Atmos. Chem. Phys. Discuss., 7, S4290–S4294, 2007
www.atmos-chem-phys-discuss.net/7/S4290/2007/
© Author(s) 2007. This work is licensed
under a Creative Commons License.

Atmospheric Chemistry and Physics Discussions

Interactive comment on "Data assimilation of stratospheric constituents: a review" *by* W. A. Lahoz et al.

Anonymous Referee #1

Received and published: 22 August 2007

General comments

This is a very good paper, which presents a comprehensive review of assimilation of observations of chemical constituents of the stratosphere, in terms of methods as well as in terms of results that have been obtained so far. It will be very useful to many readers. I strongly recommend acceptance of the paper.

At the same time, I consider a number of aspects of the paper require improvement. This is particularly true of the methodological aspects. The authors have made a commendable effort of clarification on those aspects, but improvements are still necessary. Specific comments follow.

Specific comments

1. The authors, putting assimilation in the general perspective of Bayesian estimation, write that, chiefly because of the numerical size of the problem, simplifying assumptions are necessary in practice. They write that two approaches have been followed so far, namely statistical linear estimation and ensemble assimilation (p. 9566, l. 2).

That is perfectly correct, but then the authors do not say clearly what is what among the methods they describe. They speak casually of variational and sequential methods (p. 9570, l. 7), and of the equivalence between Kalman Filter and 4D variational methods (p. 9573, l. 13), without relating those aspects to what they have said before. They later give 'justifications' for assuming that data error are Gaussian (p. 9569, ll. 11-15), without saying that it is in conditions of linearity and gaussianity that statistical linear estimation achieves Bayesian estimation. And at a later stage (p. 9593, l. 17), they refer to the Best Linear Unbiased Estimator (BLUE), without mentioning that the BLUE is precisely what statistical linear estimation is meant to produce.

All that is likely to cause some confusion in the minds of uninformed readers. A better structured approach is necessary. I suggest the following.

a. As soon as you mention statistical linear estimation, mention that it achieves Bayesian estimation when everything around is linear and Gaussian. Mention if you wish the expression Best Linear Unbiased Estimator. But if you do, say clearly that the BLUE is what statistical linear estimation is meant to produce (and concerning 'justifications' for assuming data errors to be Gaussian, the most fundamental justification, which is entirely pragmatic, is the relative simplicity and easiness of implementation of statistical linear estimation).

b. Mention that, independently of the notion of statistical estimation, two broad classes of numerical algorithms, sequential and variational, exist for assimilation. In the context of statistical linear estimation, those take respectively the forms of Kalman Filter and 4D-Var. These are two different algorithms for determining the BLUE, and are equivalent under the only condition of linearity.

ACPD

7, S4290–S4294, 2007

2. P. 9571, ll. 28-29. ‘ Ĕ the validity of the tangent linear hypothesis, required to define the tangent linear model’. The tangent linear model is defined under the only condition that the function J defined by eq (1) be differentiable. The tangent linear hypothesis is required for minimization of J to achieve statistical linear estimation.

3. P. 9569, ll. 10-11. ‘In general, in data assimilation, errors [Ĕ] are assumed to be Gaussian and unbiased’. I have already commented on gaussianity. As for unbiasedness, I do no think the authors’ statement is true. Many assimilation schemes now contain a bias correction (which, from the point of view of general estmation theory, is the proper way to deal with biased data).

4. P. 9588, ll. 15-16. It is written here that the cost of primal 4D-Var is very high compared to PSAS. That is certainly not true of 4D-PSAS, the cost of which, like the cost of primal 4D-Var, is determined by the cost of the repeated integrations of the assimilating model and its adjoint (see, e. g., Courtier, 1997, or Louvel, 2001). You must be speaking of 3D-PSAS ? (3D-PSAS is I think the only form of PSAS to have been used so far in constituent assimilation, but clarification is desirable).

5. P. 9589, paragraph starting l. 20. How can PSAS be ‘formulated so that the analysis uncertainties are determined directly’ ?

6. P. 9602. Second paragraph. It seems to be said there that, as of May 2007, ECMWF did not assimilate ozone observations operationally, in contradiction with what is written p. 9603, ll. 11-12.

7. P. 9565, ll. 19-21. ‘The model operator [Ĕ] maps the analysis forward in time to give a background state for a subsequent assimilation cycle.’ That is true only of sequential assimilation. In variational assimilation, the model operator is part of the process which produces the analysis.

8. P. 9571, second paragraph. The second and third sentences of that paragraph seem contradictory, in that the approach described in the second (uncorrelating variables) is

ACPD
7, S4290–S4294, 2007

EGU

said in the third to open up the possibility of coupling different variables.

9. P. 9573, l. 19. EnKF does not 'approximate' EKF. On the contrary, it is more general than EKF to the extent it does not require validity of the tangent linear hypothesis.

10. P. 9574, ll. 5-6. ' .. assume a diagonal B with values larger than the observation errors'. Is that really what you mean ? There is no link between B and the magnitude of observation errors. But, if B is taken diagonal (which is certainly not valid), it is reasonable to increase the diagonal terms in order to compensate for the neglect of the off-diagonal terms.

11. P. 9588, ll. 4-6. 'This [Ĕ] reduces the weight of the [Ĕ] background in the final analysis'. The general equivalence, in the linear case, between 4D-Var and KF implies that the same weight is given to all data in both methods. What is written here could be therefore be true only because of possible non-linearities. This must be clarified.

Technical corrections

The paper has on the whole been carefully written and edited. I have however noticed a few corrections to be made.

12. The instruments POAM-III (p. 9591, ll. 23-24) and SEVIRI (p. 9602, ll. 12-13) are not mentioned in Table 1. Similarly, I do not think that measurements performed by IASI (p. 9584, l. 28) have been assimilated so far, but an appropraite reference should be given about that instrument.

13. Table 3. 'Nudging' is mentioned among the assilation methods that have been used, but is not mentioned in the text. Say briefly that is it an empirical 'forcing' of the model fields towards the observed values, and can be described as an extremely simplified form of Kalman Filtering.

14. The paper by Zupanski (1997) (referred to p. 9572, ll. 2-3) is not mentioned in the list of references.

ACPD

7, S4290–S4294, 2007

EGU

15. I presume the reference to Lyster et al. (2000) (p. 9590, l. 7) should be changed to Lyster et al. (1997).

References

Courtier, P., 1997, Dual formulation of four-dimensional variational assimilation, Q. J. R. Meteorol. Soc., 123, 2449-2461.

Louvel, S., 2001, Implementation of a dual variational algorithm for assimilation of synthetic altimeter data in the oceanic primitive equation model MICOM, J. Geophys. Res., 106 (C5), 9199-9212.

Interactive comment on Atmos. Chem. Phys. Discuss., 7, 9561, 2007.