We are grateful to the editor and anonymous referees for their comprehensive and constructive comments on our work. All of these comments and concerns raised by the referee have been explicitly considered and incorporated into this revision.

Anonymous Referee #1

This paper expands upon the authors' previous aerosol-components retrieval (Zhang et al., Atm Env, 2018) by including sodium chloride as a coarse aerosol component. The authors apply their results to about 16 SONET sites all across China. The grammar is clear and for the most part the paper is very well written. This is a good paper that is suitable for publication in ACP after some modifications.

The authors cite Zhang 2018 for their methodology, but I am not exactly sure of their approach. I gather that they use the Zhang 2018 approach to determine separate complex refractive indices (CRI) for the fine and coarse modes from the SONET data. Then for the coarse mode, they use RH to determine the equilibrium mixture ratio of NaCL with water, which has a certain real refractive index (RRI). Once the RRI for the water-NaCl mixture is known, they can iterate the dust mixing ratio until they minimize the χ^2 of Eq 12. They are using a single "dust," though, so they cannot vary the IRI independent of the RRI; thus, they have limited adjustability for the spectral dependence of the CRI. This is all very reasonable, but the use of a single "dust" will sometimes increase their residuals. That is ok, though, as residual values can be monitored and retrievals can be rejected on the basis of residual values when necessary. I am comfortable with their coarse mode methodology.

I am having difficulty understanding the fine mode retrieval methodology, though, which is my biggest reservation about this paper. The authors claim to separate water-soluble organic carbon (WSOC) from ammonium nitrate (AN), but it is not clear to me how this can be accomplished without a specific assumption for the hygroscopicity parameter (kappa) of WSOC. If this is what the authors are doing, they need to specifically state this and provide the reader with the value of kappa that they chose for WSOC (as well as the rationale for using a certain kappa, and some discussion of the repercussions of using the wrong kappa in their retrieval). The authors cite (Zhang 2018), but a brief overview of the Zhang approach for the fine mode in the methodology section would be helpful.

Major Issues

1. It is not clear to me how the "derived hygroscopic parameter kappa" is obtained (p2, line 59, and Table 2). I believe the authors are deriving the Table 2 values from Equation 4, but that requires the hygroscopicity parameters of the components (κ_i); the authors say that these values can be computed by the component hygroscopic growth factors (lines 144-145 and Eq 5). However, I don't see how these component growth factors can be derived from their data, so I am assuming that they are obtaining these values from the literature. If this is the case, the authors should provide the reader with the GFi or κ_i that they use in the retrieval. Otherwise, they should provide additional details about how they obtained the κ_i with the sun photometer data.

Response: Thanks for the reviewer's comment. The values of κ_i are obtained from the literature and listed in table 2. For clarity, we have revised the manuscript as follows:

Line 156-157: "…where κ_i is the hygroscopic parameter of the *i*th component obtained from the literature (table 2), and $f_{dry,i}$ is the dry component volume fraction defined as … "

2. Figure 3:

I was confused by the “non-hygroscopic components” that are a subset of the “water-soluble components” and are the entire basis of the “water-soluble organic matter (WSOM)” â˘A˘T I am not a chemist, so I found it odd that water-soluble aerosols could be non-hygroscopic. It would be helpful to some readers (like myself) if the authors spent a sentence or two alerting the reader that water-soluble aerosols are sometimes non-hygroscopic. If they can explain the physics behind this phenomena, that would be even better.

Personally, I am skeptical about separating WSOM from AN using remote sensing techniques. From an optical standpoint, such a mixture would merely be a solution with an effective hygroscopicity parameter (kappa). Knowledge of RH and an assumed kappa allows one to derive the solute mixing ratio (and growth factor) via Eq 2, but I don’t see how one can separate the effects of multiple soluble components with the available remote sensing information (refractive index and RH) without additional assumptions (like the _i for each component).

Response: For the non-hygroscopicity of organic components, like vinegar and alcohol, they are water-soluble but more volatile than hygroscopic. In aerosols, the dicarboxylic acids are dominant in the WSOM component, and oxalic acid accounts for more than 50% in dicarboxylic acids, followed by succinic, malonic, maleic, adipic and phthalic acids (Chebbi & Carlier, 1996). Under ambient relative humidity (RH) of 10 - 90%, oxalic acid hardly deliquesce and exhibit low hygroscopicity, and others also have low hygroscopicity except malonic acid as shown in Figure R1 (Ma et al., 2013; Drozd et al., 2014; Jing et al., 2016). Considering the abundance of oxalic acid (more than 50% in dicarboxylic acids) and other low hygroscopic components in the atmospheric aerosols (Sullivan et al., 2009; Fu et al., 2013), WSOM is treated as a non-hygroscopic component in aerosol particles. We explained this point in the manuscript as follows:

Line 124-128: “It should be noted that the water-soluble property of aerosol components is not equivalent to hygroscopicity. Dicarboxylic acids represented by oxalic acid are dominant in the WSOM component but their hygroscopicity is extremely low (Ma et al., 2013; Drozd et al., 2014; Jing et al., 2016). Also other organic compounds in aerosols are less hygroscopic as shown in Zhang et al. (2018) (their figure 1). Hence, the OM components (WSOM and WIOM) are treated as non-hygroscopic components.”

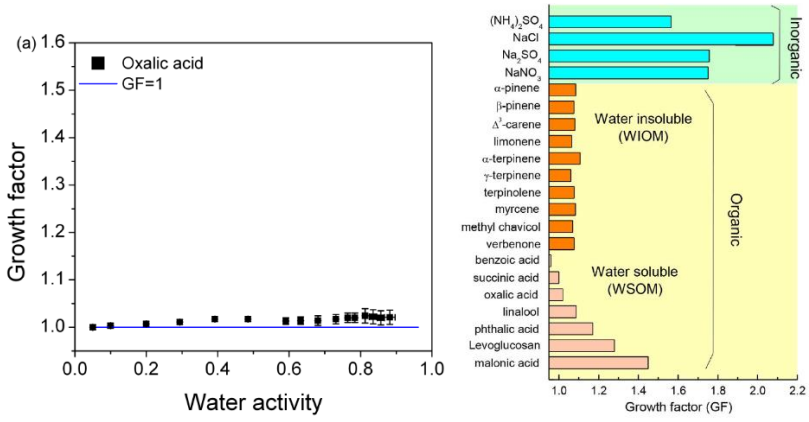

Figure R1. (a) Hygroscopic growth factors of oxalic acid from figure 2 of Jing et al. (2016); (b) Different geometric hygroscopic growth factor between inorganic and organic aerosols with RH around 85% of relative humidity from figure 1 of Zhang et al. (2018).

For the separation of the WSOM and AN components, the refractive index (RI) from remote sensing is used indirectly. From Eq. 2-5, we can get the relationship among the volume fraction of the components in the host. Because the sum of the volume fractions of these components can be equal to 1 in the host, the RI of the host can be calculated using Eq. 6-9. And then we can calculate the RI of the entire particle using the effective medium approximation (Eq. 10-12). Adjusting this fraction and comparing with the RI from the SONET observations, the volume fraction for all components can be obtained. The description and flowchart of the inversion procedure are added in section 3.4.

Ma Q., He H., Liu C. (2013). Hygroscopic properties of oxalic acid and atmospherically relevant oxalates, Atmospheric Environment, 69, 281-288.

Drozd G., Woo J., Hakkinen S.A.K., Nenes A., and McNeill V.F. (2014). Inorganic salts interact with oxalic acid in submicron particles to form material with low hygroscopicity and volatility, Atmos. Chem. Phys., 14, 5205–5215.

Chebbi A. & Carlier P. (1996). Carboxylic acids the troposphere, occurrences, sources, and sinks: a review, Atmospheric Environment, 30(24), 4233-4249.

Jing B., Tong S., Liu Q., Li K., Wang W., Zhang Y. and Ge M. (2016). Hygroscopic behavior of multicomponent organic aerosols and their internal mixtures with ammonium sulfate, Atmos. Chem. Phys., 16, 4101–4118.

Fu, P., Kawamura, K., Usukura, K., and Miura, K.: Dicarboxylic acids, ketocarboxylic acids and glyoxal in the marine aerosols collected during a round-the-world cruise, Mar. Chem., 148, 22–32, 2013.

Sullivan, R. C., Moore, M. J. K., Petters, M. D., Kreidenweis, S. M., Roberts, G. C., and Prather, K. A.: Timescale for hygroscopic conversion of calcite mineral particles through heterogeneous reaction with nitric acid, Phys. Chem. Chem. Phys., 11, 7826–7837, 2009.

3. Line 115, authors state:

"For fine mode, the water-insoluble and water-soluble components are identified using an empirical function (Zhang et al., 2018)"

How? The authors need to expand this a little. I checked the Zhang 2018, and I was not able to quickly determine how WI and WS components were separated. At a minimum, the authors should point to the specific section number in Zhang (2018), but it would be best to provide the readers with a brief recapitulation in order to best hold their interest.

Response: Thanks for the reviewer's comment. We added an explanation of the empirical function in the manuscript and add details to the supplementary as follows:

Line 115-118: "For the fine mode fraction, the water-insoluble and water-soluble components are identified using an empirical function (see section 2.2.2 in Zhang et al., 2018), which describes the ratio of the water-soluble to the water-insoluble volume fractions determined by RH, together with the parameterization of aerosol soluble volume fractions by Kandler and Schutz (2007)."

4. Section 3.3:

The forward model is described well in Section 3.2, but the inversion section (3.3) is very light. For instance, the authors cover the relationship of the real refractive index to molar refractivity in Sect 3.2, but none of that shows up in Section 3.3. Presumably the authors are using RH to partition between the soluble components and water and

also to assign a RRI to the host solution prior to the minimization procedure described in section 3.2 (which requires refractive indices of both the host solution and the insoluble inclusions). None of that is stated here, though, so as a reader I am not sure if I have this correct.

Response: Thanks for the reviewer's comment. We adjusted the structure of the methods section and added a description of the inversion procedure in both manuscript and supplementary as follows:

Line 179-216:

"3.3 Effective medium approximation

To determine the complex refractive index of a particle, i.e. including both the multi-component liquid system and water-insoluble matter, the complex refractive index ($m = n - ik$) at wavelength λ is expressed in terms of the permittivity, $\varepsilon(\lambda)$:

$$m(\lambda) = \sqrt{\frac{|\varepsilon(\lambda)| + Re(\varepsilon(\lambda))}{2}} + i\sqrt{\frac{|\varepsilon(\lambda)| - Re(\varepsilon(\lambda))}{2}} \tag{10}$$

The permittivity of the multi-component liquid system can then be calculated using equations (8) - (10). Considering the water-insoluble matter in a particle as inclusion and the water-soluble matter as the environment, the permittivity of the entire aerosol particle can be obtained by the Maxwell Garnett effective medium approximation (Schuster et al., 2005).

$$\varepsilon_{eff}(\lambda) = \varepsilon_e + 3\varepsilon_e \left[\frac{\sum_j \frac{\varepsilon_j(\lambda) - \varepsilon_e(\lambda)}{\varepsilon_j(\lambda) + 2\varepsilon_e(\lambda)} f_j}{1 - \sum_j \frac{\varepsilon_j(\lambda) - \varepsilon_e(\lambda)}{\varepsilon_j(\lambda) + 2\varepsilon_e(\lambda)} f_j} \right] \tag{11}$$

where, j is the number of water insoluble components and. $\varepsilon_j(\lambda)$ and $\varepsilon_e(\lambda)$ are the permittivities of the inclusion and its environment. The complex refractive index of the entire aerosol is estimated by aerosol component fraction using equation (10).

3.4 Inversion procedure

The flow chart for the inversion of the aerosol components is shown in figure 4. In the fine mode, the ratio of WS and WI matter is estimated using RH as described in section 2.2.2 in Zhang et al. (2018). The initial value of the host refractive index and the extreme value for the BC component are set by the calculation modules of the complex refractive index in the multicomponent liquid system (see section 3.2) and the effective medium approximation (see section 3.3), respectively. In the loop to determine the BC component, two constraints are applied to separate BC from other components. The WSOM/WIOM ratio constraint was developed by Zhang et al. (2018) based on considerations published in the literature (Chalbot et al., 2016; Bougiatioti et al., 2013; Wozniak et al., 2013; Mayol-Bracero et al., 2002; Krivácsy et al., 2001; Zappoli et al., 1999):

$$\begin{cases} f_{WSOM} \cong \alpha f_{WIOM} \\ \alpha = \dfrac{\beta \rho_{WSOM}^{-1}}{1 - \beta \rho_{WSOM}^{-1}} \end{cases} \qquad \beta \in [44\%, 77\%] \tag{12}$$

For more detail, see section 2.3.1 in Zhang et al. (2018). The volume normalization of the aerosol components in both the fine and coarse modes is used to constrain the volume fraction of the aerosol components to a reasonable range (similar as section 2.3.2 in Zhang et al., 2018)

$$\begin{cases} f_{fine} + f_{coarse} = 1.0 \\ f_{fine} = f_{BC} + f_{AN} + f_{WSOM} + f_{WIOM} + f_{AW_f} \\ f_{coarse} = f_{DU} + f_{SC} + f_{AW_c} \end{cases} \quad (13)$$

Then the inner loop of WSOM computes the CRIs of the fine mode at different BCs, and output the aerosol components of minimum χ^2. The inversion procedure for the coarse mode is simpler than that for the fine mode. There is only a loop for DU and the complex refractive index of the host can be directly calculated by equations (2) - (8) with only input of RH. The function Chi-squared (χ^2) as an iterative kernel function is expressed in the sum of the differences between the complex refractive index estimated from the forward model (m) and the retrievals (m_{rtrl}), at multiple wavelengths:

$$\chi^2 = \sum_{\lambda} \frac{\left(m_{rtrl}(\lambda) - m(\lambda)\right)^2}{m_{rtrl}(\lambda)} \quad (14)$$

λ=440, 675, 870 and 1020 nm

The retrieval is completed when the value of χ^2 reaches a minimum. The volume fractions of the aerosol components can be obtained by solving the above equations (10-12). The aerosol mass concentration in the atmospheric column is calculated using the volume and effective density of the aerosol components."

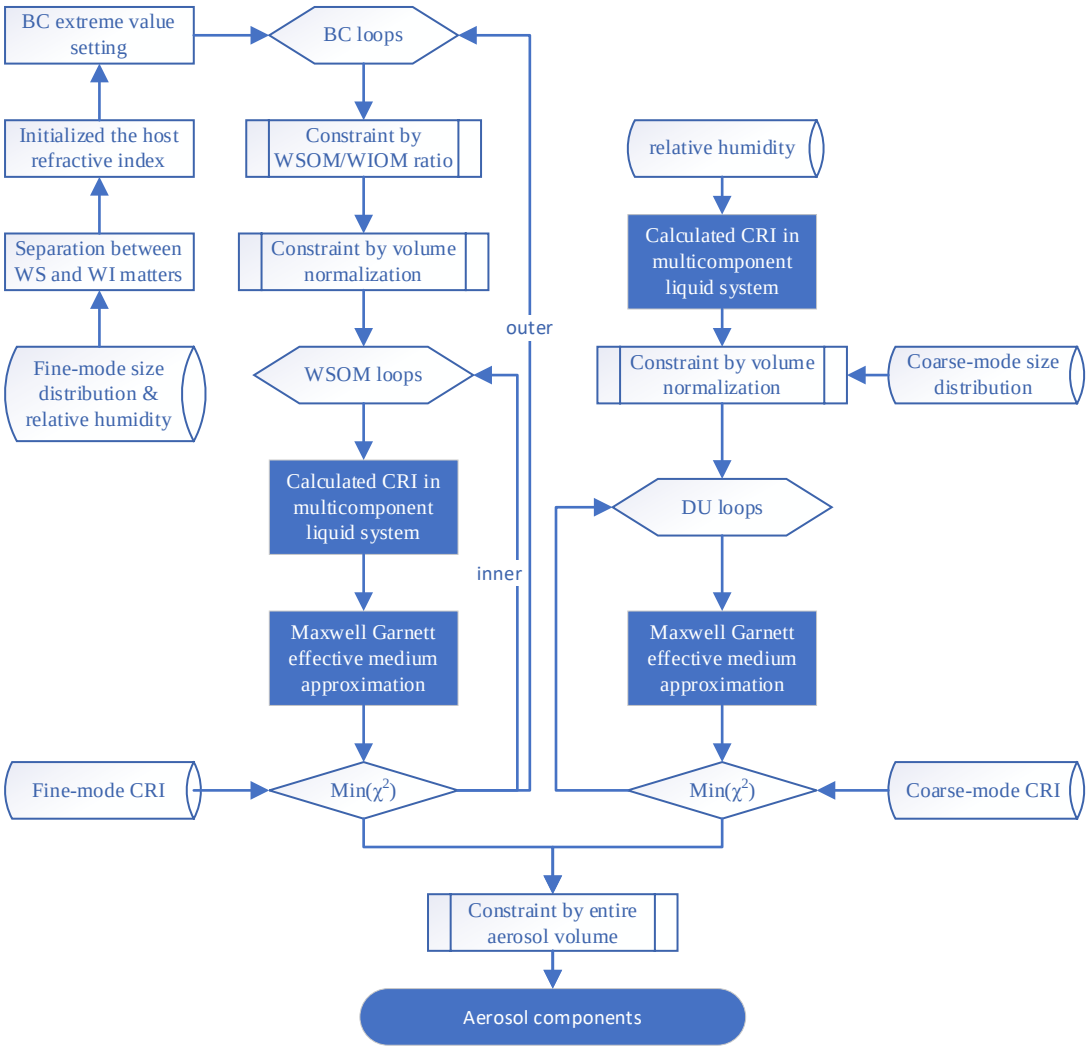

Figure 4. Flowchart of the aerosol component classification inversion algorithm.

5. Table 3:

Why is the RRI of NaCl and the coarse water so uncertain? I thought we had some good measurements of these species. Even if we didn't, how do we get 900

Response: Thanks for pointing this out. This large error (more than 900%) occurs only when the input error is large. For separation of high scattering components (NaCl and coarse water), the relative humidity (RH) and real part of CRI in coarse mode (n_c) are important dependent parameters and weakly correlated with other parameters. A large error for RH (10%) is applied to estimate the scattering components, which can be reduced because the actual observation error is usually lower than 5%. For the error of n_c summarized from Zhang et al. (2017) listed in table S2, it is large for the WS typical model, which can also affect the scattering component separation. However, if the accuracy of RH is improved, the errors of NaCl and coarse water can be effectively suppressed. Meanwhile, it should be noted that this estimation of uncertainty is for a single measurement. One of important advantages of a remote sensing approach is to perform multiple measurements quickly in a short time period. Thus, the average uncertainty of the aerosol components can be effectively reduced taking account independent errors in each observation. We illustrate this point in the manuscript:

Line 270-275: “Fortunately, the RH observed by ground-based stations is accurate, with an error which is usually less than about 5% (WMO, 2008), which can significantly reduce the uncertainty in the retrieved aerosol scattering components. It should be noted that the uncertainties in table 3 are for single measurements. One important advantage of remote sensing is that multiple measurements can be made during a short period of time. Thus, the average uncertainty of the aerosol components can be effectively reduced by taking into account independent errors in each observation. In addition, the accuracy of the retrieved n_c needs to be improved in order to deal with the aerosol component inversion.”

Minor Issues

1. Page 1, line 28, authors state:

“Optical remote sensing techniques do not provide sufficient information for a detailed analysis of chemical composition and therefore refrain to the retrieval of components describing specific properties"

My interpretation of this sentence is that we can not retrieve aerosol composition from remote sensing techniques, but I am sure that is not the authors intent (otherwise, we don’t need to read the paper). Consider rephrasing.

Response: Thanks for reviewer’s comment. This sentence was deleted.

2. Page 2, line 37:

Schuster (2009) is entitled “Remote Sensing of Aerosol Water Uptake,” and does not directly address dust.

Response: Thanks for reviewer’s comment. We revised this sentence:

Line 36-41: “In a follow-up study, Schuster et al. (2009) applied a similar procedure to determine the aerosol water fraction by fitting the real part of the refractive index of an internal mixture of water, soluble and insoluble species to observations by minimizing the cost function at all four wavelengths together. In this work the ratio of the dry volume fraction of insoluble to that of soluble aerosols was constrained by using a climatological value and the real refractive index which also prescribes the aerosol hygroscopicity. This constraint also provides a maximum insoluble fraction and the fraction of dust aerosol.”

3. Page 3, line 82:

Should also reference Dubovik, O., and M. King (2000), A flexible inversion algorithm for retrieval of aerosol optical properties from sun and sky radiance measurements, J. Geophys. Res., 105(D16), 20,673–20,696.

Response: We add this reference in manuscript.

Line 85-86: “Based on the inversion algorithm of Dubovik and King (2000) and Dubovik et al. (2000), the 440, 675, 870 and 1020 nm wavebands are used to … ”

4. Page 3, lines 87-88.

I don’t understand the meaning of these lines:

“Using these data, PVSD and CRI sub-modal parameters of atmospheric aerosols are obtained using the modal decomposition method proposed by Zhang et al. (2017). Using the sub-modal characteristics data set thus obtained, an aerosol sub-modal model was established for China by Li et al. (2019), but the submodal aerosol components have not been given."

So a sub-model model was established but not given?

Response: We revised this sentence:

Line 91-93: “Using these fine and coarse mode characteristics of the CRI, micro-physical properties of aerosols in each mode were analyzed (Li et al., 2019), but the aerosol chemical components were not determined.”

5. Page 6, line 150:

Please replace “refractive index” with “real refractive index.” Page 6, line 156:

Please tell the reader that “n" is the real refractive index.

Response: We revised as follows:

Line 162-164: “Firstly, the molar refractivity (A_e) at wavelength λ can be calculated from the real part of the complex refractive index (n_i) and the volume fraction of the individual components”

6. Equation 10:

Equation 4 uses the symbol epsilon as the component dry volume fraction, whereas here it is the permittivity. Need to change the symbol used for dry volume fraction in Eq 4 and everywhere.

Response: We change the symbol used for dry volume fraction to $f_{dry,i}$ in Eq 4 and 5. The corresponding descriptions have also been modified.

Line 153-160: “In the multicomponent liquid system, the hygroscopicity parameter κ is given by the simple mixing rule

$$\kappa = \sum_i f_{dry,i}\kappa_i \quad (4)$$

where κ_i is the hygroscopicity parameter of the ith component obtained from the literature (table 2), and $f_{dry,i}$ is the dry component volume fraction defined as

$$f_{dry,i} = \frac{V_i}{V_s}$$

”

7. Equation 12:

The numerator should be squared. Otherwise, large negative differences will produce the “best” χ^2.

Response: Thanks for pointing this out. We missed the square when wrote the formula, but we used the formula where the numerator is the square.

8. Table S1:

What is the basis for the numbers in Table S1? That is, which climatology are you using to define WS, BB, and DU?

Response: Considering that the algorithm results from Dubovik et al. (2000) are taken as input in this study, the aerosol model is consistent with Dubovik's. In table S1, the parameters of size distribution are as same as those in table 2 of Dubovik et al. (2000), and the complex refractive index in the fine and coarse modes are further calculated using Dubovik's aerosol model from Zhang et al. (2017). We give more detail on this point in the supplementary.

9. Page 9, lines 238-241:

Does it make sense to quantitatively discuss BC in the context of Fig 5? BC barely exists in that figure. I recommend adding a table or an additional figure for BC.

Response: Thanks for reviewer's comment. All component fractions are listed in the table S3 and the BC fraction is added in figure 7 (Figure 5 from the previous manuscript) as follows:

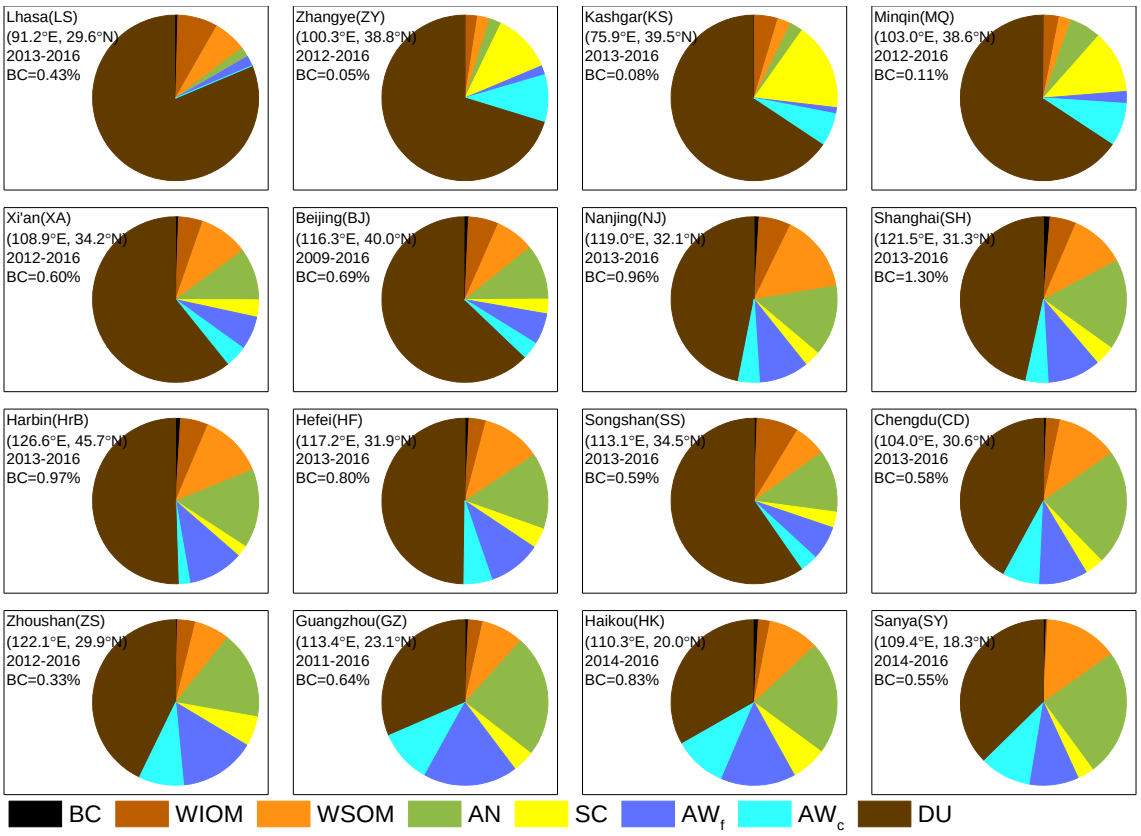

Figure 7. The averaged mass fraction of aerosol components at SONET sites. The site name, location, observation period and BC fraction are marked in each subgraph. The mass fractions of other components are listed in table S3..

10. Figure 6:

How is the color scale in Fig 6 normalized (range is 0 to 20)? Also, it is odd that some of the “estimated" values are so far off when you are using CRI as a constraint. The χ^2 must be very high in these cases. It would make sense to have a residual requirement (i.e., χ^2 < some threshold) and to throw out high values of χ^2. This should also improve your statistics (slope, intercept, bias, etc.).

Response: Thanks very much for your comments. We set a threshold and filter the results and revised as follows:

Line 303-321: “

The closure of the CRI between instantaneous optical-physical inversion and chemical estimation is examined by the data pair frequency. Figure 8 shows scatter density plots of the chemically estimated and sunphotometer-retrieved imaginary parts of the fine mode at 675 nm (k_f) and 440 nm ($k_{f,440}$) and the real parts of fine mode at 440 nm(n_f). The points are colored by the number of data pairs (Retrieved, Estimated), which are sorted according to ordered pairs in 0.0005 intervals for the imaginary parts of CRI and 0.001 intervals for the real parts. The data pairs of k_f are closely concentrated around the 1:1 line, although with slight underestimation with 94.3% of the estimated values lower than the retrieved values; only 5.3% of the data pairs have a relatively large absolute error (AE > 0.01). The mean bias is not large (-0.003), and the mean absolute value is equal to the mean absolute error (MAE = 0.003). There are two reasons for this slight underestimation in chemical estimation. On the one hand, the imaginary part of the refractive index of BC is much larger than for the other components due to its strong absorption. Thus, the inversion of the BC concentration is very sensitive to the estimation of the refractive index. As shown in table 3, although the TRE of BC is the lowest, the errors caused by k_f and $k_{f,440}$ are larger than for any other component. On the other hand, k_f is not only affected by BC in the inversion process, but also affected by organic components (WSOM & WIOM) with spectral absorption characteristics. Therefore, in most cases, k_f is underestimated in chemical estimation and $k_{f,440}$ is overestimated (Bias = 0.007). The mean relative error (RE) is 27.1%, and 62.8% of the data points are below the average relative error line. This indicates that most inversion results have good optical closure. For the closure of the real part of the fine mode, the data pairs of n_f are also concentrated around the identity line, although 76.5% of the n_f is above the identity line. Underestimation occurs mainly when n_f is larger than 1.56, because the only component with the real part of the CRI larger than 1.56 is BC, but its concentration is mainly determined by the imaginary part. The bias of the estimated n_f (Bias = 0.009) is larger than that of k_f due to the fact that the value and the range of n_f are larger than that of k_f. ”

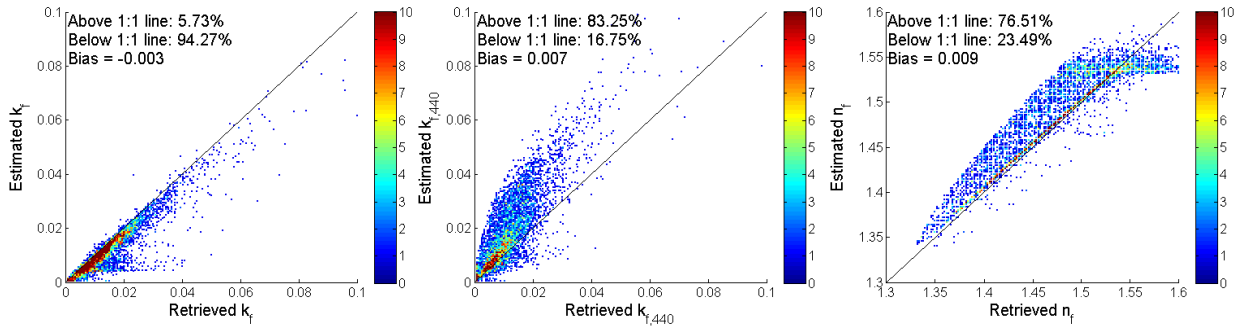

Figure 8. Data pair frequency of instantaneous imaginary parts of the complex refractive index at 675 nm (k_f), 440 nm ($k_{f,440}$), and real part at 440 nm (n_f) which are sorted according to ordered pairs (Retrieved, Estimated) in 0.0005 and 0.001 intervals for imaginary and real parts, respectively. “Retrieved” represents sub-component of CRI from the optical-physical

retrievals, and “Estimated” is estimated by retrieved chemical components. The color represents the number of cases (color bar), and the solid black line shows the 1:1 line.

11. Lines 248-249: Authors state

“As shown in table 3, although the TRE of BC is the lowest, it also causes the largest kf **and** kf;440 **errors."**

TRE is total relative error, so how can TRE cause kf and kf;440 errors? Shouldn’t cause and effect be the other way around (i.e., k errors cause TRE errors).

Response: We revised this sentence as follows:

Line 312-313: “As shown in table 3, although the TRE of BC is the lowest, the errors caused by k_f and $k_{f,440}$ are larger than for any other component.”

12. Lines 252-253, Authors state:

“This indicates that most inversion results have good optical closure, and the aerosol components retrieved by the remote sensing method used in this study should be reasonable."

This line refers to Fig 6, which is a plot of how well the component-averaged imaginary index compares to the imaginary refractive index that is used as input. Thus, a good comparison just means that you usually have good residuals (i.e., low χ^2). Fig 6 does not assure reasonableness of all components in the retrieval, though, as it only shows the imaginary RI, and most of the components of this retrieval are not sensitive to IRI. The only thing that we can claim via Fig 6 is that the retrieval might be getting BC correct. Additionally, we can’t use Fig 6 to argue that the BC mass or volume fractions are correct, as these are sensitive BC refractive index. However, you can use Fig 6 to argue that you are getting the BC AAOD correct; this is because you are using IRI as a constraint, and the IRI that you retrieve will always be the same (as long as BC has a spectrally flat IRI and your other absorbers do not).

Response: Thanks for the reviewer's detailed explanation. We revised this sentence as follows:

Line 317: “This indicates that most inversion results have good optical closure.”

13. Line 275:

I believe that the word “autumn" should be replaced with “spring."

Response: Thanks for pointing out this mistake. We revised as follows:

Line 342-343: “Low concentrations of BC in the other seasons are mainly due to the influence of frequent dust events in the spring and high aerosol hygroscopic growth in the summer.”

14. Table 3:

Total relative error is defined with 7 parameters. Presumably this is nf, kf , kf;440, nc, kc, kc;440, and RH. Are the (nf ; kf) averaged from the 675, 870, and 1020 nm wavelengths, then? I don’t recall seeing this explicitly stated.

Response: Our method of separating fine and coarse modal CRI parameters (Zhang et al, 2017) can only separate 6

of these parameters (nf, kf , kf;440, nc, kc, kc;440). Thus we assume that the imaginary parts of CRI at other wavebands except 440nm are invariable. We add a statement for CRI sub-modal parameters as follows:

Lien 90-91: "The real parts of the CRI of the fine and coarse modes (n_f and n_c, respectively) are spectrally independent, while the imaginary parts have spectral variation at 440 nm, so they are written as ($k_{f,440}$, k_f) & ($k_{c,440}$, k_c)."

15. Figure 2:

Caption should describe the timeframe of the boxplots.

Response: we revised the caption of figure 2 as follows:

"**Figure 2**. Boxplots of the relative humidities observed near each of the SONET sites. The observation periods for each site are shown in table 1. The line and the diamond represent the median and mean values, respectively, and the box shows the standard deviation (1 σ)."

16. Figure 4:

Do the pie charts correspond to both the fine and coarse modes? If so, why isn't there any AWc or SC in the WS and BB pie charts? If not, why does dust dominate over WIOM for those species?

Response: Yes, this figure shows both the fine and coarse modes. Few studies have given the fraction of water content in coarse mode which SC is closely related to, so we set it very low in the typical model.

17. Figure 7:

Throughout the text, authors use SC for sodium chloride. Here, they do not show SC but show SS (sea salt?).

Response: Thanks for pointing this out. We revised the figure.

Anonymous Referee #3

Article is well in the scope of the ACP journal and discusses a method for retrieval of aerosol chemical composition from remote sensing. Aerosol chemical composition is in a grand demand by a scientific society providing data to evaluate global atmospheric modelling and climatic models. Presented method introduced some of the novelties and meanwhile is based on a well-established and proven techniques. Paper is well referenced at least to my knowledge, although I've found some inaccuracies (see in notes below).

At the same time, I have a feeling that paper doesn't reflect fully the potential of the method and the complexity of the work performed by authors, and definitely doesn't provide enough explication to reproduce the described method. Article is very brief and dry, to the extent it sometimes hard to read and what is more important to understand concepts of the study preformed. Also, I've found out quite a number (for a short paper of 10 pages) of typos, please, consider general grammar revision, some of them I've listed below.

I would recommend this article for publication after major revision, which in my opinion should increase the impact and significance of the presented results.

Below I've highlighted some points where there is a room for improvement.

Major comments:

Formula 12: Possible typo, square is missing in numerator. Either it is missing either it is no chi-squared function. Please, clarify. If no square formula was actually used for the study, I'm not sure the method is legit as opposite sign errors in refractive indices at different wavelengths can compensate each other. Minimisation procedure of chi2 function is not described at all, at least mention how it is done. What is "iterative kernel function", Is it LUT or a version of Newtonian method? is it the same that was used in previous studies? Please clarify or refer.

Whole section 3 is rather confusing, despite of illustrations and formulas it doesn't give a clear understanding of a methodology, just some of its pieces. Also, if this method is an improvement of an existing one, it should be clearly referred and changes introduced highlighted. Please, consider re-formulating this part to make it clearer. Figure 3 and Aerosol classification in section 3.1, it is hard to understand at which part of the retrieval this classification is used. And what parameters it used to classify aerosol in groups? Is it part of forward modelling? Does it use PVD provided by SONET, other part of paper claims only refractive index is used … And section 3.3 refers to refractive index of fine and coarse mode …

Also "combined" aerosol types such as BC, OM, IS and AW are used only in figure 3 and not referenced anywhere else in the paper. Why then there are presented?

Personally, I would prefer here a flowchart of the retrieval algorithm in general, i.e. where it would be easy to understand what parameters enter it, how refractive index is modelled and fitted and what parameters are coming out.

Response: Thank you very much for pointing this out.

(1) Formula 12 is indeed a typo. We have revised it in the manuscript as follows:

Line 211:"

$$\chi^2 = \sum_{\lambda} \frac{\left(m_{rtrl}(\lambda) - m(\lambda)\right)^2}{m_{rtrl}(\lambda)}$$

λ=440, 675, 870 and 1020 nm (12)"

(2) In order to present our algorithm more clearly, we have added the flowchart in Figure 4 and corresponding descriptions

(section 3.4).

Line 192-216:

"3.4 Inversion procedure

The flow chart for the inversion of the aerosol components is shown in figure 4. In the fine mode, the ratio of WS and WI matter is estimated using RH as described in section 2.2.2 in Zhang et al. (2018). The initial value of the host refractive index and the extreme value for the BC component are set by the calculation modules of the complex refractive index in the multicomponent liquid system (see section 3.2) and the effective medium approximation (see section 3.3), respectively. In the loop to determine the BC component, two constraints are applied to separate BC from other components. The WSOM/WIOM ratio constraint was developed by Zhang et al. (2018) based on considerations published in the literature (Chalbot et al., 2016; Bougiatioti et al., 2013; Wozniak et al., 2013; Mayol-Bracero et al., 2002; Krivácsy et al., 2001; Zappoli et al., 1999):

$$\begin{cases} f_{WSOM} \cong \alpha f_{WIOM} \\ \alpha = \dfrac{\beta \rho_{WSOM}^{-1}}{1 - \beta \rho_{WSOM}^{-1}} \end{cases} \qquad \beta \in [44\%, 77\%]$$

(12)

For more detail, see section 2.3.1 in Zhang et al. (2018). The volume normalization of the aerosol components in both the fine and coarse modes is used to constrain the volume fraction of the aerosol components to a reasonable range (similar as section 2.3.2 in Zhang et al., 2018)

$$\begin{cases} f_{fine} + f_{coarse} = 1.0 \\ f_{fine} = f_{BC} + f_{AN} + f_{WSOM} + f_{WIOM} + f_{AW_f} \\ f_{coarse} = f_{DU} + f_{SC} + f_{AW_c} \end{cases}$$

(13)

Then the inner loop of WSOM computes the CRIs of the fine mode at different BCs, and output the aerosol components of minimum χ^2. The inversion procedure for the coarse mode is simpler than that for the fine mode. There is only a loop for DU and the complex refractive index of the host can be directly calculated by equations (2) - (8) with only input of RH. The function Chi-squared (χ^2) as an iterative kernel function is expressed in the sum of the differences between the complex refractive index estimated from the forward model (m) and the retrievals (m_{rtrl}), at multiple wavelengths:

$$\chi^2 = \sum_{\lambda} \frac{\left(m_{rtrl}(\lambda) - m(\lambda)\right)^2}{m_{rtrl}(\lambda)}$$

λ=440, 675, 870 and 1020 nm (14)

The retrieval is completed when the value of χ^2 reaches a minimum. The volume fractions of the aerosol components can be obtained by solving the above equations (10-12). The aerosol mass concentration in the atmospheric column is calculated using the volume and effective density of the aerosol components.**"**

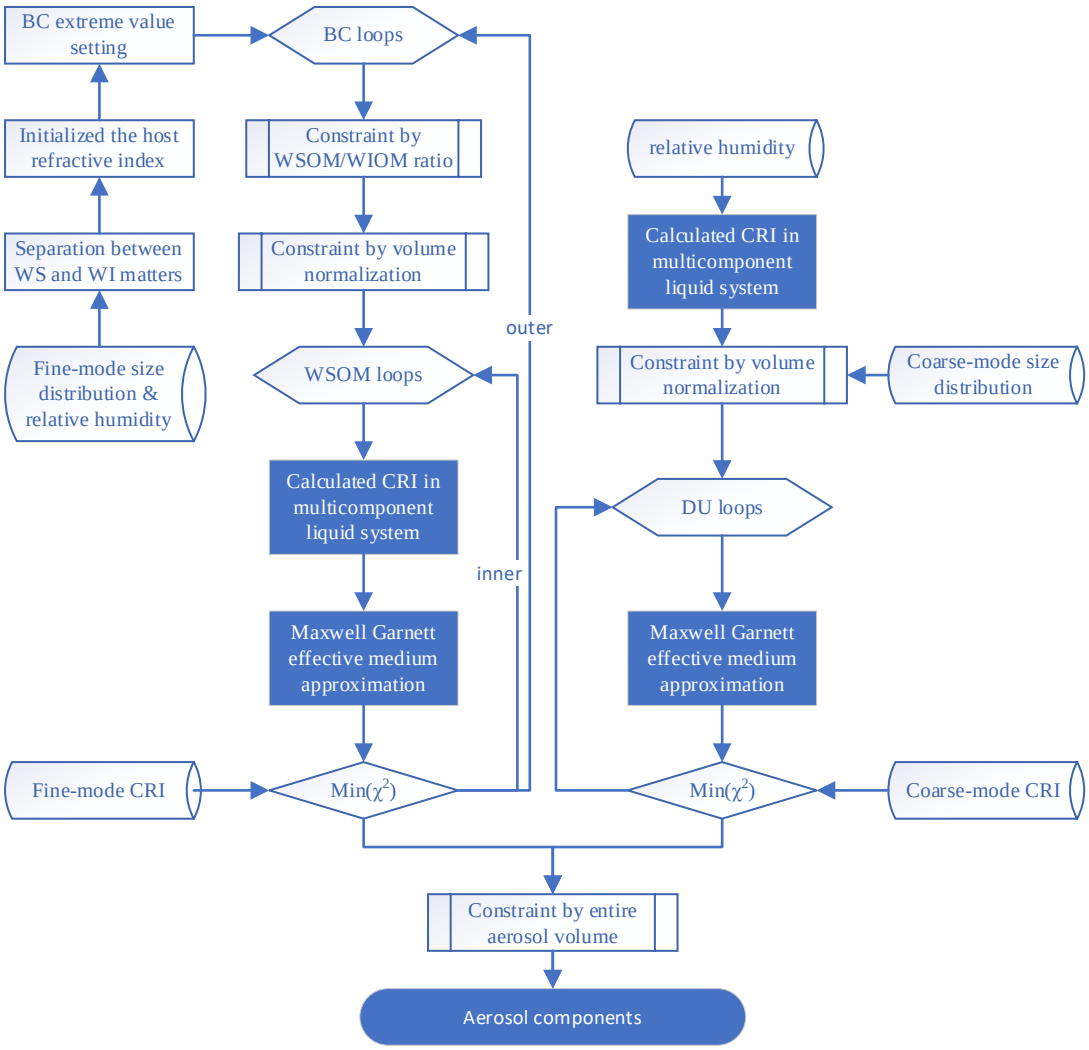

Figure 4. Flowchart of the aerosol component classification inversion algorithm .

(3) The "combined" aerosol types are used in table 3, and subsequent analysis (section 4) was associated with this combined component.

Error analysis: I would like to be convinced that method works, by showing a retrieval without any noise added, proving that it retrieves the exact pre-defined composition, i.e. calculating CRI using the forward model and retrieving it back again. Errors are analysed only from the point of view of data uncertainties, although some of them can emerge from the retrieval itself (i.e. inaccuracy of the forward model) and an obvious fitting bias in figure 6 can be an indication of that.

Response: Thank you for reviewer's comments. A set of retrievals without noise is added in the section 3.5 as follows:

Line 226-240: "The uncertainty in the retrieval results was evaluated using synthetic data, both without and with input errors added. For the first case (without input errors), a set of complex refractive indices has been obtained by calculating a set of volume fractions of the aerosol components using the forward chemical model, which was used as input for the retrieval of the aerosol components without any noisy added. For the aerosol components, the volume fraction of BC was constrained between 0.0 to 3.0% with an interval of 0.5%, and corresponding dynamic ranges for the other components with intervals of 10%, in three ambient relative humidity conditions (40%,

60% and 80%). Figure 5 shows the comparison of the aerosol component volume fractions from forward modeling used as input, and their retrieved values. The volume fractions of the retrieved aerosol components are in good agreement with the input values. For the fine mode fraction, most data pairs are located close around the 1:1 line, with the mean absolute error (MAE) of the aerosol component volume fractions of 3.0%. In five samples the difference in the AW_f is more than 20.0%, though the overall MAE for AW_f is only 5.5%. In these five samples, the BC component is low and organic matter contributes substantially to the aerosol light absorption, resulting in underestimation of the AW_f volume fraction at high RH and overestimation for moderate RH. WSOM is overall slightly overestimated and AN is underestimated by only a few percent. The correlation between the input and retrieved aerosol volume fractions in the coarse mode is even better than that in fine mode. The regression coefficient for all samples is 0.99, and the MAE is only 2.0%. These results show the very small uncertainty in the retrieved aerosol component volume fractions."

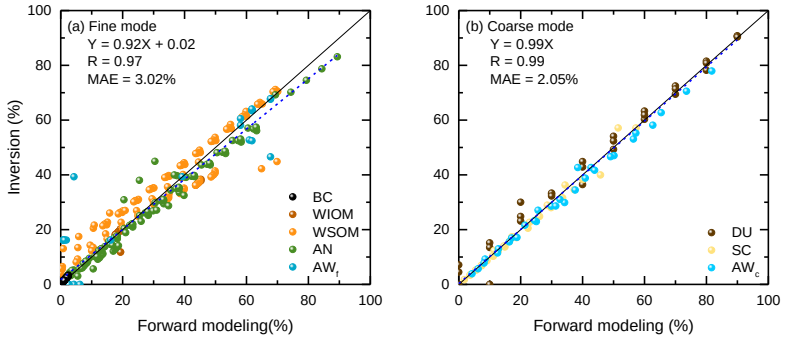

Figure 5. Scatter plots of volume fractions of aerosol components in the fine (left) and coarse (right) modes retrieved using the algorithm described in Chapter 3, versus those used as input calculated with the forward model. The solid line is the 1:1 line, and the dash line is the fitting line.

Data analysis: Again, very dry, some additional analysis (for e.g. splitting sites into several groups and analysing the seasonal averages for the groups) would be appreciated. For instance, there was a statement that higher BC in winter is because of heating in north region, splitting sites into heated/not heated regions and analysing seasonal trend would make this statement much more trustworthy and results obtained more significant. Also, a comparison to a previous method could be presented, if, of course, such comparison could be done. For e.g. comparison with OM from Zhang 2018 could be performed to illustrate improvements (if any), or discuss the similarities/differences observed.

Response: Thank you for the comment. We add the analysis about the seasonal variation of main aerosol compositions in fine mode between north and south China as follows:

Line 346-359: " The seasonal variation of the main aerosol components in the fine mode is discussed on a regional basis (Figure 10). BC concentrations in typical northern regions are higher than in southern regions, because of emissions due to winter heating only in the north. Other BC sources are vehicle emissions and biomass burning. Adverse meteorological conditions in winter result in the accumulation of BC in the atmosphere resulting in high BC values in both the north and the south. The highest BC mass concentrations in the northern region in the winter is 4.3 mg m^{-2}. OM is one of the dominant components in the fine mode, with sources similar to those of BC. The impact of biomass burning in the winter and spring over south China (Chen et al., 2017) is significant, leading to OM concentrations of more than 50.0 mg m^{-2}. In the northern region, much biomass burning occurs in the autumn (Wang et al., 2020). With the influence of heating, the OM level in the north can reach up to 80.1 mg m^{-2}. Therefore, the OM mass concentration in the northern region is only low in the summer (50.8 mg m^{-2}). AN is usually formed by secondary reactions of gaseous precursors in complex air

pollution areas. In both the northern and the southern region, AN mass concentration is larger in the summer than in other seasons, and the seasonal variation in the southern region is significantly smaller than that in the north. The mean AN mass concentration in the southern region is 8.7 mg m^{-2} higher than that in the northern region. This suggests that more AN is produced by secondary reactions in the humid climate in the south than in the northern region."

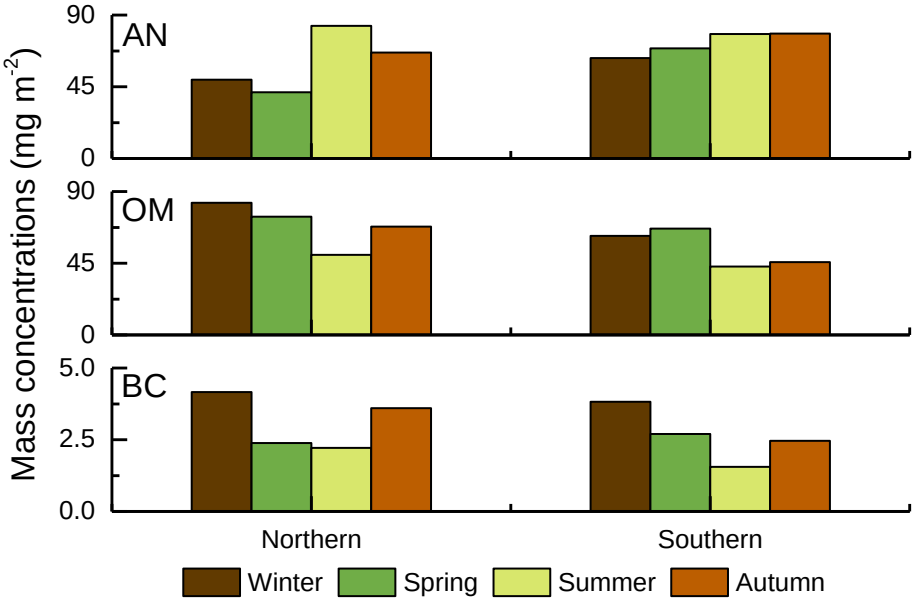

Figure 10. Comparison of aerosol component mass concentrations in northern (Xi'an, Beijing, Harbin, Hefei and Songshan) and southern China (Nanjing, Shanghai, Zhoushan, Guangzhou, Haikou and Sanya).

Compared with Zhang et al. (2018), the main advantage of the new algorithm is the increased flexibility in selecting water-soluble components. The new algorithm gives a reasonable scheme of complex refractive index (CRI) estimation, considering the mixing of multiphase solution. We added an explanation in the revised manuscript as follows:

Line 216-224: "The retrieval algorithm described here is an improvement over that described in Zhang et al. (2018). In that algorithm, the WSOM component was added to the host, but it could only be considered as a non-hygroscopic component. The proportion of solute and solution in the host mixture at different relative humidities should be measured in the laboratory, which limits the choice of aerosol components in the inversion process. Also, the real part of the CRI of the host was calculated by volume averaging, which can introduce a small error. The improved algorithm described here is more suitable for the calculation of the properties of a mixture of multiple water-soluble components as long as the hygroscopic parameter is known, which is not only convenient to measure but also independent of particle size. The hygroscopic parameter of WSOM can be varied according to the choice of mixing components instead of changing the algorithm itself. Similarly, some other water-soluble components (e.g. sulfate) can be introduced into the inversion algorithm without laboratory measurements."

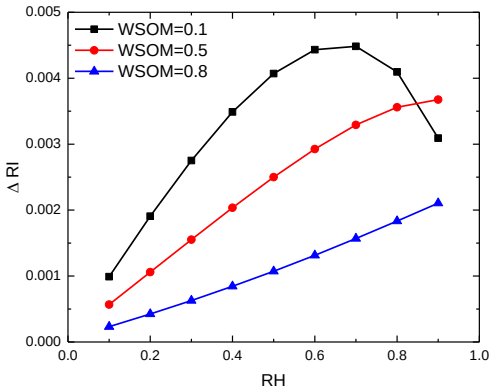

Figure R1. The differences of real part of CRI with relative humidity between the algorithm of Zhang et al (2018) and this study when the WSOM volume fraction is 10%, 50% and 80%.

There is a misfit error available as a result of the retrieval, why not use it to clean up the data a bit? This will provide more trustworthy results. Besides, as I understood estimations of the SONET retrieval errors for complex refractive index are available, why not get rid of the results whose fits are below the error bars? Authors themselves claim that ~40% of the retrieval are above the average error, i.e. these results have questionable quality and could significantly influence the statistics provided.

Response: Thanks very much for your comments. We set a threshold and filter the results and revised as follows:

Line 303-321: "The closure of the CRI between instantaneous optical-physical inversion and chemical estimation is examined by the data pair frequency. Figure 8 shows scatter density plots of the chemically estimated and sunphotometer-retrieved imaginary parts of the fine mode at 675 nm (k_f) and 440 nm ($k_{f,440}$) and the real parts of fine mode at 440 nm(n_f). The points are colored by the number of data pairs (Retrieved, Estimated), which are sorted according to ordered pairs in 0.0005 intervals for the imaginary parts of CRI and 0.001 intervals for the real parts. The data pairs of k_f are closely concentrated around the 1:1 line, although with slight underestimation with 94.3% of the estimated values lower than the retrieved values; only 5.3% of the data pairs have a relatively large absolute error (AE > 0.01). The mean bias is not large (-0.003), and the mean absolute value is equal to the mean absolute error (MAE = 0.003). There are two reasons for this slight underestimation in chemical estimation. On the one hand, the imaginary part of the refractive index of BC is much larger than for the other components due to its strong absorption. Thus, the inversion of the BC concentration is very sensitive to the estimation of the refractive index. As shown in table 3, although the TRE of BC is the lowest, the errors caused by k_f and $k_{f,440}$ are larger than for any other component. On the other hand, k_f is not only affected by BC in the inversion process, but also affected by organic components (WSOM & WIOM) with spectral absorption characteristics. Therefore, in most cases, k_f is underestimated in chemical estimation and $k_{f,440}$ is overestimated (Bias = 0.007). The mean relative error (RE) is 27.1%, and 62.8% of the data points are below the average relative error line. This indicates that most inversion results have good optical closure. For the closure of the real part of the fine mode, the data pairs of n_f are also concentrated around the identity line, although 76.5% of the n_f is above the identity line. Underestimation occurs mainly when n_f is larger than 1.56, because the only component with the real part of the CRI larger than 1.56 is BC, but its concentration is mainly determined by the imaginary part. The bias of the estimated n_f (Bias = 0.009) is larger than that of k_f due to the fact that the value and the range of n_f are larger than that of k_f. "

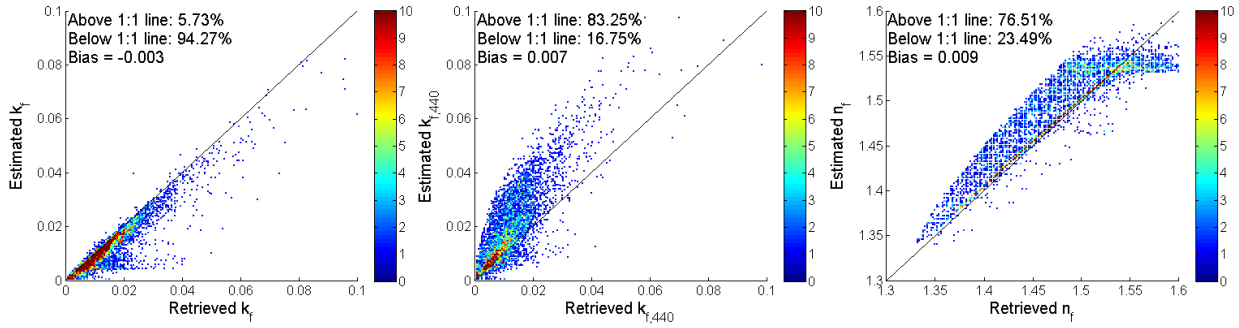

Figure 8. Data pair frequency of instantaneous imaginary parts of the complex refractive index at 675 nm (k_f), 440 nm ($k_{f,440}$), and real part at 440 nm (n_f) which are sorted according to ordered pairs (Retrieved, Estimated) in 0.0005 and 0.001 intervals for imaginary and real parts, respectively. "Retrieved" represents sub-component of CRI from the optical-physical retrievals, and "Estimated" is estimated by retrieved chemical components. The color represents the number of cases (color bar), and the solid black line shows the 1:1 line.

In addition, I think what the reviewer said "Authors themselves claim that ~40% of the retrieval are above the average error" should refer to "Moreover, the mean relative error (RE) is 29.93%, and 61% of the data points is below the average relative error line." in the manuscript. It is not a constraint on the convergence of the algorithm, but to evaluate whether the convergence condition is reasonable. Therefore, it does not affect our statistical results.

Minor comments:

Line 70: I would suggest replacing "much information" to "sufficient information"

Response: We have revised in the manuscript as follows:

Line 73-74: "These radiation and polarization measurements can provide sufficient information to calculate the columnar aerosol optical depth (AOD) and further retrieve the aerosol microphysical parameters. "

Lines 87-89: "Using the sub-modal characteristics data set thus obtained, an aerosol sub-modal model was established for China by Li et al. (2019), but the sub- modal aerosol components have not been given." I'm not sure what authors mean by this sentence, please, consider rephrasing.

Response: We rephrase this sentence as follows:

Line 91-93: "Using these fine and coarse mode characteristics of the CRI, micro-physical properties of aerosols in each mode were analyzed (Li et al., 2019), but the aerosol chemical components were not determined."

Line 94: "… the linear interpolation method is used to match the SONET observations …", please clarify if interpolation in time or alt/lat/lon or all together were used.

Response: This linear interpolation is only used in timescale. We clarify this point in manuscript:

Line 99: " The CMA stations closest to each SONET site were selected and the meteorological data were collocated in time with the SONET observations by linear interpolation between the nearest observations."

Line 210. It is claimed that error of 50% in AWf is "acceptable". No references or desired thresholds were provided to jump to such conclusion. Consider providing them, or rephrasing the sentence to a milder comparison.

Response: The revised sentence is as follows:

Line 268: "Affected by SC, the TRE of AW_c is also large due to n_c, but the TRE of AW_f is much smaller (50.05%). "

Line 308: "… improved a component inversion algorithm …", please refer to a baseline method. And what improvements authors refer to? No comparison to previous methods shown, is it complexity of the composition? Please clarify.

Response: This study improves the algorithm of Zhang et al. (2018), not only for multiphase solution calculation, but also for adding SC component. We revised this sentence as follows:

Line 389-390: "In the current study, we updated the refractive index calculation in a multi-component liquid system and improved the component inversion algorithm of Zhang et al. (2018) to retrieve atmospheric columnar aerosol components including …"

Line 353: I presume that another Dubovik 2000 paper, "Dubovik, O., and M. D. King, 2000: A flexible inversion algorithm for retrieval of aerosol optical properties from sun and sky radiance measurements. J. Geophys. Res., 105, 20 673¨C20 696." will be more suitable in the context it is referenced in the article.

Response: Thank you for your comment. We add this reference in manuscript.

Figure 5. I would appreciate to have numbers on pie charts too like figure 4, or an additional table with percentage of the components for each site.

Response: Because of the small space in the figure, we only increased the volume ratio of BC in the blank space. The fractions of other components are listed in the table S3 of supplementary as follows:

Table S3. The averaged mass fraction of aerosol components at SONET sites shown in figure 7.

Site	Fine mode					Coarse mode			AW	IS	OM
	BC	WIOM	WSOM	AN	AW_f	DU	SC	AW_c			
Lhasa	0.43%	7.85%	6.47%	1.89%	2.01%	81.34%	0.00%	0.00%	2.01%	1.89%	14.33%
Zhangye	0.05%	2.43%	2.36%	2.31%	1.79%	70.23%	11.53%	9.30%	11.09%	13.84%	4.79%
Kashgar	0.08%	4.57%	2.63%	2.51%	1.19%	65.63%	17.09%	6.30%	7.49%	19.60%	7.20%
Minqin	0.11%	2.97%	2.33%	6.13%	2.33%	65.70%	12.20%	8.23%	10.56%	18.33%	5.30%
Xi'an	0.60%	4.74%	9.62%	10.08%	6.56%	60.75%	3.37%	4.28%	10.84%	13.45%	14.37%
Beijing	0.69%	5.91%	7.67%	10.65%	6.08%	62.86%	2.85%	3.29%	9.38%	13.49%	13.59%
Nanjing	0.96%	6.26%	15.13%	13.79%	9.68%	46.94%	3.14%	4.10%	13.78%	16.93%	21.39%
Shanghai	1.30%	5.14%	10.78%	17.68%	10.33%	46.62%	3.87%	4.28%	14.61%	21.55%	15.92%
Harbin	0.97%	5.43%	12.45%	15.36%	10.80%	50.50%	2.21%	2.26%	13.07%	17.57%	17.88%
Hefei	0.80%	3.33%	11.51%	14.86%	10.34%	49.79%	3.91%	5.46%	15.80%	18.77%	14.84%
Songshan	0.59%	8.14%	6.31%	12.09%	6.63%	59.65%	3.10%	3.50%	10.13%	15.18%	14.45%
Chengdu	0.58%	2.72%	11.90%	22.54%	9.34%	42.10%	3.65%	7.17%	16.51%	26.19%	14.62%
Zhoushan	0.33%	3.55%	6.84%	17.06%	14.86%	42.82%	5.83%	8.70%	23.56%	22.89%	10.40%
Guangzhou	0.64%	2.84%	8.18%	23.80%	18.26%	31.54%	4.28%	10.46%	28.72%	28.08%	11.02%
Haikou	0.83%	2.32%	9.89%	22.03%	14.42%	33.29%	6.90%	10.32%	24.74%	28.93%	12.21%
Sanya	0.55%	0.32%	14.27%	24.78%	9.46%	37.41%	3.21%	10.00%	19.45%	27.99%	14.59%

Figure 6. Why only imaginary for fine and only 670? Fitting statistics is a necessary (but not sufficient) metric to justify that the method works. Please consider showing at least a real part at 670 too, but I would be more convinced to see all of them, all wl and fine/coarse. At least show minimal set, following the structure of Table s2 fine/coarse real/imaginary and 440 too, since they all referenced to have different level of uncertainty. Also, please, mention that it is fine mode in figure caption.

Response: We add the closure figure for the real and imaginary parts of the CRI at 440 nm. Because the spectral

changes in the real part are ignored in the CRI inversion, the optical closure at other wavelengths have a similar pattern (We don't show that in the manuscript). The Figure is as follows and its description has been mentioned in the previous response.

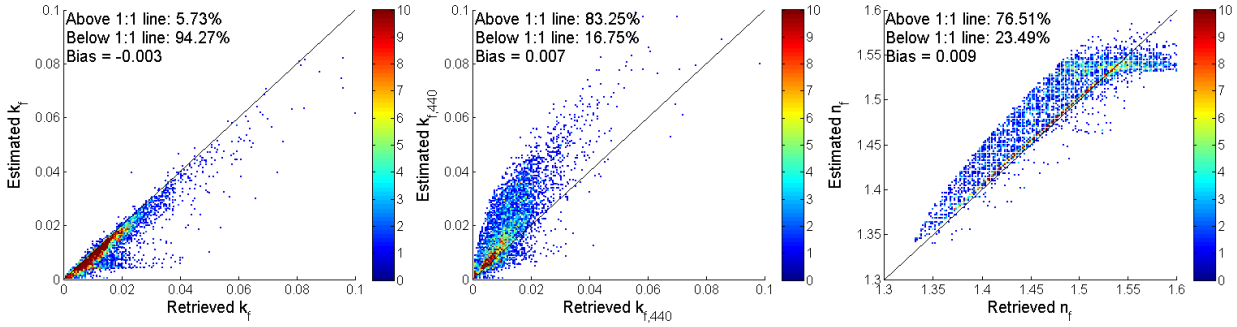

Figure 8. Data pair frequency of instantaneous imaginary parts of the complex refractive index at 675 nm (k_f), 440 nm ($k_{f,440}$), and real part at 440 nm (n_f) which are sorted according to ordered pairs (Retrieved, Estimated) in 0.0005 and 0.001 intervals for imaginary and real parts, respectively. "Retrieved" represents sub-component of CRI from the optical-physical retrievals, and "Estimated" is estimated by retrieved chemical components. The color represents the number of cases (color bar), and the solid black line shows the 1:1 line.

For the coarse mode, only the real part of CRI (n_c) is used in the inversion because the imaginary part of the CRI for both AW_c and SC are zero. Figure R2 shows the comparison of n_c between retrieved and estimated from chemical components. The residual is tiny when the n_c is less than 1.534. To the left of the 1:1 line, when n_c equals 1.534 (real part of DU), there is a set of points with a large error. That is because the relative humidity is lower than 40% leading to the AW_c is close to 0.0. Hence, the real part of the coarse mode gets the minimum of 1.534. In contrast, the real part of the CRI for the DU component is increased up to the mean value of the retrieved refractive index at all available wavelengths which is higher than the mean value of DU's n_c. The reason for doing that is that DU is a mixture. Even so, when relative humidity is high, some underestimates still occur. Fortunately, 73% of the points are not influenced by the above factors, with the mean error of 0.01. The uncertainty of the retrieved CRI also causes the large residuals.

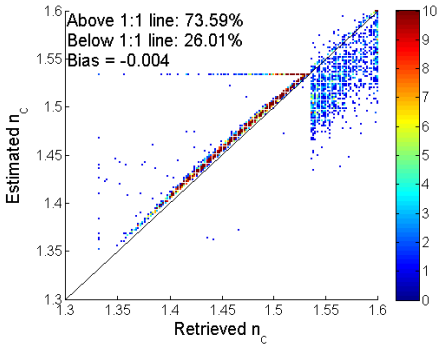

Figure R2. Data pair frequency of instantaneous imaginary parts of the complex refractive index at 440 nm (n_c) which are sorted according to ordered pairs (Retrieved, Estimated) in 0.001 intervals. The color represents the number of cases (color bar), and the solid black line shows the 1:1 line.

Figure 7. Is it possible to re-plot or at least re-paint so the color scheme will be the same as in figs 5 and 6?
Response: We re-plot this figure to match the color scheme of the other figures.

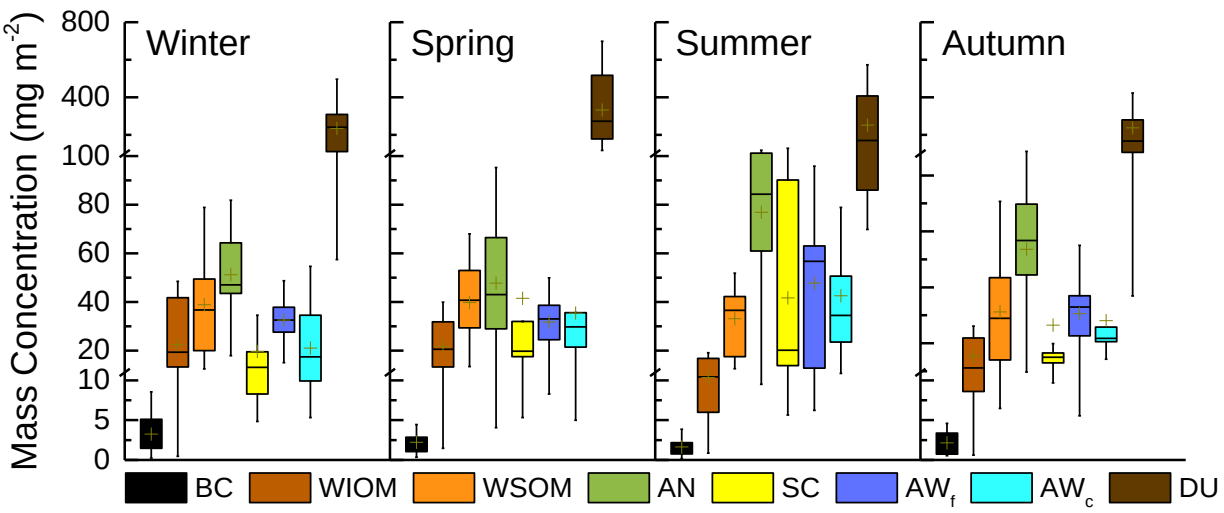

Figure 9. The mass concentrations of aerosol components in four seasons (winter, spring, summer and autumn). For the box-whisker plot, the mean value is indicated by a plus sign (+), and the median value by a short line inside the box (–). The top and bottom edges of each box represent the top and bottom quartiles (Q3 and Q1), and the corresponding whiskers are the outliers (Q3+1.5IQR and Q1-1.5IQR, IQR is interquartile range).

Table 2. Please specify what "No" in the head of the table means. If this is a number of observations, then why for quite different observation periods the numbers are so close? Was observation data filtered? Please, clarify.

Response: "No" is the meteorological station number. We explain it at the bottom of table 2.

Table s2 and s1. They are not a part of a publication provided (.pdf) and I don't see much reason why. I mean they are rather small and multiply referenced from the text, what is the reason to have one external half-page containing them?

Response: These two tables are referenced from other references to provide model input errors. For clarity, we added a description of them to the supplementary (S1).

Technical comments:

Figure 3. Size distribution axis. I believe it is μm3 *μm-2

Response: Thank you for pointing this out. We revised figure as follows:

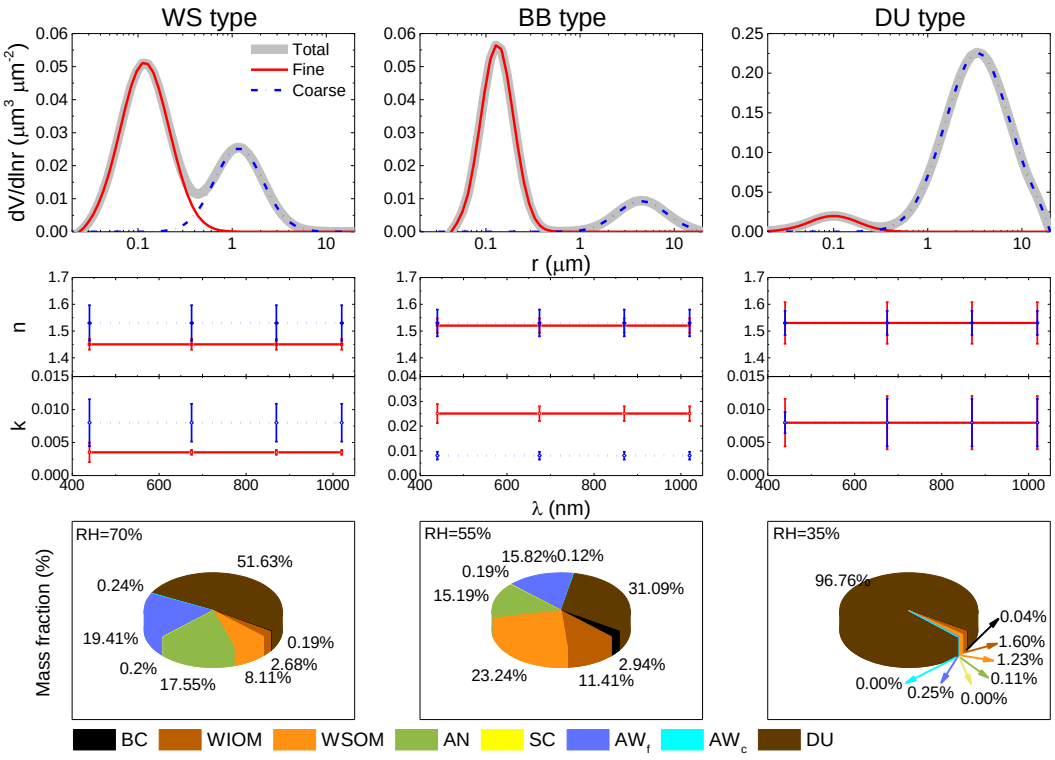

Figure 6. The fine and coarse-mode volume size distribution, complex refractive index and aerosol components of three typical aerosol models (WS: water soluble, BB: biomass burning, DU: dust).

Line 27-28: "… retrieval of the chemical composition …"

Response: Following the comment of another reviewer, we have deleted this sentence.

Line 49: "… simultaneously retrieved …"

Response: We revised this sentence as follows:

Line 52: "Zhang et al. (2018) simultaneously retrieved the WSOM and WIOM components but ignored the error in the refractive index introduced by the aerosol volume averaging method applied to the multicomponent liquid system."

Line 56: "… to solve for the refractive index in a multicomponent liquid system"

Response: We revised this sentence as follows:

Line 59: "hygroscopicity is introduced to solve for the refractive index in a multicomponent liquid system."

Line 57: "… in the algorithm …"

Response: We revised this sentence as follows:

Line 60: "The results are used in the algorithm to retrieve aerosol components … "

Line 70: "… radiation to determine the columnar water vapor"

Response: We revised this sentence as follows:

Line 73: "All bands provide both radiation and polarization measurements, except the 936 nm band which only measures radiation to determine the columnar water vapor."

Line 93-94: Probably author meant “To avoid observation uncertainties, only data from manned weather stations that are maintained regularly were used.”

Response: We modified this sentence into the form suggested by the reviewer.

Line 185: “… by the different aerosol size distribution …”

Response: We revised this sentence as follows:

Line 243: “Each type is described by the different aerosol size distribution and refractive index parameters derived from Zhang et al. (2017).”

Line 246: “On the one hand”

Response: We revised this sentence as follows:

Line 310: “On the one hand, the imaginary part of the refractive index of BC is significantly higher than for the other components due to its strong absorption.”

Line 247: “… than for the other components …”

Response: We revised this sentence as follows:

Line 310: “On the one hand, the imaginary part of the refractive index of BC is significantly higher than for the other components due to its strong absorption.”

Line 252: “… data points are below …”

Response: We revised this sentence as follows:

Line 316: “The mean relative error (RE) is 27.1%, and 62.8% of the data points are below the average relative error line.”

Line 260: “… high in spring …”

Line 260: “… to the dust transport from the northwest China …”

Response: We revised this sentence as follows:

Line 342-343: “Low concentrations of BC in the other seasons are mainly due to the influence of frequent dust events in the spring and high aerosol hygroscopic growth in the summer.”

Line 261: “… seasons indicates that the aerosol at some sites”

Response: We revised this sentence as follows:

Line 337: “The difference between the upper and lower quartiles of AW_f in the summer is larger than in other seasons indicating that in the summer the aerosol at some sites has a low hygroscopicity.”

Line 262: “… due to the observational errors …”

Response: We revised this sentence as follows:

Line 373: “The unusually high values at Shanghai in 2016 may be due to the observational errors.”

Line 304: “… changes in the meteorological conditions …”

Response: We revised this sentence as follows:

Line 385: ” Coarse mode aerosols usually derive from natural sources, and their variations can be associated with changes in the meteorological conditions.”

Improved inversion of aerosol components in the atmospheric column from remote sensing data

Ying Zhang[1], Zhengqiang Li[1], Yu Chen[2], Gerrit de Leeuw[1,3,4], Chi Zhang[1], Yisong Xie[1], Kaitao Li[1]

[1]State Environmental Protection Key Laboratory of Satellite Remote Sensing, Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing 100101, China

[2]Public Meteorological Service Center, China Meteorological Administration, Beijing 100081, China

[3]Finnish Meteorological Institute, Climate Research, Helsinki FI-00101, Finland

[4]Royal Netherlands Meteorological Institute (KNMI), R&D Satellite Observations, 3730AE De Bilt, The Netherlands

Correspondence to: Zhengqiang Li (lizq@radi.ac.cn)

Abstract. A method to calculate the complex refractive index of a multicomponent liquid system is introduced into the retrieval of atmospheric columnar aerosol components from optical remote sensing data. The aerosol components comprise five species, combining eight sub-components including black carbon (BC), water-soluble (WSOM) and water-insoluble organic matter (WIOM), ammonium nitrate (AN), sodium chloride (SC), dust-like (DU), and aerosol water in the fine and coarse modes (AW_f and AW_c). The uncertainty in the retrieval results is analyzed based on the simulation of typical models, showing that the complex refractive index obtained from instantaneous optical-physical inversion compares well with that obtained from chemical estimation. The algorithm was used to retrieve the columnar aerosol components over China using the ground-based remote sensing measurements from the Sun-sky radiometer Observation NETwork (SONET) in the period from 2010 to 2016. The results were used to analyze the regional distribution and interannual variation. The analysis shows that the atmospheric columnar DU component is dominant in the northern region of China, whereas the AW is higher in the southern coastal region. The AN significantly decreased from 2011 to 2016, by 21.9 mg m^{-2}, which is inseparable from China's environmental control policies.

1. Introduction

Atmospheric aerosol consists of a suspension of solid and/or liquid particles in the air. The aerosol particles scatter and absorb solar radiation. The chemical composition and mixing state of the aerosol particles affect their optical characteristics, which in turn influence the energy budget of the Earth-atmosphere system and thus climate (Boucher et al., 2013).

To measure aerosol composition, many methods are used including online analysis in the field, sample analysis in the laboratory, remote sensing estimation, etc. The detail of information depends on the technique used. Because of fast observation and low cost, the application of remote sensing techniques to estimate aerosol composition has developed rapidly since 2000. Satheesh et al. (1999; 2002a, b; 2005) established an algorithm for the inversion of aerosol components from remote sensing data based on the hypothesis of external mixing and assuming fixed size distributions for each component. But an external

mixture usually cannot accurately describe the natural state of aerosols. Even if the particles are individually pure when first produced, numerous processes in the atmosphere will convert an external mixture to an internal mixture (Lesins et al., 2002). Therefore, internal mixing hypotheses are widely used and multiple approaches have been developed (e.g., Zhang et al., 2018; Schuster et al., 2016, 2009, 2005; van Beelen et al. 2014; Li et al., 2013; Wang et al., 2013; Arola et al., 2011). Schuster et al. (2005) determined the volume fraction of black carbon in an internal mixture with water and a soluble component by fitting the calculated complex refractive index to retrieved AERONET values at all four available wavelengths. In a follow-up study, Schuster et al. (2009) applied a similar procedure to determine the aerosol water fraction by fitting the real part of the refractive index of an internal mixture of water, soluble and insoluble species to observations by minimizing the cost function at all four wavelengths together. In this work the ratio of the dry volume fraction of insoluble to that of soluble aerosols was constrained by using a climatological value and the real refractive index which also prescribes the aerosol hygroscopicity. This constraint also provides a maximum insoluble fraction and the fraction of dust aerosol. Brown carbon was further estimated by Arola et al. (2011) due to the large change of its absorbing characteristic with wavelength for wavelengths smaller than 550 nm, but the dust component was ignored in this study. Aerosol bimodal characteristics were used by Schuster et al. (2016) to estimate the aerosol absorbing components including BC, brown carbon and hematite in the fine and coarse modes. This method was also embedded in the GRASP (Generalized Retrieval of Aerosol and Surface Properties; Dubovik et al., 2011)) system by Li et al. (2019) for application to POLDER/PARASOL observations. The above algorithms are aimed at retrieving absorbing aerosol components, such as BC, brown carbon and iron oxides, but have only simple treatment for scattering components, especially the host of multicomponent liquids.

van Beelen et al. (2014) introduced water-soluble organic matter (WSOM) in the inversion process based on the hygroscopicity of the OM mixture, but in this study water-insoluble organic matter (WIOM) was not accounted for. Some studies separated the OM based only on the spectral changes (Xie et al., 2017; Choi and Ghim, 2016) leading to large uncertainty in the results. Zhang et al. (2018) simultaneously retrieved the WSOM and WIOM components but ignored the error in the refractive index introduced by the aerosol volume averaging method applied to the multicomponent liquid system. For other non-absorbing components, the water content and inorganic components in the fine mode are identified by the difference in hygroscopic growth between organic and inorganic matter (Zhang et al., 2018; van Beelen et al.; 2014). In the coarse mode, sea salt is identified by the aerosol sphericity in the study of Xie et al. (2017) but this parameter is difficult to observe.

Although the retrieval of aerosol components by using remote sensing methods has been greatly developed, the application of hygroscopicity to identify the weak and non-absorbing components in a multicomponent liquid system remains difficult. In this study, hygroscopicity is introduced to solve for the refractive index in a multicomponent liquid system. The results are used in the algorithm to retrieve aerosol components from data obtained from the ground-based remote sensing network SONET (Li et al, 2018). The data and method are described in sections 2 and 3, respectively. The results for the aerosol

components and derived hygroscopic parameter and effective density are presented and analyzed in section 4, and we conclude this study in section 5.

2 Measurements

2.1 Sun-Sky radiometer

The multiwavelength polarized sun-sky radiometer CE318-DP manufactured by Cimel Electronique in France, as an accurate instrument designed for long-term continuous observations in the field, can automatically measure solar and sky radiation. It consists of an optical head, a control box and a bi-axial stepping motor system. The optical head has two views: one for direct solar radiation with no focusing lens and the other for sky radiation with focusing lens. The internal optical system consists of a spectral and a polarizing filter to measure radiation in different wavebands with polarization directions. The 9 wavebands vary from the visible to the near-infrared (340, 380, 440, 500, 675, 870, 936, 1020, 1640 nm) with a full width at half maximum of 10 nm. All bands provide both radiation and polarization measurements, except the 936 nm band which only measures radiation to determine the columnar water vapor. These radiation and polarization measurements can provide sufficient information to calculate the columnar aerosol optical depth (AOD) and further retrieve the aerosol microphysical parameters.

2.2 SONET

The Sun-sky radiometer Observation NETwork (SONET) is a local observation network in China for ground-based remote sensing measurements of aerosol properties (Li et al., 2018). At present, there are 16 long-term observation sites in China, which are evenly distributed over north and south China, northwest China and the Tibetan Plateau (Figure 1). The longest time series is provided by the Beijing station, which was established in 2009. Five more stations joined in 2011 and 2012 and the network has been gradually growing to the current size. The geographical and topographical features of the long-term sites are diverse such as plateau, desert, hilly, plain and island, including three megacities, three islands and one plateau site (Table 1). SONET data provide sufficient variability, as regards length of time series, spatial coverage, climatic and topographic features and aerosol properties, for the analysis of atmospheric aerosol characteristics across China.

SONET provides continuous observations of direct sun and sky radiation measured using the multi-wavelength polarization sun-sky radiometer (CE318-DP), following the AERONET protocol (Li et al., 2018). Based on the inversion algorithm of Dubovik and King (2000) and Dubovik et al. (2000), the 440, 675, 870 and 1020 nm wavebands are used to retrieve more than 20 parameters describing the optical, physical and chemical global properties as column-integrated properties (Li et al., 2018), including particle volume size distribution (VSD), complex refractive index (CRI) and aerosol components. Using these data, VSD and CRI sub-modal parameters of atmospheric aerosols are obtained using the modal decomposition method proposed

by Zhang et al. (2017). The real parts of the CRI of the fine and coarse modes (n_f and n_c, respectively) are spectrally independent, while the imaginary parts have spectral variation at 440 nm, so they are written as ($k_{f,440}$, k_f) & ($k_{c,440}$, k_c). Using these fine and coarse mode characteristics of the CRI, micro-physical properties of aerosols in each mode were analyzed (Li et al., 2019), but the aerosol chemical components were not determined.

2.3 Meteorological data

Meteorological data provide important supplementary information for the analysis and interpretation of the SONET-retrieved aerosol information. Hourly observations from surface meteorological stations were provided by the China Meteorological Administration (CMA). Only data from manned weather stations, which are maintained regularly, were used to ensure the best possible data quality. The CMA stations closest to each SONET site were selected and the meteorological data were collocated in time with the SONET observations by linear interpolation between the nearest observations. Figure 2 shows the statistics of the relative humidity (RH) observations at each of the 16 sites. The highest mean RH occurs at the Sanya site, and the lowest value at the Lhasa site. Generally, the mean RH is relatively low at stations at northern latitudes, and often also at high altitudes. The standard deviations of wet (e.g. Sanya, Haikou) and dry sites (e.g. Lhasa, Kashgar) are smaller than at other sites.

3. Methodology

The aerosol components are determined by comparison of the aerosol microphysical properties calculated using a forward model with those retrieved from the SONET observations (Zhang et al., 2017). This is achieved by minimizing the iterative kernel function, i.e. the sum of the differences between the calculated and observed properties at each of the four wavelengths together, to find the optimum solution. The forward model includes three modules: the Maxwell Garnett effective medium approximation (Schuster et al., 2005) module to calculate aerosol internal mixing characteristics, an aerosol hygroscopic growth module to solve the hygroscopicity of water-soluble components in a multicomponent liquid system, and an organic component dynamic constraint module to keep a reasonable ratio of organic matter.

3.1 The aerosol component classification

The aerosol component classification includes five principal species (black carbon (BC), organic matter (OM), inorganic salt (IS), aerosol water (AW), dust-like (DU)). Three of these components are further sub-divided, i.e. organic matter is sub-divided into water-soluble (WSOM) and water-insoluble organic matter (WIOM), inorganic salt consists of ammonium nitrate (AN) in the fine mode and sodium chloride (SC) in the coarse mode, and aerosol water is the water content in the fine and in the coarse mode. Thus there are eight sub-components as illustrated in figure 3. All of these eight aerosol components constitute a relatively complete system comparable to those used in chemical transport models.

The aerosol components are identified following three steps. The first step is the separation of the aerosol micro-physical properties (VSD and CRI) into those for the fine and coarse modes as summarized in supplementary S1. For the fine mode fraction, the water-insoluble and water-soluble components are identified using an empirical function (see section 2.2.2 in Zhang et al., 2018), which describes the ratio of the water-soluble to the water-insoluble volume fractions determined by RH, together with the parameterization of aerosol soluble volume fractions by Kandler and Schutz (2007). Then the subcomponents are separated into inclusion (BC and WIOM) and their environment (AN, AW_f and WSOM) using their hygroscopic and optical absorption properties. It should be noted that the water-soluble property of aerosol components is not equivalent to hygroscopicity. Dicarboxylic acids represented by oxalic acid are dominant in the WSOM component but their hygroscopicity is extremely low (Ma et al., 2013; Drozd et al., 2014; Jing et al., 2016). Also other organic compounds in aerosols are less hygroscopic as shown in Zhang et al. (2018) (their figure 1). Hence, the OM components (WSOM and WIOM) are treated as non-hygroscopic components. For the coarse mode fraction, the refractive index of the mixture (AW_c and SC) is determined by their hygroscopic growth factor. Dust and hydrate in the aerosol mixture are separated by the effective medium approximation.

In these processes, the hygroscopic growth is determined by the hygroscopicity parameter κ and effective densities of the aerosol subcomponents, and the aerosol mixture refractive index is calculated by that of the subcomponents and the mixing state. Key parameters of the forward model and references are listed in table 2. We notice that the effective densities for OC and DU reported from different studies cover a wide range (Ganguly et al., 2009; McConnell et al., 2008; Wagner et al., 2012; Bond and Bergstrom, 2006) because they depend on the mixing ratios. In the current study the effective density of aerosol components is used from a widely cited study by Zhang et al. (1993).

3.2 Complex refractive index in a multicomponent liquid system

The multiple water-soluble aerosol components together with the aerosol water make up a liquid system, with increased complexity of the calculation of hygroscopic growth and complex refractive index. The κ-Köhler theory proposed by Petters and Kreidenweis (2007) can cope with the hygroscopicity of the multicomponent liquid system. In this theory, the water activity of aqueous atmospheric particulate matter can be represented by the functional form

$$\frac{1}{a_w} = 1 + \kappa \frac{V_s}{V_w} \tag{1}$$

where V_s is the volume of the dry particulate matter and V_w is the volume of the water. The activity of water in solution (a_w) is close to the relative humidity (RH) due to lower curvature effect and can therefore be replaced with RH (Tang, 1996). The hygroscopicity parameter κ is defined through its effect on the water activity of the solution. In equation (1), the ratio of V_s to V_w can be further applied to the calculation of the volume fraction

$$\sum_i f_i = \frac{V_s}{V_s + V_w} = \frac{1 - a_w}{1 - (1 - \kappa)a_w} \quad (2)$$

where f_i is the volume fraction of the ith component

$$f_i = \frac{V_i}{V_s + V_w} \quad (3)$$

where V_i is the volume of the ith component.

In the multicomponent liquid system, the hygroscopicity parameter κ is given by the simple mixing rule

$$\kappa = \sum_i f_{dry,i}\kappa_i \quad (4)$$

where κ_i is the hygroscopicity parameter of the ith component obtained from the literature (table 2), and $f_{dry,\,i}$ is the dry component volume fraction defined as

$$f_{dry,i} = \frac{V_i}{V_s} \quad (5)$$

Using equation (2) for the relationship between the volume fraction and the hygroscopicity parameter, the complex refractive index of the multi-component aerosol system can be derived using the Lorentz-Lorenz relation (Heller, 1965). Firstly, the molar refractivity (A_e) at wavelength λ can be calculated from the real part of the complex refractive index (n_i) and the volume fraction of the individual components

$$A_e(\lambda) = \sum_i f_i A_i(\lambda) \quad (6)$$

Where A_i is the molar refractivity of the ith component represented by

$$A_i(\lambda) = \frac{n_i^2(\lambda) - 1}{n_i^2(\lambda) + 2} \quad (7)$$

Then, the real and imaginary parts of the complex refractive index at wavelength λ of the multi-component liquid system, $n_e(\lambda)$ and $k_e(\lambda)$, are obtained respectively by using the molar refractivity and the imaginary part of the complex refractive index of the ith component (k_i).

$$n_e(\lambda) = \sqrt{\frac{1 + 2A_e(\lambda)}{1 - A_e(\lambda)}} \quad (8)$$

$$k_e(\lambda) = \sum_i f_i k_i(\lambda)$$

(9)

Equations (8) and (9) apply to the estimation of the complex refractive index of a multi-component liquid system with hygroscopic growth.

3.3 Effective medium approximation

To determine the complex refractive index of a particle, i.e. including both the multi-component liquid system and water-insoluble matter, the complex refractive index ($m = n - ik$) at wavelength λ is expressed in terms of the permittivity, ε (λ):

$$m(\lambda) = \sqrt{\frac{|\varepsilon(\lambda)| + Re(\varepsilon(\lambda))}{2}} + i\sqrt{\frac{|\varepsilon(\lambda)| - Re(\varepsilon(\lambda))}{2}}$$

(10)

The permittivity of the multi-component liquid system can then be calculated using equations (8) - (10). Considering the water-insoluble matter in a particle as inclusion and the water-soluble matter as the environment, the permittivity of the entire aerosol particle can be obtained by the Maxwell Garnett effective medium approximation (Schuster et al., 2005).

$$\varepsilon_{eff}(\lambda) = \varepsilon_e + 3\varepsilon_e\left[\frac{\sum_j \frac{\varepsilon_j(\lambda) - \varepsilon_e(\lambda)}{\varepsilon_j(\lambda) + 2\varepsilon_e(\lambda)} f_j}{1 - \sum_j \frac{\varepsilon_j(\lambda) - \varepsilon_e(\lambda)}{\varepsilon_j(\lambda) + 2\varepsilon_e(\lambda)} f_j}\right]$$

(11)

where, j is the number of water insoluble components and. $\varepsilon_j(\lambda)$ and $\varepsilon_e(\lambda)$ are the permittivities of the inclusion and its environment. The complex refractive index of the entire aerosol is estimated by aerosol component fraction using equation (10).

3.4 Inversion procedure

The flow chart for the inversion of the aerosol components is shown in figure 4. In the fine mode, the ratio of WS and WI matter is estimated using RH as described in section 2.2.2 in Zhang et al. (2018). The initial value of the host refractive index and the extreme value for the BC component are set by the calculation modules of the complex refractive index in the multicomponent liquid system (see section 3.2) and the effective medium approximation (see section 3.3), respectively. In the loop to determine the BC component, two constraints are applied to separate BC from other components. The WSOM/WIOM ratio constraint was developed by Zhang et al. (2018) based on considerations published in the literature (Chalbot et al., 2016; Bougiatioti et al., 2013; Wozniak et al., 2013; Mayol-Bracero et al., 2002; Krivácsy et al., 2001; Zappoli et al., 1999):

$$\begin{cases} f_{WSOM} \cong \alpha f_{WIOM} \\ \alpha = \dfrac{\beta\rho_{WSOM}^{-1}}{1 - \beta\rho_{WSOM}^{-1}} \end{cases} \qquad \beta \in [44\%, 77\%]$$

(12)

For more detail, see section 2.3.1 in Zhang et al. (2018). The volume normalization of the aerosol components in both the fine and coarse modes is used to constrain the volume fraction of the aerosol components to a reasonable range (similar as section 2.3.2 in Zhang et al., 2018)

$$\begin{cases} f_{fine} + f_{coarse} = 1.0 \\ f_{fine} = f_{BC} + f_{AN} + f_{WSOM} + f_{WIOM} + f_{AW_f} \\ f_{coarse} = f_{DU} + f_{SC} + f_{AW_c} \end{cases}$$

(13)

Then the inner loop of WSOM computes the CRIs of the fine mode at different BCs, and output the aerosol components of minimum χ^2. The inversion procedure for the coarse mode is simpler than that for the fine mode. There is only a loop for DU and the complex refractive index of the host can be directly calculated by equations (2) - (8) with only input of RH. The function Chi-squared (χ^2) as an iterative kernel function is expressed in the sum of the differences between the complex refractive index estimated from the forward model (m) and the retrievals (m_{rtrl}), at multiple wavelengths:

$$\chi^2 = \sum_{\lambda} \frac{\left(m_{rtrl}(\lambda) - m(\lambda)\right)^2}{m_{rtrl}(\lambda)}$$

λ=440, 675, 870 and 1020 nm (14)

The retrieval is completed when the value of χ^2 reaches a minimum. The volume fractions of the aerosol components can be obtained by solving the above equations (10-12). The aerosol mass concentration in the atmospheric column is calculated using the volume and effective density of the aerosol components.

The retrieval algorithm described here is an improvement over that described in Zhang et al. (2018). In that algorithm, the WSOM component was added to the host, but it could only be considered as a non-hygroscopic component. The proportion of solute and solution in the host mixture at different relative humidities should be measured in the laboratory, which limits the choice of aerosol components in the inversion process. Also, the real part of the CRI of the host was calculated by volume averaging, which can introduce a small error. The improved algorithm described here is more suitable for the calculation of the properties of a mixture of multiple water-soluble components as long as the hygroscopic parameter is known, which is not only convenient to measure but also independent of particle size. The hygroscopic parameter of WSOM can be varied according to the choice of mixing components instead of changing the algorithm itself. Similarly, some other water-soluble components (e.g. sulfate) can be introduced into the inversion algorithm without laboratory measurements.

3.5 Uncertainty analysis

The uncertainty in the retrieval results was evaluated using synthetic data, both without and with input errors added. For the first case (without input errors), a set of complex refractive indices has been obtained by calculating a set of volume fractions

of the aerosol components using the forward chemical model, which was used as input for the retrieval of the aerosol components without any noisy added. For the aerosol components, the volume fraction of BC was constrained between 0.0 to 3.0% with an interval of 0.5%, and corresponding dynamic ranges for the other components with intervals of 10%, in three ambient relative humidity conditions (40%, 60% and 80%). Figure 5 shows the comparison of the aerosol component volume fractions from forward modeling used as input, and their retrieved values. The volume fractions of the retrieved aerosol components are in good agreement with the input values. For the fine mode fraction, most data pairs are located close around the 1:1 line, with the mean absolute error (MAE) of the aerosol component volume fractions of 3.0%. In five samples the difference in the AW_f is more than 20.0%, though the overall MAE for AW_f is only 5.5%. In these five samples, the BC component is low and organic matter contributes substantially to the aerosol light absorption, resulting in underestimation of the AW_f volume fraction at high RH and overestimation for moderate RH. WSOM is overall slightly overestimated and AN is underestimated by only a few percent. The correlation between the input and retrieved aerosol volume fractions in the coarse mode is even better than that in fine mode. The regression coefficient for all samples is 0.99, and the MAE is only 2.0%. These results show the very small uncertainty in the retrieved aerosol component volume fractions.

To further evaluate the inversion results, errors were added to the synthetic data. To this end, three typical pollution cases were chosen in which the main pollutants are water soluble, biomass burning and dust aerosols, respectively, further referred to as WS, BB and DU pollution types. Each type is described by the different aerosol size distribution and refractive index parameters derived from Zhang et al. (2017). These parameters are listed in the supplementary, table S1. Note that although the acronyms of the three pollution types are the same as the aerosol component names above, it does not mean that each type includes only one single aerosol component, as illustrated below.

Figure 6 shows the aerosol volume size distribution, complex refractive index and eight aerosol components in the WS, BB and DU types used in this exercise. For the size distribution, the highest volume concentrations occur in the fine mode of the WS and BB types, whereas for the DU type the coarse mode dominates. For the complex refractive index, significant absorption occurs in the fine mode fraction of the BB type, while relatively low absorption occurs in the other models. In the WS type, the mass fraction of AW_f is close to 20% and for AN it is about 18%, significantly larger than for the other types. By comparison, the BC mass fraction in the BB type is close to 3%, and organic carbon is also high, with WSOM and WIOM mass fractions of 23% and 11%. In the DU type, the dust component is completely dominant, as expected, and the mass fractions of other components are less than 2%.

The three main sources of error in the model input parameters are the RH and the complex refractive index in the fine and coarse modes. The uncertainties due to inversion errors of the modal refractive index were discussed in detail in Zhang et al. (2017) and are directly used here to estimate their effects on aerosol components. For RH, the observation error is about 5% (WMO, 2008), in this exercise a larger error (10%) is introduced to more rigorously assess the uncertainty in the estimated

aerosol components. These typical uncertainties are listed in table S2. The total relative error (TRE), which is the propagated relative error calculated by the mean aerosol component error induced by the errors of sub-CRIs and RH in three pollution types, is used to assess the uncertainty in the aerosol composition inversion. As shown in table 3, the TRE of BC is 32.21%, less than other components in the fine mode, and the largest source of TRE is the imaginary part of the complex refractive index ($k_{f,440}$), with 25.68%. Compared with BC, the TRE of OM is larger (about 75%), primarily contributed by RH, followed by n_f. The uncertainty of the imaginary part impacts very little due to the low absorption of OM. The uncertainty of AN due to the imaginary part is low, but a very high uncertainty is caused by RH. Another component of IS is SC which usually occurs in the coarse mode. The large TRE of SC is contributed by the real part of the complex refractive index in the coarse mode (n_c), with 912.87%, leading to the largest TRE of IS. Affected by SC, the TRE of AW_c is also large due to n_c, but the TRE of AW_f is much smaller (50.05%). In the coarse mode, the TRE of DU is smallest in all of the aerosol components, only 15.79%, mainly caused by n_c. Overall, most of the uncertainties in the fine mode are from RH, and that in the coarse mode from the n_c. Fortunately, the RH observed by ground-based stations is accurate, with an error which is usually less than about 5% (WMO, 2008), which can significantly reduce the uncertainty in the retrieved aerosol scattering components. It should be noted that the uncertainties in table 3 are for single measurements. One important advantage of remote sensing is that multiple measurements can be made during a short period of time. Thus, the average uncertainty of the aerosol components can be effectively reduced by taking into account independent errors in each observation. In addition, the accuracy of the retrieved n_c needs to be improved in order to deal with the aerosol component inversion.

4. Results

4.1 Aerosol component retrievals

The averaged mass fractions of the aerosol components measured at 16 SONET sites are presented in Figure 7. Each pie chart is marked with the site name, coordinates, observation period and BC fraction. The mass fractions are also listed in table S3. The pie charts show that the coarse mode mass fraction usually dominates at the northern and northwestern sites. The mass fraction of the dust component is significantly higher than that of others, with a fraction of more than 50% at the western (Lhasa, Zhangye, Kashgar, Minqin and Xi'an), Beijing, Harbin and Songshan sites, which is different from surface observations of chemical components (Zhang et al., 2012; Liu et al., 2014). This is because sun photometers provide data integrated over the whole atmospheric column and thus include the dust transport layer near 4 km (Proestakis et al., 2018), where dust concentrations may be substantial, whereas surface observations are local point measurements. The lowest dust fractions are observed at southern sites, especially at the Guangzhou site, with a mass fraction of 31.5%. In contrast, the water content is dominant at southern sites in both the fine and coarse mode. The maximum AW (AW_f and AW_c) fraction occurs at the Guangzhou site (28.7%), and the lowest mass fractions of 2.0% and 7.5% are observed at the Lhasa and Kashgar site,

290 respectively. High AW_f occurs in the cities of east-central China due to the higher occurrence of inorganic salts with larger hygroscopicity in the fine mode at these sites, whereas the dominant AW_c in the western sites can be explained by the inorganic salt coating of larger particles in the dust source region (Rosenfeld et al., 2001). The IS fraction (AN and SC) gradually increases from north to south, which is consistent with the trend of the water content. The fraction of the AN sub-component is less than 7.0% at Lhasa, Zhangye, Kashgar and Minqin, whereas it is more than 20% at Chengdu, Guangzhou, Haikou and
295 Sanya. At the Zhoushan site also a high AN fraction is observed, up to 17.1%. For the SC component, the maximum value occurs at the Kashgar (17.1%) site. The high SC fraction at the southeast coastal sites is readily ascribed to the influence of the ocean; the high SC fraction at the Kashgar site is due to the paleo-marine source of dust over the Taklimakan Desert (Huang et al., 2010). The WIOM component fraction is high in the central sites but relatively low in the southern coastal and northwest sites. For the WSOM component, the low value of less than 3% appears only at northwestern sites (Zhangye, Kashgar and
300 Minqin). In the atmospheric column, the mass fraction of the BC component averaged over 16 sites is only 0.59%, lower than from near-surface in situ observations (usually 1% ~ 5%), which implies that the BC fraction may be reduced by the suspended layer with other components such as dust aerosol. Nevertheless, the unusually high mass fraction of BC in Shanghai could be due to observation uncertainty, also accompanied by the large error for aerosol component inversion.

The closure of the CRI between instantaneous optical-physical inversion and chemical estimation is examined by the data pair
305 frequency. Figure 8 shows scatter density plots of the chemically estimated and sunphotometer-retrieved imaginary parts of the fine mode at 675 nm (k_f) and 440 nm ($k_{f,440}$) and the real parts of fine mode at 440 nm(n_f). The points are colored by the number of data pairs (Retrieved, Estimated), which are sorted according to ordered pairs in 0.0005 intervals for the imaginary parts of CRI and 0.001 intervals for the real parts. The data pairs of k_f are closely concentrated around the 1:1 line, although with slight underestimation with 94.3% of the estimated values lower than the retrieved values; only 5.3% of the data pairs
310 have a relatively large absolute error (AE > 0.01). The mean bias is not large (-0.003), and the mean absolute value is equal to the mean absolute error (MAE = 0.003). There are two reasons for this slight underestimation in chemical estimation. On the one hand, the imaginary part of the refractive index of BC is much larger than for the other components due to its strong absorption. Thus, the inversion of the BC concentration is very sensitive to the estimation of the refractive index. As shown in table 3, although the TRE of BC is the lowest, the errors caused by k_f and $k_{f,440}$ are larger than for any other component. On the
315 other hand, k_f is not only affected by BC in the inversion process, but also affected by organic components (WSOM & WIOM) with spectral absorption characteristics. Therefore, in most cases, k_f is underestimated in chemical estimation and $k_{f,440}$ is overestimated (Bias = 0.007). The mean relative error (RE) is 27.1%, and 62.8% of the data points are below the average relative error line. This indicates that most inversion results have good optical closure. For the closure of the real part of the fine mode, the data pairs of n_f are also concentrated around the identity line, although 76.5% of the n_f is above the identity line.
320 Underestimation occurs mainly when n_f is larger than 1.56, because the only component with the real part of the CRI larger

than 1.56 is BC, but its concentration is mainly determined by the imaginary part. The bias of the estimated n_f (Bias = 0.009) is larger than that of k_f due to the fact that the value and the range of n_f are larger than that of k_f.

4.2 Seasonal variation

The seasonal variation of the aerosol component mass concentrations, averaged over 15 stations (Lhasa is not used due to lack of adequate seasonal data) and all available years, is shown as box-whisker plots in Figure 9. The top and bottom edges of each box represent the top and bottom quartiles (Q3 and Q1), and the corresponding whiskers are the outliers (Q3+1.5IQR and Q1-1.5IQR, IQR is interquartile range). The mean value is indicated by a plus sign (+), and the median value by a short line inside the box (–). Figure 9 shows that the DU component exhibits an obvious seasonality. The DU mass concentration is very high in the spring and the mean value reaches up to 332.9 mg m^{-2} due to dust transport from the northwest of China. With the weakening of dust transport and the increase of moisture, the DU fraction decreases in other seasons, with a mean value of around 240.0 mg m^{-2}. Although the DU concentration is lower in the summer than in other seasons, it is still relatively high near the dust source area, which results in a large difference between the upper and lower quartiles. In contrast, the AN mass concentration mean value peaks in the summer (76.8 mg m^{-2}), whereas a minimum occurs in the spring (47.7 mg m^{-2}). It is worth noting that although the mean value in the winter is not high (51.1 mg m^{-2}), the interval between the upper and lower quartiles of AN is the smallest in the winter. The minimum value of AN (17.9 mg m^{-2}) is higher than in other seasons (4.1 mg m^{-2} in spring, 9.5 mg m^{-2} in summer, and 11.1 mg m^{-2} in autumn). The seasonal variation of the water content is slightly different from that of inorganic salts. The low values of mean AW_f occur in the spring, while AW_c is significantly lower in the winter (21.0 mg m^{-2}) than in other seasons. The difference between the upper and lower quartiles of AW_f in the summer is larger than in other seasons indicating that in the summer the aerosol at some sites has a low hygroscopicity. The OM mass concentration is slightly higher in the winter than that in other seasons probably due to the occurrence of haze pollution in the winter, with mean concentrations of the WIOM and WSOM fractions of 22.3 and 38.8 mg m^{-2}, respectively. In the summer, the OM concentration is only about two thirds of that in the winter. The median value of the BC mass concentration is higher in the winter (3.0 mg m^{-2}), which can be related to heating in northern China. Low concentrations of BC in the other seasons are mainly due to the influence of frequent dust events in the spring and high aerosol hygroscopic growth in the summer. Similar to AN, the SC concentration peaks in the summer and has a minimum in the winter, due to the influence of the Asian monsoon. The median values in these two seasons are respectively 41.6 and 19.6 mg m^{-2}.

The seasonal variation of the main aerosol components in the fine mode is discussed on a regional basis (Figure 10). BC concentrations in typical northern regions are higher than in southern regions, because of emissions due to winter heating only in the north. Other BC sources are vehicle emissions and biomass burning. Adverse meteorological conditions in winter result in the accumulation of BC in the atmosphere resulting in high BC values in both the north and the south. The highest BC mass

concentrations in the northern region in the winter is 4.3 mg m^{-2}. OM is one of the dominant components in the fine mode, with sources similar to those of BC. The impact of biomass burning in the winter and spring over south China (Chen et al., 2017) is significant, leading to OM concentrations of more than 50.0 mg m^{-2}. In the northern region, much biomass burning occurs in the autumn (Wang et al., 2020). With the influence of heating, the OM level in the north can reach up to 80.1 mg m^{-2}. Therefore, the OM mass concentration in the northern region is only low in the summer (50.8 mg m^{-2}). AN is usually formed by secondary reactions of gaseous precursors in complex air pollution areas. In both the northern and the southern region, AN mass concentration is larger in the summer than in other seasons, and the seasonal variation in the southern region is significantly smaller than that in the north. The mean AN mass concentration in the southern region is 8.7 mg m^{-2} higher than that in the northern region. This suggests that more AN is produced by secondary reactions in the humid climate in the south than in the northern region.

4.3 Interannual variation

Figure 11 shows the interannual variations of the aerosol component mass concentrations in the atmospheric column from 2010-2016. The 16 SONET sites have been established in succession, so the number of available observations increased year by year with the longest time series from the Beijing site (see also Table 1). The annual mean mass concentrations shown in Figure 11 are averages over all sites, i.e. the number of sites was not accounted for and, in particular in the earlier years (2010-2011), the annual mean may thus be representative for one (Beijing) or a few sites. Therefore, the annual means for each site available has been plotted as well. Figure 11 shows that the annual mean mass concentrations of most of the aerosol components in the fine mode increased in most of the first years and then decreased. Influenced by China's environmental control policies, the mean AN decreased significantly from 72.4 mg m^{-2} in 2011 to 50.5 mg m^{-2} in 2016, i.e. a reduction by 21.9 mg m^{-2}. The yearly mass concentrations of AN at most sites also follow a downward trend, and AN in the southeastern coastal sites are significantly higher than that in the northwestern sites. In contrast, the mean BC mass concentration shows a peak (3.9 mg m^{-2}) in 2011, drops in 2012 to the lowest value during the whole period (2.3 mg m^{-2}), then increases somewhat to a second peak (2.7 mg m^{-2}) in 2013. After a decrease in 2014, BC climbed to 2.8 mg m^{-2} in 2016. In the southeastern coastal and northwestern sites, BC concentrations were relatively low. The unusually high values at Shanghai in 2016 may be due to observational errors. Similar to BC, AW_f also experienced a small fluctuation after a significant decline in 2012. The AW_f in aerosol measured at the southern sites are higher than that at other sites. The fine mode WIOM and WSOM components show different behaviour. WIOM reached a peak in 2013, with the peak value of 32.3 mg m^{-2}, and then showed a significant decline after 2013. WSOM also reached a peak concentration of 35.8 mg m^{-2} in 2013, which is 2 mg m^{-2} lower than the peak in 2016, and overall the concentrations increased. These results suggest that the policy of air pollution control in China is effective in controlling inorganic salts and WIOM aerosols, while WSOM still needs to be further controlled. The concentrations of the

coarse mode aerosol components fluctuate somewhat during the observation period, with a slight peak in 2013. The concentration of each component in the coarse mode at the northwestern sites is higher than that at other sites, which can be related to the high fraction of large particles. Due to the large influence of geographical factors on the coarse mode aerosol components, DU in 2010 (only Beijing site) was significantly larger than in other years. Since 2014, the mean DU mass concentration has increased, while a downward tendency is observed in the AW_c and SC concentrations since 2013. Coarse mode aerosols usually derive from natural sources, and their variations can be associated with changes in the meteorological conditions.

5. Conclusions and discussions

The accurate measurement of atmospheric aerosol components plays an important role in reducing the uncertainty of climate assessment. In the current study, we updated the refractive index calculation in a multi-component liquid system and improved the component inversion algorithm of Zhang et al. (2018) to retrieve atmospheric columnar aerosol components including black carbon (BC), organic matter (WSOM/WIOM), inorganic salt (AN & SC), dust-like (DU) and water content in the fine and coarse modes (AW_f & AW_c). This algorithm was applied to data from the SONET sun photometer network, and the regional distribution and interannual variation of atmospheric aerosol components in China were analyzed for the period from 2010 to 2016. The results show that the dust-like component is dominant in northern China, but the aerosol water content (AW_f & AW_c) is dominant in the southern coastal region. The inorganic salt (AN) in the fine mode has a significant seasonal variation, with a mass concentration of 76.8 mg m^{-2} in the summer which is significantly higher than that in other seasons. Meanwhile, the AN concentrations have significantly decreased from 2011 to 2016, which is inseparable from China's environmental control policies. However, the slight increase in WSOM and BC is still noteworthy.

As the aerosol concentrations in the atmospheric column obtained from the inversion of remote sensing data are different from those measured by in situ measurements near the surface, such as on-line aerosol mass spectrometers, the validation of the retrieval results is difficult. Proestakis et al. (2018) used data from the Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP) on the CALIPSO satellite to analyze the distribution of mineral dust over China and the results show the higher concentration of the DU component in the atmospheric column over northern China. Similarly, Huang et al. (2010) provided a basis for the high SC content at the Kashgar station due to the paleo-marine source. However, for the direct comparison of our retrievals with independent data, airborne measurements of the vertical distribution of atmospheric aerosol components are needed (Kahn et al., 2017). In future research, we will design a verification experiment to comprehensively evaluate the results from our inversion method.

The method presented can be used not only for ground-based sun-sky photometer measurements, but also for other remote sensing instruments (e.g. lidar), and even for satellite remote sensing in the future. Meanwhile, as long as measurements of

multi-wavelength extinction coefficients and aerosol particle size distributions are available, the inversion of atmospheric particulate matter composition can also be performed using comprehensive observations with multiple instruments near the surface. Therefore, this method can be widely used in low-cost and wide-area measurements in the future, providing a possibility for obtaining the global distribution of aerosol composition.

Data availability. The aerosol component data used in this study can be requested from the corresponding author (lizq@radi.ac.cn).

Author contributions. ZL conceived and designed the study. YC collected and processed the meteorological data. KL and YX collected the remote sensing data. CZ collected the DEM data and draw the map. YZ improved the aerosol component method and performed the inversions. YZ analyzed the spatiotemporal trends of aerosol component concentrations. YZ and GL prepared the paper with contributions from all coauthors.

Acknowledgments. This work was supported by the National Key R&D Program of China (grant 2016YFE0201400) and the National Natural Science Fund of China (41601386).

References

Arola, A., Schuster, G., Myhre, G., Kazadzis, S., Dey, S., and Tripathi, S. N.: Inferring absorbing organic carbon content from AERONET data, Atmospheric chemistry and physics, 11(1), 215– 225, 2011.

Bond, T. and Bergstrom, R.: Light absorption by carbonaceous particles: an investigative review, Aerosol Science and Technology, 40, 27–67, 2006.

Boucher, O., Randall, D., Artaxo, P., Bretherton, C., Feingold, G., Forster, P., Kerminen, V.-M., Kondo, Y., Liao, H., Lohmann, U., Rasch, P., Satheesh, S.K., Sherwood, S., Stevens, B. and Zhang, X.Y.: Clouds and Aerosols. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2013.

Bougiatioti, A., Zarmpas, P., Koulouri, E., Antoniou, M., Theodosi, C., Kouvarakis, G., Saarikoski, S., Mäkelä, T., Hillamo, R., Mihalopoulos, N.. Organic, elemental and water-soluble organic carbon in size segregated aerosols, in the marine boundary layer of the Eastern Mediterranean. Atmos. Environ. 64, 251 – 262, 2013.

Chalbot, M.G., Chitranshi, P., da Costa, G.G., Pollock, E., Kavouras, l.G.. Characterization of water-soluble organic matter in urban aerosol by 1H-NMR spectroscopy. Atmos. Environ. 128, 235–245, 2016.

Chen, J., Li, C., Ristovski, Z., Milic, A., Gu, Y., Islam, M., Wang, S., Hao, J., Zhang, H., He, C., Guo, H., Fu, H., Miljevic, B., Morawska, L., Thai, P., Lam, Y., Pereira, G., Ding, A., Huang, X. and Dumka, U. C.: A review of biomass burning: Emissions and impacts on air quality, health and climate in China, Science of the Total Environment, 579, 1000–1034, 2017.

Chen, Y. and Bond, T. C.: Light absorption by organic carbon from wood combustion, Atmos. Chem. Phys., 10, 1773–1787, doi:10.5194/acp-10-1773-2010, 2010.

Choi, Y. and Ghim, Y. S.: Estimation of columnar concentrations of absorbing and scattering fine mode aerosol components using AERONET data, Journal of Geophysical Research Atmospheres, 121, 13628-13640, 2016.

Drozd G., Woo J., Hakkinen S.A.K., Nenes A., and McNeill V.F. Inorganic salts interact with oxalic acid in submicron particles to form material with low hygroscopicity and volatility, Atmos. Chem. Phys., 14, 5205–5215, 2014.

Dubovik, O., and King, M. A flexible inversion algorithm for retrieval of aerosol optical properties from sun and sky radiance measurements, J. Geophys. Res. Atmos., 105(D16), 20,673–20,696, 2000.

Dubovik, O., Smirnov, A., Holben, B.N., King, M.D., Kaufman, Y.J., Eck, T.F., Slutsker, I.: Accuracy assessments of aerosol optical properties retrieved from aerosol robotic network (AERONET) Sun and sky radiance measurements. J. Geophys. Res. Atmos. 105, 9791-9806, 2000.

Dubovik, O., Herman, M., Holdak, A., Lapyonok, T., Tanré, D., Deuzé, J. L., Ducos, F., Sinyuk, A., and Lopatin, A.: Statistically optimized inversion algorithm for enhanced retrieval of aerosol properties from spectral multi-angle polarimetric satellite observations, Atmos. Meas. Tech., 4, 975-1018, doi:10.5194/amt-4-975-2011, 2011

Ganguly, D., Ginoux, P., Ramaswamy, V., Dubovik, O., Welton, J., Reid, E. A. and Holben, B. N.: Inferring the composition and concentration of aerosols by combining AERONET and MPLNET data: Comparison with other measurements and utilization to evaluate GCM output, J Geophys Res-Atmos, 114(D16), 2009.

Heller, W., Remarks on refractive index mixture rules, The Journal of Physical Chemistry, 69(4), 1123-1129, 1965.

Huang, K., Zhuang, G., Li, J., Wang, Q., Sun, Y., Lin, Y and Fu, J. S.: Mixing of Asian dust with pollution aerosol and the transformation of aerosol components during the dust storm over China in spring 2007, J. Geophys. Res., 115, D00K13, doi:10.1029/2009JD013145, 2010.

Jing B., Tong S., Liu Q., Li K., Wang W., Zhang Y. and Ge M. Hygroscopic behavior of multicomponent organic aerosols and their internal mixtures with ammonium sulfate, Atmos. Chem. Phys., 16, 4101–4118, 2016.

Kahn, R.A., Berkoff, T.A., Brock, C., Chen, G., Ferrare, R.A., Ghan, S., Hansico, T.F., Hegg, D.A., Vanderlei Martins, J., McNaughton, C.S., Murphy, D.M., Ogren, J.A., Penner, J.E., Pilewskie, P., Seinfeld, J.H., and Worsnop, D.R.: SAM-CAAM: A Concept for Acquiring Systematic Aircraft Measurements to Characterize Aerosol Air Masse," Bull. Amer. Meteor. Soc., 98, 2215–2228) doi:10.1175/BAMS-D-16-0003.2, 2017.

Kandler, K., and L. Schütz, Climatology of the average water-soluble volume fraction of atmospheric aerosol, Atmospheric Research, 83, 77–92, 2007.

Kreidenwies S. M., Petters, M. D. and DeMott, P.: Single-parameter estimates of aerosol water content, Environ. Res. Lett., 3, 035002, 1-7, 2008.

Krivácsy, Z., Gelencsér, A., Kiss, G., Mészáros, E., Molnár, Á., Hoffer, A., Mészáros, T., Sárvári, Z., Temesi, D., Varga, B., Baltensperger, U., Nyeki, S., Weingartner, E.. Study on the chemical character of water soluble organic compounds in fine atmospheric aerosol at the jungfraujoch. J. Atmos. Chem. 39, 235 – 259, 2001.

Koven, C. and Fung, I.: Inferring dust composition from wavelength-dependent absorption in Aerosol Robotic Network AERONET data, J. Geophys. Res., 111, D14205, doi:10.1029/2005JD006678, 2006.

Lesins, G., Chylek, P., and Lohmann, U.: A study of internal and external mixing scenarios and its effect on aerosol optical properties and direct radiative forcing, J Geophys Res-Atmos, 107(D10), 2002.

Li, L., Dubovik, O., Derimian, Y., Schuster, G. L., Lapyonok, T., Litvinov, P., Ducos, F., Fuertes, D., Chen, C., Li, Z., Lopatin,

A., Torres, B. and Che, H.: Retrieval of aerosol components directly from satellite and ground-based measurements, Atmos. Chem. Phys., 19, 13409–13443, 2019.

Li, Z., Zhang, Y., Xu, H., Li, K., Dubovik, O., and Goloub, P.: The fundamental aerosol models over China region: A cluster analysis of the ground-based remote sensing measurements of total columnar atmosphere. Geophysical Research Letters, 46. https://doi.org/10.1029/2019GL082056, 2019.

Li, Z. Q., Xu, H., Li, K. T., Li, D. H., Xie, Y. S., Li, L., et al.: Comprehensive study of optical, physical, chemical, and radiative properties of total columnar atmospheric aerosols over China: An overview of Sun–Sky Radiometer Observation Network (SONET) measurements. Bulletin of the American Meteorological Society, 99(4), 739–755, 2018.

Li, Z., Gu, X., Wang, L., Li, D., Xie, Y., Li, K., Dubovik, O., Schuster, G., Goloub, P., Zhang, Y., Li, L., Ma, Y., and Xu, H.: Aerosol physical and chemical properties retrieved from ground-based remote sensing measurements during heavy haze days in Beijing winter, Atmospheric Chemistry and Physics, 13, 10171-10183, 2013.

Liu, Q., Liu, Y., Yin, J., Zhang, M. and Zhang, T.: Chemical characteristics and source apportionment of PM_{10} during Asian dust storm and non-dust storm days in Beijing, Atmospheric Environment, 91, 85-94, 2014.

Ma Q., He H., Liu C. Hygroscopic properties of oxalic acid and atmospherically relevant oxalates, Atmospheric Environment, 69, 281-288, 2013.

Mayol-Bracero, O.L., Guyon, P., Graham, B., Roberts, G., Andreae, M.O., Decesari, S., Facchini, M.C., Fuzzi, S., Artaxo, P.. Water-soluble organic compounds in biomass burning aerosols over Amazonia 2. apportionment of the chemical composition and importance of the polyacidic fraction. Atmos. Chem. Phys. 107 (D20), 8091. http://dx.doi.org/10.1029/2001JD000522, 2002.

McConnell, C. L., Highwood, E. J., Coe, H., Formenti, P., Anderson, B., Osborne, S., Nava, S., Desboeufs, K., Chen, G. and Harrison, M. A. J.: Seasonal variations of the physical and optical characteristics of Saharan dust: Results from the Dust Outflow and Deposition to the Ocean (DODO) experiment, J Geophys Res-Atmos, 113, D14S05, doi:10.1029/2007JD009606, 2008.

Petters M. D. and Kreidenweis S. M.: A single parameter representation of hygroscopic growth and cloud condensation nucleus activity, Atmospheric Chemistry and Physics, 7, 1961–1971, 2007.

Proestakis, E., Amiridis, V., Marinou, E., Georgoulias, A. K., Solomons, S., Kazadzis, S., Chimot, J., Che, H., Alexandri, G., Binietoglou, I., Daskalopoulou, V., Kourtidis, K. A., de Leeuw, G. and van der, A R. J.: Nine-year spatial and temporal evolution of desert dust aerosols over South and East Asia as revealed by CALIOP, Atmos. Chem. Phys., 18, 1337¨C1362, 2018.

Rosenfeld, D., Rudich, Y. and Lahav, R.: Desert dust suppressing precipitation: a possible desertification feedback loop, Proceedings of the National Academy of Sciences of the United States of America, 98(11), 5975-5980, 2001.

Satheesh, S. K. and Krishna, M. K.: Radiative effects of natural aerosols: A review, Atmospheric Environment, 39(11), 2089-2110, 2005.

Satheesh, S. K. and Srinivasan, J.: Enhanced aerosol loading over Arabian Sea during the pre-monsoon season: Natural or anthropogenic. Geophysical Research Letters, 29(18), 211-214, 2002a.

Satheesh, S. K., Ramanathan, V., Holben, B. N., Moorthy, K. K., Loeb, N. G., Maring, H., Prospero, J. M., and Savoie, D.: Chemical, microphysical, and radiative effects of Indian Ocean aerosols, Journal of Geophysical Research, 107(D23), 4725, 2002b.

Satheesh, S. K., Ramanathan, V., Li-Jones, X., Lobert, J. M., Podgorny, I. A., Prospero, J. M., Holben, B. N., and Loeb, N. G.: A model for the natural and anthropogenic aerosols over the tropical Indian Ocean derived from Indian Ocean Experiment

data, Journal of Geophysical Research, 104(D22), 27421-27440, 1999.

Schuster, G. L., Dubovik, O., and Arola, A.: Remote sensing of soot carbon – Part 1: Distinguishing different absorbing aerosol species, Atmospheric chemistry and physics, 16, 1565-1585, 2016.

Schuster, G. L., Lin, B., and Dubovik, O.: Remote sensing of aerosol water uptake, Geophysical Research Letters, 36, L03814, doi:10.1029/2008GL036576, 2009.

Schuster, G. L., Dubovik, O., Holben, B. N., and Clothiaux, E. E.: Inferring black carbon content and specific absorption from Aerosol Robotic Network (AERONET) aerosol retrievals, Journal of Geophysical Research, 110, D10S17, doi:10.1029/2004JD004548, 2005.

Sun, H., Biedermann, L., and Bond, T.: Color of brown carbon: a model for ultraviolet 20 and visible light absorption by organic carbon aerosol, Geophys. Res. Lett., 34, L17813, doi:10.1029/2007GL029797, 2007.

Tang, I. N.: Chemical and size effects of hygroscopic aerosol on light scattering coefficients, Journal of Geophysical Research, 101(D14), 19245-19250, 1996.

Toon, O. B. and Pollack, J. B.: The optical constants of several atmospheric aerosol species: Ammonium sulfate, aluminum oxide, and sodium chloride, Journal of Geophysical Research, 81(33), 5733-5748, 1976.

van Beelen, A. J., Roelofs, G. J. H., Hasekamp, O. P., Henzing, J. S., and Rockmann, T.: Estimation of aerosol water and chemical composition from AERONET Sun-sky radiometer measurements at Cabauw, the Netherlands, Atmos. Chem. Phys., 14(12), 5969–5987, doi:10.5194/acp-14-5969-2014, 2014.

Wagner, R., Ajtai, T., Kandler, K., Lieke, K., Linke, C., Müller, T., Schnaiter, M., and Vragel, M.: Complex refractive indices of Saharan dust samples at visible and near UV wavelengths: a laboratory study, Atmospheric Chemistry Physics, 12, 2491–2512, 2012.

Wang, L., Li, Z. Q., Tian, Q. J., Ma, Y., Zhang, F. X., Zhang, Y., Li, D. H., Li, K. T., and Li, L.: Estimate of aerosol absorbing components of black carbon, brown carbon, and dust from ground-based remote sensing data of sun-sky radiometers, Journal of Geophysical Research, 118, 6534-6543, 2013.

Wang, L., Jin, X., Wang, Q., Mao, H., Liu, Q., Weng, G. and Wang, Y.: Spatial and temporal variability of open biomass burning in Northeast China from 2003 to 2017, Atmospheric and Oceanic Science Letters, DOI: 10.1080/16742834.2020.1742574, 2020.

Wozniak, A.S., Shelley, R.U., Sleighter, R.L., Sleighter, R.L., Abdulla, H.A.N., Morton, P.L., Landing, W.M., Hatcher, P.G.. Relationships among aerosol water soluble organic matter, iron and aluminum in European, North African, and Marine air masses from the 2010 US GEOTRACES cruise. Mar. Chem. 154, 24–33, 2013.

WMO, 2008. Guide to Meteorological Instruments and Methods of Observation, seventh ed. (WMO No.8).

Xie, Y., Li, Z., Zhang, Y.X., Zhang, Y., Li, D. H., Li, K. T. Xu, H., Zhang, Y., Wang, Y. Q., Chen, X. F., Schauer, J. J., Bergin, M.: Estimation of atmospheric aerosol composition from ground-based remote sensing measurements of Sun-sky radiometer, Journal of Geophysical Research Atmospheres, doi: 10.1002/2016JD025839, 2017.

Zappoli, S., Andracchio, A., Fuzzi, S., Facchini, M.C., Gelencsér, A., Kiss, G., Krivácsy, Z., Molnár, Á., Mészáros, E., Hansson, H.-C., Rosman, K., Zebühr, Y.. Inorganic, organic and macromolecular components of þne aerosol in di¤erent areas of Europe in relation to their water solubility. Atmos. Environ. 33, 2733 – 2743, 1999.

Zhang, X.Q., Mcmurry, P.H., Hering, S.V., Casuccio, G.S.: Mixing characteristics and water content of submicron aerosols measured in Los Angeles and at the Grand Canyon. Atmos. Environ. 27A (10), 1593–1607, 1993.

Zhang, X. Y., Wang, Y. Q., Niu, T., Zhang, X. C., Gong, S. L., Zhang, Y. M. and Sun, J. Y.: Atmospheric aerosol compositions

in China: spatial/temporal variability, chemical signature, regional haze distribution and comparisons with global aerosols, Atmos. Chem. Phys., 12, 779–799, 2012.

Zhang, Y., Li, Z., Zhang, Y. H., Li, D., Qie, L., Che, H., and Xu, H.: Estimation of aerosol complex refractive indices for both fine and coarse modes simultaneously based on AERONET remote sensing products, Atmospheric Measurement Techniques, 10, 3203-3213, 2017.

Zhang, Y., Li, Z., Sun, Y., Lv, Y., Xie, Y.: Estimation of atmospheric columnar organic matter (OM) mass concentration from remote sensing measurements of aerosol spectral refractive indices, Atmospheric Environment, 179, 107-117, 2018.

Table 1. SONET sites (name, location and geographical aspects) and meteorological stations used in this study.

SONET Site						Meteorological Station					
Name	**Abbr**	**Lon (◦)**	**Lat (◦)**	**Alt (m)**	**Obs Period**	**No.***	**Lon (◦)**	**Lat (◦)**	**Alt (m)**	**Geo feature**	**Geo region**
Lhasa	LS	91.2	29.6	3678	2016.03-2016.05	55591	91.1	29.7	3649	Plateau	Qinghai-Tibetan
Kashgar	KS	75.9	39.5	1320	2013.09-2016.11	51709	76.0	39.5	1289	Desert	Northwest
Zhangye	ZY	100.3	38.8	1364	2012.08-2016.10	52652	100.4	38.9	1483	Gobi & desert	
Minqin	MQ	103.0	38.6	1589	2012.02-2016.10	52681	103.1	38.6	1368	Desert & hill	
Xi'an	XA	108.9	34.2	389	2012.05-2016.11	57039	108.9	34.2	433	Half mountain, half plain	
Beijing	BJ	116.3	40.0	59	2009.12-2016.11	54399	116.3	40.0	46	Hill (megacity)	Northern
Harbin	HrB	126.6	45.7	223	2013.12-2016.11	50953	126.8	45.8	118	Plain	
Songshan	SS	113.1	34.5	475	2013.12-2016.11	57084	113.1	34.5	1178	Mountain & hill	
Nanjing	NJ	119.0	32.1	52	2013.01-2016.07	58238	118.8	32.0	35	Plain & hill	Southern
Shanghai	SH	121.5	31.3	84	2013.03-2016.04	58362	121.5	31.4	6	Alluvial plain (megacity)	
Hefei	HF	117.2	31.9	36	2013.11-2016.11	58321	117.2	31.9	27	Hill	
Zhoushan	ZS	122.1	29.9	29	2012.02-2016.11	58477	122.1	30.0	36	Islands	
Chengdu	CD	104.0	30.6	510	2013.06-2016.07	56276	103.8	30.4	461	Basin	
Guangzhou	GZ	113.4	23.1	28	2011.10-2016.11	59287	113.3	23.2	41	Mountain, hill & plain (megacity)	
Haikou	HK	110.3	20.0	22	2014.03-2016.03	59758	110.3	20.0	64	Island	
Sanya	SY	109.4	18.3	29	2014.09-2016.11	59948	109.5	18.2	419	Island	

570 * "No." is the meteorological station number.

Table 2. Growth factor derived hygroscopicity parameter (κ), complex refractive indexes ($m = n - ik$) at four wavelengths and effective density (ρ) of model components. Real and imaginary parts at four standard AERONET aerosol product wavelengths (440, 675, 870 and 1020 nm) are considered.

Component		Growth factor derived κ	Real Part				Imaginary Part		ρ (g cm^{-3})
			n_{440}	n_{675}	n_{870}	n_{1020}	k_{440}	$k_{675\sim1020}$	
OM	WIOM	0.000	1.530[c]	1.530	1.530	1.530	0.035[d]	0.001	1.547[i]
	WSOM	0.000[a]	1.530[c]	1.530	1.530	1.530	0.006[d]	0.000	
AN		0.547[b]	1.559[e]	1.553	1.550	1.548	0.000[e]	0.000	1.760[i]
BC		0.000	1.950[f]	1.950	1.950	1.950	0.790[f]	0.790	1.800[i]
AW		0.000	1.337[e]	1.332	1.330	1.328	0.000[g]	0.000	1.000[i]
DU		0.000	1.534[g]	1.534	1.534	1.534	0.002[h]	0.001	2.650[i]
SC		1.120[a]	1.562[h]	1.541	1.534	1.530	0.000[i]	0.000	2.165[i]

575 [a] Petters and Kreidenweis, 2007; [b] Kreidenweis et al., 2008; [c] Sun et al., 2007; [d] Chen and Bond, 2010; [e] Schuster et al., 2005; [f] Bond and Bergstrom, 2006; [g] Koven and Fung, 2006; [h] Toon et al., 1976; [i] Zhang et al., 1993.

Table 3. Estimated total relative errors (TRE) of aerosol component mass fractions in the three aerosol models used to evaluate the aerosol component classification inversion algorithm.

Aerosol components		RH	Fine mode			Coarse mode			TRE*	
			n_f	$k_{f,440}$	k_f	n_c	$k_{c,440}$	k_c		
BC		5.74%	0.59%	25.68%	18.57%	0.00%	0.00%	0.00%	32.21%	
OM	**WIOM**	75.82%	4.55%	5.28%	1.08%	0.00%	0.00%	0.00%	76.15%	74.73%
	WSOM	51.60%	51.92%	3.44%	2.11%	0.00%	0.00%	0.00%	73.31%	
IS	**AN**	207.00%	60.86%	7.07%	6.04%	0.00%	0.00%	0.00%	215.96%	564.42%
	SC	25.71%	0.00%	0.00%	0.00%	912.51%	2.16%	1.23%	912.87%	
AW	**AW_f**	49.77%	3.80%	3.71%	0.00%	0.00%	0.00%	0.00%	50.05%	481.32%
	AW_c	8.95%	0.00%	0.00%	0.00%	912.55%	2.10%	1.17%	912.60%	
DU		0.34%	0.00%	0.00%	0.00%	15.78%	0.04%	0.02%	15.79%	

*TRE $= \sqrt{\sum_1^7 \bar{x}_i^2}$, where $\bar{x}$ represents the mean error of aerosol components from three aerosol types. The RH is given input error of ±10% and the inversion errors of sub-CRIs is from Zhang et al., 2017 listed in Table S2.

585

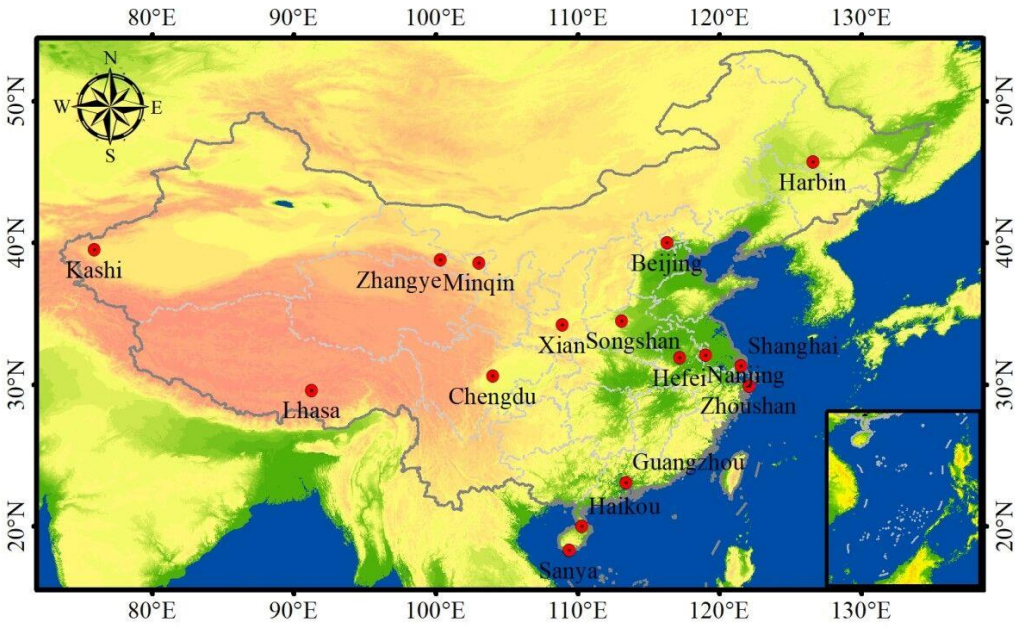

Figure 1. Locations of the 16 Sun-sky radiometer observation network (SONET) sites projected on the elevation map of China.

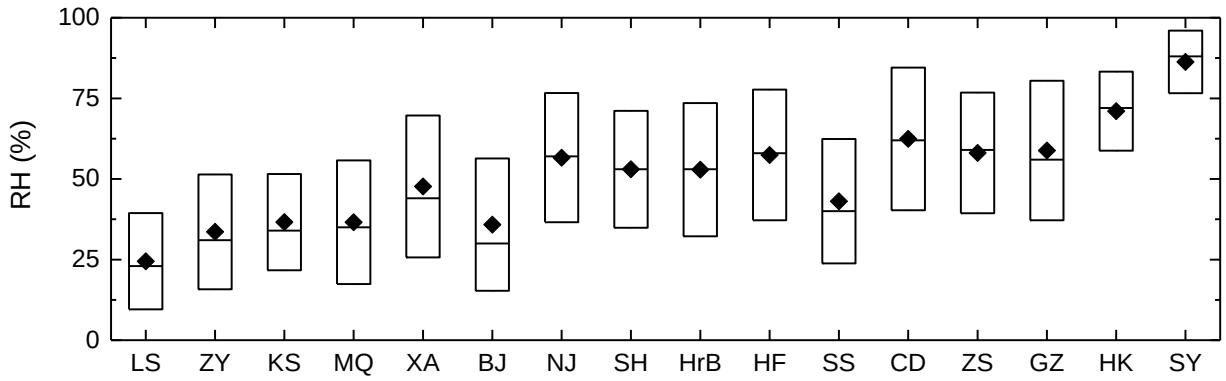

Figure 2. Boxplots of the relative humidities observed near each of the SONET sites. The observation periods for each site are shown in table 1. The line and the diamond represent the median and mean values, respectively, and the box shows the standard deviation (1 σ).

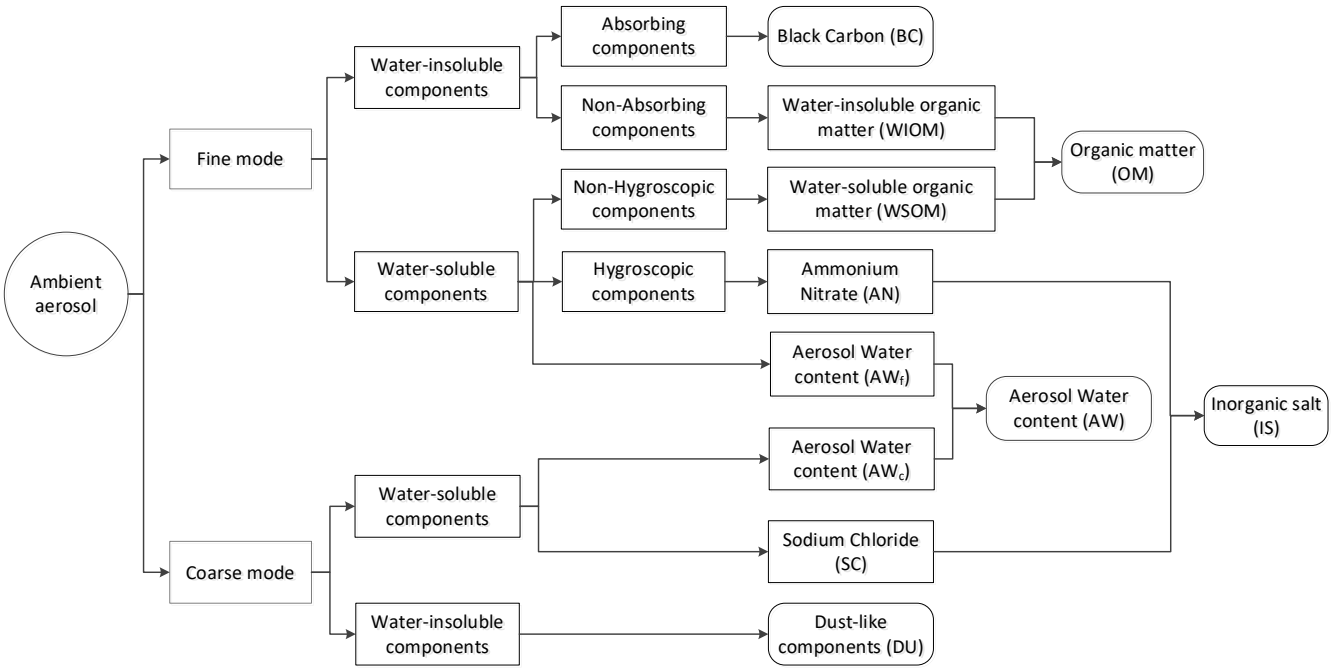

Figure 3. Aerosol component classification scheme.

600

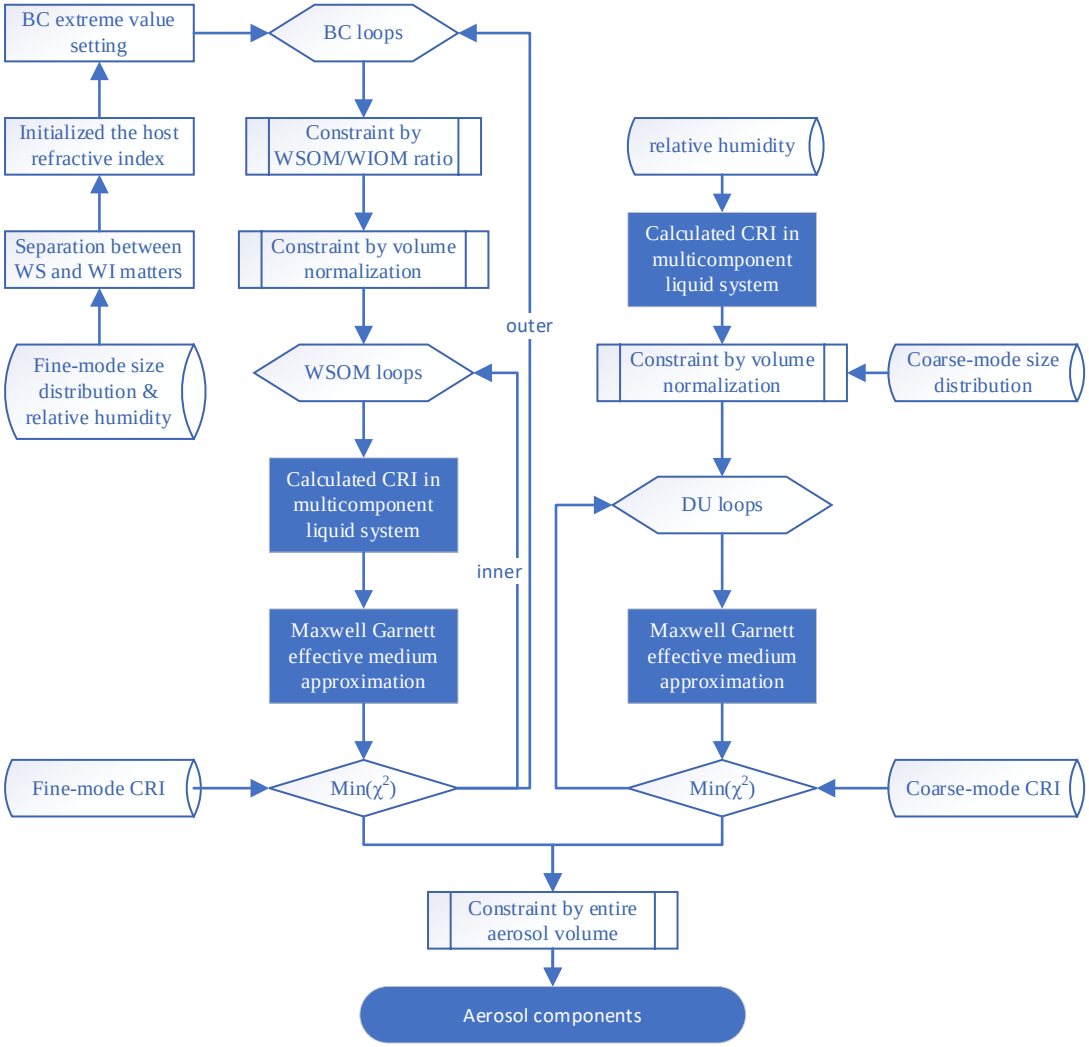

Figure 4. Flowchart of the aerosol component classification inversion algorithm.

605

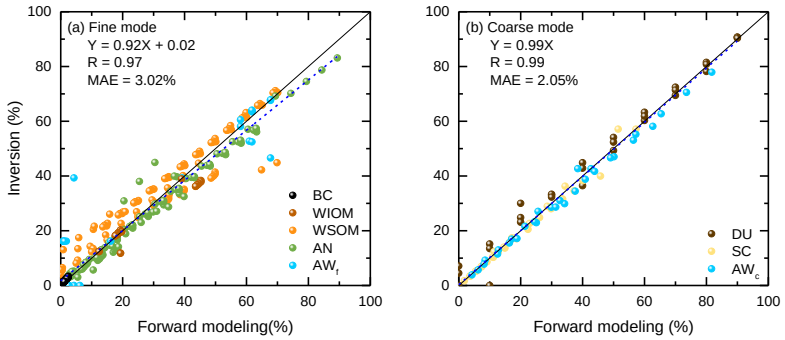

Figure 5. Scatter plots of volume fractions of aerosol components in the fine (left) and coarse (right) modes retrieved using the algorithm described in Chapter 3, versus those used as input calculated with the forward model. The solid line is the 1:1 line, and the dash line is the fitting line.

610

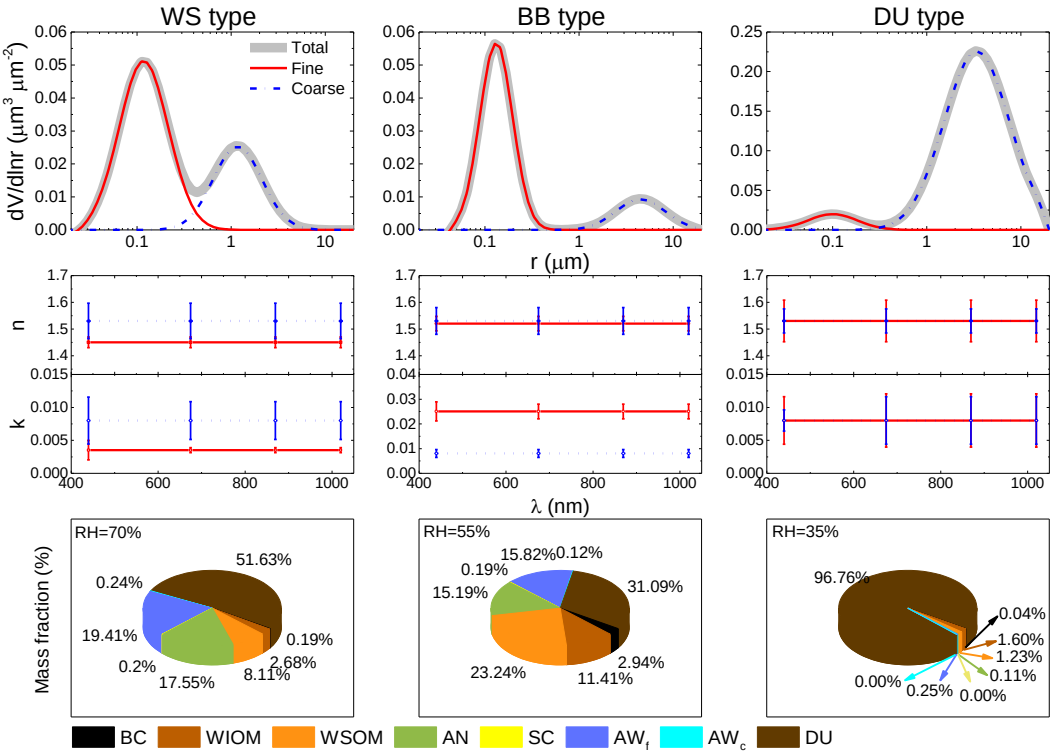

Figure 6. The fine and coarse-mode volume size distribution, complex refractive index and aerosol components describing the aerosol models used in the synthetic case study (WS: water soluble, BB: biomass burning, DU: dust).

615

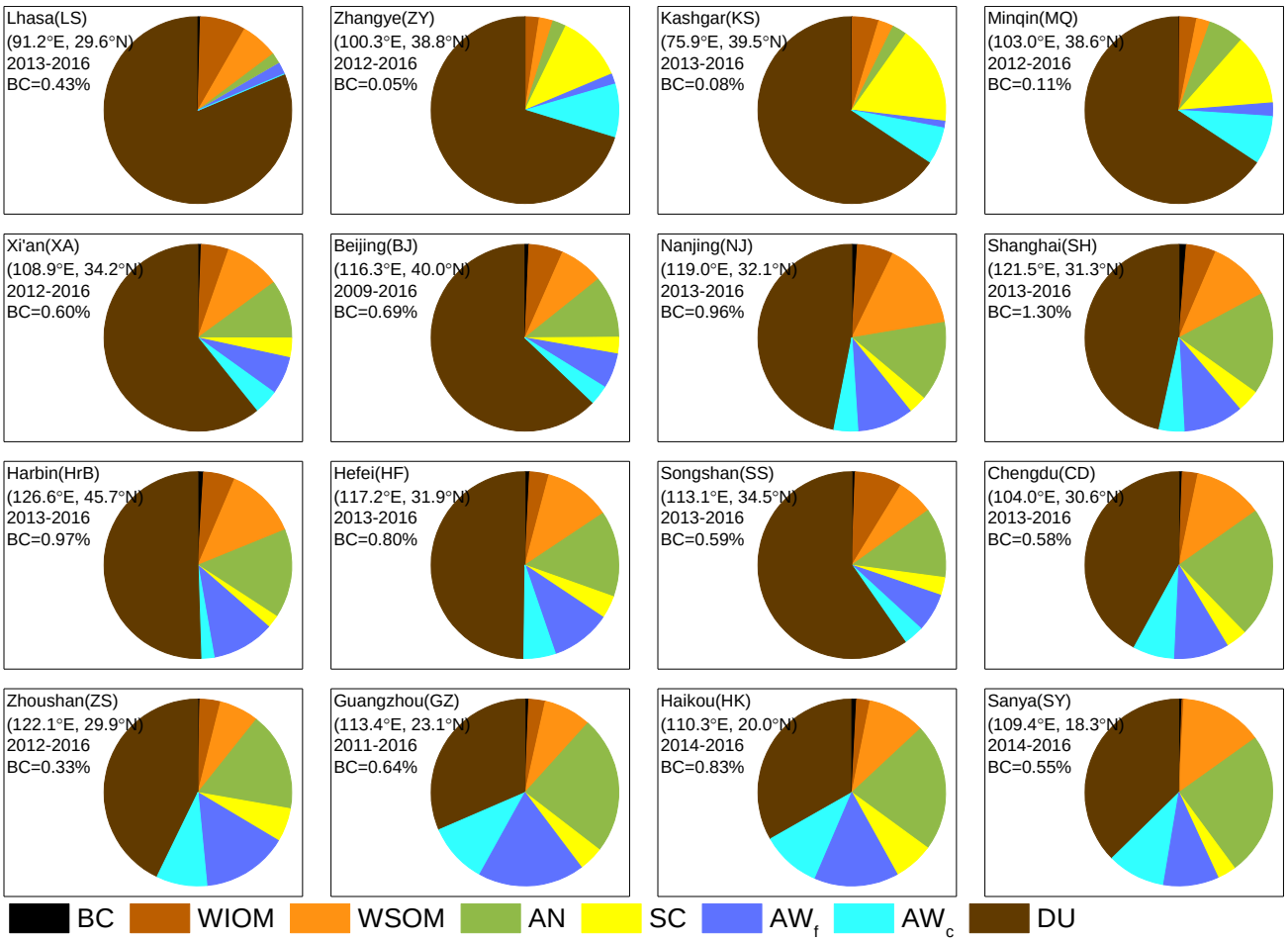

Figure 7. The averaged mass fraction of aerosol components at SONET sites. The site name, location, observation period and BC fraction are marked in each subgraph. The mass fractions of other components are listed in table S3.

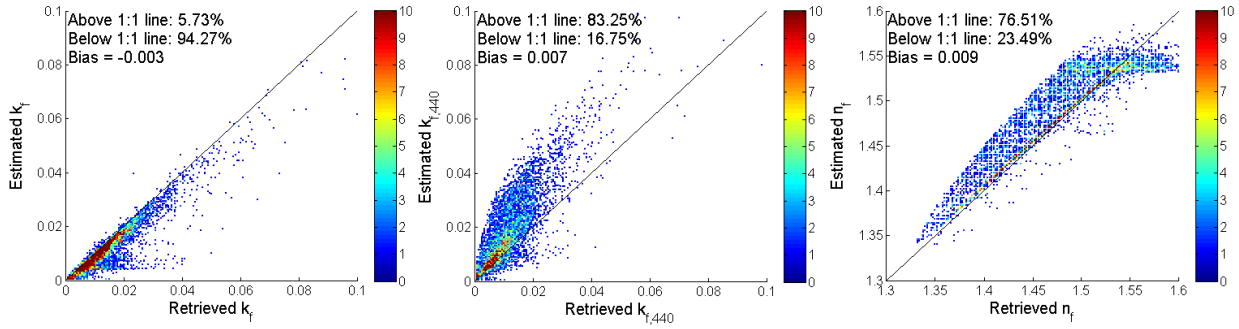

Figure 8. Data pair frequency of instantaneous imaginary parts of the complex refractive index at 675 nm (k_f), 440 nm ($k_{f,440}$), and real part at 440 nm (n_f) which are sorted according to ordered pairs (Retrieved, Estimated) in 0.0005 and 0.001 intervals for imaginary and real parts, respectively. "Retrieved" represents sub-component of CRI from the optical-physical retrievals, and "Estimated" is estimated by retrieved chemical components. The color represents the number of cases (color bar), and the solid black line shows the 1:1 line.

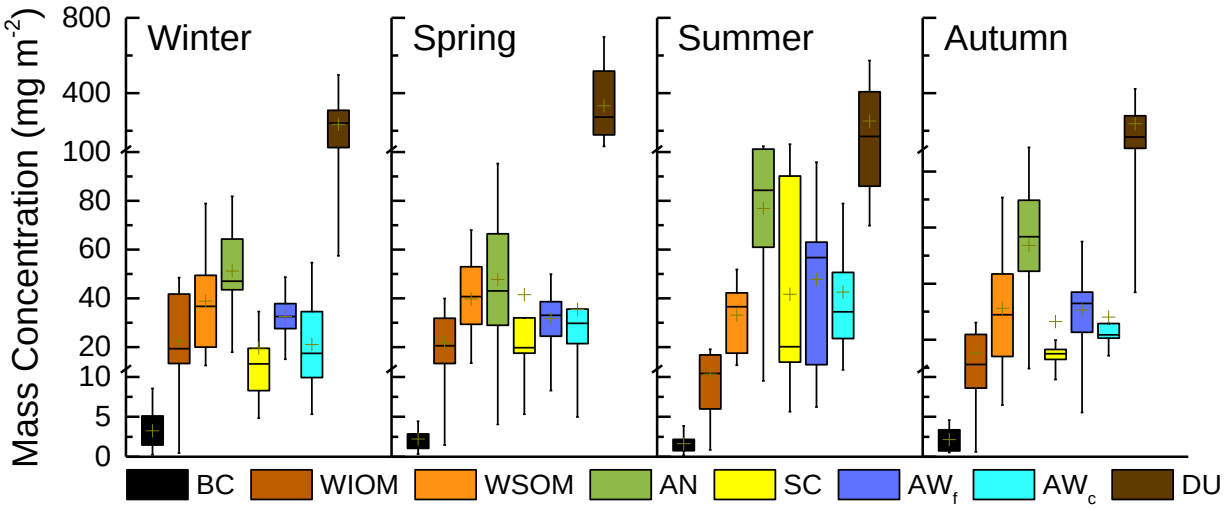

Figure 9. The mass concentrations of aerosol components in four seasons (winter, spring, summer and autumn). For the box-whisker plot, the mean value is indicated by a plus sign (+), and the median value by a short line inside the box (–). The top and bottom edges of each box represent the top and bottom quartiles (Q3 and Q1), and the corresponding whiskers are the outliers (Q3+1.5IQR and Q1-1.5IQR, IQR is interquartile range).

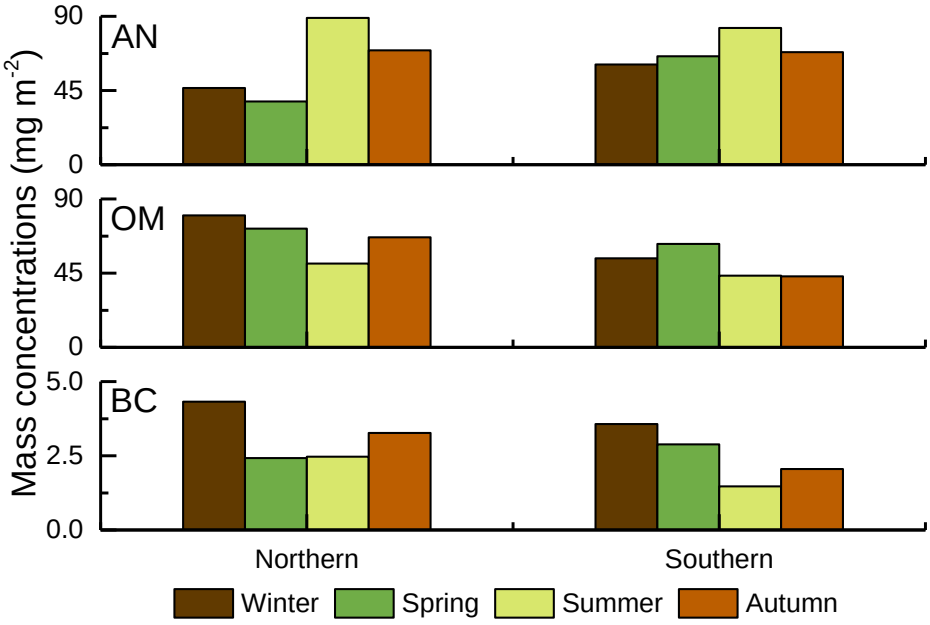

Figure 10. Comparison of aerosol component mass concentrations in northern (Xi'an, Beijing, Harbin, Hefei and Songshan) and southern China (Nanjing, Shanghai, Zhoushan, Guangzhou, Haikou and Sanya).

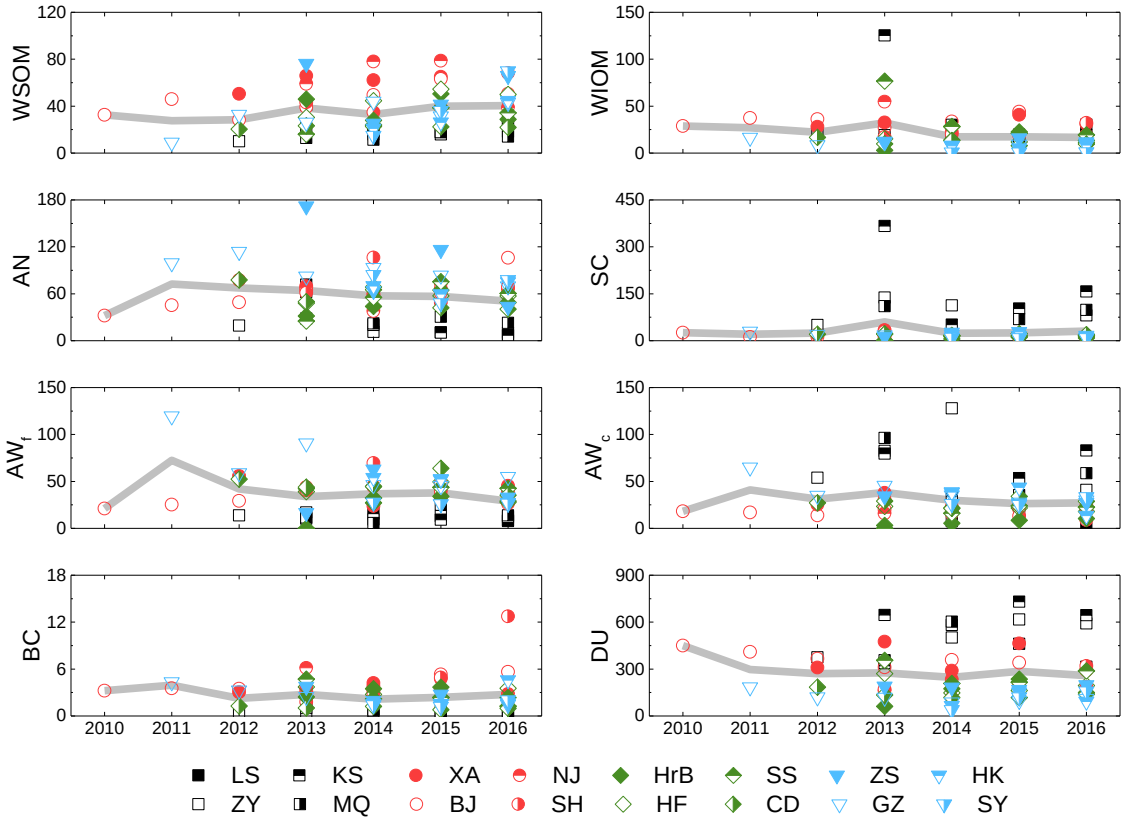

Figure 11. The interannual variations of mean aerosol component mass concentrations integrated over the whole atmospheric column with SONET sites from 2010-2016. The gray line represents the mean mass concentration of aerosol components averaged over the 16 sites; the points in each graph show the yearly value at each site. The abbreviations for the site names are from table 1.