Author's Response

B.Torres[1,2], O. Dubovik[2], C.Toledano[1], A.Berjon[1,3],
, T.Lapyonok[2], V.E.Cachorro[1], P.Goloub[2].

September 13, 2013

1.- Grupo de Optica Atmosferica, Universidad de Valladolid, Spain

2.- Laboratoire d'Optique Atmospherique, Universite Lille 1, France

3.- Izana Atmospheric Research Center, Spanish Meteorological Agency (AEMET), Tenerife, Spain.

Answer to Editor

Dear Editor,

Please find attached the answer to the two referees and the revised version of the manuscript.

Thank you

Answer to referee #1: T.F.Eck

We would like to thank Referee #1, T.F. Eck, for the detailed review and his very valuable comments. We addressed all reviewers comments in the detailed response below and we made many significant changes in several sections of the manuscript, especially in the second part of the analysis related to AERONET measurement data and in the discussion/conclusions.

POIN-TO-POINT responses:

Referee's comment 1: **First, for the simulation analyses, the authors utilized inversion results of size distributions and refractive indices from the Dubovik et al. (2002) paper that used an older version of the algorithm (particle non-sphericity was not accounted for) and with an erroneous Lambertian surface reflectance assumed (typical of green vegetation regardless of site location). Additionally the input data used in the Dubovik et al. (2002) paper were of poorer quality since interference filter stability was relatively poor prior to 1997 (resulting in larger calibration errors) and also since data quality screening was rudimentary compared to current screening.**

We agree with the reviewer that Dubovik et al.(2002) paper was based on AERONET Version 1 retrieval products and used much smaller volume of the AERONET observation that it is available now. Nonetheless, we think that climatological aerosol models from that are sufficiently reliable for the conducted sensitivity study. First, the authors of 2002 paper were aware of all potential issues of AERONET retrievals and put significant efforts into accounting for the possible uncertainties in the data selection. As the results, even presently no significant shortcoming were identified in the aerosol models of 2002 paper in any known publication. In a contrast, the most recent paper by Giles et al.(2012) that analyzed AERONET Version 2 products indicated that, generally, the values of the considered aerosol properties (in particular SSA) were rather close to those reported in 2002 paper. In our opinion, such results confirm the reliability of the Dubovik et al. (2002) results, in particular taking into account that Giles et al. (2012) used not only higher quality retrieval but also 10 year more of data. It should be noted that Giles et al. results cannot be easily used in our analysis because the paper does not provide full microphysical aerosol models that are available from Dubovik et al. (2002) paper. Finally, we would like to point out that, in the present study, aerosol data from Dubovik et al. (2002) paper were used only for theoretical sensitivity analysis. Then, we have analyzed real AERONET (Version 2) retrieval data to evaluate the effect of observations geometry. The results of that analysis did not reveal any significant discrepancies with the results of numerical tests. Therefore, we do not see any solid basis for questioning the assumptions used in the sensitivity tests.

At the same time, we have acknowledged in the paper the possibilities of some shortcomings in Dubovik et al. (2002) results and cited Giles et al. (2012) more recent climatology analysis.

Referee's comment 2: **These issues and perhaps others may have led at least par-**

tially to relatively poor matching of input AOD versus output AOD in the simulations (simulated output values $\sim 10\%$ higher than input at higher AOD levels in Table 1). Additionally, the lack of consistency in AOD levels for different sites in the simulations hinders an accurate assessment of differences between aerosol types since at higher AOD levels the influence of differences in surface reflectance and particle shape are reduced (i.e. the high AOD case for the GSFC site was 0.5 at 440 nm while for the Mongu site it was 0.8 at 440 nm). Numerous other issues with the simulation analyses are detailed in the "Specific Comments". section below.

In our opinion, the differences between AOD input and output values in Table 1, are unlikely related to the use of a different surface albedo model and other uncertainties. The Dubovik et al. (2002) aerosol models are based on the linear regression analysis. At the same time, AOD of aerosol depends non-linearily on the such aerosol parameters as real and imaginary parts of the refractive index. Therefore, no one can really expect a perfect match between input and output AODs. We think that is the main reason for the observed discrepancies. At the same time, there is a minor dependence of AOD values on the particle shape. We have included a respective discussion in the text.

In addition, we would like to point out that there is sufficient consistency in the AOD levels for the different sites: the strategy followed was to used the average values of the AOD obtained in Dubovik et al. (2002) for the sets GSFC1 ($<\tau(440)>= 0.24$ here $<\tau(440)>= 0.2$), and Zamb1 ($<\tau(440)>= 0.38$ here $<\tau(440)>= 0.4$) and twice these values for the sets GSFC2 ($2<\tau(440)>= 0.48$ here $<\tau(440)>= 0.5$), and Zamb2 ($2<\tau(440)>= 0.76$ here $<\tau(440)>= 0.8$). We agree with the referee that these differences in AOD levels have influence in the analyses and we have pointed out this fact in the manuscript.

<u>Referee's comment 3:</u> **In the analysis of simulations against real AERONET data there are also issues that were either not considered or likely bias the analysis. To begin it is puzzling that the GSFC site analyzed in the prior simulations section was exchanged with the Beijing site while the other two sites remained the same. The AOD levels of the Beijing data analyzed were much higher than at GSFC and for the other sites (see Table 4), resulting in much lower uncertainties at Beijing in retrieved optical parameters and less solar zenith angle dependence due to the much larger aerosol signal at that site.**

Following this suggestion we have replaced analysis of Beijing site data by GSFC site data.

<u>Referee's comment 4:</u> **It is also noted that identifying the Solar Village data as desert dust aerosol is not always accurate as this site more commonly exhibits mixtures of fine (sub-micron radius) and coarse mode (super micron radius) particles. The authors did not realize that the two subsets of data for the Solar Village site (bot-**

tom of Table 4) have significantly different Angstrom Exponent values (more than a factor of two difference) and therefore they are comparing mixed aerosol versus dust-dominated cases in the Desert Dust I and Desert Dust II subsets respectively. Even more importantly, for all sites studied, the data selection criteria (given on page 6879, lines 8-17) results in the chosen dates have relatively stable and homogeneous AOD (std. dev./mean < 0.1; criteria 2), and therefore having conditions that are not necessarily typical and that minimize the potential problems that Principal Plane scans have with inhomogeneous atmospheric conditions and cloud screening. Additional issues and comments regarding the real data analysis section are given below in the "Specific Comments".

We thank the reviewer for pointing on potential reasons for retrieval inconsistencies. In order to address that we have changed some days in the case Desert Dust I reducing its Angstrom Exponent values; nevertheless they are still a bit higher than for the set Desert Dust II and this fact is discussed in Table 4 and commented along the text.

Referee's comment 5: As a result of the problems noted above, I cannot agree with some of the discussion and conclusions (sections 4 and 5) presented by the authors. In the Discussion section (pages 6886-6887) the authors say: "the complete elimination of principle plane retrievals from the provided aerosol product leads to lose of valuable aerosol information, especially taking into account that both our sensitivity tests and real data analysis conducted in this study generally show high consistency between principal plane and almucantar retrievals". I argue that the analysis done in this paper does not accurately account for issues of aerosol inhomogeneity and cloud screening of the principal plane data that would result in both biases and higher random noise in principal plane retrievals relative to almucantar retrievals. Also the authors did not account for the fact that spectral AOD data (also primary input data to the retrievals) is less accurate at smaller solar zenith angles and thus may result in a significant error source for principal plane retrievals made when SZA is small (some sites/seasons at mid-day). Further details regarding issues with the Discussion and Conclusion sections are given below in Specific Comments.

The AERONET project has always placed data accuracy and the quality of products first and foremost in decisions regarding new data base versions and new products. It is for these reasons of retrieval product quality that the decision was made to provide retrievals using only the almucantar scan geometry. The AERONET team at GSFC has long been aware of the issues raised in this paper (and additional issues that were not mentioned in this manuscript) that affect the quality of principal plane retrievals. These issues are continually being analyzed and explored by the GSFC AERONET Team with the goals of providing the most consistent and accurate

database possible from AERONET observations.

Also, following the Referee suggestion we have significantly modified conclusion and discussion section with a particular effort on improving "tone" of the statements.

Answers to the specific questions/comments follow here:

Page 6854, lines 3-5: The AERONET acronym should be all CAPS (not Aeronet, as you have used throughout the paper). Also, the way this sentence is written you imply that PFR-GAW instruments measure the same quantities as AERONET and SKYNET, however the PFR only measure spectral AOD (not sky radiances) and therefore do not retrieve other parameters such as refractive indices.

Aeronet was replaced by AERONET throughout the paper. In order to make a difference between AERONET, SKYNET and PFR-GAW, we have modified the beginning of the following paragraph as follows: *AERONET and SKYNET networks provide the aerosol information from two kinds of spectral measurements: spectral data of direct Sun radiation attenuation by the atmosphere and angular distribution of diffused sky radiation; PFR-GAW only provides data of the direct Sun radiation.*

Page 6855, lines 3-6: Please also reference Holben et al. (2006; SPIE) as this reference also provides more detail on the minimum number of observations required for four different scattering angle range bins. This was an important new quality control check introduced in 2006 for the Version 2 almucantar retrieval products in order to improve retrieval robustness.

We have included the reference and we have modified the whole sentence as follows:

In AERONET network processing, the symmetry property in almucantar has been used for over a decade and those measurements exhibiting radiances differences higher than 20% *between right and left branches are eliminated (this and other quality control criteria are described in* `http://aeronet.gsfc.nasa.gov/new_web/Documents/AERONETcriteria_final1_excerpt.pdf` *and Holben et al. (2006))*

Page 6855, lines 16-17: It should be noted that operationally the maximum scattering angle in the principal plane scan is less than stated here since the observations cannot be made all the way to the horizon due to horizon obstructions and also due to inexact optical airmass computations (refraction effects, etc.). It is for the reason of difficulty of airmass computations at large airmass that the Level 2 AOD product from AERONET is limited to a maximum airmass of 5. At larger airmass values the aerosol vertical profile information is required for accurate computations.

We have added a footnote:

It should be noted that operationally observations cannot be made all the way to the horizon due *to horizon obstructions and also due to inexact optical airmass computations (refraction effects, etc.). Therefore, the maximum scattering angle in the principal plane scan is typically smaller than stated theoretically (E.g. in AERONET, the Level 2 AOD product is limited to a maximum airmass of* 5 *as the effect of athmosphere sphericity is not negligible for larger airmasses and this is not accounted in our plane-parallel radiative transfer model. Therefore,* $\Theta_M \simeq \theta_s + 80°$*).*

Page 6857: Note that the Solar Village climatology from Dubovik et al. (2002) is more representative of mixed aerosols (fine and coarse mode size mixture) than pure "desert dust aerosol" that you state here, with average Angstrom exponent (440-870 nm) ~ 0.6**. See Kim et al. (2011; ACP) and Eck et al. (2010) for more information on the selection of desert dust dominated cases and the dynamics of fine-coarse mode mixtures. Also note that the data base sample size was relatively small in 2000 (Dubovik et al.(2002)) only used data through the year 2000), and that the current data sample size is much larger for all sites analyzed and also more accurate due to consideration of particle shape, more accurate surface reflectance and better input data screening for anomalies.**

Although we agree with the reviewer that in order to select the cases of pure dust one needs to use the data corresponding very small numbers of alfa (0.2 or less), we decided to use Solar Village model because we do not expect any significant discrepancies due to choice of this model. In our opinion, the most using values of alpha 0.6 and less represent rather typical observations of desert dust and even if such data may include some mixtures these mixtures are dominated by desert dust. In particular for the examples used, $\alpha = 0.51$ in SolV1, and $\alpha = 0.33$ in SolV2.

Page 6857, lines 13-16: How do you justify selecting different AOD levels for the GSFC and Mongu sites in your simulation analyses? Both are fine mode dominated aerosol types and the use of significantly different AOD levels confounds the comparisons you make. At higher AOD levels the aerosol signal is greater and the uncertainty level of the retrieved parameters decreases, therefore it seems that your comparison begins with some built-in bias.

As it was already mentioned, we have used the average values of the AOD obtained in Dubovik et al. (2002) for the sets GSFC1 ($< \tau(440) >= 0.24$ here $< \tau(440) >= 0.2$), and Zamb1 ($< \tau(440) >= 0.38$ here $< \tau(440) >= 0.4$) and twice these values for the sets GSFC2 ($2 < \tau(440) >= 0.48$ here $< \tau(440) >= 0.5$), and Zamb2 ($2 < \tau(440) >= 0.76$ here $< \tau(440) >= 0.8$). Nevertheless, we have derive conclusions based on the average AOD levels, which are typically different for each type as shown in the key sites climatology. We have added a sentence to draw attention on the implications of such approach regarding the retrieval error at this point (see answer on next comment) and further on during the discussion of the results.

Page 6857, lines 21-24: "It should be noted that the obtained spectral aerosol optical depth does not exactly match the input values provided as reference. This can be explained by the fact that the used aerosol parameters represent the climatological regressions." This is not really a satisfactory explanation for AOD differences of $\sim 10\%$ at the higher AOD levels as shown in Table 1. Did you use the same particle shape and surface reflectance in your simulations as were used in Dubovik et al. (2002)? Note that surface reflectance in Dubovik et al. (2002) was held constant for all sites (no solar zenith angle dependence or geographical variation. All sites were assumed to have Lambertian albedo of 0.03, 0.06, 0.20, and 0.20 for 440, 675, 870 and 1020 nm respectively. Additionally the constraints on the tails of the size distribution (lowest and highest radii limits) were not as strong in the retrievals shown in Dubovik et al. (2002) as are utilized in the current (since Nov 2006) retrieval algorithm. A lack of consistency in the retrieval algorithms plus assumptions (constraints) and input data here is a problem.

We agree that inaccurate assumptions on surface reflectance values could produced some biases in the retrieval parameters. At the same time, Dubovik et al. (2002) have done a very careful data selection in their analysis with the purpose of minimizing possible uncertainties. For example, the values of complex refractive indices for desert dust were adapted to the models from the observation for those cases with high optical thickness, therefore the effect of error propagation from surface reflectance assumption was minimized. Indeed, the study by Sinyuk et al. (2007) have concluded that the significant errors in the retrieved properties can only be observed at bright desert sites for situation with low aerosol loading.

In addition, and most importantly, we would like mention, that possible uncertainties in the AERONET data can hardly affect the results of our numerical studies, because in the numerical studies retrievals are fully consistent with the forward calculation used to generate synthetic data.

For all the simulations we have used the new version of the retrieval algorithm (used since Nov 2006): the surface reflectance is approximated by a Bi-directional Reflectance Function (BDRF) using the values of $\rho_o(\lambda)$, $\kappa(\lambda)$ and $\Theta(\lambda)$ depicted in table 3 and the spheroid model has been also used during the simulation (as commented in the paper, the sphericity parameter is fixed as 0 for desert dust (i.e. all the particles are considered to be non-spherical) while for biomass burning and urban examples the sphericity parameter was set to 100 (i.e. considering all the particles as spheres)). We do agree with the referee that the use of spheroid model has an important contribution to the spectral aerosol optical depth (even though it's only been used in desert dust case) and this issue was not mentioned in the manuscript. We have included this fact in the text as follows:

It should be noted out that the obtained spectral aerosol optical depth does not exactly match the input values provided as reference. This can be explained by the fact that aerosol models from Dubovik et al.(2002) are based on the linear regression analysis. At the same time, the aerosol *optical depth depends non-linearly on such aerosol parameters as real and imaginary parts of the refractive index. Moreover, for the desert dust type there is another important element regarding the explanation of these discrepancies: The software package including the spheroids was developed later (Dubovik et al., 2006) than the analysis considered as reference here (Dubovik et al., 2002). The assumption of non-sphericity rises the aerosol optical depth as a consequence of the increase in scattering. In particular, if we considered the spherical model for the desert dust type in the forward simulations, the values for the aerosol optical depth at* 1020 *nm for SolV1 and SolV2 would be* $\tau_a(1020) = 0.294$ *and* $\tau_a(1020) = 0.493$*. These values are in a better agreement with the reference values and the differences could then be related to the mentioned non-linear dependences. Note also, that these simulated values do not depend on the "conditions" such as the measurement geometry or the solar zenith angle. Finally, we should remark here that larger AOD implies better retrieval accuracy, however we have decided to maintain the typical average values for each site (even though they are quite different) so that our study is related to the average conditions.*

Finally, we agree with the referee that the constraints on the tails of the size distribution (lowest and highest radii limits) have changed for the different versions but they are not related to the forward module so they can not be used to explain the AOD differences of $\sim 10\%$.

Page 6864, lines 4-5: You say there is no reliable information on aerosol vertical profiles. It might be more accurate to say there is little information, as CALIPSO does provide information in at least a climatological sense (if not more) on aerosol vertical profiles.

We agree with the referee. We have replaced the sentences with "There is very limited information".

Page 6868, section 2.4: You need to also mention that geographically and temporally varying (16 day averages throughout the annual cycle) surface albedos were utilized in Version 2 retrievals. These spectral surface albedos are mid-day black sky albedos from Moody et al. (2005), and are based on atmospherically corrected MODIS data averaged over a 5km radius of each AERONET site (see Eck et al., 2008; section 2.3). The BRDF model and parameters (for each ecosystem type) are used to compute the spectral reflectances at solar zenith angles throughout the day.

Thank you for your comment. We have completed the text:

The BDRF parameters are basically three: $\rho_o(\lambda)$, $\kappa(\lambda)$ *and* $\Theta(\lambda)$ *which characterize the intensity of reflectance, the anisotropy of reflectance and the forward/backscattering contribution in the total reflectance, respectively. In AERONET version 2 retrievals (Holben et al. (2006), Eck et al. (2008)), the land BRDF parameters are adopted from MODIS Ecotype generic BRDF models and mixed by the ecosystem map of Moody et al. (2005): geographically and temporally varying (16 day averages throughout the annual cycle) surface albedos are utilized. These spectral surface albedos* *are mid-day black sky albedos from Moody et al. (2005), and are based on atmospherically corrected MODIS data averaged.*

Page 6869, lines 3-10: Can you quantify the errors in spectral albedo that results from this variation in BRDF model parameters that you assume? Otherwise it is difficult to assess whether you have analyzed a realistic range of spectral surface albedo uncertainty.

As the referee states, among other issues, differences between retrievals from Version 1 and 2 of AERONET due to surface albedo are studied in Sinyuk et al. (2007) and Eck et al. (2008). Both studies are of great interest as, by means of real examples, the authors compare the aerosol properties retrieved using two different approximations to characterize the surface albedo: by a Lambertian surface (AERONET version 1) and by a Bi-directional Reflectance Function (BDRF version 2). The aim of the present study is, however, to check whether there is a more favorable measurement geometry, between principal plane or almucantar, with respect to errors in the description of surface albedo. With this goal, we could have done a similar analysis comparing the results from AERONET Version 1 and Version 2 for both geometries. But giving the number of years that Version 2 has been operative with satisfactory results, this kind of analysis would have been outdated. We understand that the actual model correctly describes the surface albedo and that the BDRF parameters are quite accurate for every site at any time in the year.

On the other hand, as we do not have any "a priori" information about possible errors in the BDRF parameters, the best way to characterize them is assuming a random noise. We reckon that a biased error is controllable and subject of being updated at anytime in AERONET-retrievals, so if there had been any "known bias" they would have been already corrected.

The description of these random errors is done as follow in the text: *Specifically, the errors in* ρ_o *will be relative errors generated randomly from a normal distribution with mean* 0 *and standard deviation* 15%*, with a limit up to* ±30%*. The errors in* $\kappa(\lambda)$ *and* $\Theta(\lambda)$ *will be absolute errors generated randomly, and with values of the standard deviations of* 0.05 *and* 0.025 *respectively; the error limits will be established as* ±0.1 *for* $\kappa(\lambda)$ *(with* $\kappa(\lambda) > 0$*) and 0.05 for* $\Theta(\lambda)$*.*

In order to quantify the simulated errors, in figure 1, three histograms with the values of the "erroneous" BRDF parameters are represented for the example of Solar Village site. This figure would likely help in the description of the simulated error scheme but might be too long. We prefer not to include it but we are open to change this depending on the referee's opinion.

Page 6869, lines 11-13: For your analysis of the effect of surface reflectance errors on the retrievals you only show results for the largest aerosol AOD for each site. However this minimizes the effects of spectral surface albedo effects and BRDF in the retrievals since errors due to surface reflectance are much larger for low to mid levels of AOD. Therefore your analysis of this error source tends to minimize the potential

errors and does not give the range of more typically encountered values due to this perturbation.

That's true the effects due to an error of BRDF in the retrievals are much larger for lower levels of AOD. However, in order to summarize we needed to make a choice and we took the AOD with the highest level (in the same way that we have only taken two solar zenith angles for the analysis). At bottom and specifically for this particular section, we are much more interested in knowing if there is a more favorable geometry regarding BRDF errors rather than in developing a detailed analysis of the differences obtained in the retrievals.

Page 6870, lines 5-7: The authors state: "The analysis of the results leaded to two main conclusion. First, the introduction of the random errors in the surface reflectance assumption did not results to any retrieval bias." This is misleading, since in reality the assumed surface reflectance can have a bias and this bias will result in biases in the retrieved optical properties. See Sinyuk et al. (2007) and Eck et al. (2008) for examples.

As commented, in the studies Sinyuk et al. (2007) and Eck et al. (2008), they compare the results obtained from AERONET Version 1 and Version 2. They know that the surface reflectance model used in Version 2 was more accurate than the one used in Version 1, and that the model used in Version 1 was biased respect to the new model. Here, we just state that the introduction of random errors does not produce any significant bias in the retrievals. In any case we have reformulated the sentence as follows: *There is not a clearly defined tendency in the mean of the differences as it could be expected because in our analysis only random errors were considered. (...) Therefore, the analysis needs to be done in terms of the standard deviation which contains, in this case, the information about the dependency of the retrieved products on a random error in the surface reflectance.*

Page 6871, lines 16-18: "In summary, the observed differences in the optical parameters are relatively small considering the magnitude of the errors that we have introduced in the surface reflectance assumption." I argue that the uncertainty in the retrieved parameters due to surface reflectance uncertainty has not been adequately characterized in this paper since: 1) high AOD cases only were studied, which minimizes the surface effects, and 2) it is not possible to tell whether you have used realistic errors in surface reflectance in this study.

In the answers above the issues regarding the error assumptions and AOD levels have already been discussed. In addition, we have modified the conclusions of this specific subsection as follows: *In summary, the introduction of a random noise in the BDRF parameters affects more at* $\theta_s = 45°$ *than at* $\theta_s = 75°$ *for all the analyzed aerosol properties. This is more evident for the case of the principal plane, which presents higher differences at* $\theta_s = 45°$ *than the almucantar geometry,* *while at $\theta_s = 75°$ the differences are similar for both geometries being even slightly smaller for the principal plane.*

Page 6872, lines 17-21: I am puzzled as to why you only analyze the retrieval errors that result from the maximum pointing error (0.4 degrees) and not the more typical value of less than 0.1 degrees. As a result it is not possible from this section to know what the effects of a typical pointing error has on the retrievals.

In the work Torres (2012) the retrieval errors for 0.2° and 1° are also analyzed. The differences in the retrievals can be neglected in the case of 0.2°; this result is added in the article, in order to settle clear that the typical pointing errors do not have any significant influence on the retrievals. The results obtained for the case of 1° were dismissed for the present study as it does not represent a realistic error. The case of 0.4° is the maximum possible error that could appear, given the geometry of sun-photometer, without affecting the direct sun measurements ($1.2° - 1.3°$ of field of view). In the study "Measurements on pointing error and field of view of Cimel-318 Sun photometers in the scope of AERONET" from Torres et al. 2013 (recently published in AMT), it is shown that for some punctual measurements, errors above 0.3° can be committed, even though the typical pointing errors are under 0.1°. Therefore, to show such extreme case is necessary to show the pointing error effects more clearly knowing that typical errors do not produce significant differences.

Page 6873, lines 5-8: It should be noted in the text here that in actuality the left and right sides of the almucantar data scan are averaged for AERONET processing. This averaging is not just a hypothetical possibility, as suggested in this sentence, but in reality this is operational and has always been applied in AERONET data processing (based on a decision made over a decade ago as a result of much analysis). This results in numerous benefits versus the principal plane scan as you have stated previously.

The authors did not want to mean in this sentence that averaging is a possibility, just that almucantar geometry offers this possibility. As it was mentioned, in the introduction we specify clearly that symmetry property has been used to process almucantar data for more than a decade. We have added a small clarification in the text: *the errors are initially not symmetric but due to the possibility of averaging the left and the right branches (as it is done operationally in AERONET), they become symmetric.*

Page 6877, lines 1-4 and lines 17-19: You state that the real part (and imaginary part in lines 17-19) is more stable for the biomass burning aerosol type than for the urban aerosol, but you need to mention that the AOD levels are not the same for these simulations (the high AOD cases have GSFC at 0.5 and Mongu at 0.8 at 440

nm). Therefore these comparisons are biased from the start since several sources of uncertainty diminish as AOD increases.

We have included that the higher AOD used for Mongu could also affects in this result.

Page 6879, lines 5-6: Why did you change from GSFC to Beijing for the urban aerosol category while the other two sites (Mongu and Solar Village) remained the same? You need to justify and fully explain why you did not maintain continuity with the same urban aerosol site used in the analysis done in the previous (simulations) section.

Simulation and real data analyses were in the beginning two different studies. At some point we decided that they could complement each other but the sites were already chosen. Anyway, we have followed your recommendation and replaced Beijing with GSFC.

Page 6879, lines 8-10: "...not asked for in retrievals..." is awkward wording and confusing, so please rewrite this sentence. Even though you did not use Level 2 retrievals (since you want to analyze data for SZA<50 degrees) you should still apply data quality checks such as sky error $<5\%$, ensuring good fit between measured sky radiances and those computed from the retrieved aerosol parameters. Additionally the symmetry check should be applied to the measured almucantar radiances to screen for clouds and other non-homogeneities. Was there any attempt to screen the principal plane scans for clouds? Did you only use cases where AOD>0.4 at 440 nm and if you did not then justify why lower AOD cases were analyzed.

We have changed the confusing parenthesis by: *(though we did not only use Level 2 retrievals as we are interested also in almucantar data with* $\theta_s < 50°$ *and* $\tau_a(440) < 0.4$*).*

We have applied symmetry check for cloud screening (20%) and we have only taken principal plane data "cloud-free" from visual inspection. We have eliminated 8 almucantar/principal plane pairs as they were overpassing the sky error limit established at 5% (only principal plane retrievals and none of them overpassed 8%). All of these cases were for desert dust inversions and we have needed to replace some days to assure a minimum of 4 pairs of data per day. We have taken advantage of this fact to decrease the Angstrom exponent in some examples of the group called Desert-dust I as it will be commented later.

Page 6879, lines 18-22: The analysis of real data is for a relatively small sample size, only 6 to 12 days per site/subsite data set and therefore possibly not very representative of all conditions, especially for partial cloudiness and aerosol inhomogeneity since you have only selected days with stable AOD and therefore less cloudy and more homogeneous conditions.

It is true that we have considered only favorable conditions and this fact has been added in the discussion. A more general analyses could be made in a different study though it would not be immediate as there is not a cloud screening implemented for principal plane.

Page 6880, lines 3-4: You say all the data from Beijing is from fall and winter, yet Table 4 shows data from August (2 days; summer) and one in March (meteorological spring) also, so there is an inconsistency here. Also you need to mention in this section that the very high AOD levels at Beijing results in significant reduction in retrieval errors since the larger aerosol signal dominates all of the sources of uncertainty. Therefore comparison of Beijing to the other sites begins with a built-in bias towards better results from Beijing.

We agree. However, the discussion here is not necessary since we have replaced Beijing with GSFC in the analysis of real data in order to increase robustness of the study

Page 6880, line 4-6: You state that an Angstrom Exponent (alpha) threshold of higher than 1.7 was used for Beijing, however this is not a typical Angstrom exponent value for this site even for months when there are no desert dust transport events. The alpha values in Beijing are typically significantly lower than 1.7 even for pollution events since he fine mode (sub-micron radius) particle size is relatively large (high AOD results in greater coagulation rates) and there is a larger coarse mode background value in Beijing than other urban sites, partly due to fly ash from coal combustion (see Eck et al. (2010) and references therein).

Due to this and other issues, we have replaced Beijing with GSFC.

Page 6880, lines 10-19: You need to show the mean Angstrom exponent in Table 4 for each day (in addition to the mean AOD), especially for the Solar Village data that you split into two populations. The Desert dust I (Solar Village) population has mean Angstrom values ranging from 0.42 **to** 1.1 **except for one day with 0.20, therefore this population should be described as "Mixed fine/coarse" rather than "Desert dust". On the other hand, the Desert dust II population has Angstrom exponent ranging from** 0.11 **to** 0.24**, except one day that alpha of** 0.47**, therefore this population has true desert dust aerosol (see Kim et al. (2011) for dust selection criteria using AERONET data). Therefore differences between retrievals from these two data populations for this site can in part be attributed to differences in aerosol type and you need to mention this in the text. Additionally, data quality issues are apparent in some of the days in the Desert dust II population. For example, on some dates in the Desert dust II population, there are spectral crossover of 440 and 500 nm AOD which indicates AOD calibration uncertainty resulting in larger AOD errors at small solar zenith angles (SZA). See Hamonou et al. (1999; JGR) equation 1 showing**

that the error in AOD due to calibration is proportional to $1/m$ (m =optical airmass; $m \sim 1/cos(SZA)$). Therefore it is likely that retrieval errors (biases) for both scans are larger at smaller solar zenith angles. It is necessary to include some discussion of the uncertainty in AOD as a function of SZA in this section, and in other sections of this paper. Other errors including sky radiance inconsistencies at 6 degrees azimuth angle (indicating sky radiance calibration issues; these will be quality-control screened in a future AERONET data base version) are evident in some of the days in the Desert dust II data but not in the Desert dust I data. Therefore input data quality also varies between the two Desert dust (Solar Village site) data populations.

The Angstrom exponent has been added in Table 4.

We have changed some of the selected days in Solar Village as the sky error for several principal plane retrievals were a little higher than 5% (between 6% and 8%). With the new days selected, the mean Angstrom exponent of Desert dust I goes from 0.20 to 0.73 and for Desert dust II from 0.11 to 0.55. In any case as the mean Angstrom exponent values of the second group are still lower than in the first one, we have mentioned that the differences between retrievals from the two data populations for this site can in part be attributed to differences in aerosol type (this is added in subsection 3.3).

The influence of the AOD errors has not been analyzed in the simulation analysis since we were interested only in the geometry of radiance Measurements and AOD direct sun observations are done in the same way for both PPL and ALM geometries. Undoubtedly it would be quite interesting to analyze the effect of a sun calibration error, moreover, when this error is not constant along the day as Hamonou et al. (1999; JGR) states. For our study, we agree with the referee that there are several places along the text where it should be mentioned that the uncertainty in the AOD has important consequences in the retrieving process so we did it in subsection 3.3 and in the discussion.

Page 6881, lines 27-28: Please explain why Beijing was chosen in this section in place of GSFC, since this confounds the ability of the reader to assess the differences in both simulations and real data, as can be done for both Mongu and Solar Village sites.

Already discussed.

Page 6882, lines 1-2: I disagree with your suggestion that the mean spectral single scattering albedo (SSA) values for Beijing are similar to other urban sites shown in Dubovik et al. (2002). See Figure 19 in Eck et al. (2005) that shows less wavelength dependence of SSA at Beijing than the other strongly absorbing urban sites. Additionally the fine mode particle radius is significantly larger at Beijing (at the high AOD levels you studied) than for GSFC, Mexico City, Maldives and Cretiel (urban

sites shown in Dubovik et al. (2002)).

This text was erased in the manuscript as we have not considered Beijing case.

Page 6882, lines 15-16: It should be noted that the 0.03 uncertainty in SSA for AERONET almucantar retrievals (Dubovik et al., 2000) is largely due to sky radiance calibration uncertainty, and the 0.03 value is for AOD of ~ 0.4 **at 440 nm, while for the Beijing data analyzed in this paper the SSA uncertainty is much lower due to much higher AOD of the cases studied.**

We agree with the referee. But as Beijing data were erased from the analysis this fact does not need to be considered.

Page 6883, lines 12-14: It seems that you are confusing uncertainty in almucantar retrievals due to combined effects of calibration, pointing error, surface albedo, etc. with the consistency between almucantar and principal plane retrievals. Therefore it seems that there is no "improvement" to talk about.

It is true. We have deleted the sentence from the text.

Page 6883, lines 20-28: You say that the discrepancies between Desert dust I and Desert Dust II cannot be attributed to differences in the aerosol measured, but a quick check of the data for the dates you analyzed show that there are significant differences in Angstrom exponent and therefore it is to be expected that there should be differences in aerosol properties for these two Solar Village data subsets. You compared size distribution parameters in this section but did not look at any measure of the relative contribution of fine versus coarse mode in AOD (fine mode fraction (FMF) or Angstrom exponent). Eck et al. (2010) and Eck et al. (2008) have shown significant variation in SSA and size distribution as a function of FMF and/or Angstrom exponent.

As it was previously discussed, we have made several changes in this subsection. We have noted that there are some intrinsic differences in the data sets regarding Angstrom exponent (even though we have changed some days in the analysis) as well as we pointed out some uncertainties in the AOD for the case Desert Dust II.

Page 6885, lines 11-14: It should be mentioned here that the better agreement between almucantar and principal plane retrievals for larger solar zenith angles is also due in part to more accurate AOD input data to the retrievals at larger SZAs.

We have added the following paragraph:

It should be noted here that, apart from the error sources considered along the study, the AOD errors are of an importance in the retrieving process (Dubovik et al., 2000). We have not focused on them in the present study, since in principle the AOD measurements are equivalent in the PPL *or ALM geometries. However, if AOD errors are significant they may have different effects on aerosol retrievals using PPL and ALM observations, specially in those cases with larger differences in the information content, i.e. for small solar zenith angles where the scattering angle coverage in the almucantar is significantly smaller. Furthermore, the error in AOD due to calibration is proportional* $1/m$ *(*m *=optical airmass;* $m \sim 1/cos(SZA)$*, see Hamonou et al.(1999)) which particularly enlarges the AOD errors for short solar zenith angles. The combination of both factors can explain the better agreement between the two geometries at larger SZA where the AOD is more accurate. In fact, some uncertainties in the AOD calibration have been found in the set Desert dust II data (e.g. spectral crossovers of 440 and 500 nm).*

Page 6886, lines 4-6: Averaging of the left and right branches of the almucantar scan not only results in more "stable" measurements but also a more representative spatial (angular) distribution of sky radiances due to aerosol inhomogeneity. This should be added to your discussion.

Thank you for the comment, it has been added.

Page 6886, Line 25: Please note that SZA also exceeds 50 degrees at mid-day in winter in subtropical to mid-latitudes, not just at high latitudes as you suggest.

Also added.

Page 6886 line 28 through Page 6887 Line 7: You suggest here that valuable information is being lost by not providing the principal plane retrievals as a data product within AERONET. However the only advantage that principal plane retrievals have over almucantar retrievals is a larger solar zenith angle range (allowing for mid-day retrievals in the tropics and in summer in mid-latitudes), due to sky radiance measurements made over a greater range of scattering angles. However your reasoning does not consider a balanced approach since there are numerous disadvantages that the principal plane scan has relative to the almucantar scan. Lack of rigorous cloud screening and lack of ability to account for aerosol inhomogeneity are two significant factors that hinder the robustness of the principal plane scan retrievals. Your "real" data analysis presented in this paper does not assess the effects of these issues since you deliberately selected only dates where the AOD was very stable thereby excluding many cases where the principal plane would be disadvantaged. You studied additional disadvantages of the principal plane such as greater sensitivity to pointing errors and aerosol vertical profile than almucanatrs, but yet did not consider in the analysis presented that spectral AOD (a key algorithm input parameter) is less accurate at smaller solar zenith angles. In summary you need to at least modify your statements to account for the fact that the principal plane retrievals will often be of lower quality

(larger uncertainty), and especially so at lower solar zenith angles.

We agree with the referee and we have changed deeply the discussion section including the recommendations that he has suggested.

Page 6887 lines 8-11: It is quite inappropriate for you to be mentioning the new hybrid scan concept in this paper since the idea originated at GSFC by the AERONET group after much analysis, and there are no GSFC co-authors on this manuscript. Additionally, even then it is premature to mention in this paper since this is still in the early development phase.

The hybrid scenario was internally presented in the internal AERONET workshop in Lille-2011, and this reference should have been included in the present manuscript (probably as personal communication).

The unique intention of the authors for mentioning the hybrid scan was to highlight that AERONET community is aware of most of the results/problems commented in the manuscript and how they are improving the products in Version 3. In any case and as you state, there are no GSFC co-authors on this manuscript so we have erased hybrid scan concept from the discussion.

Page 6887, lines 13-27: The opening statement in the conclusions that considers only the ideal case is a poor way to begin discussion on principal plane versus almucantar scan results. Additionally a failure to mention the advantages the almucantar has when the there is partial cloud cover is a serious omission here in the Conclusions section. Another significant omission in the Conclusions and the entire paper is the lack of consideration of AOD accuracy and how it varies as a function of solar zenith angle.

We agree with the referee and we have also changed the conclusion section adding the comments suggested here.

Page 6888, lines 16-19: You did not realize that there are significant differences in Angstrom exponent and FMF between the Desert I and Desert II subsets signifying different aerosol properties, therefore this conclusion is invalid and that part of the paper needs to be re-analyzed or rewritten to take consideration of these issues.

We have also added this fact in the conclusions.

Page 6888, lines 26-28: It is somewhat surprising that you consider this a new conclusion (need to limit almucantars to $SZA > 50°$), as it was well known for over a decade.

It was not our intention that this fact was understood as a new conclusion from our work. We just wanted to emphasize that we obtained the same results as in previous AERONET analysis. We have changed this sentence in order to settle it more clear.

Technical corrections:

We have corrected all the technical corrections indicated by the referee.

<u>Answer to referee #2</u>

Me and the co-authors of the paper thank the positive review of anonymous referee #2. We agree in clarifying the points that were requested including the pertinents comments along the manuscript:

POIN-TO-POINT responses:

<u>Referee's comment 1:</u> **P6860, line 6-13: If possible, please explain physically why there is a strong connection between the retrieved fine mode and the real part of the refractive index. Isn't there a connection between coarse mode distribution and real part of refractive index? "There is a strong connection between the retrieved fine mode and the real part of the refractive index and this fact will be recurrent in the next studies. Nevertheless, the disagreements are more striking in the coarse mode."**

Both the real part of the refractive index and the fine mode are very sensitive to large angles of phase function ($> 20-30°$), while for shorter angles the coarse mode has a greater influence ($< 20-30°$). Therefore, both parameters are bound and variations in the retrievals in one of them have consequences in the other. Thus, this effect has been already observed in previous studies analyzing the effects on the inversion process of non-sphericity (Dubovik et al. 2002, Fig. 3), effect of surface albedo (Sinyuk et al. 2006, Fig. 17 and Fig. 18) and (Li et al. 2009). This has been included in text.

<u>Referee's comment 2:</u> **P6864, line 25-26: Why is the effect of multiple scattering important for principle plain observation? If possible, please explain a little more. "Only small differences can be observed in the coarse mode for desert dust at $\theta_s = 75°$, where the effect of multiple scattering is more important. "**

From Eg 1 to Eq. 12, we have explained that in single scattering approximation only principal plane measurements may be affected by a heterogeneous aerosol vertical distribution. In the discussion from P6864 we have indicated two facts for measurements at $\theta_s = 75°$: first that the presence of heterogeneous aerosol vertical distribution affects almucantar measurements for large solar zenith angles as multiple scattering is more relevant. Second, principal plane measurements are more affected by the presence of heterogeneous aerosol vertical distribution as this presence affects not only in multiple scattering terms but also in single scattering approximation.

<u>Referee's comment 3:</u> **P6866, line 16-17: It is better to add some reference to support following sentence. "in the real atmosphere molecular scattering generally dominates at the altitudes above 5 km."**

We have added two references: Figure 6 in Elterman (1966) and Table 4.10 in D'Almeida et al. (1991). It can be seen, how in general, between 3 and 10 km the contribution of Rayleigh to the extinction is more important that the aerosol one.

<u>Referee's comment 4:</u> Section 2.5: Did the author discuss the reason why the opposite results occur the negative and positive pointing error, and also about difference between principle plane and almucantar? Please explain or discuss why the opposite error occurs? And if possible, the meaning of the following sentence is not clear. It means that the short wave is more affected by the fine mode particles, right? Please explain. (P6876 line 27-28) "it becomes greater as the wave length is shorter, or on the other words, is larger at wavelengths more affected by the fine mode."

Yes, we have discussed briefly the differences between positive and negative and also between principal plane and almucantar. With this aim, we have used the figure 11 as reference which contains the radiance relative differences. Maybe it was not too clear and we have changed some sentences as the one in - P6876 line 27-28 - as follows: *it becomes greater as the wavelength is short as short wavelengths are more affected by the fine mode particles.*

<u>Referee's comment 5:</u> P6882, line16 and 25-26: Please add some references for the following sentences: "the uncertainty given by Aeronet for ω is 0.03" "the accuracy given for this parameter is 0.03 for biomass burning and urban and a bit higher, 0.05, for desert dust;"

They are given in the following url: `http://aeronet.gsfc.nasa.gov/new_web/Documents/inversions.pdf` and the criteria are given following the recommendations of the paper by Dubovik et al. (2000). The url is added in the manuscript.

Technical corrections:

We have addressed the technical corrections indicated by the referee. However, regarding the comment 4, we have preferred to keep the term backward code in figure 2 as it makes the description more visual. For the same reason, we also prefer to keep $R(\Theta, \lambda)$ instead of $R(\Theta, \lambda, n, k, r)$ because in the position that has in the chart it represents the "measures" and in this case it depends only on Θ and λ (the direction and the channel we are using) while n, k and r should be retrieved.

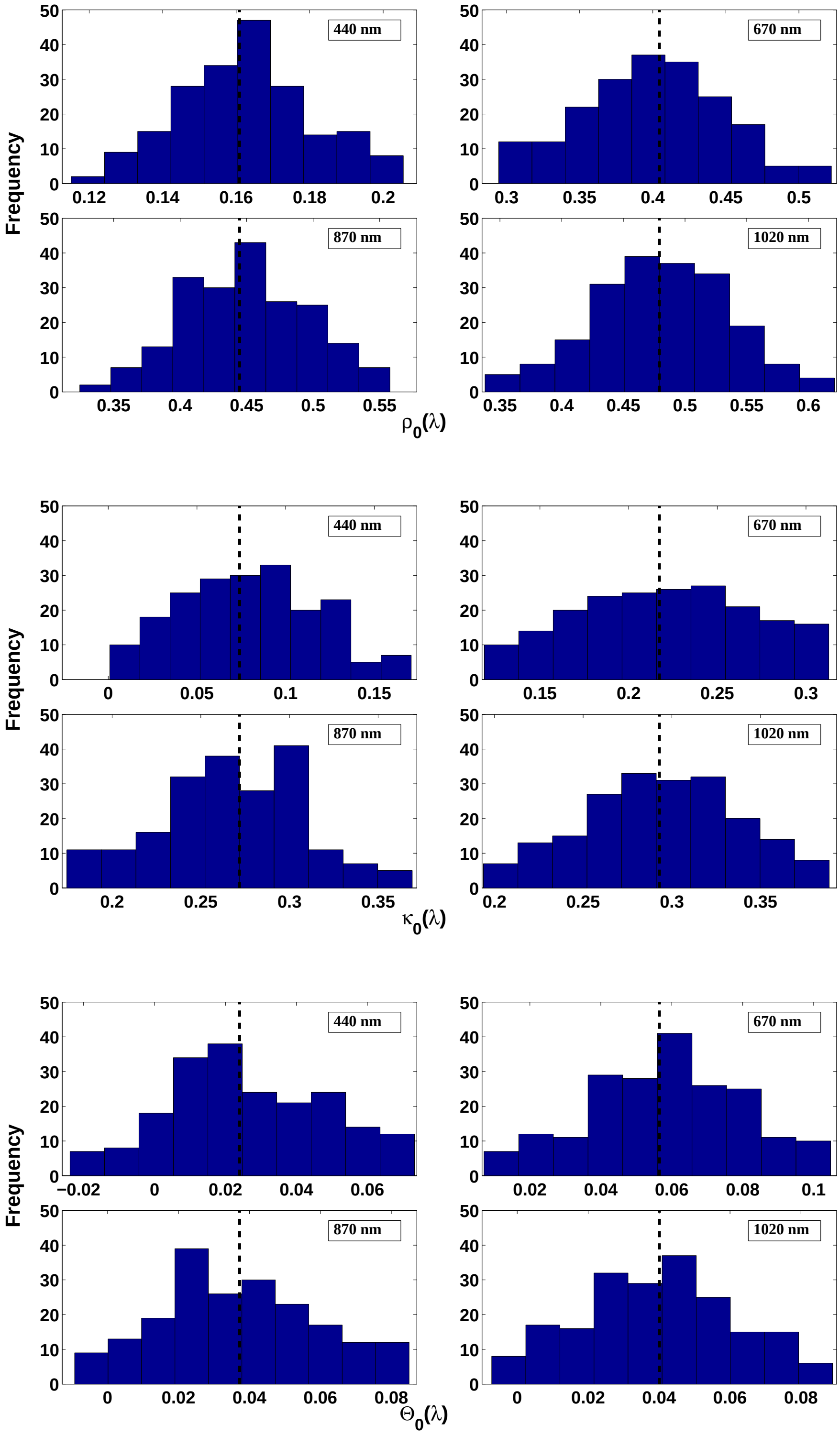

Figure 1: Histograms of the used BRDF parameters in order to simulate random errors in the description of the surface albedo for the case of Solar Village: the values ρ_o are represented on the top, the values of $\kappa(\lambda)$ in the middle and the values of $\Theta(\lambda)$ at the bottom for the different wavelengths. The dashed line in every subfigure represents the original value of the corresponding BRDF parameter.