Atmos. Chem. Phys. Discuss.,
doi:10.5194/acp-2016-914-AC4, 2017

Atmospheric Chemistry and Physics
Discussions
Open Access

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

Michael Kahnert and Emma Andersson

michael.kahnert@smhi.se

Received and published: 2 February 2017

Below the reviewer comments are marked in blue, our response is marked in black.

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.

We are grateful for this encouraging assessment of our work, as well as for the insightful comments and suggestions. It is obvious that the reviewer has devoted considerable time into studying the manuscript and providing constructive criticism on various aspects of the content and organisation of the paper. Our detailed response to these comments follows.

1 Main comments

1. 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.

   We have, indeed, only cited studies on aerosol data assimilation. Most of the studies cited by the reviewer are concerned with numerical weather prediction (NWP). We will add a paragraph to discuss related NWP studies and include the citations suggested by the reviewer.

2. 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.
This is a very valid point that was also brought up by other reviewers. We will include these results and revise the figures and discussion accordingly. In particular, we intend to replace Fig. 2 by two new figures. The first figure will, similarly to the old Fig. 2, show the total mass concentration of different aerosol species, but now for both the constrained and the unconstrained analysis. The second figure will show a similar comparison of a selection of aerosol species in specific size bins. We anticipate that this comparison will illustrate that the unconstrained analysis yields more erratically varying vertical profiles (i.e., results that vary more wildly in the null-space).

3. 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.
This is a tricky point. We put some thought into this before writing the paper, and we concluded that the appendix is, indeed, most interesting for readers who are mainly interested in data assimilation methodology, and for those who are very eager to learn something about it. But other readers, e.g. lidar instrument

developers, will most likely be deterred from reading the paper if we merge the entire appendix with the main body of the paper. However, the reviewer's criticism is very valid, and it has been brought up by several reviewers. We believe that reviewer 2 has suggested a very good compromise, namely, to state and explain the main results (equations) from the appendix in the methodology section of the paper, while retaining the derivations and more detailed explanations in the appendix. This alleviates the problem that parts of the main text are hard to understand without the information given in the appendix. At the same time, we avoid the risk of making the paper inaccessible (or just too boring) for those who do not mainly work with data assimilation methodology.

We therefore propose to follow the suggestions of reviewer 2 in this point. It seems to us that this will also adequately address the main point of criticism brought up by reviewer 4.

2 Minor comments

1. 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 relevant. However the representativity component is not insignificant in many contexts.
Agreed. We will add a discussion of the representativity error in the text accompanying table 2, where we will make it clear that in this numerical test we have neglected this source of error.

2. Observation standard deviation was reported in percentage, but it was unclear

what this was a percentage of. Please clarify.
It is a percentage of the observed backscattering coefficient or extinction coefficient. We will change the text in Sect. 2.5 from “We assumed an observation error standard deviation of 10 %” to “We assumed that the observation error standard deviation is 10 % of the measurement value.”

3. 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.
Agreed.

4. 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).
We will add an explanation of the terminology to Sect. 2.4. Following the suggestions of reviewer 2, we will provide key equations of the appendix with explanations in the main text. This will also apply to Eq. C12, which will be provided in the methodology section. Thus the explanation and definition of the term “signal degrees of freedom” will appear much earlier in the revised paper. (Note that in the remote sensing community the number of signal degrees of freedom is also known as the “effective rank” of the problem.)

5. 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.
Agreed.

6. 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.

We do cite the study by Burton et al. (2016) (L64), although we do so in the introduction. We will add two more references that analyse the information content of lidar observations, namely, the papers by Veselovskii et al. (2004) and Veselovskii et al. (2005). However, these papers analyse the information content with respect to particle size and refractive index, not with respect to chemical composition. Therefore, the citation of these papers fits better into the introduction, together with the citation of Burton et al. (2016).

7. 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?
OK, we will replace this with “using 40 terrain-following coordinates”.

8. 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).
OK, we will replace this with “follows similar principles to the NMC method”.

9. Line 129: I would suggest replacing “Given m observations of, e.g., m1 different parameters at m2 different wavelengths, so that m1 m2=m, how many...” with “Given m observations (e.g., m1 different parameters at m2 different wavelengths, so that m1 m2=m), how many...”
Agreed.

10. 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.

This was a bit confusing. We will reformulate this sentence, and we will add a more detailed explanation of the terminology *signal degrees of freedom*.

11. 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"
Yes. Actually, this text will be extended with a lot more explanations, and it will provide the main equations from the appendix. We will follow the reviewer's suggestion and introduce the term *observability matrix* for the scaled Jacobian.

12. Footnote 2, page 5: I found this distinction a bit cryptic. Please consider rephrasing.
There seem to be two fractions in the community. One that uses *data analysis* and *data assimilation* almost interchangeably, and another that insist on keeping these two concepts apart. We are mostly guilty of belonging to the first one, but we do not want to make a big deal out of mere questions of terminology (which is why we put this into a footnote rather than into the main text). However, we will do our best to clarify the text in the revised manuscript.

13. 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.
Agreed; we will add this information in the revised manuscript.

14. 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.

Agreed (replacing "observations. (We" by "observations (we". In addition, in response to an earlier request to be more specific what me mean by "10 %" (percent of what?), we will replace the text in parenthesis with "(we assumed that the observation error standard deviation is 10 % of the measurement value)".

15. Line 162: What is "Nd:YAG"? Please clarify. I suspect that this is some error with the bibliography manager.
It is no error. "Nd:YAG" is the standard abbreviation for "neodymium-doped yttrium aluminium garnet" laser, one of the most commonly used solid-state lasers in remote sensing. We will add this information.

16. 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 ..."
Agreed. However, reviewer 2 has suggested to replace the cases considered in Table 1 with different cases that are more closely associated to combinations of wavelengths and parameters that are technologically feasible and common. Thus the text accompanying Table 1 is likely to change considerably.

17. Line 171: a missing full stop after the right parenthesis.
Agreed.

18. Table 1, caption: the "Nd:YAG" term appears again.
See our earlier response.

19. Line 181: I believe that "weak constrains" should be "weak constraints".
Yes.

20. Line 189: See my comment above about the representativity component to the observation error.
Agreed, see our earlier response.

21. Figure 1: I think it would be interesting to see the increment as an additional panel in this figure.
Figure 1 has been criticised by several reviewers. In fact, this figure is not particularly useful in the context of our paper. We are not discussing any aspects of regional modelling or horizontal information spreading in the assimilation algorithm. The model merely serves us to provide us with a test case. So, we will remove this figure in the revised manuscript (see also our response to reviewer 2).

22. 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.
See the previous item.

23. Figure 2: The units appear to be “mixing ratio [ppb-m]”. Do you mean mass mixing ratio? Please clarify.
Yes. This will be corrected in the revisions.

24. Figure 4: Do we need all panels? Why not just show the first three or four, and then a selection of the remaining terms.
Agreed, we will show 10 instead of 20 panels. Following the comment by reviewer 2, we will run a $3\alpha + 2\beta$ test case, in which case we will have 5 signal degrees of freedom. Thus we will show the first 5 signal-related transformed increments, and 5 out of the 15 noise-related increments.

25. 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 ...”
Agreed.

26. Line 282: “It also appeared that among the original model variables, secondary inorganic aerosol components were most faithfully retrieved by the inverse mod-

elling solution” – why is this? why SIA? Do they have specific optical properties to make them more observable by such LIDAR pseudo-observations?
This question has been brought up by several reviewers. We follow the suggestion of reviewer 2 and add an analysis of the linear coefficients that transform the elements in model space to the signal-related control variables. We will add a figure and a discussion — see our detailed response to reviewer 2.

27. Line 293: I would suggest the following change: “The present study should be extended...” becomes “The present study could be extended...”
Agreed.

28. 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.
OK, we will replace the text with “Another important issue concerns the choice of ...”.

29. 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?
Worst of all, this model is rather unpredictable, since its accuracy depends on the size, refractive index, and shape of the aerosols. Also, it may, in some cases, give reasonable results at one wavelength and for one specific parameter, and fail at other wavelength or for other optical parameters.

    There is a large body of work concerned with aerosol optics and the shortcomings of simplified model particles. Some of these studies focus on specific types of aerosols, others on specific morphological properties, such as non-sphericity, inhomogeneity, surface roughness, or chemical heterogeneity. It is difficult to pick just a few of such studies as representative citations. So, perhaps the best we can do is to cite a recent review paper on aerosol optics modelling that discusses the strengths and shortcomings of various morphological models (Kahnert et al., 2014).

30. Paragraph beginning at line 307: It may be worth making it clear that y is not observed, but a model equivalent of the observations
We will insert the following sentence: "The operator $\hat{H}$ maps from model space into observation space, which allows us to compare model output and observations."

31. Lines 324 and 326: I would suggest replacing all instances of "3-dimensional" with "three-dimensional"
Agreed.

32. Paragraph beginning 336: I would suggest mentioning that the assumption of unbiased background and observation errors
Agreed.

33. Footnote 6, page 15: See my comments above about the representativity component of the observation error.
OK; see our earlier response.

34. 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)"
Agreed.

35. Paragraph beginning at line 368: Please comment on the role of spatial and inter-species correlations, particularly in light of the comment "if we allow all model

variables to be freely adjusted" (line 374).
OK. We will add the following footnote (after "(within the given error bounds).": By solving the equation $\nabla J|_{\vec{x}=\vec{x}_a} = \vec{0}$ for the analysed state $\vec{x}_a$ it can be shown that the solution to the inverse problem is given by $\vec{x}_a = \vec{x}_b + \mathbf{K} \cdot (\vec{y} - \hat{H}(\vec{x}_b))$, where $\mathbf{K} = \mathbf{B} \cdot \mathbf{H}^T \cdot (\mathbf{H} \cdot \mathbf{B} \cdot \mathbf{H}^T + \mathbf{R})^{-1}$ is known as the gain matrix. This illustrates that the analysis updates the background estimate $\vec{x}_b$ by mapping the increment $(\vec{y} - \hat{H}(\vec{x}_b))$ from observation space to model space by use of the gain matrix. The correlations among the model variables enter into the gain matrix through the matrix $\mathbf{B}$ In our case the vertical correlations are rather weak in comparison to correlations among different aerosol species.

36. Line 369: It might be worth noting that δx is not constrained to ensure that all components of x remain positive in the analysis.
There is no such constraint in the minimisation process itself, but we do post-process the results for δx such that negative concentrations would be set to zero. In practice, this rarely ever happens.

37. Line 386: The phrase "rather tricky" strikes me as somewhat colloquial. I would suggest the following change: "However, to actually make such a comparison israther tricky" becomes "However, to actually make such a comparison poses two problems."
Agreed.

38. 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.
OK, we will add a formal definition of the expectation value for discrete variables in a footnote.

39. Footnote 8, page 17: Should $A \cdot A = B$ be $A^T \cdot A = B$?
Yes!

40. Line 425: Should "(C7)-(C9)" not be "(C6)-(C9)"? As far as I can see, Eq. (C6) is required here.
Yes.

41. Line 434-435: Please state which particular sections/chapters of Rodger (2000) the reader is referred to.
Agreed.

42. 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)
This is not generally true. The summation goes from 1 to $\min\{m, n\}$, where n is the dimension of model space, and m is the dimension of observation space. We will add these summation limits to the sums.

43. 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?

In physics one usually learns about holomorphic constraints in theoretical mechanics, often by considering a point mass moving on a hypersurface. So, this is often the mental picture one invokes when dealing with constrained problems. For instance, a point mass in three-dimensional Euclidean space with a single holomorphic (i.e. strong) constraint can be pictured as moving on a two-dimensional surface. Thus this constraint reduces the dimension of the manifold on which the the point mass can move from three to two. One would therefore *naively* expect that one is now dealing with a two dimensional problem. The reason why this is naive is because a nonlinear constraint will correspond to a *curved* manifold. To characterise this manifold requires additional equations. Only if we have *linear* constraints, then the hypersurface is simply a tilted plane, which, by a suitable rotation-translation, can be brought into coincidence with, e.g., the xy plane. In

such cases, and only in such cases, can the dimension of the problem actually be reduced, as one would naively have expected.

44. 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.)."
Agreed.

45. Appendix: For all unit and zero matrices (and vectors), I would suggest indicating the dimension as a sub-script.
Agreed. We will change this throughout the manuscript.

46. Line 498: I would suggest adding a subscript to clarify with respect to what the differentiation refers (i.e. replace ∇ with ∇_ξ).
Agreed.

47. Paragraph beginning line 515: how was this tuning done in practice?
As it is explained in the text. When the error variance is too large, one can see that the analysis is close to the unconstrained one. When it is too small, the analysis lies very close to the background estimate. One varies the variance until one obtains an analysis that departs from the background without drifting over to the (often noisy) unconstrained analysis.

48. 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"
We will replace the text with "Equation (D18) can be expected to give a sharper transition from unconstrained to constrained model variables than Eq. (D19)."

Our tests, so far, showed that the differences between these approaches are not quite as dramatic as we expected.

We will replace the text after Eq. (D21) with " The test we performed, so far, showed that these different approaches often yield analysis results that are quite similar. However, in each approach the free parameters σ_G and c may assume different values. If they are not well tuned, then the analysis tends either toward the background estimate or the toward the unconstrained analysis."

49. 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"?
This seems to be a misunderstanding. What we mean by "chemical components" is "chemical components in the aerosol phase". Since our paper is exclusively concerned with aerosols, we thought that there was no risk of misunderstanding. Thus, by "chemical subspace" we mean "subspace of aerosol components". We will revise the text accordingly and replace all instances of "chemical components" by "aerosol components", and similarly for "chemical subspace". Also, we will revise the text in response to point 51 (see below).

50. Line 570: Full stop missing after N_c.
OK.

51. 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 formula-

tion"? 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.
We *are* using the spectral formulation for the minimisation of the cost function. However, we formulate the weak constraints in a subspace of physical space, as explained above. The Cholesky decomposition is only applied to the reduced B-matrix in the formulation of the weak constraints. We do not go into the details of spectral data assimilation, since these questions are rather specific to our particular implementation, while the paper is not restricted to spectral methods. However, we will rewrite this entire subsection and explain the reduced subspace approach in much more detail. We will also add a short footnote on how to incorporate this into the spectral formulation.

3 Minor formatting issues

1. References with parentheses inside parentheses: lines 33, 268, 269, 376
This will be corrected.

2. 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.
Our latex program seems to insert spaces when using the eqnarray environment, but not when using the equation environment. We trust that the copy editor will take care of this problem.

4 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

Agreed; these will be added to and discussed in the introduction.

References

Burton, S. P., Chemyakin, E., Liu, X., Knobelspiesse, K., Stamnes, S., Sawamura, P., Moore, R. H., Hostetler, C. A., and Ferrare, R. A.: Information content and sensitivity of the $3\beta+2\alpha$ lidar measurement system for aerosol microphysical retrievals, Atmos. Meas. Techniques, 9, 5555–5574, 2016.

Kahnert, M., Nousiainen, T., and Lindqvist, H.: Review: Model particles in atmospheric optics, J. Quant. Spectrosc. Radiat. Transfer, 146, 41–58, 2014.

Veselovskii, I., Kolgotin, A., Griaznov, V., Müller, D., Franke, K., and Whiteman, D. N.: Inversion of multiwavelength Raman lidar data for retrieval of bimodal aerosol size distribution, Appl. Opt., 43, 1180–1195, 2004.

Veselovskii, I., Kolgotin, A., Müller, D., and Whiteman, D. N.: Information content of multiwavelength lidar data with respect to microphysical particle properties derived from eigenvalue analysis, Appl. Opt., 44, 5292–5303, 2005.

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