Date: January 19, 2018
Paper Review: A novel method for calculating ambient aerosol liquid water contents based on measurements of a humidified nephelometer system
Authors: Ye Kuang et al.

The authors present both empirical and machine-learning methods for determining an aggregate or effective volume and the aerosol water content from the nephelometer Angstrom and backscattering coefficients. The empirical method uses cross-correlations between the aerosol scattering coefficient, measured volume, backscatter fraction, and Angstrom exponent to estimate the aerosol volume. The machine-learning method uses backscatter and Angstrom exponents to mimic the aerosol scattering to volume ratio. The machine-learning method offers a valuable tool that could be applied to many aspects of atmospheric aerosol and chemical predictions. The paper needs further development. I think it's important to refine the methodology, improve the paper organization, and clarify some of text to make this a stronger paper. The section linking scattering hygroscopic growth to volume hygroscopic growth doesn't follow a valid analysis method. The fRH and gRH data come from different measurement sites. I suggest leaving out the sections on volume hygroscopic growth as well as the discussion on kappa-Kohler. I don't recommend publication until the paper is restructured. I suggest resubmitting the paper after removal of the sections on hygroscopic growth and water content.

The paper need better organization and clear, step-wise presentation of the methods and results. The methodology is scattered throughout the paper. Description of the results is vague and tends to gloss over important features.

Please edit the English grammar and word order. Avoid run-on sentences. You've a tendency to state conclusions without providing supporting evidence. State the methods used and then the data results. Presenting methods and results are interspersed and add to confusion.

Introduction

1. Although the methodology is different from other inversion techniques, such as those of Ziegar et al., that calculate an "effective" gRH from fRH, the methodology is similar to that used in Aeronet retrievals of the aerosol effective radius from the AOD, Angstrom exponent and asymmetry parameter. There was a paper that attempted to calculate fRH or aerosol water using Aeronet data, but it suffered from low signal and spatial resolution. Can the authors describe how their method is similar to or different from the Aeronet retrievals and also speculate if this method could be used with remote sensing AOD measurements? Below is a link that has links to their "spectral deconvolution algorithm".

https://aeronet.gsfc.nasa.gov/new_web/Documents/Inversion_products_V2.pdf
or
Atmos. Meas. Tech., 10, 695-708, 2017, https://doi.org/10.5194/amt-10-695-2017

2. The introduction needs to state more about the methodology other than saying it's "a novel method". Add a paragraph describing the two techniques. Describe the empirical model use of

size-dependent parameters (backscatter and Angstrom) to predict the ratio of scattering/volume. Describe how machine-learning methods, using large data sets, over a long-time period as input, mimic the system behavior via feed-back loops to predict an output.

3. In Section 2.1 Transfer Table S1 to Section 2.1. You can leave the detailed measurement description in the supplement.

Give general information on the breakdown of the aerosol composition between organics/sulfate/nitrate/dust at this time of year.

Sections 2.2 to 3.2 need to be reorganized.
I suggest segregating Section 2 into
Section 2.2 is closure, 2.3 is Mie theory, 2.4 Machine learning method

I suggest renaming section 2.2 as "Closure Calculations". Show the scattering closure between the measured and calculated scattering coefficient. The integrity of the volume and fRH closure depends on the scattering closure. Figure 2 would validate the measurements better if 2a showed the scattering closure and 2b showed the scattering vs volume.

Figure 2 shows 2 branches or subsets of the data; one above and a 2nd below the fit line. Is this behavior present in the scattering closure? Do these two branches represent 2 different aerosol types or multiple size modes?

The application of an average Rvsp to estimate the volume doesn't add to the paper and distracts from the other methods. I suggest removing it from the analysis.

Section 2.3 Mie
Move equations 5 and 6 to the start of section 2.3 and explain how you relate scattering to volume and the assumptions in this approach. Move the discussion of Mie theory from section 3.1 to section 2.3. State your adaptation of Mie Theory in a clear, stepwise, logical fashion. Show that going from equations 5 > 7 >6 assumes that Q is roughly linear with r such that Q=k*Q(m).

Describe Mie model using simulated data with 4 aerosol types and results in Figure 3. Describe limitations of assumption that Q is linearly proportional to r.

Describe what the dotted red lines represent in the Figure 3 caption.

Section 2.5 Machine learning

Describing an alternate method of machine learning using size-dependent scattering parameters; Angstrom and backscatter to mimic the measured scattering/volume ratio.

Give some background on “machine learning” and the algorithm name. Can you add a simplified algorithm decision tree with basic logic steps or diagram that would help explain the process?

Results and Discussion:
Section 3.1 Empirical method
Refer back to Figure 2 and need for estimating a variable kscat or R. Introduce the empirical method of determining R from HBF and Angstrom exponent. Explain figure 4.

In your description of the results, use and simple and direct language. Leave out extra information on the impact of BC and mixing state on the HBF and Angstrom exponents until the discussion of the data fit lines as these are secondary contributions.

Explain Figure 5a and variance about fit line in relation to HBF. This variance likely stems from the Angstrom exponent and HBF describing a fraction of the PNSD. Show a plot of HBF(450, 550, 700nm) and Angstrom exponents(450/550, 450/700 and 550/700) versus r for a lognormal size distribution. The plot will show the sensitivities of these parameters to aerosol size.

Section 3.2 Machine learning
Explain how machine method uses 6 parameters to describe the aerosol volume relative to the scattering. Explain figures 5b and Figure 6 and how machine learning is an improvement over the empirical method.

3.3 fRH and Vrh

The method of simulating the Kd size distribution from variations of the average and then applying this to the measured size distribution to obtain a 4 modelled volume growth values isn’t valid. The PNSD shape will change with aerosol type as will the Kd size distribution. Aerosol size-dependent growth varies with size such that multiplying an entire Kd distribution by a constant won’t reproduce the Kd distribution of a different aerosol types. In addition the Kd and frh data come from two different measurement sites.