Atmos. Meas. Tech. Discuss.,
doi:10.5194/amt-2017-468-RC2, 2018

Atmospheric Measurement Techniques Discussions

Interactive comment on "A Machine Learning Approach to Aerosol Classification for Single Particle Mass Spectrometry" *by* Costa D. Christopoulos et al.

Anonymous Referee #2

Received and published: 11 April 2018

Overall, this paper strives to explain a particular application of a machine learning approach to the classification of single particle mass spectra and to describe the performance of the approach to a particular "blind" test. The dataset that is used to train and test the machine learning approach is a good one for such a study, containing a variety of different particle types. The major shortfall of this paper is that the authors neither explain the details of the machine learning approach fully nor do they fully engage with the aerosol classification results, leaving the reader somewhat confused in both realms. In addition, the authors do not attempt to address the performance of their approach in terms of time, or give information about how applicable it would be to ambient data sets where particles would not necessarily be of such distinct types.

Finally, they given no metrics for success – how good is good enough performance for this approach? How good are other methods, compared to that presented here?

I would recommend that this paper be significantly revised, in such a way that a) the machine learning approach can be fully described and choices made justified with data, and b) the aerosol particle classification results can be fully examined and compared to other methods.

Specific Comments: 1. The paper reads as if it was written by two separate people, one for the algorithm discussion and one for the aerosol particle classification discussion. This should be addressed as a final version (or versions) is developed. For example, on p. 4, the transition between lines 12 and 13 is abrupt and jarring. 2. The authors refer to "volatile" components of aerosol particles multiple times in the paper (first p. 3, line 13). I believe they mean semi-volatile, or at least "more volatile" than other components. Volatile species would not be expected to be found in particles. 3. In section 2.2, where the training data set is introduced, the authors need to discuss the applicability of this dataset to any "real" experiment. Would these particles be a good representation of ambient particles, for example? 4. In the discussion of the data presented in Table 2, the authors state that the columns labeled "broad" are applicable to the categorization of the particles when they are lumped together into broad chemically-similar categories. These categories should be defined in this context (and not just in the context of the confusion matrices), presumably in Table 1. Any interpretation of the differences should be discussed in the results section. In addition, the paragraph on p. 11 about the 59+ ion observed in some samples (which ones?) should be moved into the results section. Finally, is it certain that 59+ is Co rather than an organic contaminant? 5. The authors provide no information about the average mass spectra of the individual particle types and the variability within "identical" particles or between particle types. This would seem to be an important parameter in determining how well the algorithm can do to separate them. Based on the two peaks shown in Figure 1, this is an important factor. 6. The discussion of confusion matrices was confusing. Essentially, these matrices

represent normalized counts of the sorting of known particles into the available classes. This can be stated much more cleanly than the three page description provided on pp. 12 – 14. This is another example of a section that is trying to be both an algorithm and a particle chemistry paper, and not mixing the two effectively. 7. Figure 5 and the discussion of the "blind" test in Section 3.2 are key to the goals of this paper, but are confusing in their presentation. Regarding the text, why do the authors not know the number of particle types that were used in the challenge (p. 15, line 2 "3 -4 aerosol types . . . were aerosolized")? How well can the results be evaluated if the test conditions are not known? The authors describe a probabilistic correction for the mis-labeling that they observe in their confusion matrices (p. 15, lines 14 – 18), and say that the results "better represent the underlying aerosol population." (p. 15, lines 17 – 18), but they don't provide the data to evaluate this claim. The data presented in Figure 5 do not make the case that the authors are trying to make. While the two models (positive and negative) show relatively good agreement with each other, the representation of the particles introduced into the chamber is poor. The authors show the breakdown of soot, SOA, and mineral particles introduced in Figure 5, state that the soot particles are too small to see with their instrument, and then compare against the soot-containing dataset anyway. If the pie that represents "Aerosols Reported by AIDA" were renormalized to include only observable particles in this experiment, SOA would represent 44% of the pie and mineral would represent 56% – assuming that the "Aerosols Reported by AIDA" pie is also representing number of particles, rather than mass of introduced particles (this is not stated). If this pie represents something other than number, there is no comparison to the blind sample possible in this figure, only a comparison between the two models.

Technical Corrections: 1. P. 10, line 17: remove the word "rows" from the line. 2. P. 15, line 3: "AIDA" is written as "ADIA." 3. Figure 1 is very difficult to read. The black points for the "Blind AIDA Aerosol" are only visible on the right-most part of the graph, in the region of (0.3, 0.04) and the similar colors are hard to differentiate. Consider a figure like this that is broken out into the broad categories of particle types. 4. Figure

5, negative model, the small pie includes 5 wedges but only 4 labels.

Interactive comment on Atmos. Meas. Tech. Discuss., doi:10.5194/amt-2017-468, 2018.