Geosci. Model Dev. Discuss.,
https://doi.org/10.5194/gmd-2017-220-AC3, 2018

Geoscientific Model Development Discussions

Interactive comment on "The degree of freedom for signal assessment of measurement networks for joint chemical state and emission analysis" *by* Xueran Wu et al.

Xueran Wu et al.

x.wu@fz-juelich.de

Received and published: 30 August 2018

Dear Reviewer,

Thank you for your attention to our manuscript, and your valuable comments on the research, as well as the suggestions for improving the paper. We have tried to address your concerns. Details on the changes are below.

Major comments:

- 1. Please explain in detail how the methodology proposed in Section 5 is different than the following work (if it is not, explain similarities and cite appropriately):

(1)https:// rmets.onlinelibrary.wiley.com/ doi/ pdf/ 10.1002/ qj.123
(2) https:// www.sciencedirect.com/ science/ article/ pii/ S1877050912002347
(3)http:// http:// www.met.rdg.ac.uk/ phdtheses/ Information%20content%20of%20observations%20in%20variational%20data%20assimilation.pdf

Thanks for reminding us about those references. We will definitely cite them in the following version of our paper.

The central difference of our study to those cited is the problem of model control by two (and generalizable to multiple) markedly different parameters: here initial values vs. emission rates (boundary condition) for chemistry transport models. The resulting outcome is expressed as partial DFS, related to each parameter, and dependent on the measurement configuration and transport (weather conditions). The proposed algorithm quantifies each sensitivity. More specific differences to the papers to be considered include:

The paper with the link (1) studied an ensemble-based data assimilation approach for Kalman filter and 3D-VAR. The studies of our paper are based on Kalman smoother and 4D-VAR.

The paper by Singh et al. (2013) we cited is actually a more complete version of the paper with the link (2). They used the information metrics to qualify the value of measurements in frame of ensemble runs based on 4D-VAR. However, in section 5 of our paper, all derivations are based on an ensemble Kalman smoother, which are different from 4D-VAR for nonlinear models. Besides, we are not only interested in the DFS of the total system, but also focus on improvement of each grid point from any given observation network and give metrics for concentrations and emissions separately. Further, we apply the singular value decomposition in Eq (53) to calculate $\tilde{P}$ without the calculation of Hessian of the cost function based on 4D-VAR.

In the thesis with link (3), the author studied the information content of obser-

vation based on 3D-VAR and 4D-VAR. But it is hard to see the ensemble-based studies in that thesis. Besides, the author of the thesis used the SVD technique into the observability matrix, which is different from Eqn (50) in our paper.

- 2. Please explain the computational cost of the methodology. How does the cost scale with the number of ensembles? With the data set? With the model size? With the assimilation window length?

Thanks for your comments. It shows in our tests that the computational cost is linearly increasing with the increasing of ensemble number and the assimilation window length. Besides, since the dimensions of the data set and model size are much larger then the ensemble numbers, the computation cost of our approach related to SVD (Eqn. 49) mainly depends on the numbers of ensembles. The explanation about the computational cost has been added in the example with the content " It has been tested that the computation cost of our approach is linearly increasing with the number of ensembles.(P14, line 16) " and "The computation times are approximately 8.1s, 28.5s and 39.4s in our tests with the above three different assimilation windows, from which we can verify that the computation cost is nearly linearly increasing with the length of data assimilation window. (P14, line 28)"

- 3. There is no comparison between the results obtained with the authors' approach and other existing approaches in the literature, e.g., Zupanski 2007. The numerical experiments would be more convincing if such a comparison was included.

Thanks for your suggestion. We didn't compare the results with other works just because it is novel to extend the transport model with emissions in our paper and we more focus on the quantitative balance between the original model state

and emissions. At the same time, we completely agree that it would be more convincing if we would compare our results with the others, once studies with a similar objective are available. We believe that it will be more significant to do the comparison if the approach is applied into the real atmospheric models. In fact, this is exactly what we are doing now (P21, line 16).

- 4. There is no discussion on how this can be applied to nonlinear systems. Adding some nonlinear chemistry to one of the test problems may help argue that this methodology, while developed under linear assumptions, can be in fact useful for nonlinear systems as well.

Thanks for your valuable advise. We mainly focus on the linear model and only give a linear example in this paper. We are so sorry that we didn't clearly explain that this approach can be also applied into nonlinear case. In general, the scope and limits of our approach is exactly the same as 4D-var and Kalman(-Bucy)-filters apply for non-linear (chemical) models. The central assumption is the validity of the linearized (tangent-linear) model. In the context of atmospheric chemistry modelling, the simulation must "be in the right chemical regime". (For a detailed discussion see papers Goris and Elbern, 2013 and 2015.) However, for critical cases of photochemistry with, say, observations of ozone as non-emitted product species, driven by emitted nitrogen oxides and volatile organic compounds, it can be safely expected that still the measurement network and transport conditions invoke the same results as in case of a passive tracer. A short discussion is given on Page 12 , 13 with Eqn (56) in the revised manuscript.

Minor comments:
Eqn. (25) does not seem to be the traditional 1-norm of a matrix. Please clarify the notation. Also, clarify what matrix square root is used, as there are infinitely many

possibilities. Eqn. (38), for example, does not seem to follow from the current (25) unless we are more specific.

Thanks for your reminding and sorry for the unclear notation. This is the nuclear norm of a matrix. Avoiding the confusion with the induced matrix 1-norms, we directly called it as the nuclear norm in the paper (Please see Page 7, line 30 in the revised manuscript). Besides, the definition of matrix square root is clarified in the revised version (Please see Page 6, line 25).

*Equation (20), /*we* define a matrix P: : :/ This matrix is the standard starting point in the definition of DFS, cf. Fisher 2003, Singh 2013. The discussion of Eqn (28) is confusing. The well accepted meaning of Ptilde(j) in the literature is: how much have we learned about variable xj from the data: from 0 (nothing) to 1 (everything). This is the amount learned about one degree of freedom (xj) out of n; the total number of degrees of freedom informed by the data/signal is the sum over all variables.*

Thanks for the comments and sorry for unclear expression. We have rewritten this part. Please see Page 6, from line 21 to line 2 on Page 7.

Editorial comments: Please carefully revise the writing of the manuscript (English correctness) as well as the spelling. There are hard to read (in English) formulations such as /aspiring a means/ (?). There are also typos in the manuscript. For example: /to what extend/ should be /to what extent/ in the Abstract, or /anfdWu/ should be /andWu/ in the Introduction, etc. Please avoid embedding URLs in text, they are best deferred as citations referring to web pages.

Thanks for the comments. We went through the whole paper to correct the grammar mistakes and improve the phrasing and attempt to provide a satisfactory writing in the next version. Due to the shortness of introduction, the URLs

have been deleted.

Sincerely,

Xueran Wu (corresponding author), Hendrik Elbern and Birgit Jacob