EARTH SYSTEM SCIENCE
INTERDISCIPLINARY CENTER

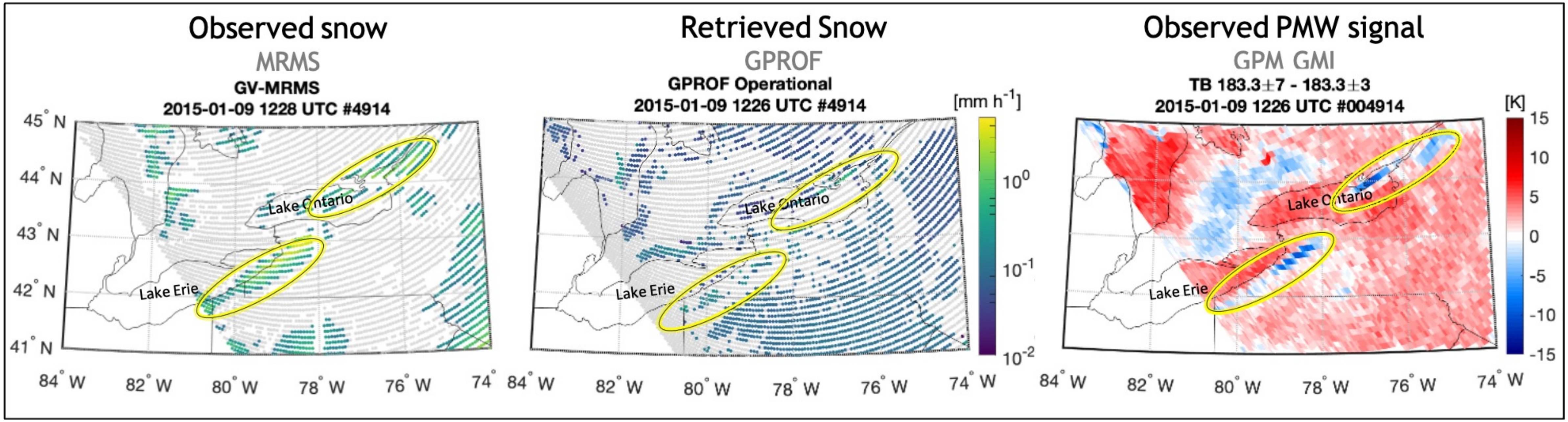
Retrieving Snowfall Class from Satellite Passive Microwave Observations

Veljko Petković
Yulan Hong
Lisa Milani

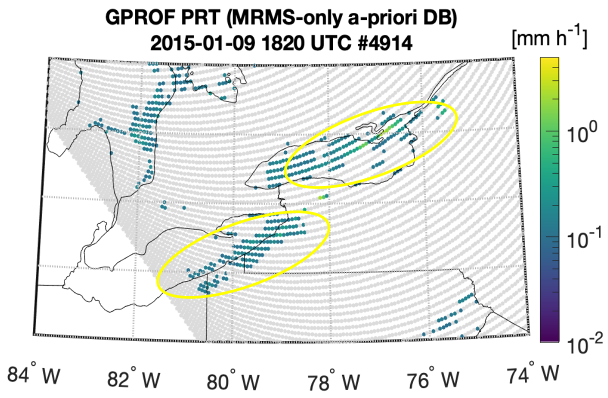
NASA PMM Project
PI: Lisa Milani, ESSIC/UMD, GSFC NASA
Co-I: Veljko Petkovic, ESSIC, University of Maryland
Collaborators from FMI, NPS, NASA GSFC, UW-Madison, UH, NOAA/NESDIS, OU

UNIVERSITY OF MARYLAND

Passive Microwave Retrievals



Specific *a priori* database

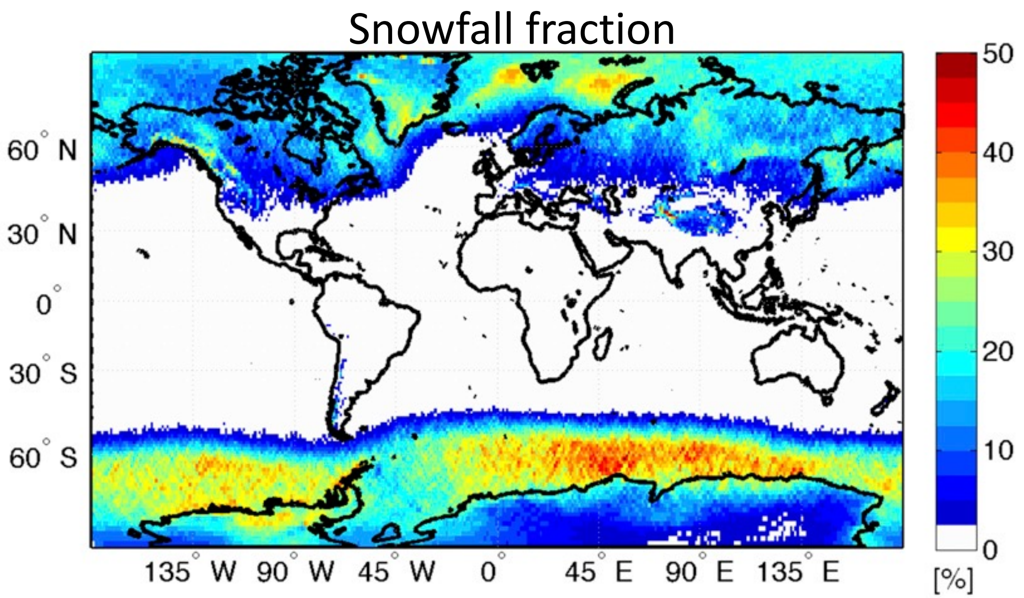


Factors influencing Passive Microwave (PMW) snowfall rate retrievals (e.g. GPROF): Surface type/coverage, ice crystal shape/orientation/density, atmospheric water vapor content

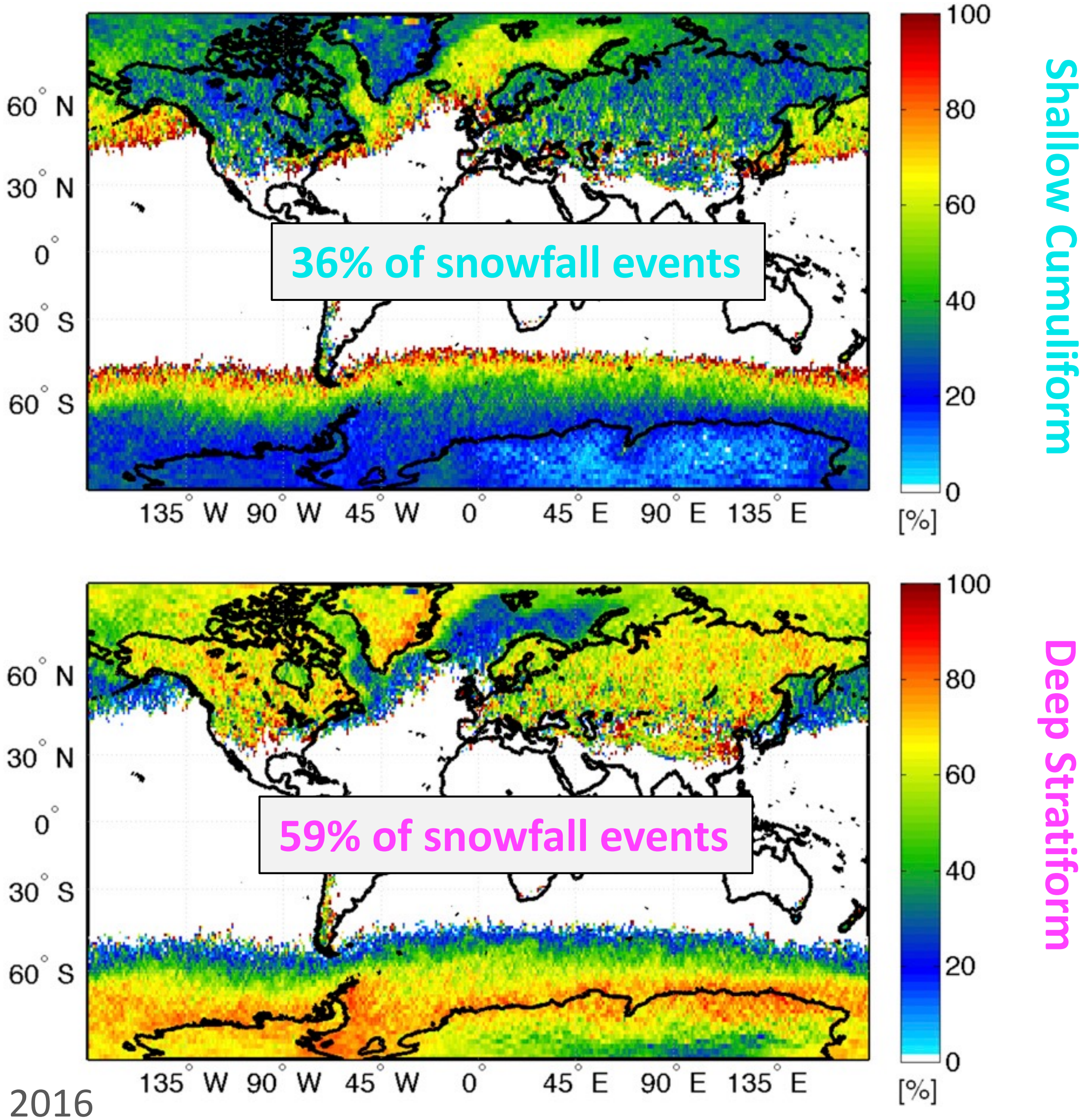
Retrievals need an *a priori* (i.e., training) database to constrain the solution (in most cases). A priori knowledge must be "complete".

Previous work (e.g., Milani et al. 2021) confirms the value of snowfall regime information in PMW snowfall retrievals

Global Snow Distribution



> 60% of snowfall fraction is shallow cumuliform:
- North Atlantic Ocean
- Labrador Sea
- Sea of Japan
- US Great Lakes Region

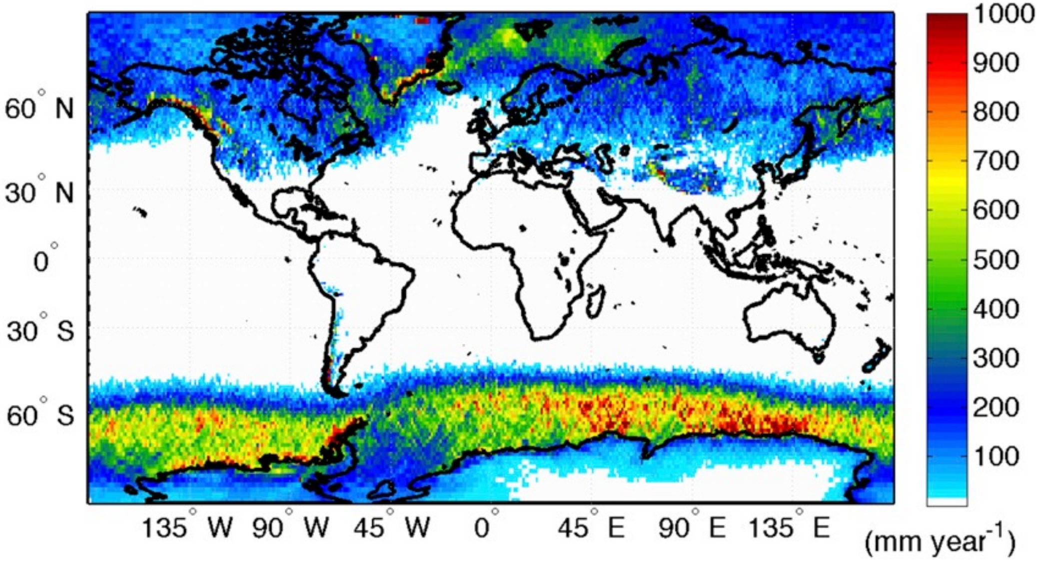


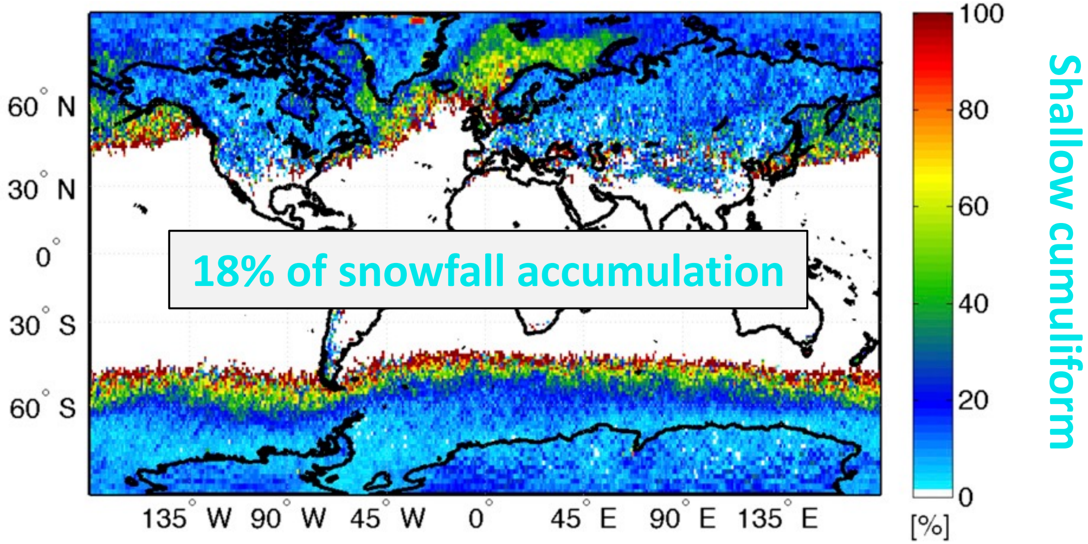
Kulie et al. 2016

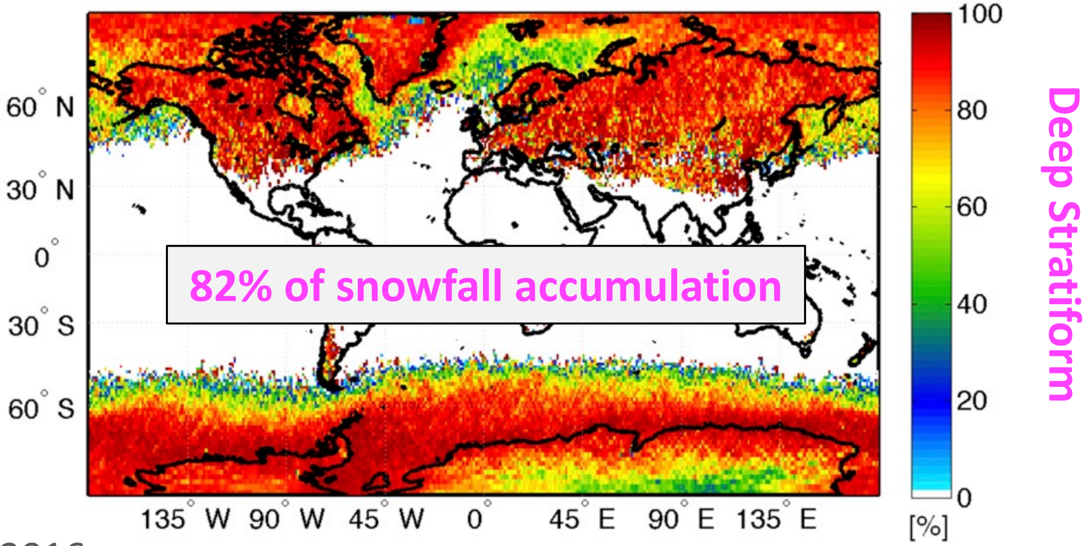
Global Snow Distribution

~ 50% of shallow cumuliform accumulation:
- North Atlantic Ocean
- Labrador Sea
- Sea of Japan
- US Great Lakes Region

Mean annual snowfall





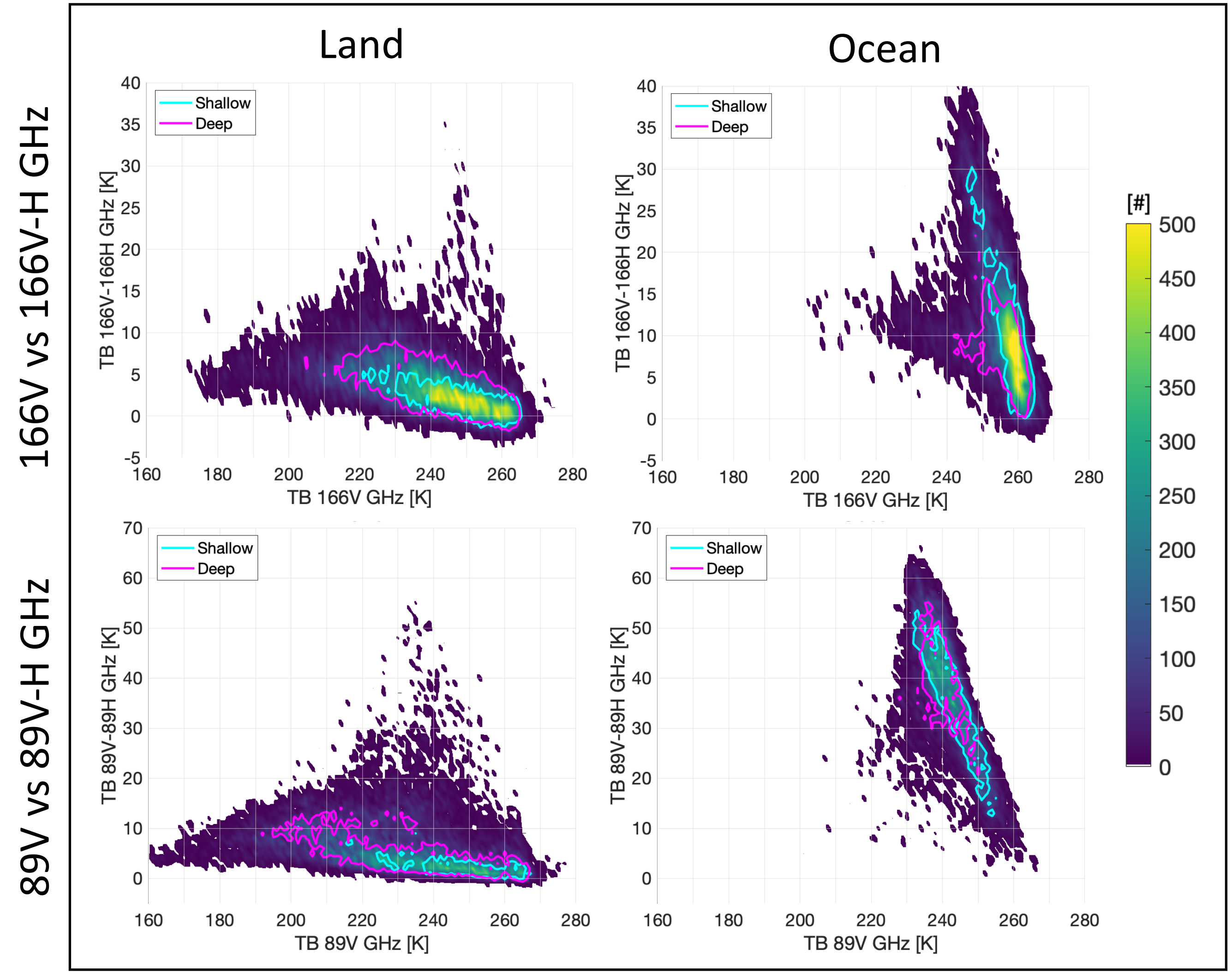


Kulie et al. 2016

A Challenge and a Proposed Solution

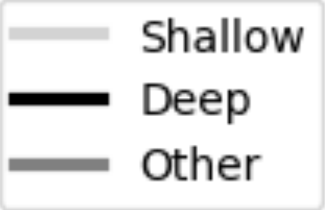
- Different snowfall regimes often appear radiometrically similar, preventing the retrieval to converge.
- Non-linear relationships between Brightness Temperatures (TBs) at different frequencies, or combination of frequencies
- Machine learning techniques can enable finding the “hidden” relationships and help with the PMW classification of snowfall regimes, taking advantage of the entire range of information carried by the PMW channels

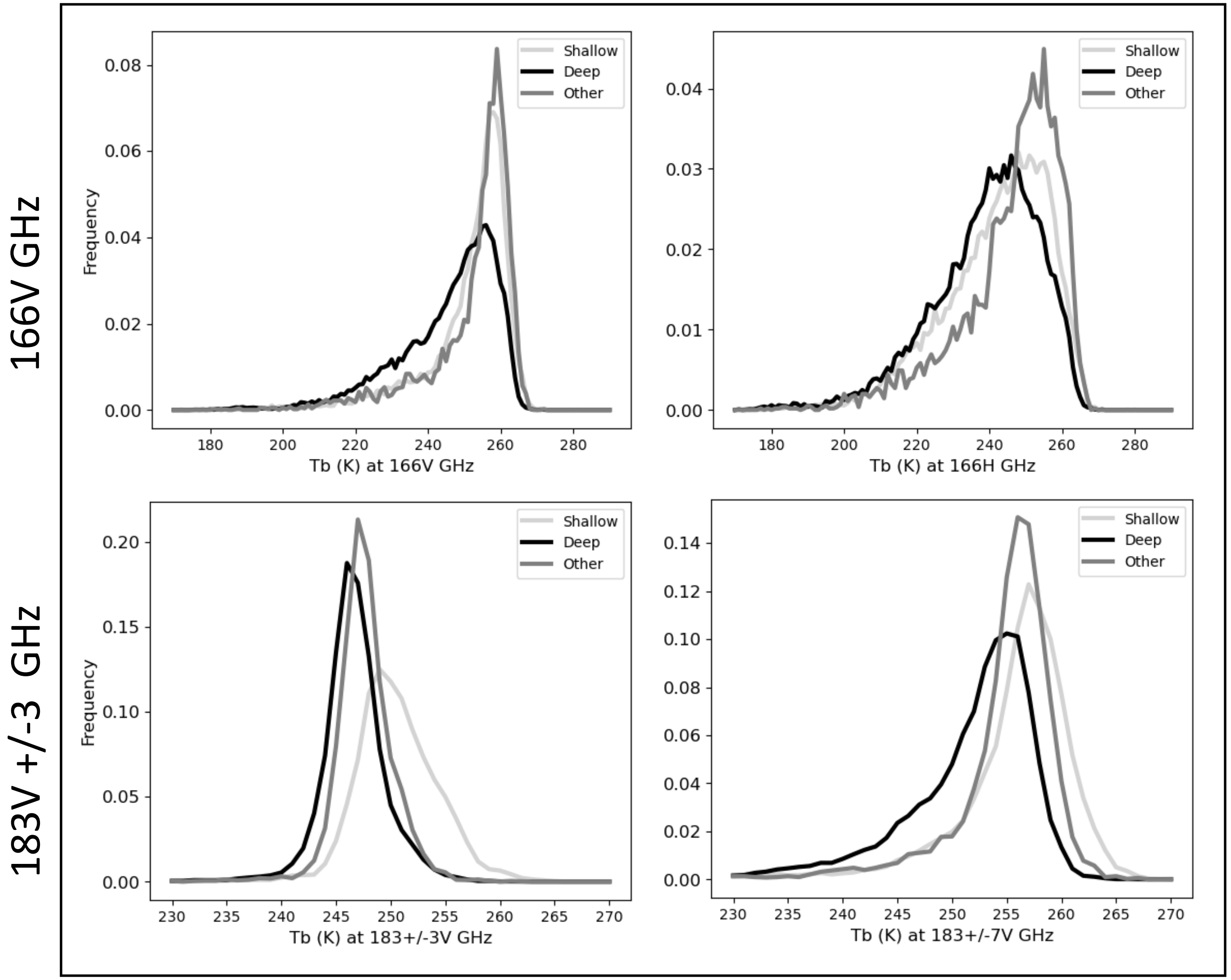
Shallow
Deep



A Challenge and a Proposed Solution

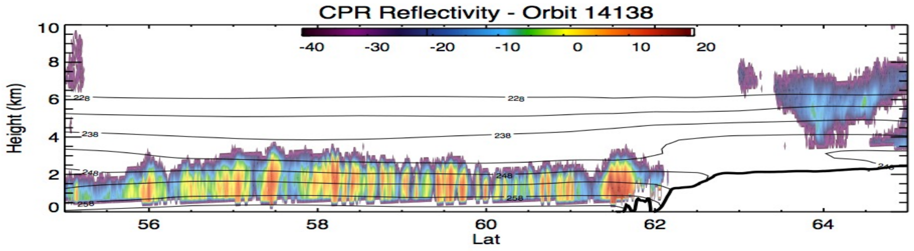
- Different snowfall regimes often appear radiometrically similar, preventing the retrieval to converge.
- Non-linear relationships between Brightness Temperatures (TBs) at different frequencies, or combination of frequencies
- Machine learning techniques can enable finding the "hidden" relationships and help with the PMW classification of snowfall regimes, taking advantage of the entire range of information carried by the PMW channels





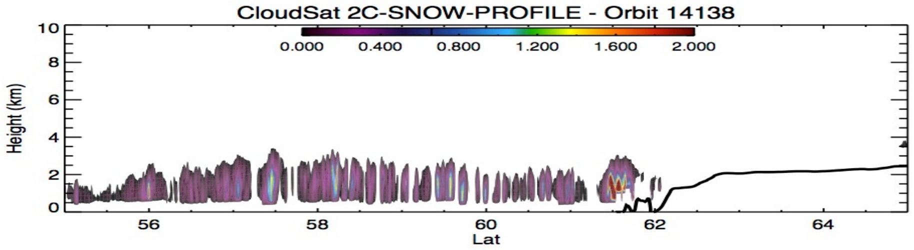
Classification scheme

Cumuliform Snow = 2B-CLDCLASS Stratocumulus or Cumulus + 2C-SNOW > 0



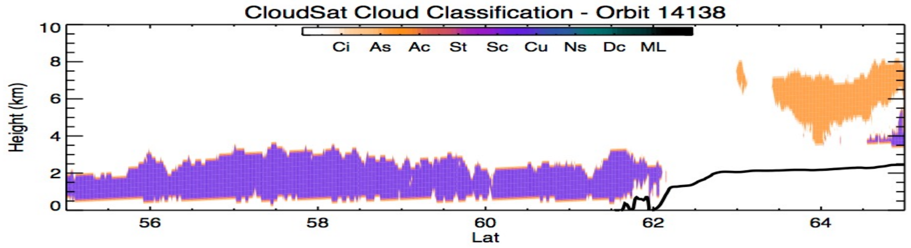
2B-GEOPROF

Reflectivity Profiles



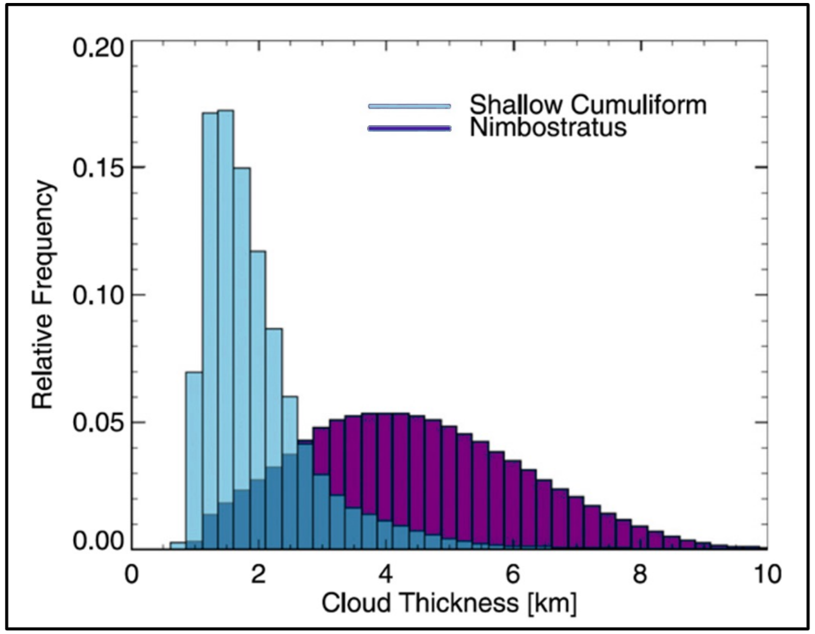
2C-SNOW-PROFILE

Snowfall rate retrievals



2B-CLDCLASS

Cloud classification

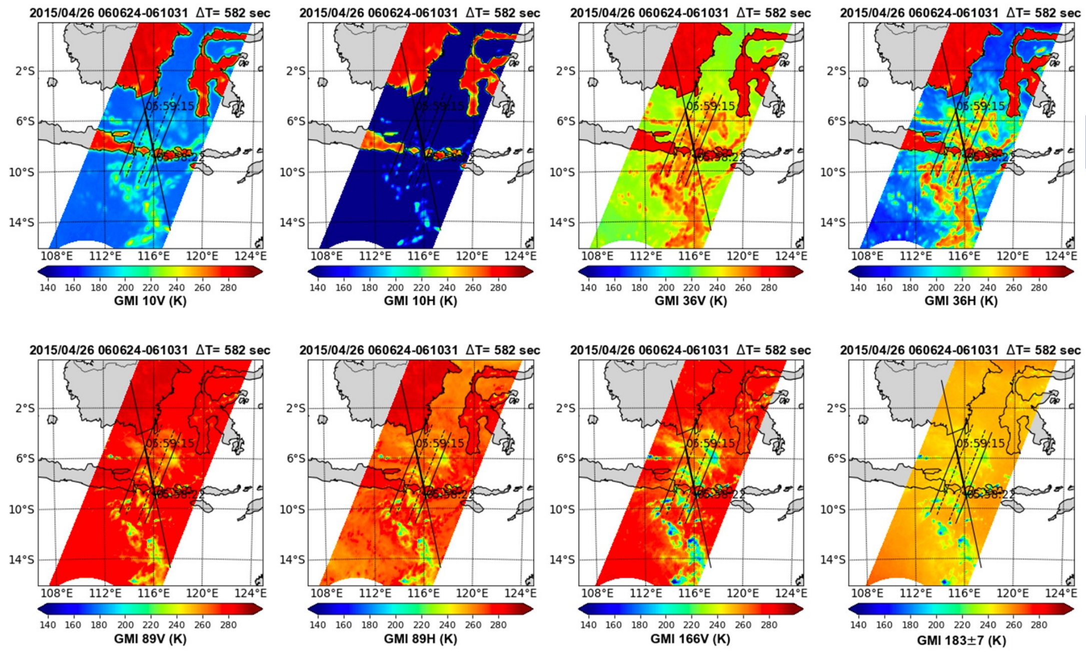


Kulie et al. 2016

EARTH SYSTEM SCIENCE
INTERDISCIPLINARY CENTER

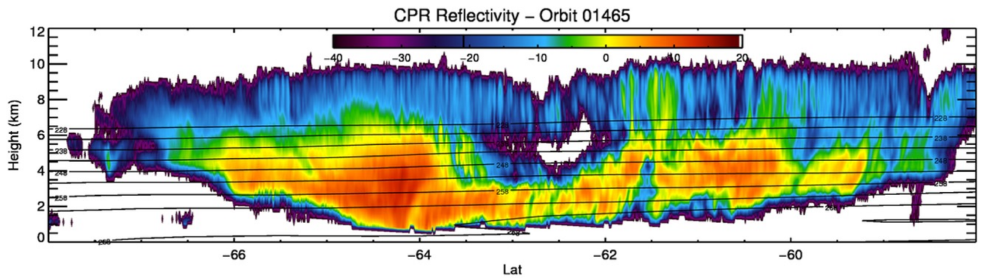
Method – Building the Training Dataset (*a priori* knowledge)

CloudSat-to-GMI Matches (2B-CSATGPM)

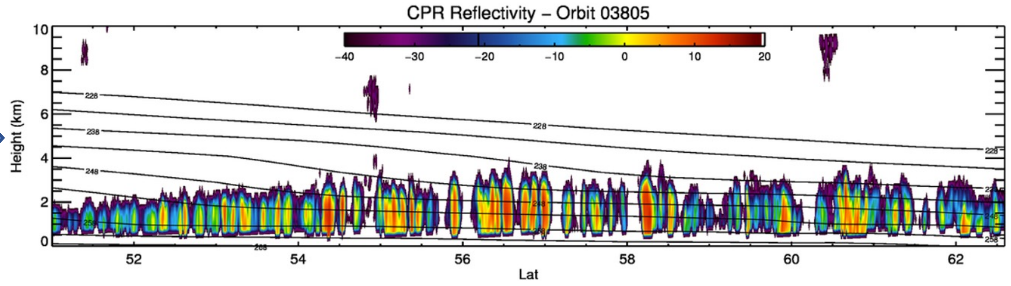


Turk et al. 2021

Deep Stratiform

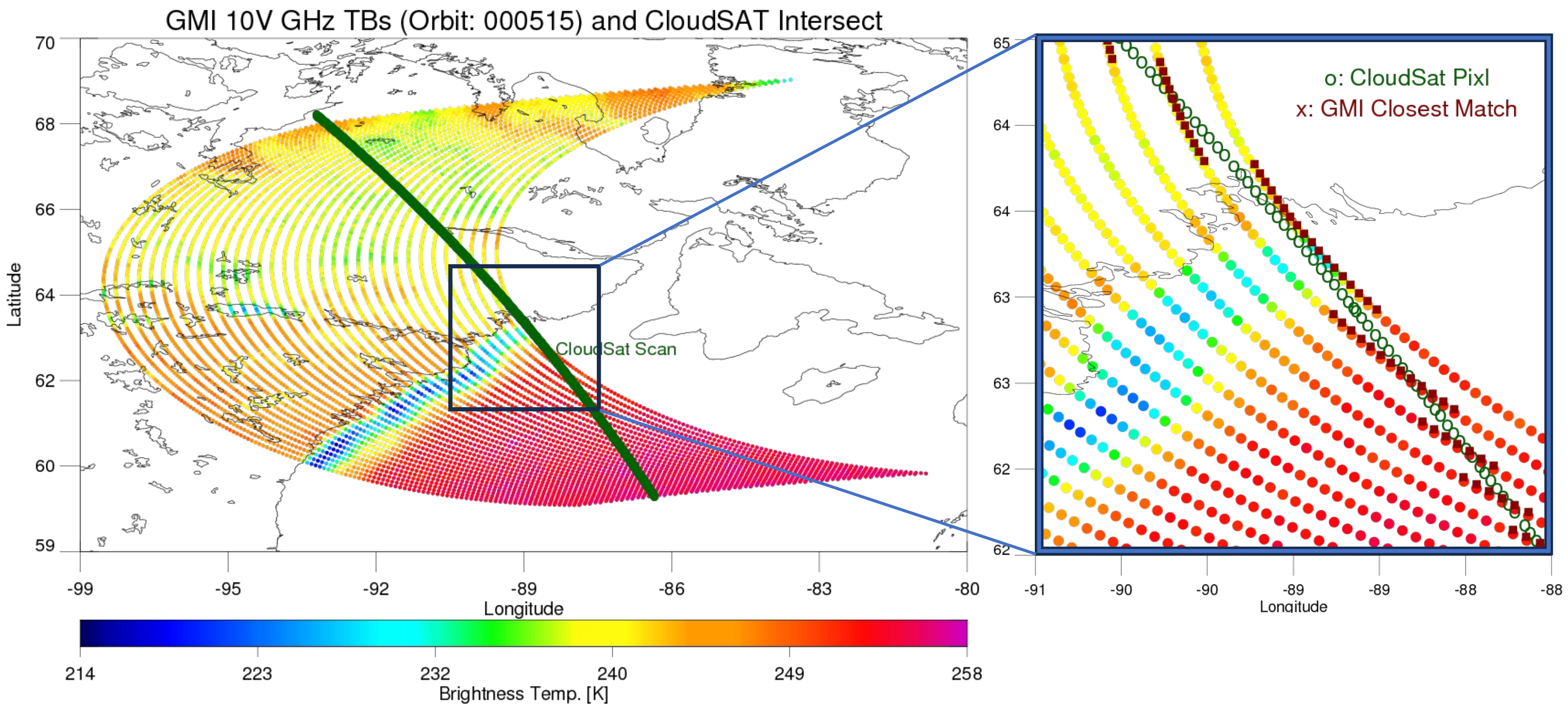


Shallow Convective



Curtesy of Dr. Mark Kulie

Method – Building the Training Dataset (*a priori* knowledge)



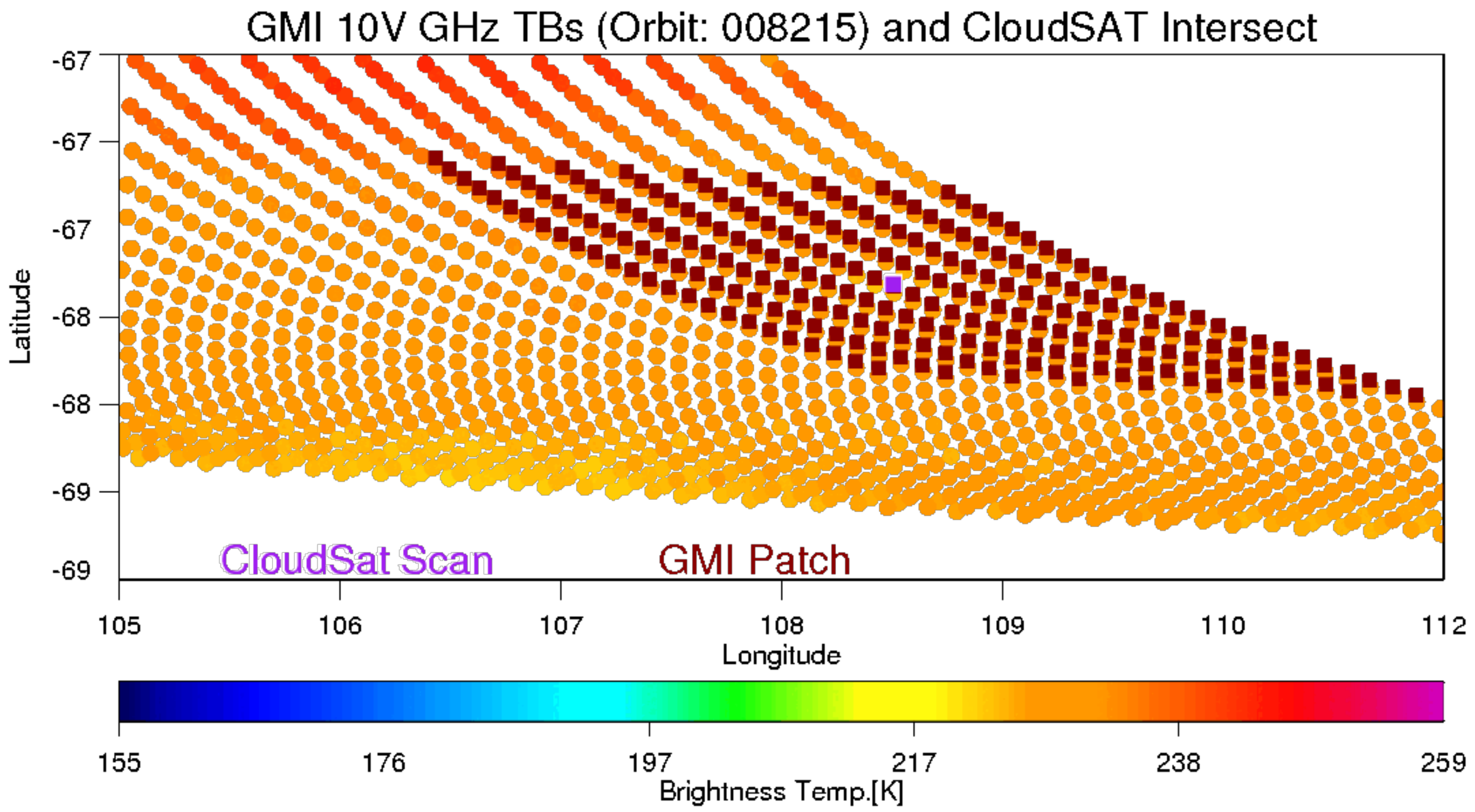
Input feature(s):
- GMI 13 channels TBs
- TBs polarization diff.
- 3x5 to 9x25 patches
- TPW
- T2m

Target (label) feature(s):
- CloudSat snowfall regime class

Time period:
- 2014–2018 to train
- 2019 to test

Total number of scenes:
- ~ 60k

Method – Building the Training Dataset (*a priori* knowledge)



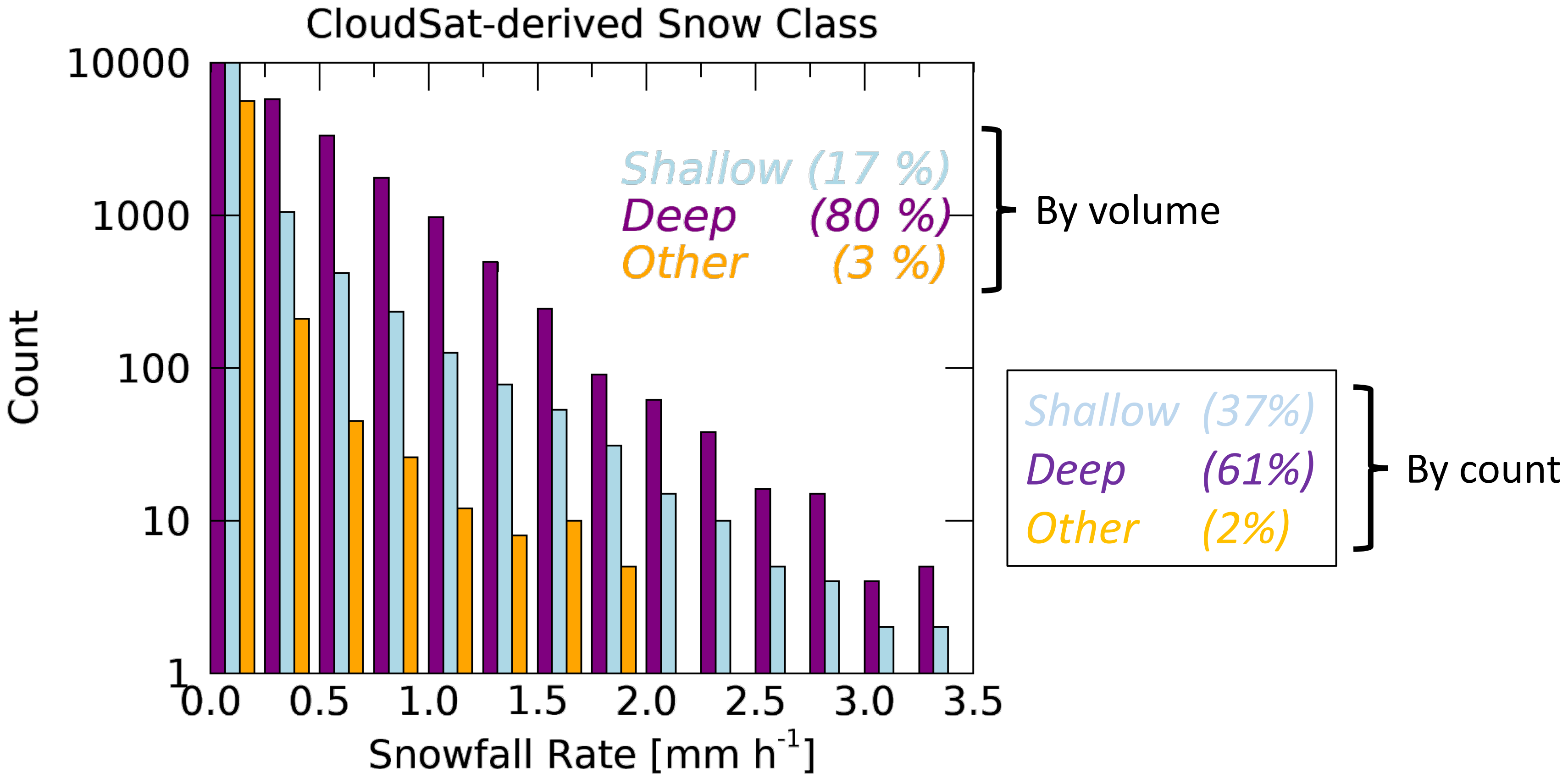
Input feature(s):
- GMI 13 channels TBs
- TBs polarization diff.
- 3x5 to 9x25 patches
- TPW
- T2m

Target (label) feature(s):
- CloudSat snowfall regime class

Time period:
- 2014–2018 to train
- 2019 to test

Total number of scenes:
- ~ 60k

Method – Building the Training Dataset (*a priori* knowledge)



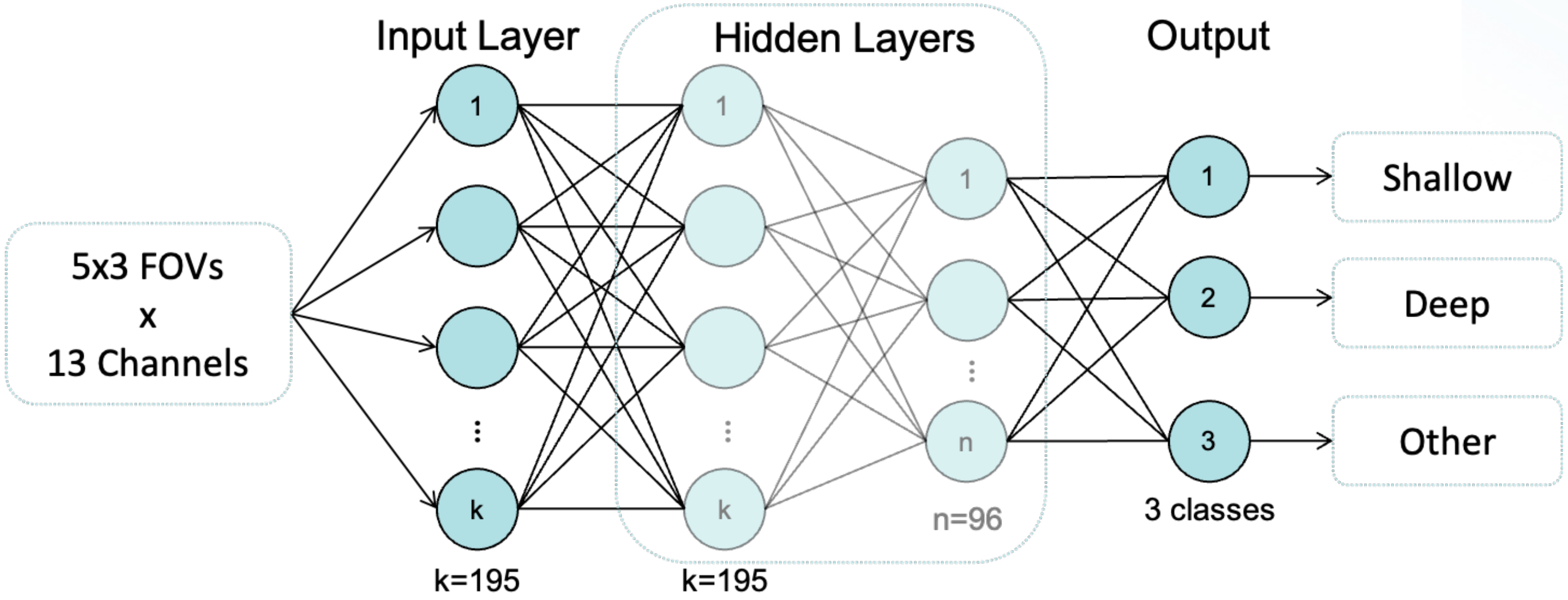
Input feature(s):
- GMI 13 channels TBs
- TBs polarization diff.
- 3x5 to 9x25 patches
- TPW
- T2m

Target (label) feature(s):
- CloudSat snowfall regime class

Time period:
- 2014–2018 to train
- 2019 to test

Total number of scenes:
- ~ 60k

Fully Connected Neural Network



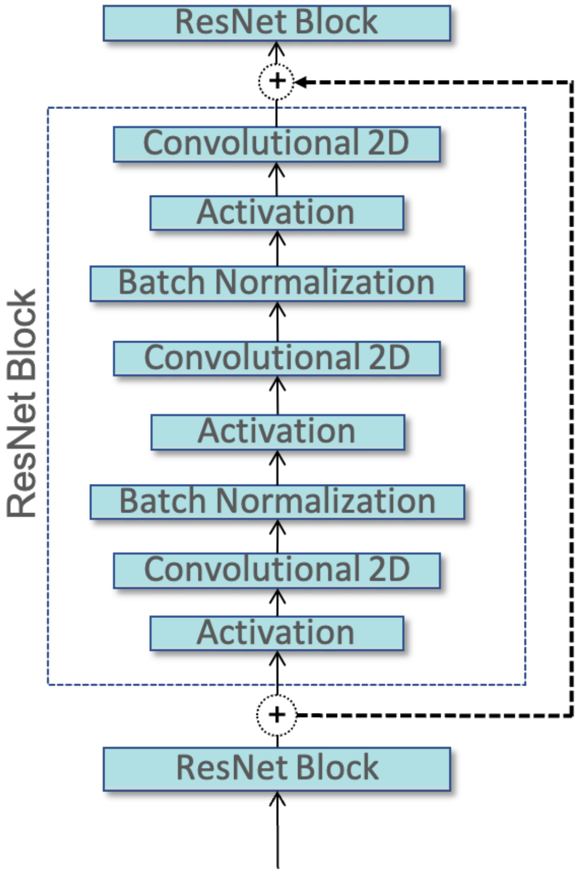
Train on 5 years of global overpasses.
Testing for different:

- Number of hidden layers (3)
- Number of neurons (195-96-96)
- Number of neurons (2925-1920-96)
- Activation function (ReLU)
- Optimizations (Adam)

- Batch norm is applied to input
- Uses standard Softmax function

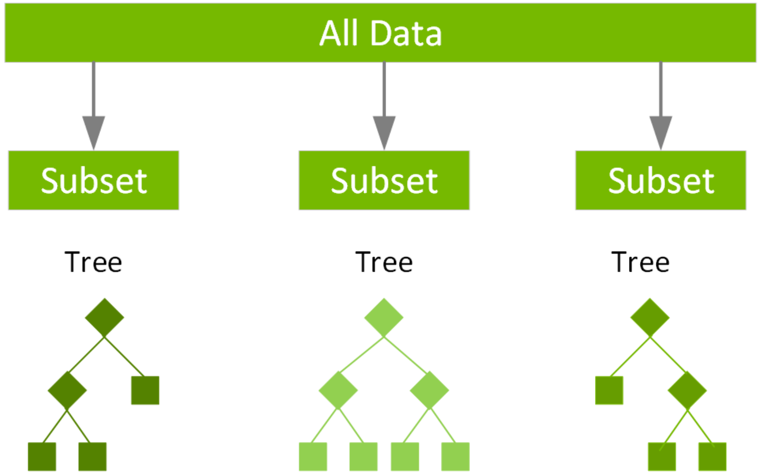
- Fully connected model is a basic ML architecture
- With a goal to achieve optimal performance, ResNet model is considered (next slide)

Residual Network Deep Model



- Convolutional neural networks (CNNs) are the best performing models in large-scale image recognition tasks
- Residual connections have the ability to train deeper CNN models.
- Developing residual network (ResNet) architecture to train for both deterministic and probabilistic models on the classification task.

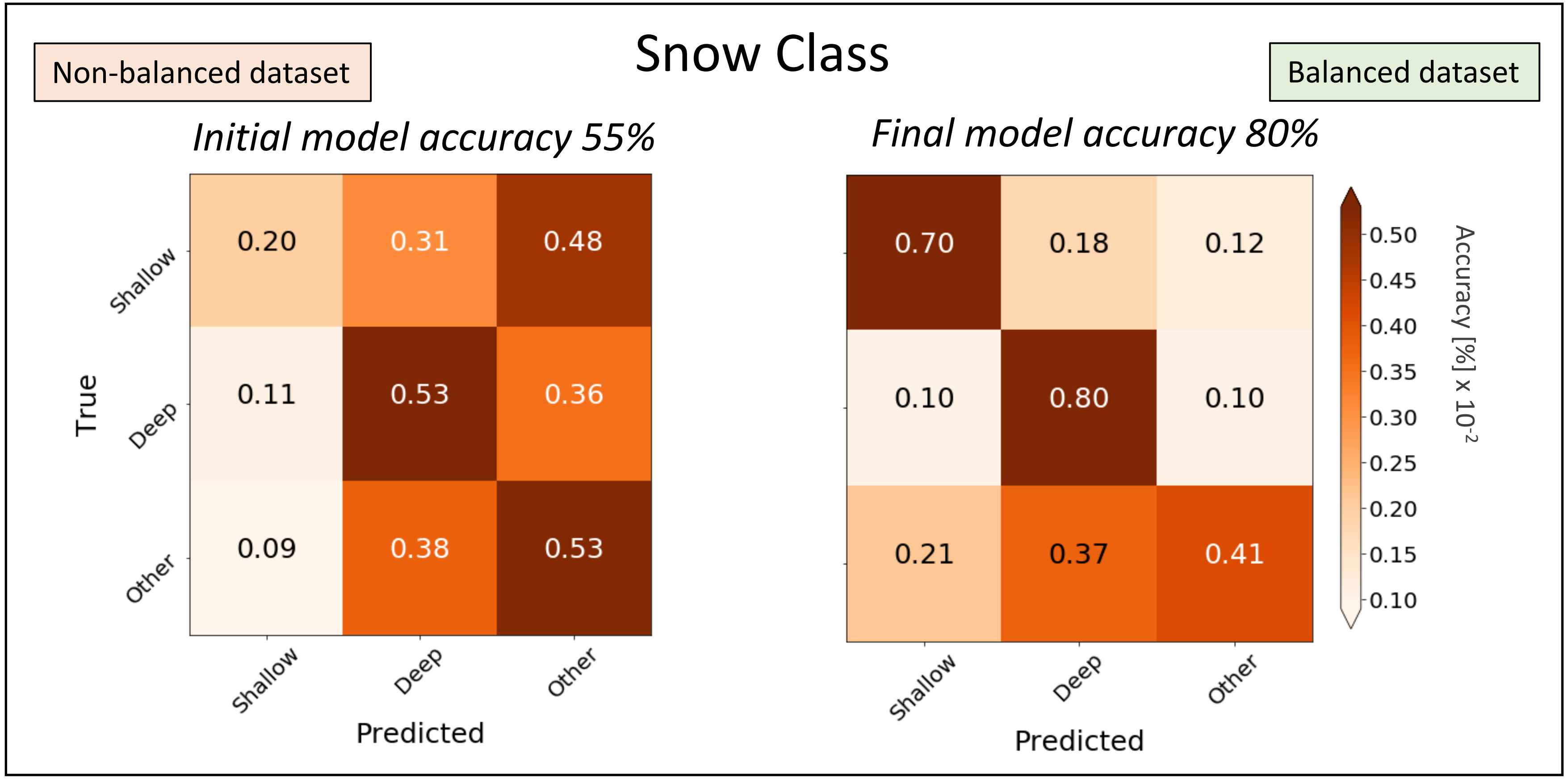
Extreme Gradient Boost – XGBoost Model



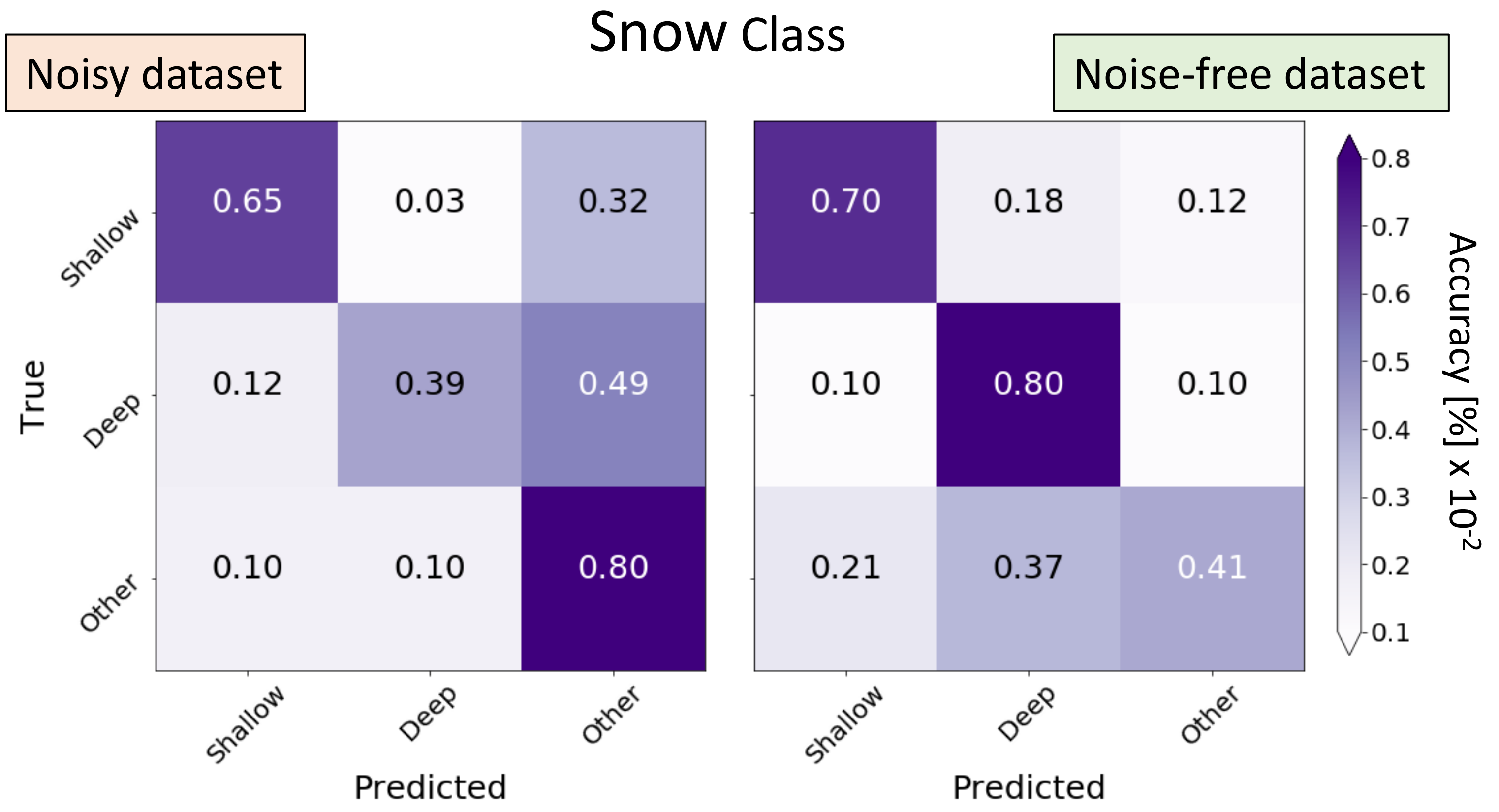
- Distributed gradient-boosted, decision tree models
- Provides parallel tree boosting. Seen as the leading machine learning library for regression, classification, and ranking problems.
- Builds on supervised machine learning, decision trees, ensemble learning, and gradient boosting.
- Predicts the label by evaluating a tree of if-then-else true/false feature questions to assess the probability of making a correct decision, for both classification and regression tasks.

EARTH SYSTEM SCIENCE
INTERDISCIPLINARY CENTER

Results – Fully Connected Neural Network

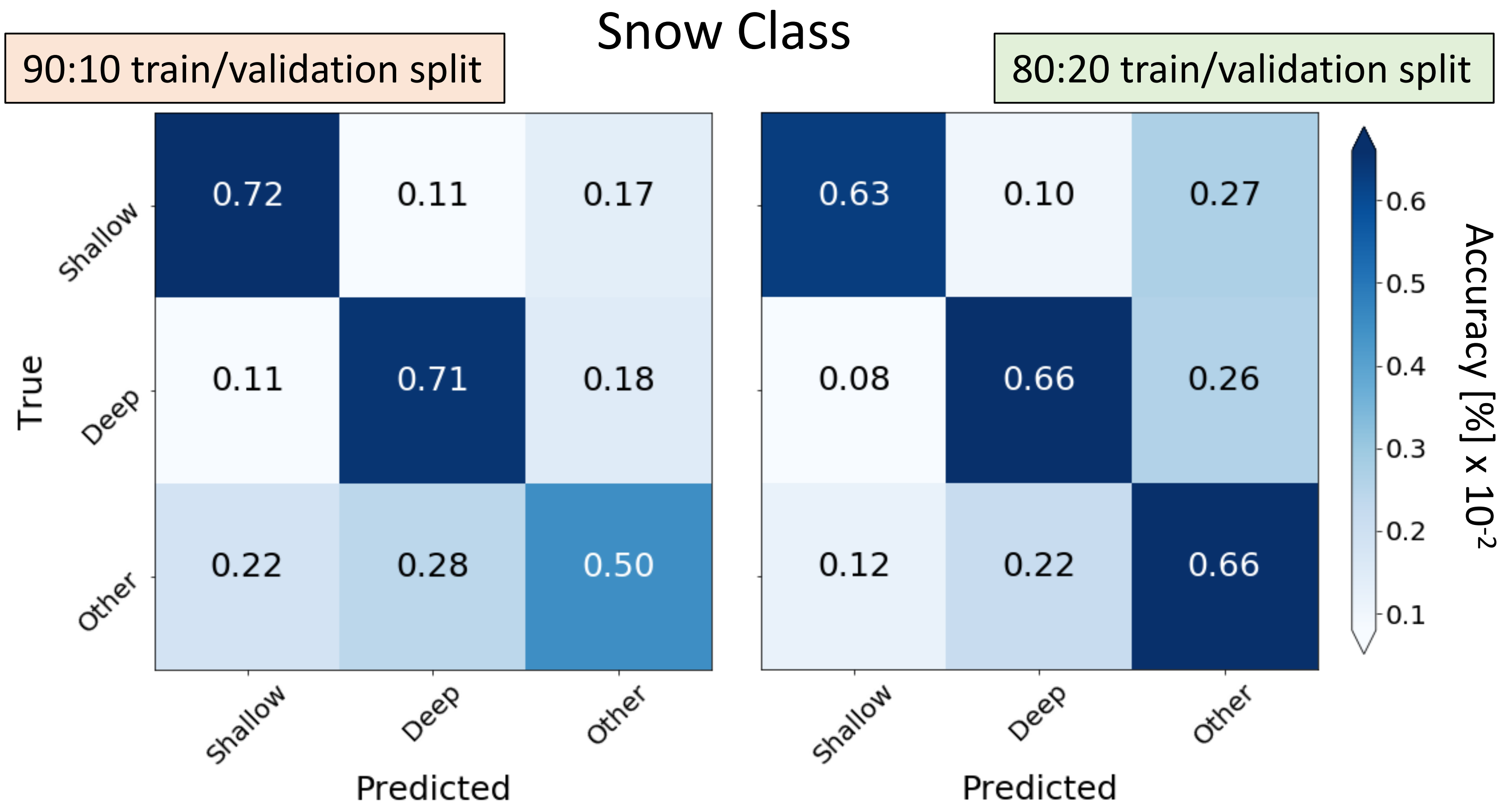


Lessons Learned: Removing the noise



The effect of missing Tbs (i.e., noise) is found to be significant. The noise present at as little as 8% of the high frequency Tbs (edge of the level1C-R swath) results in significant reduction in overall accuracy (Left: noise; Right: Noise-free).

Lessons Learned: Training to validation ratio



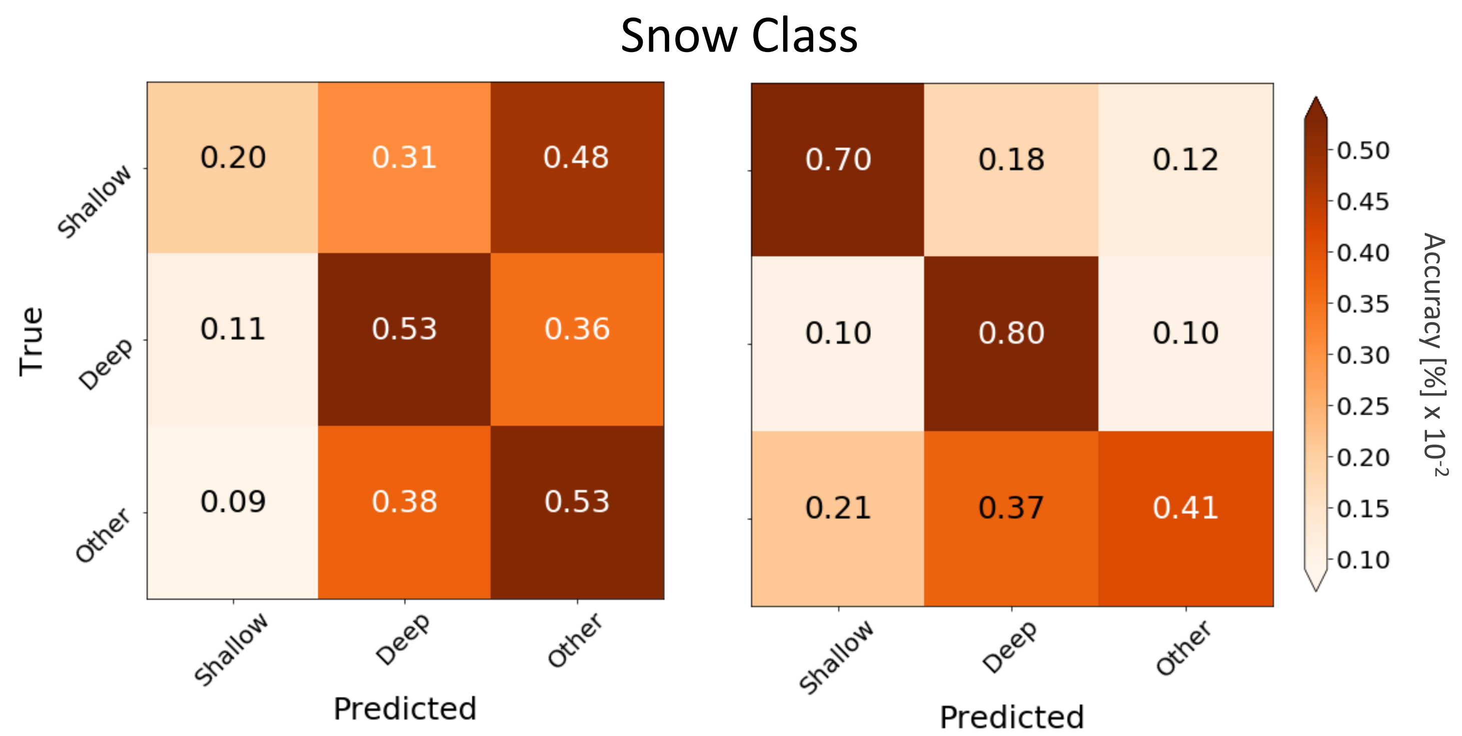
The size ratio between the training and validation datasets affects the accuracy of the model in an expected manner

Training to validation ratio:

- (left) 90:10
- (right) 80:20

EARTH SYSTEM SCIENCE
INTERDISCIPLINARY CENTER

Results – Fully Connected Neural Network



Initial overall accuracy 55% *Final overall accuracy 80%*

Noise-free
Balanced
Optimally trained model
80:20 train:validation ratio

Overall accuracy achieved: 80%

Shallow class: 70%
Deep class: 80%
Other class: 41%

Results – XGBoost

Input features: T2m, tpw, 166h-166v, 89h-89v, 36.5vh, 89h, 166vh, 183(3), 183(7)

Training (accuracy: 90%)

	Recall	F1-score	Support
Shallow	0.27	0.39	16777
Deep	0.68	0.72	27312
Other	0.13	0.22	4782
Dry	0.98	0.95	261795
Wght. avg	0.90	0.89	310666

Testing (accuracy: 89%)

	Recall	F1-score	Support
Shallow	0.18	0.28	2653
Deep	0.50	0.55	3326
Other	0.04	0.08	717
Dry	0.98	0.94	42169
Wght. avg	0.89	0.87	48866

Training	Shallow	Deep	Other	Dry
Shallow	**69**	3	8	4
Deep	4	**76**	10	3
Other	3	4	**69**	1
Dry	24	17	13	**92**

Testing	Shallow	Deep	Other	Dry
Shallow	**62**	6	18	4
Deep	9	**61**	7	3
Other	4	5	**41**	1
Dry	25	28	35	**91**

Conclusions

- PMW snowfall retrievals will benefit from snowfall regime knowledge
- The matched GPM–CloudSat dataset (2B-CSATGPM) allows to build a specific training/a-priori dataset for ML uses
- Challenges: limited dataset size
- Lessons learned: eliminating noise, ensuring balanced training, and optimizing learning process are essential when working with small data samples
- Fully Connected NN alone delivers models capable to correctly label 80% of global snowfall
- XGBoost shows good ability to handle suboptimal representation of regime distributions offered by the training dataset
- Achieved accuracy sets the scene for implementation to PMW retrieval

NASA PMM Grant #80NSSC22K0611