Using Neural Networks for Bayesian Precipitation Retrievals

from GPM Passive Microwave Observations

Simon Pfreunschuh[1]
Teodor Norrestad [2]
Patrick Eriksson [1]
Inderpreet Kaur [1]
Christian D. Kummerow [3]

[1] Chalmers University of Technology
[2] previously Chalmers University of Technology
[3] Colorado State University

- **Global Precipitation Measurement (GPM)**: Constellation of microwave radiometers providing frequent ($T \leq 3\mathrm{h}$) global measurements of precipitation
- **Goddard Profiling Algorithm (GPROF)**: The retrieval algorithm used to retrieve precipitation profiles.

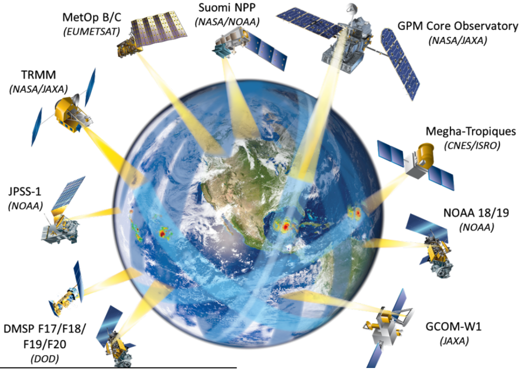


Image source: NASA

Research question:

Can we use a neural-network-based retrieval in the next version of GPROF?

Motivation

- Simpler and faster
- Greater flexibility with respect to input variables
- Better retrievals

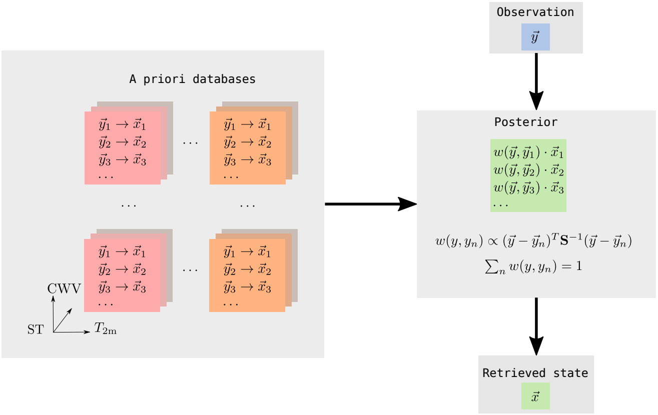
This study

- Developed a neural-network-based surface precipitation retrieval
- 1-to-1 comparison to the current GPROF version

References, code, and models

- Check out the project page: `https://github.com/see-mof/regn`

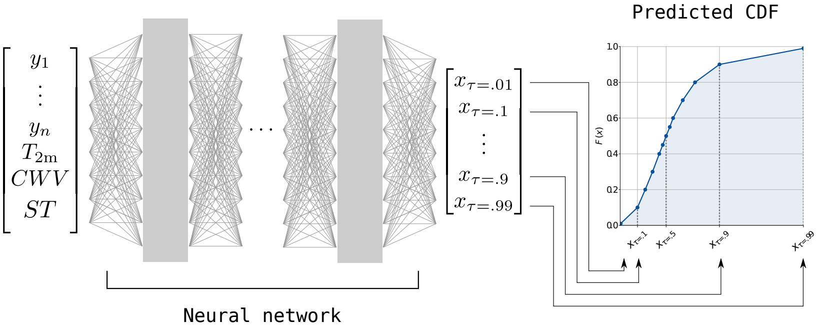
- Bayesian retrieval method: Monte Carlo Integration (MCI)
- Posterior approximated using importance sampling based deviation from obs.
- A priori databases stratified by surface type, two-meter temperature, total column water vapor



Quantile regression neural network (QRNN)

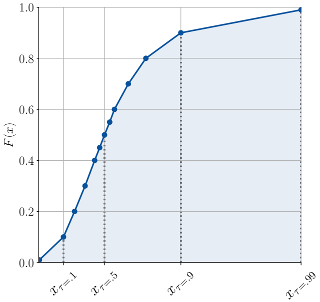
- Trained to minimize quantile loss function (skewed absolute error):

$$\mathcal{L}(x_\tau, x) = \begin{cases} \tau|x_\tau - x| & x_\tau \leq x \\ (1-\tau)|x_\tau - x| & \text{otherwise} \end{cases} \tag{1}$$



CHALMERS
UNIVERSITY OF TECHNOLOGY

Predicted posterior



Derived statistics

- Point predictors:
  - Median: $x_{\tau=.5}$
  - Posterior mean: $\mathbf{E}(x|y) = \int x \, dF$
- Confidence intervals:

$$P(x_{\tau=.45} < X \leq x_{\tau=.55}) = 10\%$$
$$P(x_{\tau=.35} < X \leq x_{\tau=.65}) = 30\%$$
$$\ldots$$

- Classifier:

$$P(X > x) = 1 - F(x)$$

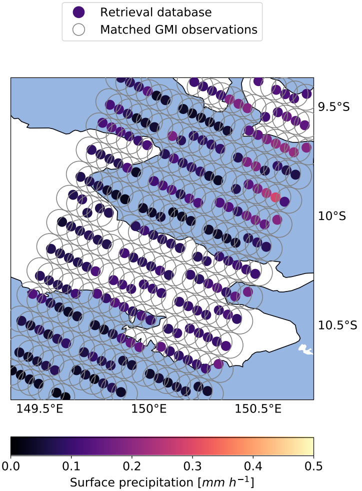
Can QRNN replace Monte-Carlo integration?

Training

- Data: GPROF V5 a-priori database (GMI/DPR combined observations)
- QRNN: Fully-connected NN, 10 layers, 128 neurons, ReLU activations

Evaluation

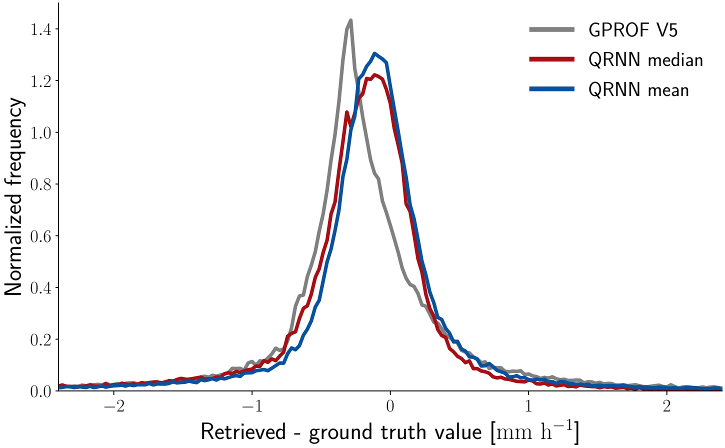
- Day 1 and 2 of each month left out for testing
- Real GMI observations matched to nearest sample in retrieval database
- Data subsampled to decrease redundancy



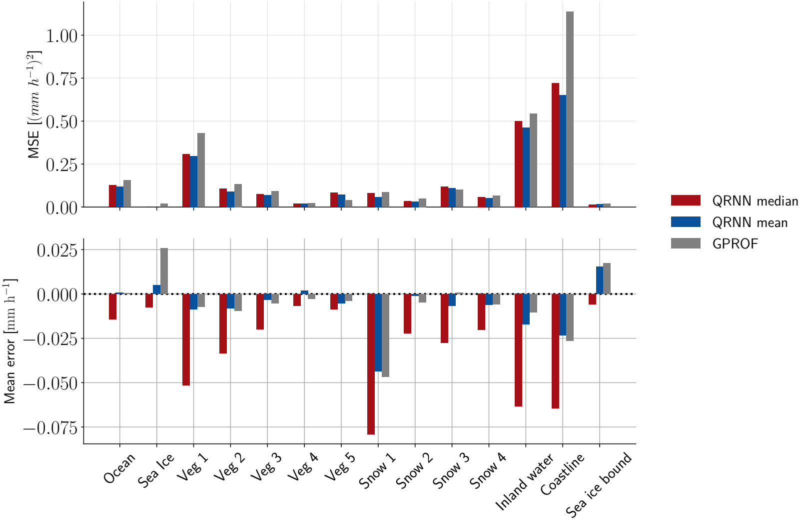
CHALMERS UNIVERSITY OF TECHNOLOGY

- Reduced bias in QRNN results
- Minor reductions in MSE and MAE

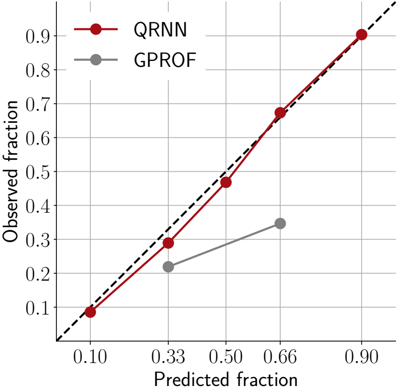
Deviations for rain rates $> 0.3\,\mathrm{mm\ h^{-1}}$**:**



- Reductions mostly consistent across surface types

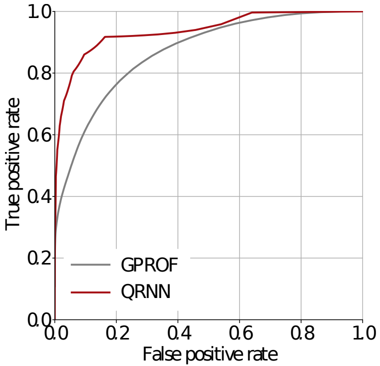


- Predicted confidence intervals consistent with observed deviations
- Not the case for GPROF 1st and 3rd precipitation terciles



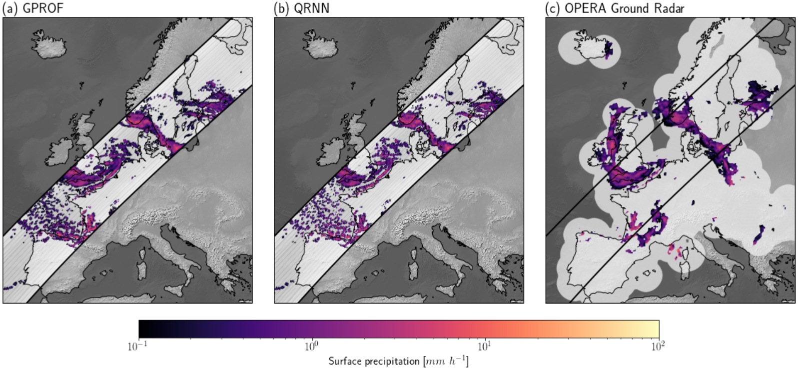
Receiver operating characteristic

- Raining/Non-raining classification, threshold = 0.01 $\mathrm{mm\ h^{-1}}$
- QRNN predictions more reliable than GPROF probability of precipitation



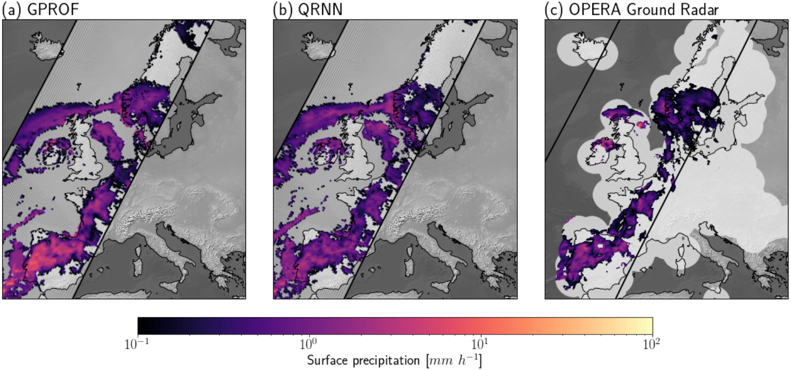
CHALMERS
UNIVERSITY OF TECHNOLOGY

GMI



MHS

- QRNN works equally well for cross-track scanning sensor
- One network can handle all viewing angles



Can QRNNs replace Monte Carlo integration?

- Results, so far, indicate yes

Potential improvements

1. More accurate point estimates
2. Better estimate of the posterior distribution (uncertainty, classification)
3. One network can handle different surface types and viewing angles
4. Potential path towards making use of spatial information

Next steps

- Run QRNN-based GPROF version in parallel with MCI-based version
- Evaluate QRNN-based GPROF in production

Open questions

- Profile retrievals
- Correlated errors

More information, code and models

- Check out the project page: `https://github.com/see-mof/regn`