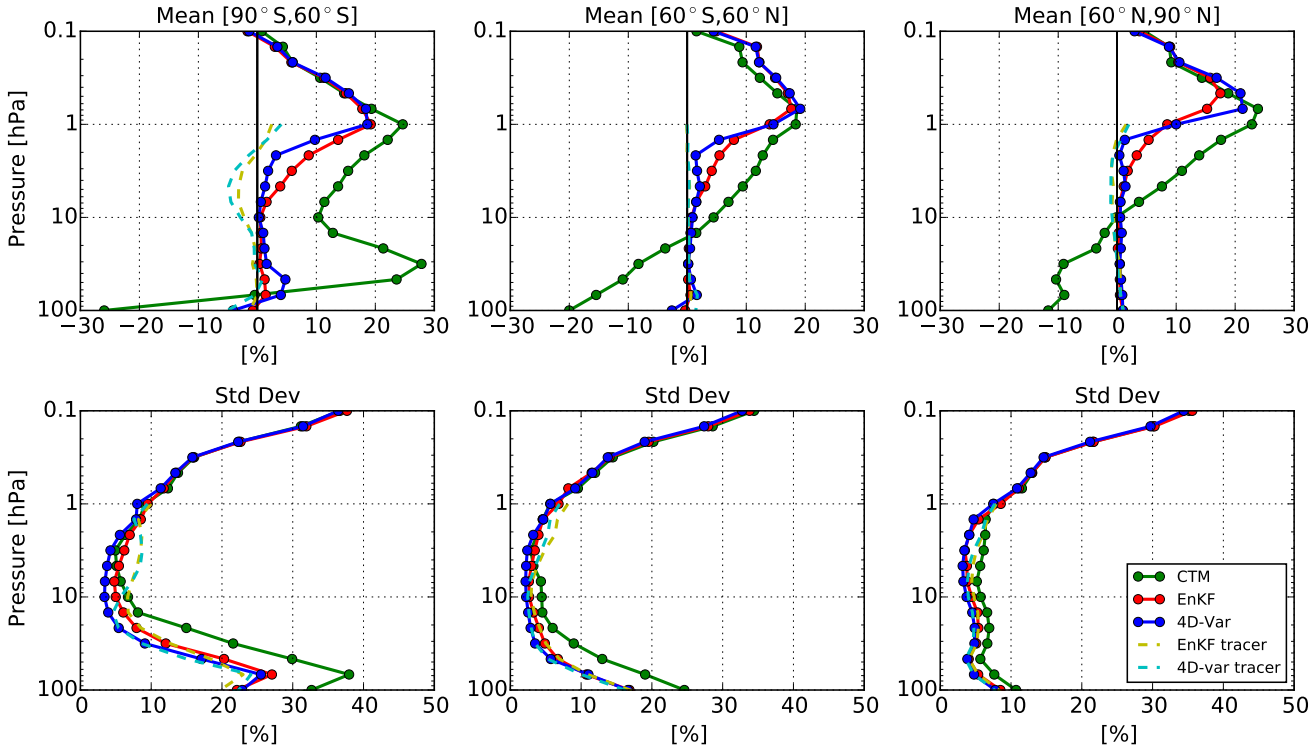

Figure 4. O_3 OmF bias (top) and standard deviation (bottom) computed for the full chemistry CTM (green), EnKF (red), 4D-Var (blue) based on the same model. The chemical tracer EnKF (dashed yellow) and 4D-Var (cyan) are also shown. OmF statistics is computed in percent with respect to the assimilated EOS Aura MLS data for the period September-October 2008 and for three latitude bands (from left to right: South Pole, Tropics - middle latitudes and North Pole).

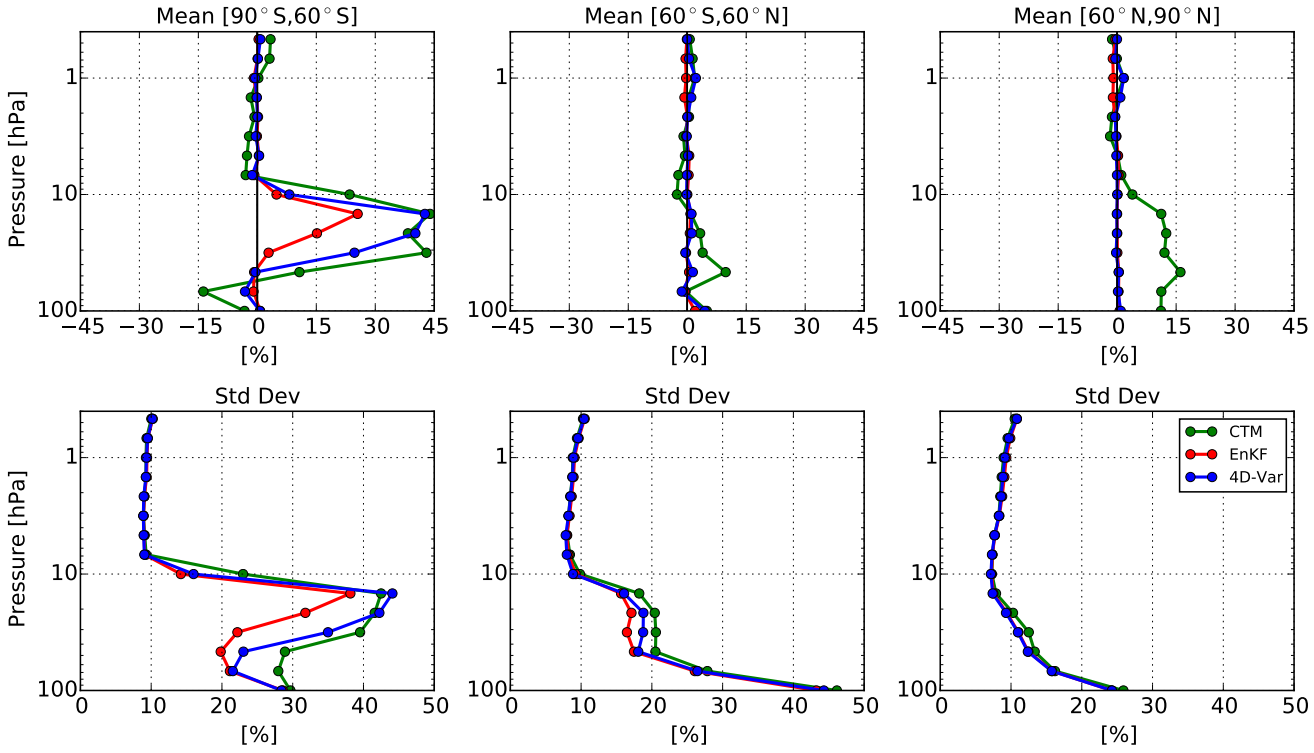

Figure 5. HCl OmF bias (top) and standard deviation (bottom) computed for the full chemistry CTM (green), EnKF (red) and 4D-Var (blue). OmF statistics is computed in percent with respect to the assimilated EOS Aura MLS data for the period May-June 2008 and for three latitude bands (from left to right: South Pole, Tropics - middle latitudes and North Pole).

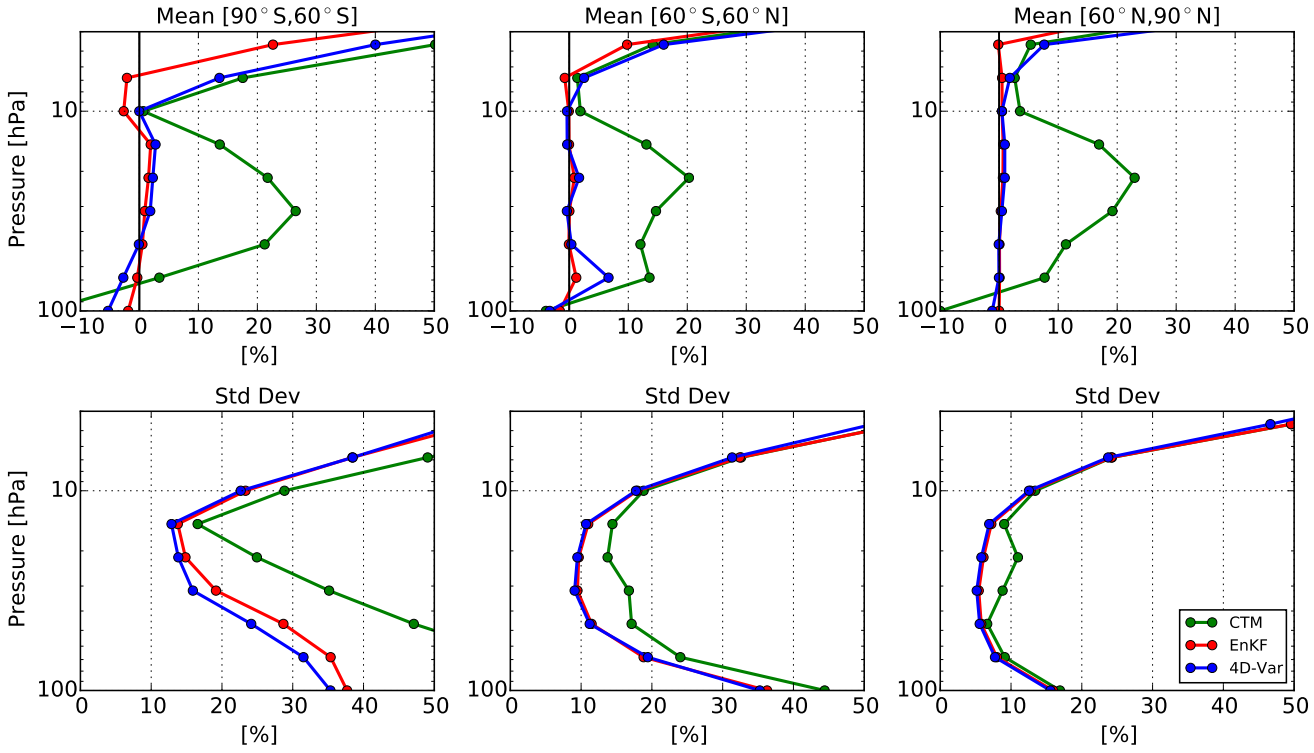

Figure 6. HNO_3 OmF bias (top) and standard deviation (bottom) computed for CTM (green), EnKF (red) and 4D-Var (blue). OmF statistics is computed in percent with respect to the assimilated EOS Aura MLS data for the period September-October 2008 and for three latitude bands (from left to right: South Pole, Tropics - middle latitudes and North Pole).

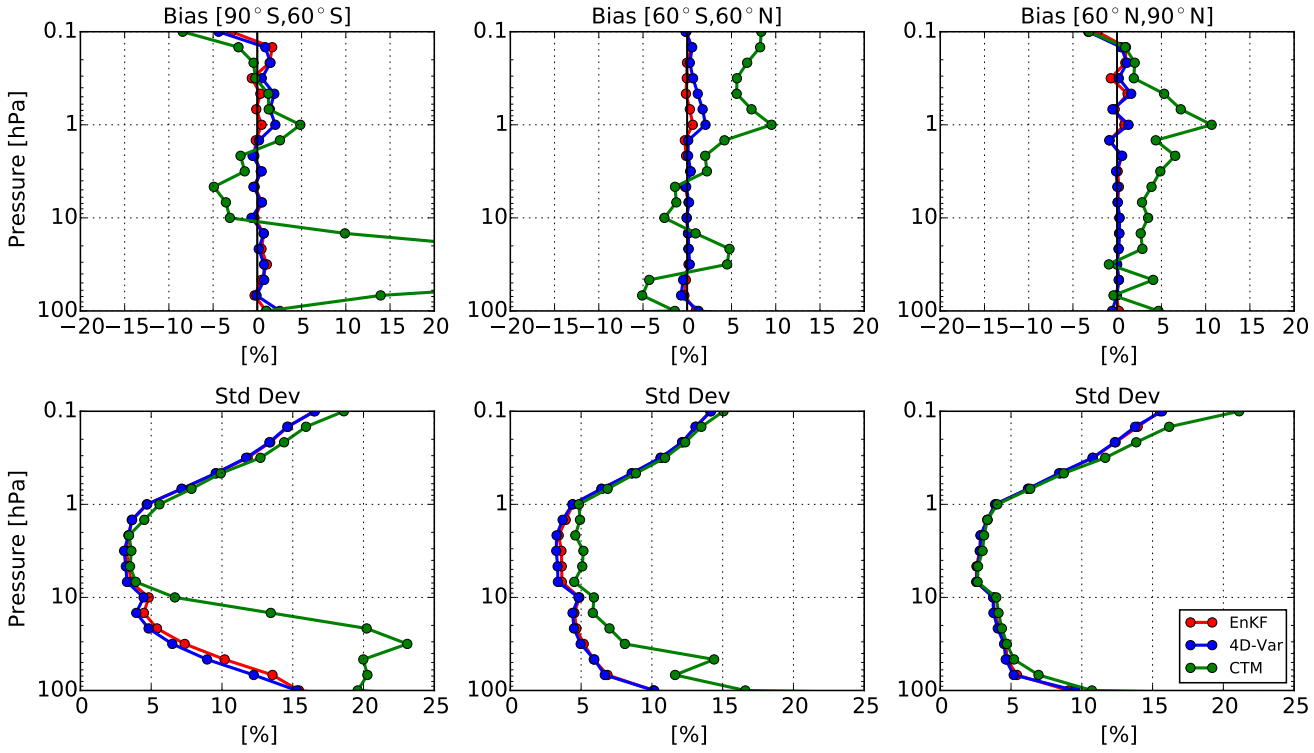

Figure 7. H_2O OmF bias (top) and standard deviation (bottom) computed for the full chemistry CTM (green), EnKF (red) and 4D-Var (blue). OmF statistics is computed in percent with respect to the assimilated EOS Aura MLS data for the period September-October 2008 and for three latitude bands (from left to right: South Pole, Tropics - middle latitudes and North Pole).

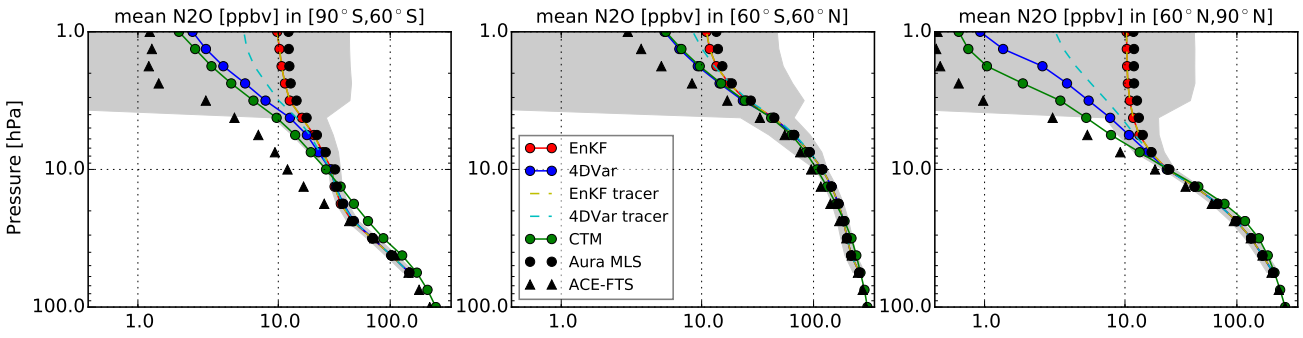

Figure 8. Mean N_2O from the full chemistry CTM (green), 24 h forecasts from EnKF (red) and 4D-Var (blue), based on the same model, and Aura MLS (black dots) and ACE-FTS data (triangles). 24 h forecast from chemical tracer EnKF (dashed yellow) and 4D-Var (cyan) assimilation are also shown. The grey area shows the precision of Aura MLS data. The statistics are computed for September-October 2008 and for three latitude bands (from left to right: South Pole, Tropics - middle latitudes and North Pole).

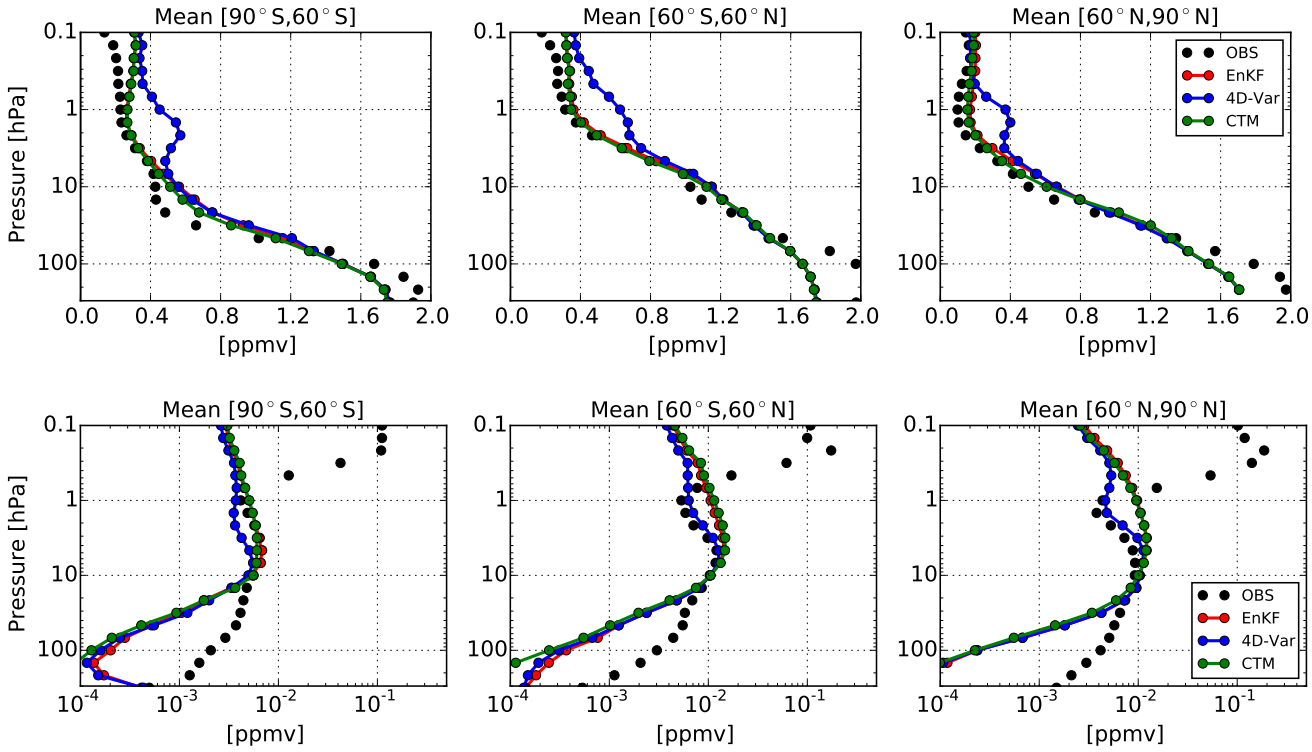

Figure 9. Verification of non-observed species from CTM (green), 24 h forecasts from EnKF (red) and 4D-Var (blue) assimilation against MIPAS IMK data (black dots): mean CH_4 (top) and mean NO_x (bottom) profiles. The statistics are computed for September-October 2008 and for three latitude bands (from left to right: South Pole, Tropics - middle latitudes and North Pole).

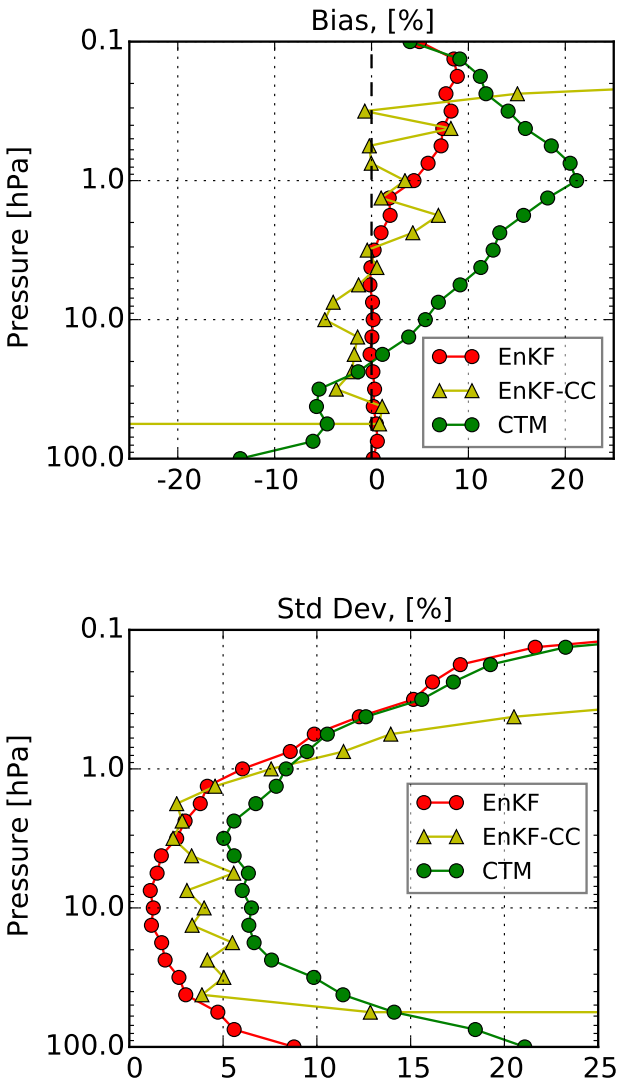

Figure 10. OmF bias (top) and standard deviation (bottom) between Aura MLS data and O_3 analyses of EnKF (red), EnKF-CC (yellow) and CTM (green), see text for acronym definition. The statistics are computed during 24h on September 15 2008.