Interactive comment on “Evaluating the accuracy of NO\textsubscript{x} emission fluxes over East Asia by comparison between CMAQ-simulated and OMI-retrieved NO\textsubscript{2} columns with the application of averaging kernels from the KNMI algorithm” by K. M. Han et al.

Anonymous Referee #3

Received and published: 8 August 2014

The authors present two main topics: (1) the necessity of using averaging kernels when comparing model simulations of NO\textsubscript{2} tropospheric columns with satellite retrievals. (2) the sensitivity of the modeled NO\textsubscript{2} columns to simulation parameters (seasonal cycle; alternative emission inventory; reaction rate N2O5). In its current form I do not consider it fit for publication in ACP. The paper needs serious restructuring to better present its scientific relevance. The revision should be more concise.
GENERAL COMMENTS

I miss a well described motivation for the presented study. Does it concentrate on the importance of the use of averaging kernels in comparison studies (which is obvious for the satellite community, but apparently less obvious for the modelling community)? Does it want to correct previous work (e.g. by Han et al., 2011) which did not take into account the AKs? Does it want to show that emission inventories in East Asia are wrong or out-dated? Does it want to contribute to other sensitivity studies analyzing the model error in reproducing NO2 columns (e.g. Lin et al., 2012)? Does it want to do so to improve future top-down emission estimates with satellite observations? These motivations are all hidden in the text, but should be stated more clearly. The addressed scientific questions should dominate the structure of this paper and its analysis.

Satellite retrievals are not the truth, and can also be biased. So the cause of differences between model and observation can be found in both. Uncertainties in satellite retrievals are first mentioned in the end Section 3.2.4, but should be given a more prominent place in the analysis of the results.

The method of applying the averaging kernels on the model simulations is only briefly explained at the end of section 2.2; Figure 4 is hardly discussed. What I understand is that the OMI observations are horizontally gridded to the model grid. Why not doing the opposite: interpolating the model values to each satellite footprint? This makes a fairer comparison as the averaging kernel is associated with the footprint area of the observation, and not with a model grid cell.

SPECIFIC COMMENTS

Title: The title represents only one side of the study, and neglects the performed sensitivity analysis.

Abstract, P17587, line 1: Please mention which inventories are used for the evaluation.
Abstract, P17587, line 1: Some indication of the geographical extent of the used East
Asian domain would be nice.

Abstract, P17587, line 13: “28%”: using which emission inventory? Can the difference be attributed to wrong emissions?

1, P17588, line 14: “Han and Song, 2012” were not the first ones to find out about the importance of this removal process. Maybe an earlier reference is more appropriate.

1, P17588, line 4-6: To compensate for height dependent sensitivities, the column retrieval algorithm depends on cloud information and an assumed NO2 profile. If this NO2 profile reflects reality, the retrievals can be compared directly with simulations (and are in that sense “real” or “true”). If the true NO2 profile is different, the averaging kernel of the retrieval method should be applied to the simulation to compensate for this effect.

1, P17589, line 9. Consider a definition $\Omega$ instead of $\Omega_{NO2}$ throughout the whole paper, as NO2 columns are the only columns studied in this paper. This will increase the readability of the symbolized quantities used in the text.

2, P17590, line 1-5: Leave out. This should be clear by now.

2.1, P17590, line 8-9: “because relatively (…) this year”. Better: because INTEX-B was compiled for this year.

2.1, P17592, line 1-2: “modeling conditions” $\rightarrow$ model setup

2.1, P17593, line 20: Change $x \bar{I}C-x_a$ to $x \bar{I}C-x \hat{I}C_a$, to differentiate $x \hat{I}C_a$ being a column quantity and $x_a$ a vector quantity.

2.1, P17594, line 10: The change of the AKs over the seasons can also be related to cloud climatology, especially because it is not clear from the text that the observations in Figure 3 have been filtered for cloud radiance.

3.1.1, P17595, line 3. Consider writing the section title in words instead of symbols.
3.1.1, P17595, line 20-21: “possibly” can be left out. “Han et al., 2009” were not the first ones to find out about this removal process. Maybe an earlier reference is more appropriate.

3.1.1, P17596, line 6-11: Apparently this an important motivation to conduct this study. Therefore, it should be given a more prominent position, for instance in the Abstract or Introduction.

3.1.1, P17596, line 13-14: “(... correct previous conclusions”. This conclusion should therefore also be mentioned in the Conclusion section.

3.1.1, P17596, line 28: Mention that the NME is defined in Table A1.

3.1.2, P17597, line 10-11, Figure 7: I do not understand why I see in each panel so few scatter points. With a model resolution of 30 km2 and comparable OMI footprint resolution each focus region contains dozens of grid cells / observations, which sum up in a three-month period to hundreds of data pairs. Please explain in more detail how a data pair is established.

3.1.2, P17597, line 20-26: A lot of different statistical quantities are introduced here; not all of them are familiar to everyone. Different quantities highlight a different aspect of how model and observation compare. Maybe it is an idea to describe in Table A1 (or elsewhere in the Appendix) for each quantity its specific use in comparing model with observation.

3.1.2, P17598, line 6: “between 2 x 1015” → “between -2 x 1015”

3.1.2, P17598, line 6-7: My main interpretation of the MB results would be that for all seasons the mean bias is negative almost everywhere (except for CEC2), i.e. CMAQ,AK is smaller than OMI, a strong indication that the used NOx emission inventory is underestimating the real emissions.

3.1.2, P17598, line 17-19: In the abstract and conclusion the underestimation of NOx emissions is estimated to be around 28%. This is the section about statistical analysis.
so here it should give more detailed information on how this number is derived.

3.2: This is definitely not the first sensitivity analysis. Previous work, such as by Lin et al. (2012) should be properly discussed. May be parts of 3.2.4 can be included in such an overview. The choice of why investigating the sensitivity to parameters in Case 2,3,4 should be clearly explained. And finally, how do the sensitivity results compare or add to existing results?

3.2.1, P17599, line 3-4: It is unclear if the imposed seasonal variation is taken the same for each emission sector.

3.2.1, P17599, line 5-7: The larger difference found in winter time could also indicate a NOx lifetime issue of the model in colder/darker environments.

3.2.1, P17599, line 18-19, Figure 1: Indicate more clearly that the monthly variation of INTEX-B is taken from Zhang et al. (2009).

3.2.2, P17600, line 2: Which version of the REAS inventory is used, for which base year?

3.2.2, P17601, line 6-7: Some words about satellite derived emission inventories seem appropriate here.

3.2.3, P17601, line 12-14: Not necessarily true. In winter time the increased lifetime transports NOx further away from its source. This make you more sensitive to the correctness of the meteorological fields (e.g. winds).

3.2.4, P17603, line 8-12: “Although not shown (…) should be investigated further”. Leave out.

3.2.4, P17603, line 13-16: “it can be suggested”, “will/may be able to help”. This can be stated stronger.

3.2.4, P17603, line 26: How do k1, k2, k3, and k4 relate in magnitude? This gives information about to which mixing ratio the balance between NO2 and NO is especially

C5717
sensitive.

3.2.4, P17604, line 15-16: “The uncertainties (…) to some degree”. Vague. Clarify or leave out.

3.2.4, P17604, line 28-29: “This is why we said that the summer was not a season of major interest in this study.” Similar statements have been at earlier points in the text. Why not centralize them (e.g. in the Introduction or in Section 2), and list all considerations to focus only on the winter months?

3.2.4, P17605, line 1-12: This paragraph on retrieval uncertainties could be moved forward to Section 2.2 where the satellite product is first discussed. Maybe include some description of the retrieval error to better interpret the statistical study in Section 3.1.2: tropospheric column retrievals typically have a dominant absolute error (∼0.5 \(10^{15}\) molec/cm\(^2\)) at low values, and have a dominant relative error (30-40%) at high values.

4, P17607, line 3-15: In my opinion, this is not a conclusion of the presented study. Instead it is an important motivation to do the sensitivity analysis in Section 3, where this text could be included in the introduction. Sensitivity studies as presented in this study improve the model (or at least improve the understanding of the model error and bias) to reproduce NO2 columns. This is very important to improve the accuracy of top-down emission estimates made with satellite observations.

Table 1, P17618: Indicate reference year 2006

Figure 5 and Figure 6: Consider merging the two figures in a 4 x 5 panel

Figure 7: Indication of units

Figure 8: Use a neutral (white) color for a value range around 0.