Authors' Response to Reviewer 2

We thank the Reviewer for the throrough revision of our manusrcipt and constructive comments. These suggestions improved the presentation and quality of the original manuscript. Please find below Authors' response (**A**) to Reviewer's comments (**R**):

General Comments:

R1: This paper is a straightforward application of the existing technique of Kulmala et al., ACP (2011) to estimate nucleation mode particles over South Africa from MODIS and OMI data products. Since nucleation mode particles are too small to be directly detected via ground-based or space-borne optical instruments, a proxy is computed that balances retrieved concentrations of gas-phase particle precursors against the existing aerosol condensational sink as represented by particle concentrations in the larger, optically-active size range. The difference between the two papers is that this study focuses on South Africa, while Kulmala et al., ACP (2011) focuses both on Hyytiala as well as extends the work to create global nucleation proxy maps. The present paper also develops a slightly different method for approximating the condensational sink (CS) from ground-based measurements and AERONET retrievals. The new method for approximating condensational sink is found to not substantially improve over the existing assumption that CS=AOD. In reviewing Kulmala et al., ACP (2011), that Referee #1 noted that "After some changes, [that] manuscript could be suitable for publication in ACP. Not for the goodness of the results but in order to encourage developing more suitable satellite products for the analysis of fine particles. [Her] suggestion is that the main conclusion should be reformulated such that it is not possible to get adequate estimation results for nucleation mode particles with current satellite products." These comments also apply to the present manuscript by Sundström et al., which also shows the extremely poor skill of this technique when using the MODIS and OMI data (Table 3 and Figures 5 and 10). Yet, statements in the abstract and conclusions sections imply that these satellite-based proxies are rather good at showing the potential for nucleation events – this conclusion is not supported by the results! Publishing null results is important because it prevents wasteful duplication of effort and it motivates future work to either improve the proxy method or supplant it with another technique. As such, a paper discussing the reasons for the poor proxy skill and perhaps linking these results to sources of uncertainty and proxy sensitivity would be a welcome addition to ACP. However, with the abstract, discussion, and conclusions as presently stated in this manuscript, I cannot recommend publication.

A1: Based on the Reviewer's comments we agree that we need to clarify the manuscript in many parts. For example, the slightly different method for approximating the condensational sink (CS) from ground-based measurements and AERONET retrievals has been pointed out several time by the Reviewer as one of the major results, which was not our intention. Based on Reviewer's comments it is also obvious that we need to emphasize some of the differences between this study and Kulmala et al. (2011), and modify the text accordingly so that it is clear to the reader that this study is not just a "wasteful duplication of effort" over different location, but instead this study could provide some useful information about using satellite data in such applications. The major differences to Kulmala et al. (2011) study are that while they derived the formulas for the satellite-based proxies and showed preliminary satellite-based proxy maps, we use the actual satellite-data and compare it against in situ measurements. Such analysis was not presented in Kulmala et al. (2011). All the comparisons carried out in their paper between proxies, nucleation mode number concetrations, CS, and AODs had been obtained using in situ data, not actual satellite data. Hence, this is the first manuscript where the satellite-based proxies are compared against in situ measurements.

The Reviewer also raised good points especially related to the in situ-based proxies and the uncertainties related to the satellite proxy-approach (later in the specific comments), which we now have added to the manuscript. As a summary, the major changes made in the manuscript are:

- Abstract has been completely rewritten
- Introduction (Sect. 1): the text has been somewhat modfied and obvious spelling mistakes have been corrected.

- Data, Sect. 2: New paragraph is added where the uncertainties of the OMI NO_2, SO_2 and UV-products are described. Also other parts of text has been somewhat modified.
- Theoretical concepts in Sect 3: Proxies, condensation sink, and aerosol extinction. Some modifications in the text are made.
- Section 4: Results. This Section has undergone major changes and it is now re-organized. E.g. the detailed description of the emission sources as well as the seasonal variation of the different satellite parameters are removed, and the focus is now more on the performance of proxies obtained both using in situ and satellite data. Also the comparison between nucleation event- and non-event days is removed since the statistics were too much skewed (towards the event-days).
    - Subsection 4.1., comparison of CS and AOD: the text has been modified.
    - 4.2., In situ proxies and comparison with nucleation mode number concentration N_{nuc} . This is a new subsection where we calculate the proxies from the in situ data and evaluate their performance in predicting N_{nuc}.
    - 4.3., Proxies using satellite data. This Section has been further divided into three Subsections. In Sect. 4.3.1. the spatial pattern of the satellite NO_2, SO_2 and AOD as well as the proxies is defined from the four years of satellite data. In Sect. 4.3.2. the satellite parameters are compared with in situ data, and in Sect. 4.3.3. proxies calculated using satellite data are compared with in situ N_{nuc} .
- Section 5, Conclusions are rewritten completely.

Specific Comments:

R2: *Throughout the manuscript, linear regression is used to compare the proxy and measurement variables, and goodness of fit is assessed using a Pearson correlation coefficient, r, and associated p-value. First, while the correlation coefficient (r) does give us a metric for assessing the linear dependence between the proxy and measurement variables, the coefficient of determination (R^2) is more meaningful in evaluating the skill of the proxy by representing the proportion of total variation in the measurement captured by the proxy. In all instances in the manuscript except for the comparison between dry scattering coefficient and condensation sink (Figure 3), the R^2 values are around 0 - 0.3. The most direct and important comparison (Figure 10) shows an R^2 value of 0.10. This means that the proxy is only able to explain 0-30% of the observed variability in the measurement variables, and that a majority of the variability remains unexplained. Given this high level of uncertainty, slight improvements in the correlation coefficient do not really show an increase in the skill of the proxy, as is suggested, e.g., on Pg. 25846, Lines 6-8. While it's easy to square the values of r now reported, I suggest the authors report the R^2 values instead throughout the manuscript as a more direct metric for assessing the ability of the proxy to represent the observations.*

A2: Initially we choose to report the R- and p-values similarly as in Kulmala et al. (2011) to be able to compare the results. However, we have now replaced those values with R^2, as the Reviewer suggested.

We agree that even though the correlation between nucleation mode number concentration and satellite-based-proxies slightly increased when using the CS-estimate from the York fit, it did not significantly improve the overall performance of the satellite-based proxy, which lead to the impression the reader might get e.g. from the sentences in p. 25846, lines 6-8. Since this is not a major improvement, we have removed this part of the text. The major issues when estimating CS using satellite AOD are still the ocurrence of elevated aerosol layers and the relative uncertainty related to the rather low AOD values often observed over South Africa. A better improvement could potentially be obtained if there would be a sufficient number of coincident vertical aerosol extiction profiles available.

R3: *In addition, scatter plots of all one-to-one fit comparisons with regression lines should be uploaded as supplementary information. As stated by Referee #1 in reviewing Kulmala et al. (2011): "Drawing a line through a random sample and claiming that there exists significant correlation is bad statistics and in*

some cases even deceptive." Being able to visualize the regressions used to generate Tables 2-3 is essential for understanding how the data are distributed, and including them in the SI ensures that they don't clutter the main paper.

A3: The scatterplots have been added as supplementary material.

R4: *Second, the p-value tells us whether or not we can reject the null hypothesis that r = 0. As illustrated in the plot below (computed assuming the test statistic follows a chi-squared distribution for simplicity), the minimum value of r needed to reject the null hypothesis rapidly decreases as the sample size increases. This can mean that for N greater than about 50-100 points, the correlation can be very weak (i.e., lacking scientific or explanatory significance), but still be statistically significant. As such, this p-value statistic is not really meaningful and should be removed. Instead, the number of points, N, used in the regression should be reported in both the tables and figures.*

A4: We refer here to the answer A2; we have removed the p-values and added the number of points as suggested. Overall, when using actual satellite data in comparisons with the in situ measurements that have been carried out during some limited time period, the number of coincident in situ-satellite observations can become quite low, as in many cases in this study, which is not good from the statistical point of view. We agree that the number of observations should be pointed out more clearly.

R5: *There needs to be a little more honesty in assessing the performance of the proxy. The following statements are not supported by the current results, and therefore should be removed or new supporting data or reanalysis be included to support the claims:*
Page 25826, Lines 10-12: "However, when the AOD in the proxy sink was replaced by an estimate from linear bivariate fit between AOD and CS, the agreement with the actual nucleation mode number concentration improved somewhat." I presume that this relates to the YORK fit vs. the LSQ fit in Figure 5. It's true that the LSQ fit completely fails (presumably due to extreme outliers outside of the plot area), while the YORK fit seems to follow the data better, but it is not clear that this translates into a meaningful improvement in capturing the nucleation mode number concentration.

A5: We refer here to answer A2. One of the major issues is still the fact that AOD is a measure over total atmospheric column and CS is defined from the in situ measurements. Based on the Reviewer's general comments the abstract has been now rewritten and this sentence has been removed.

R6: *Page 25826, Lines 16-19: "Best agreement between the satellite and in situ based proxies were obtained for NO2/AOD and UV-B/AOD2, whereas proxies including SO2 in the source term had lower correlation." Using the numbers from Table 3 to compute R2 values, the NO2/AOD and UV-B/AOD2 proxies are able to explain 10-12% and 3-6% of the observed variation in nucleation mode number concentration, respectively. This is contrasted with the metrics including SO2, which are able to explain at 1-5% of the variability. So, while this statement is technically correct, I find it to be misleading to the reader in that it suggests that there is indeed agreement between the proxy and the observed nucleation mode number concentration. There is only 10-12% agreement at most.*

A6: The sentences in the abstract **(**Page 25826, Lines 16-19) referred to the agreement between in situ and satellite-based proxies, not correlation between the proxies and nucleation mode number concentration. However, based on the Reviewer's general comments the abstract has been now rewritten and this sentence has been removed.

R7: *Page 25845, Lines 8-10: "A distinct improvement in the quality of the proxy components was obtained when different satellite products were selected to those utilized by Kulmala et al. (2011)."*
This statement does not appear to be true. r values from K2011 and present study as follows:
K2011: UV-SO2/CS^2 = 0.54, UV/CS^2 = 0.49, UV-SO2/AOD^2 = 0.25; UV/AOD^2 = 0.23
Present Study (Table 3): UV-SO2/AOD^2 = 0.09-0.21; UV/AOD^2 = 0.17-0.25

A7: The the correlation coefficients of R=0.54 and R=0.49 in Kulmala et al. (2011) have been obtained using in situ data measured at Hyytiälä, not satellite data. Also, the correlation coefficients of 0.25 and 0.23 have been obtained using in situ UV- and SO_2 -measurements, and AOD from the AERONET sunphotometer, not from satellite. Hence, in this work we show the first comparisons of the satellite-based proxies and the nucleation mode number concentration using actual satellite data.

R8: *Page 25846, Lines 6-8: "Some improvement, however, was obtained (0.21 < R < 0.34) when the AOD was replaced by the estimated sink from the York fit (Fig. 5)." Using the numbers in Table 3 to compute R2 values, the use of the York fit versus AOD improves the percentage of explained variance from 10%≻ 12%, 1%≻ 5%, 1%≻ 4%, and 3%≻ 6% for the four different proxies in the order given in Table 3, respectively. It is not clear to me that this is a meaningful improvement.*

A8: We refer again to answer A2. Despite the slight improvement by replacing AOD with the York-fit estimate, the elevated aerosol layer and the relative uncertainty of AOD are still the most significant factors affecting the performance of AOD as a substitute to CS. We have removed this part of the text.

R9: *Page 25846, Lines 20-21: "In general this study showed that the satellite proxies seem to be able to show the potential for nucleation events in a statistical sense. Actual data from non-event days would have been needed to carry out such study." This sentence is just not supported by the data. The statistics are heavily skewed toward event days with no data from non-event days. Therefore, it's not possible to be able to validate or invalidate the ability of the proxy in distinguishing event from non-event days. The second sentence is correct – actual data from non-event days would be needed to draw a conclusion.*

A9: Since the statistics are indeed heavily skewed toward the event days and conclusions can not be made whether the proxies could predict new particle formation or not, we have removed this chapter.

R10: *Page 25846, Lines 22-24: "More studies of the satellite based proxies in different type locations and environments are needed to improve the proxies, and especially the sink term, further." This study and that of Kulmala et al. (2011) have shown that these proxies have very low skill when using column-integrated satellite measurements. This is probably due in large part to the uncertainties and coarse resolution associated with using these column-integrated measurements, and less due to regional peculiarities that might be uncovered by the authors' suggested path forward. Consequently, I doubt that increasing the number of locations and environments studied while following the same set of methods as these two studies will actually improve the proxies. Rather, I would think the best way to improve the proxies would be to improve the satellite inputs. I'd like to encourage you to include some discussion in the paper on measurement uncertainty and the sensitivity of the proxies to these uncertainties, which could serve to underpin future measurement design considerations.*

A10: Both Hyytiälä (Finland) used in Kulmala et al. (2011) and South Africa are locations where the AOD is often very low (< 0.1). In such locations where the relative uncertainty of satellite based AOD is high, and where an elevated aerosol layer can easily double the columnar AOD value, the estimation of (surface) CS is difficult, as is seen in this study. Carrying out a comparison of CS and AOD over an area where the observed satellite-based AOD range is wider, and overall the AOD is higher (and hence the relative uncertainty smaller), we might get somewhat better estimate of the CS. Also in such environment an elevated aerosol layer most probably would not have that large contribution to the total column AOD as it has over this study area. On the other hand, it is true that the satellite NO_2 and especially SO_2 product might still not have the sufficient accuracy that would be needed to improve the performance of the satellite based proxies.

R11: *Page 25846, Lines 24-26: "The next step is to study the satellite based proxy approach in China, where, in addition to the elevated NO2 and SO2 column densities, the AOD signal is also strong." I don't understand the meaning and purpose of this statement, which both singles out the entire country of China as being particularly polluted (which is not supported by any discussion in this paper), and seems to imply that*

the proxies have worked so well in South Africa that no further work is needed and it's time to move on to more complicated regions -- China. This statement should be removed.

A11: We have modified this part of the text as it is clearly misleading. Referring to answers A10 and A15, the relative uncertainties especially in satellite AOD and SO_2 are very high over South Africa since the observed values are low over the major part of the area. E.g. within our study area the typical MODIS AOD values are abt. ~0.1, which would convert to the relative uncertainty of abt. ~65%. China was mentioned here because we have there similar in situ measurements available that were carried out in South Africa. It has been shown in several studies that over China the AOD over densely populated areas could vary around 0.5. In those cases the relative uncertainty of AOD would be ~ 25%, which would also reduce the relative uncertainty associated to the satellite-based proxies. Another question is then how well overall the proxies would work over China for predicting the nucleation mode number concentrations.

R12: *Table 2 and the discussion on Page 25841 indicates that there are weak correlations between the in situ and the satellite-based proxies. If you use only the ground-based, in situ NOx, SO2, UV-B, and N(Dp>100nm) to compute the proxies (Equations 1-3), how well does it correlate with the nucleation mode number concentration? This sort of analysis must be included in the discussion because it places an upper limit on the skill of the proxy in capturing nucleation mode number concentration. If the collocated in situ measurements don't produce reasonable proxies of nucleation mode number concentration, then the much more uncertain, coarser satellite retrievals will not be able to do any better.*

A12: This is a good suggestion. We have now added a new section (the in situ proxy – nucleation mode number concentration comparisons to the manuscript, and discussion considering what could be expected from the satellite based proxies based on the results with the in situ proxies, as well as the diurnal variation of each in situ-based proxy parameter.

R13: *Kulmala et al. (2011) discuss multiple potential proxies corresponding to different assumptions related to the exponent, n, in their equations 6-12. What is the reason for only exploring a single regional proxy as given by Equation 1 in this work?*

A13: Due to the uncertainties related to the satellite data we choose to consider only the cases of Nn,1. In the n^{th} order proxy terms the uncertainties would get even higher than they are now.

R14: *There is a lot of detail about specific point sources given on Pages 25838-25839 including some discussion of the ore types in the smelters. This level of specificity and discussion doesn't seem relevant to this paper, which is concerned with characterizing the satellite proxies using ground-based data. This section should be tied more directly into how it informs the proxy analyses or it should be removed.*

A14: The text on p. 25838-25839 has been modified, and too specific discussion has been removed.

R15: *There needs to be a more extensive discussion of uncertainties and sensitivities (as mentioned above). What are the uncertainties of each of the satellite measurements that feed into the proxies? How do these uncertainties translate into the overall proxy uncertainty through error propagation? The regional proxy could be more sensitive to errors in AOD because it's a higher-order term – is that what dominates the overall proxy uncertainty or is it UV or SO2? What kind of measurement precision or accuracy is needed from next-generation satellite sensors in order to achieve reasonable proxies?*

A15: This is an important point. According to Tanskanen et al. 2006 the OMI UV irradiance uncertainty is about 7%, but can increase during some episodic aerosol plumes up to 20 %. The OMI NO_2 tropospheric column uncertainty has been resported as ~ 0.75 x 10^{15} molec./cm^2 (Boersma et al., 2011, Bucsela et al., 2013). Over South Africa this would convert to about 7 -25 % uncertainty, depending on how high the NO_2

column values at each location are. For OMI PBL SO_2 the uncertainty for one column observation is very high, the noise reported by Krotkov et al. (2008) can be even 1.5 DU, but when averaging over longer time period, and/or larger spatial area the noise can be reduced to 0.3-0.6 DU. Over South Africa this means e.g. that over the background areas the observed SO_2 column densities can be about the same magnitude as the noise, and over hot spots the uncertainty can be abt. 60% for a single observation, and abt. 20% when averaged over longer time period. The AOD uncertainty according to Levy et al. is 0.05+15%, which means that the relative uncertainty for AOD=0.1 would be 65%, and for AOD=0.25 35%. Hence, over background areas where both AOD and SO_2 are low, the satellite-based SO_2 UV/AOD² -proxy can have an uncertainty of over 90%. On the other hand, over source areas where both NO_2 and AOD are slightly elevated the NO_2/AOD proxy would have an uncertainty of ~50%. Overall over South Africa the uncertainty in satellite-based proxies is high. Over areas where e.g. both NO_2 and AOD are elevated, the relative uncertainty in the satellite-based proxy values would be expected to be somewhat lower than in this study.

We have added this discussion to the text.

Minor Comments:

R16: *Tables 2 and 3 and all inline text: Report as R2 instead of R. Remove p values. Add scatter plots w/ regression lines for each correlation coefficient to the supplementary material.*

A16: R2 is reported instead of R, and p values are removed as suggested. Scatter plots with regression lines have been added to the supplementary material.

R17: *Figure 3: Add a histogram plot for each showing the relative error centered about the regression line, since the log-log plot makes it hard to see how the points are distributed about the regression line. Report as R2 instead of R. Also, is there really only one significant figure in the regression pre-exponential constant?*

A17: We have reported R^2 instead of R, but we did not quite understand what was meant by the histogram plot. We feel that the figure, as it is now, is informative enough and the needed differences can be seen on the log-log-scale too.

R18: *Figure 4: Almost all of the points fall below AOD=0.8 and sigma=200. Please rescale the axes so that this is more clear. Report as R2 instead of R.*

A18: Changed as suggested.

R19: *Figure 5: Again, almost all the points are less than AOD=0.5 now. Please rescale the figures. Are there points not shown that would skew the regression line in the Botsalano panel? I wouldn't know how the LSQ curve would diverge from the visible data without some extreme outliers. Also, add R^2 values to each of the regression lines and consider if it makes any sense to report equation coefficients for regression lines that do not, at a minimum, explain a majority of the variance.*

A19: There are no points that skew the regression line in the Botsalano panel. Now that the axes are rescaled, the scatter of the points can be seen more clearly. R^2-values have been added to the regression lines, and the equations have been removed.

R20: *Figure 6: Add interquartile ranges to each median profile. For the low MODIS AOD cases, why are there no points above 3.3 km?*

A20: The interquartile ranges have been added to the profiles as suggested. To avoid the figure becoming too busy, the calipso profiles were vertically averaged into 200 m height bins. The points in the red profile above 3.3 km are missing since they did not have the required quality control flag.

R21: *Figure 7: Emphasize in the caption that the SO2 density is only in the PBL, while the NO2 column*

density is the entire troposphere, and the AOD is presumably over the entire atmospheric column. The caption does not indicate that these are not all over the same vertical scale. Also, please list the locations that the points correspond to.

A21: The OMI SO_2 Planetary Boundary Layer (PBL) product is a retrieval of SO_2 total column density throughout the whole atmosphere, and ”PBL” refers only to the *a priori* profile assumed in that product (the *a priori* profile has SO_2 predominately in the PBL). This has been now clariefied in the text (in the caption as well as Sect. 2). The names of the in situ measurement stations have been added to the caption.

R22: *Figure 10: Report as R2 instead of R. Remove p-value.*

A22: Changed as suggested.