<u>Review of Zhang et al. (2016)</u>: *"Continuation of long-term global SO_2 pollution monitoring from OMI to OMPS"*

Zhang et al. provide a comparison of the current operational NASA principal component analysis (PCA) SO_2 retrieval results for the OMI and OMPS satellite instruments, focusing on three global regions affected by anthropogenic pollution during 2012-2015. Based on daily and annually averaged data, the results indicate that the currently operating (and upcoming) OMPS instrument(s) allow to continue the long-term global SO_2 pollution monitoring by OMI.
The paper is well written. However, besides considering the following specific comments, I think the authors should put a little more effort into analysing the remaining differences between both satellite instruments (e.g. spatial resolution vs. daily coverage) before final publication in AMT. While the comparison of the OMPS and OMI SO_2 results shows a general agreement between both instruments (which is good to see, but not surprising), not much information is given about the discrepancies and how good both datasets really agree if they would be used for a quantitative long-term monitoring of anthropogenic emissions. Especially for the SO_2 emissions over China it would have been nice to see a more detailed case study, e.g. including a plot with monthly total SO_2 masses for OMI 2004-2015 directly compared to OMPS 2012-2015.

<u>Specific comments</u>:

<u>Abstract, l.18-22</u>: At this point of the paper it remains unclear what the authors mean by the terms 'temporal' and 'spatial' correlation. It would be helpful to clearly state somewhere in the first sections of the manuscript that a linear regression fit is applied to the <u>daily</u> SO_2 masses for an extended time period (with the results here called 'temporal' correlation) and to the single grid pixels of yearly averaged SO_2 column (results here called 'spatial' correlation).

<u>Abstract, l. 22-23</u>: The choice of $r > 0.6$ (and only focusing on the Mexican region) appears a little bit arbitrary in order to demonstrate how 'good' both satellite instruments agree. Why did you use this specific value? As a first time reader I would have the impression that the numbers for $r > 0.7$ (or higher) are just too low to mention in the abstract. The results presented in Fig. 6 further indicate that only roughly half of the measurements have a correlation coefficient $r > 0.6$ for eastern China and the South African region, which is surprisingly low at a first glance for two instruments evaluated by using the same algorithm and with only some minutes temporal difference. Maybe the main conclusion here is that there are substantial differences for the two satellite instruments if the SO_2 columns are compared pixel by pixel on a daily basis (probably because of the different spatial resolution) and this should be discussed in more detail in the manuscript.

<u>p. 2, l. 14-17</u>: Could you be a little more specific on that? Isn't it self-explanatory that the main limiting factor in detecting point sources is spatial resolution?

<u>p. 4, l. 15-16</u>: What is the reason for not applying an additional empirical background correction over the Pacific? Figure 1 indicates that there are small (but noticeable) systematical differences for the two satellite instruments.

<u>p. 4, l. 26-27</u>: It is difficult to distinguish the volcanic SO_2 emissions from Popocatépetl from the anthropogenic emissions in the Mexico City area. Based on daily regional OMI SO_2 column maps

(e.g. available via http://so2.gsfc.nasa.gov/) and because of the increased sensitivity of the satellite instruments to SO_2 at higher altitude, it is quite likely that the main emissions in the considered region (or at least a large fraction) are of volcanic origin. It's however interesting to see that the results agree best over Mexico compared to the other two scenarios (again probably because of the increased sensitivity). Certainly, it would have been better to focus on another region with strictly anthropogenic emissions instead of Mexico (e.g. the copper smelters in Norilsk, Russia), especially because there is a similar (more comprehensive) AMTD paper by the authors that strictly focuses on volcanic SO_2 emissions. These problems should be discussed in more detail in the revised version.

Figure 2: Unfortunately, the colorbars are all restricted to ± 2 and ±0.5 DU. Please scale to maximum/minimum values for each region. Otherwise the main enhancements are always shown in red/white and it remains unclear how large the differences really are.

Please indicate that the lower figures show "OMPS - OMI".

The most remarkable discrepancies appear for the highest columns. What are possible reasons? For the Mexican region, I think this is another indication that the main part of the enhanced SO_2 columns are caused by volcanic emissions. The location of the volcano as well as the city centre should be marked in Figure 2.

p. 4, l. 32: As already mentioned in the previous comment: The OMPS SO_2 columns are only lower than those from OMI for the highest OMPS/OMI values for Mexico/South Africa. For large parts of Mexico, the OMPS columns appear higher. Please discuss possible reasons.

p. 5, l. 1: What data are included in the 'annual regional averaged SO_2 columns'? Is it the data from the area shown in Figure 2? If so, please name coordinates of the considered areas.

p. 5, l. 2ff: Please explain the terms 'spatial/temporal' correlation.

p. 5, l. 2: Please give the exact number (not just '>0.9') as you did for the following 2 cases. The r value for China is even the highest in 2012 with 0.96.

p. 5, l. 3: The patterns appear to be all similar, even for lower VCDs. However, it is difficult to compare the data only by taking a look at the given maps, as the maximum values are all in red because of the chosen colorbar. Please scale to maximum/minimum values for all 3 regions.

p. 5, l. 4ff: How large is the variation of the values if you include a wider area (how smooth is the background signal)? It would be good to have a reference value over an unpolluted area nearby, as no empirical offset correction is applied.

p. 5, l. 9-11: Especially in the case of Mexico, the peak OMI SO_2 columns are much higher than those from OMPS. Probably this is the case because of Popocatépetl as the dominating (point) source in the region.

p. 5, l. 14-15: The given numbers indicate that you should definitely adjust the minimum/maximum of the colorbars in Figure 2: Given a maximum SO_2 difference of -2 DU, probably all Mexican negative values are 3-4 times larger than your minimum colorbar value of -0.5 DU.

As your upper limit for the colorbar is +2 DU for the averaged OMI/OMPS data maps, it would be even possible that some of the clearly enhanced measurements differ by a factor of 2

How large are the positive differences?

p. 5, l. 22-23: It would be very helpful to show a total SO_2 mass timeline 2012-2015 for all 3 regions in an additional figure as well as maps for the other years (2013-2015).

p. 5, l. 26: The difference in sampling time should be less than 30 minutes, right (typically about 10 minutes)? Do we really expect here large differences due to transportation?

p. 5, l. 30-31: It would be better to give the total (regional integrated) SO_2 mass instead of the averaged SO_2 column.

p. 6, l. 1: How significant is the change of the total SO_2 mass if you change the area considered for the calculations (e.g. ± 1° lat/long)? Some emissions over China are certainly left out because of the currently selected area.

p. 6, l. 31-33: I don't really get the argument. Do you think that the correlation decreases because the remaining SO_2 columns are close to the detection limit and therefore scatter around a remaining offset? Otherwise I don't see why the decrease of the averaged SO_2 column should lead to more significant discrepancies. Please reformulate.

p. 6, l. 17: This is only the case for 2014/2015. For 2012/2013 the slope decreases while the r value remains quite stable...

p. 6, l. 18-19: What are the total SO_2 masses per year for the individual regions? In my opinion, it would make more sense to give the total emissions instead of the averaged daily SO_2 mass for each year.

p. 6, l. 21-23: Same argument as in p. 6, l. 31-33: Please make this point more clear.

p. 7, l. 3: Which days were excluded? All days in April/May or just the ones exceeding a certain SO_2 threshold over South Africa?

p. 7, l. 16-17: Why did you decide to show only data for August in each year? Including all individual months would further indicate if there are some seasonal differences between OMPS and OMI.

Minor comments:

p. 2, l. 4-5 and 10-11: Please add Theys et al. (2015). This is one of the most recent and comprehensive papers about an OMI SO_2 retrieval.

Theys, N., I. De Smedt, J. van Gent, T. Danckaert, T. Wang, F. Hendrick, T. Stavrakou, S. Bauduin, L. Clarisse, C. Li, N. Krotkov, H. Yu, H. Brenot, and M. Van Roozendael (2015), Sulfur dioxide vertical column DOAS retrievals from the Ozone Monitoring Instrument: Global observations and comparison to ground-based and satellite data. J. Geophys. Res. Atmos., 120, 2470–2491. doi: 10.1002/2014JD022657.