Atmos. Meas. Tech. Discuss.,
doi:10.5194/amt-2019-327-AC1, 2020

Atmospheric Measurement Techniques Discussions

Interactive comment on "Low-level liquid cloud properties during ORACLES retrieved using airborne polarimetric measurements and a neural network algorithm" *by* Daniel J. Miller et al.

Daniel J. Miller et al.

daniel.j.miller@nasa.gov

Received and published: 11 February 2020

Author Response: AMT-2019-327

We would like to thank the reviewers for their thoughtful comments. In response to the reviewer's suggestions we will make several important clarifications in the manuscript text. In addition to this, we will update the manuscript to include links to the public dataset.

Specific Comments:

1) We will revise this sentence to read as follows: "In this study we developed a neural

network (NN) that can be used to retrieve cloud microphysical properties from multi-angular and multi-spectral polarimetric remote sensing observations."

2) P3, L33-34: Agreed, and this is a very good point. Our intention was for this sentence to indicate that the NN does not impose a specific analytical interpretation of the training data. We will revise this sentence as follows: "... in a manner that is independent of any imposed parameterized relationship between geophysical variables and the observations ..."

3) P7, L29. There is still some limited information about the cloud droplet size in the angular and spectral dependence of the total reflectance, but the SWIR bands provide the greatest sensitivity to droplet size information. The primary reason we wanted to develop a NN that mixed total and polarized reflectance information is to achieve a future objective (not the focus of this paper) of performing simultaneous aerosol and cloud retrievals using a similar framework.

4) P8, L9. In the revision we will modify to: "The HSRL-2 screening criteria were removed..."

5) P12, L10-11. The cloud top height was originally arbitrarily fixed to limit the size of the training dataset. Also, the cloud top to observation altitude (flight altitude) separation is actually the more important difference impacting the cloud retrievals. By fixing the cloud top height, which is relatively consistent throughout the campaign, the difference in cloud top and observation altitude is modeled.

6) P13, Table 1. As in the previous work (Segal-Rozenhaimer et al (2018)) we assumed a black ocean, and therefore no dependence upon wind speed or ocean color. Considering that the smallest cloud optical depth in our training set was 2.5, changes from this assumption would have a minimal impact on the observed radiances, and an even smaller impact on the polarimetric observations. Furthermore, the expression of the cloud bow in the polarimetric signal is quite different than that of a glint (peaked in the reflected sun direction) or ocean color (likely isotropic) reflection. For these rea-

sons, we concluded that training for those parameters as well would not be worth the computation effort. Segal-Rozenhaimer et al (2018) described these approximations in section 4.1, and it wouldn't hurt for us to summarize in this paper. So, we added the following text to the end of section 3.1:

"For reasons of computational efficiency, the training set simplifies some aspects of nature. Compared to the predecessor paper (Segal-Rozenhaimer et al., (2018), we used a larger set of geometries and wider range of parameter values, but many of the same approximations. For example, this training set assumes plane parallel radiative transfer (neglecting 3D effects) and a 'black' ocean surface with no reflections due to sun glint or ocean color. The former is beyond our computing resources and desired level of parameterization, while the former is expected to be heavily attenuated by the cloud."

7) P13, L7, "principle" was amended to "principal". Thank you for catching this error.

8) P14, L7-9. We have attempted to parse this more carefully in the revised manuscript.

9) P14, L17-20. Our approach to standardization was performed as a two-step process: the first is the standardization based on the uncertainty and the scale of the different inputs, and the second is the linear scaling step, to put the inputs in a range between about -1 and 1. We chose to perform these steps separately since we know the uncertainty model for our inputs. Indeed, although the common practice is to scale/normalize all inputs similarly (scaling all covariances to the same values), in the case where some inputs have different contribution (in our case, the reflectance data has larger uncertainty so we wanted it to be less "visible"), it is advisable to scale the inputs differently (LeCun et al., Efficient BackProp, 1998). One might argue that we could have done this in one step but we felt that this approach has a better physical basis (i.e. using the uncertainty model to rescale the inputs and then the "standard" scaling to assist in the convergence of the system, as in the common practice.

10) P14, Section 3.3. While we agree, the evaluation of those retrievals has been

performed before on other similar simulated datasets (e.g., Alexandrov et al. 2012a,b and Miller et al. 2018), and the behavior of our training set can likely be approximated from those previous studies – which both used the same forward radiative transfer model.

11) P14, L25. We have compared numerous architectural choices in the process of arriving at the one used in this research (including the architectures we experimented with throughout the development of our previous paper, Segal Rozenhaimer et al. (2018) as you mention). We tested architectures with many fewer hidden nodes (between 40 and 1024), including varying the number of hidden layers and got best results with the architecture we are currently using. Indeed, although PCA theoretically represent the majority of the variance, it isn't always a one-to-one physical representation, which seemed to be important in our case (PCA reff results were less generalized than the current one), since the RSP has a large number of viewing zenith angles, and their full representation seemed to add more information. The number of hidden layers and nodes is often a trial-and-error parameter for each problem. Hence, we performed cross-validation on the various architectures tested and got the best training and test results for the selected architecture. We tested drop-offs as well, but did not get better results, or greater generalizability.

12) P14, L26. We have revised this to say: "This network is instead trained using a mini-batch method, where a batch of samples (128) is presented to the network and the weights on each of the hidden layers are only updated after each batch has been processed.

13) P15, L12. Noise is generated each training cycle and is representative of radiometric uncertainty. However it is not based on the instrument uncertainty model used in our uncertainty standardization process described earlier in the paper.

14) P15, L22-23. Even in our previous network we were not getting reliable results for ve (compared to PP retrieval of ve). The most likely explanation is that when the NN is

simultaneously optimized for tau, re, and ve the retrieval the behavior of the ve retrieval suffers the most because there is simply less sensitivity to this parameter than to re or tau. Also, the PP retrieval can often struggle to retrieve ve as well (refer to figure 4 of Miller et al. 2018). I think a NN result trained for only a single variable retrieval, such as in the work of Di Noia et al. (2019), would perform a more accurate ve retrieval. However, because the objective of our network is to eventually disentangle the coupled retrieval of aerosol and cloud properties, we have attempted to avoid single variable retrievals.

As far as angular sampling is concerned, we are indirectly using the same criteria as the PP retrieval – because we have conditioned our retrieval on the comparison to an existing PP retrieval. This means that we are always seeing rainbow angles in the results shown here. We used to have a stricter angular sampling condition (only near the principal scattering plane), but we have relaxed that requirement and increased the size of the training set in this study. This relaxation of angular requirements increases the amount of observations we can retrieve but makes the retrieval problem more complicated.

15) P16, L3-4. Sorry, this is a confusing attempt to distinguish the output of the NN and the linearly adjusted output of the NN. We have attempted to clarify the language in the revision.

16) P16, L20. We have reframed this statement. It was originally intended to convey the sense that machine learning can sometimes feel like a black box. However, this is something that we've come to learn isn't entirely the case and there have been significant recent advancements in the field of "explainable AI" (e.g., McGovern et al. 2019).

Your recommendation regarding the relationship between the forward model and the NN retrievals is an interesting one. We are indeed using the same forward model in both cases. However we have not looked at this sort of inverse problem analysis

because the greatest source of error we expect from this retrieval comes in the from assumption biases (3-D radiative effects, and other un-modeled features in the data) and less from forward modeling biases themselves.

17) P18, L13-15, in reference to sentence "An evident feature . . . below each PDF". I assume you were referring to the re histograms here. It's possible the "spikes" in the NJK histogram are actual features of the distribution of microphysical properties because I also see similar features in the RFT and PP (lower bin resolution) histograms.

We agree, the current network implementation poorly generalizes from the limited LUT grid. We have modified this approach for our future studies to randomly sample a continuous distributions of physical parameters, but we have not made this change for this particular study.

18) P20, L15. "I guess "the behavior of" should not be there." Thank you, we have corrected this in the revision.

19) P21, L4-5. This was perplexing, but I think the explanations for it are two-fold. First, it could be that mixing total and polarized reflectance information is still not working well in our network – the polarized reflectance information is still not weighted strongly enough. Second, we think that this could be a result of the network lacking an explicit understanding of the joint relationship between the spectral and angular dependence. Without this shared dependence the network is treating measurements at different wavelengths and angles as independent pieces of information without necessarily relating them to one another. To this end, our future work on this topic will likely involve a convolutional neural network approach, where the multi-angle and multi-spectral dependence of the training data can be expressed as an image – explicitly providing the neural network with context for the joint dependence of the dataset on spectral and angular information.

As mentioned previously, we were attempting to avoid a total or polarized reflectance only retrieval because our final objective (aerosols above clouds) will require us to disentangle information from both total and polarized reflectance. Perhaps for the microphysical portion of this problem though, it may make sense in the future to perform the re/ve retrieval on polarimetric only measurements.

20) P22, L1. SWIR bands also increase in brightness with increasing optical thickness, but they are generally less sensitive to this change than the VNIR bands. They also saturate at lower optical thicknesses. Later in the paper, I was attempting to refer to the other wavelength bands that are not normally used in the bispectral retrieval (0.410 - 0.555 μm). We will attempt to clarify the sentence that says that in the revised paper to avoid future confusion.

21) P25, L12, "I guess there should be an “of” between “impact” and “atmospheric absorption” " Thanks, this was corrected in our revision.

22) P28, L3. I guess “While” should be removed. We have removed while and replaced it with “Whereas”

23) P28, L17, statement “it lacks a clearly traceable relationship between observations and retrievals”. These are good points. We will revise the section to be less aggressive about our usage of physical traceability when discussing the Neural Network approach. We will attempt to make a subtler point in this section.

REFERENCES Reviewer Bishop, C. M. (1995), Neural Networks for Pattern Recognition, Oxford University Press.

Di Noia, A., et al. (2019), Retrieval of liquid water cloud properties from POLDER-3 measurements using a neural network ensemble approach, Atmos. Meas. Tech, 12, 1697-1716, doi: 10.5194/amt-12-1697-2019. Nocedal, J., and Wright, S. J. (1999), Numerical Optimization, Springer.

REFERENCES Author

Alexandrov, M. D., Cairns, B., Emde, C., Ackerman, A. S., and van Diedenhoven, B.: Accuracy assessments of cloud droplet size retrievals 10 from polarized reflectance

measurements by the research scanning polarimeter , Remote Sensing of Environment, 125, 92–111, 2012a.

Alexandrov, M. D., Cairns, B., and Mishchenko, M. I.: Rainbow Fourier transform, Journal of Quantitative Spectroscopy and Radiative Transfer, 113, 2521–2535, 2012b.

Miller, D. J., Zhang, Z., Platnick, S., Ackerman, A. S., Werner, F., Cornet, C., and Knobelspiesse, K.: Comparisons of bispectral and polarimetric retrievals of marine boundary layer cloud microphysics: case studies using a LES–satellite retrieval simulator, Atmospheric Measurement Techniques, 11, 3689–3715, 2018.

McGovern, Amy and Lagerquist, Ryan and John Gagne, David and Jergensen, G. Eli and Elmore, Kimberly L. and Homeyer, Cameron R. and Smith, Travis, Making the Black Box More Transparent: Understanding the Physical Implications of Machine Learning, Bulletin of the American Meteorological Society, 100, 11, 2175-2199, 2019, 10.1175/BAMS-D-18-0195.1

Segal-Rozenhaimer, Michal, D.J., Miller, K. Knobelspiesse, J. Redemann, B. Cairns, M.D. Alexandrov, Development of neural network retrievals of liquid cloud properties from multi-angle polarimetric observations, Journal of Quantitative Spectroscopy and Radiative Transfer, 220, 2018, 39-51, DOI:10.1016/j.jqsrt.2018.08.030.

Interactive comment on Atmos. Meas. Tech. Discuss., doi:10.5194/amt-2019-327, 2019.