Atmos. Chem. Phys. Discuss.,
doi:10.5194/acp-2016-914-RC1, 2016

Atmospheric Chemistry and Physics Discussions

Interactive comment on "How much information do extinction and backscattering measurements contain about the chemical composition of atmospheric aerosol?" *by* Michael Kahnert and Emma Andersson

Anonymous Referee #4

Received and published: 11 November 2016

Summary

The authors consider the case of assimilation of remote-sensing data (specifically aerosol extinction and backscattering coefficients) applied to aerosols fields within a chemical transport model. They describe how an additional term can be added to the 3D-var cost-function so that the assimilation adjusts only those components (in a transformed space) for which the observations provide information. The additional term relies on the singular value decomposition of the scaled observation operator. In this way, the assimilation automates the choice of control variables in an otherwise highly

under-constrained inverse problem.

Verdict

The paper is very well written and is surprisingly clear, given the subject matter. The manuscript introduces a potentially very powerful concept for variable selection into the field of aerosol data assimilation. The authors have probed the idea in a minimal test case, which assists in understanding the effects. I found that the shortcomings of the paper were relatively minor. I felt there was insufficient discussion of the literature of related treatment. I was unsure about whether the organisation of the material was optimal (see the "Main comments"). Finally, a counter-experiment without the addition of the new constraint in the 4D-var cost function was, in my opinion, lacking. All in all, I believe that the paper should be published, pending the minor revisions suggested below.

Main comments

- There was little or no discussion of literature on related treatments. I have not the time to read all of these myself, however I have included a list at the end of articles that may be relevant, for example those that deal with information content of observations in data assimilation or those that refer to the singular value decomposition of the observability matrix.
- I believe that a small counter-experiment was lacking. In the results presented in section 3.2, I would suggest also presenting results for the assimilation experiment which did *not* include the additional constraint in the 3D-var cost-function.

- I was unsure whether the organisation of the material was optimal - I highlight this as an issue that the editor may wish to take up. The introduction concludes by urging the reader to read the Appendix before proceeding onto the rest of the methods and results section. Much of the interesting methodology is contained within the Appendix, and we agree that it would be difficult to make sense of the main part of the paper without a good understanding of the contents of the Appendix. As such, I would suggest incorporating the Appendix into the main body of the text. At one level, this is really a matter of taste, and thus I leave it to the editor.

Minor comments

- When describing observation errors, there was no reference to the component from "representativity errors" (i.e. measurements are made at a point, or over a small area in the case of remote sensing, while model grid-boxes are typically in the order of kilometres across in the horizontal dimensions). All of the discussion about observation errors was in terms of the measurement error and errors in the observation operator, both of which are relavent. However the representativity component is not insignificant in many contexts.
- Observation standard deviation was reported in percentage, but it was unclear what this was a percentage of. Please clarify.
- I would suggest replacing all instances of the term "costfunction" with "cost function" (or "cost-function"). The latter is about 15 times more common (on the web, at least). Similarly, I believe that the compound word "nullmatrix" is used in German (capitalised, that is) whereas it is "null matrix" (or "zero matrix") in English.

- I could not find a definition to the term "signal degrees of freedom". Please include this somewhere (preferably at first usage, or in the Appendix).
- Line 48: Please replace "This is a rather bold approach that largely disregards ..." with "This is approach largely disregards ..." – please use argument rather than rhetoric to explain what is wrong with the work of others.
- Line 54: The reference to Kahnert (2009) is used to show that several optical properties at multiple wavelengths may allow constraining more than just the total mass concentration. Surely other authors have looked into this. If so, please summarise other work done. If not, please say so.
- Line 98: "... using 40 eta-layers with variable thickness depending on the underlying topography" – do you just mean that this is a terrain-following coordinate? Or is there something more sophisticated about this?
- Line 125: "The background error covariance matrix of the model a priori is modelled with the NMC method ..." . I checked the reference (Kahnert, 2008), in which I believe this is described. If the implementation is the same here as in the 2008 article, then I believe that it is best to say that it "follows similar principles to the NMC method" or "is inspired by the NMC method". If it is indeed the NMC method, the authors should clarify the difference to methodology laid out in Kahnert (2008).
- Line 129: I would suggest replacing "Given m observations of, e.g., m_1 different parameters at m_2 different wavelengths, so that $m_1 \cdot m_2 = m$, how many..." with "Given m observations (e.g., m_1 different parameters at m_2 different wavelengths, so that $m_1 \cdot m_2 = m$), how many..."
- Line 130: "... we can constrain to better than observation error" – do you mean "model error"? If not, please explain that the transformation makes the (rescaled) observation errors and (rescaled) model variables comparable.

- Line 134: Please replace "... a singular value decomposition of the Jacobian of the observation operator ..." with "... a singular value decomposition of the Jacobian of the scaled observation operator ..." or something similar. By the way, this scaled observation operator appears to have a name: "the observability matrix"
- Footnote 2, page 5: I found this distinction a bit cryptic. Please consider rephrasing.
- Line 150: I realise that this is something that is clarified later on, but I would suggest saying a few words at this point about the synthetic observations; namely, what kind of observations they were and how many observation points there were.
- Line 154: I would suggest the following change "thus providing nearly perfect observations. (We assumed an observation error standard deviation of 10 %) The only ..." becomes "thus providing nearly perfect observations (we assumed an observation error standard deviation of 10 %). The only...". See also my comment about about describing the units for the observation error standard deviation.
- Line 162: What is "Nd:YAG"? Please clarify. I suspect that this is some error with the bibliography manager.
- Line 168: I would suggest the following change: "... two wavelengths. (Compare, e.g., cases 1., 2., and 3. to cases 4., 5., and 6.) Hence ..." becomes "... two wavelengths (compare, e.g., cases 1., 2., and 3. to cases 4., 5., and 6.). Hence ..."
- Line 171: a missing full stop after the right parenthesis.
- Table 1, caption: the "Nd:YAG" term appears again.

- Line 181: I believe that "weak constrains" should be "weak constraints"
- Line 189: See my comment above about the representativity component to the observation error
- Figure 1: I think it would be interesting to see the increment as an additional panel in this figure
- Figure 1: The text on the scale is a bit too small. I would suggest having one scale, rather than three, and enlarging the scale so that the labels can be read.
- Figure 2: The units appear to be "mixing ratio [ppb-m]". Do you mean mass mixing ratio? Please clarify.
- Figure 4: Do we need all panels? Why not just show the first three or four, and then a selection of the remaining terms.
- Line 262: I would suggest the following change " ... dramatic decrease in both the entropy and signal degrees of freedom ..." becomes " ... dramatic decrease in both the entropy-change and signal degrees of freedom ..."
- Line 282: "It also appeared that among the original model variables, secondary inorganic aerosol components were most faithfully retrieved by the inverse modelling solution" – why is this? why SIA? Do they have specific optical properties to make them more observable by such LIDAR pseudo-observations?
- Line 293: I would suggest the following change: "The present study should be extended..." becomes "The present study could be extended..."
- Line 295: I believe that the expression "highly underrated" is somewhat dramatic and relatively colloquial, and does not fit with the tone in the rest of the paper. The authors are encouraged to use argument rather than rhetoric to make their point.

- Line 297: Regarding the statement "There is little one can put forward in defence of this model other than pure convenience". Some justification is required (e.g. some references) to demonstrate why this model is untenable. There's a saying (attributed to George Box) "All models are wrong, some models are useful". Does this model give significantly worse results than representations, or is it just inaccurate in its assumptions?
- Paragraph beginning at line 307: It may be worth making it clear that y is not observed, but a model equivalent of the observations
- Lines 324 and 326: I would suggest replacing all instances of "3-dimensional" with "three-dimensional"
- Paragraph beginning 336: I would suggest mentioning that the assumption of unbiased background and observation errors
- Footnote 6, page 15: See my comments above about the representativity component of the observation error.
- Footnote 7, page 16: I would suggest the following change: "The observation errors are often uncorrelated" becomes "The observation errors are often assumed to be uncorrelated (this is not always true)"
- Paragraph beginning at line 368: Please comment on the role of spatial and interspecies correlations, particularly in light of the comment "if we allow all model variables to be freely adjusted" (line 374).
- Line 369: It might be worth noting that δx is not constrained to ensure that all components of x remain positive in the analysis.
- Line 386: The phrase "rather tricky" strikes me as somewhat colloquial. I would suggest the following change: "However, to actually make such a comparison is

rather tricky" becomes "However, to actually make such a comparison poses two problems."

- Paragraph beginning at line 390: Please introduce the meaning of the angle-bracket notation. I believe that this is common in physics, but other disciplines (e.g. statistics) often use different notation for the expectation.
- Footnote 8, page 17: Should $A \cdot A = B$ be $A^T \cdot A = B$?
- Line 425: Should "(C7)-(C9)" not be "(C6)-(C9)"? As far as I can see, Eq. (C6) is required here.
- Line 434-435: Please state which particular sections/chapters of Rodger (2000) the reader is referred to.
- Equations C12, C15: I would suggest showing the range of the summation to indicate that it is a summation over observations (i.e. i ranges from 1 to m)
- Line 479: "Naively, one may have expected that the dimension would, on the contrary, be reduced to $n - k$" – why? is this because the number of unknowns remains the same but the number of equations to be solved has increased by k?
- Line 486: I would suggest the following change: "(Note that the covariance matrices and their inverses are symmetric, i.e., $R^T = R$, etc.)" becomes "Note that the covariance matrices and their inverses are symmetric (i.e. $R^T = R$, etc.)."
- Appendix: For all unit and zero matrices (and vectors), I would suggest indicating the dimension as a sub-script
- Line 498: I would suggest adding a subscript to clarify with respect to what the differentiation refers (i.e. replace ∇ with ∇_ξ)
- Paragraph beginning line 515: how was this tuning done in practice?

- Line 549: “It turns out that Eq. (D18) gives a relatively sharp transition from unconstrained to constrained model variables, while Eq. (D19) gives a very gentle transition” – this can be seen from the equations. I would suggest replacing the sentence with “It can be seen that Eq. (D18) gives a relatively sharp transition from unconstrained to constrained model variables, while Eq. (D19) gives a very gentle transition”

- Paragraph beginning line 567: I found that this went too fast and skipped a bit too much detail, after what was otherwise a very well-written paper that included a fair bit of theory. In particular, can you please explain in further detail the reduced matrices. The phrase “we are primarily interested in constraining the chemical components” was surprising, since I thought the authors were mainly interested in the aerosol components. What does it mean to “restrict ourselves to the chemical subspace”?

- Line 570: Full stop missing after N_c

- Paragraph beginning 574: similar to the above comment, I found that this skipped over too much detail. Please add further explanation. The authors state that in their present study, they use a Cholesky decomposition of the B-matrix. Is this what was used in Kahnert (2008), or is this described as the “spectral formulation”? If it is different, it may be relevant to understand why the Cholesky decomposition was preferable to the author’s previously presented methodology. This is mainly to understand the requirements and limitations of the proposed methodology.

Minor formatting issues

- References with parentheses inside parentheses: lines 33, 268, 269, 376

- Some of the in-line equations appeared to be missing spaces on one or both sides of the equals sign – this only appeared in the appendix. See lines: 391, 403, 404, 425, 515. I might just be imagining it. The paper was otherwise very well laid out.

References the authors may wish to consider

- Qin, X. Measuring information content from observations for data assimilation: relative entropy versus shannon entropy difference. Tellus: Series A. 59, 2, 198-209, 2007.

- J. Joiner, A. M. da Silva. Efficient methods to assimilate remotely sensed data based on information content. Q. J. R. Meteorol, SOC. (1998), 124, pp. 1669-1694

- C Cardinali, S Pezzulli, E Andersson. Influence-matrix diagnostic of a data assimilation system. Q. J. R. Meteorol. Soc. (2004), 130, pp. 2767-2786. doi: 10.1256/qj.03.205

- C. Johnson, N. K. Nichols; B. J. Hoskins. Very large inverse problems in atmosphere and ocean modelling. Int. J. Numer. Meth. Fluids 2005; 47:759-771.

- M Bocquet, 2009: Toward Optimal Choices of Control Space Representation for Geophysical Data Assimilation. Mon. Wea. Rev., 137, 2331-2348, doi: 10.1175/2009MWR2789.1.

- F Rabier, N Fourrie, D Chafai, P Prunet. Channel selection methods for Infrared Atmospheric Sounding Interferometer radiances. Q. J. R. Meteorol. Soc. (2002), 128, pp. 1011-1027

• C Johnson, B. J. Hoskins, N. K. Nichols. A singular vector perspective of 4D-Var: Filtering and interpolation. Q. J. R. Meteorol. Soc. (2005), 131, pp. 1-19 doi: 10.1256/qj.03.231

Interactive comment on Atmos. Chem. Phys. Discuss., doi:10.5194/acp-2016-914, 2016.