Neural Network (NN) retrievals of Low level Liquid cloud properties from multi-angle polarimetric observations during ORACLES 2016-17

Michal Segal Rozenhaimer[1,2], Kirk Knobelspiesse[3], Jens Redemann[1], Brian Cairns[4], and Daniel Miller[3]

[1]NASA Ames Research Center, Moffett Field, CA 94035, USA, [2]BAER Institute, Moffett Field , CA 94035, USA, [3]NASA GSFC, Maryland, USA, [4]NASA GISS, New-York, USA

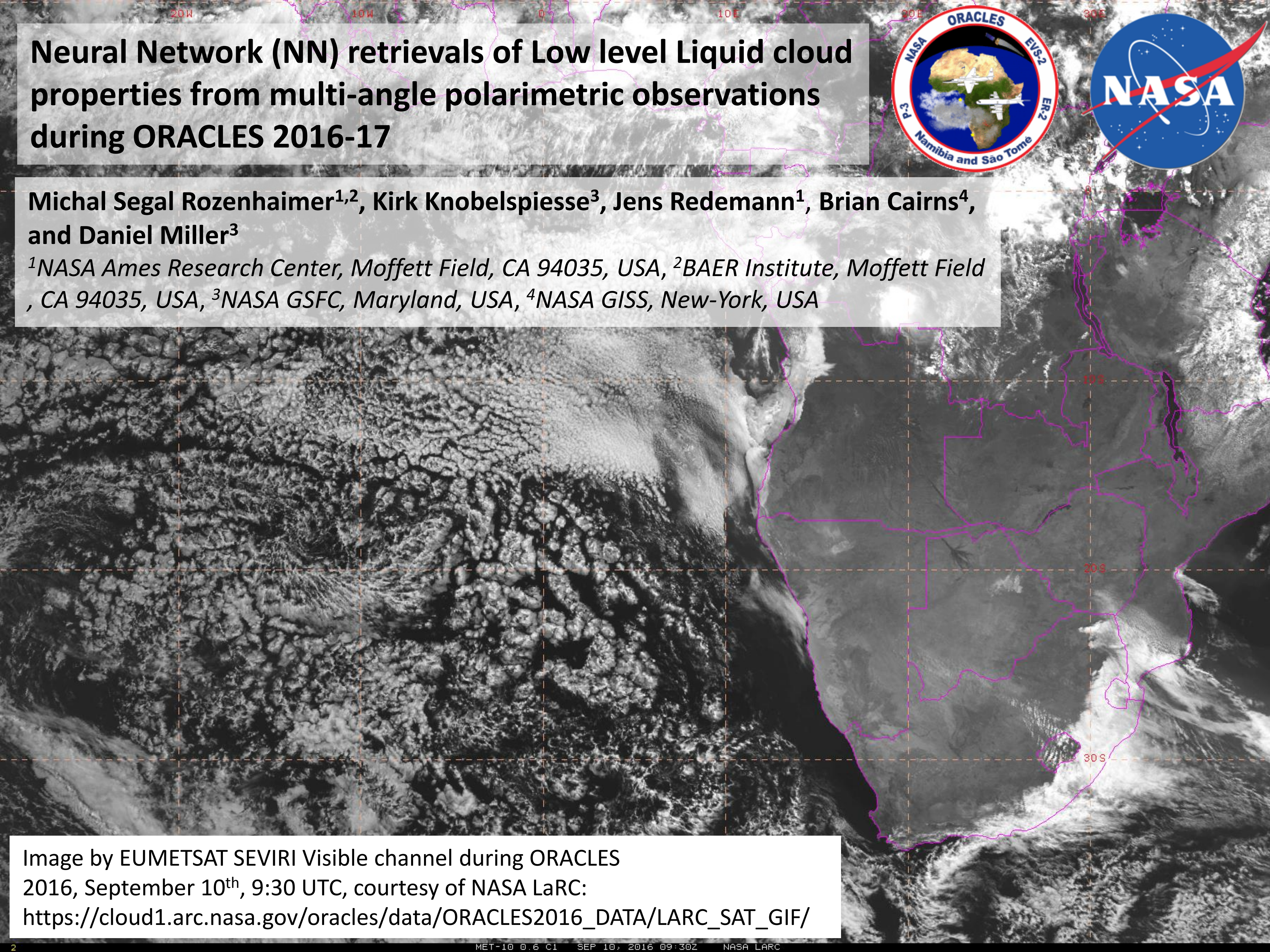


Image by EUMETSAT SEVIRI Visible channel during ORACLES 2016, September 10th, 9:30 UTC, courtesy of NASA LaRC: https://cloud1.arc.nasa.gov/oracles/data/ORACLES2016_DATA/LARC_SAT_GIF/

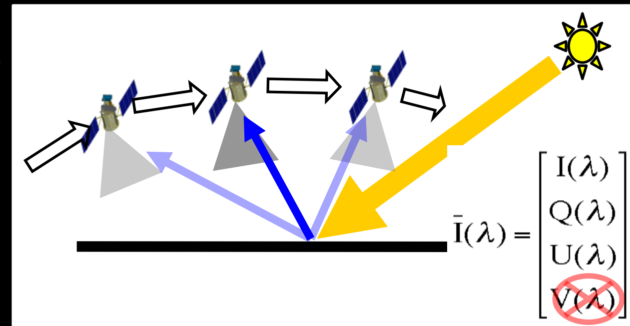
Outline

- *Multi-angle polarimetric measurements during ORACLES*
- *Goals*
- *Retrieval approach – Neural Network*
- *Retrieval approach – implementation and different schemes*
- *Validation results from the two different schemes*
- *Results from ORACLES 2016-2017*
- *Summary, Conclusions & Future work*

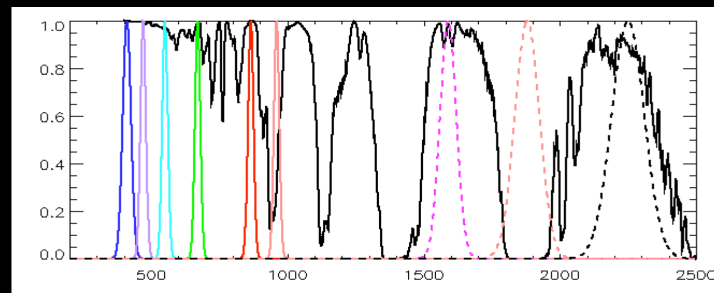
Research Scanning Polarimeter (RSP) in ORACLES

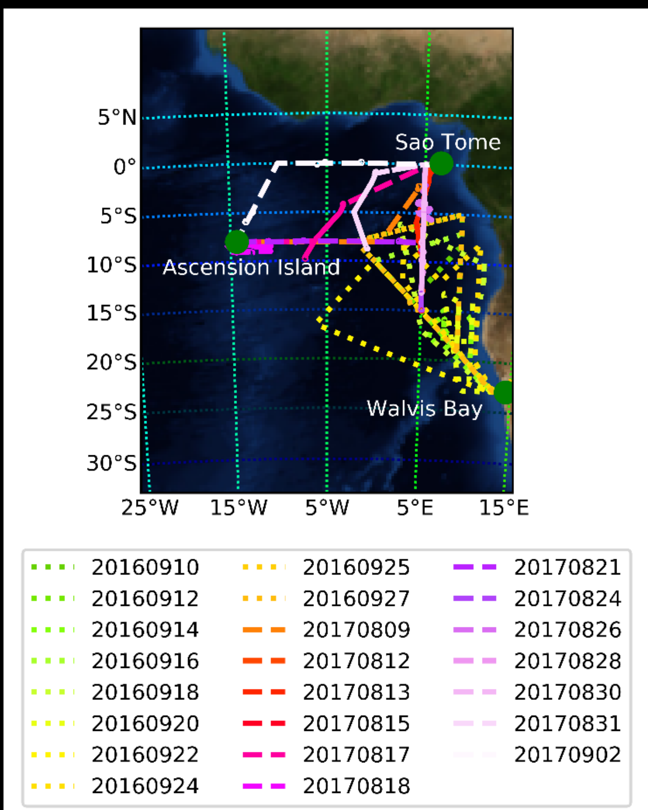
<u>RSP: multi angle, passive polarimeter</u>

152 view angles x
9 (7 clouds) spectral channels x
3 polarization states =
~ 4000 observations per pixel
(but highly correlated)



Liquid clouds products
Droplet size distribution at top of cloud
Droplet number concentration
Liquid water path
Cloud optical thickness





ER-2 - 2016

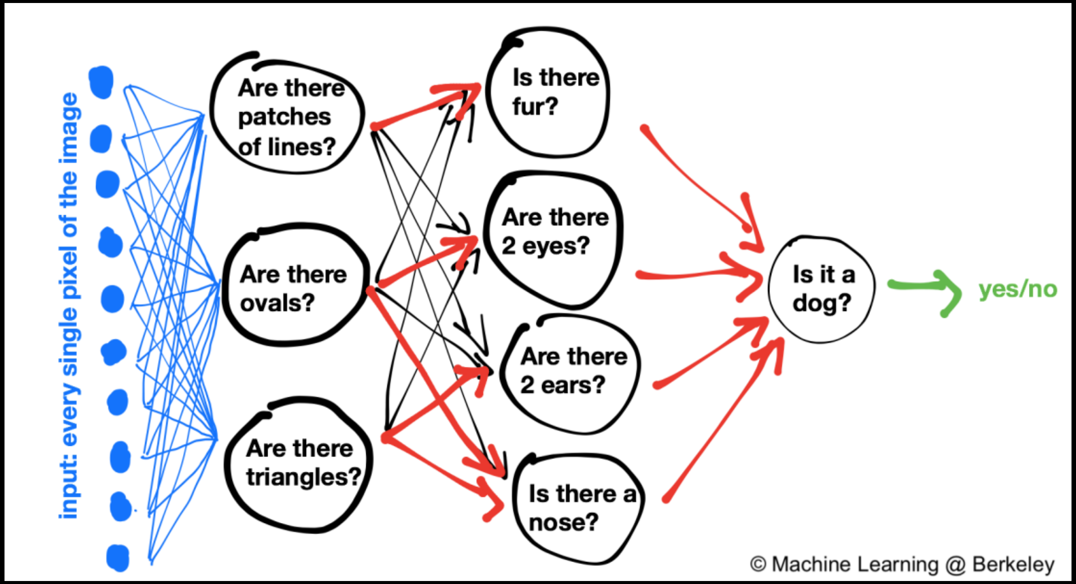
P-3 - 2017

Goals

- *To develop a new and combined retrieval from multi-angle polarimetric measurements for low-level clouds and above cloud aerosols (ACA)*
- *Our current work focuses on the low-level cloud scheme (better validation opportunities, allows to develop insights on the algorithm bahaviour)*

Retrieval approach – Neural Networks (NN)

What is a Neural Network?



Why Neural Network?
(Advantages/disadvantages)

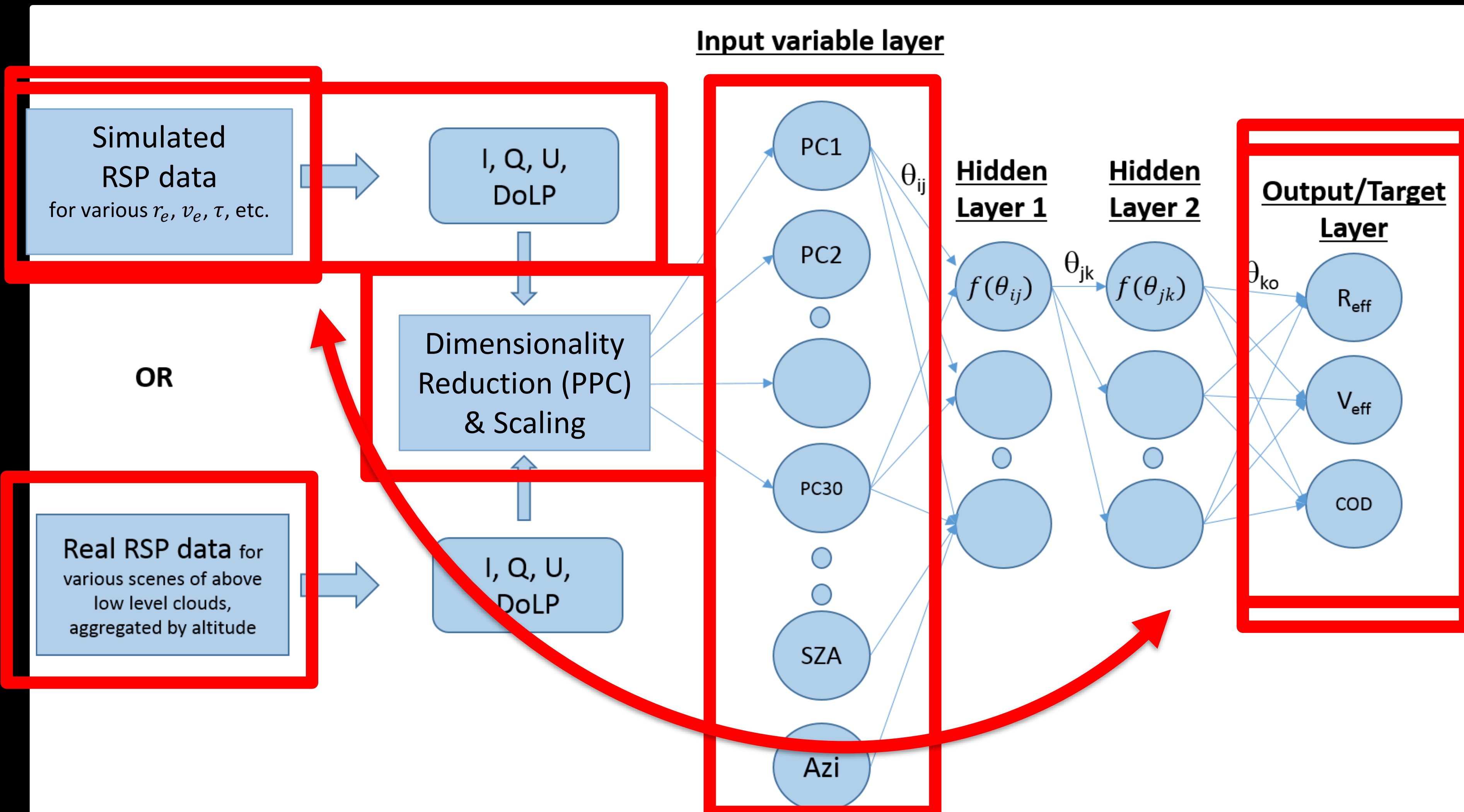
- Can represent any type of function, especially non-linear ones, given enough parameters (layers and nodes), where the training process creates a “transfer function” between inputs and outputs.
- Hard to assess and interpret the network’s function in absolute physical meanining.
- For continuous output networks, can interpolate and extrapolate beyond the training set LUT, unlike best fit methods.
- If convergence fails or results don’t do well, hard to assess and needs a lot of trial/error tweaking.
- Gaining large popularity among the remote sensing community (e.g., Del Frate et al., 2005; Cardena et al., 2007;   Milstein and Blackwell, 2015; Di Noia et al., 2015), as it can sometime supersede purely physical based model retrievals (especially in remote sensing imagery interpretation).

Retrieval approach – NN scheme (stage I)

(1) Creating training set (simulated data) (2) Preparing the inputs (dimensionality reduction)

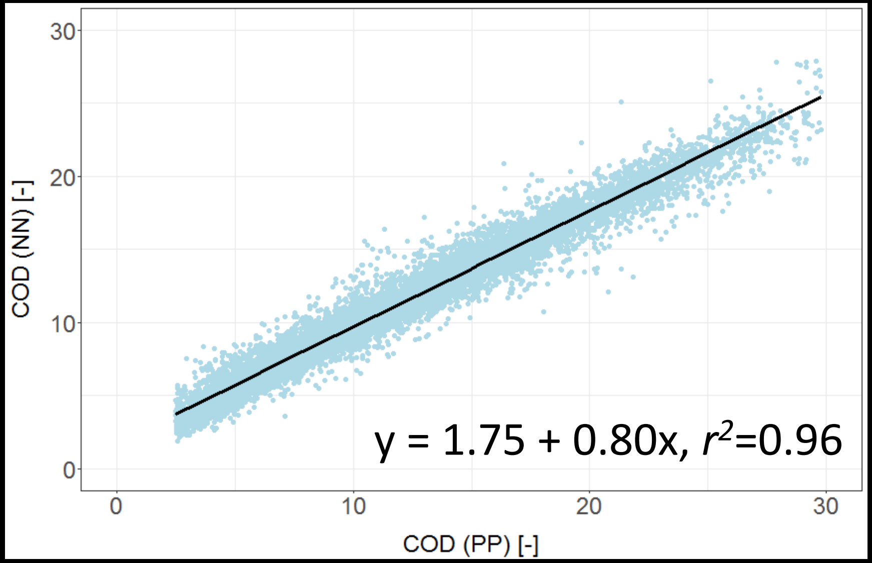
(3) Training (optimizing network weights and architecture) – train/validation/test sets

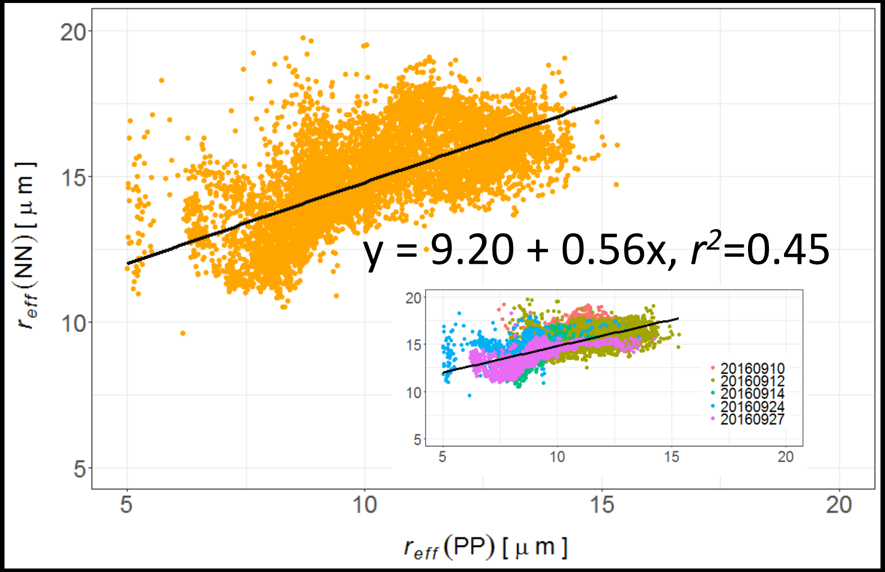
(4) Application of the network on real data (prediction)



Validation of NN scheme (Stage I)

We validated our NN with field observations from ORACLES 2016 ER-2 observations that use the standard RSP algorithms:
Parametric (PP), Nakajima-King (NK), and Rainbow Fourier Transform (RFT)





Segal-Rozenhaimer et al. (2018), *submitted to JQSRT*

- Good correlation and almost no bias with COD
- Lower correlation and positive bias with r_{eff} (NN behaves more like NK method and gives more weight to the total reflectance input)
- The reason why more weight is given to the total reflectance stems from the dimensionality reduction procedure (not taking into account the true uncertainty of each of the inputs)

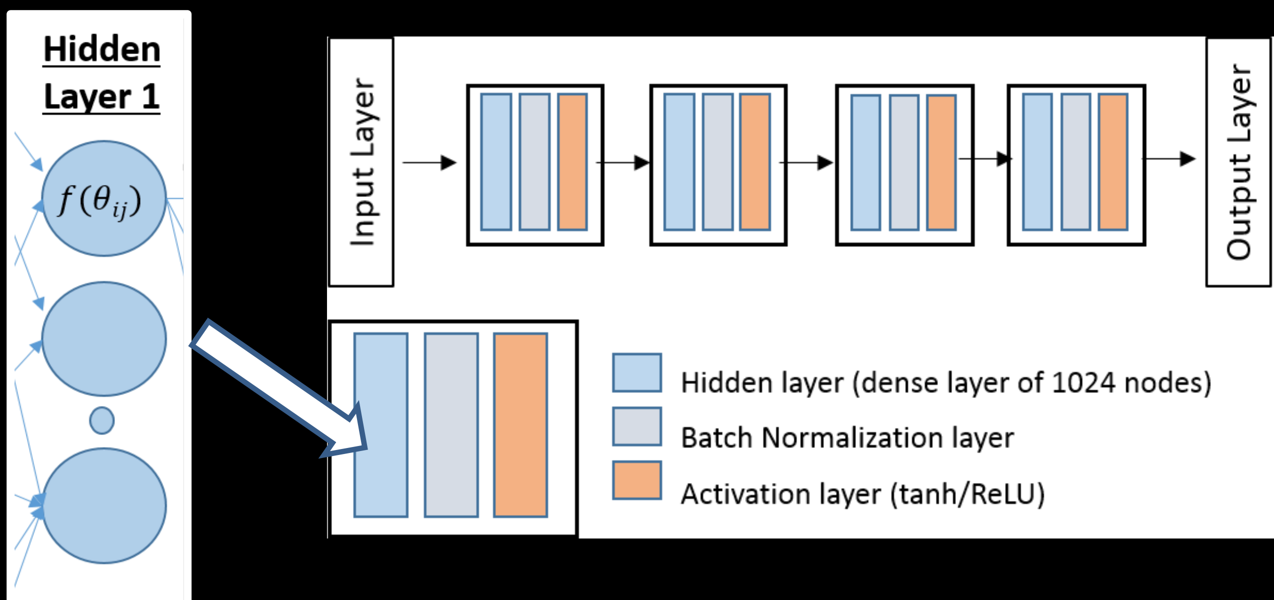
Retrieval approach – NN scheme (stage II)

- **Goal**: Improved uncertainty representation of each input (I – 3% vs. DoLP or Q – 0.2%)
- **Method**: data is now standardized by subtracting the mean measurement value divided by the expected measurement uncertainty for that mean over a given range of geometries:

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

(Knobelspiesse et al, 2018, *in prep.)*

- **Modifications needed for the new NN scheme**:
  - ✓ 1024 nodes in each layer (4) Instead of 40 nodes in 2 layers
  - ✓ Better optimization and gradient descent schemes (TensorFlow/Keras)
  - ✓ ReLU activation function (faster computation)
  - ✓ Mini-batch training (instead of online, i.e. updating weights after each training point)



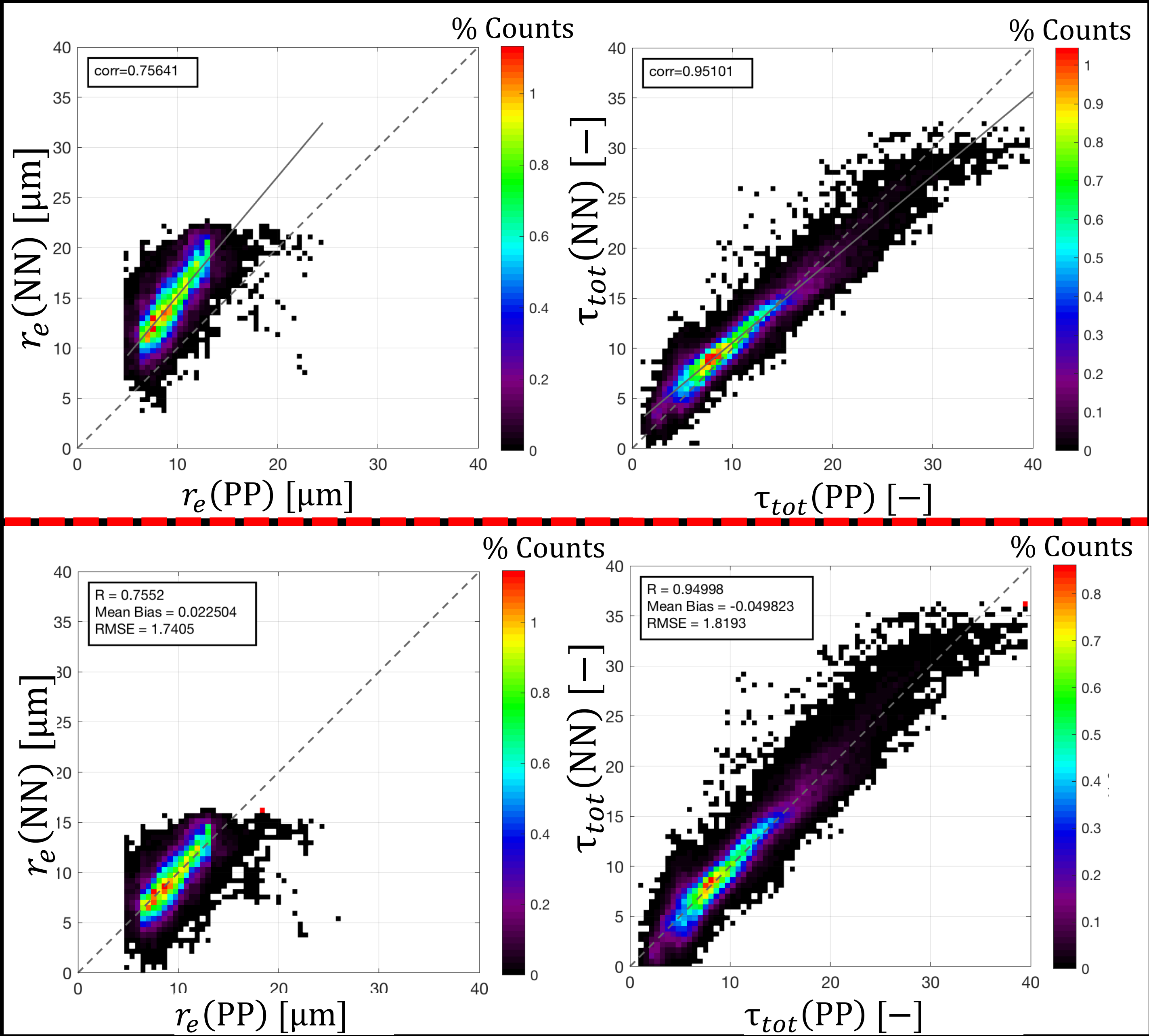
Validation of NN scheme (Stage II)

A linear rescaling is applied to raw output to correct a side-effect of our approach to standardization (weighting by uncertainty).

Original Output

Linearly Corrected

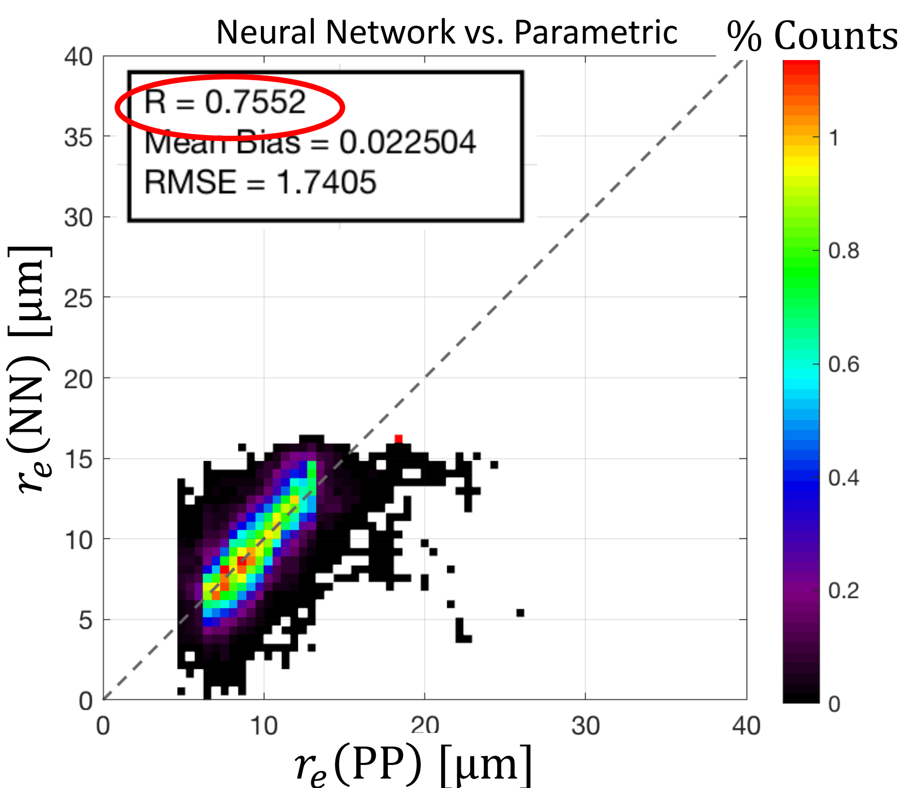
This correction is obtained from a linear fit of a partial subset of observations. Afterwards, it is applied to the full dataset (2016 ER2).

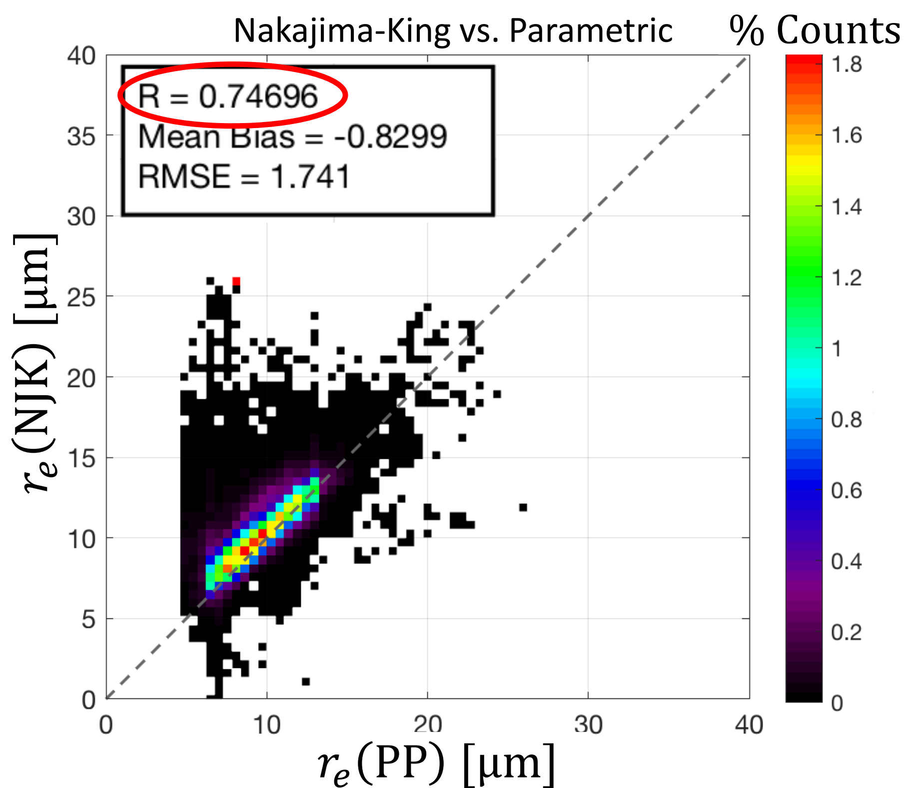


Evaluation

How do we do relative to other retrieval comparisons?

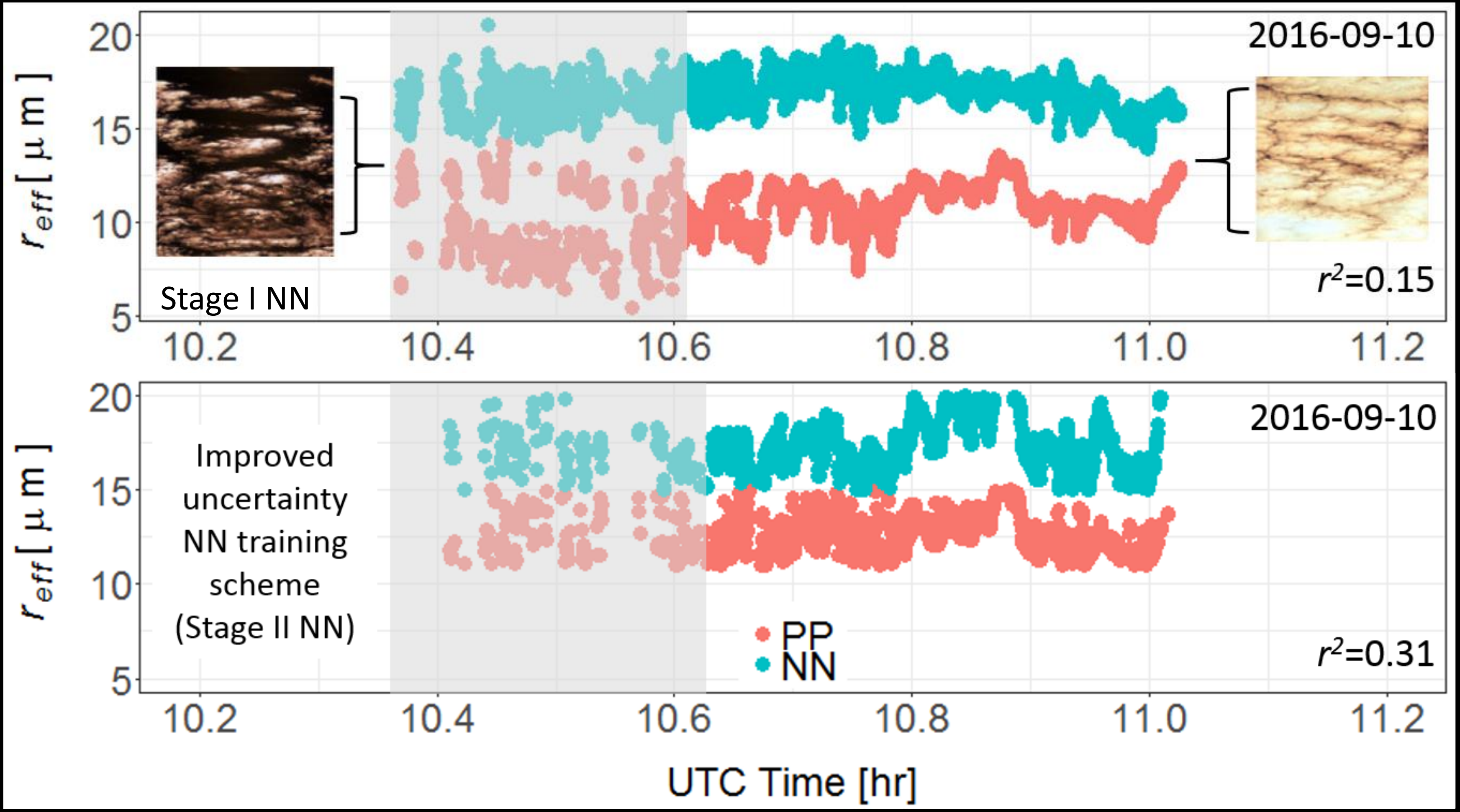
- Retrieval comparison is a tricky business
  - Different state space sensitivities...
  - Different uncertainties...
- Relative to other retrieval comparisons (NJK vs. PP) we have similar performance.



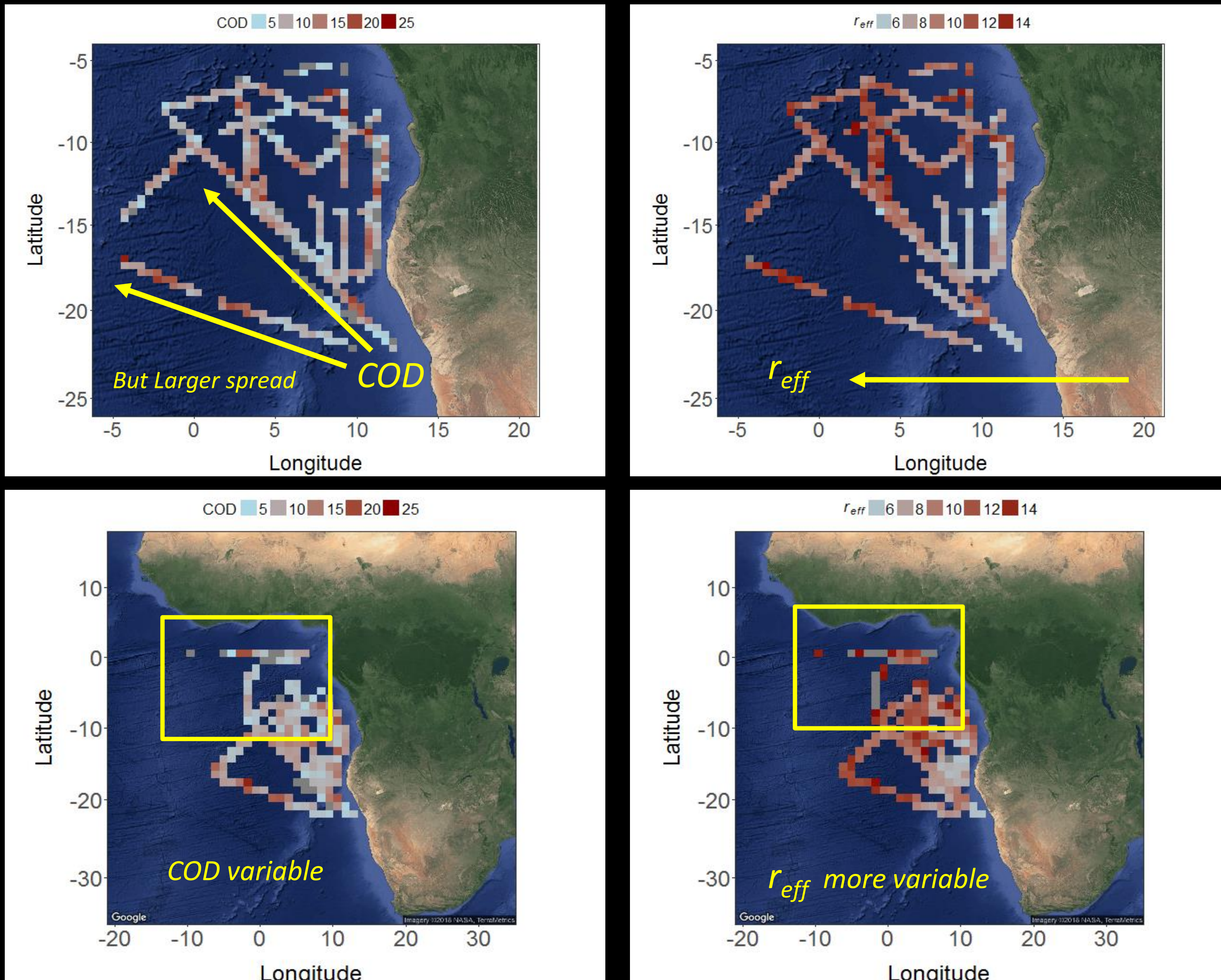


Comparing NN schemes (Stage I vs. Stage II)

Time Series results from Stage I NN scheme



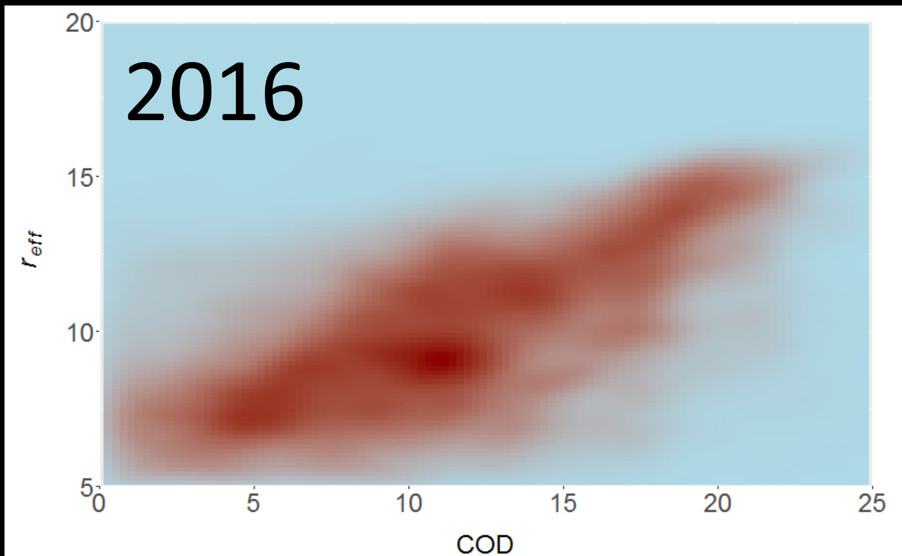
Results from ORACLES 2016-2017

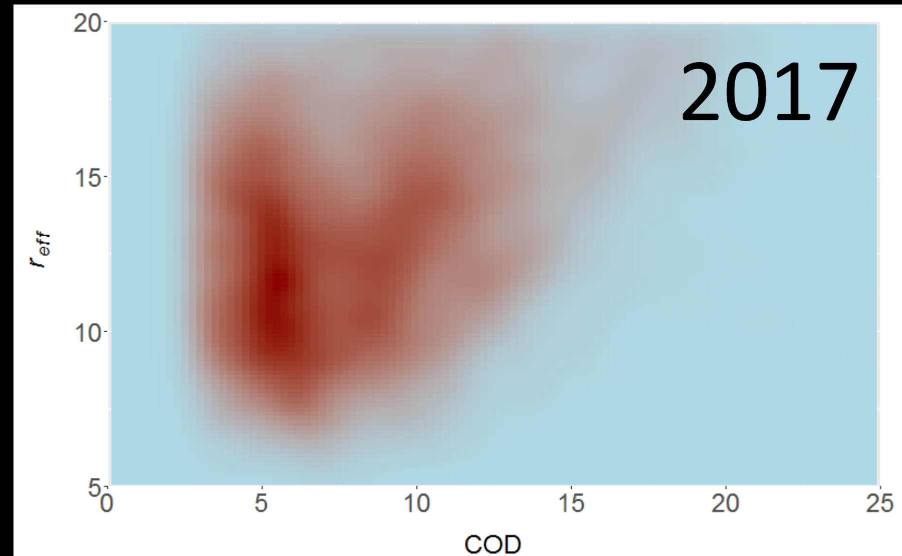


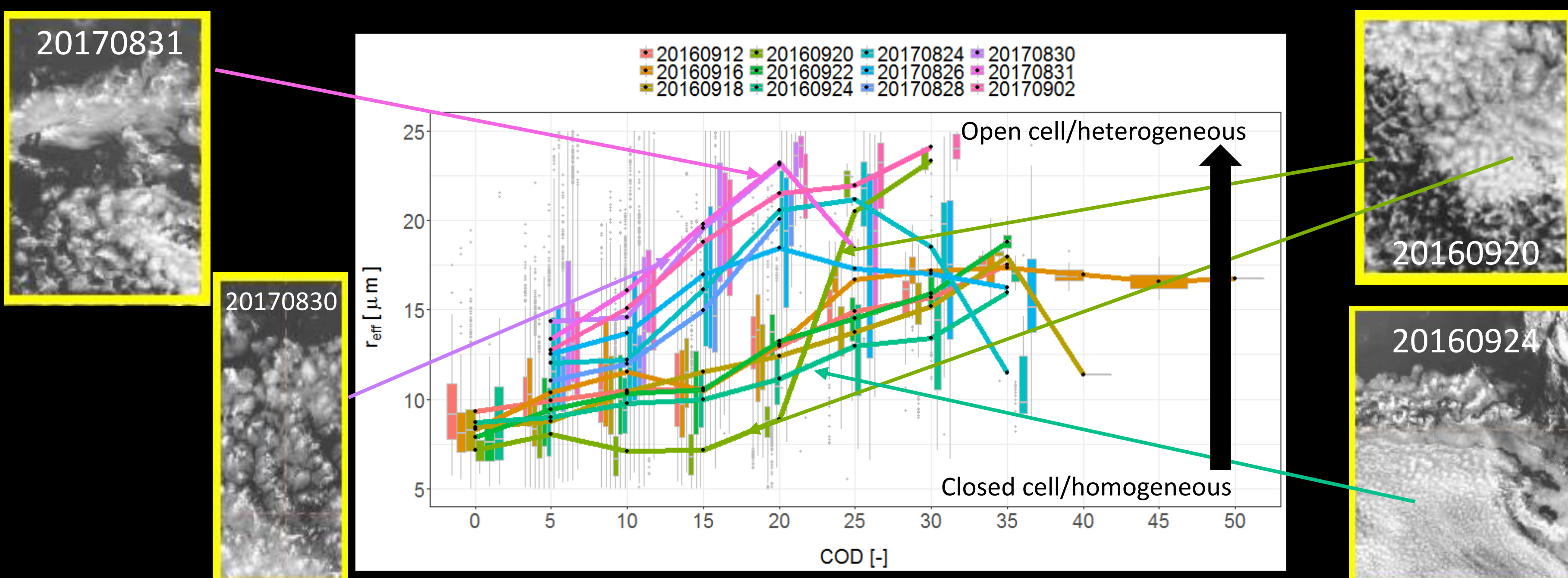
2016

2017

Results from ORACLES 2016-2017







Summary & Conclusions

- A Neural Network was trained, tested and validated for retrievals of liquid cloud property from RSP measurements.
- Application of the correct uncertainty model for the various inputs had an effect of goodness of the retrieval scheme (latter scheme is better).
- Comparison with standard RSP cloud products using ORACLEs data show good agreement between the two methods.
- ORACLES 2016 show increase of r_{eff} further from the Namibian coast, while COD show increase as well, but is more variable overall.
- There seem to be a correlation between r_{eff} and COD for both 2016 and 2017, but the correlation strength and slope depends on cloud macrostructure (higher variability in r_{eff} is seen in open cell forms)

Future work

- Developing Above Cloud Aerosols NN scheme: simulations will probably use the cloud only training set as a ‘surface’ under the aerosols, and a new approach of 2D inputs (wavelength x VZA) will be tested.

Thanks!

Data products can be found here:

Stage I NN:

http://data.giss.nasa.gov/pub/rsp/ORACLES_2016/NeuralNetworkCloud/

Stage II NN:

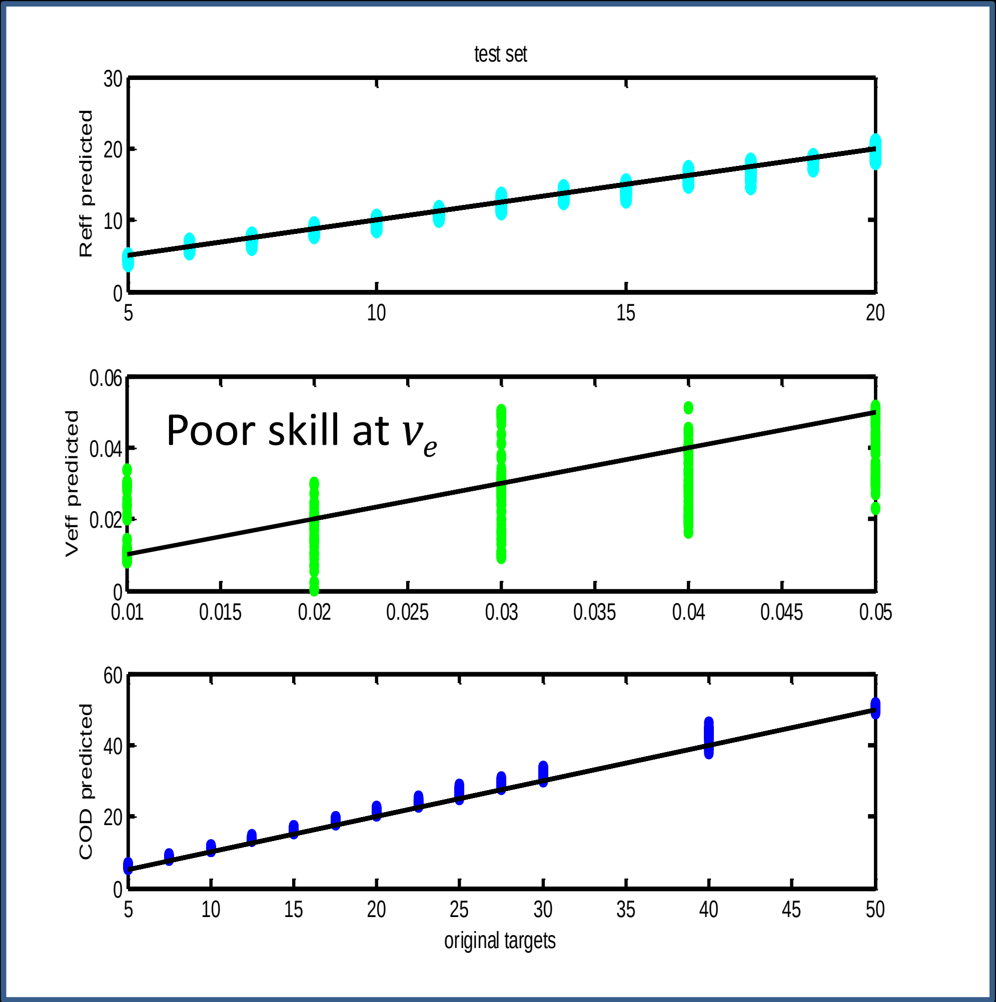
http://data.giss.nasa.gov/pub/rsp/ORACLES_2016/

Backup slides

Implementation

Step IV – optimize network inputs & weights

to get the best target predictions on the training and test sets – tested with reserved simulations not in training set



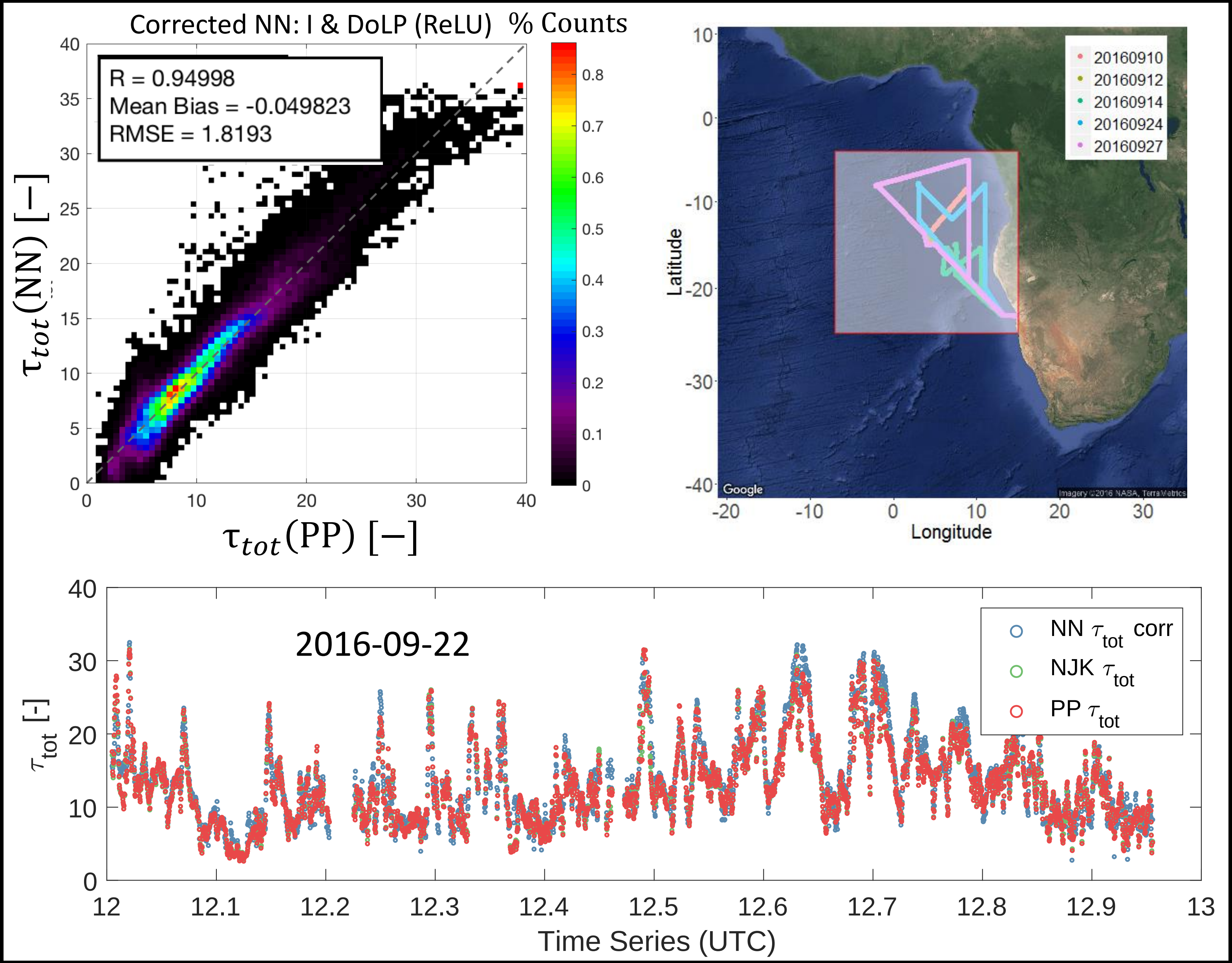
Input Label	# input variables	R_{eff} RMSE	V_{eff} RMSE	COD RMSE
Ri	30	1.01	0.016	2.21
Qi	30	0.93	0.008	9.04
Rp	20	0.74	0.006	10.85
DoLP	100	0.54	0.006	1.71
Ri-Qi	60	0.78	0.010	0.45
Ri-Rp	50	0.60	0.009	0.97
Ri-DoLP	130	0.37	0.006	1.16
Qi-Rp	50	0.80	0.005	6.88
Qi-DoLP	130	0.35	0.004	0.81
Rp-DoLP	120	0.45	0.004	1.18

Here we test the value of different ways of slicing & dicing the data.

This helps us explore best way to manage measurement uncertainty

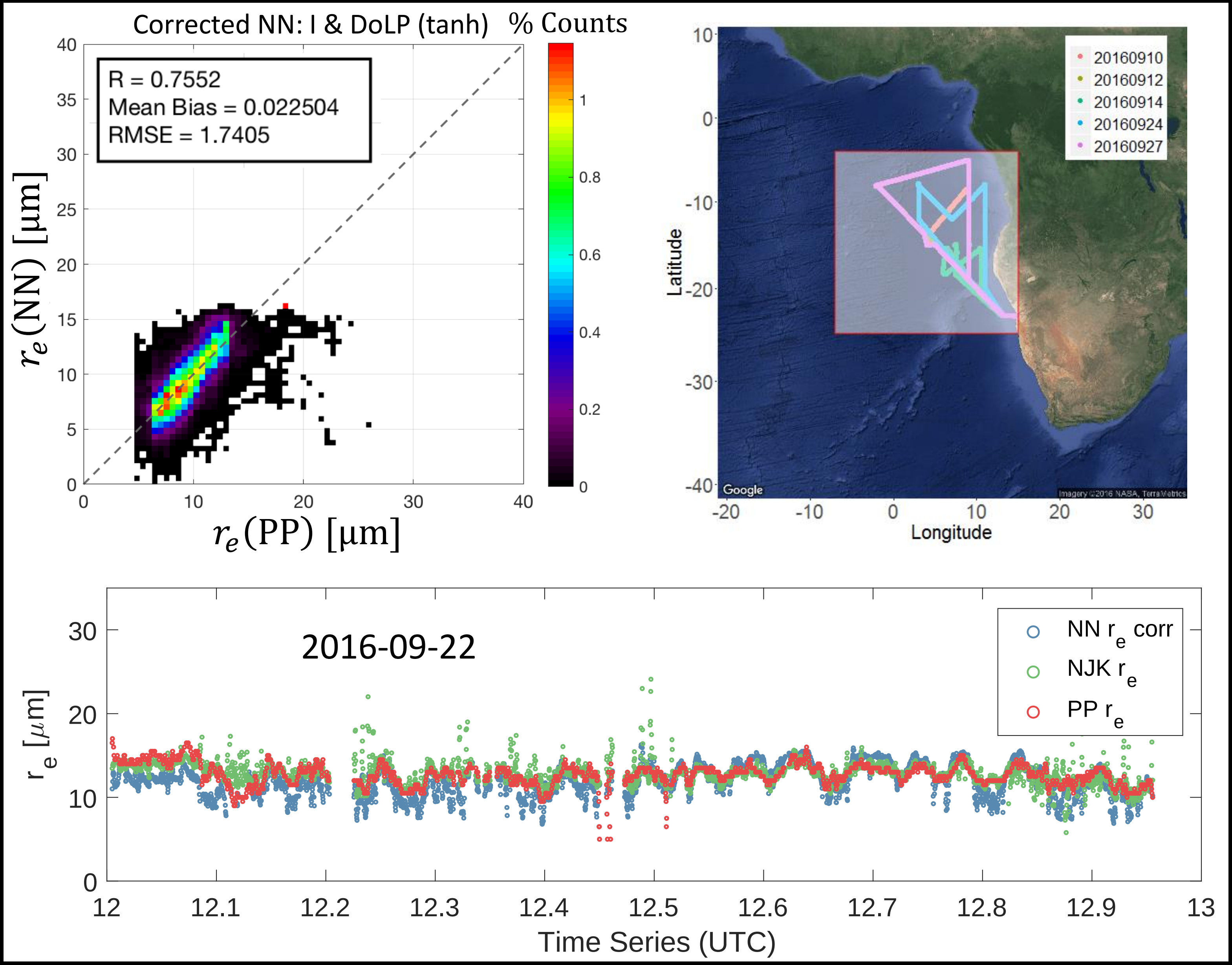
Step VI – Retrievals & Validation: Cloud Optical Thickness

Validate with field observations that use standard RSP algorithms



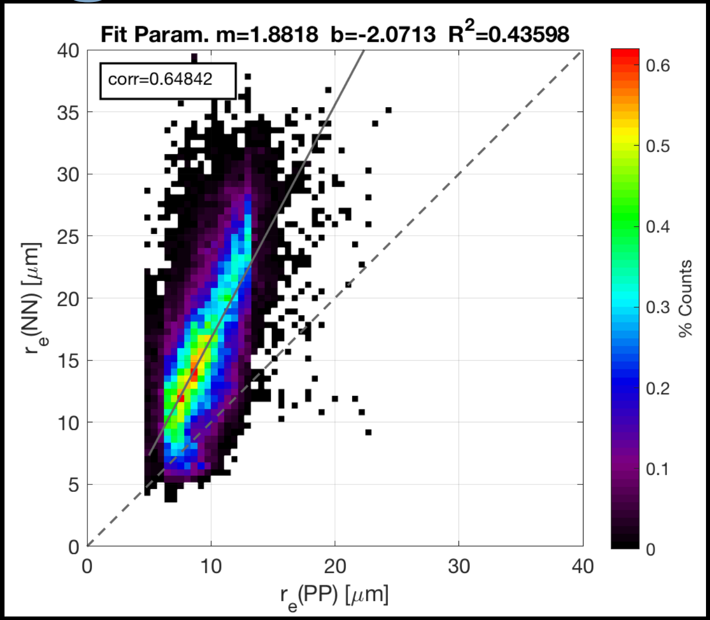
Step VI – Retrievals & Validation: Cloud Droplet effective radius

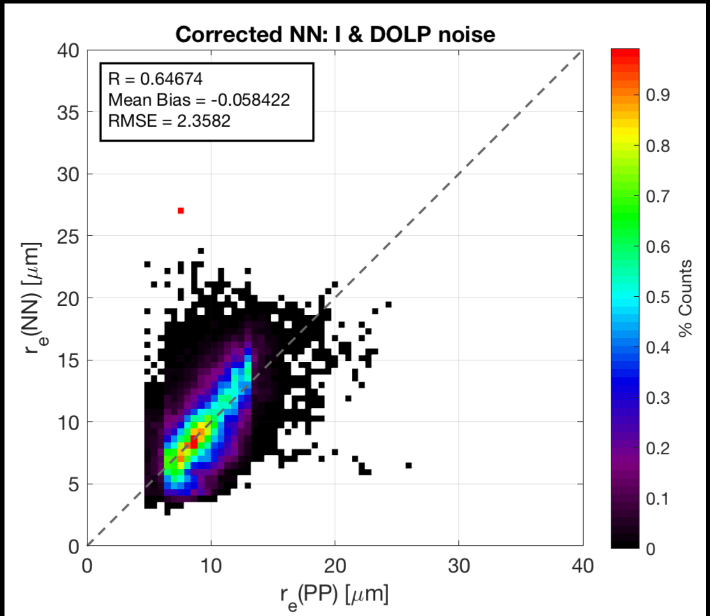
Validate with field observations that use standard RSP algorithms



A small change… A big difference

- Results for tanh d

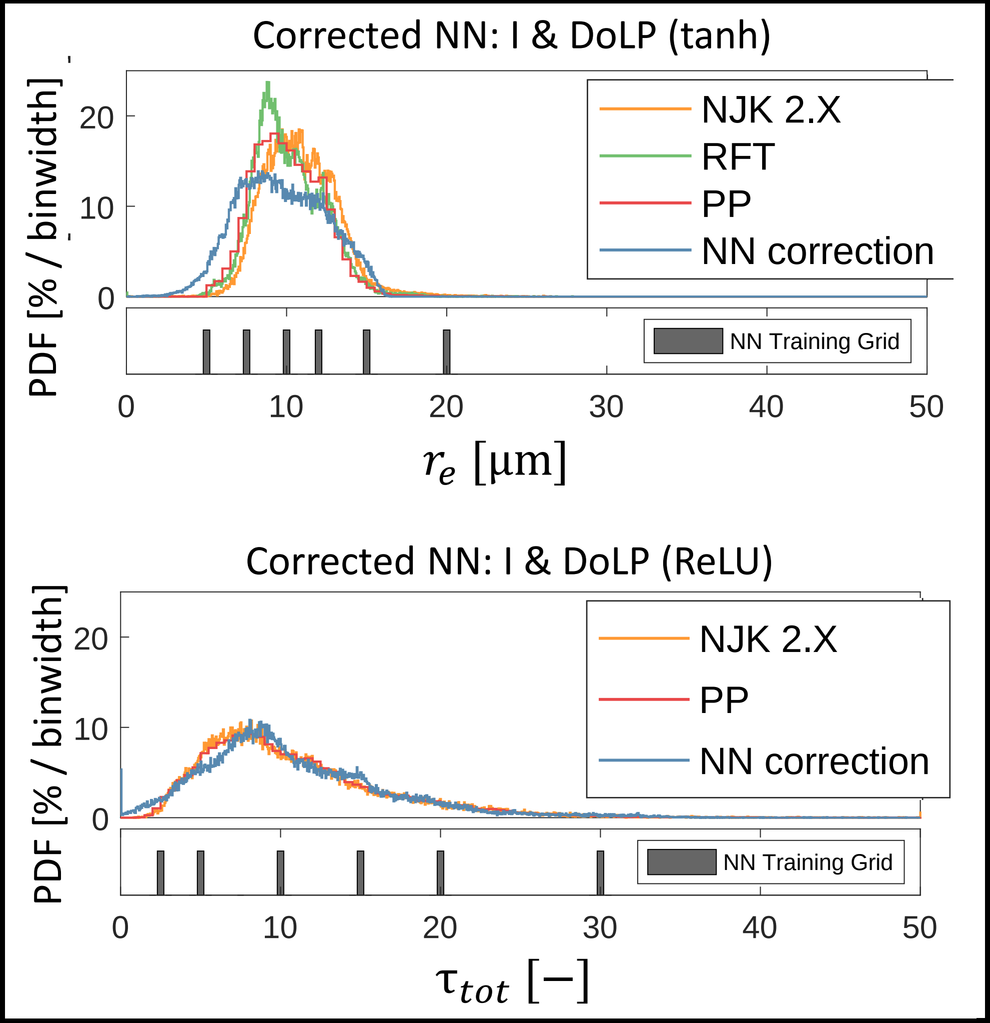




Evaluation

Step VI – Retrievals & Validation

Validate with field observations that use standard RSP algorithms:
Parametric (PP), Nakajima-King (NJK), and Rainbow Fourier Transform (RFT)

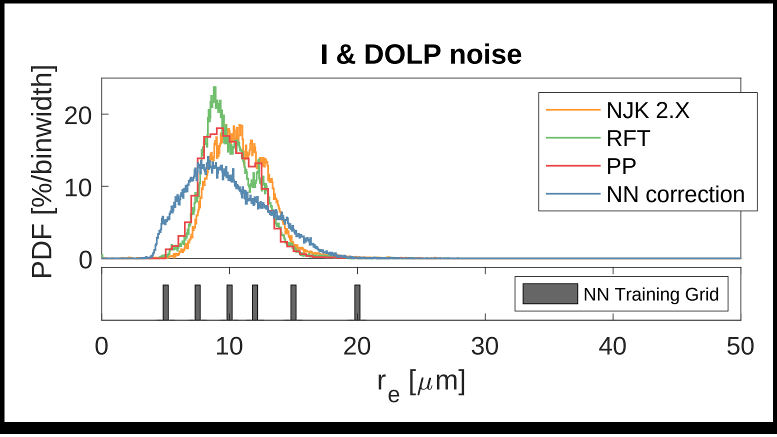


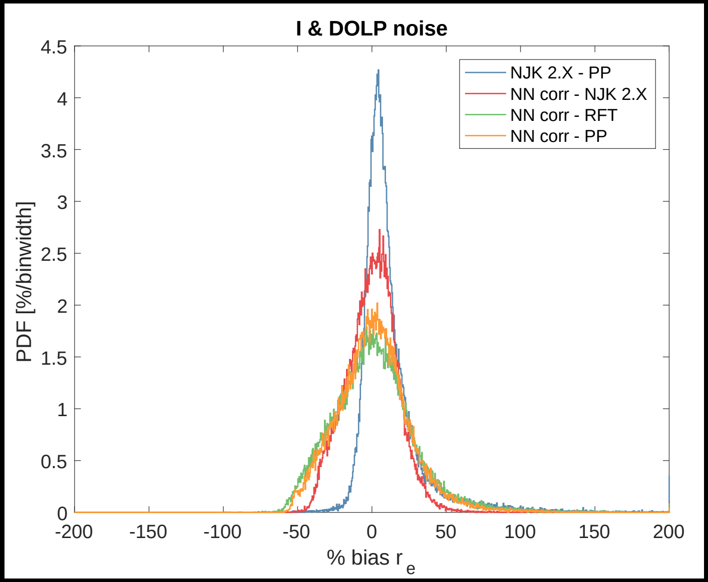
Take away:

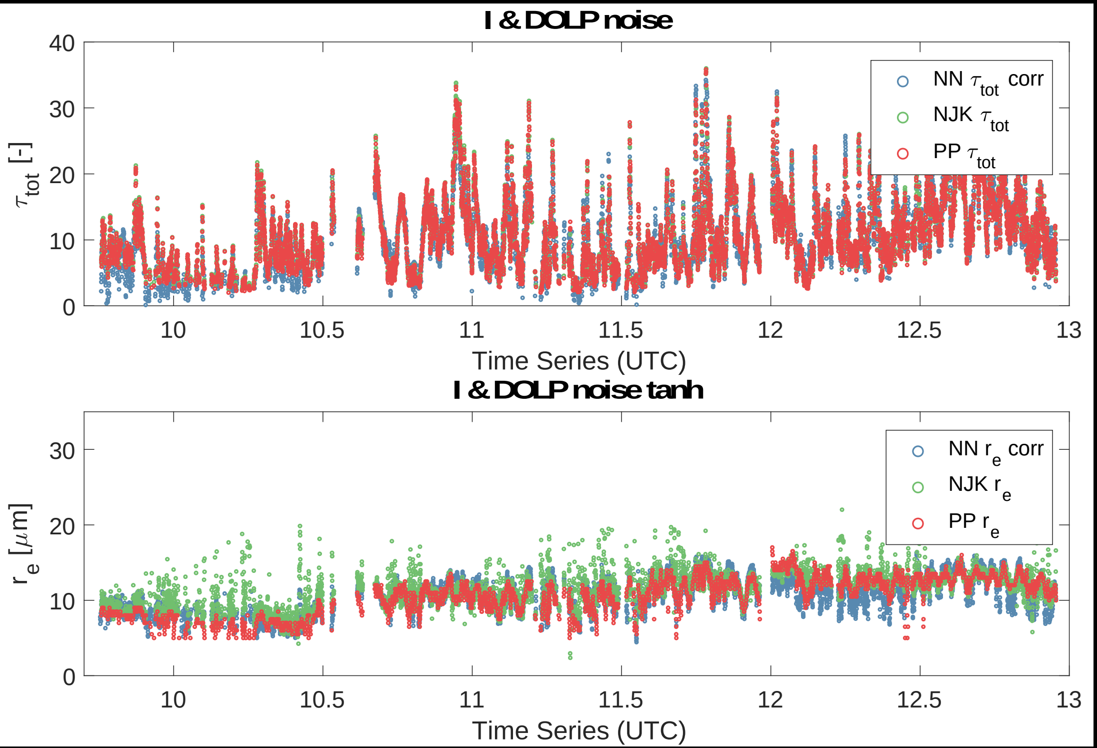
With a relatively simple forward model training set, the Neural Network approach appears to perform fairly well.

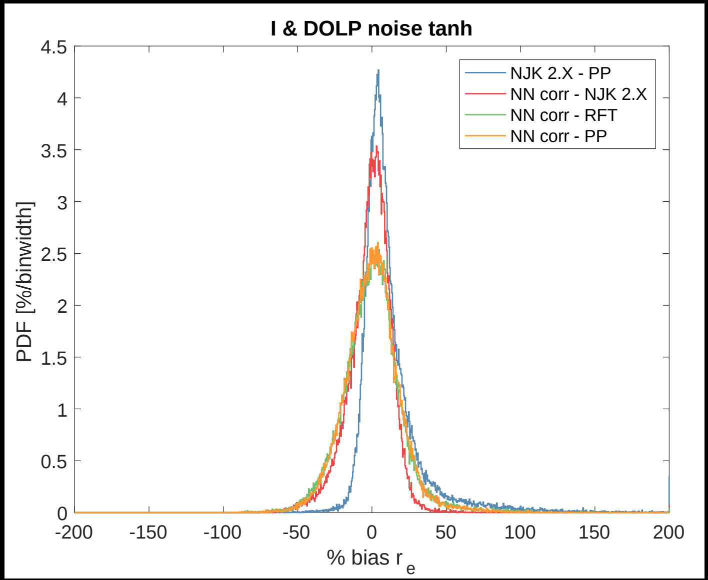
The network can be trained with a denser training grid and possibly be improved.

Other Histograms and Time Series





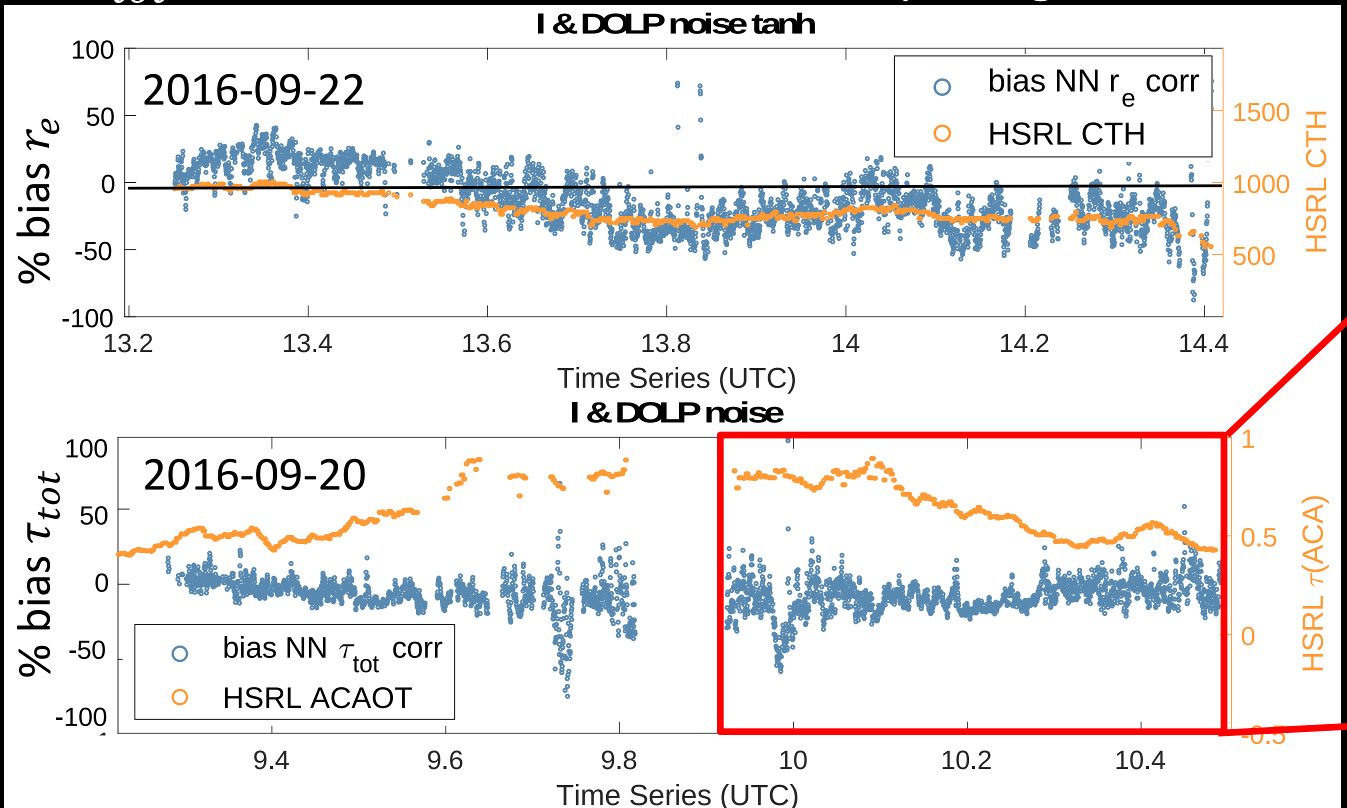


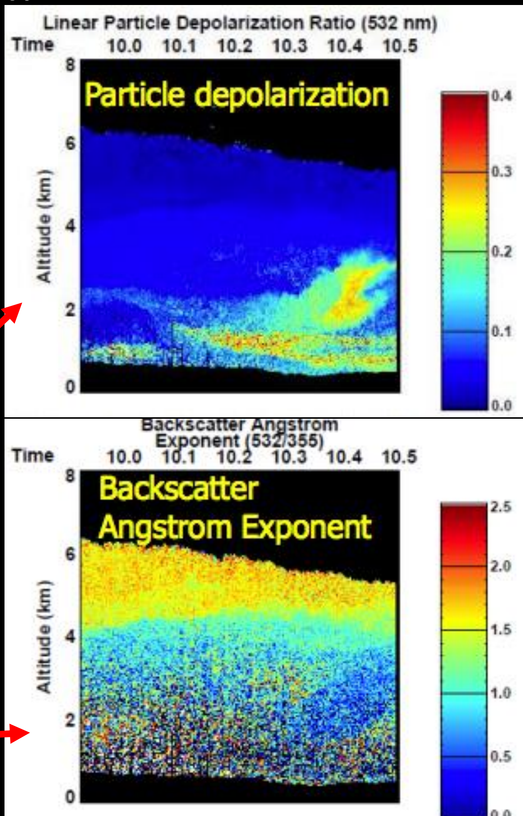


Remaining Issues

Are there any explanations for remaining biases and variability?

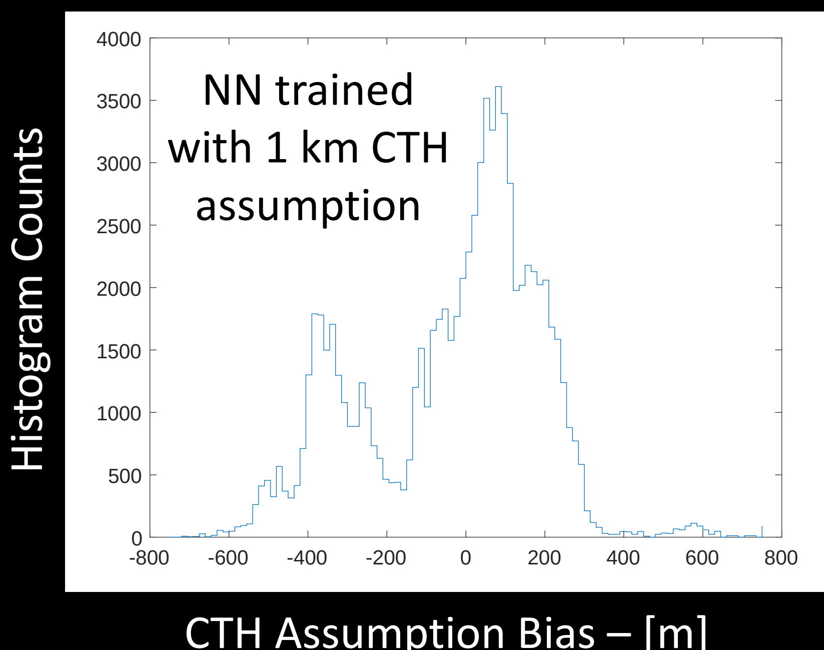
- Lower clouds tend to have low biased r_e retrievals
  - Difficult to tease out physical relationship, but indicates that the neural network should be trained using a variable cloud top height training set.
- Above cloud aerosol optical thickness (ACAOT) might be anti-correlated to the τ_{tot} bias, but trend is not obvious despite significant ACAOT.

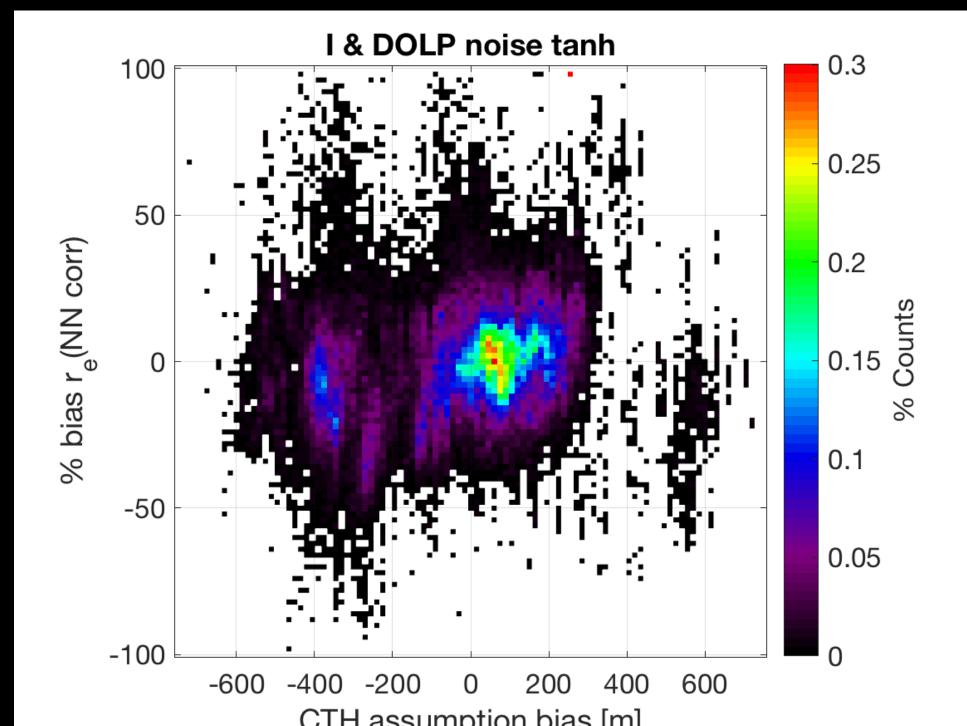




Sensitivity to Cloud top height

- Bias 2d histograms: confounding issues

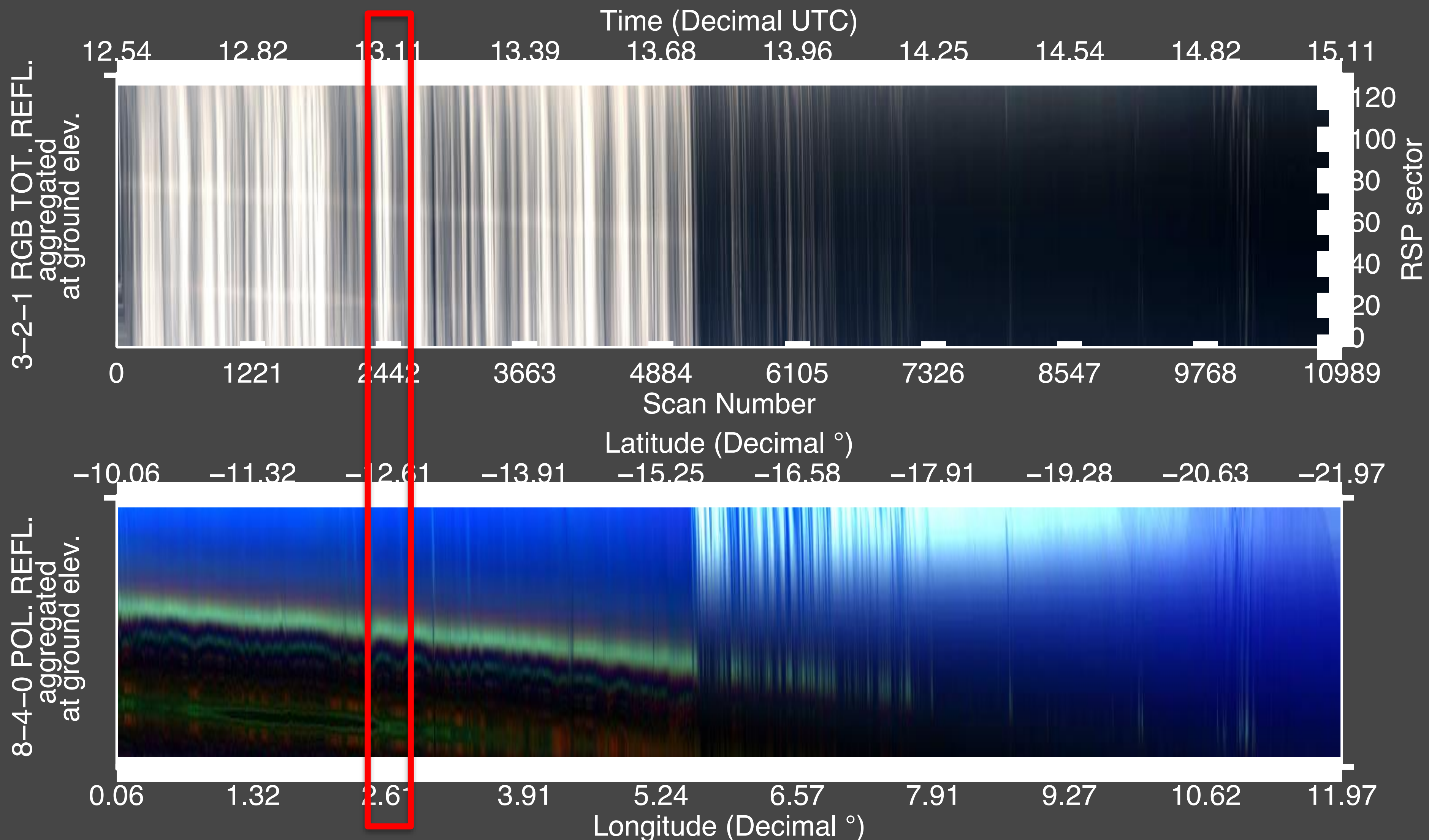




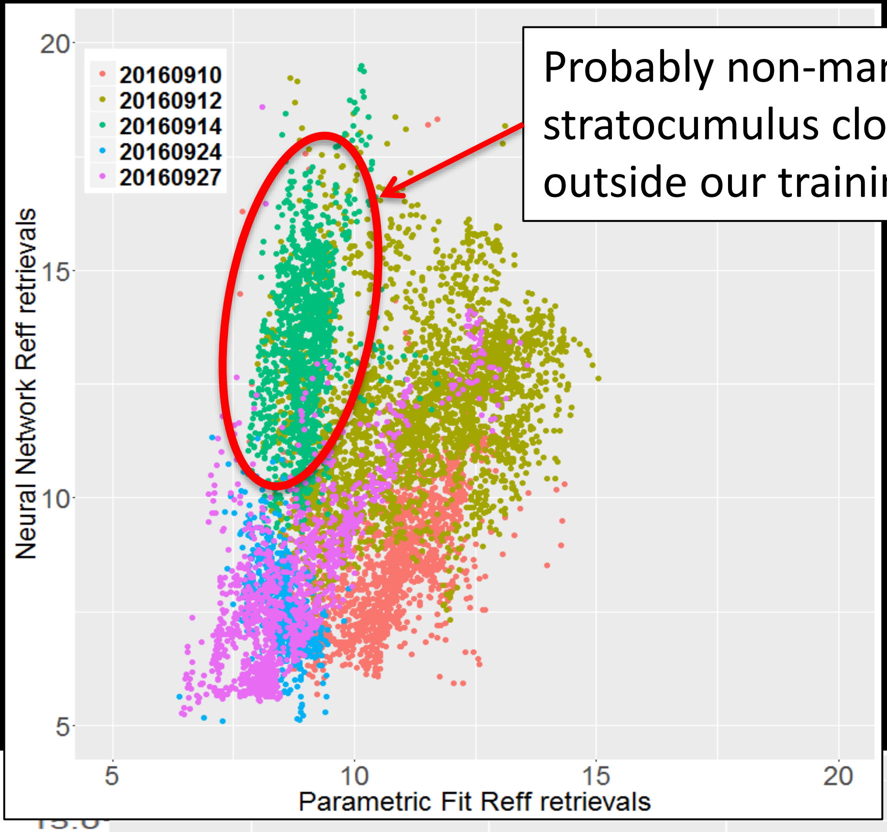
Sept. 12 ER-2 Return To Base leg

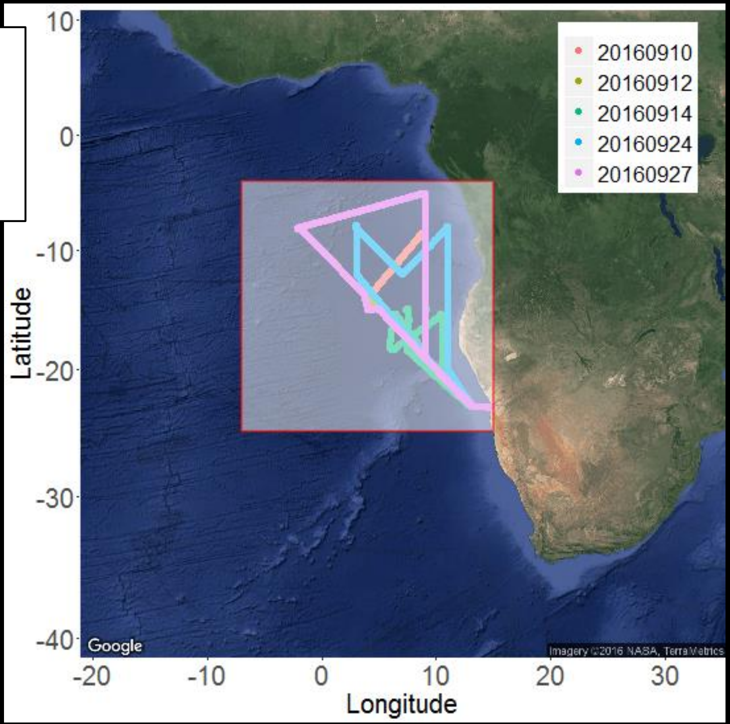
2016/09/12, 10989 scans: 12:32:15 to 15:06:21 UTC

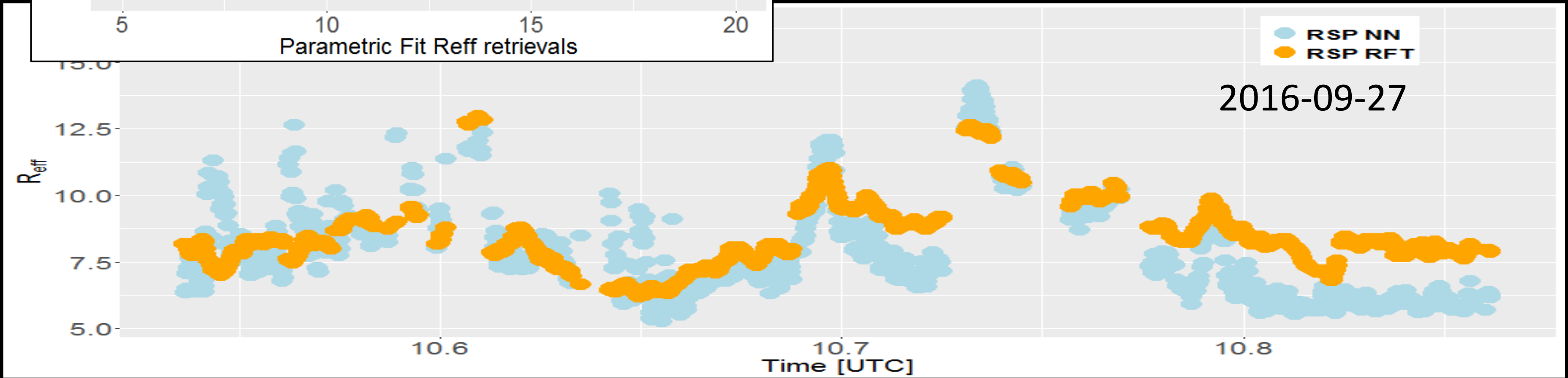
Average Rel. Azimuth in central 60% of scans: –16°; Scattering angle range: 51°–178°



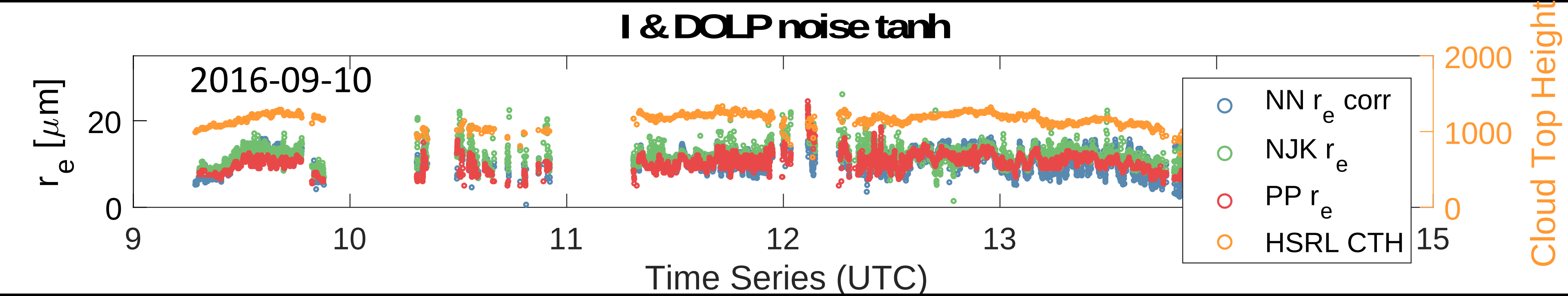
Step IV – retrievals & validation: Cloud Droplet effective radius
Validate with field observations that use standard RSP algorithms

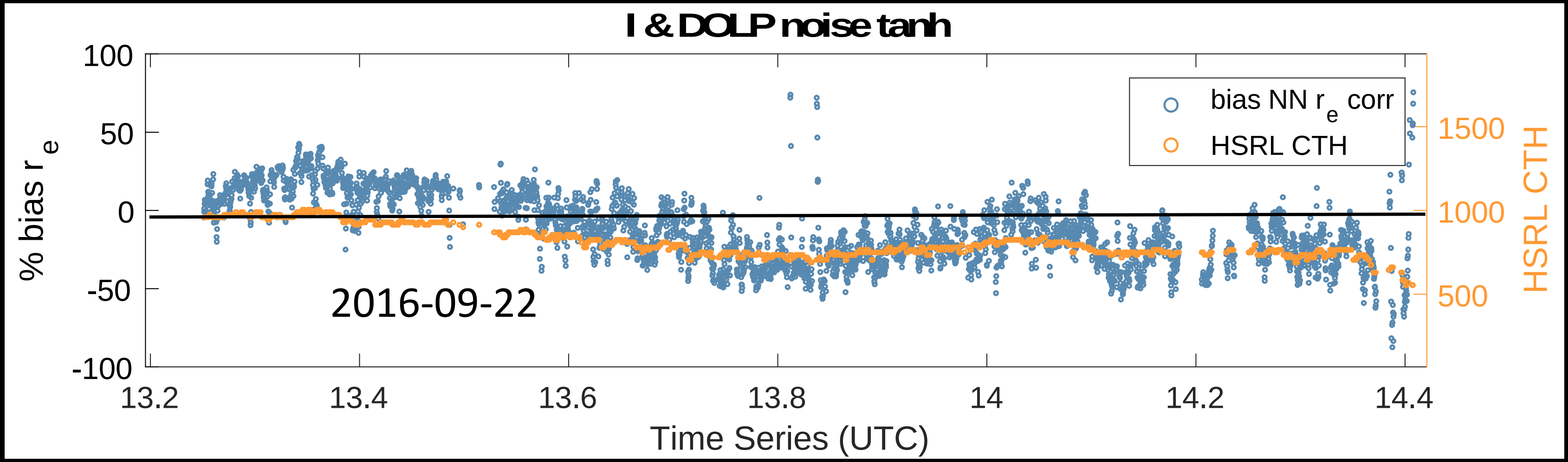


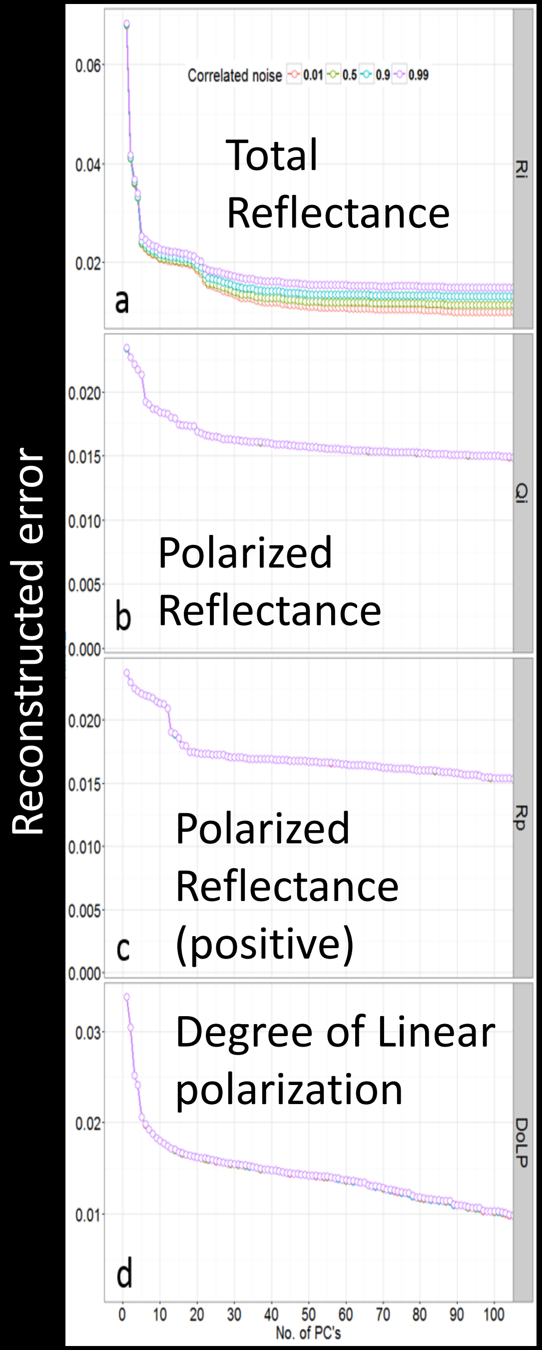




Natural Variability of r_e with CTH







Implementation

Step II:

- dimensionality reduction to create input layer nodes

y = noise free simulated data

y'= noisy simulated data

P = Projection obtained from PCA on y

P_n is number of PC to be retained

y'_r = reconstructed vector of noisy simulated data = $P^T_n P_n y'$

Reconstructed error = $||y'_r - y||$

\# of PC's are chosen to minimize the reconstruction error

Polarization basics

Descriptive terms

Reflectance R_I, R_Q, R_U

$$R_I = \frac{I p r_o^2}{F_o \cos\left(q_s\right)}$$

Similar for Q, U & V radiances [w/m^2 sr]
r_o – solar distance [AU]
F_o – Exo-atmospheric irradiance [w/m^2]

Polarized reflectance R_P

$$R_P = \sqrt{R_Q^2}$$

Linearly polarized component of reflectance
Always positive, polarization directionality lost

Degree of Linear Polarization DoLP

$$DoLP = \frac{\sqrt{Q^2 + U^2}}{I} = \frac{R_P}{R_I}$$

Polarized reflectance fraction, always positive
Often less sensitive to calibration, but…
…expresses both total and polarized interactions