Geosci. Model Dev. Discuss.,
doi:10.5194/gmd-2016-95-AC5, 2016

Geoscientific Model Development Discussions

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

S. Skachko et al.

sergey.skachko@aeronomie.be

Received and published: 6 June 2016

June 6, 2016

The author's responses are marked in blue. We would like to thank the anonymous referee 3 to the useful remarks.

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 $\mathbf{A}$ and $\mathbf{B}$ are, respectively L_A and L_B, the correlation length scale of $\mathbf{C}$ 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

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