**Atmospheric
Chemistry
and Physics
Discussions**

Interactive comment on "Ozone data assimilation with GEOS-Chem: a comparison between 3-D-Var, 4-D-Var, and suboptimal Kalman filter approaches" *by* K. Singh et al.

Anonymous Referee #2

Received and published: 3 October 2011

Singh et al. present tropospheric ozone analyses obtained using TES observations, the GEOS-Chem model and three assimilation schemes based on, respectively, 3D-Var, KF and 4D-Var. The evaluation of the analyses is carried out by means of comparisons against independent ozonesonde observations. At this stage, I found the description of the experiments not clear enough. As a result, the review of the paper was difficult. Whatever the accuracy of my understanding of the paper, I suggest that you clarify it or improve the experimental setup. This will certainly need an additional iteration in the review. However such a study is very valuable and I encourage the authors to not give up.

General Comments

1. The setup of the different systems and the use of the analyses do not seem to be done properly. As I found the text not very clear, I will first present my understanding of your work and my point of view on it.

   - TES data are assimilated from the 1st of August initialized by a GEOS-Chem free model run. Two comparisons are performed considering TES data for 5 days and 2 weeks. Four assimilation runs are performed: one 3D-Var, one KF and two 4D-Var. The first two are conducted with an assimilation frequency of 4 hours while the 4D-Var experiments are performed with an assimilation window of 5 days and 2 weeks, respectively. On P22266/L17-18, it is mentioned, for sequential methods, that "as we move forward in time, the analysis field agrees better with the true state of the atmosphere." This is because the initial conditions used on 1st of August are too far away from the data. Before comparing the different schemes, you should run several cycles of assimilation such that the differences between the forecast and the observations reach a certain constant value. The analyses during this period, which is called the spinup period, should not be used in the comparison. Typically, this period is between one and two weeks, depending on the data coverage. Note that this period also exists for 4D-Var and you should also perform several 4D-Var cycles to ensure that the spinup period is not taken into account in the comparison. This could explain the difference between the two 4D-Var analyses using a 5-day and 2-week assimilation window, respectively.

   - The comparison of the analyses should be done at their optimal time. For 3D-Var and KF, this time corresponds to the analysis time. 4D-Var analyses, at the beginning of the assimilation window, behave more likely as the observations (noisier analyses). On the other hand, 4D-Var analyses become

more likely a pure model at the end of the assimilation window (smoother analyses). Thus, the optimal time of 4D-Var analyses is at the middle of the assimilation window, not the end of the assimilation window as shown in Figs. 3, 4, 6 and 7.

- It is also not clear why two assimilation windows are used in the case of 4D-Var and why the two assimilation windows are so long. Such long windows might be justified if the aim of using 4D-Var is to constrain unobserved region or unobserved species. As only one analyses could be used within a cycle (see my previous comment), your setup provides only one analyses every 5 days or every 2 weeks. If the final goal of this study is to use GEOS-Chem to provide tropospheric O3 analyses, this frequency seems too long with respect to the transport time scale in the troposphere. Maybe using a one-day assimilation window would be a better choice (insofar as this is a natural choice that corresponds to a diurnal cycle).

- It is stated in the paper that comparisons between the analyses follow the method presented in Geer et al. However, Geer et al. make the comparison for a much longer time period, using a bigger set of independent data with an evaluation of the analyses in different latitude and altitude regions. I found the domain of the comparison used in this paper to small to make any definitive conclusion (ozonesondes about Northern America). The authors might consider a longer period of assimilation (outside the spinup) and make comparison with ozonesondes in different regions. Using ozonesonde data from another network, such as NDACC, might increase the number of independent ozone profiles.

2. The description of the model used in this study is not very clear, probably because the information is split in different places in the paper. I would recommend that you merge the information in Sec. 3 and the first § of Sec. 6.1 in Sec. 5. Moreover, Sec. 5.1 does not aim at describing the different schemes. It is

C9700

rather devoted to a discussion on the setup of the assimilation window. Perhaps this section could be renamed accordingly. Most important: I still do not know whether the chemical scheme of GEOS-Chem used in this study is based on linoz (see P22261/L9), SMVGEAR or KPP (see Sec 6.1). This is also important to understand Sec. 6.3.1 (which I failed; see my specific comments).

3. I found the discussion in Sec. 6.3.1 not very clear. I would suggest that the authors follow the method elaborated by Fisher and Lary (1995) to derive the potential influence of one observed species on the other non-observed species included in the model.

4. At this stage, I have not yet evaluated the Sec 6.3.2

5. I found numerous typos and misleading sentences. I strongly recommend that the authors take some extra time to reread carefully their paper before the next submission.

Specific Comments

P22250/L15-18: The first papers that discussed chemical variational assimilation are: Fisher and Lary (1995); Elbern et al. (1997); Khattatov et al. (1999); Errera and Fonteyn (2001). They could be mentioned.

P22250/P19: replace "Khattatov et al. (2000)" by "Khattatov et al. (1999)" as the first one only discussed suboptimal KF.

P22250/§4: Lahoz and Errera (2010) provide a review of chemical data assimilation and worth being cited.

P22251/L7: typo: "implementations", "performances" with "s".

P22251/L8: Replace "," by "." in "interest, A".

P22251/L16: Replace "tropospheric" by "stratospheric" as Geer et al. is mainly concerned by the stratosphere.

P22252/L16-18: replace "by minimizing" by "that minimize". Add also at the end of the sentence "... and taking into account the error covariances of the observations and the forecast."

P22253/L15: replace "state space" by "model state vector".

P22255/L15: I guess $\mathbf{H}$ must be replaced by $\mathcal{H}$.

P22257/L3-4 and next §: The sentence says that $\mathbf{P}^f$ is discussed in sec 5.2 while it is discussed in the next §. Please, clarify.

P22257/L21: Remove "gas phase" because I suppose your model also includes photodissociation and heterogeneous reactions.

P22258, title of sec 4: This section discusses mostly the observation operator. This should be mentioned in the title of the section.

P22258/L18-19: I could deduce the connection between the thermal contrast and the high latitudes but I would like the sentence be more explicit.

P22259/L6: what is "The corresponding TES observation operator 3"? Is "3" referring to a version number? Please, clarify.

P22259/L21: "The points ...": do you mean the model grid points or the observation location? To avoid any misleading, use observation "locations" instead of "points" in the paper.

P22260/L16: Section 3 mentions GEOS-Chem v7.2.3 while Section 5 mentions v7-04-10. This is confusing and must be clarified.

P22261/L9: In this study, does GEOS-Chem use a full chemical scheme or does it use

the linoz ozone parameterization?

P22261/L26-28: I don't understand this sentence, especially with the fact that the $\mathbf{B}$ matrix is supposed to be diagonal.

P22262/L9-11: On which base are the two assimilation windows for 4D-Var chosen?

P22262/L15-17: Why is the cost function in 4D-Var updated every 4 hours and not at every model time step? Is the model time step 4 hours, too?

P22262/L17-20: This sentence is poorly written and partly wrong. It is stated in the first part of the sentence that: "Contrary to ... at the initial time and". It is not exactly true to say that 4D-Var analyses are produced only at the initial time. It would be more appropriate to write that 4D-Var produces new initial conditions and that these initial conditions can be used by the model to get the analyses at any time within the assimilation window. In the second part of the sentence, it is written: "and accounts ... in the assimilation window". In fact, this is the case for the three methods. What makes the difference between sequential and variational methods is more likely the fact that sequential methods suppose that the atmospheric state is static during the assimilation window, while in variational methods, a model is used to make the time intepolation of the model initial condition at the observation time.

P22262/title of sec 5.2: By "Specification of error variances", I understand the error variance of the background and observation. This section only discusses the background errors. Accordingly, the title must be clarified.

P22262-22263/§1 of sec 5.2: Do you mean that the background variances are set identical for each model level and set as 50% of the mean ozone field at that level taken from the initial guess? If so, this § can be shortened.

P22264/L7: I suppose that the vertical correlations are also neglected. If so, replace "horizontal correlations" by "spatial correlations".

P22265/L4: replace by "..., we set the KPP parameters RTOL=...".

P22265/L11: Do you mean "We first consider a comparison between the analyses for the case where the 4-D-Var assimilation is based on an assimilation window of five days."?

P22265/L15-16. Can you specify if the 4D-Var system requires a huge CPU time or memory (or both)?

P22265/L16-19: Model and adjoint runs require an substantial CPU time. Is the storage of the model history an issue? If not, the sentence could be terminated by "forward and adjoint model runs".

P22265/L28: add a "," between ".html" and "Thompson".

P22266/L7-10: A verb is missing in the sentence "Forecast scoring ...".

P22267/L16: replace "(Nassar et al., 2008)" by "Nassar et al. (2008)".

P22267/L22: as far as I understand the sentence, I would replace "ozone distribution" by "ozone variability".

P22267/L26/27: "... at the end of the assimilation window". See my general comments.

P22268/L26: Do you mean "We next consider comparing the analyses produced by the 4-D-Var assimilation method with an assimilation window of two weeks."?

P22269/L9: "4-D-Var still provides" and not "provided"

P22269/L11: "it performs well" instead of "performed"

P22269/L11: remove "as well".

P22270/L3-6: the word "case" is used too often in this sentence.

P22270/L7-10: The sentence is too long, which makes it unclear. Please rewrite.

P22271/L17-20: Why "in principle"? If your 4D-Var system based on KPP, the backward integration of the model (using the adjoint of the KPP chemical solver) will provide

C9704

information on the non-observed species if they are chemically coupled to the observed species. Is it not the case in your 4D-Var formulation?

P22272/L2: replace "Sandu et al. (2005a)" by "(Sandu et al., 2005a)"

P22272/L5-10: what you call the "adjoint sensitivities" seems to be the gradient of the cost function for a given species. Is that right? If so, use this terminology instead. In the case where KPP is the chemical solver, it is surprising that the system does not also reduce the gradient for these unobserved species. Moreover, the best way of measuring the influence of observed species on non-observed species is to use the influence function introduced by Fisher and Lary (1995).

P22272/L26: "comparison" is repeated twice.

References

Elbern, H., Schmidt, H., and Ebel, A.: Variational data assimilation for tropospheric chemistry modeling, J. Geophys. Res., 21, 15,967–15,985, 1997.

Errera, Q. and Fonteyn, D.: Four-dimensional variational chemical assimilation of CRISTA stratospheric measurements, J. Geophys. Res., 106, 12,253–12,265, 2001.

Fisher, M. and Lary, D. J.: Lagrangian four-dimensional variational data assimilation of chemical species, Q. J. R. Meteorol. Soc., 131, 1681–1704, 1995.

Khattatov, B. V., Gille, J. C., Lyjak, L. V., Brasseur, G. P., Dvortsov, V. L., Roche, A. E., and Water, J. W.: Assimilation of photochemically active species and a case analysis of UARS data, J. Geophys. Res., 98, 18,715–18,737, 1999.

Lahoz, W. and Errera, Q.: Constituent Assimilation, in: Data Assimilation: Making sense of observations, edited by Lahoz, W., Kahattatov, B., and Ménard, R., pp. 449 –490, Springer, 2010.