Atmos. Meas. Tech. Discuss.,
doi:10.5194/amt-2019-106-RC2, 2019

Atmospheric Measurement Techniques Discussions

Interactive comment on "Classification of iron oxide aerosols by a single particle soot photometer using supervised machine learning" *by* Kara D. Lamb

Anonymous Referee #2

Received and published: 29 April 2019

The manuscript amt-2019-106 by K. Lamb describes the application of a machine-learning algorithm known as random forests to the data set produced by the NOAA SP2. The work is thorough and the writing is of an excellent calibre. The contribution to the field is significant as this work may significantly influence future SP2 data analyses (hopefully without making them more opaque, which is the inevitable shortcoming of machine learning). I am happy to say that I have only minor requests for information/modified graphs. The manuscript should be published in AMT after the following minor corrections, most of which request language clarification or additional details.

Comments on the abstract:

The abstract specifies that conditional probabilities of each class are provided but then later refers to 'correct identification'. This change of language from probabilistic to absolute identification confused this reviewer on the first read, and the text could be slightly changed to be consistent with a probabilistic perspective (including a definition of what "correct" means in terms of probability... is it a probability of 90%? is it when one class was more than twice as likely as the next? this discussion could also be mentioned in the main text).

Similarly, please explicitly define the "broader class" approach in the abstract. For such technical work, many readers will only read the abstract.

Finally, it should be made clear in the abstract that you are not using an "SP2" but a "modified SP2". This work is not extensible to the standard SP2. Conversely, this work should motivate other SP2 users to modify their SP2s, therefore the modifications must be highlighted.

Minor comments / Requests for information:

I found that the author's decision to include a large amount of detail on the basics of machine learning helpful, and since this is an interdisciplinary journal it is appropriately detailed. However, the author may also consider moving sections of that text to an Appendix to allow the main text of the paper to focus on the essentials.

Page 8, line 16. Why should only the red PMT alignment be sensitive? Is it due to a unique physical configuration of the filter/PMT? Can the author please either speculate or state that this is as surprising to her as it is to the reader?

On page 11 the author mentions that various other machine-learning algorithms were tested with negative results. I think many readers would appreciate more information on these negative results (too often we only report successes). I suggest including a brief appendix (a few paragraphs or table) describing what was done. Surely the author compiled metrics on the different algorithms before deciding to focus on random

forests; this information would be of value to readers who need to know whether their data sets might be significantly different to this one. This would also provide objective support for the manuscript's focus on random forests.

Page 12, it is unclear to me what happened to particles with no valid position-sensitive detector information. Were they rejected?

Please refer to Figure 4 in the legend of Table 2, for the benefit of the non-linear reader. Please also define x4 (post incandescent scattering) more precisely; that is, specify what time interval after scattering was used. Is it defined when incandescence returns to zero? Is it defined for a fixed distance from x8, the position in the laser? Is it possible that this definition influenced the results?

In Section 4.1, several statements such as "most important feature" and "significantly worse classification" were used. It would be helpful if these were quantified numerically, as the reader does not know how to interpret them otherwise. Also please clarify "reduced by approximately 1/3rd", does this mean "reduced by a factor of 0.33" or "...0.7"?

In Section 4.1, a reduced set of training features was justified because "there was a clear break in the relative importance of different features". Presumably the author compiled statistics on prediction accuracy when sequentially removing features, otherwise this statement could not be made. This would be very informative to include as a table or figure.

Page 18 line 29. The private communication with S. Kaspari must definitely be expanded on as it is a very important part of the data interpretation. How did Kaspari prove that the particles were rBC and not dust? Microscopy? Can a quantitative analysis be made?

Section 4.3, the author comments on the low number of dust particles impacting accuracy at small sizes, can this fact be placed in the context of expected FeOx size

distributions? I initially thought it would be insignificant but then a paper on penetration into the brain is cited later.

Page 23 line 26, "misidentifying" based on what? How do you know the true class? It seems like you are somehow convinced that these particles are truly rBC which the algorithm cannot identify – if so, can you please explain why?

Figure 9 legend. Coatings do not allow particles with a smaller rBC mass to be detected (in the incandescence channel). Perhaps the real reason for more smaller particles being detected here is simply more were available (nebulizing fullerene soot produces larger particles than combustion engines). A limit of quantification for the color ratio (eg 0.8 fg) should be defined and discussed.

Figure 9. In my own experience with extremely dense "point plots", I have found that it is impossible to visualize the histogram (or pdf) once the overlap becomes as severe as in this figure. The same problem will occur in Figure 2, but is not misleading (or easy to improve) there. For Figure 9, please add a panel showing the histogram of color ratios for each class, in the region of constant color ratio (>2 fg), or please change to 3 panels of joint PDFs (cumulative count instead of overlapping points), or 3 panels of transparent points.

Page 25 line 5, presumably the author has data to prove this strong dependency? Please show it.

Very minor comments:

Please add abbreviations to Table 1. (Rather than in the legend of Figure 2.)

Contractions such as "it's" are normally discouraged; I leave the details of this to the AMT editing staff.

Page 8, line 12-14. This sentence is grammatically flawed and I can't see what it should be corrected to; please revise.

Figure 3 legend, expand "PS" to position sensitive like in Figure 4.

I would suggest changing "L-II" to "LII" because the latter is an established acronym, and because hyphenated words typically do not retain their hyphens when abbreviated.

Page 11 line 6, change eg to ie.

Page 15 line 8, after "subset" state "discussed below in section ..." for the reader's benefit.

Page 22 line 15, here and later the word "aerosols" starts to creep in to the lexicon, which I find confusing (the author seems to be using "aerosols" as "collection of particles" rather than "suspension of particles in a gas", perhaps "particle ensemble" or "sample set" would be clearer).

Interactive comment on Atmos. Meas. Tech. Discuss., doi:10.5194/amt-2019-106, 2019.