Response to Anonymous Referee #4

We would like to thank Referee #4 for his/her comments. We have done our best to address each of the points as detailed below.

#1. *There is one area where I wonder if Tesche has misunderstood, and the authors could reword to make sure others don't have the same issue, which is the Cattrall et al. (2005) reference. On first reading, I also thought this was wrong. The authors reference Cattrall et al. (2005) in the context of direct measurements of lidar ratio, while the study is not direct measurements but models based on AERONET inversions. However Cattrall et al. (2005) do also include a table of results from field campaigns which were (I think) direct measurements. I think that table is what the authors mean by 'direct measurements', not the Cattrall analysis itself. So the authors should be careful how they reference this study, to make sure people don't get the wrong impression.*

It was/is our intention to reference the table in the study published by Cattrall et al. (2005). We have tried our best to clarify the manuscript so that readers will not have the same confusion. We modified the introduction of the manuscript to say:

"To date, experimental techniques for directly measuring the lidar ratio include the use of High Spectral Resolution Lidar (HSRL, Eloranta, 2005; Hair et al., 2008) and Raman Lidar (RL, Ansmann et al., 1990). These instruments are capable of measuring aerosol backscatter and extinction parameters independently and therefore do not require the lidar ratio to be prescribed (e.g., Shipley et al., 1983; Grund and Eloranta, 1991; Piironen and Eloranta, 1994; Müller et al., 2007; Amiridis et al., 2009; Tesche et al., 2009a,b; Burton et al., 2012). On the other hand, Cattrall et al. (2005) use AERONET size distributions inverted from sun photometer data (Holben et al., 1998) to calculate the lidar ratio and then compare their indirect to literature reported direct measurements. They determined that their indirect method (28±5) compared well to the literature average of direct retrievals (29±5) (see Tables 3 and 4 in Cattrall et al., 2005). Direct measurements do not suffer the same limitations as indirect ones which require assumptions on size distribution and chemical composition or a molecular extinction profile. The supplementary Table S1 summarises available retrieval methods and values of some experimentally determined lidar ratios over marine regions. Currently, most lidars do not yet have Raman or high spectral resolution capability and CALIPSO is the only lidar that provides aerosol data at the vast spatiotemporal resolution required for global climate model comparison."

#2. *I also feel it's important to point out the limitations of the AERONET estimates of lidar ratio; it is a calculated value based on an inversion (i.e. a retrieval result). It is sensitive to both coarse mode size but also fine/coarse partition, so there are several effects to disentangle. This links into points Tesche raises about error and the conclusions of the analysis. I have some more suggestions about this below. But, as Tesche notes, I think it's important to be more explicit about the limitations of these indirect measurements, as others may not be aware of them.*

Here we address the reviewer's concerns over the community's understanding of indirect and direct retrievals of the lidar ratio. We modified the manuscript to explicitly point out the difference of indirect and direct lidar ratio retrievals by including a short comment on the issue in the introduction (part of the response to the reviewer's first concern). However, we would like to point out the table in the supplementary information as a reference for the different methods of lidar ratio retrievals. The reviewer's concern on limitations like derived coarse and fine mode partitioning are addressed there in the table footer.

#3. *For the SODA data, based on the HSRL comparison, the authors state that for AOD > 0.05 the lidar ratio uncertainty should be <50%. What fraction of the data used for the main study are for AOD < 0.05? From Figure 3 it is hard to tell because these are scatter plots not scatter density plots, but it looks like there are a lot of points there, especially at lower wind speeds. And from Figure 6 I think saying 'Additionally, Fig. 6b illustrates that the relative uncertainty in the SODA retrieved Sp is less than 50% for AODs > 0.05' is misleading because it could be understood as an upper bound; it might be better to say a 'typical uncertainty of order 50%' because there are a whole bunch of points with errors approaching 100% for AOD in the range 0.05-0.1.*

To answer the reviewer's question, "*What fraction of the data used for the main study are for AOD < 0.05?*", we determined that there are a total of 278 out of 13,481 (2.0%) AOD < 0.05 instances in this study, with majority of these instances occurring at the low wind speed regime. Furthermore, we agree that the statement we made previously on the 50% error could be misleading. We have therefore modified the text to both include the number of retrievals below an AOD of 0.05 as well as reconcile any ambiguity on the errors we expect from the HSRL vs. SODA comparison. We also redrew Fig. 3 as a scatter density plot with reported R^2 values. The manuscript now reads as follows:

"Additionally, Fig. 6b illustrates that the relative uncertainty in the SODA retrieved S_p is typically below 50% for AODs > 0.05. In our study, the bulk of AODs measured by SODA (98%) exceed this value under the quality control criteria discussed in Sec. 2.4."

#4. *I partially agree with Tesche's comments on the scatter plot of SODA AOD vs. HSRL. The authors should color-code points in some way to indicate where they were taken, and/or the aerosol type. I would imagine that the type-dependence of SODA AOD validation is weak but we don't know unless we can see the data. Also, as well as R2, the*

RMS error and also any bias would more useful (as they will alias into scatter and bias in the lidar ratios too).

The data for the scatter plot of SODA AOD vs. HSRL are all taken near the eastern coast off of North Carolina. The details on the location for these data have been reported in Josset et al. (2011). We cannot, unfortunately, give more quantitative information on this dataset as this is a contributed dataset (directly from Josset et al., 2011) and only includes what has already been reported. The best that can be said is that the majority of these data is of marine origin and is directly comparable with HSRL measurements collocated in space and time with the CALIOP overpass. The specific aerosol classes are explicitly called marine and pollution or marine and dust (Table 1, Josset et al., 2011). For the second part of this comment, we have reported the RMSE of 0.04.

#5. *In their response the authors point out only 13 coincidences with the Maritime Aerosol Network and say this is not enough to validate the SODA method. I think this data would still be valuable to include, even if only 13 points, as another point of comparison. In fact the small data volume might be an advantage in some ways because if there is some point which matches very well or very poorly it is easy to identify and dig into the data to figure out why. This would be a good backup to the current evaluation against HSRL.*

As the reviewer requested, we have added a SODA comparison with the Maritime Aerosol Network. We decided to do a comparison with the MAN dataset since we agree with the reviewer that it could potentially illustrate some inconsistencies between the two retrieval methods. We began by employing a modified collocation scheme from Smirnov et al. (2011) and Kleidman et al. (2010). We required that the SODA retrieval be within ± 30 minutes of the MAN retrieval as well as within a circle with radius of 25 km around the MAN measurement. A map of the retrieval locations as well as 50 km (enlarged for visibility; red) circles is shown below. Figure 1 includes an enlarged inset to see the satellite track (dashed line) that falls within the MAN circle (red). There were 51 matching locations that passed the collocation screening. Since the satellite passing will match with MAN measurements that are within ± 30 minutes, there are usually multiple MAN measurements within approximately the same vicinity that will compare to one SODA retrieval. So, error bars on Fig. 2 indicate the maximum and minimum values of the MAN AOD to illustrate the range for the closest SODA retrieval. For all SODA retrievals, the variance in the reported AOD was less than 0.02 and so the closest point was used to compare with MAN data. The scatter plot of MAN and SODA comparison is also shown below with the 1:1 relation. The majority of the points fall reasonably close to the 1:1 line. The correlation is 0.59 and the RMS error is 0.03. There appears to be a low bias of SODA with respect to MAN, but it is still considered good agreement. However, we do note, that a negative bias in AOD would alias into the lidar ratio.

We think that this comparison is useful for the manuscript, but adds a level of complexity that we would prefer remain in the supplement. Therefore, the following two figures have been moved to the supplement and consolidated to Fig. S3. We have also included some discussion on the colocation scheme and maritime aerosol network itself.

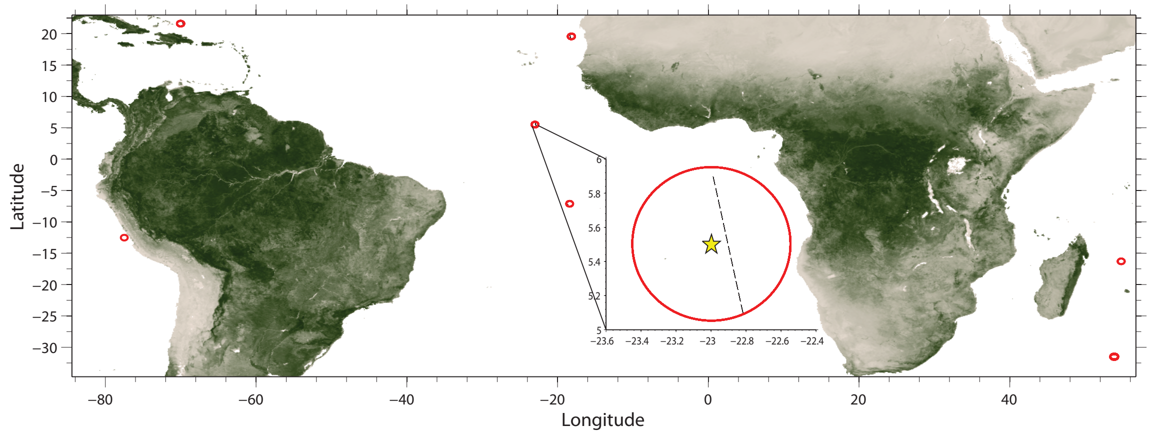

Figure 1 – Map of collocated SODA and MAN measurements for the entire measurement period (December 2007 – November 2010). The red circle indicates the region that a SODA retrieval must reside. The MAN observation location is shown with a yellow star and the SODA track is indicated by a black dashed line.

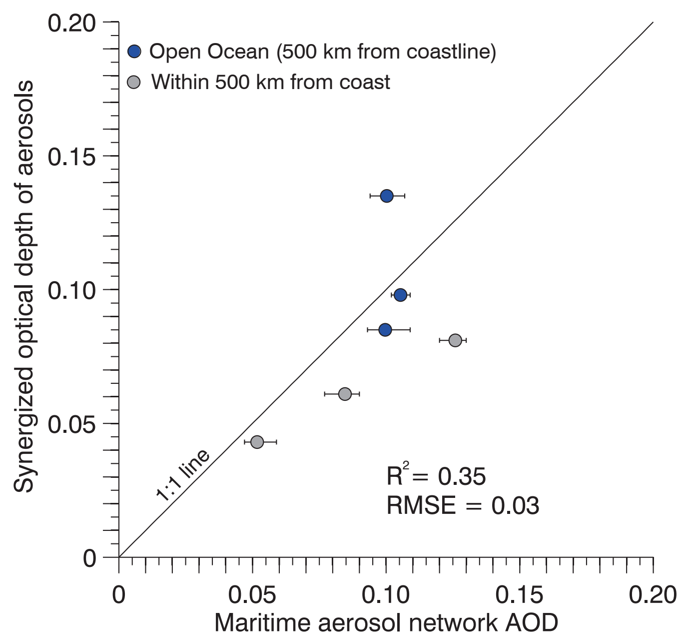

Figure 2 – Scatter plot comparison of MAN and SODA retrieved aerosol optical depth for the collocated regions shown in Fig. 1. The error bars indicate the minimum and maximum values for the MAN retrievals in the ± 30 minute SODA window. R^2 is 0.35 and the RMSE is 0.03. Blue points indicate retrievals made at least 500 km from the nearest coastline.

#6. *As the reviewers and authors note, there are some spurious regions in Figures 1 and 2 (e.g. Asian coastlines, but also off the west coast of South America, which might be sulfate from mines, and Australia, in some seasons) where the AOD and lidar ratio are elevated, which are probably not marine aerosol cases. I suspect this will affect some things like the PDFs of lidar ratio in Figure 4, and the whole analysis based on wind speed. I therefore suggest that the authors make some further data cut to be more confident that they are looking at marine aerosols, and redo the relevant later analyses based on that. Perhaps take some large boxes over parts of the open oceans, and exclude the data in these coastal regions. That would make the results more convincing, and if these coastal cases were spurious non-marine data, I expect it would bring out any marine signal in the remainder more clearly.*

We assume the reviewer notices these spurious regions from Fig. 2. We agree that these regions appear as outliers, however, we also note that there are very few instances in which these regions report valid retrievals (see supplementary Fig. S1). To minimize the confusion we decided take this chance to reiterate our strict quality control criteria that is already set in place and instead reference our supplement for the spatial location of the frequency of retrievals (see manuscript for details). While it is true that these data near the regions where the reviewer is concerned could be contamination, it is also true that these data exhibit depolarization ratios representative of marine aerosol, are located with high retrieval SNR (due to nighttime data as well as 70% shot-to-shot data in the 5 km average) and are located 2 km and lower in a single aerosol layer, as reported by CALIOP feature classification flags. Referencing supplementary Fig. S1, we also point out that the number of retrievals for these regions will not significantly impact the PDFs or the wind speed analysis, as the lidar ratios in question are in the tails of the distribution and our wind speed analysis focuses on the mode of each PDF, not the distribution as a whole. For these reasons, we feel the suspect regions noticed by the reviewer should be noted in the manuscript, but not excluded from our analysis. As such, no figures or data has been altered to exclude these points. We finally note that during a revision, the 2 km layer top screening criteria was inadvertently omitted and has been reinserted into section 2.4.

#7. *Figure 3: This should be redrawn as a scatter density plot for clarity. Also, the least-squares results should be removed because this technique is not mathematically appropriate for finding regression relationships for this type of data (even though people do it a lot). This is because errors are not Gaussian (in low-AOD conditions the error distribution is truncated because AOD cannot be negative), because both datasets have non-negligible error (least squares regression assumes the x-axis data are 'correct'), and because the error is AOD-dependent (so treating all points equally is not appropriate). Just plotting a 1:1 line over a scatter density histogram would be clearer to see the main point of the figure, and less statistically problematic.*

We agree with the reviewer that the figure should be redrawn as a scatter density plot. The same data has been reformatted into aerosol optical depth bins of 0.01 and color coded by frequency (see Fig. 3 below). Also, the least squares slope line has been removed. The resulting figure follows and has been modified in the text

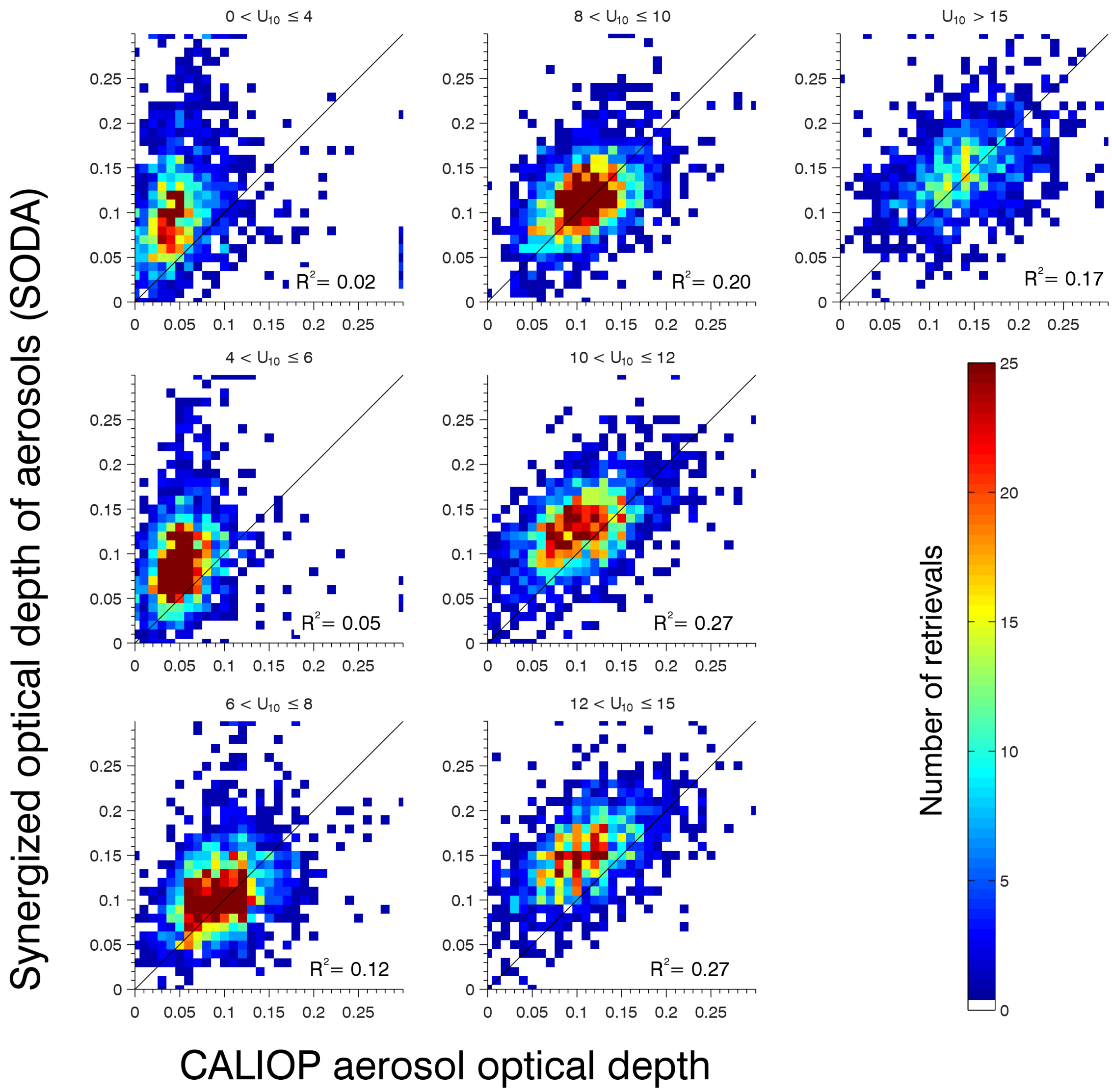

Figure 3 - Scatter density plot for grid cell median aerosol optical depth from SODA (y-axis) and CALIOP (x-axis).

#8. *Figure S2: This should really be in the paper, not in the supplement. People often don't read supplements, and it is a plot on which a lot of discussion rests. I expect it may look a bit different if the coastal regions are excluded, but really it highlights some of the issues Tesche and I have. For example, the wind speed distribution has very few low or high values (see also Table 2) and yet these are the wind bins which will have the strongest influence on the linear fit because they are the extrema. And, especially for wind speeds < 5 m/s, many of the bin medians and means lie quite far from the 95% confidence interval of the linear fit. This is clear evidence that the linear model or error estimate on it are not appropriate (perhaps the true model is not linear, the data are not independent of each other, error characteristics of the data are different between low-wind and high-wind conditions, etc). One way to (partially) deal with this would be to* *chop the data into e.g. 10 equally-populated bins, and do the regression along that wind speed range.*

Further, one can draw a roughly flat line through this shaded confidence interval from 0 m/s to about 10 m/s, or about 5 m/s to 15 m/s; either way this is saying that any real change in lidar ratio with wind speed over this data range cannot be distinguished from zero with 95% confidence. So the authors should be careful not to overstate their finding. Perhaps excluding the coastal outliers will help with this.

We would like to avoid putting the parameterization in the manuscript. We feel there is not enough evidence for an actual wind speed parameterization to be gleaned from this data. We understand the reviewer's point that the linear fit is influenced heavily by the extrema, but without the extrema, the fit is indeterminable with negligible slope. For this reason, we refrain from the parameterization being in the main manuscript as the point we are trying to get across is the utility of a new method to retrieve a lidar ratio for marine aerosol. The main points are that the method produces results in the range of other investigators and is useful globally by not being confined to low spatial resolution in-situ data. The wind speed dependence is not the main focus of the current paper. We have revised the text of the manuscript briefly to reiterate this. For clarity, our wind speed analysis highlights if nothing else, the need for better satellite retrievals for low aerosol optical depth and low wind speeds (to reduce possible systematic error) and/or better understanding of marine aerosol properties (optically, physically and chemically, see last paragraph section 3.2).

#9. *I am also curious, if the authors attempt a direct correlation between raw (i.e. unbinned) lidar ratio and wind speed, what is the value of R^2? If it is very low then perhaps it would be better to look at one or two parameters which are explaining a higher proportion of the variance in the lidar ratio instead (this kind of echoes Tesche's statement about the utility of this part of the analysis). Or at least make an estimate of the variance in the lidar ratio caused by their retrieval uncertainty, if possible how much is random vs. systematic, and how this compares to the wind-induced variability. I am not yet convinced that the results are not entirely spurious here. For example higher wind means higher AOD (on average), and so lower uncertainty on lidar ratio. This may in part be the cause of the wide distribution in lidar ratio in Figures 4 and S2 in low-wind conditions. So if a lot of the error in low-wind conditions is systematic and not random, it is possible all we are seeing in this wind analysis is the difference between some positive bias in low-wind conditions, and smaller errors in high-wind conditions. Or it could be that coastal regions with lower wind speed also have contamination from non-marine (continental) aerosol sources which have higher lidar ratios. Or it could be a combination of these factors. It seems like a bit of a waste if the authors do not dig into this more deeply.*

Here, the reviewer mainly appears to be skeptical that there is any discernable correlation between the lidar ratio and wind speed, and therefore any value of the parameterization shown in supplementary Fig. S2. The reviewer's first request on the value of R^2 for the entire dataset on lidar ratio and wind speed is unfortunately difficult to answer since the dataset used in the analysis has been binned and analyzed as histograms. However, bivariate histograms were used to look at the relationship of lidar ratio to wind speed and there appears to be little to no relationship between 5 and 15 m/s (slope < 0.005) as the reviewer suspected. It is clear that the variability in the lidar ratio with the available parameters in our analysis is controlled by integrated attenuated backscatter, aerosol optical depth, layer top, and layer thickness, respectively. Other parameters have also been investigated and appear to contribute less variation to the lidar ratio so errors in the retrieval of integrated attenuated backscatter and aerosol optical depth could be the main driver of the observed variability. We feel the bivariate histograms discussed are outside the scope of this paper, however. These histograms attempt to draw comparison between microphysical processes and their impacts on the lidar ratio. Unfortunately, this is currently unrealistic since any systematic error in integrated attenuated backscatter and aerosol optical depth cannot be decoupled from wind or other parameter dependencies. An example of some of the bivariate histograms we looked at is shown below but is not included in the revised manuscript. We acknowledge the reviewer's concern, but we feel we have exhausted all the resources available to show the robustness of our results. We therefore return to our data screening procedures and error estimates on aerosol optical depth and integrated attenuated backscatter as reassurance.

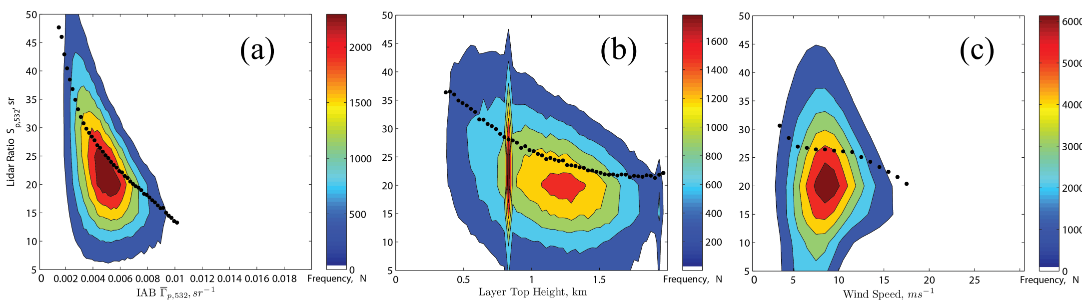

Figure 4 - Bivariate histograms comparing derived lidar ratio with integrated attenuated backscatter (a), layer top (b) and wind speed (c). Means are shown as filled black circles.

References

Smirnov, A., Holben, B. N., Giles, D. M., Slutsker, I., O'Neill, N. T., Eck, T. F., Macke, A., Croot, P., Courcoux, Y., Sakerin, S. M., Smyth, T. J., Zielinski, T., Zibordi, G., Goes, J. I., Harvey, M. J., Quinn, P. K., Nelson, N. B., Radionov, V. F., Duarte, C. M., Losno, R., Sciare, J., Voss, K. J., Kinne, S., Nalli, N. R., Joseph, E., Krishna Moorthy, K., Covert, D. S., Gulev, S. K., Milinevsky, G., Larouche, P., Belanger, S., Horne, E., Chin, M., Remer, L. A., Kahn, R. A., Reid, J. S., Schulz, M., Heald, C. L., Zhang, J., Lapina, K., Kleidman, R. G., Griesfeller, J., Gaitley, B. J., Tan, Q., and Diehl, T. L.: Maritime aerosol network as a component of AERONET – first results and comparison with global aerosol models and satellite retrievals, *Atmos. Meas. Tech.*, 4, 583-597, doi:10.5194/amt-4-583-2011, 2011.

Kleidman, R.G., Smirnov, A., Levy, R.C., Mattoo, S., Tanre, D.: Evaluation and Wind Speed Dependence of MODIS Aerosol Retrievals Over Open Ocean, *IEEE T Geosci Remote*, vol.50, no.2, pp.429,435, Feb. 2012, doi:10.1109/TGRS.2011.2162073, 2010.

Response to Anonymous Referee #5

We thank the referee for his/her commentary and have addressed the concerns (in italics) below.

#1. *Method section should include section 4 about uncertainties.*

We feel that in the methods section we have listed the steps necessary for others to successfully reproduce the lidar ratios. Section 4 discusses a post-lidar ratio error analysis in which the error on the integrated attenuated backscatter is estimated. This section's aim is to provide information on the uncertainty with the retrievals throughout the study, and is not necessary to reproduce our calculations. For this reason, it is our opinion that the methods section should not contain this post measurement error analysis. We would like to refrain from changing the structure in the suggested way.

#2. *Do not discuss wind effect on AOD when describing fig.1. I would recommend to rather focus on the first result of the paper related to potential biases between CALIOP retrieved AOD and SODA consistent with previous work (e.g. MODIS).*

The discussion of the wind effect in our Fig. 1 acts as a transition to the next section on the examination of a potential wind dependence. We would prefer to leave the text as is; however, we would like to contribute more to the validation of SODA to CALIOP AOD comparisons in our supplementary information (in an effort to address the reviewer's concern but preserve the focus of our manuscript). In the supplement a following write up is included to address the reviewer's comment.

"These plots show the distribution of errors between SODA and CALIOP aerosol optical depth retrievals. We show scatter density plots in Fig. S4 (a) along with the distribution of errors in Fig. S4 (b-c). These plots compare the grid cell medians for all of the data used in the manuscript and total 13,481 points. The distribution of errors show that the RMS error is 0.06. Redemann et al. (2012) stated that for RMS error < 0.1, the combination of instruments can be used to obtain further information on aerosol optical properties. The distribution of errors is evaluated to obtain the median error of 47%. The CALIOP retrieved AOD is directly a function of the prescribed lidar ratio and is a major contributor to the bias shown in Fig. S4 (a). An increase in the prescribed lidar ratio would mitigate some of this bias."

#3. *After, I would recommend showing a scatter plot of CALIOP and SODA's AOD to highlight this difference without segregating the data according to wind regimes.*

We have plotted the data as suggested by the reviewer (see the Fig. 1 below). According to Fig. S4 (a), there is a consistent positive bias of SODA relative to CALIOP. This is expected since CALIOP prescribed lidar ratio of 20 sr for marine regions will effectively always inhibit the aerosol optical depth. Even so, the RMS difference is 0.06 which points to evidence that the SODA product can be used in conjunction with CALIOP to retrieve lidar ratios as RMS differences below 0.1 are generally accepted for quantitatively combining products to get additional information on aerosol features (Redemann et al., 2012). The distribution of relative error for the entire dataset is shown in Fig. S4 (b-c).

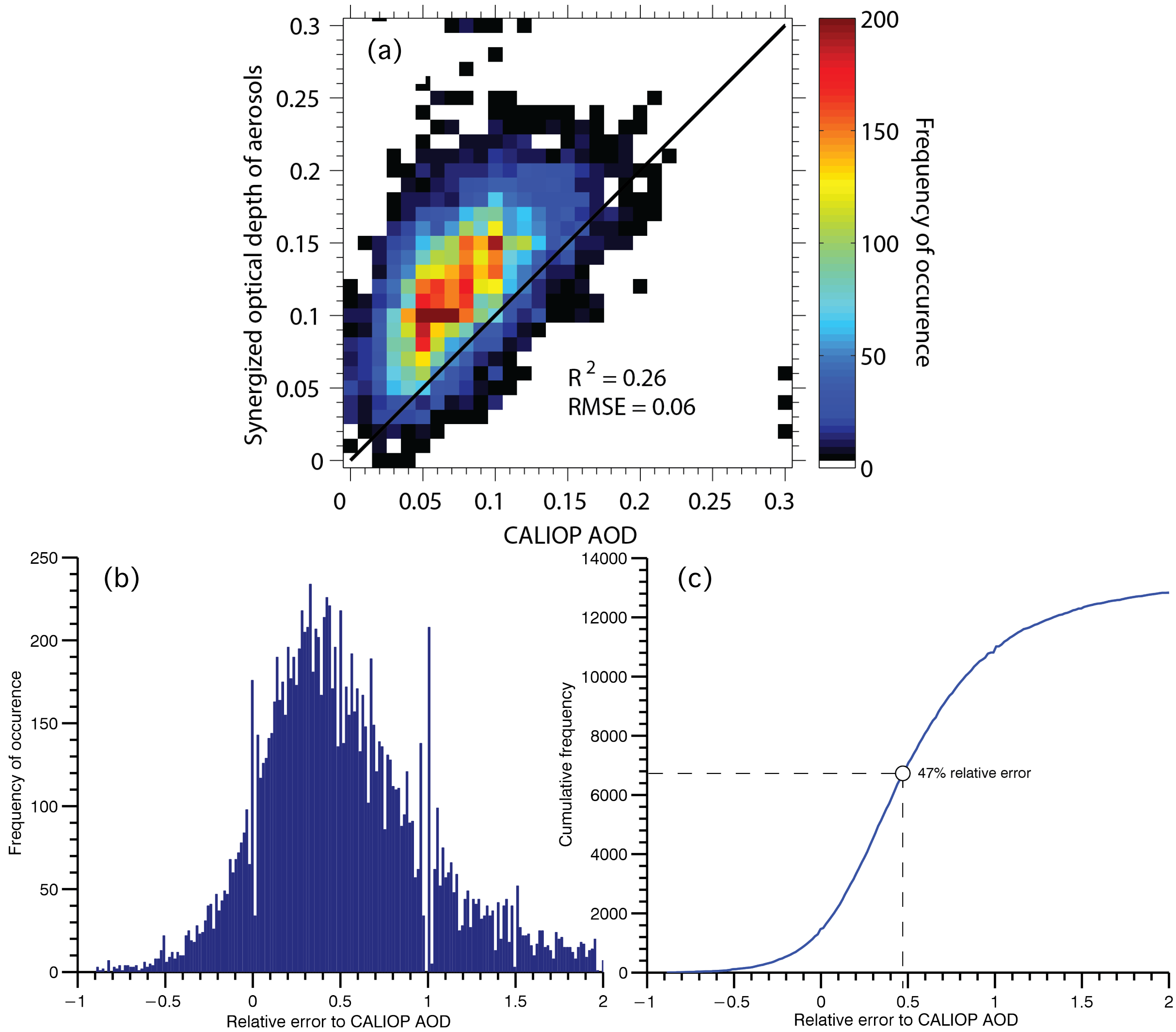

Fig. S4. (a) Scatter density plot of all available (13,481 occurences) SODA to CALIOP aerosol optical depth data. Each point indicates a grid cell median from the spatial maps shown in the main manuscript. The solid black line is the 1:1 relation. The R2 value is 0.26 and the RMS error is 0.06. (b) Histogram of the relative error of SODA compared to CALIOP for each of the points indicated in (a). (c) Cumulative error with the median value reported at 47%.

#4. *Following this part, seasonal maps of lidar ratio from CALIOP and SODA could be showed.*

Our seasonal maps of lidar ratio for SODA (Fig. 2 in the manuscript) are the only useful plot to show in this instance. Due to the type of aerosol scenes used in this study, the CALIOP lidar ratio rarely deviated from 20 sr and can be regarded as a constant value. As such we do not see the need for another plot on CALIOP lidar ratio.

#5. *After that, I would recommend a new section on wind vs. Lidar ratio relationship by showing seasonal maps of surface wind speed which could be compared with lidar ratio* *maps. I propose to show Scatter plots of lidar ratio vs wind speed for different regions (remote places which are likely not influenced by pollution (e.g. Oceans in the Southern Hemisphere) and places which can be influenced by the transport of dust (Spring, Tropical N. Atlantic) and biomass burning aerosol (Summer, Southern Atlantic Ocean) and Pollution (Bay of Bengal, North Eastern Pacific)*

We believe that the scatter plot of lidar ratio versus wind speed is just another way to visualize the PDFs we have already included in the manuscript and another way of interpreting our supplementary Fig. S2. So, we don't really see the increased value in adding figures in this way. We agree with the reviewer that dedicated case studies for the relationship between lidar ratio and wind speed over different parts of the oceans and seasons are needed. However, such studies are outside the scope of the current paper. Moreover, looking arbitrarily at regions could inadvertently bias the results of the study. Effects of different atmospheric processes as well as ocean physicochemical/biological composition on optical properties of marine aerosol are highly uncertain and it would be incorrect to attribute to location alone.

#6. *The second important factor which could eventually influence the lidar ratio of marine aerosol and not mentioned in the paper would be the relative humidity. Trying to find RH data from ship measurements could be interesting. The third factor influencing marine aerosol lidar ratio would be the production of sulfate precursors (e.g. DMS) which could then mixed up with sea salt aerosol and modify the lidar ratio. Satellite data of chlorophyll might be interesting to investigate that.*

Please see our response to comment #5.

References:

Redemann, J., Vaughan, M. A., Zhang, Q., Shinozuka, Y., Russell, P. B., Livingston, J. M., Kacenelenbogen, M., and Remer, L. A.: The comparison of MODIS-Aqua (C5) and CALIOP (V2 & V3) aerosol optical depth, Atmos. Chem. Phys., 12, 3025-3043, doi:10.5194/acp-12-3025-2012, 2012.