Atmos. Chem. Phys. Discuss.,
doi:10.5194/acp-2016-1043-RC2, 2017
© Author(s) 2017. CC-BY 3.0 License.

Atmospheric Chemistry and Physics Discussions

Interactive comment on "Evaluation of ACCMIP ozone simulations using a multi-constituent chemical reanalysis" *by* Kazuyuki Miyazaki and Kevin Bowman

Anonymous Referee #1

Received and published: 27 January 2017

Evaluation of ACCMIP ozone simulations using a multi-constituent chemical reanalysis by Kazuyuki Miyazaki and Kevin Bowman

Overview:

The authors use a new chemical reanalysis based on multiple satellite observations to evaluate ACCMIP ozone simulations for the period 2005-2009. This evaluation is juxtaposed with an evaluation using ozone sondes. Thereby, the authors quantify the sampling biases of the ozone sonde observations and its impact on the evaluation results. The evaluation with the ozone re-analysis, which was shown to be in good agreement with the ozone sonde observations, give a more comprehensive picture of the model deficiencies than ozone sondes.

General remarks:

Chemical reanalyses are newly emerging data sets. The authors present an interesting and convincing application of these data set for the evaluation of atmospheric chemistry models. To make this point they show that the evaluation with ozone sondes can suffer from biases because of the spatial coverage and representativeness as well as the specifics of the temporal sampling of the ozone sonde network. The work is timely and of high interest to the scientific community.

The paper would benefit from a more stringent focus on the ozone sonde sampling biases and its impact on the evaluation because this is the actual novelty aspect of the paper. The ACCMIP models have been evaluated. So only the differences of the new evaluation approach with previous work is of interest. I would recommend to add to the title "- focus on ozone sonde sampling biases" or similar. The sampling biases should be mentioned and discussed in abstract and introduction more clearly.

To get a better understanding of the sampling biases, i.e. the difference in the mean over area averages using all grid points at regular intervals or only the stations locations at the time of the observations, it is recommended to show the sampling biases not only for the differences between ACCMIP models and the reanalysis but also for the Re-analysis and the model runs, including the control run itself. It would be interesting to see to what extent they differ as the reanalysis may also be effected by the "sampling biases" of the assimilated observations. A strong sampling biases for model result will help to convince modellers to use reanalysis data for model evaluation.

The authors should aim to provide a better understanding of the reasons of the sampling biases. Do they come more from spatial heterogeneity or the variable temporal sampling. The latter can be estimated by comparing re-analysis means at a 2 hourly resolution or only at the ozone sonde observing times.

For the sake of consistency the quantification of the sampling biases should be carried out for one set of latitude bands in the same way as in the more regional analysis

presented in section 5.

Section 5 “Impact of Sampling on model evaluation” discusses the regional biases and the general problem in a lot of detail but sections 4.2 and 4.3 discuss already the sampling biases for the latitude bands. I recommend moving the introduction of the sampling biases to an earlier section (2).

The discussions section, in particular 6.3, does not discuss the direct results of the paper but gives an outlook on other potential aspect of the usefulness of the evaluation with chemical re-analyses. However, the positive impact on species not directly assimilated has not be demonstrated in the paper. Also, the four year comparison is not long enough to infer trends and longer re-analysis of atmospheric composition are likely to suffer from temporal artefacts because of the changing observational system. I would therefore not discuss in detail these aspects in the paper as there is not enough evidence given to support them.

The used ozone sondes observations need to be clearer identified and their sampling discussed. A table of the used ozone sondes, their sampling frequency and outage in the period and mean should be summarised in a table not only for the regional areas but also for the latitude bands. It should be made clear which stations are used for global/hemispheric stratification in Figures 1 to 7 and the more regional stratification Figure 9-10.

Specific remarks,

PL1: 5 Please “the” before instrument names

P1L5: Please add a sentence on the advantages of using a 3D re-analysis rather than ozone sondes for the model evaluation.

P1L6: Please ad here or at L 12 the problem of the ozone sampling biases

P1L12: Please state more clearly the differences in the evaluation results when using the re-analysis as complete field and on the the ozone sonde observation locations

and times only.

P1L24: better “transport”

P2L8: Please add some references for these evaluation studies

P2L13: there is a “First” and a “Third” (L19) but I did not find a “Second”

P2L18: The sentence starting with “However, . . .” is a strong motivation for the paper. Please elaborate and also mention that the climatologies do no capture the temporal variability of the observed ozone.

P2L30: Please consider citing overview papers such as Bocquet et al. (ACP 2015) or Sandu et al. (Atmosphere, 2011)

P4L4: Please comment how this is related to resolution of the evaluated ACCMIP models.

P4 L22: Please explain how the ensemble is constructed, i.e. what parameters are varied to get a different ensemble members in the EnKF. This information is important because you later use the ensemble spread partially as indicator of the analysis error.

P4 L23: “satellite retrieval operator”? This implies radiances i.e. Level 1 were assimilated, which is perhaps not the case. Please clarify.

P5L4: Please mention if it could be shown that the modulation of the lifetimes was an improvement.

P5 L14: Please provide a table with the assimilated retrievals and additional information such as assimilated height range, temporal data coverage and an indication of observation errors statistics.

P5L21: Please elaborate on the period and the meteorological input for this time-slice setup. It is important to know what sort of realism can be expected from the simulation if they are compared against observations.

P6L9: Please clarify which station were used for the comparison. The ones listed in table 4 ? If so mention it here. Provide information about station numbers and individual temporal coverage as this may vary greatly and contribute to the ozone sonde sampling bias.

P6 L17-23: This description of the model changes may better put in the model description section.

P6L30: Please clarify what the differences in the assimilated observations are between this data set and the previous one.

P7L10: Please clarify what temporal averaging the temporal correlation is based on (i.e. monthly means, annual means, instantaneous values etc.). It is good practise to de-seasonalize the time series to get a more meaningful information about the temporal correlation. On the other hand, 5 years might be too short to obtain a robust information about seasonality and year-to-year variability.

The numbers in figure 1 indicate a very good reduction in biases but far less so for variability measures. (The reduction in RMSE seems dominated by the bias component and temporal correlation is less improved). This seems to contradict the theoretical basis of data assimilation, which is meant to reduce the error variance assuming bias free model. A further discussion would be helpful.

P7L14: As you also show the ozone sondes in the sections on seasonal variation and hemispheric gradient, it seems odd not to show the ozone sondes observations in Figure 3. Please add a further panel with colour dots at the station location.

P7L18: It is confusing that you choose a different latitude bands for table 1 and table 2. Please use the same selection of latitude bands through the paper.

P7L18: Please, clarify how spatial r was calculated (only using the 5 year mean, based on lat-long grid points or area-weighted grid-points etc. Consider filtering small scale noise by averaging over areas corresponding to the resolution of the reanalysis.

The spatial correlation coefficient presented here seems less suited to express agreement in spatial patterns, which would be meaningful for the understanding of the model performance. Spatial r might be too much effected by the underlying spatial variability of the actual fields, thereby penalizing fields with greater more random variability i.e. standard deviation.

P7L19: The lower spatial correlation coefficient at p=500 hPa in NH could simply be caused by a different transport patterns (winds) and larger heterogeneity than in SH. Good correlation at the surface could be simply because a good match of emission patterns. High correlation at 200hPa in extra-tropics could mean that the transition in to the stratosphere agreed reasonably well. So are the different spatial r really helpful to distinguish model performance?

P8L4: Please clarify again how the statistical variables shown in the Taylor diagram were computed. Given my scepticism about the meaning of the spatial correlation, I would consider omitting Figure 4 and shortening the discussion.

P9L6: Please clarify how the seasonal amplitude was calculated. How was made sure that “noise”, i.e. unstructured variability, was not attributed to the seasonal amplitude.

P9L12: In section5 you discuss the sampling bias with respect to the regional areas. You should also discuss the sampling biases w.r.t to the selected latitude bands. This is needed because you also discuss model performance for the latitude bands. As mentioned in the general remarks, please also indicate the difference between the model results sampled at ozone locations and observation times and the area averages.

P9L13: Please confirm that the average is area weighted and not based simply on lat-long grid boxes, which decrease in size towards the poles.

P10L8: Please clarify how you exactly calculate the hemispheric gradient both for the gridded fields and for the ozone sonde observations.

P10L15: add missing “At” before “Around ”

P10L18: How do these value compare to values from the literature?

P10L26: As you already discuss sampling biases it a bit inconsistent to put the section at this place. This very good introduction to sampling biases (p10L27 – p11L15) should come earlier in the paper, i.e. in the part when you discuss the methods (section2)

P10L15: The sampling biases depends on the averaging area and the selection of ozone sondes. The sampling biases estimated by using your re-analysis should be presented for the Tilmes regions as well as for the latitude bands (choose one set only) in and uniform way. As model results are often evaluated for the latitude bands , this information would be very interesting for the scientific community.

P10L24 Please add also the stations used for the latitude bands averages in table 3.

P12L1: I think there is would be very good to compute the sampling biases also directly for the re-analysis i.e. the difference between the re-analysis sampled as ozone and as area-time averages. This information would be in my opinion of more general meaning than the values for the ACCMIP error.

P12L1: Why does table 5 show the median whereas otherwise only the ensemble mean is discussed or shown. (Using only the median would be perhaps a better option overall)

P13L8: The ozone network in SH high-latitudes is actually quite high because of the need to monitor the ozone hole. The launch frequency varies for some stations a lot because more sondes are launched during the ozone hole season.

P13L24: Please add also the sampling biases for the latitude bands.

P14 L14: Please clarify how much of the analysis ensemble spread depends on sometimes arbitrary choices to cause spread between the ensemble members.

P14L15: Please clarify to what extent the analysis uncertainty is controlled by the uncertainty of the assimilated observations.

P14L21: How does the ensemble spread relate to the spread of the ensemble in the EnKF. Could the ACCMIP ensemble spread be used to verify the EnKF ensemble spread ?

P15L3: Please see my general comment on this chapter. Improvement on species not directly assimilated needs to be demonstrated. Long-term reanalysis could suffer from artificial jumps because of the change in the observing system (for example degradation of TES after 2010).

P15L29: I don't understand this conclusion at all. Re-analysis are only valid for present day conditions when observations are available. They cannot be used to for pre-industrial estimates nor the differences with today's values.

P16L4: Please mention that you (only) consider ozone sondes as reference in this paper.

P16L6: Please mention the advantage of using a re-analysis, i.e. a gridded field. Please mention that the biases of the re-analysis against ozone sondes are small.

P16L17: Please add a statement if these finding are consistent with other evaluation studies, i.e. the Young et al. paper.

P16L20-30: Please give some numbers for the sampling biases. Also include the sampling bias w.r.t to latitude bands. p17L4: Please add a statement that it will be a challenge to combine all these observations in a consistent way in a more long term re-analysis.

Interactive comment on Atmos. Chem. Phys. Discuss., doi:10.5194/acp-2016-1043, 2016.