Atmospheric
Measurement
Techniques
Open Access

Discussions

AMTD

Interactive
comment

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

Anonymous Referee #1

Received and published: 20 April 2019

The author demonstrates for the first time the usefulness of supervised machine learning algorithms for post-processing of the waveform data acquired by the single-particle soot photometer, with particular attention to the classification of iron oxide aerosols. First, the author provides a detailed review of the previous works and clarifies the issues to be solved/mitigated in this work. Second, the author defines the (physical and mathematical) features embedded in the signal waveforms and explain the machine learning algorithm applied to them. Finally, the author shows the suggested algorithm can reduce the chance of misclassification of the iron oxide aerosols than the conventional simpler algorithm. Along with the presentation of the results, the author also fully explains the limitation of the applicability due to a particular selection of laboratory

samples used to train the algorithm. The manuscript is very logically written, and all the figures are easy to understand. Considering the superior quality of discussion and presentation, I can recommend the publication of this work. However, I request minor revisions to improve the readability and influence to a broader audience (including other SP2 users).

Minor comments: Most of the contents in sections 3.3-3.5 look like an overview of "the established theory" of machine learning. If so, the author could shorten these sections (or moved to the supplementary information).

p.2, line 8. nitrogen - > nitrate ?

p.8, line 24. real part -> imaginary part ?

p.18, line 9-10. "retaining the 11 most important features for the 6-class case and the 9 most important for the 3-class case"

Please refer Table3 in this sentence. Otherwise, readers could not follow which features are used here.

p.24, line 6-7. "This method improves upon the performance of previous classification methods using only 3 or 4 features derived from the single particle signals"

Please clarify which features the author mention here.

p.25. line 4-5. "we recommend acquiring samples for training data sets with the same instrument, optical configuration, and operating conditions as the data sets to be processed."

To my opinion, it is better to mention this point as "important requirement" rather than "recommendation".