Atmos. Meas. Tech. Discuss.,
doi:10.5194/amt-2016-7-RC1, 2016


Atmospheric Measurement Techniques Discussions
Open Access
EGU



# *Interactive comment on* "Retrieval of Intensive Aerosol Microphysical Parameters from Multiwavelength Raman/HSRL Lidar: Feasibility Study with Artificial Neural Networks" *by* M. Mustafa Mamun and Detlef Müller

**Anonymous Referee #3**

Received and published: 28 March 2016

Retrieving aerosol microphysical parameters from lidar data is an active topic of aerosol research. The manuscript of M.M. Mamun and D. Muller explores the feasibility to retrieve such properties using Artificial Neural Networks, which seems to be a promising approach for future research. Unfortunately, the current form of the manuscript has several weaknesses and omissions that should be addressed before being published.

Major comments:

The authors should give better motivation for their approach. They mention that ANN could "increase data processing speed and quality". However, they are not mentioning



Printer-friendly version

Discussion paper

other techniques for the retrieval of aerosol properties, like the linear estimator approach of Veselovskii et al. 2012. This method in particular is fast and can provide the same parameters like the ANN, namely aerosol effective radius and refractive index. Given this, what are the current (or future) benefits of the ANN approach?

The authors do not address the effect of measurement uncertainties in the quality of the results. This is a crucial omission in the analysis. In ill-posed problems, such as the one tackled by the authors, a small change in input data could lead to completely different retrieved solution. Stabilizing (regularizing) the solution is one of the main outcomes of original work like Muller 1999a, b etc. The authors should perform at least some sensitivity test to show that their solutions are stable. Given the processing speed of NN, this should be a straightforward task. Without this, all the results presented by the authors are of limited usefulness. If the results are unstable, the authors should explain how this can be improved in the future (e.g. training the network with perturbed data etc.).

The authors give a very detailed analysis of the network performance for different particle radius and angstrom exponents etc (figs 6 to 13), but they lack an overview of the performance. I would like to see first some overview plots and tables describing the overall performance of the network. These could focus on the "easy" case of 3a + 2b dataset (Case A). In the current form, it is very hard for the reader to extract any useful conclusions.

For the 10 – 100nm range, the authors state that "results and discussions will be shown in our future study". This is strange! Why should part of the analysis be omitted? Are the results very interesting or very bad? The authors should present the results and clearly state the problems and opportunities that they present. Otherwise, the conclusion that the ANN method can model particles from 0.01um to 0.50um is not valid.

Minor comments [line number in brackets]:

[93] Explain what are the requirements of climate modeling community. Why is the processing speed of e.g. Muller et al 2014 not sufficient for climate modeling?

[136] Revise this sentence. Feedback connections are against the definition of feed-forward NN!

[211 – 223] This part reads like a training manual for MATLAB functions. The authors should focus on the algorithms and not on the specific implementations. These information should be moved to an Appendix.

[261] Give a brief description of the computer resources you use. This will give context to the reader concerning the computational effort needed for your approach.

[268 – 275] These numbers should be better summarized in one sentence or a small table.

[290] In a previous paragraph you mention that you use 70% for training, 20% for testing, and 10% for validation. Why do you use an extra 50% of the original data for blind-testing? Give supporting evidence/references that fifty-fifty sharing is a good approach.

[281] Give link to Matlab in the first use of the name. Also, give proper reference to the NN toolbox of Matlab.

Technical comments:

[63] Provide a reference for EarthCARE mission.

[208] The learning algorithm is the LM algorithm., Trainlm is just the MATLAB implementation of this algorithm.

[Tables] You define the scenarios (A, B, . . .) in table 2. You can omit the details in all other tables (“Input combination” column).

[Tables] Take care that all values are provided with the proper significant digits.

[Figures] The font is too small, it should be increased to a legible size. Consider summarizing your results in fewer, clearer plots.

---

Interactive comment on Atmos. Meas. Tech. Discuss., doi:10.5194/amt-2016-7, 2016.