A neural network approach to polarimetric observations of aerosols above clouds:

Design, demonstration, and comparison to existing algorithms

Daniel J. Miller[1*]
Michal Segal-Rozenhaimer[2,3]
Kirk Knobelspiesse[1]

1. NASA Goddard Space Flight Center
2. NASA Ames Research Center
3. Bay Area Environmental Research Institute

This research is supported by NASA programs:
NASA ORACLES Field Campaign
NASA Postdoctoral Program Fellowship

AGU Fall Meeting 2018 A33H-3239

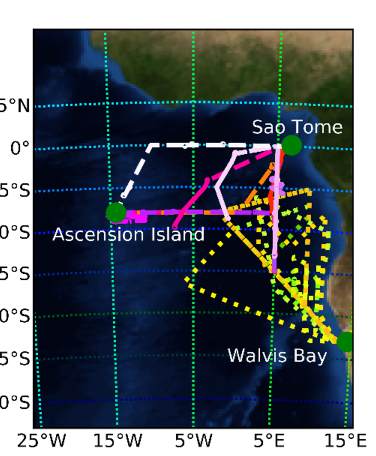
NASA POSTDOCTORAL PROGRAM

National Aeronautics & Space Administration

Introduction

The impact of aerosols on clouds constitutes one of the greatest sources of uncertainty in the understanding of Earth's climate. Above cloud aerosol (ACA) scenes in particular can impact the radiative budget (direct effect), cloud development (semi-direct effect), and microphysics (indirect effects). Passive remote sensing retrievals of ACA scenes is difficult because traditional retrieval approaches can be slow and complicated, due to the large state space exploration required. This is especially true for optimal inversion approaches, where a growth in the number of dependent variables can drastically complicate and slow the retrieval search. One way way to improve the speed and convergence of such retrievals is to provide a better 'first guess' obtained by a neural network (NN) for the retrieval to search around [1]. In this study we aim to develop and improve a neural network (NN) based algorithm for the retrieval of cloud properties in ACA scenes [2]. Our NN retrieval is applied to data from the airborne Research Scanning Polarimeter (RSP), which measures both polarized and total reflectances in nine visible and shortwave infrared bands, with each pixel observed from numerous viewing angles. We apply the NN algorithm to RSP field campaign data from the ObseRvations of Aerosols above CLouds and their intEractionS (ORACLES) 2016 and 2017 campaigns and compare to results obtained from other standard algorithms. We will evaluate these retrievals using ORACLES data satisfying the following criteria:

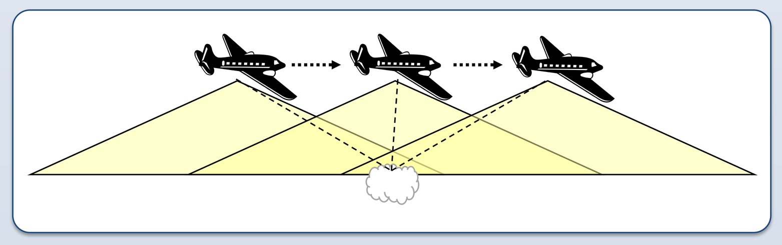
- Cloudy scenes as identified by other RSP retrieval methods
- Successful RSP retrievals using all other techniques
- Coincident RSP and HSRL data for cloud top height definition
- Instances with HSRL cloud top height below 2 km

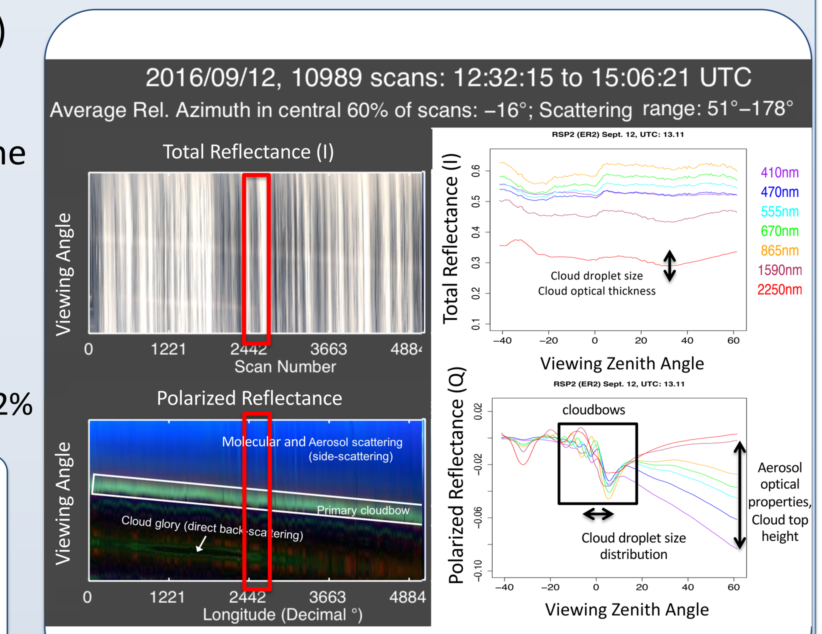


Dashed lines are for 2017 flight.
Dotted lines are for 2016 flights

The Research Scanning Polarimeter

- Continuous along track scanner (not an imager)
- Can "see" the same point from different views
- ±60° from nadir w/ 152 viewing angles per scene
- 9 bands in visible and shortwave infrared:
  - 410, 470, 555, 670, 864, 960, 1593, 1880, 2263 nm
- Simultaneous measurements of Stokes vector I (intensity), Q and U (2 linear polarization)
  - Measurement uncertainty: $\mathrm{d}I \cong 3\%$ & $\mathrm{d}DoLP \cong 0.2\%$

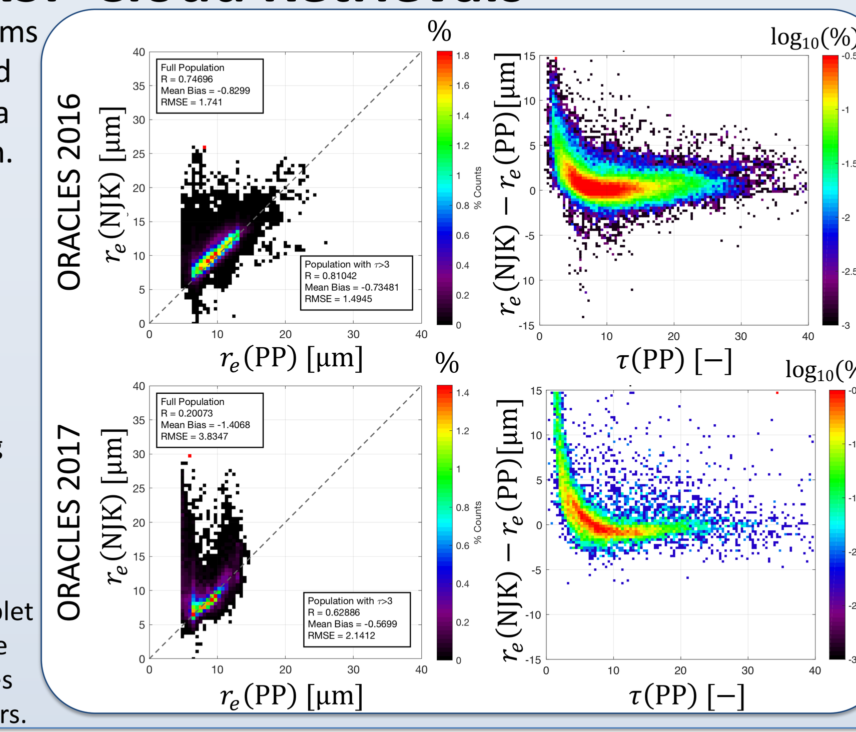




Standard RSP Cloud Retrievals

RSP has three standard cloud retrieval algorithms that all provide different perspectives on cloud retrievals [6,7]. Brief descriptions of each and a comparison of NJK and PP retrievals are shown.

- Nakajima-King: (NJK) [3]
  - A nadir total reflectance retrieval based on the well known Nakajima-King approach. A look up table of SWIR and VNIR band pairs are used to retrieve r_e and τ uniquely.
- Parametric Polarimetric: (PP) [4]
  - Curve fitting approach that fits single-scattering phase functions to polarized reflectances to retrieve distribution parameters r_e, v_e.
- Rainbow Fourier Transform (RFT) [5]
  - Resembles the Fourier transform between droplet size distribution space and polarized reflectance space using special basis set functions. Retrieves full droplet size distribution - not just parameters.



Neural Network Architecture

The disparate uncertainties of total and polarized reflectances from RSP observational (and training) data required us to weight network input relative to an instrument uncertainty model.

$$\widehat{x}_i(\vartheta_s,\vartheta_v,\varphi,\lambda) = \frac{x_i(\vartheta_s,\vartheta_v,\varphi,\lambda) - \bar{x}(\vartheta_s,\vartheta_v,\varphi,\lambda)}{\sigma\left(\bar{x}(\vartheta_s,\vartheta_v,\varphi,\lambda)\right)}$$

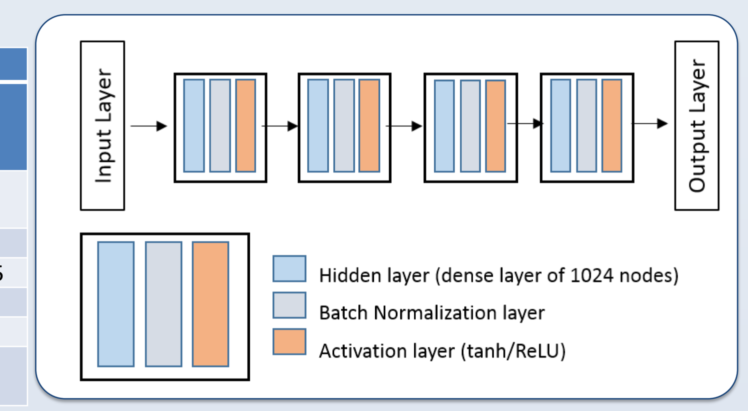
Parameters [units]	ORACLES 2016		ORACLES 2017	
	# of grid points	Training Grid	# of grid points	Training Grid
Altitude [m]	N/A	N/A	3	5000, 6000, 7000
r_e – [μm]	6	5, 7.5, 10, 12.5, 15, 20	6	5, 7.5, 10, 12.5, 15, 20
v_e [-]	6	0.01, 0.03, 0.05, 0.07, 0.1, 0.15	6	0.01, 0.03, 0.05, 0.07, 0.1, 0.15
τ [-]	6	2.5, 5, 10, 15, 20, 30	6	2.5, 5, 10, 15, 20, 30
SZA [°]	12	10 to 65 in increments of 5	13	5 to 65 in increments of 5
RAA [°]	17	0, 2, 4, 6, 8, 12, 16, 20, 24, 28, 32, 40, 50, 60, 70, 80, 90	31	0 to 90 in increments of 3

We experimented with different networks for the datasets available because observational platform differences (high altitude 2016 and low altitude 2017) required it. The primary difference between ORACLES 2016 and 2017 networks stems from the training set grids defined above.

Networks using a tanh activation function estimated r_e well (and τ poorly) and networks using a rectified linear unit (ReLU) activation function estimated τ well (and r_e poorly). As a consequence separate networks are used for each NN retrieval.

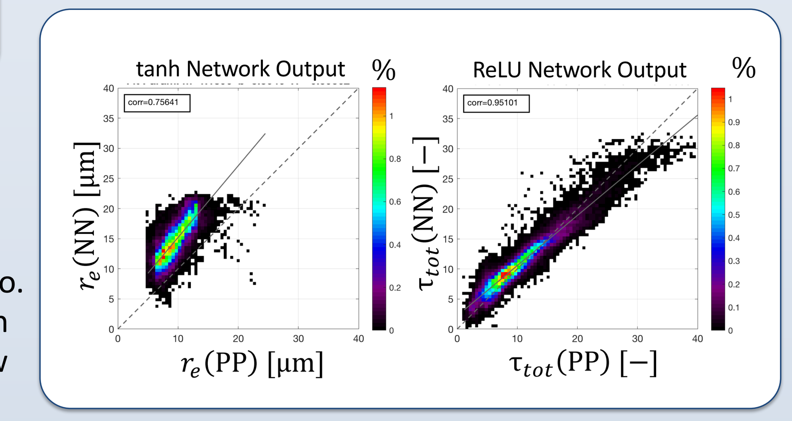
Parameter [units]	R^2	|Bias|	RMSE
τ [-]	0.999	0.016	.021
r_e [μm]	0.987	0.044	0.076
v_e [-]	0.941	0.094	0.16

The evaluation dataset showed that the network was well trained for r_e and τ retrievals, but not for v_e (RMSE=range/2).



- Deep network structure with four hidden layers
- Normalization layers re-standardize data
- Activation layer either tanh or ReLU
- Trained in a mini-batch mode with a rate of 0.0001 and 100 epochs for each training scenario.
- Network optimized using Adam algorithm within the Keras Python API making use of a TensorFlow backend.

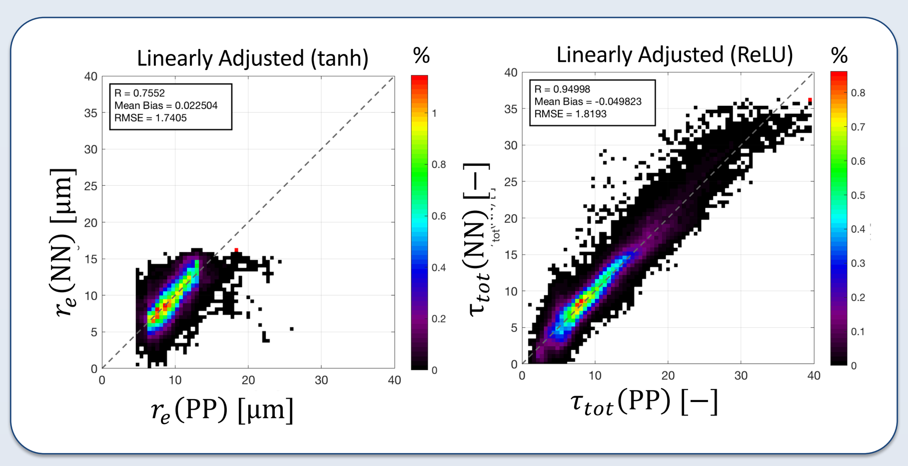
The initial output of the network exhibits clear linear biases when compared to the other RSP retrievals. This linear offset was absent during our training evaluation exercise, and the high correlations of these retrievals imply that the NN retrieval is otherwise generally performing correctly. The source of this linear offset is still an open question, though the current hypothesis is that it is associated with the preprocessing of input data.

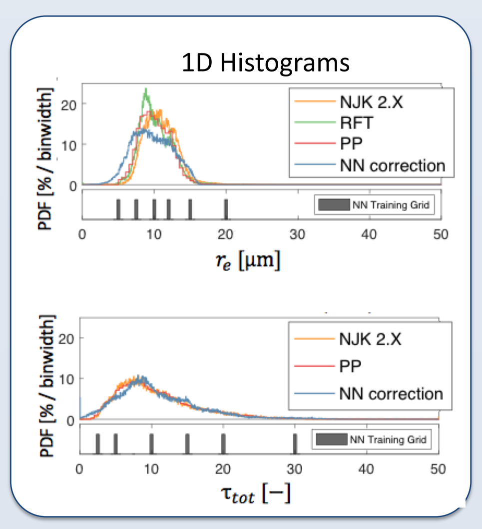


ORACLES 2016

Density regression plots of the linearly scaled NN retrievals from the ORACLES 2016 dataset reveal strong statistical evidence that the NN retrieval behaves well on average.

In fact, these correlation and RMSE characteristics are similar to those observed when comparing two different standard RSP cloud products (shown previously).

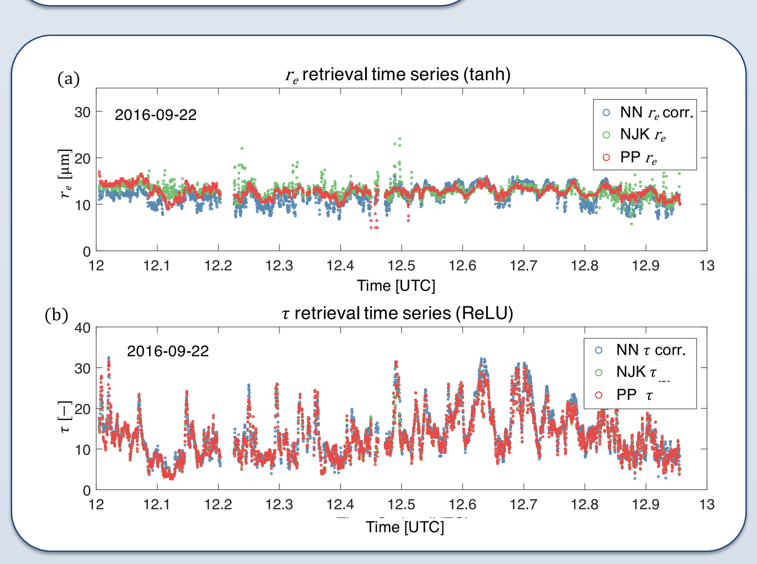


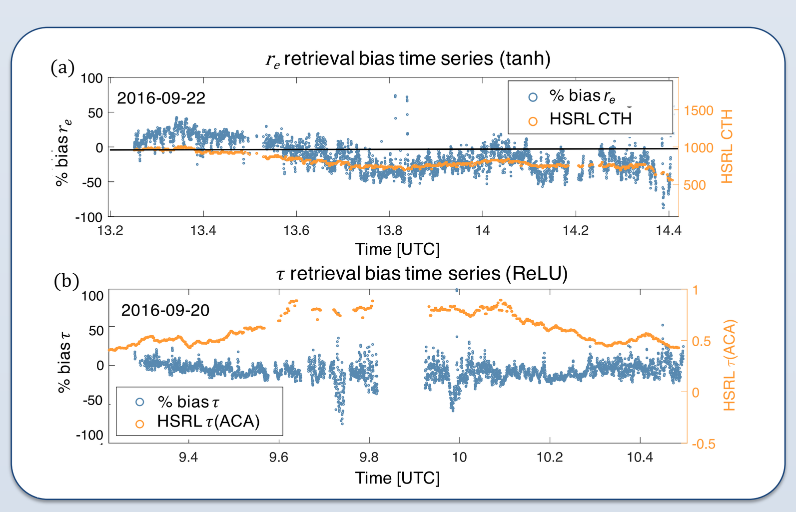


The histograms of each retrieval reveals more detail, retrievals of τ (and to a lesser extent, r_e) are clustering around the locations of training set grid points.

The time series below show that the spatial variability of the NN retrieval behaves similarly to the existing algorithms.

One useful finding was that Found that results for re retrieval depended on cloud top height. This inspired us to change the training set for the network training set during the analysis of the 2017 dataset when the RSP instrument would be flying at variable altitudes. Additionally, there was a less consistent and weaker dependence of the τ retrieval on the above cloud aerosol optical thickness.

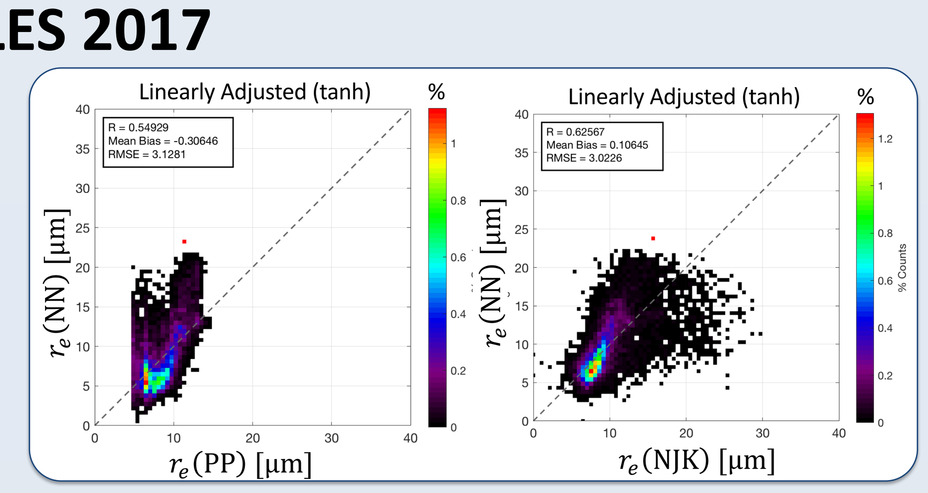


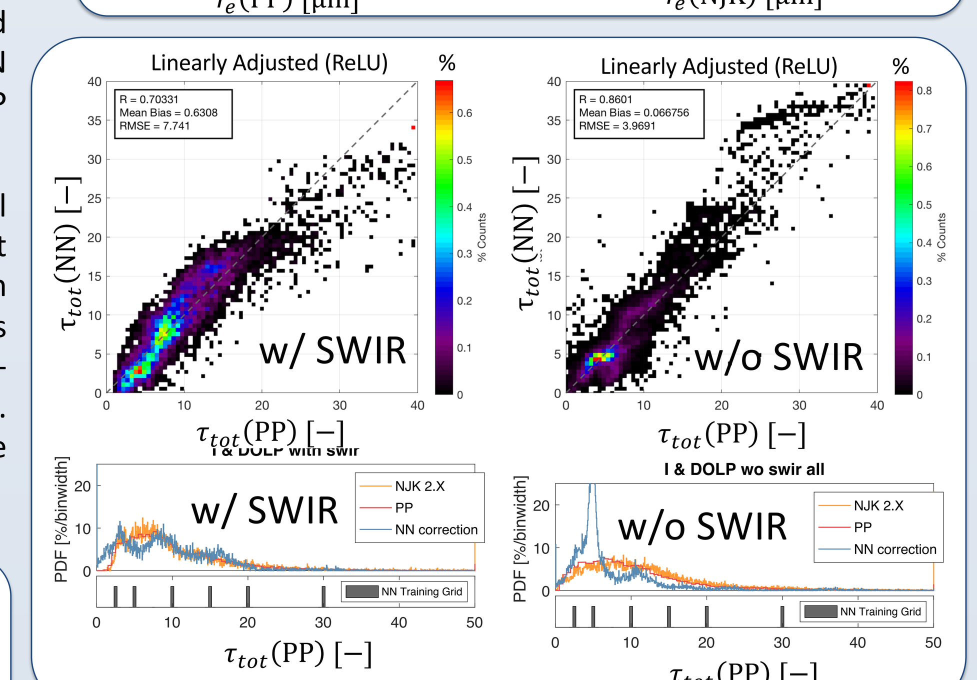


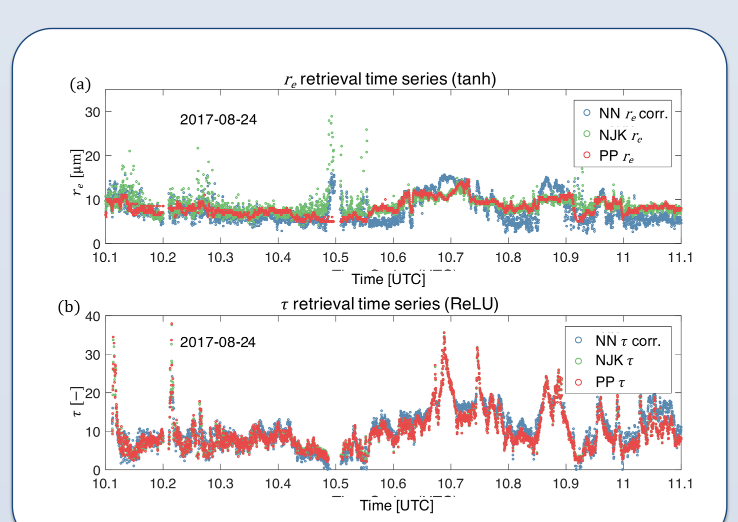
ORACLES 2017

The NN retrievals perform more poorly for the ORACLES 2017 dataset despite changes intended to improve results. Regression of the linearly scaled NN retrievals from ORACLES 2017 dataset reveal more complicated and nonlinear relationships than those observed for 2016.Most notably though, the comparison of the r_e retrievals to PP or NJK retrieval shows different behaviors. Because the ORACLES dataset 2017 dataset contains more instances of low-τ, it is probably that the issue here is related to biased total reflectance that result in biased NN retrievals that ingest them (whereas the PP retrieval uses no polarized reflectance.

During ORACLES 2017 RSP often lacked useful SWIR data, as a consequence we explored what the consequences of excluding SWIR data from training and input. We found that r_e retrievals were very poor, but τ retrievals were less non-linear but highly clustered around training bins. Perhaps with a better training grid, SWIR-free NN retrievals may have performed better.







The example time series for ORACLES 2017 exhibits a poorer spatial variability match to RSP retrievals This is largely attributable to deviations that occur near dramatic increases/decreases in τ or gaps in the cloud layer. Both of these could be a possible indication of the influence of inhomogeneous clouds and 3D radiative effects on total reflectances that are subsequently driving NN retrieval biases.

Summary & Outlook

Summary of Findings:

- We got something that works pretty well despite not being trained with ACA layer information.
- Appears to work better for 2016 data than for 2017 data, the change in the population of cloud properties (more thin and more broken clouds in 2017) is likely associated with this behavior.
- The NN behaves more like polarimetric retrieval when clouds are homogeneous, and when clouds are broken and inhomogeneous the retrieval behaves like the NJK retrieval.

Strengths and Limitations of our Approach:

- We created an algorithm that attempts to meld both total and polarized reflectance information into the same retrieval, to date there is no significant effort to hybridize retrievals of cloud microphysics like this. As a consequence our results are somewhere between polarimetric results and bispectral results.
- The need for the linear correction using a validation dataset is concerning. During evaluation of the training dataset we did not observe this.
- One of the strengths of a NN retrieval is that it can explore retrievals in situations where other retrievals are not performed for one reason or another – the linear scaling requirement hurts that argument.
- What we are comparing to is *not* truth and what we are training with is not considering a large component of the observed system (the presence of the aerosol above the cloud).

Future Research Goals:

- Embed an aerosol above cloud layer in the NN training set and train a network to retrieve aerosol and cloud properties.
- Use these NN retrievals as first guess estimates to accelerate an optimal estimation above cloud aerosol retrieval.

Citations

1. Di Noia, A., O. P. Hasekamp, L. Wu, B. van Diedenhoven, B. Cairns, and J. E. Yorks (2017), Combined neural network/Phillips-Tikhonov approach to aerosol retrievals over land from the NASA Research Scanning Polarimeter, *Atmos. Meas. Tech. Discuss.*, 1–27, doi:10.5194/amt-2017-148.
2. Segal-Rozenhaimer, M., D. J. Miller, K. Knobelspiesse, J. Redemann, B. Cairns, and M. D. Alexandrov (2018), Development of neural network retrievals of liquid cloud properties from multi-angle polarimetric observations, *Journal of Quantitative Spectroscopy and Radiative Transfer*, *220*, 39–51, doi:10.1016/j.jqsrt.2018.08.030.
3. Nakajima, T., and M. D. King (1990), Determination of the Optical Thickness and Effective Particle Radius of Clouds from Reflected Solar Radiation Measurements. Part I: Theory, *J. Atmos. Sci.*, *47*(15), 1878–1893, doi:10.1175/1520-0469(1990)047<1878:dotota>2.0.co;2.
4. Alexandrov, M. D., B. Cairns, C. Emde, A. S. Ackerman, and B. van Diedenhoven (2012), Accuracy assessments of cloud droplet size retrievals from polarized reflectance measurements by the research scanning polarimeter, *Remote Sensing of Environment*, *125*, 92–111, doi:10.1016/j.rse.2012.07.012.
5. Alexandrov, M. D., B. Cairns, and M. I. Mishchenko (2012), Rainbow Fourier transform, *Journal of Quantitative Spectroscopy and Radiative Transfer*, *113*(18), 2521–2535, doi:10.1016/j.jqsrt.2012.03.025.
6. Miller, D. J., Zhang, Z., Platnick, S., Ackerman, A. S., Werner, F., Cornet, C., and Knobelspiesse, K.: Comparisons of bispectral and polarimetric cloud microphysical retrievals using LES-Satellite retrieval simulator, Atmos. Meas. Tech., https://doi.org/10.5194/amt-2017-325, in review, 2017.
7. Miller, D. J., Z. Zhang, A. S. Ackerman, S. Platnick, and B. A. Baum (2016), The impact of cloud vertical profile on liquid water path retrieval based on the bispectral method: A theoretical study based on large-eddy simulations of shallow marine boundary layer clouds, *J. Geophys. Res.*, *121*(8), 4122–4141, doi:10.1002/2015JD024322.