

## ***Interactive comment on* “The direct inversion method for data assimilation using isentropic tracer advection” by M. N. Juckes**

**Anonymous Referee #2**

Received and published: 30 September 2005

Review of “The direct inversion method for data assimilation using isentropic tracer advection”, MS-NR acpd-2005-0192, by Juckes.

General comments:

This is a potentially important piece of work, and it is important that ACPD are providing a forum for the data assimilation community to discuss this paper. However, I think that before this work is accepted for ACP, the author should address a number of points. These points are of a general and specific nature. General points concern what the author wishes to convey in the paper and the level of detail; specific points concern a number of items which require clarification or further explanation.

Full Screen / Esc

Print Version

Interactive Discussion

Discussion Paper

I flag the paper "accept subject to major revision" because, in my view, the changes required concern a change of focus/emphasis on how the paper is currently written. However, I think these changes should not take too long.

My general concerns about the paper are:

I do not think the method and its advantages come through clearly in the paper. In my view, this is largely because of the mathematical detail in the paper. Reducing and/or condensing this detail should help in this. For example, are all the appendices needed?

In discussing the method, the author should also provide clear evidence of what is novel about it, and how it compares with other more established data assimilation approaches. The specific comments indicate parts of the paper where this could be done. I think it is important to establish this in the paper.

It is not clear to me what the aims of the paper are. Is the paper about a novel method, or is it about evaluation of MIPAS data? I think the author wishes to inform the community of this novel method but this does not come through as clearly as it should.

Specific comments (page and line numbers refer to the ACPD paper):

P. 8880:

Abstract: There is too much detail in the abstract. Do we need to know in the abstract the number of sondes or the number of profiles?

Introduction: The first paragraph paper states "This paper assesses the quality of the MIPAS ozone observations". This suggests that the paper is about evaluation of MIPAS data, rather than about a (potentially) novel method for doing data assimilation. If the paper is about the method, this should be stated up front in the introduction.

P. 8881:

L 6: I think it is "El Serafy and Kelder".

Full Screen / Esc

Print Version

Interactive Discussion

Discussion Paper

L. 7: Which “situation”?

L. 8: “observation” -> “observations”

L. 11: What do you mean by “additional data”. Additional to what?

L. 14: I think it would be appropriate to explain what is meant by “background field”.

L. 15: I think it would be appropriate to provide a reference for the Kalman smoother.

L. 18: I think it would be appropriate to explain what is meant by “control variables”.

P. 8882:

L. 3: Perhaps say: “The method used in this paper is”

L. 6-7: What is the importance (if any) of the method preserving the quasi-elliptic nature of the problem.

L. 10: The 0.5 degree resolution is, presumably, latitude-longitude?

L. 13-14: How come “there is clearly more information in the field that could be obtained from simple interpolation of the observations.”?

L. 16: What “information”? Please specify.

L. 17: What “timescales”? Please specify.

L. 21: I think it would be helpful to indicate that by independent observations it is meant observations not used in the assimilation.

P. 8883:

Eq (2): Perhaps indicate that “E” is the “expectation” operator. I am assuming this is what it is. If it is not this, the author should indicate what it is.

L. 15: I would have thought “M” acts as an operator. Perhaps this should be mentioned.

P. 8884:

Full Screen / Esc

Print Version

Interactive Discussion

Discussion Paper

L. 2: Does “random” mean Gaussian pdfs? It is common in data assimilation to assume Gaussian pdfs.

Eq. (5): I think that, roughly speaking, the 3 terms on the RHS of the eqn. are, analogous to, JO (observations penalty term), JQ (model penalty term) and JB (background penalty term), that would arise in weak constraint 4d-var. Is this indeed the case? If so, it would clarify the meaning of Eq. (5) to state this (the comments on Eq. (6) hint that this is indeed the case, but I think more explanation is needed).

L. 19: "random, the value"

P. 8886:

L. 1: It would be help to state what the analysis in relation to Eqs (5)-(9) is; is it “chi”?

L. 15-19: Is this obvious? Perhaps further explanation is needed.

P. 8887:

L. 20: Would it be worthwhile speculating what the likely effect of wind variability is? Should we worry that this has not been tested?

P. 8888:

L. 2: I think the smoothing term plays a similar role to JB in variational methods (see above). I think it would be helpful to mention this.

L. 16: The “solution” of what?

L. 19-24: 4th order in what? in  $r$ ? I find the discussion unclear. Is it obvious that a 6th order term guarantees sufficient smoothness? I think the author needs to provide more explanation.

P. 8889:

L. 4: "non-dimensionalised by rate"

Full Screen / Esc

Print Version

Interactive Discussion

Discussion Paper

L. 11-15: Why these choices of the parameters?

P. 8890:

L. 3: Has the concept of the “error covariance matrix” been introduced in the paper?

L. 5-6: Explain the statement: “the error in real-time assimilation will be larger than in a retrospective one.” Is it to do with how the information is used?

L. 12: What is the difference between Figs. 1 and 2a? Is Fig. 1 needed?

L. 17: “low ozone values”

P. 8891:

L. 2: Why do you expect a loss of accuracy near the end of the analysis windows?

L. 3: “looking at the difference”

Eq. (15) (and elsewhere): explain what  $dA$  is. Presumably it is an element of area on an isentropic surface.

L. 12: What “information”? Is this incorrect information? The text suggests that this information is detrimental.

L. 14: At the ends of which “overlap period”?

L. 17: Is the difference in accuracy between the Kalman filter and the Kalman smoother related to the backward integration in time in the KS?

L. 17-18: It is not clear to me how you get an e-folding time of 2-3 days from Fig. 2.

L. 20: I understand that the value of 4-6 days is related to the e-folding time of 2-3 days, but the way they are related is not clear to me. Have I missed something?

L. 22: I think more explanation is required on why the information content of the KS is ~2 times that of the KF. This whole paragraph (L. 14-26) is difficult to follow.

Full Screen / Esc

Print Version

Interactive Discussion

Discussion Paper

P. 8892:

L. 6: Where does the figure of  $\sim 0.23$  ppmv come from?

L. 14: This is the first mention of the Laplace equation. The text suggests that Eq. (7d) (which does not appear in the text) is related to the Laplace equation. If it is the case that Eq. (7) is related to the Laplace equation, this should be mentioned when Eq. (7) is introduced.

L. 22-24: Presumably this is because the ratio cycles/control variable is  $\sim$  constant?

P. 8893:

Eq. (18): It would help if it was indicated in the text that B and V are, respectively, the bias and the variance.

This page: I see no discussion of Fig. 4 before a discussion of Fig. 5. Is this an oversight?

L. 19: Perhaps state that Migliorini et al. found a negligible bias between MIPAS ozone data and ozonesondes.

P. 8894:

L. 12: The phrase “For SBUV and OSIRIS the observations, measurements minus MIPAS analysis” is a bit unclear. Does the author mean observations minus analysis (O-A)?

L. 18: “shows the means”

P. 8895:

L. 11: “Still enough” for what?

L. 13: Explain the term “nearest neighbour” plots?

L. 15: How does the assimilation account for the spatial and temporal variation of the

Full Screen / Esc

Print Version

Interactive Discussion

Discussion Paper

ozone field? By filling in the gaps in an intelligent manner?

L. 20: Where is Fig. 4a?

L. 23: Where do you get the value  $\sim 0.35$  ppmv? From a particular Fig.? If so, indicate which one.

L. 25: “to be independent”

P. 8896:

L. 3: The difference in resolutions in Figs. 1 and 2 should be discussed earlier when they are introduced. This should make it clear to the reader what are the differences between Figs. 1 and 2. Above, I question whether Fig. 1 is needed - perhaps if their differences and purpose was brought more to the fore when they are first introduced in section 3, having the two Figs. could be justified.

L. 15-17: Is there an implication for computer cost?

L. 22: The background term is also used in 3d-var. I think this should be discussed when Eq. (5) is introduced (see above comments).

L. 26-27: It seems to me that this difference between 4d-var and the model described in the paper is important. I think it warrants more explanation. Is this difference an advantage of the method described?

P. 8897:

L. 8: Why “this cannot happen here”?

L. 11-13: I think it would be helpful if weak constraint is defined (i.e. model errors are allowed). What do you mean by “demonstrated to work”? This is vague. I think more specific comments are needed.

L. 17-19: The text suggests a cost advantage for the method described. If this is indeed the case, this is an advantage and should be brought more to the fore.

Full Screen / Esc

Print Version

Interactive Discussion

Discussion Paper

L. 20-21: Describe the evidence in the paper supporting this statement.

P. 8898:

L. 3-6: I think this is a good example of the way data assimilation “adds value”. This point should be emphasized.

L. 7-9: How are these values estimated? Here 2.5 days is mentioned, but in the main body of the paper 2-3 days are mentioned.

Before appendix A: It would be useful to have an indication of planned future work.

P. 8899:

L. 1-4: Do this data refer to Table 4? I see no mention of Table 4 in the text.

L. 7-13: Inconsistent spelling of discretised/discretized and variants.

P. 8900:

Eq. (B1): Why are these choices made?

P. 8991:

L. 6: What is the physical meaning of the relation  $\text{lap} \ll \text{ld}$ ?

P. 8902:

L. 7: Write in full: European Centre for Medium-range Weather Forecasts

P. 8905:

My overall view of appendices is that they have too much information. Granted that most readers would not look into them in detail, nevertheless, I think the author should seriously consider reducing the content in the appendices. Besides being difficult to follow, they do distract the reader from what I think should be (is?) the purpose of this paper, viz., the discussion of a potentially novel approach to data assimilation, and what its advantages are.

Full Screen / Esc

Print Version

Interactive Discussion

Discussion Paper



P. 8908:

L. 7: “Serafy” -> “El Serafy”.

P. 8909:

Table 2 caption: I think this should be rephrased to immediately identify that the mean and variance are in units of %, with an indication of what the % is relative to.

P. 8918:

Fig. 4: This is too busy and difficult to interpret. Can the Fig. be simplified? At what isentropic level are the data plotted?

P. 8919:

Fig. 5: As Fig. 4., this is too busy. The % are relative to what variable? What is SAGE III(L)?

P. 8920-8921:

Figs. 6-7: Are both these Figs. needed?

P. 8922:

Fig. 8: It would be helpful if the colour of the lower line could be identified in the caption.

---

Interactive comment on Atmos. Chem. Phys. Discuss., 5, 8879, 2005.

Full Screen / Esc

Print Version

Interactive Discussion

Discussion Paper