

## **Reply to editor, including a list of all relevant changes made in the manuscript**

We appreciate thoughtful comments given by the reviewers. Our answers to the comments are attached from page 4 (same as those posted online).

The manuscript was revised accordingly. The introduced changes are listed below. We hope that we adequately answered to the comments and improved the manuscript accordingly.

1. Abstract, line 29. Upon suggestion by reviewer #1, we mentioned that this is the first synthetic data analysis of the MADRAS network observations.
2. Abstract, lines 41–43. We explained “aerosol shielding effect” plainly, as a possibility of incomplete accounting of NO<sub>2</sub> near the surface under relatively high aerosol conditions for the satellite observations.
3. Abstract, lines 45–49. We removed the previous sentences mentioning that urban type seasonality was found at most sites and now included two new sentences with respect to the weekly and diurnal variations as follows: Weekend reduction in the TropoNO<sub>2</sub>VCD found at Yokosuka and Gwangju was absent at Hefei, implying that the major sources had different weekly variation patterns. While the TropoNO<sub>2</sub>VCD generally decreased during the midday hours, it exceptionally increased at urban/suburban locations (Yokosuka, Gwangju, and Hefei) during winter.
4. Section 2.2. We shortened the lengthy sentences with respect to the deployment layout of the instruments in section 2.1 of the original manuscript. Now the sentences are (lines 240-248):

At each site, the light-receiving unit of the MAX-DOAS instrument was located on the rooftop of single to five-story building. Three types of deployment layout were used with respect to the spectrometer (type A, B, and C, see Table 1). For type A and B, the spectrometer was located indoors in a customized thermoelectrically controlled refrigerator (type A) or in a temperature stabilized case (type B). The controlled temperature was 20 or 25 °C for type A and 40 °C for type B for all seasons. For type C, the spectrometer was located outdoors in a light-receiving unit and its temperature was controlled to seasonally-changing levels (25 or 30 °C in winter and 35 or 45 °C in summer). The precise temperature stabilization within ±0.2 °C (using KT4; Panasonic, Kadoma, Japan)...

The types of the instruments are mentioned in Table 1. Table 1 also includes types of the observational sites (e.g., urban, suburban, rural and remote).

A new sentence (lines 236-237) mentioned that the Fukue site is remote, away from major cities (e.g., 100 km from Nagasaki, population 440,000).

5. Section 2.4. All information of satellite data set used for comparison with MAX-DOAS is now summarized here, upon suggestion by the reviewer #1.

### **2.4 OMI satellite data products for comparison**

In this section OMI satellite data products of TropoNO<sub>2</sub>VCD, to be compared with our MAX-DOAS

products in section 3.1, are summarized. The OMI is a UV/vis nadir viewing spectrometer on the National Aeronautics and Space Administration (NASA) Aura satellite on a sun synchronous orbit launched in 2004. An OMI pixel size is  $13 \times 24 \text{ km}^2$  or larger. We used two different products, i.e., one derived from the algorithm developed by the NASA and the other from the algorithm (Dutch OMI  $\text{NO}_2$  (DOMINO)) developed by Koninklijk Nederlands Meteorologisch Instituut. The NASA data set was the ver. 2.1 release of the gridded OMNO2d daily level 3 products (OMNO2d.003), with cloud screening at 30%, at a resolution of  $0.25^\circ \times 0.25^\circ$ , available from the NASA Giovanni website ([http://gdata1.sci.gsfc.nasa.gov/daac-bin/G3/gui.cgi?instance\\_id=omi](http://gdata1.sci.gsfc.nasa.gov/daac-bin/G3/gui.cgi?instance_id=omi); Bucsela et al., 2013). The latter data set was the monthly DOMINO ver. 2.0 collection 3, at a resolution of  $0.125^\circ \times 0.125^\circ$ , available from the Tropospheric Emission Monitoring Internet Service (TEMIS) website ([http://www.temis.nl/airpollution/no2col/no2regioomimonth\\_col3.php](http://www.temis.nl/airpollution/no2col/no2regioomimonth_col3.php); Boersma et al., 2011). The data at the nearest grid were used for both products. For the DOMINO algorithm, the results at eight adjacent grids were included (gray lines in Fig. 7) in addition to the nearest grid, to represent the spatial inhomogeneity of  $\text{NO}_2$  over the range  $0.375^\circ \times 0.375^\circ$ .

The two algorithms subtract stratospheric  $\text{NO}_2$  component as simulated by a chemical transport model (DOMINO ver 2.0) or determined directly from satellite data (NASA ver. 2.1) from the total and then determine Tropo $\text{NO}_2\text{VCD}$  using tropospheric air mass factors. For both satellite data products, air mass factors were computed as average of clear and cloudy conditions weighted by the cloud radiation fraction, and therein the aerosols are implicitly taken into account similarly to clouds (Boersma et al., 2011; Bucsela et al., 2013). They both adopt the vertical profile shapes of  $\text{NO}_2$  simulated by global chemical transport models at relatively coarse resolutions (TM4 at  $2^\circ \times 3^\circ$  for DOMINO and GMI at  $2^\circ \times 2.5^\circ$  for NASA) but with down to monthly time resolution. The uncertainty for individual retrievals of Tropo $\text{NO}_2\text{VCD}$  was estimated to be  $1.0 \times 10^{15} \text{ molecules cm}^{-2} + 25\%$  (Boersma et al., 2011), and on the order of  $10^{15} \text{ molecules cm}^{-2}$  (Bucsela et al., 2013) for clear-sky conditions. Observational information content with respect to vertical profiles of  $\text{NO}_2$  and the amount of aerosols is less than the case of MAX-DOAS observations; advantages of MAX-DOAS are that 1) observations of  $\Delta\text{SCDs}$  of  $\text{NO}_2$  at multiple axes are available, 2) simultaneous determination of aerosols is enabled using  $\text{O}_4$  absorbance determined in the same axes, and 3) the determined aerosol quantities are explicitly taken into account in the  $\text{NO}_2$  retrievals, although data at fixed locations with the instruments are only available.

6. Page 19, lines 489-497. More explanation was added, with respect to the observational information content for the comparison between MAX-DOAS and satellite data products.

This is less likely to be explained by overestimation by MAX-DOAS at high AODs, where larger observational information content regarding aerosols (multiple axis measurements of  $\text{O}_4$ ) were used in the derivation of  $A_{\text{box}}$  and Tropo $\text{NO}_2\text{VCD}$ . All of the data ( $n = 1834$  from the six sites) were subdivided into two groups of equal size, based on AOD values (i.e., two groups with high and low AOD values) and a Welch's t-test was applied to test the statistical significance of the difference between the two means. The results suggested that the OMI(NASA)/MAX-DOAS ratio was significantly lower for the group with higher AODs at the 95% confidence level. Similar tests for individual sites led to the same conclusion for Fukue, Zvenigorod, and Gwangju.

7. Page 20, lines 511-519. We rephrased the explanation, regarding the possible underestimation by the satellite observations. We mentioned previous theoretical work (Leitao et al.) here.

Considering that the variance of the ratio was only partly explained by AOD (Fig. 10a), one could argue that the effect of aerosols was almost successfully removed even in the current satellite data retrieval. However, a weak dependence of the ratio on AOD was still discernible, suggesting that the retrieval could be improved by an explicit treatment of the aerosols. Leitão et al. (2010) theoretically demonstrated that such underestimation by satellite observations could occur when the aerosol layer extends to relatively higher altitudes than NO<sub>2</sub>. Recently, Shaiganfar et al. (2011) and Ma et al. (2013) implied that the shielding effect of NO<sub>2</sub> by aerosols could be the cause of low biases in OMI observations over India and China.

8. Pages 21-22, lines 546-550.

Inspired by the comment by the reviewer #1, a new paragraph as follows was added to show that the satellite observation successfully captured day-to-day variations in addition to the monthly variations, although the sensitivity was consistently small. Figure 12 was inserted accordingly.

Deviations of pixel-level OMI(NASA) TropoNO<sub>2</sub>VCD from monthly mean values positively correlated with those for coinciding MAX-DOAS observations (Fig. 12, with  $R^2$  values of 0.45) with a slope of 0.40, a similar value of those in Fig. 9. This suggested that the satellite observation successfully captured day-to-day variations in addition to the monthly variations, although the sensitivity was consistently small.

9. Summary, lines 670-671. A sentence that future satellite observations with smaller footprint sizes may help distinguishing the causes is moved from section 3.1 to here.

10. Figure 5. X-axis titles were changed from EL to elevation angle. (The acronym EL was removed from text too).

11. Figure 13. Data on selected months (Jan, Feb, Apr, Jul, Oct, Nov, and Dec) were only shown here to improve visibility while highlighting the behavior in winter (mentioned in text).

## Reply to the 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.

Comment 1: Abstract: Overall, the abstract is the worst part of an otherwise great manuscript. I probably wouldn't have read this paper based on the poor abstract. The authors are doing themselves a disservice.

1)(line 41): I doubt many will know what an "aerosol shield effect" is. Really you are just saying that aerosols weren't adequately accounted for in the radiative transfer calculation. Please consider rewording this sentence.

2)I would make it clear that this is the first study with MAX-DOAS in Russia and Asia.

3)(line 45): "The prevailing : : : " This sentence doesn't say anything. You know that the sites are located in urban settings.

4)(line 46): "The presence : : : " This sentence is not clear at all. Why "diesel vehicles" and what "other sources"?

Reply 1: The abstract is rewritten. Instead of "aerosol shielding effect" we will use a phrase "incomplete accounting of NO<sub>2</sub> near the surface under relatively high aerosol conditions" for the satellite observations. We will also mention that this is the first synthetic data analysis for the MADRAS network observations. The previous sentences "The prevailing seasonal patterns with a wintertime maximum implied the dominance of anthropogenic emissions around our sites. The presence of weekend reductions at Yokosuka and Gwangju suggested the dominance of emissions from diesel vehicles, with significant weekly cycles, whereas the absence of such a reduction at Hefei suggested the importance of other sources." were removed. We now state more generally that "Weekend reduction in the TropoNO<sub>2</sub>VCD found at Yokosuka and Gwangju was absent at Hefei, implying that the major sources had different weekly variation patterns." One sentence regarding diurnal variations as follows will be added: "While the TropoNO<sub>2</sub>VCD generally decreased during the midday hours, it exceptionally increased at urban/suburban locations (Yokosuka, Gwangju, and Hefei) during winter."

Comment 2: Section 2.3 Retrieval algorithms & Summary.

As you discuss later, some of the discrepancies between the data from MAX-DOAS and OMI are associated with different assumptions used in the two OMI and MAXDOAS retrieval algorithms.

Could you describe the two OMI retrieval algorithms, say in a new Section 2.4, and highlight the main differences between the two OMI algorithms and the algorithm used for the MAX-DOAS instruments? You could also comment on the limitations of satellite data, such as spatial coverage, which should be accounted for when comparing to surface observations.

Reply 2. As suggested, we will describe the two OMI retrieval algorithms in a new Section 2.4 as follows, highlighting differences in the observational information content regarding aerosols and vertical profiles between MAX-DOAS and satellite observations:

## **2.4 OMI satellite data products for comparison**

In this section OMI satellite data products of TropoNO<sub>2</sub>VCD, to be compared with our MAX-DOAS products in section 3.1, are summarized. The OMI is a UV/vis nadir viewing spectrometer on the National Aeronautics and Space Administration (NASA) Aura satellite on a sun synchronous orbit launched in 2004. An OMI pixel size is  $13 \times 24 \text{ km}^2$  or larger.

We used two different products, i.e., one derived from the algorithm developed by the NASA and the other from the algorithm (Dutch OMI NO<sub>2</sub> (DOMINO)) developed by Koninklijk Nederlands Meteorologisch Instituut. The NASA data set was the ver. 2.1 release of the gridded OMNO2d daily level 3 products (OMNO2d.003), with cloud screening at 30%, at a resolution of  $0.25^\circ \times 0.25^\circ$ , available from the NASA Giovanni website ([http://gdata1.sci.gsfc.nasa.gov/daac-bin/G3/gui.cgi?instance\\_id=omi](http://gdata1.sci.gsfc.nasa.gov/daac-bin/G3/gui.cgi?instance_id=omi); Bucsela et al., 2013). The latter data set was the monthly DOMINO ver. 2.0 collection 3, at a resolution of  $0.125^\circ \times 0.125^\circ$ , available from the Tropospheric Emission Monitoring Internet Service (TEMIS) website ([http://www.temis.nl/airpollution/no2col/no2regioomimonth\\_col3.php](http://www.temis.nl/airpollution/no2col/no2regioomimonth_col3.php); Boersma et al., 2011). The data at the nearest grid were used for both products. For the DOMINO algorithm, the results at eight adjacent grids were included (gray lines in Fig. 7) in addition to the nearest grid, to represent the spatial inhomogeneity of NO<sub>2</sub> over the range  $0.375^\circ \times 0.375^\circ$ .

The two algorithms subtract stratospheric NO<sub>2</sub> component as simulated by a chemical transport model (DOMINO ver 2.0) or determined directly from satellite data (NASA ver. 2.1) from the total and then determine TropoNO<sub>2</sub>VCD using tropospheric air mass factors. For both satellite data products, air mass factors were computed as average of clear and cloudy conditions weighted by the cloud radiation fraction, and therein the aerosols are implicitly taken into account similarly to clouds (Boersma et al., 2011; Bucsela et al., 2013). They both adopt the vertical profile shapes of NO<sub>2</sub> simulated by global chemical transport models at relatively coarse resolutions (TM4 at  $2^\circ \times 3^\circ$  for DOMINO and GMI at  $2^\circ \times 2.5^\circ$  for NASA) but with down to monthly time resolution. The uncertainty for individual retrievals of TropoNO<sub>2</sub>VCD was estimated to be  $1.0 \times 10^{15} \text{ molecules cm}^{-2} + 25\%$  (Boersma et al., 2011) and on the order of  $10^{15} \text{ molecules cm}^{-2}$  (Bucsela et al., 2013) for clear-sky conditions. Observational information content with respect to vertical profiles of NO<sub>2</sub> and the amount of aerosols is less than the case of MAX-DOAS observations; advantages of MAX-DOAS are

that 1) observations of  $\Delta$ SCDs of  $\text{NO}_2$  at multiple axes are available, 2) simultaneous determination of aerosols is enabled using  $\text{O}_4$  absorbances determined in the same axes, and 3) the determined aerosol quantities are explicitly taken into account in the  $\text{NO}_2$  retrievals, although data at fixed locations with the instruments are only available.

Comment 3: Section 3.1. It is easy to get a great correlation between OMI and MAX-DOAS since they both will reproduce the seasonal cycle in  $\text{NO}_2$ , which is associated with its chemical lifetime. I recommend deseasonalizing the data for a fairer assessment of how well the MAXDOAS and OMI data agree.

Reply 3: For both MAX-DOAS and OMI, the seasonality in Tropo $\text{NO}_2$ VCD arose from the observed SCDs themselves, without using external information (e.g., a priori modeled seasonality, chemical lifetime etc), although seasonality in the vertical profile SHAPES was only used for OMI. Therefore the shown correlation (e.g., Fig. 8) is not automatically expected and is worth to be discussed. On the other hand, comparison of deseasoned data as suggested by the reviewer is important to test if the satellite captures the day-to-day variations. In the revised manuscript, we will add one figure (attached, to be inserted as Fig. 12 in the manuscript) and explanation for it at the last part of Section 3.1 as follows:

Deviations of pixel-level OMI(NASA) Tropo $\text{NO}_2$ VCD from monthly mean values positively correlated with those for coinciding MAX-DOAS observations (Fig. 12, with  $R^2$  values of 0.45) with a slope of 0.40, a similar value of those in Fig. 9. This suggested that the satellite observation successfully captured day-to-day variations in addition to the monthly variations, although the sensitivity was consistently small.

Comment 4. Line 465: This paragraph is very unclear.

Line 473: Describe how "the effect is already taken into account more adequately."

Line 485: Huh? How are the vertical profiles treated differently in the retrieval algorithms?

Reply 4. The Lines 465, 473, and 485 of the submitted manuscript are in the paragraph starting from line 11 of page 2901 of the typeset ACPD manuscript. In the revised manuscript, several unclear sentences in the previous manuscript are revised and more explanation is given. The improved paragraph will be as follows:

Another possibility would be that systematic underestimation by satellite observations arises from assumptions in the vertical profiles and aerosol treatment. Figure 10a shows that low OMI(NASA)/MAX-DOAS ratios (using a gridded data set for OMI) are associated with high AODs (as observed by MAX-DOAS); although the median ratio is near unity at low AODs ( $\sim 0.1$ ), it becomes lower ( $\sim 0.7$ ) with AODs as high as 1. In this study, only data with more than  $1 \times 10^{15}$  molecules  $\text{cm}^{-2}$  for both MAX-DOAS and satellite observations are used. This suggests the possibility that the satellite

observations underestimate TropoNO<sub>2</sub>VCD when aerosols are densely present. This is less likely to be explained by overestimation by MAX-DOAS at high AODs, where larger observational information content regarding aerosols (multiple axis measurements of O<sub>4</sub>) were used in the derivation of  $A_{\text{box}}$  and TropoNO<sub>2</sub>VCD. All of the data ( $n = 1834$  from the six sites) were subdivided into two groups of equal size, based on AOD values (i.e., two groups with high and low AOD values) and a Welch's t-test was applied to test the statistical significance of the difference between the two means. The results suggested that the OMI(NASA)/MAX-DOAS ratio was significantly lower for the group with higher AODs at the 95% confidence level. Similar tests for individual sites led to the same conclusion for Fukue, Zvenigorod, and Gwangju.

Figure 10b shows that the OMI(NASA)/MAX-DOAS ratio for TropoNO<sub>2</sub>VCD had a weak decreasing trend with the retrieved parameter  $\nu_1$ , the fraction of NO<sub>2</sub> present in the lowest 1 km. The median values decreased from around unity to 0.67 as  $\nu_1$  increased from  $\sim 0.6$  to 0.9. Welch's t-tests applied to two groups of data sorted by  $\nu_1$  values suggested that the ratio was significantly lower for the group with higher  $\nu_1$  values when using data from all six sites and when using data from Yokosuka and Hefei individually, at the 95% confidence level. This suggests that the underestimation occurs when NO<sub>2</sub> is mostly present near the surface. These analyses, in combination, imply that the lower values from the satellite could be partly caused by the assumptions made regarding the vertical profiles and aerosol treatment in the satellite data analysis. This may be important at clean sites, where the spatial inhomogeneity cannot be responsible for the difference. For both satellite data products, air mass factors were computed as average of clear and cloudy conditions weighted by the cloud radiation fraction, and therein the aerosols are implicitly taken into account similarly to clouds (Boersma et al., 2011; Bucselo et al., 2013). Considering that the variance of the ratio was only partly explained by AOD (Fig. 10a), one could argue that the effect of aerosols was almost successfully removed even in the current satellite data retrieval. However, a weak dependence of the ratio on AOD was still discernible, suggesting that the retrieval could be improved by an explicit treatment of the aerosols. Recently, Shaiganfar et al. (2011) and Ma et al. (2013) suggested that the shielding effect of NO<sub>2</sub> by aerosols could be significant for OMI observations, resulting in similarly low values. Lin et al. (2013) suggested that concentration of aerosols at the top of the boundary layer increased retrieved NO<sub>2</sub> by 8%.

Comment 5: Figure 3. Why is OMI NO<sub>2</sub> much higher over the ocean near Japan than over Japan's cities?

Reply 5: The map of the country borderline did not match precisely. An improved figure will be used.

Comment 6: Figure 7. Why are the error bars on the MAX-DOAS data so much larger than those on the OMI data? It seems from your conclusions that the error bars on the OMI data should be much

larger.

Reply 6: The major cause of the large bars is that data with relatively large values are included in MAX-DOAS. It should be noted that the bars represent 1-sigma ranges of the MAX-DOAS and satellite data, and do not include systematic uncertainty. Satellite observations sometimes fail to reproduce data with relatively large values, especially when high  $\nu_1$  values are associated. Thus, our conclusion does not contradict this figure.

We thank the reviewer for their comments helping us to improve our manuscript.

## Reply to the referee #2

We appreciate the comments given by the reviewer after careful reading of our manuscript.

Comment 1: Was there an intercomparison with all instruments part of MADRAS at the same site? Even the instruments are quite similar some important parts are not, which might lead to systematic differences. For section 2.1 I prefer to have in addition a table illustrating more clearly the differences between the setups instead of a lengthy and difficult to read description. A similar comment to section 2.2: Please add information on site characteristics to Table 1.

Reply 1. We have run two instruments together at the same site. The typical difference in the derived NO<sub>2</sub> DSCDs was <5%, indicating that the instrument-specific difference was small.

We will classify the instruments into three types (type A, B, and C), according to the deployed location of the spectrometer and temperature settings (indoors (in a refrigerator, in a temperature-stabilized box) and outdoors in a light-receiving unit) and explain them in text (this part will be moved to section 2.2). The type of each instrument will be shown in Table 1. Table 1 will also include types of the observational sites (e.g., urban, suburban, rural and remote). In section 2.2 of the revised text, only brief information of the observational sites will be included.

Comment 2: Please avoid as much acronyms as possible. Since the text is already quite technical, in particular unusual acronyms like EL (for elevation angle!) and also TropoNO<sub>2</sub>VCD (why not just name it “tropospheric NO<sub>2</sub> column”?) do not improve the readability.

Reply 2: Upon suggestion we will use elevation angle instead of EL. After consideration, we retain the acronym TropoNO<sub>2</sub>VCD, used more than 50 times in the text and in axis titles in many figures, believing that the acronym leads to conciseness of the manuscript.

Comment 3: Comparison between MAX-DOAS and satellite data: Please refer in your discussion on pp 2901 and 2902 to the study by Leitao et al., (Leitão, J., Richter, A., Vrekoussis, M., Kokhanovsky, A., Zhang, Q. J., Beekmann, M., and Burrows, J. P.: On the improvement of NO<sub>2</sub> satellite retrievals – aerosol impact on the airmass factors, *Atmos. Meas. Tech.*, 3, 475-493, 2010) where in detail the impact of aerosols on the satellite observations has been investigated. P. 2902, l21: Please remove or rephrase the statement: “Future satellite observations with smaller footprint sizes will improve the analysis.” This would be the case, if spatial inhomogeneity is the major reason for the underestimation by satellite instruments. But the outcome of this (and other) study (ies) is quite clear, that this is not the case.

Reply 3: Thank you for the important suggestion. In the revised manuscript we will mention that

Leitão et al. (2010) theoretically demonstrated that such underestimation by satellite observations could occur when the aerosol layer extends to relatively higher altitudes than  $\text{NO}_2$ . The sentence regarding future satellite observations will be deleted here. Instead in the summary section, we will mention that future satellite observations with smaller footprint sizes may help distinguishing the causes.

Comment 4: Figure 12: This plot is too busy. Please select either months representative for the different seasons or use seasonal averages (e.g. December to February for winter like in Figure 17).

Reply 4. We will leave data on selected months (Jan, Feb, Apr, Jul, Oct, Nov, and Dec) to improve visibility while highlighting the behavior in winter (mentioned in text).

We thank the reviewer for the comment helping us to improve our manuscript.