

Interactive comment on “First quantitative bias estimates for tropospheric NO₂ columns retrieved from SCIAMACHY, OMI, and GOME-2 using a common standard” by H. Irie et al.

Anonymous Referee #1

Received and published: 3 August 2012

General comments

This manuscript focuses on quantitative bias estimates of the tropospheric NO₂ retrieved from satellite-based measurements with their original ground-based measurements. The validation for such satellite-based products is quite important, and the purpose of this manuscript has a large scientific significance. At this stage, however, my assessment is that this manuscript needs a major change for acceptance, as there are many unclear and inconsistent descriptions.

1. The authors gave a lot of arbitrary values (0.05, 0.10, 0.15, 0.20, 0.25, 0.30, 0.35,

C1711

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.

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 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.

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.

C1712

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.

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.

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

C1713

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.

-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.

-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.

-Section 4

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

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 .

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

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

C1714

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).

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 $2\text{-}3 \times 10^{15}$ molecules /cm² shown in Figure 2. The values are close to the error of the OMI data (1×10^{15} , which is shown at line 10 on page 3957).

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

Page 3959, line 20: What is “this”?

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.

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.

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?

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

C1715

comment.)

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.

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.).

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.

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

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.

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.

C1716

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.

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).

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

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

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.

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

References

Bucsela, E. J., et al. (2008), Comparison of tropospheric NO₂ from in situ aircraft measurements with near-real-time and standard product data from OMI, *J. Geophys. Res.*, 113, D16S31, doi:10.1029/2007JD008838.

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