Response to Anonymous Referee #2

General comments:

The authors introduce a new correction method to correct for the truncation error of the Aurora3000 nephelometer. This method uses the Angstrom exponent and also the hemispheric backscattering coefficient and is based on training a random forest machine learning model. To the reviewer's knowledge, the method is new and could be a step forward.

However, the reviewer has major concerns about the description of the model and the presentation and interpretation of results.

Response: Thank you very much for your review of our manuscript. We sincerely appreciate the efforts you have put in the review process and have revised our manuscript according to your comments. Below we will respond to your comments one by one. Your comments are in bold italics, and my responses are in plain text. All the changes have been included in the newest version of our manuscript.

The role of the field measurements in this manuscript is not clear. In the absence of a complete, albeit simple, classification of the field data using SAE and SSA, it is questionable whether the data provide a sufficient basis for initialising the model. For example, it is not clear how strong the light absorption of the simulated aerosols are. Any information on single scattering albedo or the imaginary part of the refractive index are missing.

Response: All the field campaign sites are located in the Northern China Plain and the field data can represent the background aerosol properties there. As shown in Fig.S1, the single scattering albedo (SSA) of eight datasets varies between 0.235 and 0.997, and the scattering Ångström exponent (SAE) also covers a wide range, indicating that the field measurements used in this manuscript is sufficient to initialize the model.

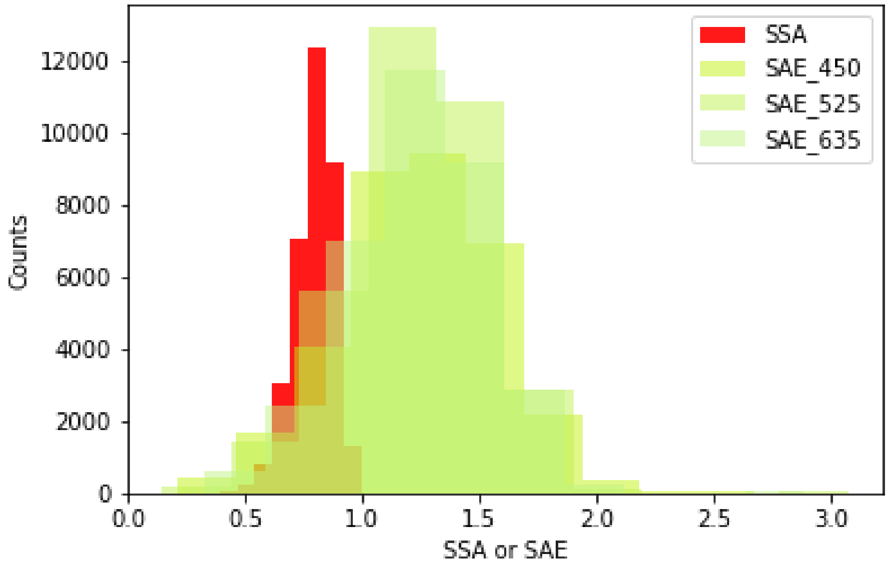

Fig.S1: The probability density distribution of SSA and SAE.

Furthermore, the reviewer finds it difficult to distinguish between measurement (PNSD) and speculative assumptions (refractive indices or kappa) in both models (dry and different RH conditions). For simplicity, the simulation study could also have been carried out with synthetic data for clearly defined aerosol types. The description of the model calculations are often imprecise, as important parameters such as refractive indices used from etc. are not specified.

Response: Both the measurement and speculative assumptions are needed in order to derive parameters to train the model. The reason why we use the in-situ measurements rather than synthetic datasets is that we want to train the machine learning model for the in-situ nephelometer correction. With these in-situ datasets, we can obtain a model that can be better representative of in-situ measurements, hence we can obtain better correction results. We assumed that the aerosols were composed of absorbing black carbon and non-absorbing materials, and their refractive index is set to be 1.80-0.54i (Ma et al., 2012) and 1.53-10^{-7}i (Wex et al., 2002), respectively. Refractive indices and single scattering albedo are addressed in the new manuscript.

Chapter 3 points out the performance of the new model. Why are results of the new model only shown for data from the Gucheng measurement campaign? Why were experimental data not used to show the performance of the new algorithm for dry conditions by a closure experiment of the light scattering coefficient? And more

important, why haven't the authors shown how their approach compares to the simple linear parameterization shown in Mueller et al., (2011)?

Response: When training the machine learning model, we need to choose the training datasets first. The more training datasets we use, the better the trained model can perform. Therefore, we split eight datasets into seven training datasets and one test dataset (Gucheng), and we can only use the test dataset to verify our new model. A closure experiment of the light scattering coefficient has been done and for more details please refer to Fig.S1 in the supplement of Yu et al. (2018). Our revised manuscript has added the comparison with the simple linear parameterization shown in Müller et al. (2011), and it can be seen that this linear regression method is less accurate than our method.

Specific comments:

1. Line 36: What parameter is mentioned?

Response: It is the "scattering correction factor" in the sentence.

2. Line 39: What methods have been proposed in Mueller et al (2011)?

Response: In Müller et al. (2011), the two methods, also mentioned in the introduction of our manuscript, are correction using measured size distributions and correction using scattering Angström exponents. Müller et al. (2011) did not put forward a new correction method and what they focused on is to provide parameterizations for angular sensitivity functions of the Aurora 3000.

3. Line 56: Figure 5 in Mueller et al (2011) suggests that a simple linear function is not sufficient. Unfortunately, this was not discussed further in Mueller et al. (2011).

Response: Yes. Müller et al. (2011) followed the simple linear regression method put forward by Anderson and Ogren (1998). They only utilized one parameter (SAE) to do the regression analysis. In the Sect.3.1 of our revised manuscript, using Gucheng data, we find out that this linear method lacks accuracy. Inspired by their work, the paper uses more relating parameters to predict CF and the verification results show better performance.

4. Chapter 1: In general, the description of the state of the knowledge is little vague. How large are uncertainties when using the simple parameterizations of Anderson

(1998) and Mueller (2011)?

Response: Anderson and Ogren (1998) and Müller (2011) used the same method that is correction using scattering Angström exponents. As for different datasets, the uncertainties could be different. We have used Gucheng data to compare this linear regression method with ours and added this part to Sect. 3.1 in our revised manuscript.

5. Line 70 and Figure 1: Just taking a large set of total number concentrations as an argument that a large number of possible aerosol types have been covered is not sufficient. Furthermore, no evidence of a coarse mode particle can be seen in the particle size distributions. The large range of scattering Angström exponents (see Figure 2) suggest that could be are cases with a significant coarse mode volume fraction.

Response: As seen in Fig.S2 (also added as Fig.2 in the revised manuscript), SSA of eight datasets varies a lot and the measurements can represent the background aerosol properties in the Northern China Plain. Our datasets barely include the coarse mode particles taking the location and date of campaign into account. Due to the limitations of Mie model, our method is not suitable for coarse mode particles (discussed at the end of the paper), too. We have deleted “including most continental aerosol types” to avoid misunderstanding.

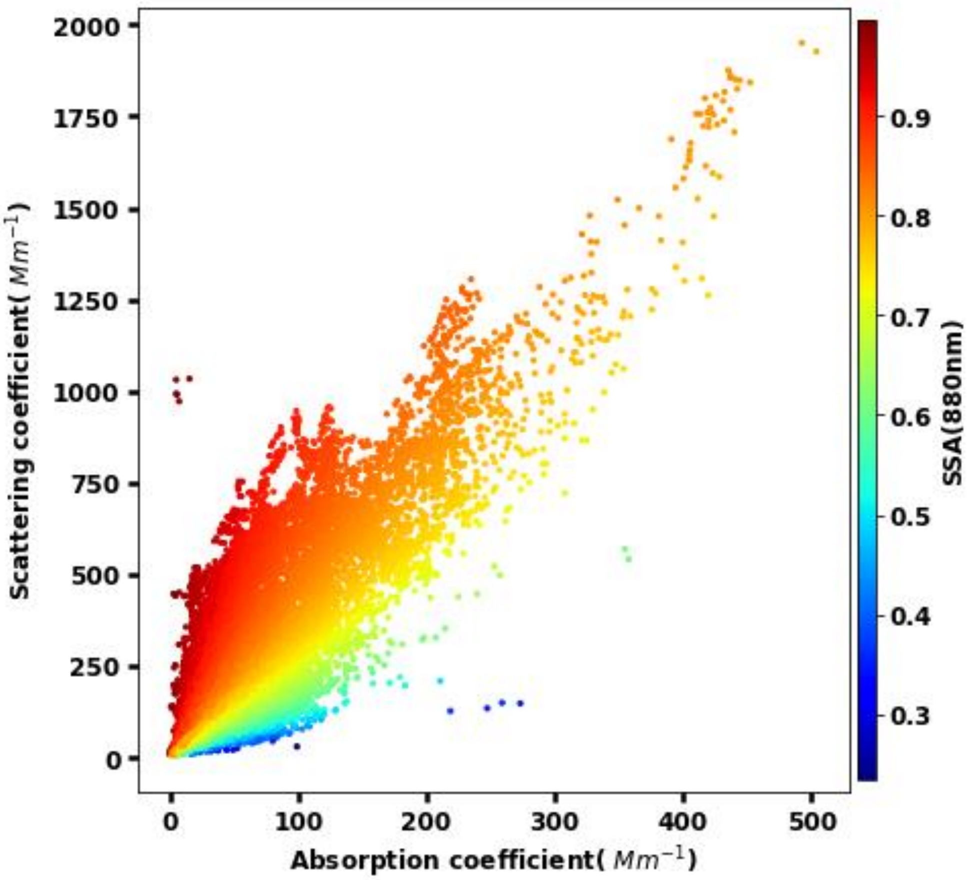

Fig.S2: The SSA of eight datasets.

6. Line 89: A core radius of 35 nm might be too small to represent internally mixed aged particles. Furthermore, a constant core size also means that the volume fraction of absorbing material and the single scattering albedo decreases with increasing particle size. What does this mean for the interpretation of Figure 3? What range of single scattering albedos is covered with this model?

Response: As shown in Fig.S1 and Fig.S2, the single scattering albedo (SSA) of eight datasets varies between 0.235 and 0.997. This paper made the assumption of three independent cases including all scattering particles, all absorbing particles, and all core-shell mixed particles of which core diameter is 70 nm. Ma et al. (2012) pointed out that the freshly emitted LAC particles are assumed to be distributed with geometric average diameter of 50 nm. Therefore, diameter of 70 nm can represent the internally mixed aged aerosols to some extent. Here we also compare the results between core diameter of 70 nm and that of 100 nm (Fig.S3) and their general trend is similar.

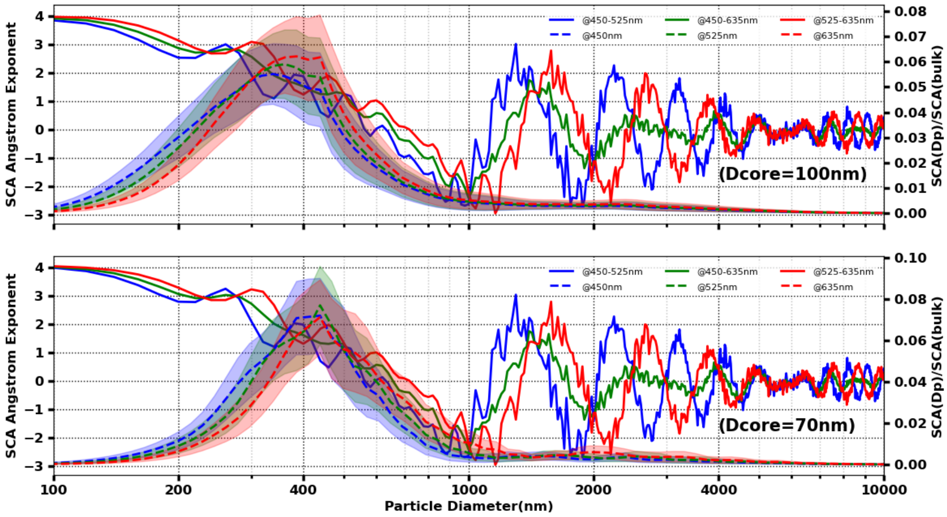

Fig.S3: The SAE change of core-shell mixing particles of core diameter 100 nm and core diameter 70 nm with the change in particle diameter (solid line). The dashed lines represent the ratio of scattering at a certain diameter relative to the total scattering.

It means that the variations of core diameter between 70 and 100 nm do not influence our conclusion, because small particles contribute little to the total scattering. Based on the assumption, the purpose of that figure is to distinguish the particle size range where the change of SAE varies greatly. That is to say, we aim to find out the size range where the variability of SAE is mostly sensitive to the concentration of particles.

7. Line 88: What refractive indices are used for absorbing and scattering materials?

Response: For the absorbing materials, we assume that they are black carbon aerosols and the refractive index is 1.80-0.54i (Ma et al., 2012); for the scattering materials, refractive index is set to be 1.53-10^{-7}i (Wex et al., 2002). The information has been added in the new manuscript.

8. Line 91 and throughout the manuscript: Replace ‘band“ by „wavelengths”.

Response: Thanks for your comment. We have changed all the “band” to “wavelength” in the modified manuscript.

9. Line 92: Mie model: The description of the optical model should include how large the truncation angles were and how the imperfect Lambertian light source was taken into account. Are calculated values shown in Figures 3 and 4 for in ideal nephelometer or simulating the output of Aurora3000?

Response: Thanks for your comment. It is stated in the manuscript that calculated values shown in these two figures are the simulation output of Aurora3000. Brief information about the nephelometer light has been added to the new manuscript. More details come as follows. According to Müller et al. (2011), as for Aurora3000, the angle range of light is limited from $10^{\circ}-171^{\circ}$ and the real angular intensity distribution can be described by :

$$Z_{\mathrm{ts}}(\theta)=\begin{cases}0 & 0^{\circ}<\theta<10^{\circ}\\ \beta_1\times(sin\theta)^{\beta_2} & 10^{\circ}<\theta<171^{\circ}\\ 0 & 171^{\circ}<\theta<180^{\circ}\end{cases}.$$

Where β_1 is 1.01 and β_2 is 1.190.

10. Figures 3 and 4: The reviewer believes that all measured size distributions (referred to as 'bulk' in Figure 3) served as the basis for the calculations. This should be mentioned in the text. Furthermore, it is not clear how the ratio of size resolved to total scattering is calculated. Was the size resolved scattering calculated for a constant size interval on linear scale or constant on logarithmic scale?

Response: Thanks for your comment. The phrase of "all measured size distributions" has been added to the new manuscript. The size resolved scattering calculated for a constant size interval was constant on logarithmic scale.

11. Line 123: Can the authors explain why R ext is sensitive to HBF?

Response: HBF is influenced by the mixing state of BC (Ma et al., 2014). Rext is the ratio of externally mixed black carbon to total black carbon. The change of Rext means that the mixing state of BC changes and thus HBF can change accordingly.

12. Figure captions 3 and 4 : "absorbing particles (b)"

Response: Thanks for your comment. Change made.

13. Line 141: Please specify "Conditions of nephelometer light source"

Response: We have added that "for more details please refer to Müller et al. (2011)". The detail conditions of nephelometer light source come as follows. According to Müller et al. (2011), as for Aurora3000, the angle range of light is limited from $10^{\circ}-171^{\circ}$ and the real angular intensity distribution can be described by :

$$Z_{\mathrm{ts}}(\theta)=\begin{cases}0 & 0^{\circ}<\theta<10^{\circ}\\ \beta_1\times(sin\theta)^{\beta_2} & 10^{\circ}<\theta<171^{\circ}\\ 0 & 171^{\circ}<\theta<180^{\circ}\end{cases}.$$

Where β_1 is 1.01 and β_2 is 1.190.

14. Line 149: "RF predictor" not defined.

Response: Thanks for pointing this out. We have added the definition.

15. Line 150: Why are the model results just checked for data from Gucheng and not for the other stations?

Response: Because other campaign data are used to train the random forest machine learning model and we cannot use the training data to test the derived model. The more training datasets we use, the better the trained model can perform. Therefore, we split eight datasets into seven training datasets and one test dataset (Gucheng).

16. Line 170: Specify "assumed size distributions of kappa"

Response: Thanks for your comment. This part has been specified in the new manuscript.

17. Figure 6: y-axis, "CRH" should be C(RH)?

Response: Yes, it is. We have changed it to C(RH).

18. Figure 6: Do not split legend to subplot (a) and (b)

Response: Thanks for pointing this out. Revised as suggested.

19. Figure 6: How can f(RH) and Fb(RH) be negative for low RH?

Response: At low RH, particles have not taken up water and theoretically f(RH) and f_b(RH) equal 1. In practice, there may be some small measurement errors of nephelometer. Therefore, the scattering coefficient under the condition of high humidity could be slightly lower than that under the dry condition. Although the two values are negative for low RH, their values approach 1.

20. Lines 230: The reviewer can not follow the conclusion on the strength of the absorption. The authors missed to give any information on the strength of absorption like single scattering albedo or complex refractive index.

Response: The refractive indices and single scattering albedo of eight datasets have been added in the revised manuscript. The single scattering albedo varies from 0.235

to 0.997 and our machine learning method can still be accurate whether including absorbing aerosol or not. However, Bond et al. (2009) pointed out that the linear regression method of Anderson and Ogren (1998) is in error by 1-5% for absorbing particles.

21. Line 310: The reviewer thinks it would be better to reword the paragraph, since the study could also be done with synthetic datasets. With synthetic data, also simulations of e.g. desert and marine aerosol types could be done.

Response: As for our method, we apply the Mie model to simulate the nephelometer's measurements, then the results are used to train the random forest machine learning model. Desert and marine aerosol particles are non-spherical rather than spherical, and hence the limitations of Mie model restrict the application of our new method to these aerosol types.

Reference:

Anderson, T. L., and Ogren, J. A.: Determining aerosol radiative properties using the TSI 3563 integrating nephelometer, Aerosol Sci. Tech., 29, 57-69, doi:10.1080/02786829808965551, 1998.

Bond, T. C., Covert, D. S., and Müller, T.: Truncation and angular-scattering corrections for absorbing aerosol in the TSI 3563 nephelometer, Aerosol Sci. Tech., 43, 866-871, doi:10.1080/02786820902998373, 2009.

Ma, N., Birmili, W., Mueller, T., Tuch, T., Cheng, Y. F., Xu, W. Y., Zhao, C. S., Wiedensohler, A.: Tropospheric aerosol scattering and absorption over central Europe: a closure study for the dry particle state, Atmos. Chem. Phys., 14, 6241-6259, doi: 10.5194/acp-14-6241-2014, 2014.

Ma, N., Zhao, C. S., Müller, T., Cheng, Y. F., Liu, P. F., Deng, Z. Z., Xu, W. Y., Ran, L., Nekat, B., van Pinxteren, D., Gnauk, T., Müller, K., Herrmann, H., Yan, P., Zhou, X. J., and Wiedensohler, A.: A new method to determine the mixing state of light absorbing carbonaceous using the measured aerosol optical properties and number size distributions, Atmos. Chem. Phys., 12, 2381-2397, doi:10.5194/acp-12-2381-2012, 2012.

Müller, T., Laborde, M., Kassell, G., and Wiedensohler, A.: Design and performance of a three-wavelength LED-based total scatter and backscatter integrating nephelometer, Atmos. Meas. Tech., 4, 1291-1303, doi:10.5194/amt-4-1291-2011, 2011.

Wex, H., Neususs, C., Wendisch, M., Stratmann, F., Koziar, C., Keil, A., Wiedensohler, A., and Ebert, M.: Particle scattering, backscattering, and absorption coefficients: An in situ closure and sensitivity study, J. Geophys. Res.-Atmos., 107, 8122, doi:10.1029/2000jd000234, 2002.

Yu, Y.L., Zhao, C.S., Kuang, Y., Tao, J. C., Zhao, G., Shen, C. Y. and Xu, W. Y.: A parameterization for the light scattering enhancement factor with aerosol chemical compositions, Atmos. Environ., 191, 370-377, doi:10.1016/j.atmosenv.2018.08.016, 2018.