We appreciate the reviewers' valuable comments and constructive suggestions which help us improve the quality of the manuscript. We have carefully revised the manuscript according to these comments. Point-to-point responses are provided below. The reviewers' comments are in black, our responses are in blue and changes in manuscript are in red.

Reviewer:

Authors have a made a huge effort by rerunning the algorithms in order to improve the study, also a lot of details about the two measuring sites and their atmospheric conditions has been added. However, my major concerns have not been answered and although a lot of information has been added, I can quote from my earlier review:

"However, the manuscript lacks of explanations on the causes of differences between the v4.2 and v5.0 retrievals and sufficient evidence on the actual seasonal variability of aerosols in the two regions. Algorithms of both versions are clearly described, but it is crucial to pinpoint and discuss the way the differences between the versions affect the retrievals. Since the two algorithms are not treated as a "black box", it should be more clear which physical processes affects the retrievals and which atmospheric conditions could lead to highest uncertainty. At least some discussion on the uncertainty of each variable in each approach should be provided. Also, the part about the seasonal variability of aerosols properties, results are presented but not investigated and discussed in the level expected for a scientific study."

I think nothing has been done in this direction. Also, information added are not used in order to interpret the results and explain the findings of the present study. Also, some discussion is needed to conclude the validity or higher quality of v5.0 retrievals before using them for aerosol characterisation in section 3.2

I suggest a major revision of the manuscript before considering for publication in AMT.

Response: Thank you for your valuable comments and constructive suggestions.

V4.2 is based on inversion scheme of the Phillips-Twomey type solution of the first kind of Fredholm integral equation with homogeneous smoothing constraint, and V5.0 is based on the second kind of the equation with inhomogeneous constraint with a priori information for aerosols (Twomey, 1963) to retrieve the inherent aerosol optical properties.

The most different physical process between V4.2 and V5 is a derivation of particle size distribution. V4.2 doesn't have a constraint for the size distribution. On the other hand, V5.0 has a constraint for it using the term ($S_a^{-1}(x_k - x_a)$) in Eq. 4 and gives a strong constraint for the edge of size distribution, so the edge of size distribution in V5.0 close to zero, but V4.2s' looks jumped.

We have compared the differences between retrieved SSAs at 500 nm byV5.0 and V4.2 when set the coarse model radius rm2 in Eq. (5) as rm2 = 1.5, 1.8, 2.0(default), 2.5 and 3.0 in Skyrad.pack V5.0 based on the measurements in 2014. Based on the sensitivity tests, we found the correlation coefficient between SSAs by V5.0 and V4.2 was the highest when setting rm2 as 2.0 (the default value) in V5.0 at the two sites.

We assumed an error of ±5% for calibration constant F0, ±5% for solid view angle SVA, ±50% (±0.05) for ground surface albedo Ag. We compared the differences in retrieved SSA values at a wavelength of 0.5 μm between cases with and without the assumed errors. On the basis of the sensitivity tests, it is concluded that an error in the calibration constant (F0) causes an error in both retrieved SSA and AOD. The averaged differences in retrieved SSA values due to ±5% error in F0 varied from 3% to 5%. An error of ±5% for solid view angle SVA introduced about ±2% differences in retrieved SSA values both by V4.2 and V5.0. Overestimation or underestimation in the Ag results in underestimation or overestimation of the SSA. An error of ±50% for ground surface albedo Ag caused about 1% averaged differences in retrieved SSA values both by V4.2 and V5.0. With the atmospheric pressure PRS increased by 1%, 2%, 3% and 4%, the averaged differences in SSAs didn't exceed 0.8%.

We also investigated whether the total amount of aerosols in the atmosphere were linked to the differences in SSA between the two versions. It could be said that the condition of low AOT affected the retrieval accuracy of SSA, especially when AOT is less than 0.4.

Specific comments

P3l25 Is the same calibration used also for sky radiance measurements?

Response: The calibration method for sky radiance measurements is different from the calibration method for the direct solar irradiance measurements. The solar disk scan method has been routinely used in the SKYNET measurement of the SVA of the sky radiometer by scanning a circumsolar domain (CSD) of ±1° around the sun with every 0.1° interval (Nakajima et al., 1996; Uchiyama et al., 2018). (P4 Line 1- 5)

P6 l23-25 Other differences should at least be mentioned, before referring to Hashimoto et al work. At the end the present manuscript is about these differences.

Response: Other differences between V5.0 and V4.2 are mainly the cloud screening algorithm. The cloud screening method in V4.2 relies heavily on the global flux test and needs global irradiance data, V5.0 poses a condition regarding the magnitude of the coarse mode of the SDF as follows:

$$C_v \times v(2.4\mu m) < \max\{v(7.7\mu m), v(11.3\mu m), v(16.5\mu m)\}.$$

where Cv is a threshold coefficient to be determined for optimum rejection of cirrus contamination, v(r) is vertically integrated aerosol SDF, as a function of particle radius, r. Based on the analysis of data at the Pune and Beijing sites (Hashimoto et al., 2012), Cv is set as 2 in V5.0 to reject most cirrus contamination cases and pass through dust cases. It is necessary to determine Cv after collecting more cirrus contamination data and dust day data. (P7 Line 5- 15)

Section 3.

AOD is derived from Direct Irradiance measurements. In earlier text only the algorithms for inversion products is described. Are there differences in direct sun algorithms between 5.0 and 4.2? Since AOD is presented these details should be explained. Also, even this slight differences of AOD could propagate considerable uncertainties, thus it should be more clear if and how AOD is used in inversion calculations in both versions.

Response: There are nearly no differences in direct sun algorithms between 5.0 and 4.2. As shown in Fig.2, there were very slight differences between AODs by V5.0 and V4.2, this is mainly caused by the very small differences in calibration constant F0. F0 in V4.2 and V5.0 are both determined from sky radiance data by the Improved Langley method. V5.0 adopts more rigorous data processing and cloud detection methods. The sky radiance measurements which involved in F0 calculation are a little different in V4.2 and V5.0. (P7 Line 27- 30)

In inversion calculations in V4.2 and V5.0, AOD are both used as indicative values in the first step of the loop but are updated at each iteration. The AOD data can be given different weights with respect to the normalized diffuse sky flux data, according to their reliability. (P8 Line 1- 4)

Section 3.1.1

As it is stated now, it seems that the differences are caused by just the different approaches in the versions and the functions that the data are fitted to. Is that the conclusion? Are these differences connected to atmospheric conditions? Additionally, these approach of plotting just mean monthly values hides the real picture. At least some basic statistics (average and standard deviation of the differences at each bin) should be presented and discussed. Currently there is no information on the scattering of differences and if these differences are permanent biases.

Response:

Based on the study, we found that the most different physical process between V4.2 and V5 is that V5.0 introduced a constraint for the size distribution for it using the term (Sa-1(xk – xa)) in Eq. 4 and gives a strong constraint for the edge of size distribution, the values of the retrieved size distribution of the smallest size classes (r<0.05 μm) and the largest size classes (r > 10 μm) by V5.0 were close to zero.

We have added the scattering of VOL differences in the manuscript as shown in Fig.4 and Fig.5. We have also added some basic statistics (average and standard deviation of the differences at each bin) in Table 1 and Table 2. The percentage difference of the volume size distribution between SKYRAD V5.0 and V4.2 were larger than 50% at smaller size (r <0.025 μm at Qionghai, r <0.017 μm at Yucheng) and larger size (r >10μm at both sites). When the radius is between 0.17-5 μ m, the size distributions retrieved by V5.0 were in good agreement with those by V4.2. (P9 Line 5- 14)

Section 3.1.2

At this section there is no discussion on the causes of the differences. g. Are they explained strictly algorithmically or is there some natural process driving them? In most cases v5.0 retrieves lower SSA values suggesting the presence of more absorbing aerosols. Is there any evidence on that?

Response: In most cases V5.0 retrieving lower SSA values couldn’t suggest the presence of more absorbing aerosols.V5.0 tends to underestimate the SSA due to underestimation of the coarse aerosols when the a priori SDF for constraint tends to be close to zero for radii larger than 10μm.

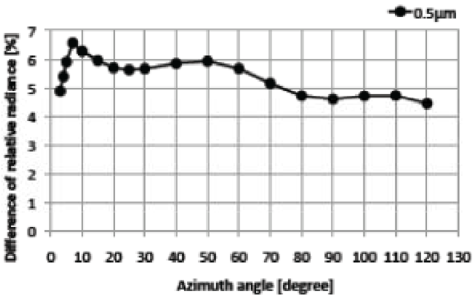

Fig. 8. The percentage difference of the relative radiances at 0.5 μm for each scattering angle between SDFs with and without particles over 10 μm in radius.

The above figure from the reference Hashimoto et al. (2012) shows that the difference between the relative intensity with and without a cut above 10 μm for the SDF (R is the relative radiance, △R=[R(cut above 10μm)- R(no cut above 10μm)]/ R(no cut above 10μm)). From this result, the lack of a large coarse part in the SDF causes overestimation of sky radiance at all observation angles. It is likely that V5.0 works to decrease the SSA value to dim the sky radiance in the calculation when a tight constraint on the SDF for particles with radius over 10 μm is applied. (P11 Line 17- 23)

P8l12-16 Why at Qionghai are only relative differences presented, while for Yucheng both relative and absolute are mentioned? Absolute differences are more important generally and it should be mentioned for both datasets. Also, as mentioned above, at least a standard deviation of these differences should be presented. From the scatter plots it is clear that the deviations are very diverse in each dataset.

Response: We have updated the tables, as the former table was got based on all the measurements, the new tables are based on the simultaneous observations of V4.2 and V5.0, which is consistent with the scatter plots. The standard deviations of absolute differences were added. (P32 and P33)

Figure 4. At Qionghai at 1020nm, it seems there two groups of measurements. One with values close to 1 and one with values lower than 0.8. Is there a physical explanation for this? It clearly needs more investigation this behavior. The same behavior at 1020 nm is also presented at figure 5 for refractive index.

Response: We haven't found a clear reason for this behavior. We need further investigation in the future work.

Section 3.1.3

Same as above. At least some discussion on the scattering of the data, since real part appears to have almost random differences. . Are the differences explained strictly algorithmically or is there some natural process driving them?

Response: In the retrieval, V4.2 found the optimum complex refractive index by trying several refractive indices. Complex refractive index in V4.2 can only be chosen from the predefined set of values in V4.2. In V5.0, complex refractive index were directly included in the state vector x, including constraints on the complex refractive index. As a priori estimation, mr usually be set as 1.5. (P12 Line 18- 21)

At present, we also haven't found a clear reason for the fact that real part appears to have almost random differences. We need further investigation in the future work.

Section 3.1.4

This sensitive test is an appropriate way to understand the algorithmical differences. But still no conclusion is drawn from this test and nothing is discussed in respect to the findings of previous sections.

Response: We have added some sensitivity tests for the main causes of error in the SSA and AOD retrieval by V5.0 and V4.2.We assumed an error of ±5% for calibration constant F0, ±5% for solid view angle SVA, ±50% (±0.05) for ground surface albedo Ag. We compared the differences in retrieved SSA values at a wavelength of 0.5 μm between cases with and without the assumed errors. On the basis of the sensitivity tests, it is concluded that an error in the calibration constant (F0) causes an error in both retrieved SSA and AOD. The averaged differences in retrieved SSA values due to ±5% error in F0 varied from 3% to 5%. An error of ±5% for solid view angle SVA introduced about ±2% differences in retrieved SSA values both by V4.2 and V5.0. Overestimation or underestimation in the Ag results in underestimation or overestimation of the SSA. An error of ±50% for ground surface albedo Ag caused about 1% averaged differences in retrieved SSA values both by V4.2 and V5.0. With the atmospheric pressure PRS increased by 1%, 2%, 3% and 4%, the averaged differences in SSAs didn't exceed 0.8%. (P14 Line 6- P16 Line 12)

Base on the inter-comparison results in Section 3.1 and the sensitivity tests in Section3.2, we couldn't get the conclusion that V5.0 is definitely better than V4.2. We haven't yet got other measurements in the two sites to help us prove that V5.0 is better than V4.2. The most different physical process between V4.2 and V5 is a derivation of particle size distribution. On the one hand, V5.0 tends to be robust to the cloud contamination, owing to inversion constraint by a priori SDF which filters out coarse particles to simulate cloud-scattered radiation. Some tests by Hashimoto et al (2012) showed that the SDF setting in V5.0 was useful for detecting ill-conditioned data caused by cirrus contaminations, horizontally and/or temporally inhomogeneous aerosol stratification, and so on (Hashimoto et al., 2012). On the other, due to a priori SDF for constraint tends to be zero for radii larger than 10μm, V5.0 will underestimate the coarse mode aerosols when a large amount of coarse particles of the dust-like aerosol type with radius greater than 10 μm exits. Estellés et al. (2018) found underestimation of the coarse aerosols by the V5.0 in African dust storm cases, whereas V4.2 retrieved coarse mode SDF similar to the observed one (Estellés et al., 2018). (P18 Line 4- 16)

Considering that V5.0 adopts more rigorous data processing and cloud detection methods, and the SSA and mi had high correlation coefficients between V4.2 and V5.0 with default the coarse mode radius rm2 value in V5.0 based on the above comparison results, we chose the retrieved results by V5.0 to analyze the seasonal variability of the aerosol optical properties over Qionghai and Yucheng. (P18 Line 17- 21)

We have added the above comments in the revised manuscript.

Figure 8. This approach also helps to increase the understanding of the algorithm. I suggest to plot real differences instead of absolute, because the sign is important to understand weather the version over or underestimates compared to the previous one. Also the lack of discussion on the uncertainty of the retrievals in the manuscript, makes it harder to interpret which range of difference is in the expected uncertainty.

Response: Following the reviewer's suggestion, we have replaced the absolute with real differences. (P17 Figure 10)

P13 l1-2. Also, why selecting v5.0 for studying the seasonal variability should be explained here.

Response:

Base on the inter-comparison results and the sensitivity tests, we couldn't get the conclusion that V5.0 is definitely better than V4.2. We haven't got other measurements in the two sites to help us prove that V5.0 is better than V4.2. On the one hand, V5.0 tends to be robust to the cloud contamination, owing to inversion constraint by a priori SDF which filters out coarse particles with radius greater than 10 μm. Some tests by Hashimoto et al (2012) showed that the SDF setting in V5.0 was useful for detecting ill-conditioned data caused by cirrus contaminations, horizontally and/or temporally inhomogeneous aerosol stratification, and so on (Hashimoto et al., 2012). On the other, due to a priori SDF for constraint tends to be zero for radii larger than 10μm, V5.0 will underestimate the coarse mode aerosols when a large amount of coarse particles of the dust-like aerosol type with radius greater than 10 μm exits. Estellés et al. (2018) found underestimation of the coarse aerosols by the V5.0 in African dust storm cases, whereas the version 4.2 retrieved coarse mode SDF similar to the observed one. (P18 Line 4- 16)

Considering that V5.0 adopts more rigorous data processing and cloud detection methods, and the SSA and mi had high correlation coefficients between V4.2 and V5.0 with default the coarse mode radius rm2 value in V5.0 based on the above comparison results, we chose the retrieved results by V5.0 to analyze the seasonal variability of the aerosol optical properties at the two sites. (P18 Line 17- 21)

Section 3.2.1

I honestly have a difficulty understanding why these section for PM2.5 have been added Surely it makes more clear the local emissions types, but it seems unlinked with the rest of the study. Findings mentioned here are nowhere used to explain anything about SKYNET retrievals and their behavior. Also, it is an unexplained decision to study PM 2.5 while from all SKYNET retrievals there is a picture of constant dominance of larger particles in both regions. Unless you could integrate the findings to the discussion in the rest of the study, linked differences found between the two sites and preferably even connect the deviations between the two versions with the types of aerosols, I suggest removing this section.

Response: Thank you for your kind comments. We have removed this section.

Section 3.2.2

Which months are considered in each season should be defined. Also, the discussion about the humidity is not clear. It seems that aerosol loads are generally in the same order throughout the year and humidity causes the variations of AOD. This needs more evidence to support it and a lot of discussion and data are needed to prove it. If this is not the case, please restate to make clear the finding of this analysis.

Response: Four seasons were considered in this paper (i.e., spring (March-May), summer (June-August), autumn (September-November), and winter (December-February)) to investigate the seasonal variations of the aerosol optical properties over Qionghai and Yucheng. (P18 Line 24-26)

We have restated the analysis as follows:

The maximum AOD average of 0.99 occurring in summer, several factories which produced inorganic and organic fertilizers located, the stronger sunlight in summer accelerated the photochemical reaction and enhanced the formation of fine particulate nitrate (Wen et al., 2015), the humidity in summer is higher than other seasons over Yucheng (Meng et al., 2007), high humidity combined with large fractions of hygroscopic chemical components (e.g. sulfate, nitrate, ammonium, and some organic matters) can enhance light extinction and haze intensity the scattering coefficient of secondary inorganic aerosols (such as sulfate, nitrate and ammonium) (Tao et al., 2017) .(P19 Line10-16)

Section 3.2.3

The discussion in this paragraph is not consistent with the next section. Since lower SSA values in winter are explained by the presence of carbonaceous particles, why winter SDF are dominated by coarse mode in both regions?

Response: The volume size distribution in both sites presented bimodal patterns with a 0.1-0.2 μm fine particle mode and a 3.0-6.0 μm coarse particle mode in four seasons, and the volume of the coarse aerosol particles relative to the whole was larger, especially in Yucheng. Carbonaceous particles have higher values than other seasons in both sites due to winter heating and regional transport, black carbon aerosols are dominated by absorption effect, so the seasonal SSAs in winter are lower than other seasons in winter, but it hadn’t changed the fact that the two regions are dominated by coarse particles. With winter heating, due to incomplete combustion, in addition to black carbon aerosols, there will also be some dust.