Atmos. Chem. Phys. Discuss., 15, C1996–C2002, 2015
www.atmos-chem-phys-discuss.net/15/C1996/2015/

Interactive comment on "A tropospheric chemistry reanalysis for the years 2005–2012 based on an assimilation of OMI, MLS, TES and MOPITT satellite data" *by* K. Miyazaki et al.

Anonymous Referee #1

Received and published: 28 April 2015

1 Overview

The manuscript by Miyazaki et al. is a tour de force of contemporary satellite data assimilation. The authors use remote sensing measurements of multiple different species to constrain both concentrations and emissions over an 8 year timeframe. The topic and use of models and measurements is suitable for ACP, and the paper is overall well written and clear. The scientific endeavor itself is quite challenging, yet the authors take a lot of care in evaluating the results of their assimilation, primarily through comparison to independent observations from other measurement platforms and sensitivity studies.

Overall, the results of the assimilation are quite encouraging. Nevertheless, unravelling the reason behind many of the observed changes proves challenging, given the tremendous number of degrees of freedom for an inversion such as this and the sheer volume of results to analyze. I wish there were more instances where the assimilation results could be translated directly into an improved understanding of some underlying deficiency in the model's transport, chemistry, or deposition. Further, there are a few areas where the inversion doesn't preform as well (such as for NO_2 concentrations in polluted areas) or where the results just don't make sense (such as the inference of large increases in NO_x and CO emissions in the US and Europe). The paper will be suitable for publication after the authors make revisions to address these and other comments provided below.

2 Major comments

- 242: It seems a bit out of place to bring up advantages of this approach over another here. But if you want to discuss this, then it also seems a bit odd to only mention the advantages of one approach – surely there are disadvantages as well.
- 260: It might be clearer to say: X^b is the ith row (or column? I'm not sure) of an $N \times k$ matrix X^b, where …
- 260: What is the size of N here? Is it just the size of the state vector (38), of the state vector times the physical system dimension?
- Eqn 3: To be consistent with Eq (1), shouldn't it be Y_i^b?
- Section 2.2.1: Something is missing from this section, namely the application of the forecast model itself. Where does that come into play? Should it act on x^b, or x^a from the previous step?

- Section 2.3: The explanation of the state vector is not clear. Line 300 implies that the following discussion pertains to both emission and concentration scaling, but then the description that follows on lines 302 - 304 is only for emissions scaling.
- Section 2.5: Why is the notation here for x and y different than other sections, where they are not italicized?
- Section 2.7: I realize that inclusion of scaling factors in each surface grid cell for emissions and each grid cell for concentrations is made feasible through the localization step (otherwise the system dimension would be too big). Still, the details of how this is setup are not clear. Are different sets of ensembles used within each localization region, or are there only 30 ensemble members spanning the entire globe? If the latter, this seems like a tremendously large space to span by so few members. The scale L seems to be of order of a few grid cells in the horizontal. But CO emissions have an impact on concentrations several km away. How is the setting of L to only 600 km justified?
- 593: Or this could indicate mis-specification of R?
- 604: This is a bit at odds with the figure and following sentence, which show that χ^2 is not constant. Perhaps rephrase?
- 701: One need not hypothesize about the information content of the TES data – the DOFs (trace of averaging kernel) will quantify this directly and could be used to check your explanation here.
- Section 4.3: One aspect that warrants discussion is the difference between the distribution of the analysis increment and the changes in OmF between the control run and reanalysis. Granted transport, chemistry, and the observation operator make these not correspond 1:1; however, I had a hard time rationalizing what I saw. For example, presumably much of the improvement in MLS ozone occurs

in the upper troposphere. How is it that the OmF for MLS, which was mostly negative south of 50 S but positive and negative north of 50 N be improved by an increment that is positive nearly everywhere?

- Also, line 717 implies that the drastic differences in ozone increments at different vertical levels is a sign that the system is working well. I would consider another possibility that the system is under-constrained, despite the results of the χ^2 test, and that the increments are exhibiting high frequency oscillations that lie in null space of your forecast model (rather than the observation space). The χ^2 test makes assumptions about the normality of the state parameters and observations, which may not hold.
- Fig 7: Maybe something is mis-labeled, but I had a hard time following the description of this figure in the text, in comparison to looking at the figure itself, which seems to show pretty much no improvement between the control and reanalysis for O_3. This is further confused by the Table which is referenced describing the errors in ppb, while the text describes them in %, which makes it a bit hard to follow.
- Fig 9: I also found it difficult to identify in these figures the features described in the text. Please instead provide a plot of the model minus observation for the control and reanalysis results.
- 854: Is the author refering to other previous studies (if so, please cite them) or the present work?
- 879: however, the performance was significantly worse at the surface
- 931: I'm not sure what is meant by “emissions constraints provided at the ground surface”. Also, it's not clear why the improvements with respect to the IAGOS data are attributed to emissions rather than the adjustment of the concentrations directly.

- 971-977: This is interesting. I wish there were more analysis like this showing how careful analysis of the model performance can be used to parse errors in emissions from errors in concentrations.
- 1052: I wonder to what extent the poorer performance for the NO_2 concentrations in urban areas is owing to the adjustments made only to NO_x and CO and not VOCs.
- 1055: This argument is a bit dangerous to make, since the perturbations from the model's natural dynamic chemical state are being forced by the state vector increments. So this would seem to be a potential pitfall of adjusting both concentrations and emissions simultaneously.
- 1070: The impact of the observations at one hour of the day should be much longer reaching, given the substantial role of NO_x on the chemical state. Despite the short lifetime of NO_2 itself, this impact can be multiple days, reaching several hundred km.
- 1172: It seems like if this still "remains an important issue" then you might find a more recent paper on the topic than 2003.
- Section 6: Several studies have investigated trends in NO_2 over the US and Europe in the past decade, with direct inferences for NO_x emissions trends being largely negative (e.g., Russell et al., ACP, 2012). Unfortunately, the results don't look anything like the trends shown here with increasing NO_x emissions. This is probably the weakest aspect of this paper, along with similar issues for the CO emission trends – how do you reconcile these results with known improvements in combustion efficiency / control technology in developed nations? This needs to be seriously addressed.
- 1335: But there is no improvement visible actually at the surface? Or is the scale of the plot just such that this improvement isn't visible? While I don't doubt that including emissions leads to changes in the mid-trop, it seems it should though at least make an equal or lager improvement directly at the surface.

- 1342 - 1348: Shouldn't they both lead to the same O_3 trends, the one which matches the observations, just by different means?
- Table 6: The caption should state what is shown (percent error? ppb?), and if the values are model minus observations or vice versa.
- Section 7.4.1: All though I'm not quite sure what the numbers in Table 6 represent (see previous comment), it seems odd that in the NH the differences at 200 hPA (16.3 vs 13.2) would be larger than differences at the surface (0.1 vs - 2.3) for the test using different prior emissions. Why would the impact of changing the prior emissions be greatest in the upper trop? Perhaps the results of this test would be better served by showing a plot of the difference in the posterior emissions between the standard and HTAP-based reanalysis.
- Section 7.6: Why not just report the uncertainty reduction as measured by the posterior / prior error, rather than the ensemble spread?
- Section 7.7: I found the first half of this section (1620 - 1639) quite speculative. It could be removed, given the paper is already quite long.
- Conclusions: While the overall results of this work are indeed impressive, I feel there were a few aspects which need to be made more transparent, such as the inability to improve surface NO_2 concentration in polluted areas, or the strange trends in NO_x and CO emissions in the US and Europe. A bit more balanced evaluation of all strengths and weaknesses would be good.

3 Corrections

- 4: retrieval data → data
- 53: to develop → in developing
- 55: information, → information
- 56: the estimates → estimates
- 121:), by →) by
- 247: the analysis performed → performing the analysis
- throughout: number % → number%
- 578: no new paragraph
- 619: too large → excessive
- 1151: while → while it

Interactive comment on Atmos. Chem. Phys. Discuss., 15, 8687, 2015.