**Atmospheric
Chemistry
and Physics
Discussions**

Interactive comment on "Evaluation of a regional air quality forecast model for tropospheric NO$_2$ columns using the OMI/AURA satellite tropospheric NO$_2$ product" *by* F. L. Herron-Thorpe et al.

Anonymous Referee #1

Received and published: 14 February 2010

The paper discusses a comparison between NO2 column measurements made by the AURA-OMI instrument and AIRPACT air quality model results. A reasonable agreement is found in winter months. In summer an interesting large difference is found related to wildfires, where the model produces much higher NO2 values than OMI.

To my opinion the paper in it's present form is not ready for publication, and a major revision is needed. The main reasons for this judgement are the following:

1. The authors should be more careful in the formulation of the method and results.

C10831

My (many) detailed comments given below illustrate this point. Discussions are often unclear, not well formulated and sometimes not based on evidence.

2. The conclusion from comparisons with and without kernels leaves the reader confused. Is it important to account for the kernels? What are the quantitative errors if the kernels are not used? The discussion is mentioning several issues (computationally expensive, unrealistic high values) which seem to be quite specific for the kernel implementation used by the authors and can not be generalised to other NO2 comparisons. Kernels are applied only for one month, and not for the full 18 months. The abstract mentions: "applying the averaging kernel in cloud free conditions has little effect". but fig 3 suggests that the a-priori profile has large limitations, which should influence the cloud-free results. How do these findings match?

3. The OMI product from KNMI also contains kernels. It is a pity that these kernels have not been used. For instance it would be interesting to compare the kernels from the two products. Although many results are also shown for the KNMI product, a description of the two algorithms is not provided in a balanced way. It is not clear what the main differences are between the two algorithms.

4. The discrepancies observed during wildfires are interesting, but are not really discussed. What are the problems in modelling these fire emissions and what is the error estimate on this? Is the model overestimating? The new experiences with "BlueSky" may help to illustrate model uncertainties. What are the issues in the OMI retrievals which may result in a large underestimate of the NO2 from fires? What is the most likely cause for the discrepancy (OMI or model)?

5. There is too much discussion about emission estimation. On this topic there are no results presented, and therefore only a brief mention of the inversion plans (and literature on inversions) is justified. The introduction and the end of the conclusions section would be a good place for this. The authors seem to be too optimistic about the inversions: from the uncertainties, the wildfire issues and differences between the

retrievals it is not clear to me that one should even attempt to use OMI for emissions estimates.

Detailed comments:

"SCHIAMACY" should be SCIAMACHY (in references and introduction)

p. 66, l15: "but unfortunately the NRT product does not represent the best calculation provided in the official data collection". Please explain why the NRT product is less accurate.

p 67, l13: "to the height of stratosphere-troposphere exchange". Is it meant in this way? I would expect "to the height of the tropopause".

p 68, l20: "require the use of radiative transfer models and a geo-referencing scheme to identify areas where there is commonly pollution in the boundary layer." Radiative transfer models are needed in most remote sensing retrievals (rather trivial statement). The geo-referencing refers to the masking applied to estimate the stratosphere, I guess. But this depends on the approach and is not required. For instance a reference sector (over pacific) approach has been widely used which I would not classify as geo-referencing (it is a more simplified assumption).

Please motivate the choice for < 35% cloud cover ? Other choices have been made by other groups. Are the results in the paper sensitive to the exact limit for cloud cover?

p 69, l9: "for pixels not dominated by point sources". Why is the analysis not relevant for point sources? Clearly the footprint size of the satellite should be accounted for, but when the appropriate averaging is applied I do not see a problem.

p 69, l10: "Our overall goal is to make future NOx emissions adjustments based on previous OMI/AIRPACT comparisons". This is not the subject of this paper, so irrelevant for the discussion. I suggest to remove this line.

p69, l22: "(http://www.doas-bremen.de/doas glossary.htm)". It is uncommon to refer to

such websites.

p69, l23: "An 18 month trend analysis". "Trend" is a word commonly used for changes over time, for periods > 20 years. The use of the word "trend" is confusing.

p70, eq 1: Why is this called "density". It is the total amount of molecules in the layer per 1cm by 1cm surface area.

p70, eq1-2-3-4: These equations are standard, and almost trivial. It would suggest to remove these equations and to simply mention that the layer concentrations have been converted to layer column amounts, in molecules/cm2.

p70, eq 1-4: It is a bit surprising that the density is calculated by the ideal gas law, eq 2. It would be more accurate to use a total air mass per model grid box. Is this not available in CMAQ/MCIP ?

p71, sec 2.2: Since two OMI products are compared in this paper, it would be useful to the reader to provide a summary of the main differences between the two retrievals (and possible impacts of those differences if known). Section 2.2 only describes the NASA approach, and does not discuss the KNMI approach.

p71, l6: I believe that these details on the NASA-OMI algorithm are documented. So why do you refer to "Eric Bucsela, private communication"? It is better to cite a refereed paper, in this case Bucsela 2006.

p71, l14: "where corresponding points from AIRPACT are masked from averaging as well.": How is this done? Is this based on spatial collocation, or on the cloud fraction in the AIRPACT model system? Is it done for AIRPACT or for the spatially averaged AIRPACT results?

p71, l15-22: "This can be a relevant source of error when comparing model results to satellite derived columns." This paragraph discusses meteorological variables and their impact on the OMI retrieval, but the conclusion is vague. Are there papers that discuss this effect ? Is number density really needed in the retrievals, and, if so, is it a

large source of error (provide order of magnitude)? If the authors can not answer these questions it is better to remove this paragraph.

p71, sec 2.3: NASA averaging kernels are used. The KNMI product also includes averaging kernels. Is there a reason for not using those as well? It would be interesting to see a comparison between the two kernels.

Fig.2: Please make a clear distiction which boxes have been computed by the authors, and which boxes are computed by NASA (KNMI) as part of the retrieval process. In the averaging kernel computation there are quantities dAMFCld and dAMFClr. Where do those come from? The text mentions "IDL routines and lookup tables". What do these routines require as input, and what do they provide as output? Please provide such links between the main text and the boxes in Fig.2. The authors use quite some text and a figure to explain the processing, but the actual computing steps remain unclear.

p72, l8: "However, the latest version of KNMI data includes the averaging kernels in the daily level 2 data files." Why have not these kernels been used? Apparently these kernels are provided as a matrix, and it would be numerically cheap to apply them!

p72, l15: "used by OMI algorithms". This suggests that both the NASA and KNMI retrievals use GEOS-CHEM ?

p73, l12: "Despite the numerical intensity involved with applying the NASA averaging kernel, spatially averaging the model results to the daily OMI swath requires only a simple additional function in scripting as compared to independent comparisons." This is an incoherent sentence: what is the link between the first and second part? Perhaps this line can be removed.

p73, l15: "It is a useful and efficient method to adjust model results based on the variance in OMI footprint size throughout the swath." I do not understand this remark. Please explain (or remove).

p74: Is the kernel applied to the original AIRPACT or to the averaged AIRPACT data?

C10835

Please provide details how correlations are computed: is it a spatial correlation for the monthly-mean AIRPACT and OMI distributions? What formula is used? Computing correlations is standard, but nevertheless in my experience different expressions are used by different groups (differing especially in the calculation of the means/references).

p74, Fig 4: It would be nice to include a panel showing the lat-lon distribution of the kernel-applied columns. This will help to appreciate the impact of the application of the kernel.

p75, l2: "This sometimes leads to very large values if the product of the layer mixing ratio and pressure equals the layer above it." I do not understand how this can be a problem?! Very large values of what? The averaging kernel?

p75, l9: "Overall, the same general trends are found in the modeled columns". Please be more precise. What is meant by "trend" here? As mentioned before, the word "trend" is used in several places and is confusing to me (apparently has a different meaning to the authors).

Fig. 7: Concerning the model range (min-max values): please specify if this is calculated for the non-averaged, averaged or kernel-applied model values.

p75, l25: "In general, the interpretation of bias trends changes much more when we spatially average AIRPACT results, as opposed to applying the averaging kernel and masking the erroneous data." I do not understand this statement. Please reformulate.

p76, l7: " the spatially averaged AIRPACT values give a better representation of what should be directly compared to OMI". This is a statement which is out of place. Spatial (horizontal) averaging is unrelated to the kernel issue (which is related to the vertical profile), and the neglect of the kernels can not be compensated by spatial averaging, as this sentence seems to suggest. It is also not fully clear to me if the spatial averaging is really a better representation of OMI. (The figures seem to suggest it is, but there

C10836

may be other reasons for a better comparison)

p76, l20: teem -> team

p76, l20: It is well known that instruments like OMI are able to observe biomass burning signals. For instance the African biomass burning and it's seasonality is well observed and in reasonable agreement with models. Why should it be different for these wild fires? Again, referring to a discussion at the science meeting is not convincing: is there evidence in the literature to confirm this statement that fires can not be observed by OMI.

p77, l7: "This is in contrast to the slow periodicity of high NO2 values retrieved by OMI". What do you mean by "slow periodicity"?

p77, l 22: "July through December shows much higher OMI". It is useful to mention the difference between the NASA and KNMI product in this case.

p78: It is mentioned that the KNMI product has negative values and suffers from "stripes". I was a bit surprised that these effects are not smoothed out in the monthly averages, and have such an impact on the correlation (which are spatial correlations computed for the monthly mean field if I understand correctly).

p78, l17: "current collection 3 of OMI tropospheric NO2 provided by NASA seems to cause a systematic trend of higher values in the summer." Is this referring to e.g. Fig. 6, or is this taken from the literature (if so, provide a reference)?

p78, l22: "However, all long-term NASA timelines show a clear anti-correlation with season." Same comment, please provide reference if available.

p79, l7: Same comment: Are these conclusions drawn from this work. If not, is there a reference? (Where do these conclusions come from?)

p79, l16: "However, this effect may be expected naturally: less sunlight is incident on the airshed during cloud covered cool months, so less NO2 is photolyzed to NO." This

C10837

is too simple. The lifetime of NOx, conversion to NOy species and removal through e.g. HNO3 play a crucial role in the difference between summer and winter.

p80, l1: "wildfire emissions would require a 4-D-var analysis using near real-time NO2 retrievals". Why is near-real time needed? Why is 4D-Var needed? When the fire location and timing is known, it seems that a scaling of the source strength can be applied to match OMI as a cheap way to do the inversion. However, from the earlier discussions on the fire comparisons it seems to me that the biggest problem is the characterisation/understanding of the OMI measurements: which fraction of the fire-produced NOx is actually observable by OMI (given the smoke and clouds).

p80, l3: "DOAS satellite retrievals do not readily resolve boundary layer concentrations". I do not understand this remark which contradicts the results presented. My impression is that most of the features observed in e.g. figures 4 and 5 are to be interpreted as originating from the boundary-layer. At least the city hotspots should be interpreted as NO2 in the boundary layer.

p80, l15: "Ultimately, we have decided it would be best to make adjustments based on the average of the two data sets." This choice is completely arbitrary it seems ...

p80, l22: " long term trends". The word "trend" has been used several times in the paper and is confusing. What is "long term" in this case: 18 months, 5 years, 1 month ?

p81, l3: "computationally expensive". This is not a general problem, but is quite specific for the way the authors derive the kernel for the NASA product it seems.

p81, l3: "erroneously large". I guess the OMI NO2 column cannot be trusted as well in these cases ?!

p81, l9: "significant number of trends". Please use other words, e.g. "conclusions"

p81, l23: "BlueSky". Are there already indications that the quantitative emissions are substantially smaller in this model ?

C10838

C10839