Responses to Referee #1's Comments

Referee's comment: *This paper classifies dominant aerosol types at the Anmyon AERONET site, as well as 6 other AERONET sites. Although the paper is quite readable, I don't believe that it should be published in ACP.*

Response: We thank the referee for carefully reviewing the manuscript and providing valuable comments. Since your comments are serious and significant, we have carefully prepared the following responses.

To begin with, we'd like to mention that we will modify the title as "Identification of column-integrated dominant aerosols under high aerosol optical depth conditions using the data set from a single AERONET site". We will add two phrases, "under high aerosol optical depth conditions" and "using the data set from a single AERONET site". The first phrase is to indicate that we identified column-integrated dominant aerosols when AOD ≥ 0.4. We provided the total occurrence rate, including the low AOD conditions (AOD < 0.4) to compare it with the occurrence rates of dominant aerosols. However, there seems to be confusion, and thus we decided to clarify in the title that we used the data of AOD ≥ 0.4 for dominant aerosols.

The situation is similar for the second phrase. One of our main objectives is to demonstrate that we can obtain good information on dominant aerosols using the AERONET data set just from a single site. As in other works, we also analyzed worldwide AERONET sites which have distinct source characteristics, but our purpose was to evaluate the validity of the results from the Anmyon site.

The following are our responses to your specific comments. Your comments are shown in italics as seen above and are numbered for convenience. Note that in the responses, we used page and line (P and L) numbers in the discussion paper (Choi et al., Atmos. Chem. Phys. Discuss., 13, 26627–26656, 2013) when referring to the original manuscript. We also added P and L numbers in the parentheses to your own page and line numbers so that we as well as you can easily find the location without the original manuscript you had referred to.

1. *My main issue with this paper is that the methods section does not have nearly enough details. The authors results cannot be reproduced on the basis of what is presented here. How do the authors classify as MD, MD+ carbon, or mixed coarse when FMF >0.5? This is not described anywhere.*

We separated MD, MD+ carbon, and mixed coarse particles from coarse mode aerosols (FMVF ≤ 0.5) and secondary inorganic ions, BC, and OC from fine mode aerosols (FMVF > 0.5). Here, FMVF is the fine mode volume fraction (please note the response to comment 17). All these were done using the cluster analysis as described in the caption of Fig. 2 on P26652 and P26633, L21-26. However, as you mention several times in your comments, our description on the use of the cluster analysis was not sufficient. As a result, the description to explain the procedure of the cluster analysis (P26633, L21-28) will be extended substantially as follows:

"(5) Types of absorbing aerosols in fine and coarse modes were distinguished, respectively, using the K-means clustering method with parameters of AOD_{440}, SSA_{440}, $AAE_{440\text{-}1020}$, and FMVF (Table 1). Note that AOD was used for classifying aerosol types although it depends on aerosol amount rather than type. This is different from other works in Table 1 except Omar et al. (2005) who used coarse and fine mode volume concentrations. We used AOD for a cluster analysis since we obtained the most plausible results with AOD, which means that

aerosol types are not completely separable with their amounts.

K-means clustering is a method to partition a data set into the prescribed number, *K* groups (Jain, 2010). To determine the best *K*, we coupled the K-means clustering with discriminant analysis which classifies the data into the given clusters (Romesburg, 2004; Aczel and Sounderpandian, 2009). We performed the K-means clustering with different values of *K* and chose the best *K* by examining the hit ratio, which is a measure of how correctly the discriminant analysis classifies the data set into the same groups given by the cluster analysis. When the hit ratio is 100%, the classification by the discriminant analysis completely coincides with that from the clustering method.

In this work, the commercial software code SPSS (Version 12.0; http://www.spss.com) was used for clustering and discriminant analysis. For the coarse mode, the hit ratios were 98.4%, 99.5%, 95.9% and 96.4% when *K* was varied from 2 to 5. As a result, we selected $K = 3$ and the clusters were designated as mineral dust (MD), MD+carbon (MD mixed with carbon), and mixed coarse particles by comparing their properties with those from the previous works, which will be discussed later. For the fine mode, the hit ratios were 100%, 98.9%, and 97.9% when *K* was varied from 2 to 4. We selected $K = 2$ and the clusters were designated as BC and OC. The discriminant analysis was also used to classify dominant aerosols at worldwide AERONET sites into the aerosol types obtained from the Anmyon site."

Aczel. A. D., Sounderpandian, J.: Complete Business Statistics, 7/e, McGraw-Hill, ISBN: 0073373605, 2009.

Jain, A.: Data clustering: 50 years beyond K-means, Pattern Recognition Letters, 31, 651-666, 2010.

2. *SSA > 0.95 has to be secondary ions with their scheme, but what are the authors including in this set? Anmyon is close to the coast –does this include sea salt?*

We thought of ammonium sulfate and nitrate for secondary ions as explained on P.26633, L7-12.

It is true that Anmyon is close to the coast and is surely influenced by sea salt. However, as mentioned on P26633, L16-20, we did not separate sea salt from the coarse mode because AOD is typically less than 0.1 for pure maritime environments (Smirnov et al., 2002) while we tried to identify dominant aerosols when AOD was sufficiently high to obtain valid optical and microphysical properties.

3. *What about secondary organics? What do they mean when they say that BC and OC are the dominant aerosols when FMF < 0.5 and SSA < 0.95?*

Atmospheric aerosols consist of a myriad of species with various types of mixtures. Many elaborated instruments have been used to identify individual species, which require much time, labor, and expense. Estimation of aerosol species from optical measurements resides on the other side. Although not so precise and accurate, it does give information automatically in real time. As seen in Table 1, most research works have provided information on aerosols optically classified (non-absorbing, slightly absorbing, etc., Lee et al., 2010; Logan et al., 2013) or those related to sources (desert dust, biomass burning, etc., Omar et al., 2005; Russell et al., 2010).

In the present work, we obtained information on chemical species such as mineral dust, BC, OC, and secondary ions. These species are probably the maximum number of species available so far, which have been estimated from optical measurements. It is certain that identification of secondary organics could be an interesting topic for future researches.

4. *What can't mixtures of BC with secondary ions produce the same FMF and SSA?*

We are sorry but we are not sure what this question means. If this question means whether mixtures of BC with secondary ions can produce the same FMVF and SSA with varying the mixing ratio, the answer is no. This is because optical properties of BC and secondary ions are quite different, as shown in Table 2.

5. *Basically, the authors have not presented any data to demonstrate that the scheme outlined in Figure 2 works at all. Afterwards, they have a nice discussion of the results, but the method description is so imprecise that the reader cannot conclude anything from the results. The authors need to elaborate the methods section extensively so that others have enough information to reproduce their results, and for the reader to draw meaningful conclusions about the results.*

Please see the response to comment 1.

Major issues

6. *Page 2, line 15 (P26631, L14-19), the authors state: "An automatic tracking sun and sky-scanning radiometer, CE 318 (CIMEL Electronique; also called a sunphotometer) measures direct radiation on the principal plane (with fixed azimuth angle and varied zenith angle) and diffusive radiation on the almucantar plane (with fixed zenith angle and varied azimuth angle up to 180 in both sides) at 8 wavelength channels (340, 380, 440, 500, 675, 870, 940 and 1020 nm) (Holben et al., 2001)." This is not correct, as direct radiation is NOT measured along the principle plane. Both scan modes (almucantar and principle plane) measure the diffuse field. "Direct" radiation is obtained only when the instrument is pointing at the sun. Also, only 4 channels are measured in the almucantar (not 8). Brent Holben certainly knows this.*

We will correct the paragraph (P26631, L14-26) as follow: "An automatic tracking sun and sky-scanning radiometer, CE 318 (CIMEL Electronique; also called a sunphotometer) measures direct and diffuse radiation at AERONET sites (Holben et al., 1998). Direct radiation is measured with a 1.2° full field of view every 15 min at 340, 380, 440, 500, 675, 870, 940, and 1020 nm (nominal wavelengths), and aerosol optical depth (AOD) is retrieved at all wavelengths, except at 940 nm which is used for retrieving column water vapor (Holben et al., 2001; Eck et al., 2010). Diffuse radiation is measured on the principal plane (with fixed azimuth angle and varied zenith angle) and the almucantar plane (with fixed zenith angle and varied azimuth angle up to 180^{o} in both sides) using 4 wavelength channels (440, 675, 870, and 1020 nm) (Holben et al., 2001). Diffuse radiation in the almucantar geometry is measured at optical air masses of 4, 3, 2, and 1.7 both in the morning and afternoon, and once per hour in between (Eck et al., 2010). Using almucantar measurements, volume size distribution and complex refractive indices are determined by comparing AOD with that from direct radiation measurements, and other parameters, such as SSA, are retrieved (Dubovik et al., 2000, 2006)."

7. *Page 7, line 21 (P26633, L21): This clustering scheme is the meat of your paper, but you're glossing over it with a single paragraph and referencing a 258 page book. What criteria go into this black box???*

Please see the response to comment 1.

8. *Page 9, line 1 (P26635, L8): "BC and OC do not show significant differences in AOD, SSA and FMF."... Again, what goes into the clustering, though? The authors are classifying everything with SSA < 0.95 and FMF > 0.5 as BC/OC dominant, so it is not too surprising*

that you don't see differences in these parameters. How do you discriminate BC from OC in your scheme? Where is the data to back up this statement (and other statements like this)?

It is true that BC and OC do not show significant differences in AOD, SSA and FMVF in this work as shown in Table 2. However, Table 2 also shows that AAE of BC is obviously lower than OC (P26635, L8-9). The rest of the paragraph (P26635, L9-19) demonstrates that AAEs of BC and OC in Table 2 are in a range similar to the previous works. Regarding the cluster analysis, please refer to the response to comment 1.

9. *Page 10, line 1 (P26636, L1): "SSA of secondary ions is high because of their colors which hinder absorption..." I don't understand this statement.*

That statement was miswritten. It will be replaced with: "SSA of secondary ions is high because their size is similar to incoming radiation, and increases with water vapor by hygroscopic growth (Malm, 1999; Eck et al., 2005)."

Malm, W. C.: Introduction to Visibility, Cooperative Institute for Research in the Atmosphere, NPS Visibility Program, Colorado State University, 1999.

10. *Page 10, line 1 (P26636, L2-3): "...This is confirmed in Fig. 3..." If you want to "confirm" that SSA increases with water vapor, why don't you plot it directly (instead of plotting Reff vs water vapor and SSA vs Reff)? This figure did not confirm anything for me.*

First of all, we will replace the figures showing individual values with those showing means and standard deviations to demonstrate more clearly that SSA of secondary ions is high at high water vapor as follows:

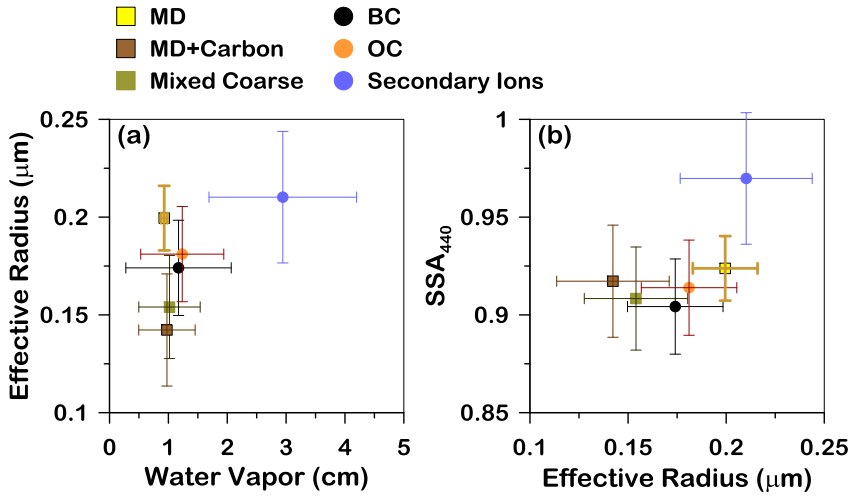
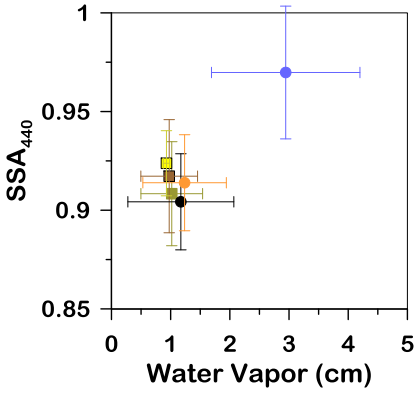


Fig. 3. Plots of (a) fine-mode effective radius vs. column water vapor and (b) SSA_{440} vs. fine-mode effective radius. Symbols and error bars represent means and standard deviations, respectively.

The figure on the right-hand side directly shows the relationship between SSA and water vapor as you suggested. However, we didn't present this figure despite your suggestion. This was because we'd like to show that SSA of secondary ions was high at high water vapor due to hygroscopic growth but the figure on the right-hand side alone does not provide information on the reason for this phenomenon.



11. *Page 10, line 5 (P26636, L4-5): "The effective radius (Reff) is defined as (http://aeronet.gsfc.nasa.gov/new_web/Documents/Inversion_products_V2.pdf)" Effective radius was defined long before AERONET existed: Hansen, J., and L. Travis (1974), Light scattering in planetary atmospheres, Space Sci. Rev., 16, 527–610. I believe that this article cites an even older article for Reff as well. You can also find it in Seinfeld and Pandis, of course.*

Thank you for your suggestion. We replaced an informal AERONET document with Hansen and Travis (1974) and Schuster et al. (2006).

Hansen, J. E., and Travis, L. D.: Light scattering in planetary atmospheres, Space Science Reviews, 16, 527-610, 10.1007/bf00168069, 1974.

12. *Page 11, line 2 (P26637, L2-5): The authors state "Here, we presume that diffuse radiation was measured at one hour intervals. The number of daytime hours was counted on a monthly basis by examining the earliest and latest times at which diffuse radiation was measured." This is a faulty approach. There is a long period in the middle of the day (SZA < 50) where almucantar scans are not included in the level 2 dataset. Thus, the authors are grossly overestimating the number of daytime hours during which diffuse radiation could be measured (at least in the summer months). This once again begs the question – did Brent Holben read this?*

We thank you for your clarification. On the basis of your statement, we recalculated the occurrence rate by introducing the concept of the instrument working hours as follows:

"The occurrence rate was calculated by dividing the occurrence number of the aerosol type by the total number of raw data for diffuse radiation. Here, raw data for diffuse radiation indicate the almucantar raw data. We used those for SZA between 50^o and 80^o from which inversion products were obtained (García et al., 2008). This means that the occurrence rate represents how often dominant aerosols occur during the instrument working hours when the instrument measures diffuse radiation that can be used for inversion products."

García, O. E., Díaz, A. M., Expósito, F. J., Díaz, J. P., Dubovik, O., Dubuisson, P., Roger, J. C., Eck, T. F., Sinyuk, A., Derimian, Y., Dutton, E. G., Schafer, J. S., Holben, B. N., and García, C. A.: Validation of AERONET estimates of atmospheric solar fluxes and aerosol radiative forcing by ground-based broadband measurements, Journal of Geophysical Research: Atmospheres, 113, D21207, 10.1029/2008JD010211, 2008.

The above description will substitute that from P26636, L25 to P26637, L5. Because of this change the occurrence rates (%) at the study sites will increase as follows:

	Original Works		Revised Works	
	Total	Dominant Aerosols	Total	Dominant Aerosols
Anmyon	4.7	1.2	21.1	5.4
Beijing	10.6	6.3	25.7	15.2
Mexico City	2.9	1.1	8.5	3.1
GSFC	11.6	0.9	25.7	2.0
Mongu	8.7	2.2	23.1	5.8
Alta Floresta	2.8	0.8	10.6	2.9
Cape Verde	6.0	2.2	14.4	5.3

Regarding your last question, we had sent the original manuscript to B. Holben as a coauthor before the submission. We have also sent him these responses to the referees' comments along with a revised manuscript before posting them.

13. *Page 12, line 19 (P26638, L19): The authors are using Cape Verde as typical dust, but that location also sees biomass burning in the winter. How do the authors filter out biomass burning?*

We didn't have to filter out the effect of biomass burning because it is not distinguished in the present work. However, since some studies indicated it, we will mention it on P26640, L11 as follows:

"Although some studies indicated that Cape Verde was influenced by biomass burning emissions from the Sahel in winter (Tesche et al., 2009; Kim et al., 2011) and we also observed an increase in the fraction of mixed coarse particles in winter (not shown), such effects are not distinguished in Figs. 5 and 6."

Tesche, M., Ansmann, A., Müller, D., Althausen, D., Engelmann, R., Freudenthaler, V., and Groß, S.: Vertically resolved separation of dust and smoke over Cape Verde using multiwavelength Raman and polarization lidars during Saharan Mineral Dust Experiment 2008, J. Geophys. Res., 114, D13202, doi:10.1029/2009JD011862, 2009.

14. *Page 12, line 24 (P26638, L24-26): The authors state "Heavy aerosol loading at Beijing is confirmed by a high occurrence rate of dominant aerosols, which is 6.3%, the highest among seven sites, more than five times that at Anmyon." So basically, the authors are only able to classify the dominant aerosol type 6.3% of the time at your most favorable site? How is this important? Some discussion would be helpful.*

As mentioned in the response to comment 12, we redefined the occurrence rate; the occurrence rate of dominant aerosols at Beijing becomes 15.2%. Nevertheless, it is apparent that the occurrence rates of dominant aerosols are low mainly because they are distinguished when AOD $\geq$ 0.4. To explain this, we will add the following explanation after the description given in the response to comment 12, with which the two sentences on P26637, L6-8 will be replaced:

"In Fig. 4a, the monthly total of the occurrence rates varies from 28% in March and November to 15% in June and July. It is noticeable that the level of the occurrence rate is low, particularly for dominant aerosols. On the annual basis, the total occurrence rate was 21.2%, which is the sum of 5.4% for dominant aerosols and 15.8% for low AOD aerosols. Annual mean occurrence rate of 5.4% for dominant aerosols is so low but could be plausible in the sense that high AOD conditions of AOD $\geq$ 0.4 are not common (although we distinguished aerosol types under these conditions). However, it is also not high even when including the occurrence rate of low AOD aerosols, that is, without AOD restrictions. This is because the level 2 products which are cloud screened and quality assured are also unusual even in comparison with the instrument working hours."

15. *Page 14, line 16 (P26640, L16) and Figure 5: 14% BC at GSFC is huge! In fact, is looks quite large everywhere in Figure 5, except Cape Verde.*

At GSFC the fraction of BC is high, but only among dominant aerosols (Fig. 4b). The occurrence rate of BC is the lowest among the study sites except Cape Verde where that of BC was minimal (Fig. 4b). On the other hand, at GSFC, the occurrence of secondary ions is the highest. Due to the prevalence of secondary ions, dominant aerosols at GSFC are influenced by secondary ions as described on P26640, L18-23, P26642, L5 (summary and

conclusions), and P26628, L19-20 (abstract).

16. *Table 2 and Figures 4 and 5: What is the timeframe for the data, and how many data points at each site? Is this all available data at all these sites? What years?*

We will add the time frame to the captions of Table 2 and Figs. 4 and 5:

Table 2. Properties of dominant aerosols (mean ± standard deviation) at Anmyon for the period 1999-2007.

Fig. 4. Monthly variations in (a) the occurrence rates of dominant aerosols along with aerosols of low AOD and (b) the fractions of each aerosol type among dominant aerosols at Anmyon for the period 1999-2007.

Fig. 5. (a) The occurrence rates of dominant aerosols along with aerosols of low AOD and (b) the fractions of each aerosol type among dominant aerosols at selected AERONET sites in comparison with those of the Anmyon site. Study periods: Anmyon (1999-2007), Beijing (2001-2011), Mexico City (1999-2010), GSFC (1993-2011), Mongu (1995-2009), Alta Floresta (1993-2010), and Cape Verde (1993-2010).

We will also add on P26638, L21, "Study periods are different by site because we used all available data at each site."

Minor issues

17. *Page 6, line 17 (P26632, L17): I'd avoid using the nomenclature "FMF" to represent the fine fraction of the volume distribution, as it has become customary to use the term FMF for fine mode *optical depth* fraction (as in the MODIS retrievals).*

Thanks for your suggestion. We will use "FMVF (fine mode volume fraction)" instead of FMF to avoid confusion.

18. *Page 7, line 10 (P26633, L10-12): I don't understand... Are you computing the SSA for the fine mode, here, or is this the SSA of both modes when FMF < 0.5?*

In this case, FMVF > 0.5, and SSA is largely from fine particles. On the other hand, SSA of MD in Table 2 is from mineral dust type aerosols, which are coarse particles with little of fine particles (FMVF = 0.08).

19. *Page 8, line 24 (P26634, L23-24): See also Chun (PNAS 2013) and Bahadur (PNAS 2013)*

We will add them to the references as follows: "Cluster analyses of AERONET major sites revealed that AAE of desert dust aerosol was mostly 1.5-2.6 (Russell et al., 2010; Giles et al., 2012; Chung et al., 2012; Bahadur et al., 2012)." We will not mention them separately because their AAE values are 2.2-2.6 and 2.2±0.5, which are within the range already given in the manuscript.

Bahadur, R., Praveen, P. S., Xu, Y., and Ramanathan, V.: Solar absorption by elemental and brown carbon determined from spectral observations, Proceedings of the National Academy of Sciences, 109, 17366-17371, 10.1073/pnas.1205910109, 2012.

Chung, C. E., Ramanathan, V., and Decremer, D.: Observationally constrained estimates of carbonaceous aerosol radiative forcing, Proceedings of the National Academy of Sciences, 109, 11624-11629, 10.1073/pnas.1203707109, 2012.