Geosci. Model Dev. Discuss., 5, C401–C420, 2012
www.geosci-model-dev-discuss.net/5/C401/2012/

Geoscientific Model Development Discussions

Interactive comment on "Assimilation of OMI NO_2 retrievals into the limited-area chemical transport model DEHM (V2007.0) with a 2-D OI algorithm" *by* J. D. Silver et al.

J. D. Silver et al.

jds@dmu.dk

Received and published: 8 July 2012

The original point-by-point reply to reviewers (posted in this interactive discussion as a supplementary file) did not include the reply to point 4 of reviewer 2. A number of minor oversights were also noted, and these have been corrected in the text below. The following version of this point-by-point reply replaces the version uploaded on 02 July, 2012.

We are grateful to the reviewers for their many instructive comments, questions and suggestions. Based on the insights gained from the reviews, we have revised many different aspects of this work. The following is a summary of the changes made in the

manuscript, the methods used, and other relevant points.

- The two-dimensional formulation has been replaced with a three-dimensional formulation.
- The background error covariances were estimated using the NMC method, which makes no use of the satellite retrievals. The revised correlation model now allows for seasonal and diurnal cycles, spatially variation in the variance field, and non-separable, horizontally homogeneous and isotropic correlations.
- The revised version of the manuscript includes an examination of seasonal, diurnal and vertical patterns in the background error covariance parameters.
- Observation errors now have the appropriate units (they were incorrectly specified in some of the assimilation experiments in the original work).
- The Hoelzemann et al. (2001) correction to the background covariances is not applied.
- Averaging kernels were applied in the observation operator.
- Forecasts were performed to assess the potential for use of the analysed concentration fields in operational chemical weather forecasts. This replaces the comparison of a pair of simulations started with different initial conditions.
- The manuscript has been restructured to deal with method and results thematically, rather than separately.
- The comparison with ozone sondes has been removed, but averaged NO_2 and O_3 profiles are presented.

- Three assimilation experiments were run to examine the issue of correlated observation errors, instead of the five experiments in the original manuscript (which were presented with the aim of finding a good configuration of the OI scheme).
- For stations measuring NO_2 by chemiluminescence, the modelled equivalent is taken as the sum of NO_2 and nitrates.
- Figures 4-6 in the original manuscript have been removed.
- Four regions within the model domain were defined, and time-series and verification statistics are presented for each of these.
- The version of the chemistry-transport model DEHM used in the study has been upgraded.
- The most recent version of the retrievals was used, rather than an older version of these data (as described in the initial version of the manuscript).
- A great deal of the text has been rewritten to describe the new results and the revised methods.

1 Reply to reviewer 1

1. *The major driving force of DA is to improve the model forecast capability. Seeing the short memory of chemical initial conditions in CTM, it will be more interesting to show how DA will change chemical simulation in a short time period, e.g., 3 days, instead of an annual simulation.*

   The updated manuscript includes the results from short forecasts.

2. *More seriously, the authors did not report how the vertical profile used in the OMI retrieval compared to the modeled NO2 profile. The total amount of gas retrieved from satellites is very sensitive to the assumed vertical profile in the retrieval algorithm. One cannot simply compare the modeled total column values to the satellite retrieval without regard to its vertical structure.*

   The assimilation scheme has been upgraded from a two-dimensional to a three dimensional scheme, such that the implicit assumption of a perfect profile shape is not required. The upgraded observation operator takes account of the OMI averaging kernels, thus accounting for the sensitivity of the retrieved value to NO_2 at different heights.

3. *The paper is not particularly well organized, making it unnecessarily long. For example, all the figures and tables were mentioned in section 3 without detailed explanations. Then they were presented again in section 4. This is not a good flow. Please consider re-organize these two sections to reduce the redundancy and improve the readability.*

   The paper was originally organised so as to separate the methods, the results and the discussion. We have reorganised sections 3 and 4 to combine methods and results thematically as suggested.

4. *Section 3.3, Figure 4 is not necessary and can be removed.*

   This figure has been removed.

5. *The authors replicated the missing data to calculate the annual/seasonal average. Suggest only including the model results when observation is available to calculate the modeled average.*

   We did not replicate any missing data in these calculations; instead, we replicated the pattern of missingness in the modelled data, which is precisely what

the reviewer suggests.

6. *Section 4.1, line 15, suggest changing from "..is not so simple.." to "..is not so straightforward..".*

   The sentence in question does not appear in the revised manuscript.

7. *Section 4.1, line 25, Figure 3 hardly leads to the conclusion that "Without DA, the model underestimates . . . and does not capture . . .". A probability density function plot may be a better option.*

   This sentence has been removed.

8. *It should be clearer to show the NO2 difference (model âĂŽÄì OMI) maps in Figures 5 and 6.*

   These plots have been removed.

9. *Sections 3.5 and 4.4, it appears less related to the presentation but emphasizes the importance to compare short term DA results.*

   The simulations originally presented in this section have been replaced with an examination of short forecasts.

10. *Figure 7 is not very clearly presented. Consider modifying.*

    The comparison with ozone sondes has been removed, although spatially- and temporally-averaged modelled vertical profiles are shown in figure 12 of the revised manuscript.

2 Reply to reviewer 2

1. *The authors did not make full use of the OMI data product. Averaging kernels provided in this product have not been used. Also the error estimates available in the product have been used only in experiment 4 (5), which apparently is not even the optimal choice (which is exp 3 with constant errors). The vertical sensitivity (or averaging kernel) of the measurement technique is strongly height dependent, and residual clouds have a strong impact on what can be observed from space. In particular OMI is extra sensitive to NO2 above the PBL, and the amount of free troposphere NO2 modelled by DEHM should be investigated, reported and compared to other studies (see references). Why have the kernels not been used?*

   We thank the reviewer for drawing our attention to this point, which we agree is important. We were not aware previously aware of the role of the averaging kernels. The updated manuscript describes a three-dimensional treatment of the assimilation scheme, and the OMI averaging kernels have been used to calculate coefficients of the linear observation operator.

2. *The implementation of the observation covariance R is not clearly explained, see Table 1. What does R_{ii} = 1 mean? How are exp 4 and 5 implemented?*

   We agree that this was not clearly explained. The basic units for the retrievals and the observations was 10^{15} molecules/cm^2. Thus instead of R_{ii} = 1, it should have been stated that R_{ii} = (1.0 $\times$ (10^{15} molecules/cm^2))^2. Exp 4 and 5 of the first draft used incorrectly specified observation errors, as discussed below. The discussion of observation errors forms a central part of the revised manuscript, and we have ensured that these units have been correctly specified in the revised assimilation scheme and experiments.

3. *The authors mention that negative estimates of the tropospheric column are excluded. However, such negative values in the OMI product seem to represent the uncertainty in the retrieval in cases where the amount of NO2 is less that the error bar. Excluding those negative numbers will generally result in positive biases in clean area. Does this explain part of the positive adjustments made in the assimilation process?*

   We are grateful to the reviewer for making this instructive point. In the simulations presented in the revised manuscript, negative estimates were included. We mention this point in section 2.1, which discusses the OMI retrievals.

4. *The lifetime is an important issue. If (in Summer) the NOx lifetime is only a few hours, the impact of data assimilation will be lost within the same few hours. Since OMI is only available once per day, how can the authors claim that OMI can be used for reanalyses using the OI technique? Fig. 8 seems to suggest longer lifetimes than suggested by the introduction (between 3 and 13 hours). Please explain this difference. Also, the difference between the reference run and exp 3 seems to be tiny (the NO2 curves are basically on top of each other). Is this correct? It seems incompatible with the considerable adjustments seen in figures 4-6. What is the mechanism to keep the NOx adjustment information in the model for 24-48 hours.*

   The curves mentioned in figure 8 (of the initial version of the manuscript) show the relative difference, defined as the magnitude of the difference *relative* to the magnitude of the difference in the initial conditions. As can be seen in the time-series of spatially averaged concentrations (the second set of curves, which basically overlap), there are only minor differences the different simulations during the summer months, hence the differences present in the initial conditions were rather small. Also, the spatial averaging smoothes out spatial variation, thus minimising apparent differences.

   Other features may be contributing to the effectively longer relaxation time (24-48 hours) than the atmospheric lifetime of NO_x reported in the introduction (3-13 hours), namely flow-on effects resulting from the correction of NO_2 concentrations. The initial conditions used for the two simulations reported here were taken, respectively, from the free run (i.e. no assimilation) and from a run with assimilation. The assimilation of NO_2 impacts O_3 concentrations indirectly through the chemistry, and differences in O_3 levels in the initial conditions may contribute to the longer-than-expected relaxation time.

   As for the feasibility or usefulness of the assimilation, it is shown that it yields a better temporal correlation with observations. This appears to be primarily due to better resolution of the winter-time peak concentrations, associated with stable atmospheric conditions. It is during winter that the lifetime of NO_x is longest, and thus in this period the effects of the assimilation have a longer impact in the simulation.

5. *The Hollingsworth and Lonnberg approach to estimate model and observation contributions to the error covariance is not simply applicable to satellite observations. The main assumption in this approach is that errors in the measurements are uncorrelated in space, which is often a reasonable assumption for well separated surface stations, but generally does not apply to satellite data products, given their sensitivity to e.g. cloud fields and surface scattering properties. When errors are correlated the estimate of the background error weight θ_b (eq.2) is no longer meaningful. The error correlation above the intercept could still be due to the model, e.g. wrong local (within one grid cell) emissions. On the other hand, part of the spatial correlations could be due to the satellite observation errors.*

   The reviewer has brought to our attention a serious flaw in the approach used to calculate the error covariance matrices in the initial version of the manuscript. In the revised version, the NMC method was used to estimate background error

covariance parameters. In the revised manuscript we also discuss the problem of correlated observation errors and the experiments presented address this issue.

6. *The authors do not present a-posteriori validation for their OI approach. Are the observed differences (observation minus forecast) compatible with the covariance matrices used (chi^2 test)?*

   This suggestion has been incorporated into the revised version of the manuscript.

7. *It seems that the main impact of the assimilation is a correction of an overall negative bias in the model. Would it be better to replace the OI scheme by a bias-correction scheme (described in papers by e.g. Dick Dee)?*

   In work not presented here, we implemented the sequential bias correction scheme described in Eq. 17-18 of Dee (2005, QJRMS). However the estimated bias field appeared to grow in magnitude with successive assimilation cycles. We suspect that this was due to incorrectly specified bias covariance parameters (which were estimated based on the analysis increments from Exp. 2 of the revised manuscript, using a similar estimation procedure as for the climatological background covariances. However limitations on time have prevented further investigation into this area. The issue of bias-correction is raised in the discussion, and it is a potential extension to the OI scheme presented in the revised manuscript.

8. *The authors compare with profile information based on ozone sondes. I fail to see why this is very relevant for this paper. Instead it would be good if the authors can do some "confidence building" to show that the NOx/NO2 profile in the model is reasonable up to the tropopause, by comparing for instance with other models and/or with available observations. Just showing the model profile would already be of use.*

   The comparison with ozone sondes has been removed from the manuscript, since on reflection we agree that it adds little to the study. The revised manuscript includes an illustration of temporally and spatially averaged NO_2 profiles from the different simulations. However they are presented without reference to observations, partly due to the lack of available NO_2 profiles and partly because the differences in the mean profile between the simulations using assimilation were relatively small.

   We are not aware of regular measurements of NO_2 profiles with in situ measurement techniques. It appears that such data are only collected in some measurement campaigns, and as such are difficult to access. There are ground-based measurements of total column NO2 (from UV visible spectrometry or infrared Fourier-transform spectrometers), however these doe not provide the concentration profile. We applied for access to the MOZAIC database (containing measurements from aircraft of certain trace gases, including total nitrogen oxides), however there has not been sufficient time to do justice to these data, which require careful collocation in time and space. We note that the DEHM does not currently include emissions from aircraft, and these data are collected along flight corridors.

9. *p312-313, intro: The part on NOx chemistry and impact on health is basic background knowledge and can be summarised in a few lines with references. I would have expected this part at the beginning of the introduction. The lifetime seasonality remark is relevant and should be kept.*

   We have abbreviated this section as suggested.

10. *p314, l26: Negative estimates of the tropospheric column are not an artifact of the retrieval, but represent the uncertainty in the retrievals.*

    In the simulations presented in the revised manuscript negative estimates were not systematically excluded.

11. *p315, l12: "if any"*

    This sentence does not appear in the revised manuscript.

12. *p315, l14: "unreliable" what is the criterium?*

    In retrospect, this was a poor choice of words. In the original version, pixels with negative retrievals or with a cloud radiance fraction greater than 50% were excluded. In the revised work, we excluded pixels with a cloud radiance fraction greater than 50% or surface albedo greater than 0.3; both these criteria are stated as recommendations in the the DOMINO user guide. Negative tropospheric column concentrations were included, assuming the aforementioned quality control criteria were satisfied.

13. *p316: Have the DEHM free troposphere (NOx) concentrations been validated? Are relevant processes for the free troposphere (lightning, convection) included in the model?*

    As mentioned above, we have not had access to measurements of NO_2 or NO_x vertical profiles (apart from the MOZAIC data, which, as mentioned above, require careful treatment and there has simply not been time to go into this). Thus DEHM free troposphere NO_x concentrations have not been validated. The DEHM participated in the recent AQMEII model-intercomparison, in which a comparison with the MOZAIC data-set is being undertaken, however the examination of the MOZAIC data has not yet been completed, as initial efforts focussed on analysis of surface measurements.

    Natural emissions of NO_x (from lightning, soil and wild-fires) as well anthropogenic emissions (mainly combustion processes) are included, but aircraft emissions are not. We have now clarified this in the text.

14. *p317: Eq 2: B has the dimension of x-squared, so the column concentration squared. Is this dimension included in θ_b ?*

    All variables in Eq. 2 (of the original version) are scalars, and it is applied to calculate entry B_{ij}. The parameter e_b (the background error weight) is also a scalar. The reviewer refers to a θ_b, which does not appear in the original discussion paper and is probably a misreading of e_b, since the two look very similar in the font of the GMDD manuscript.

    The parametrisation of the background error covariances presented in the initial manuscript has been substantially reworked, and the parameter e_b does not feature in the revised manuscript.

15. *p318, eq 3: ? Please provide additional motivation for this form. Why is model-forecast error correlation length depending on the number of observations? From a theoretical point of view it should be completely independent of the density of the newly added observations, but there can be a (often weak) dependence on the observations of the previous day. Given the short residence time of NOx I would expect this dependence to be very weak.*

    On reflection, we agree with this comment. The correction to the correlation length-scales does not feature in the specification of background covariances presented in the revised manuscript.

    We agree with the reviewer that from a theoretical perspective, the background error covariances should be independent of the observation density. The adjustment, suggested by Hoelzemann et al. (2001, Phys. Chem. Earth), is designed to give less smoothing in regions of the model domain where there is a high observation density, thus providing higher-resolution increments in such regions.

16. *p318, l10: Why would the length scale for ozone be relevant for NOx ?*

    On reflection, the comparison was perhaps not very useful. It was simply in contrast to the previous work (Frydendall et al, 2009, ACP) done with the same

two-dimensional OI assimilation scheme as presented in the initial version of the manuscript.

17. *p318, Sec 2: Is the model bias-free compared to OMI: long-range correlations will mess up the length scale parameter estimates. Has the estimate of the innovation correlation been debiased?*

    In the initial manuscript, no such debiasing was applied. The NMC method for estimating background covariances, as presented in the revised manuscript, does not depend on the OMI columns, so the correlation length-scales cannot be contaminated by correlated observation errors.

18. *p318. l10: See my general remark on the Hollingsworth-Lonnberg approach.*

    As mentioned above, we have replaced the Hollingsworth-Lonnberg approach with the NMC method in the revised version of the manuscript.

19. *p318, l25: Why is a binning to grid points needed? The procedure to compute correlations is described very extensively (six steps). I would suggest to summarise this in 1-2 lines.*

    The description of the binning procedure has been removed as it was part of the Hollingsworth-Lonnberg method, which is not used in the revised manuscript. The detailed description of the binning strategy was suggested by one of the co-authors, who found the original 1-2 line description unclear.

20. *p319, eq 4: Please explain the r^2 normalisation in this formula. Does this account for the increasing number of pairs at larger distances?*

    The normalisation factor is required to capture the decay, which is often quite short, with respect to the maximum separation distance considered. Greater weight is thus placed on points of interest (i.e. in the region of correlation decay, with shorter separation distances). This has been clarified in the text.

21. *p320, l22: R_{ii} should scale like the observation error squared. Has this been done, or is $R_{ii} = \sigma$, which would not be consistent with the Kalman filter equations. Please be more explicit on the exact form of R_{ii} in exps 3, 4 and 5. What does R_{ii}=1 mean?*

    The reviewer highlights some shortcomings both in the presentation of the observation errors and in some of the experiments. The meaning of The R_{ii}=1 has been explained above. In Exp. 4, we used $R_{ii} = \sigma_i$ where we should have used $R_{ii} = \sigma_i^2$. This mistake has been corrected in the simulations presented in the revised manuscript.

22. *p320, l27: Again, the form y/σ has no dimension, while R_{ii} should scale like the observation error squared. How is this implemented?*

    The reviewer is correct to point out that these observation errors have no dimension. It was implemented exactly as was stated, and was thus incorrect. This flawed experiment does not feature in the revised manuscript. The observation errors used in the experiments presented in the revised manuscript have units of the square of the observations themselves, that is $(10^{15}$ molecules/cm$^2)^2$.

23. *Sec 3.2: Are the EMEP observations representative for the grid cell of DEHM? Please discuss this issue. Or would one expect offsets? NO2 surface measurements often contain other nitrates like HNO3, PAN. Has this been corrected for?*

    Within the EMEP network, a number of different techniques are used to measure NO_2, namely chemiluminescence, spectrophotometry and the Griess-Saltzman reaction. We have discussed the matter with a specialist in air quality monitoring, and were informed that measurements of NO_2 from the chemiluminescence are contaminated by other nitrates (thus chemiluminescence-detected NO_2 measurements should be compared with modelled NO_2 plus the sum of modelled nitrates), however data from the other techniques can be compared directly with the

modelled NO_2. We have adjusted the verification accordingly. This is described in section 3.2 of the revised manuscript. We thank the reviewer for bringing this to our attention.

24. *p321,322: It would be useful to provide a short explanation how R2, NMSE and FB are defined.*

    In the revised manuscript only the bias and correlation coefficient are used (as opposed to the fractional bias or normalised mean-squared error). Both of these are widely used verification statistics in air quality modelling, and we do not deem it necessary to define them here (other than to clarify which direction the bias was calculated, bias = modelled minus observed, which we state in the text).

25. *Sec 3.4: replace "ozonosondes" by ozone sondes or ozonesondes (several times in text)*

    The results showing ozone sondes have been removed from the revised manuscript (since the difference in ozone concentrations between the different assimilation experiments was found to be rather small), along with this spelling mistake.

26. *p322: Why are De Bilt and Legionowo chosen?*

    There were 7 NDACC sites reporting ozonesonde data for 2005, and these were just the first two on the list. There were too few sites for a thorough analysis, so they are just presented as illustrative. As mentioned above, the results showing ozone sondes have been removed from the revised manuscript.

27. *Section 4. Sections 3 and 4 should be merged into one section, e.g. combine 3.2 with 4.1 etc. . . The split is artificial and is not helping the reader.*

    The paper was originally organised so as to separate the methods, the results and the discussion. We have reorganised sections 3 and 4 to combine methods and results thematically.

28. *p323, l15: It is a bit disappointing that constant errors seem to perform better that using the error bars in the OMI product. Adding more information should help.*

    As mentioned in the manuscript (both in the original and the revised version), the retrieved tropospheric column concentration is highly correlated with the reported error in the DOMINO product. This is may be an appropriate description from a measurement perspective, however it can potentially be problematic in an assimilation context. If the higher retrieved values are assigned higher observation errors, then these data will have less influence on the analysis, resulting in negatively biased analysis increments. This is discussed in Sec. 3.1 of the revised manuscript and motivates one of the experiments.

29. *325, sec 4.4: The relaxation time of 2-3 days is long compared to other lifetime estimates. See above.*

    Other features may be contributing to the effectively longer relaxation time reported here, namely flow-on effects resulting from the correction of NO_2 concentrations. The initial conditions used for the two simulations reported here were taken, respectively, from the free run (i.e. no assimilation) and from a run with assimilation. The assimilation of NO_2 impacts the O_3 concentrations indirectly through the chemistry, and it may be that the differences in the O_3 in the initial conditions are responsible to the longer-than-expected relaxation time.

30. *325, bottom: It is not clear to me why the lower correlations are resulting from a spread of information to cloudy parts. Please explain more clearly. There can be multiple other reasons.*

This was just a hypothesis, and we agree that other explanations are possible. However the issue at stake, of understanding lower O_3 correlations, is not relevant in the revised manuscript; the whole OI system has been reworked, and the resulting O_3 levels show higher temporal correlation than for the free run in four of the five regions considered. Hence the point about information-spreading has not been raised in the revised manuscript.

31. *326: The NO2 profile shape is kept constant. I do not see how this can be validated with ozone. Please provide a discussion on the quality of the NO2 model profiles (see above).*

    On reflection we agree that this was not a useful comparison, and the ozone sonde results do not appear in the revised manuscript. The three-dimensional assimilation scheme presented in the revised manuscript does not keep the profile shape constant and does not depend on the assumption that the initial profile shape is accurate. We include a discussion of the NO_2 model profiles, although without further measurement data, it is difficult to comment on their quality.

32. *326, l21: The lower background weight: See remark on the Hollingsworth and Lonnberg approach above.*

    As mentioned above, the Hollingsworth and Lonnberg method for estimating error covariances has been replaced by the NMC method.

33. *327, l20: It would have been good to discuss the forecast impact in more detail, as was done in Wang.*

    This is addressed in section 3.3 of the revised manuscript.

34. *Table 2 and 3: The bias of the model compared to OMI and surface observations is 5, quite large (38%). It seems that a bias correction scheme could be more appropriate, see general notes above.*

    These estimates were not calculated properly, since averaging kernels were not applied. Bias correction is mentioned in the discussion section of the revised manuscript.

35. *Fig 2: Density plot: it would be better to use colors and include a log-scale legend to provide actual counts. In this way one can judge the spread of values in a better way.*

    Given that the revised manuscript describes the NMC method, rather than the innovation statistics approach, for estimating background covariances, this particular figure has been removed. A similar figure (Fig. 3c) has been included in the revised manuscript, and the actual counts in each cell are indicated by a legend beside the density plot, following the reviewer's suggestion. The sampling of points has been done slightly differently (compared to the original manuscript), resulting in an even distribution of correlations for a given separation distance. Hence the log-scale suggested is not required.

36. *Fig 3: There is apparently very little seasonal variability in model? Why? As mentioned in the intro the lifetime strongly depends on the season. What does "calculated" refer to, i.e is this the reference run or one of the assimilation runs?*

    The lack of seasonal variation in the modelled time-series appears to be due to limitations on the vertical resolution of the boundary layer in the meteorology from the Eta model, which uses the η (step-mountain) coordinate. In effect, there is a lower limit of mixed-layer heights that can be resolved. This is important during the winter months, when stable conditions and nocturnal inversion layers result in accumulation of directly-emitted components. A large proportion of the seasonal variation not captured by DEHM can be attributed to difficulties in capturing low mixed-layer heights. This is discussed in the revised manuscript, and results from a DEHM simulation using meteorology from MM5 (which uses terrain-following σ vertical coordinates) are presented. The nature of the σ vertical coordinates

allows for a higher vertical resolution of the boundary layer, which is critical for diagnosing low mixed-layer heights.

The term "calculated" referred to the modelled results (which was either the reference run or the assimilation run depending the title of the plot, which appears in the top-left corner). This figure does not appear in the revised manuscript but are replaced with similar figures (Figs. 7-9).

37. *Fig 4: How can there be such large increases at spots where no OMI data is available (grey)?*

    As mentioned above, this may be due to the indirect effects from the chemistry. The changes to NO_2 influence O_3 levels, and this in turn can influence NO_2 both locally and downwind of the grid-cells adjusted by the assimilation.

38. *Fig 6: Is this consistent with fig 3? There seem to be large increases also in spring-summer in the data assimilation. The simulation in Fig 3 seems to suggest that the bias in spring-summer is small.*

    Given that averaging kernels were not used in the work reported in the first version of the manuscript, we realise that the results presented in Figs. 4-6 (showing maps of OMI retrievals and DEHM column concentrations) are not valid. Hence it is hard to say whether the results in Fig. 6 are consistent with those in Fig. 3.

39. *Fig 7: Could be more clear if the horizontal scale is blown up (lower stratosphere is not so interesting for this paper) so that lines in the lower troposphere are more clearly separated.*

    As mentioned above, the comparison with ozone sonde data has not been included in the revised manuscript. However modelled NO_2 and O_3 columns are presented, and a logarithmic scale is used to on both the x- and the y-axes, thus emphasising differences in the lower troposphere, as suggested by the reviewer.

40. *Fig 8: Is the relative difference consistent with the lines? The two model simulations seem to overlay almost perfectly, while I would have expected a significant difference?*

    The "relative difference" refers to the magnitude of the difference *relative* to the magnitude of the difference in the initial conditions. As can be seen in the time-series of average concentrations, the differences between the different simulations during the summer months are relatively, hence the differences present in the initial conditions were rather small. In hindsight, it would have been more informative to initialise the "perturbed" simulation with a larger perturbation, such as the differences that might be present during the more variable winter-time concentrations. Figure 8 of the initial version of the manuscript has not been included in the revised version.

Interactive comment on Geosci. Model Dev. Discuss., 5, 309, 2012.