Response to reviewer#1

Thanks for the reviewer's helpful suggestions! The comments are addressed point-by-point and responses are listed below.

Comments: Summary: In this paper, the authors use measurements of dry and humidified aerosol properties to validate a new machine learning algorithm for determining the asymmetry parameter based on routinely measured properties. The asymmetry parameter (g) is important in some radiative forcing models that are based on the Henyey-Greenstein approximation of aerosol scattering phase function, a function only of g. The algorithm for determining g using dry and humidified nephelometer measurements is validated with the CCD-LADS, a new instrument that directly measures the aerosol scattering phase function. An attempt is then made to connect these results with radiative forcing models.

Reply: We thank the anonymous reviewers' comments and suggestions.

Comments: General comments: It would be useful to have more information about the different aerosol properties at the three different ground sites. For example, how do the PNSD and BC fraction vary between the different sites, as well as seasonally and diurnally? In particular, examples of the PNSD (dry and ambient) averaged for the different sites would be a useful figure to have since particle size is such an important factor in determining asymmetry

Reply: Thanks for the comments. We give the characteristics of the measured mean PNSD in section 5 of the supplementary and in section 4.1.1 of the main manuscript.

Fig. S6 gives the measured distribution of the aerosol particle number size distrubion (PNSD) at Gucheng and PKU. The red line gives mean results of aerosol PNSD and the red line gives the median aeorosl PNSD. At the same time, the aerosol PNSD at the ambient relative humidity are calculated by using the measured aerosol PNSD, the ambient RH and the aerosol hygroscopic growth factor κ, which is derived from the datasets of the humidified nephlometer (Kuang et al., 2017). At the same time, we calculated the aerosol scattering coefficient distibution under the dry

condition by using the Mie scattering theory and the measured aerosol PNSD. There are totoal 8613 and 5298 different aerosol PNSDs are measured at Gucheng and PKU respectively.

From fig. S6, the peak diameter of the mean and median PNSD at Gucheng locates around 150nm. However, the peak diameter of the mean and median PNSD at PKU locates at around 100nm. At the same time, there are large partitions of small particles that are lower than 60nm at PKU. However, these particles, which are lower than 100nm, contribute little to the total aerosol scattering. The aerosol PNSD at Gucheng is more dispersed than that of the PKU, which corresponds to a larger variation in the g values. When compared with the dry aerosol PNSD, the calculated ambient aerosol PNSD at Gucheng varies more than that of the PKU, which is resulted from the high relative humidity of the ambient condition.

The size distribution of the aerosol scatter coefficient at around 500nm contributes more to the scatter coefficient at Gucheng than at that of the PKU. Thus these particles with the diameter larger than 500nm contribute less to the aerosol scattering coefficient. As g increases with the aerosol diameter, the aerosol g under dry conditions at Gucheng tends to be smaller than that at PKU.

We are willing to show the measurement results of the aerosol PNSD and mass of the BC concentrations to the reviewers. Fig.R1 and fig.R2 give the measurement results of the Gucheng and PKU respectively.

However, we are preparing the characteristics of the measurement of aerosol PNSD and BC fraction in other works. We decide not to give the detail discussion about these measurements results in the main manuscript, which is not the main scope of this work.

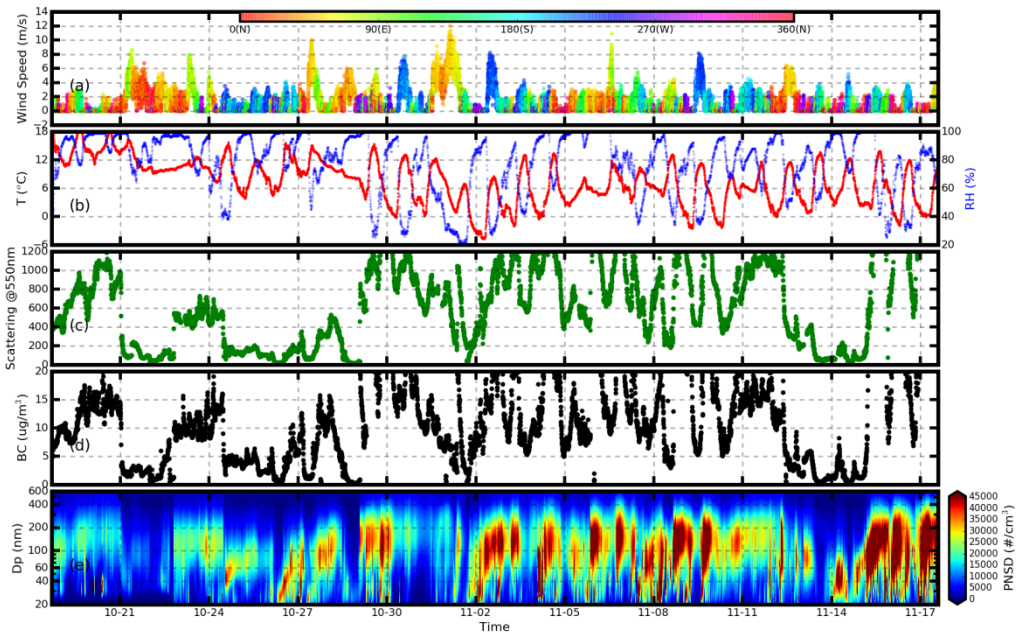

Fig. R1. The measurement data at the station of Gucheng. The time series in the figure represents (a) the wind speed and wind direction, (b) the temperature (in red), the relative humidity (in blue), (c) the scattering coefficient at 550 nm, (d) the BC mass concentrations and (e) the aerosol PNSD.

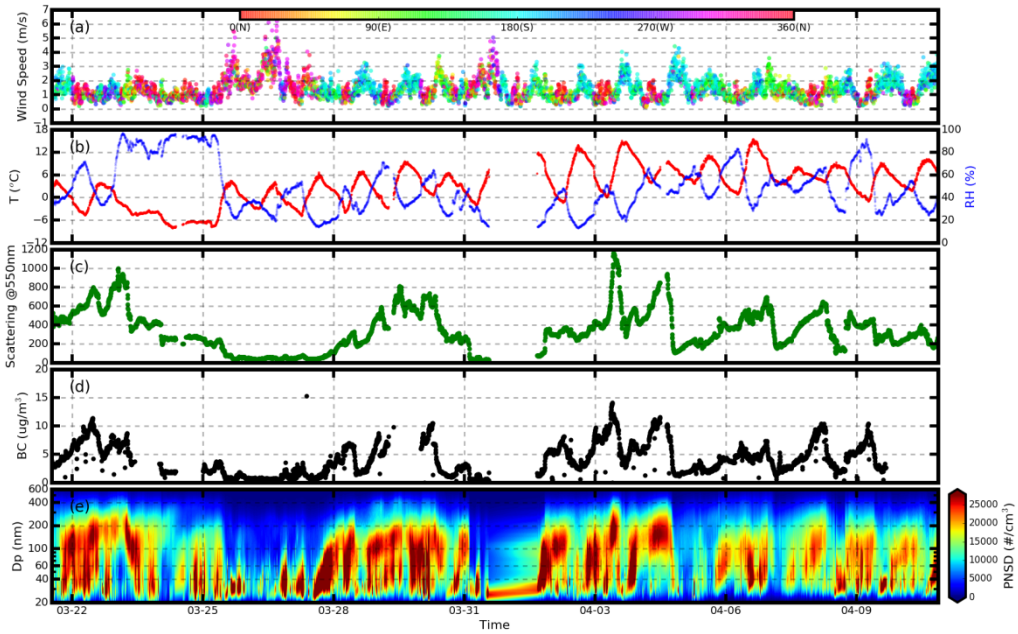

Fig. R2. The measurement data at the station of PKU. The time series in the figure represents (a) the wind speed and wind direction, (b) the temperature (in red), the relative humidity (in blue), (c) the scattering coefficient at 550 nm, (d) the BC mass concentrations and (e) the aerosol PNSD.

Comments: The weakest part of the study is the machine learning algorithm section. The use of the random forest algorithm is not well-justified, nor is there sufficient discussion of the parameterizations used.

Reply: Thanks for the comments. We have made some revisions according to the suggestion. Some discussions were added in the introduction in the manuscript. At the same time, section 2 in the supplementary material is added and gives some detail discussion of using the machine learning algorithm.

Comments: I find the label "calculated phase function" and "measured phase function", used extensively in the discussion and throughout the figures, confusing. Please clarify whether referring to theoretical phase functions determined using the PNSD, Mie theory, and assumed refractive index, or phase functions directly measured by CCD-LADS and be consistent with how these are referred to.

Reply: Thanks for the comments. We have made some revisions according to the suggestion. Most of the revisions are done in section 4.3.1. At the same time, we also made some revisions about the g values by using different method. The calculated g value by using the Mie theory are labeled as g_{Mie}. The value predicted by using the random forest machine learning model is labeled as g_{ML}. The calculated value by the phase function measured from the CCD-LADS are labeled as g_{CCD}.

Comments: I think it would strengthen the discussion to include estimations of the uncertainties of the various measured properties. These can then be used for sensitivity analysis instead of arbitrary deviations.

Reply: Thanks for the comments. We think the reviewer gives a good view of conducting the uncertainties test. The method of estimating the uncertainties is reconstructed following the reviewer's opinion. The uncertainties from the measured properties are analyzed and the corresponding uncertainties are estimated by using the Monte Carlo simulations. The results are mainly discussed in section 4.2.2 of the manuscript.

Comments: The authors should be careful about the input parameters of the random forest algorithm, particularly for their sensitivity analysis. The measured quantities σ_{sca} and β_{sca} for a specific wavelength are related to each other, and therefore I think it is important to question whether they can be treated as independent input variables.

Reply: Thanks for the comments. We agree with the reviewer's idea and reconstructed the method of conducting the sensitivity analysis. Some discussions were added in section 2 of the supplementary material to determine the input parameters.

Comments: Please ensure that axes labels on figures legible (text size difficult to read on e.g. Figure 7)

Reply: Thanks for the comments. We have made some revisions according to the suggestion.

Comments: Specific comments: P1 L15-16: Does this range only apply to dry aerosol?

Reply: The reported g values apply to the dry aerosol. We made some revisions at the corresponding line. For the g values under different RH values, it can have a wider range as shown in the fig. 2 of the manuscript.

Comments: P1 L16-17: Specify what types of aerosol this enhancement applies to

Reply: The enhancement value 1.2 of the g value applies to the continental aerosol. We made some revisions at the corresponding line.

Comments: P1 L25: Reference desired for this information.

Reply: Thanks for the comment. We have made the revision according to the comments.

Comments: P1 L26-28: HG reference

Reply: Thanks for the comment. We have made the revision according to the comments.

Comments: P1 L27: "The HG phase function ($P_{HG}(\theta)$) is..."

Reply: Thanks for the comment. We have made the revision according to the comments.

Comments: P2 L34: "$P(\theta)$ is the normalized scattering phase function"

Reply: Thanks for the comment. We have made the revision according to the comments.

Comments: P2 L37: More references would be useful here.

Reply: Thanks for the comment. We have made the revision according to the comments.

Comments: P2 L39: "few studies" – which ones?

Reply: Thanks for the comments. We have made the revision according to the comment. Pandey and Chakrabarty (2016) measured the phase function of fractal black carbon aerosols and compared the difference of the direct aerosol radiative forcing between using the measured phase function and the HG phase function. Boucher (1998) and Wiscombe and Grams (1976) simulated the difference by simulating the aerosol phase function using the Mie scatter theory. However, there is no study, to our knowledge, that uses the filed measurement of the aerosol optical properties to estimate the difference.

Comments: P2 L49: What is meant by "no available method to measure g directly" – surely a measurement of $P(\theta)$ is a fairly direct measurement technique? (i.e. Bian, Dolgos instruments)

Reply: We agree with the editor's suggestion and make the revision according to the comments. The corresponding text is changed as 'There are many methods to derive

the aerosol g for the dry and ambient condition.'

Comments: P2 L55: Define backward hemisphere angles

Reply: The hemispheric backscatter ratio b is the fraction of the scattered intensity that is redirected into the backward hemisphere of the scattering particle (Marshall et al., 1995), with the definition of

$$b = \frac{\int_{\frac{\pi}{2}}^{\pi} P(\theta)\cdot sin\theta\cdot d\theta}{\int_{0}^{\pi} P(\theta)\cdot sin\theta\cdot d\theta}. \qquad (1).$$

The main advantage of the backscatter ratio is that it can be measured with an integrating nephelometer equipped with a backscatter shutter (Charlson et al., 1974). We also made the revision at the corresponding line.

Comments: P2 L56: It would be appropriate perhaps to also cite Horvath et al, 2016 (J. Aerosol Sci.)

Reply: Thanks for the comment. We have made the revision according to the comments.

Comments: P3 L57-63: How were these values determined – observations (if so, which instruments used?) or models?

Reply: In the previous study, almost all of the aerosol asymmetry factor values are determined by using the Mie scattering theory. We have made the revision according to the comments at the beginning of paragraph 3.

Comments: P3 L64-66: Can you discuss difficulties measuring g under ambient RH vs dry conditions?

Reply: Thanks for the comments. We give some of the difficulties of measuring g under ambient RH vs dry conditions below.

Aerosol g is derived by measuring the aerosol phase function. However, it is very difficult to measure the aerosol phase function accurately.

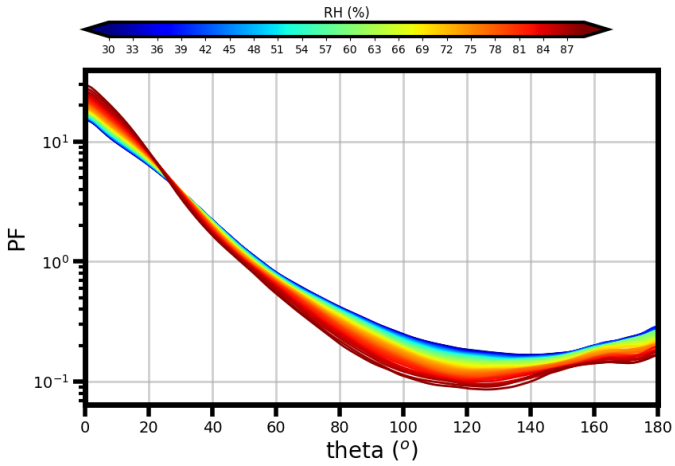

Figure R3. The aerosol phase functions under different RH conditions. Different colors represent the different RH.

Fig. R3 gives one example of the calculated phase function under different relative humidity by using the Mie scattering theory and the measured mean aerosol PNSD at the measurement site of PKU. Detail of calculating different phase function under different relative humidity can refer to the main manuscript and Zhao et al. (2017). From fig. R3, the aerosol phase function under dry condition can change from 0.2 to 12 at different angle. For the calculated phase function at relative humidity of 90%, the phase function can change from 0.1 to 30. For some other studies, the aerosol phase function can be as low as 0.002 (Horvath et al., 2016). As the aerosol phase function can change significantly at different angle, measurement of aerosol

phase function has very high requirement of the instrument. Up to now, there are few commercial instruments that can measure the aerosol phase function such as nephelometer aurora 4000 (Chamberlain-Ward and Sharp, 2011). However, these instruments are not widely accepted and used. Measurement of the phase function is very difficult.

At the same time, there is no method to measure the aerosol phase function covering all the angles from 0 to 180 degree. The phase function changes significantly at the range of 0~10 degree and 170~180 degree. Many of the studies derive the aerosol phase function by interpolation.

As for measuring the phase function of the ambient RH, the most important requirement is that the RH of the aerosol should not be changed when conducting the measurement. The method developed by Bian et al. (2017) can be used to measure the ambient phase function.

At the same time, the aerosol phase function change significantly with the aerosol morphology, aerosol size, aerosol complex refractive index and the incident light wavelength. Knowledge of the aerosol phase function is very limited. The calculation of the ambient aerosol phase function by using the Mie scattering theory requires more information such as the aerosol PNSD and aerosol hygroscopic growth properties.

Comments: P3 L64-66: It would be appropriate to cite Andrews et al. again here

Reply: Thanks for the comments. We made the revision according to the comments.

Comments: P3 L68: Particles do not absorb water –they can take up water, or water can condense on them

Reply: Thanks for the comments. We made the revision according to the comments.

Comments: P3 L73-76: Define Mwater

Reply: Thanks for the comments. We made the revision according to the comments.

M_{water} is the molecular weight of water.

Comments: P3 L79-80: Should specify that Mie theory only applies to spherical particles – g can also vary by morphology

Reply: Thanks for the comments. We made the revision according to the comments.

Comments: P3 L84: Can you use subscript for f_{RH} to reduce confusion (instead of fRH)?

Reply: Thanks for the comments. We have changed all of the fRH to f_{RH} in the manuscript.

Comments: P4 L87: Define back-scattering coefficient.

Reply: Thanks for the comments. We give the definition of the back-scattering coefficient below. However, back-scattering coefficient is widely used for those who are familiar with the instrument of the nephelometer. We decide to not add the definition of the back-scattering coefficient to the main manuscript.

The back-scattering coefficient is defined as

$$\beta_{sca} = \int_{90^\circ}^{180^\circ} F_\lambda(\theta) d\theta \qquad (2),$$

where the angular scattering function of the particle population $F_\lambda(\theta)$ is defined as

$$\mathrm{F}_\lambda(\theta) = \int f_\lambda(\theta, m_\lambda, d_p) \cdot \frac{dn(d_p)}{dlog(d_p)} \cdot \pi \cdot \left(\frac{d_p}{2}\right)^2 \cdot dlog(d_p) \qquad (3),$$

where $f_\lambda(\theta, m_\lambda, d_p)$ is the angular scattering function of the individual particles with complex refractive m_λ and diameter d_p and $\frac{dn(d_p)}{dlog(d_p)}$ is the aerosol particle number size distribution. More details of the definition of the back-scattering coefficient can refer to Müller et al. (2011).

Comments: P4 L106-110: Please provide seasonal information about these field campaigns, as well as whether they covered the full diurnal cycle.

Reply: Thanks for the comments. We have made the revision according to the comments at paragraph 4. As shown in fig. R1, theses datasets covered the full diurnal

cycle.

Comments: P4 L110-111: For people who are unfamiliar with Beijing, please describe the location in terms of either distance from the centre or relative population density – some more general metric

Reply: Thanks for the comments. We have made the revision according to the comments at paragraph 4 and section 1 in the supplementary material. PKU station is 11km from the center of the Megacity Beijing, which is adjacent to Hebei Province and the megacity Tianjin. In the above three cities, the industrial manufacturing has led to heavy air pollution.

Comments: P4 L110-P5 L116: Can you also describe the general wind patterns (i.e. are the sites downwind or upwind of the urban centres?)

Reply: Thanks for the comments. We have made the revision according to the comments at paragraph 4 and section 1 in the supplementary material. At the same time, the wind patterns can be recognized by fig.R1, fig.R2, and fig.S1.

Comments: P5 L118-119: Type of impactor? Is 10 μm the 50% cut-point?

Reply: Thanks for the comments. We have added the information about the impactor. 10 μm is the 50% cut-point of the ambient aerosol particles.

Comments: P5 L119: How is RH measured?

Reply: Thanks for the comments. At the Gucheng site, the RH of the sample aerosols is measured by the temperature and humidity probe, which is installed after the Nafion drying tube and the RH is all the way lower than 30%. For the RH at PKU, it is measured at the beginning of the field measurement using the same way as Gucheng, and the RH is lower than 30%. One day later, the probe was removed. We believe that the RH at PKU is all the way lower than 30% because the Nafion drying tube works well during the filed measurement and the RH at PKU is significantly lower than that of Gucheng.

Comments: P5 L121-122: Are particles assumed to be spherical for SMPS and APS retrievals of size?

Reply: The SMPS measure the aerosol PNSD by scanning the electrical mobility. The particles are assumed to be spherical when relating the aerosol electrical mobility with the aerosol electrical mobility diameter. The APS 3321 is a time-of-flight spectrometer that measures the velocity of particles in an accelerating air flow through a nozzle. It measures the aerodynamic size. The retrievals of the size of the APS do not require assumption of the particle to be spherical. However, when relating the aerodynamic size and the electrical mobility diameter, the particles are assumed to be spherical.

Comments: P5 L119: How low was RH? How sensitive are results to variations in RH at these low levels (presumably it is not sensitive due to low gf in this region but would be useful to say this explicitly)?

Reply: The RH may change when the RH and temperature of the ambient aerosol changes. However, we just need to confirm that the measured sample aerosols are lower than 30% (In some fields measurement when the ambient RH is high and the temperature is high, the RH of the sample aerosol may be acceptable to be dried to below 40%). When the RH of the sample RH is lower than 30%, the aerosol physical properties change little because the growth factor is very low. We assume that all of the measurement results measured by the instruments are at dry condition when the RH of the sample aerosol is lower than 30%, which is widely accepted when conducting the measurements.

Comments： P5 L119-132: Please provide error/uncertainty margins on key instrumentation (i.e. RH ±1%?)

Reply: Thanks for the comments. We have made the revision according to the comments. The uncertainties of the key input parameters are given in section 4.2.2 when doing the sensitivity test.

Comments： P5 L129: Provide laser wavelength and polarization characteristics (circular or linear)

Reply: Thanks for the comments. We have made the revision according to the comments. The wavelength of the laser is 532nm and a quarter-wave plate was mounted in front of the laser emitter to change the polarization state of the laser from linear to circular.

Comments: P5 L126: Please be consistent about f(RH) vs fRH

Reply: Thanks for the comments. We have made the revision according to the comments. We have changed all of the f(RH) and fRH to f_{RH}

Comments: P5 L127: This is a very long measurement period – can you ignore changes in aerosol population over this timescale?

Reply: Thanks for the comments. We agree with the editor's opinion that the aerosol population over 45 minutes may changes and can't be ignored. We made some quality control when retrieving the aerosol hygroscopic growth factor κ. The aerosol scattering coefficient and backscattering coefficient are recorded every minute. The scattering coefficient can catch the variation of the aerosol population. The hygroscopic growth factor κ are retrieved when the maximum and the minimum value of the measured scattering coefficient are within the ranges of 1.4 and 0.6 times of the mean measured scattering coefficient. For most of the cases, the scattering coefficients do not have large variations. Fig. R4 gives one example of the measured scattering coefficient at dry and a given relative humidity conditions. We can see that the scattering coefficient is relative stable during the period. Fig. R5 gives part of the measured time series of κ. Form fig. R5, the retrieved κ values can reflect the variation of the aerosol hygroscopicity.

At the same time, some revisions are made accordingly at paragraph 5.

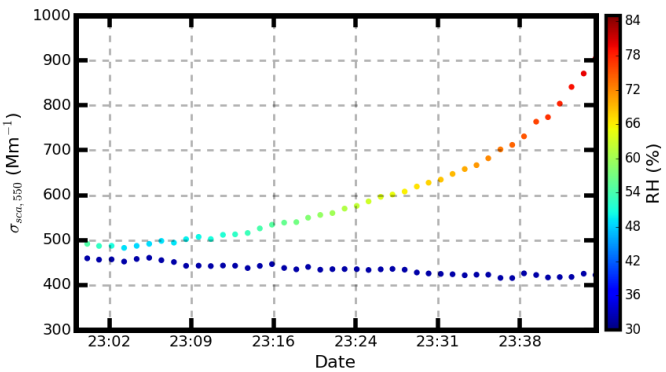

Figure R4. The measured scatter coefficient time series at dry condition and the given relative conditions.

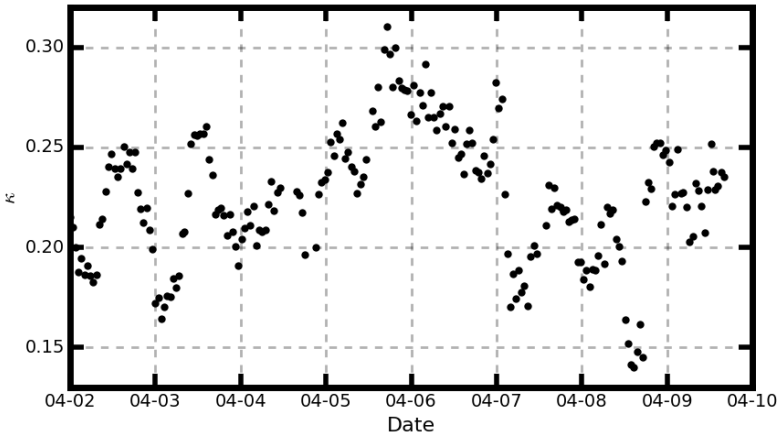

Figure R5. The time series of the measured κ during the field measurement at PKU.

Comments: P5 L128: What is time resolution of CCD-LADS?

Reply: Thanks for the comments. We have made the revision at paragraph 5. The time resolution of the CCD-LADS during the UCAS field measurement was set to 5 minutes.

Comments: P5 L136: Please clarify BC mixing state and BC fraction determination earlier.

Reply: Thanks for the comments. We have made the revision accordingly at section 3.1.

Comments: P5 L127: Need more details about "self-constructed humidified nephelometer system" – what range/steps of RH? Is the population sufficiently consistent? Is 45 min for range of RH or single RH set point? How was RH altered? How was RH monitored? Is there another reference with more details about this instrument?

Reply: Thanks for the comments. Some revisions are made at paragraph 5. Some revisions are made in the manuscript accordingly. The detail information of the humidified nephelometer is described elsewhere (Kuang et al., 2017) and we give some brief description here. The humidifier is used to control the RH of the sample

aerosol and σ_{sca} is measured for each of the controlled RH. The sample aerosol is humidified through a Gore-Tex tube, surrounded by a circulating water layer in a stainless steel tube. The RH is changed by changing the temperature of the circulating water, which is controlled by the water bath and software. For each cycle, the RH points are set to range from about 50% to about 90% over 45 minutes. For most of the cases, the aerosol PNSDs are consistent over the cycle. These cycles of f_{RH} values are abandoned when the measured maximum and the minimum σ_{sca} value are beyond the range of 1.4 and 0.6 times of the mean measured scattering coefficient of each cycle. Two combined RH and temperature sensors (Vaisala HMP110; accuracy of 0.2 and 1.7% for RH ranges from 0 to 90 %, respectively, and accuracy of 2.5% for RH ranges from 90 to 100% according to the manufacturer) are placed at the inlet and outlet of the wet Neph, and the measured RHs and temperatures are defined as RH1/T1 and RH2/T2, respectively. The dew points at the inlet and outlet of wet nephelometer were calculated using the measured RH1/T1 and RH2/T2, and the average value was considered as the dew point of the sample air. The sampled RH can be calculated through the derived dew point and the sample temperature, which is measured by the sensor inside the sample cavity of the nephelometer.

Comments: P6 L148: Why use volume mixing ratio vs, for example, Maxwell-Burnett or mass?

Reply: Thanks for the comments. The method of deriving the refractive index by using the volume mixture ratio is widely used in the previous work (Kuang et al., 2017; Wex et al., 2002; Zhao et al., 2017) and justified by Wex et al. (2002). However, we didn't measure the complex refractive index during the field measurement and can't decide whether the method of using the Maxwell-Burnett or mass mixture ratio is better for our study.

Comments: P6 L150: If using a core-shell model with BC in the centre, do you assume water only mixes with shell, or with both core and shell?

Reply: Thanks for the comments. In our work, we assume the BC to be

non-hygroscopic and the water mixes only with the shell. We make some revision accordingly at section 3.1.

Comments: P6 L147: How is the corresponding ambient aerosol PNSD at given RH computed?

Reply: Thanks for the comments. The ambient aerosol PNSD at given RH were computed by using the measured dry aerosol PNSD, aerosol hygroscopic factor κ, derived from the dataset of humidified nephelometer, and formula (4) in the manuscript. For each aerosol, the growth factor can be calculated by using κ and formula (4). Then the corresponding aerosol diameter of the dry aerosol diameter can be calculated and then the ambient aerosol PNSD at given RH were computed. Fig. R6 gives one example of the calculated aerosol PNSD at different RH.

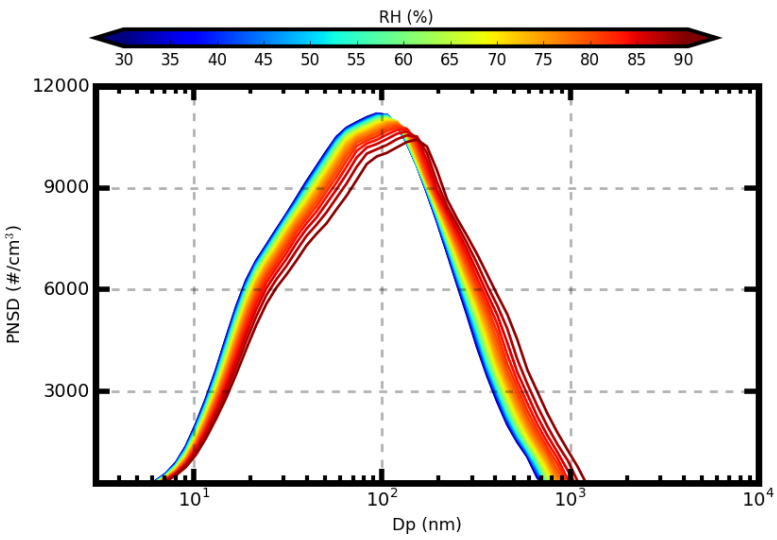

Figure R6. The changes of aerosol PNSD with RH.

We have made some revisions accordingly in the main manuscript.

Comments: P6 L157-158: Can you provide references going into more detail about random forest model algorithms?

Reply: Thanks for the comments. We made the revision according to the comments. In the manuscript, we added the reference that can be used to know more detail about random forest model algorithms such as (Breiman, 2001; Huttunen et al., 2016;

Pedregosa, 2011). However, the random forest machine learning model is a new technique that can be used for the regression. There are few studies that use the random forest machine learning model.

Comments: P6 L158-162: Why is the random forest model appropriate for this specific example?

Reply: Thanks for the comments. We have made some revision at section 3.2 and the introduction section. Random forest machine learning model is a powerful technique that can be used for classification and non-linear regression (Breiman, 2001; Hu et al., 2017; Huttunen et al., 2016). This model is a widely used nonparametric machine learning algorithm that has several strengths. First, it involves fewer assumptions regarding the dependence between observations and outcomes when compared with traditional parametric regression models. Second, strict relationships among variables are not needed before implementing the random forest model. The relationship between the aerosol optical properties measured by humidified nephelometer and g is highly non-linear. Thus, the random forest machine learning model is used in this work to study the calculation of g based on the datasets of the humidified nephelometer system.

Comments: P6 L164: How was the number of trees determined? What is the sensitivity of the results to npre and ntree?

Reply: This comment is replied together with the next comments. Please refer to the next reply of the next comments.

Comments: P6 L163-168: Can you provide more justification for your chosen parameterization? Does the connection between measured scatter and backscatter coefficients affect the suitability of the algorithm, or the accuracy of the results?

Reply: Thanks for the comments. We ignore that there are many factors that may be influenced by the number of trees. For different number of trees, the accuracy of the random forest machine learning model and the time for running the model would

change. For different numbers of n_{pre}, the accuracy and the time for running the model would change too.

The influence of the n_{tree} on the accuracy of the model and the times for training the model is studied to determine the best n_{tree} for the model. The input data of the model come from the dataset of Gucheng and the test data comes from the dataset of PKU. Difreent values of n_{tree} are used in the model and then the time for running the model the accuracy of the model is compared. The accuracy of the model is estimated by comparing the calculated g values (g_{Mie}) and the predicted g values by the random forest machine learning model (g_{ML}) of PKU. The correlation coefficient (R^2) between g_{Mie} and g_{ML}, the mean relative differences between g_{Mie} and g_{ML}, and the standard deviation (Std) of the relative differences between g_{Mie} and g_{ML} are studied. The results are shown in fig. R7. From fig. R7, we can see that the R^2 increases with the increment of n_{tree} from 0.934 to 0.95 when n_{tree} is lower than 32 and changes slightly when n_{tree} is larger than 32, which means that the random forest machine learning model can increase the accuracy of predicting g with the increment of n_{tree} when is lower than 32. At the same, the mean relative difference between g_{Mie} and g_{ML} is all the time lower than 1%, which means that the g can be unbiased estimated by the random forest machine learning model. The Std fluctuates between 1% and 1.2% which is also a very small value. The time of training the model increases with the increment of n_{tree} form 0.2s to 1.2s when the n_{tree} increase from 2 to 100. However, the time for predicting the data is slightly changed with the increment of n_{tree}. Therefore, the time of running the model is not a main concerning when choosing the $n_{tree.}$ With the discussion above, the n_{tree} is chosen to be 32 with the accuracy and time of running the model taken into consideration.

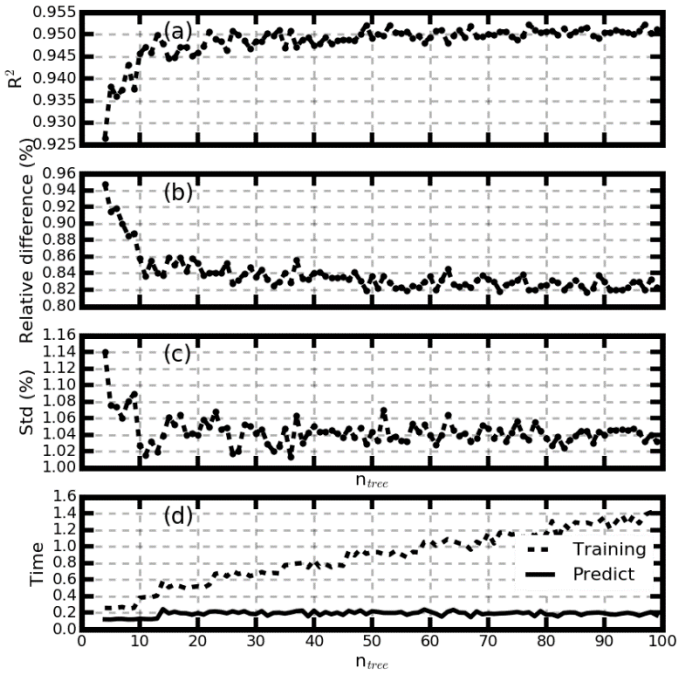

Figure R7. The comparison of g_{ML} and $g_{Mie,dry}$ under different values of n_{tree}. The panels give (a) the correlation coefficient, (b) the mean value of the difference, (c) the standard deviation of the difference valued between g_{ML} and $g_{Mie,dry}$ and (d) the times for running the machine learning model.

Different input parameters can result in different behavior of the machine learning model. The number of n_{pre} is also changed to test the performance of the model by changing the input parameters. From section 3.1.2 of the manuscript, it is obviously that the κ and RH is necessary when predicting the g. We discussed the accuracy of the model by using different combination of the scattering coefficient and back-scattering coefficient as the input of the random forest machine learning model. There are total six group of tests were carried out. These tests contains (1) all of the three σ_{sca} and three β_{sca}; (2) two σ_{sca} and two β_{sca}; (3) one σ_{sca} and two β_{sca}; (4) three σ_{sca}; (5) three β_{sca} and (6) single parameters of the σ_{sca} or β_{sca}. The details of the tests and the results are shown in Table. R1. These results include the R^2, the mean relative differences and the Std of the relative differences between g_{Mie} and g_{ML}. If all of the three σ_{sca} and three β_{sca} are used as the input, the model can work

Table R1. The influence of the input parameters on the model performance.

N	$\sigma_{sca,635}$	$\sigma_{sca,525}$	$\sigma_{sca,450}$	$\beta_{sca,635}$	$\beta_{sca,525}$	$\beta_{sca,450}$	R^2	Mean±Std (%)[*1]
1	True[*2]	True	True	True	True	True	0.955	0.8±1.28
2	True	True	False	True	True	False	0.936	0.9±1.36
	True	False	True	True	False	True	0.939	0.8±1.36
	False	True	True	False	True	True	0.962	1.3±1.32
3	True	False	False	True	False	False	0.864	0.7±1.79
	False	True	False	False	True	False	0.961	1.7±1.32
	False	False	True	False	False	True	0.936	2.2±1.25
4	True	True	True	False	False	False	0.613	1.4±2.75
	False	False	False	True	True	True	0.817	2.3±1.99
5	True	True	False	False	False	False	0.568	1.4±2.90
	True	False	True	False	False	False	0.588	2.3±2.82
	False	True	True	False	False	False	0.624	2.3±2.75
6	True	False	False	False	False	False	0.401	2.6±3.3
	False	True	False	False	False	False	0.386	2.7±3.39
	False	False	True	False	False	False	0.355	2.8±3.49
	False	False	False	True	False	False	0.323	2.9±3.58
	False	False	False	False	True	False	0.344	2.8±3.52
	False	False	False	False	False	True	0.325	2.9±3.56

[*]1. The mean and the standard deviation of the relative difference between the g_{ML} and g_{mie}.

[*]2. If the value is True, the corresponding parameter is used as the input parameter of the random forest machine learning model.

well as the R^2 can reach 0.949 and the relative difference is 0.8±1.28. For those of tests (2), the best results come from the test that using the σ_{sca} and β_{sca} at 525 and 635nm with R^2=0.962 and mean value of 1.3±1.32. Despite the fact that the R^2 is

slightly higher than that of test (1), both the mean and the standard deviation of the relative difference between g_{Mie} and g_{ML} is larger than that of (1). However, the test that uses the σ_{sca} and β_{sca} at 525 only at (3) gets almost the same result. The results of the other tests are much worse than test (1). From test (4) and (5), the corresponding R^2 is much lower when the σ_{sca} and β_{sca} are not used as the input at the same time. It is concluded that the backscattering coefficient is very important for estimating the g.

For these tests of test (1), the third one in test (2) and the second one in test (3), it is hard to decide which group to be best one as the input parameters. However, it is concluded the test (1) is the most stable when comparing the mean$\pm$std of the relative difference. Thus, the three σ_{sca}, three β_{sca}, κ and RH are chosen as the input of the model and the n_{pre} is set to be eight.

Comments: P6 L165: scattering coefficients for dry aerosol? Or humidified? If humidified, isn't there a relationship between κ and the scattering coefficients?

Reply: Thanks for the comments. The three scattering coefficients and backscattering coefficients are all for dry aerosol. We made the revision accordingly at the manuscript.

Comments: P L170-171: Why did you use separate data sets for training vs testing? Why not use a subset of each, or subset of one and test on the other subset?

Reply: Thanks for the comments. It does not matter to choose different subset for training and testing. The more amounts of effective data are used as training data, the better the model can be used to predict the result. Thus we picked some data as testing data and all of the rest data as the training data. By using the datasets as training data at Gucheng site and the datasets as testing at PKU site, we can say that the random forest machine learning model can be used to predict the g values for the continental aerosols in the North China Plain.

At the same time, we conducted the random forest machine learning model by using a subset data as the test data and the other part as the training data at the field

measurement of Gucheng. The results are better than the results shown in the main manuscript. However, the aerosol population of the testing subset might be very similar to the aerosol population of the training data as both of the dataset come from the same measurement sites and thus the results using this method might be not that convincing.

Comments: P6 L173: Specify humidified vs dry nephelometer

Reply: Thanks for the comments. We have made the revision according to the comments. The scattering coefficient and backscattering coefficient are all come from the dry nephelometer.

Comments: P7 L189-191: What is ceiling (maximum altitude) of your SBDART model?

Reply: For our study, the aerosol distribution in the vertical direction is parameterized based on Liu et al. (2009). There is almost no aerosol above 50km. So the ceiling of the SBDART model in our study is 50km.

Comments: P7 L191-192: Assume HG phase functions or Mie phase functions in SBDART?

Reply: Thanks for the comments. We have made some revision accordingly to make this sentence more clearly. The profiles of the aerosol extinction coefficient, aerosol single scattering albedo and g are the results from the calculating of the Mie scattering model. The input of the Mie scattering Mie model comes from the parameterization of the aerosol vertical distributions. At the same time, we add some supporting information in section 5 of the supplementary to detail the vertical distributions of the aerosols.

Comments: P7 L194-195: Mean results from a specific span of time? Different times of day? Please be more specific.

Reply: Thanks for the comments. We have changed the manuscript accordingly.

Atmospheric meteorological parameter profiles come from the results of the intensive radiosonde observations at the Meteorological Bureau of Beijing (39°48' N, 116°28' E) at the local time of 13:30 from July to September in 2008. Kuang et al. (2016) studied these measured profiles and found that the vertical distributions of these parameters, which include profiles for water vapor, pressure and temperature, can be used as a good representation of the meteorological parameter profiles in the NCP during the summer. The mean results during the measurements, which is shown in fig. S4, are used in this study.

Comments: P7 L199: Please provide reference for SBDART defaults values

Reply: Thanks for the comments. We added the reference for the SBDART defaults values. At the same time, the default value can found in the manual of the SBDARF model, which can be inquired from the internet https://www.paulschou.com/tools/sbdart/.

Comments: P7 L192-199: I am a bit confused – were radiosonde data from the summer combined with albedo data from the winter? Can you explain?

Reply: Thanks for the comments. We changed the surface albedo value to the mean results of the surface albedo of Beijing from Jul to September in 2008, which is corresponded to the time and the location of the field measurements of the meteorological profiles used in this study.

At the same time, we found the results of our study are scarcely influenced because we are concerning the variation in the DARF due to the change of the aerosol asymmetry but not the absolute values of the DARF.

Comments: P8 L232: Does setting RH and κ values give different results from removing them as input variables and setting npre=6?

Reply: Thanks for the comments. The two methods of using the input of the model result in almost the same results. Fig. R8 gives the results of using the method. Fig. R8(a) gives the results of without the RH and κ values and fig. R8(b) gives the results

with the RH and κ values equaling to zeros. From fig. R8, it's concluded that using this two methods has almost no different results.

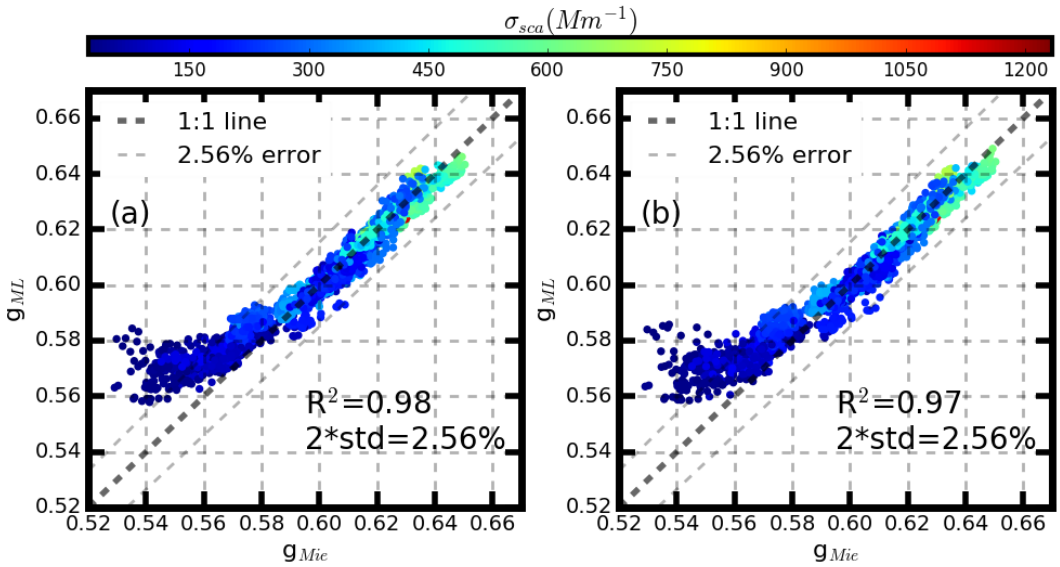

Figure R8. Comparison of the calculated g_{Mie} values and the machine learning predicted g_{ML} values (a) without the RH and κ values and (b) with the RH and κ values equaling to zeros.

Comments:P9 L238-239: Sensitivity of calculated g values to input RH and κ?

Reply: Thanks for the comments. We just want to show the results of the comparison of the g_{ML} and g_{Mie} at different RH. This corresponding paragraph is reconstructed.

Comments:P9 L247: Should mention that other empirical relationships have been given between g and b (i.e. Sheridan and Ogren, 1999 (JGR); Moosmüller and Ogren, 2017 (Atmos.); Marshall et al, 1995 (Appl. Opt.)) Perhaps one of the others is more accurate in these cases?

Reply: Thanks for the comments. We have add some discussion about the mentioned empirical relationships between g and b. The relationship between g and b used in this study is the most widely accepted parameterization scheme. Sheridan and Ogren (1999) used the relationship between g and b and up scatter fraction based on the work of Wiscombe and Grams (1976), which is basically, the same as the one in our discussion. Moosmüller and Ogren (2017) summarized the recent development of

relating the g and b. They give a corrected scheme of Wiscombe and Grams (1976) as:

$$g = -6.347b^3 + 6.906b^2 - 3.859b + 0.9852 \qquad (4).$$

Another simplified parameterization scheme is also given in this study as:

$$g=1-2b \qquad (5).$$

The comparison of g_{Mie} and the parameterized g using formula (4) and (5) are shown in fig. R9. Results show that neither of the schemes works well for the continental aerosols in the NCP. However, the parameterization scheme by Marshall et al. (1995) depends on the assumption about the aerosol size distribution, which is not suitable in this study.

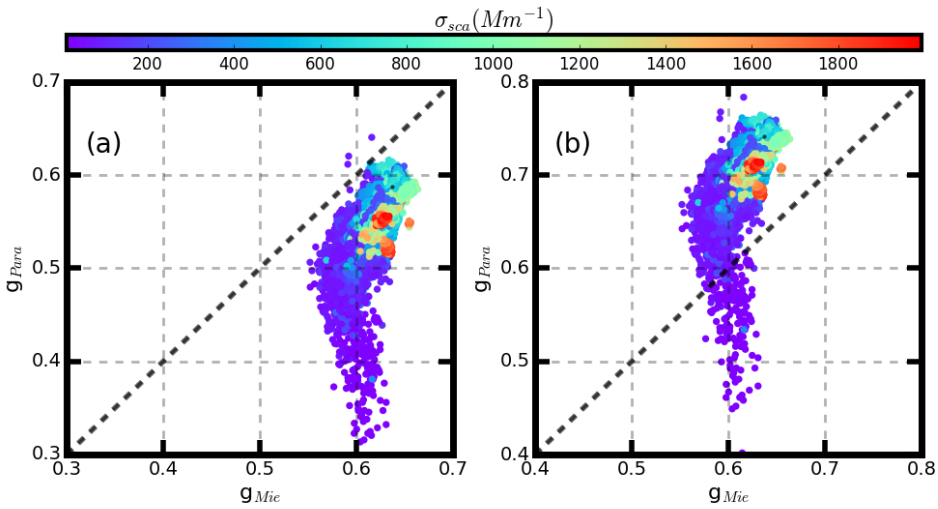

Figure R9: Comparison of the g_{Mie} and the parameterized g by using (a) formula (4) and (b) formula (5).

Comments:P9 L256-261: Why random/arbitrary deviations in input values? Can you use the specified/measured uncertainties in the measurements themselves?

Reply: Thanks for the comments. The reviewer gives a good view of doing the sensitivity test. We have made the revision accordingly. We give the uncertainties of the measured parameters, which are used as the input of the machine learning model. The Monte Carlo tests were carried out to study the sensitivity of predicted g to each input parameter. Details of the discussion are shown in section 4.2.2 of the manuscript.

Comments:P10 L269: Based on your own comments earlier in the paragraph isn't the model simply more sensitive to completely independent input variables, which happen to be RH and κ in this example? What happens if you only input σsca, or specify that σsca and βsca must vary in the same way?

Reply: Thanks for the comments. We have changed the method of conducting the sensitivity test. The Monte Carlo tests were carried out to study the sensitivity of predicted g to the uncertainties of each input parameter. We find that κ is the most important factor that may influence the accuracy of predicting g because the κ has great uncertainties as the input parameter. More details of the discussion are sown in section 4.2.2 of the manuscript.

Comments:P10 L274-276: Can you clarify what you mean by "actual calculated aerosol phase function" – does this imply those measured directly by CCD-LADS, or those calculated using Mie theory? If the latter, can you also compare DARF using directly measured phase functions using CCD-LADS?

Reply: Thanks for the comments. We have made some revision to make the discussion more clear. The 'actual calculated aerosol phase function' implies the phase function calculated by using the Mie theory.

We don't think it is proper to using the measured phase functions by CCD-LADS at this section for study. As shown by Bian et al. (2017), the measured phase function

by CCD-LADS shows good consistence with the calculated phase function, which to some extent means that the calculated phase function by using Mie theory is reliable. However, there still exist some uncertainties of the measured phase function. The phase function calculated by using the Mie theory is a better choice for testing the uncertainties of DARF in the using of HG phase function.

Comments:P10 L289: Suggest citing McComiskey et al, 2008 (JGR)

Reply: Thanks for the comments. We have made the revision accordingly.

Comments:P10 L289-290: It is unclear how this section ties into the rest of the paper, or what novelty comes from the discussion. As you point out, previous studies have also undertaken to study how modeled DARF varies with input parameters like asymmetry parameter. Can you add anything unique to the literature regarding g and DARF?

Reply: Thanks for the comments. The revisions were made accordingly in section 4.3.2. The logic of this section is briefly described here. In this study, we use the random forest machine learning model to calculate the g. The uncertainties of the g values from the input parameter is estimated to be 1.95% when predicting g and the total variation in running the random forest machine learning model is estimated to be 4.47%. At the same time, the g can varies about 10% for different aerosol PNSD and can be enhanced by 20% with the increment of RH from 30% to 90%. It is very important to know the extent of the variation in DARF corresponding to the uncertainties from g. At the same time, it is very important to know the uncertainties of the DARF when using the random forest machine learning to predict g.

Comments:P10 L290-298: Would it be perhaps more appropriate to vary g according to a typical RH profile? You show that g tends to be higher when RH >90%, conditions which are only likely to occur in specific layers of the atmosphere depending on the local meteorology.

Reply: Thanks for the comments. We have changed the method of estimating the

influence of RH on g and DARF. It is estimated by calculating the DARFs with the g values calculated from the dry parameterized aerosol population profile and g values calculated from the RH and aerosol population profiles.

Comments:P10 L291: Where does 2.3% come from?

Reply: Thanks for the comments. We have made the revision accordingly. We attend to estimate the uncertainty of DARF form the uncertainties of the random forest machine learning model. It is worthy note that the values 2.3% is changed because the method of estimate the uncertainties of predicting g is changed.

Comments:P12 L339: Regarding "actual phase function", please clarify (see general comments)

Reply: Thanks for the comment. We have made the revision accordingly.

Comments: Figure 1: What about Hauirou? Can you also provide graphs of average PNSDs for each site?

Reply: Thanks for the comments. We have added the information of the Huairou (UCAS) in the figure and corresponding text was changed.

Comments: Figure 3: Clarify which data used for these calculations – which site.

Reply: Thanks for the comments. We have made the revision at section 4.2.1 and the corresponding section.

Comments: Figure 4: Please explain in text a possible justification for the higher dependence on $\sigma_{sca,635}$ compared to other wavelengths. Also, why is g very dependent on 5% variation in κ, but not significantly affected by greater deviations?

Reply: Thanks for the comments. We reconstructed the method of conducting the sensitivity study. We found some of the result unacceptable in the figure. And the figure is removed from the manuscript.

Comments: Figure 5: Clarify “calculated phase function” (see general comments)

Reply: Thanks for the comments. We have made the revision accordingly.

Comments: Figure 7: Please clarify figure in caption by changing order: gCCD,Cal as a function of gMachine,cal

Reply: Thanks for the comments. We have made the revision accordingly.

Comments: Technical comments

P6 L157: Please correct “tress” to “trees”

Reply: Thanks for the comments. We have changed the manuscript accordingly.

Bian, Y., Zhao, C., Xu, W., Zhao, G., Tao, J., Kuang, Y. (2017) Development and validation of a CCD-laser aerosol detective system for measuring the ambient aerosol phase function. Atmos. Meas. Tech. 10, 2313-2322.

Boucher, O. (1998) On Aerosol Direct Shortwave Forcing and the Henyey–Greenstein Phase Function. Journal Of the Atmospheric Sciences 55, 128-134.

Breiman, L. (2001) Random Forests. Machine Learning 45, 5-32.

Chamberlain-Ward, S., Sharp, F. (2011) Advances in nephelometry through the Ecotech Aurora nephelometer. ScientificWorldJournal 11, 2530-2535.

Charlson, R.J., Porch, W.M., Waggoner, A.P., Ahlquist, N.C. (1974) Background aerosol light scattering characteristics: nephelometric observations at Mauna Loa Observatory compared with results at other remote locations. Tellus 26, 345-360.

Horvath, H., Kasahara, M., Tohno, S., Olmo, F.J., Lyamani, H., Alados-Arboledas, L., Quirantes, A., Cachorro, V. (2016) Relationship between fraction of backscattered light and asymmetry parameter. Journal of Aerosol Science 91, 43-53.

Hu, X., Belle, J.H., Meng, X., Wildani, A., Waller, L.A., Strickland, M.J., Liu, Y. (2017) Estimating PM2.5 Concentrations in the Conterminous United States Using the Random Forest Approach. Environ Sci Technol 51, 6936-6944.

Huttunen, J., Kokkola, H., Mielonen, T., Mononen, M.E.J., Lipponen, A., Reunanen, J., Lindfors, A.V., Mikkonen, S., Lehtinen, K.E.J., Kouremeti, N., Bais, A., Niska, H., Arola, A. (2016) Retrieval of aerosol optical depth from surface solar radiation measurements using machine learning algorithms, non-linear regression and a radiative transfer-based look-up table. Atmospheric Chemistry and Physics 16, 8181-8191.

Kuang, Y., Zhao, C., Tao, J., Bian, Y., Ma, N., Zhao, G. (2017) A novel method for deriving the aerosol hygroscopicity parameter based only on measurements from a humidified nephelometer system. Atmos. Chem. Phys. 17, 6651-6662.

Kuang, Y., Zhao, C.S., Tao, J.C., Bian, Y.X., Ma, N. (2016) Impact of aerosol hygroscopic growth on the direct aerosol radiative effect in summer on North China Plain. Atmospheric Environment 147, 224-233.

Liu, P., Zhao, C., Zhang, Q., Deng, Z., Huang, M., Xincheng, M.A., Tie, X. (2009) Aircraft study of aerosol vertical distributions over Beijing and their optical properties. Tellus Series B-chemical & Physical Meteorology 61, 756–767.

Marshall, S.F., Covert, D.S., Charlson, R.J. (1995) Relationship between asymmetry parameter and hemispheric backscatter ratio: implications for climate forcing by aerosols. Applied Optics 34, 6306-6311.

Moosmüller, H., Ogren, J.A. (2017) Parameterization of the Aerosol Upscatter Fraction as Function of the Backscatter Fraction and Their Relationships to the Asymmetry Parameter for Radiative Transfer Calculations. Atmosphere 8, 133.

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

Pandey, A., Chakrabarty, R.K. (2016) Scattering directionality parameters of fractal black carbon aerosols and comparison with the Henyey-Greenstein approximation. Opt Lett 41, 3351-3354.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion,B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg,V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot,M., and Duchesnay, E.: (2011) Scikit-learn: Machine Learning in Python. J. Mach. Learn. Res. 12, 2825-2830.

Sheridan, P.J., Ogren, J.A. (1999) Observations of the vertical and regional variability of aerosol optical properties over central and eastern North America. Journal of Geophysical Research: Atmospheres 104, 16793-16805.

Wex, H., Neusüß, C., Wendisch, M., Stratmann, F., Koziar, C., Keil, A., Wiedensohler, A., Ebert, M. (2002) Particle scattering, backscattering, and absorption coefficients: An in situ closure and sensitivity study. Journal of Geophysical Research: Atmospheres 107, LAC 4-1-LAC 4-18.

Wiscombe, W.J., Grams, G.W. (1976) The Backscattered Fraction in two-stream Approximations. Journal Of the Atmospheric Sciences 33, 2440-2451.

Zhao, G., Zhao, C., Kuang, Y., Tao, J., Tan, W., Bian, Y., Li, J., Li, C. (2017) Impact of aerosol hygroscopic growth on retrieving aerosol extinction coefficient profiles from elastic-backscatter lidar signals. Atmospheric Chemistry and Physics 17, 12133-12143.