Atmos. Meas. Tech. Discuss., 5, C2280–C2291, 2012
www.atmos-meas-tech-discuss.net/5/C2280/2012/

Atmospheric Measurement Techniques Discussions

Interactive comment on "First quantitative bias estimates for tropospheric NO_2 columns retrieved from SCIAMACHY, OMI, and GOME-2 using a common standard" *by* H. Irie et al.

H. Irie et al.

hitoshi.irie@chiba-u.jp

Received and published: 19 September 2012

Reply to anonymous referee 1

We thank the reviewer very much for reading our paper carefully and giving us valuable comments. Detailed responses to the comments are given below.

General comments:

1. The authors gave a lot of arbitrary values (0.05, 0.10, 0.15, 0.20, 0.25, 0.30, 0.35, 0.40, 0.45, 0.50, 0.60, 0.70, 0.80, 0.90, and 1.00) of latitude/longitude grid size as *coincidence criteria in the regression analyses between satellite and ground measurements. After the analyses, they finally found the difference of the behaviors of slopes and correlation coefficients between the China and Tokyo cases, and then they attributed the difference to the spatial inhomogeneity of the tropospheric NO2, referring to the averaged distribution shown in Figure 1. However, the authors would have immediately recognized such a spatial inhomogeneity exists in Figure 1 without conducting the regression analyses, and without giving the arbitrary values they could have directly estimated the representative spatial scales of the tropospheric NO2 inhomogeneity in each region. The authors should estimate the typical spatial scales for each region before the regression analyses. They can directly find the values for the criterion x to evaluate the statistics they need.*

Reply: Following the reviewer's comments, we have revised the manuscript to include discussions about spatial scales for each region without using regression analyses or referring to Figure 1. For this, we conducted a detailed analysis of satellite data only and two new figures (Figures 6 and 7 of the revised manuscript) relevant to this additional analysis have been added accordingly. It turned out that the unique determination of typical spatial scales, which the reviewer is thinking of, was difficult at least partly because sufficient satellite data are unavailable on spatial scales that we are testing (particularly at small x values). Also the spatial scale varies with time significantly (Figure 7 of the revised manuscript). So, we think that testing various x values (as coincidence criteria or spatial scales) would still be worthwhile for better bias estimates.

Indeed, there is an inconsistent description which suggests that the authors should rearrange the construction of the manuscript wholly; at lines 23-25 on page 3960, the authors say: "These suggest that the spatial distributions of tropospheric NO2 VCDs around the Chinese sites during the observation periods were rather homogeneous and therefore appropriate for bias estimates". This sentence means that the authors had recognized the importance of the homogeneity of the tropospheric NO2 before the

AMTD

5, C2280–C2291, 2012

Interactive Comment

regression analysis. If it is true, however, the authors should look for such a homogeneous condition first of all for the best bias estimations. They had never examined the inhomogeneity of the spatial distribution of the tropospheric NO2 before the regression analysis. It looks very strange.

Reply: According to the comments, the revised manuscript now includes discussions about the inhomogeneity of the spatial distribution for each region without using regression analyses, as mentioned above.

In the manuscript, the authors apparently discarded the results of the Tokyo case for the final bias estimation and adopted only the results for the China case. However, if the authors estimate the representative spatial scale of the tropospheric NO2 for each case (China and Tokyo) and use the scale as the criterion x in the regression analysis, the authors can estimate the biases not only for the China case but also for the Tokyo case. This will be a large advantage of the major revision.

Reply: We also hoped that we could make use of data from the Tokyo case. As mentioned above, we tried to uniquely estimate the representative spatial scale of the tropospheric NO2 but found it very difficult. In particular, for the Tokyo case, it is most likely that the representative spatial scale is smaller than our smallest latitude/longitude grid size ($\sim$0.05°) or the OMI pixel size. We realize, however, that completely ignoring the Tokyo case was not a good idea, as pointed out by the reviewer. So, the revised manuscript now includes the results from the Tokyo case in the discussion of the bias estimates, where the Tokyo case is used to support the results for the China case.

2. The authors focus largely on the effect of the spatial inhomogeneity of the tropospheric NO2 on the regression analyses (e.g., Figures 4 and 5). However, the manuscript does not have any description on the spatial structures which should exist in the products of SCIAMACHY, OMI and GOME-2 or any comparison among those products. Figure 1 could be a part of such descriptions, but it is only for GOME-2. In the introduction, the authors describe that OMI has an equator crossing time different *from other two sensors. Is there any possibility that the NO2 spatial structures derived from OMI are different from others? If so, the authors should use a different criterion for OMI. They should examine whether there is no difference of spatial scale among the data set of SCIAMACHY, OMI and GOME-2.*

Reply: Following these comments, we have analyzed the satellite data and found that the NO2 spatial structures derived from OMI are very similar to those from SCIAMACHY and GOME-2 (see Figure 6 of the revised manuscript), particularly around Tsukuba, Yokosuka, and Hedo, where relatively large amounts of satellite data in the MAX-DOAS observation time period were available. The NO2 spatial structures derived from OMI could be different than SCIAMACHY and GOME-2 occasionally around the other sites. It seems difficult to quantify this with a limited number of satellite data in the short time period of the MAX-DOAS observations there. So, we have decided to use the same criterion for all three sensors. Discussions of the spatial structures are now made using Figure 6 in the revised manuscript.

3. The authors conclude that the difference of Figures 4 and 5 is caused by the spatial inhomogeneity. However, the length of the time period of the data obtained in the China case is totally different; it is much shorter than the one in Tokyo. Is there any possibility that the difference affects the results? The authors should discuss the effect of the difference of the sampling periods on the regression analyses.

Reply: For the Tokyo case, we made similar regression analyses only for the season, when the China case data were available. This resulted in similar numbers of comparison pairs with those of the China case. We found that the spatial inhomogeneity effects reduced the slope of regression line similarly for both the limited and entire time periods. This is now stated in the revised manuscript. Also, revision has been made to include additional discussion of the spatial inhomogeneity with Figure 7 (a new figure in the revised manuscript).

Specific comments

-Title

The authors use the word "first" in the title, but I would say this is an exaggerated expression. The previous studies have already performed various validations for the satellite-based tropospheric NO2 data which are used in this manuscript (e.g., Bucsela et al. (2008) for OMI), and even one of the authors wrote a paper on similar analysis as well [Irie et al., 2009]. Indeed, the authors would insist that the data set of the combinations of SCIAMACHY, OMI and GOME-2 have never experienced the bias estimations together, but there could be many combinations of sensors. For example, the first bias estimation for GOME, SCIAMACHY and OMI could be possible. I do not think this is fair. The authors should include some phrase like "east Asia" in the title, because they utilize the data only for China and Japan. This is also one of the reasons why I think the present title is exaggerated. The authors mention the effect of the spatial inhomogeneity of the tropospheric NO2 on the validation of satellite data. The authors could include some phrases related to the spatial inhomogeneity into the title.

Reply: Following the reviewer's suggestions, the title has been changed to "Quantitative bias estimates for tropospheric NO2 columns retrieved from SCIAMACHY, OMI, and GOME-2 using a common standard for East Asia." We tried to include "the effect of the spatial inhomogeneity of tropospheric NO2 on the validation of satellite data" but we now think that this is too detailed and the sentiment is already included implicitly in the phrase "using a common standard".

-Section 2

Page 3957, lines 8-9: Are the models (and any inputs to them) which calculate the AMF exactly same among all the products? If there are differences of them, AMF could be different, and it should cause the difference of the NO2 amounts even though the real NO2 amounts are exactly identical.

Reply: For all the sites, we use the same radiative transfer model but different inputs. The inputs differing among the sites are the surface elevation and the altitude at which

AMTD

5, C2280–C2291, 2012

Interactive Comment

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

the telescope is located. This is for retrieving NO2 amounts at the same quality for different sites, so the effect should not be as important as the reviewer thinks.

-Section 3

Page 3958, lines 24-25: In the MAX-DOAS retrieval, did the authors use the same cross sections of NO2 as was used in the satellite measurements? If they used the cross sections at the different temperature and/or pressure, the VCD retrieved would be affected.

Reply: We agree with the reviewer. Its potential influence has been stated in the text as part of the systematic error in the MAX-DOAS NO2 retrieval.

-Section 4

Page 3959, line 8: Why do the authors take the value "0.20" for the criterion in Figure 2? Show the reason.

Reply: Figure 2 is an example of our correlation analyses and basically we can take any value. The smaller value is better as a stricter coincidence criterion is used, but at the same time the statistical significance drops. Considering the two effects, we have chosen the value of 0.20 as an optimal number for the readers to understand our correlation analyses.

Does this criterion mean a square of 0.20° of latitude and longitude? I do not think this is appropriate, and the authors should take a circle instead of a square, e.g., a circle with the radius of the criterion x.

Reply: Yes, it means a square of 0.20° of latitude and longitude. We choose this approach, because the instantaneous field of view of the satellite observations is rectangular. We realize that a circle can be another option and hope to investigate this in future studies.

Page 3959, lines 11 and 12: It seems strange that the site Hedo is included in the both *cases of China and Tokyo.*

Reply: Our target is to quantify the bias over East Asia (as the reviewer pointed out above) and not to quantify the biases separately for the China and Tokyo cases. For both cases, Hedo data play a critical role in drawing a regression line with an intercept close to zero, as discussed in the manuscript.

Page 3959, line 16: "as their data are distributed over a wide range of NO2 values" Which does "their data" mean? Is the data for China or Tokyo? When I first read the manuscript, I thought it means the data for China. If so, the phrase shown above looks inconsistent with Figure 2; the red points, which indicates the data for China, have a much smaller range of NO2 than for other groups (e.g., Tsukuba).

Reply: It means the data for all sites except Hedo. We want to say that the range of the NO2 values is wide compared to that of Hedo. This sentence has been revised accordingly.

Page 3959, line 18: A small magnitude could be just a noise component. Therefore, the authors cannot say that the result of the measurements is "reasonable" only by the fact that a small magnitude is observed in remote areas. The values of OMI's VCD in the site Hedo are just 2-3x10^{15} molecules /cm^{2} shown in Figure 2. The values are close to the error of the OMI data (1x10^{15}, which is shown at line 10 on page 3957).

Reply: We understand the reviewer. We have revised the sentence to "For comparisons over Hedo (shown in green), which is located in a remote area, both satellite and MAX-DOAS data consistently show very small NO2 VCD values, compared to the other sites."

Page 3959, line 19: What are the same features? Does this mean a qualitative agreement between satellite data and MAX-DOAS data?

Reply: It means that both satellite and MAX-DOAS data show very small NO2 VCD values, compared to the other sites. The sentence has been revised to state this more

AMTD

5, C2280–C2291, 2012

clearly.

Page 3959, line 20: What is "this"?

Reply: It means that both satellite and MAX-DOAS data show very small NO2 VCD values, compared to the other sites, for any cases we investigated. This has been made clearer in the revised manuscript.

Page 3959, line 20: The authors force the intercepts to be zero without any quantitative discussion or justification, but just for simplicity. This manuscript, however, aims to evaluate the biases quantitatively. I would show the values of the intercepts even though the values were small. Of course the authors can finally neglect the intercepts if the values are really small, but the values themselves should be shown.

Reply: According to the reviewer's comment, we have added a quantitative discussion about the intercept in the revised manuscript.

Page 3959, line 23: I do not think the China case is really "excellent", as it still includes some deviations from the line of unity. The authors would compare it with the Tokyo case. Apparently, the deviation from unity in the Tokyo case is much larger than in the China case. However, the number of the samples for the China case is much smaller than that of the Tokyo case. The simple comparison is not fair.

Reply: We have removed the word "excellent" from this sentence.

Page 3959, lines 23-25: How did the authors calculate the standard deviation of the slopes? Did you calculate several slopes and then calculate the standard deviation of them?

Reply: Many statistical software packages and some graphing calculators now provide the standard error of the slope as a regression analysis output. We used the IGOR Pro software package, which calculates the standard deviation based on the Numerial Recipes in C (http://www.haoli.org/nr/; see the Chapter 15.2 "Fitting Data to a Straight line" for more details).

AMTD
5, C2280–C2291, 2012

Page 3960, line 1: I just want to confirm what "the difference" means. Does it mean the difference between the slopes of China and Tokyo case? (I suppose it in the next comment.)

Reply: Yes, it does. This is now clearly stated in the revised manuscript.

Page 3960, lines 1-2: Why did the authors decide to take various coincidence criteria here suddenly? Before this paragraph, the authors did not describe anything on the importance of the spatial distribution for the regression analysis. Isn't there some other reason which would cause the difference of the slopes in the China and Tokyo cases? The authors should examine other reasons.

Reply: Following this comment, to note the importance of the spatial distribution in advance, we have added one sentence "The coincidence criterion x of 0.20° is first tested here, while the comparison results could be affected by the choice of x according to the spatial distribution of NO2 around observation sites, as discussed later."

Page 3960, lines 23-25: The description suggests that the authors assume a constant bias in any regions. Is this justified? The tropospheric NO2 amounts include the error of AMFs, which should contain the regional dependence (uncertainties of emission inventory, aerosol loading, albedo etc.).

Reply: We agree with the reviewer. Considering the limited MAX-DOAS observations and the significant spatial inhomogeneity effects, we think that the estimate of the representative bias for each satellite sensor is most appropriate at this stage. We hope to identify such a regional dependence using more robust validation data.

Page 3960, lines 26-27: How did the authors calculate simply averaged slopes? Did the authors use the slopes of the fifteen criterion x for the simple average? If so, it is strange. The data with smaller criterion x would be added repeatedly. Without averaging, the authors could immediately obtain one slope value using one criterion x in one regression analysis.

AMTD

5, C2280–C2291, 2012

Reply: We understand the reviewer. It was difficult to uniquely determine the most appropriate criterion (as the reviewer suggests) due to limited satellite data and spatial inhomogeneity effects, as mentioned earlier. Also, we think that the stricter criterion is better in terms of a smaller spatial inhomogeneity effect, but at the same time it can lead to worse statistics. To consider these and the reviewer's comment, we give, in addition to averaged slopes over 15 coincidence criteria (at $x < 1.00°$), averaged slopes at $x < 0.50°$ and the original slopes (in Figs. 3 and 4). Furthermore, we added a sentence stating that the slopes are all less than 20% irrespective of the choice of x, in response to the reviewer's concern.

Page 3960, lines 27-29: How did the authors obtain the biases? Suppose the slope is 0.9, is the bias 10%?

Reply: We say a bias of -10%, when the slope is 0.9. This relationship is now clearly stated in the revised manuscript.

About the error of the biases: All of the errors shown in the text are 14%. Is that by chance? Otherwise, did the authors just show the uncertainties of MAX-DOAS NO2 retrieval? If so, why did the authors use the phrase "mostly"? I was confused as I expected other error sources.

Reply: We have revised the manuscript to state that the error of the bias is derived as the root-sum-squares of the slope's uncertainty and the uncertainty of the MAX-DOAS NO2 retrieval.

The errors of the biases should be the uncertainties of the slopes in one regression analysis. I do not think the uncertainties of the slopes can be simply replaced by the uncertainties of MAX-DOAS NO2 retrieval. The authors should show the slopes' uncertainties derived from the regression analysis. Otherwise, justify the replacement in the text.

Reply: The slope's uncertainties are now given in Figures 2 and 3. Also, the revision has been made to state that the error of the bias does not come from the uncertainty of the MAX-DOAS NO2 retrieval only but is derived as the root-sum-squares of the slope's uncertainty and the uncertainty of the MAX-DOAS NO2 retrieval.

Page 3961, lines 1-4: Why is the criterion of 0.50° strict enough? I think the number of 0.50 has no foundation. Show the reason that the criterion of 1.0° is not strict enough but the criterion of 0.50° is enough. The authors concluded that a smaller criterion x is better here. From Figure 5, however, the authors concluded that the statistics in the China case have a small dependence on the criterion x. The two conclusions look inconsistent with each other.

Reply: We think that the stricter criterion is better, in terms of a smaller spatial inhomogeneity effect, but at the same time it can lead to worse statistics. For the China case, while the dependence of the slope on the criterion x is small as a result of the NO2 being distributed homogeneously there (compared to the Tokyo case), we still think that the estimated slope is more accurate at smaller x values, if a sufficient number of comparisons is available. As the reviewer points out, we realize that there is no very strong foundation for the number of 0.50. We have added a discussion of this point in the revised manuscript. Furthermore, we added a sentence stating that the slopes are all less than 20% irrespective of the choice of x.

Technical corrections

There are some subjective expressions in the text (e.g., "excellent" at line 23 on page 3959 and "high quality" at line 7 on page 3961).

Reply: These have been changed to more objective words, as suggested.

Page 3957, line 10: "+30%" might be "(+30%)".

Reply: We have confirmed that "+30%" is correct.

Page 3959, line 14: "The slopes. . .are controlled by comparisons. . ." This sentence sounds strange to me; do "comparisons" control the slopes??

Reply: This sentence has been revised to "For the Tokyo case the slope is controlled mainly by data from Tsukuba and Yokosuka, ..."

Page 3959, line 15: "for the China case the three Chinese sites" Does this mean "(the slopes) for the China case (controlled by) the three Chinese sites"? This sentence would need rearrangement.

Reply: This sentence has been revised to "Similarly, for the China case the slopes are controlled by data from the three Chinese sites."

Page 3961, lines 23-24: The word "thus" is used twice successively.

Reply: Corrected.

Interactive comment on Atmos. Meas. Tech. Discuss., 5, 3953, 2012.