Atmos. Chem. Phys. Discuss., 13, C4439–C4443, 2013
www.atmos-chem-phys-discuss.net/13/C4439/2013/

Atmospheric Chemistry and Physics Discussions

Interactive comment on "An evaluation of the CMAQ reproducibility of satellite tropospheric NO_2 column observations at different local times over East Asia" *by* H. Irie et al.

Anonymous Referee #2

Received and published: 3 July 2013

This paper describes a model evaluation of CMAQ against NO2 tropospheric columns derived from 3 satellite instruments, looking at the sensitivity of model NO2 columns to different parameters, including model resolution and emissions. The paper builds upon a previous work (Irie et al., AMT 2012), where they concluded that biases between satellite retrievals are insignificant (<10%), and allow for analysis combining these satellite data. The current paper illustrates that the model is generally able to capture NO2 during the July period, but not in December. Also discrepancies between SCIAMACHY and GOME-2 retrievals are observed, when averaged over larger regions (e.g. CEC in July), which tend to be significantly larger and of different sign than the

earlier reported biases (Irie et al., 2012). These large biases are to some respect worrying, and seem to limit the validity of the remaining study, i.e. the evaluation of the model diurnal variation. The authors also end concluding that even though suggested by Irie et al. (2012), additional validation of the satellite retrievals is required before they can be used for evaluation of the diurnal variation.

While the paper is compact, and to the point, and interesting in its attempt to confront model simulations against different retrieval products in a consistent manner, at its current stage it misses interpretation of the various biases, which leaves the reader largely unsatisfied. In the current shape it is merely an account of the various retrievals, and the corresponding sensitivity runs. Hence I believe that this study should be seriously rewritten to provide such additional information, before it can be accepted for publication.

General comments:

P14042, l8-10: Your selection criterion for the sensitivity experiments seems ad-hoc and limited. It would be good to put this study better into perspective of other sensitivity studies using retrievals, such as Lin et al. (2012) and others.

P14044, l26: "The largest difference of NO2 VCDs with respect to the value at a cloud fraction of 20 % is found to be < 30 %, which is much smaller than the quoted uncertainty in the satellite retrievals". I understood from P14043, l12 that uncertainty in satellite retrievals is actually ∼30%, so in the order of magnitude of the satellite retrieval uncertainty. Please explain. Furthermore, here you provide statistics for the largest region (CEC), where one may expect some smoothing from the regional averaging. I wonder to what extent numbers change if smaller regions with large pollution sources, like 'BEI', are selected. Such sites may additionally contain a considerably less amount of individual observations due to combined aerosol / cloud shielding. The question is whether these (small) regions contain sufficient pixels with observations to be able to provide quantitative statements.

ACPD
13, C4439–C4443, 2013

P14045,l10: "The associated error bars represent simple averages of quoted uncertainties in the satellite swath data used for monthly-mean calculations.": It appears from the error bars in the figures that the uncertainty for GOME and SCIAMACHY observations is ∼60%, rather than 30%, as discussed in the introduction. Could you give more information on these numbers and provide actual uncertainty estimates for the various instruments?

P14045, l28:"the difference is likely insignificant": As you probably rightly acknowledge that the difference between GOME and SCIAMACHY is insignificant, this also suggests that the difference between NO2 VCD's from those instruments and OMI is insignificant. Hence it is very hard to make quantitative statements on the diurnal variation. In fact, why not validate the models directly using MAX-DOAS if you are interested in diurnal cycles. Can you comment?

P14047, l5: You conclude that diurnal cycle in emissions do not always produce better agreement with measurements. Could you give some more comments on this? I would believe that including a better representation of NOx emissions would be rather important to get better match to the retrievals. But from your analysis this seems not the case, or hardly anything changes.

P14048, l2: "are reproduced well for all 12 regions": This seems not true: E.g. for regions ECS, SOJ and PRD, where diurnal variation is quite different. It might be better to introduce the quantitative evaluation (Tables 4/5) in the respective section (i.e., here) to diagnose the diurnal variation.

P14048, l13: "larger" shouldn't this be "smaller"?

P14049, l13-14: "negatives": The fact that there are 'some negatives' is not sufficient to suggest that there is an issue with the retrieval algorithm, and should be removed. Individual pixels are allowed to give negative values, to compensate for pixels with too positive values. This is simply due to the natural scatter in the observations, and relate to the observational uncertainty. At this stage one might equally well conclude that the

ACPD
13, C4439–C4443, 2013

model has problems with getting the NO2 right in winter. Please correct.

P14049, l20: "It is thought. . .": It would be good to get more indications of what changes in CMAQ chemistry when increasing the resolution. Now things are unclear and do not add to the understanding. E.g., do you see more O3 production? Do you better capture spatial variation in observed NO2 (see comment on correlation statistics)? Again, why not directly compare to MAX-DOAS observations to obtain a clear evaluation of the diurnal cycle, and impact of increased resolution?

P14050, l9: If the authors conclude that satellite observations are insufficiently constrained to be used for quantitative statements on the diurnal variation then it is questionable how we should interpret the current results, and what we can learn from this. Please comment.

P14050, l20: "Quantitative agreement . . . are taken into account": what do you mean with this sentence? In what sense did you take model uncertainties into account for the quantitative agreement? It is also unclear which simulation is performing best or whether you find suggestions for possible model biases or biases in the satellite retrievals.

It is disappointing that no reasons for the discrepancies are identified. Would you suggest that there are biases in the emissions? Biases in the retrieval? Biases in CMAQ? By just reporting the discrepancies it is difficult to learn something from this. How do other models perform? Is it a general feature? Could there be issues with representativity of the results due to small sampling areas combined with short time periods for analysis? With so many open questions it is difficult to accept this manuscript for publication.

For example, it would be good to see some more basic statistics, including maps of satellite NO2 from the 3 products for June and December, and corresponding numbers for the spatial correlation of the monthly mean fields, in order to assess how this changes (improves?) with increased resolution and for different instruments, having all

ACPD
13, C4439–C4443, 2013

amongst others different pixel sizes. It could give an indication of whether the three instruments see the same features.

References:

Lin et al., Modeling uncertainties for tropospheric nitrogen dioxide columns affecting satellite-based inverse modeling of nitrogen oxides emissions, Atmos. Chem. Phys., 12, 12255–12275, 2012

Interactive comment on Atmos. Chem. Phys. Discuss., 13, 14037, 2013.