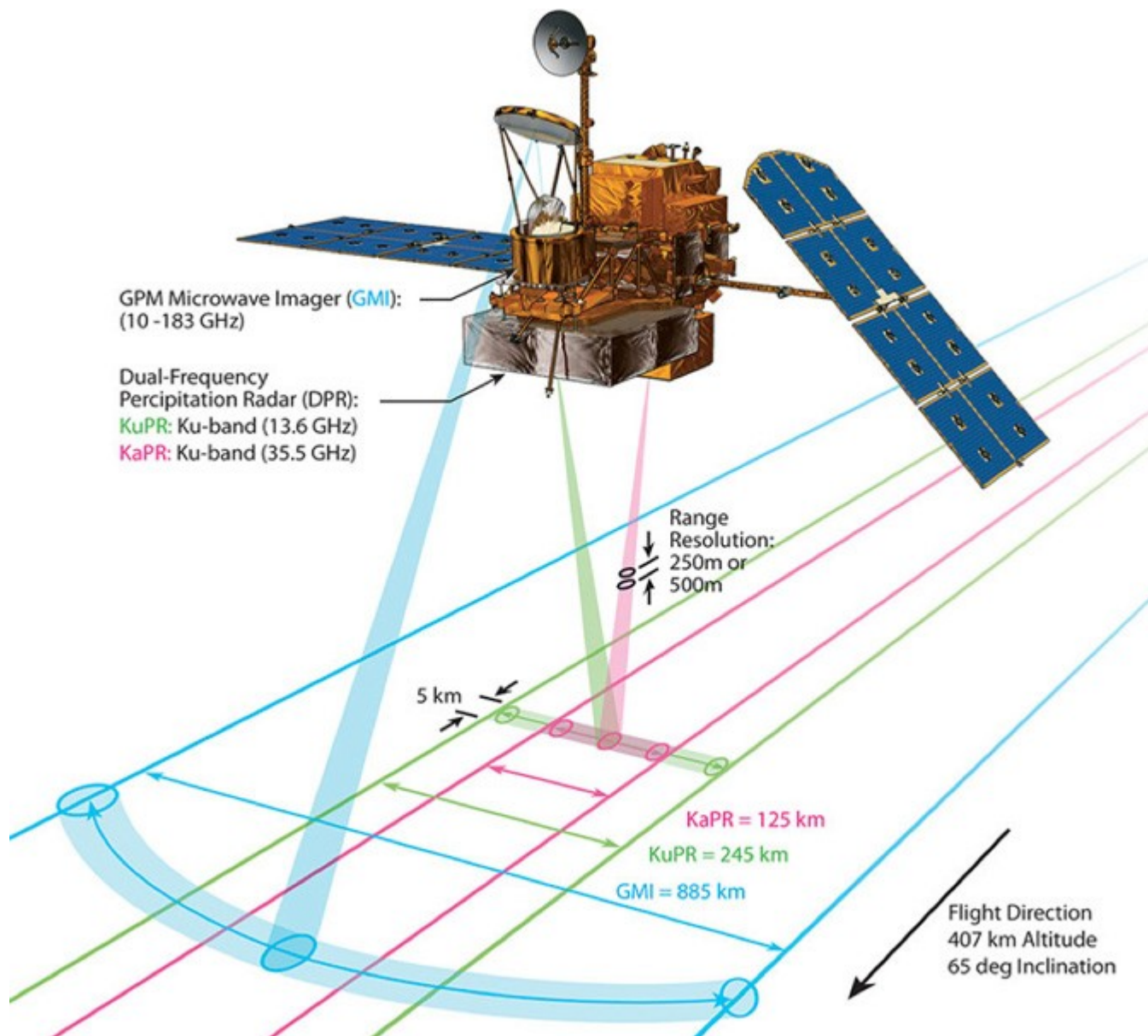


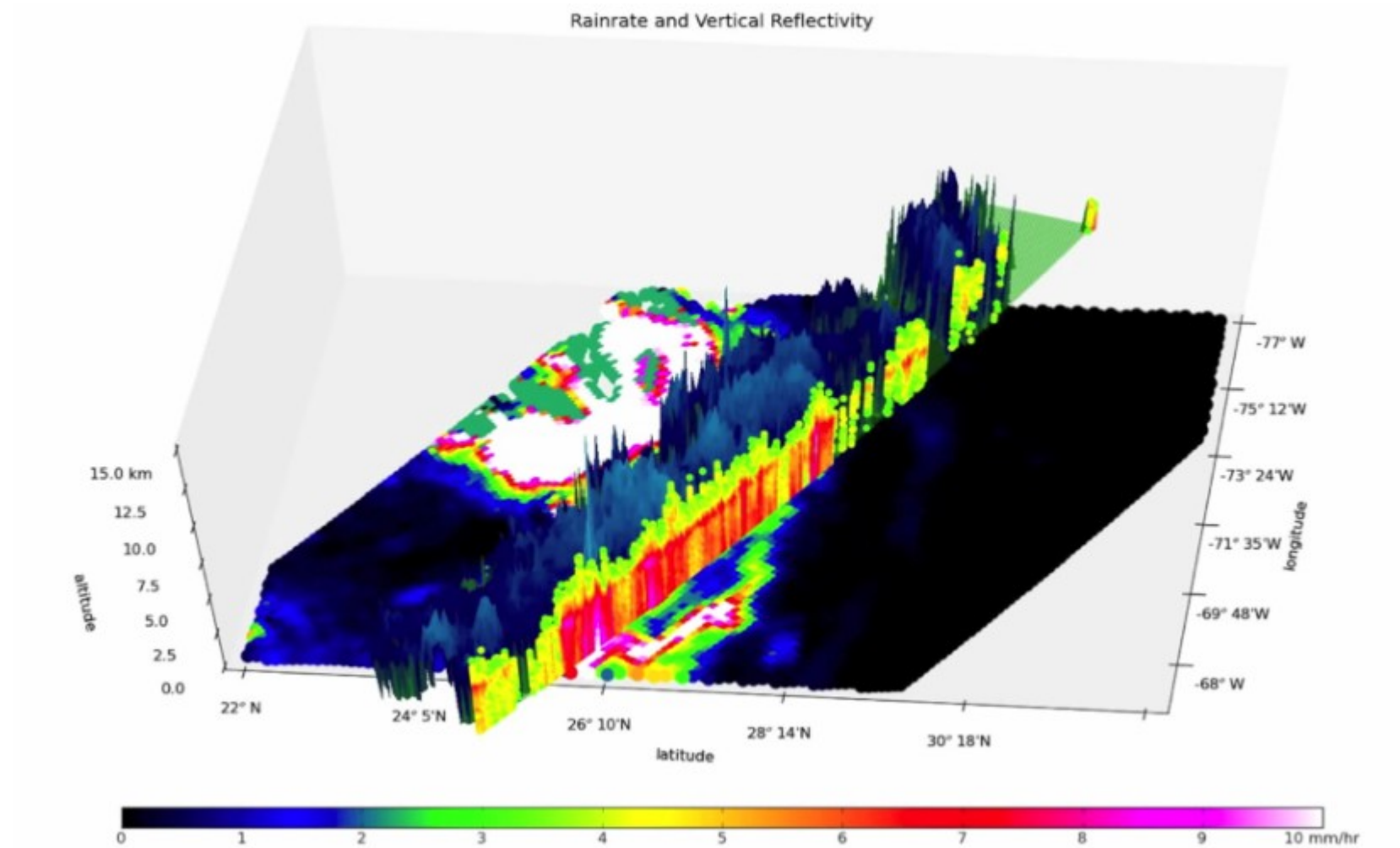


Machine Learning Based State-of-the-art Hurricane Rainfall Prediction

Ke Li



Synopsis



Content

- Data
- Bayesian Based Algorithm
 - K-means Clustering
 - Dimension Reduction
 - Naive Bayes Classifier
- Neural Networks

Data

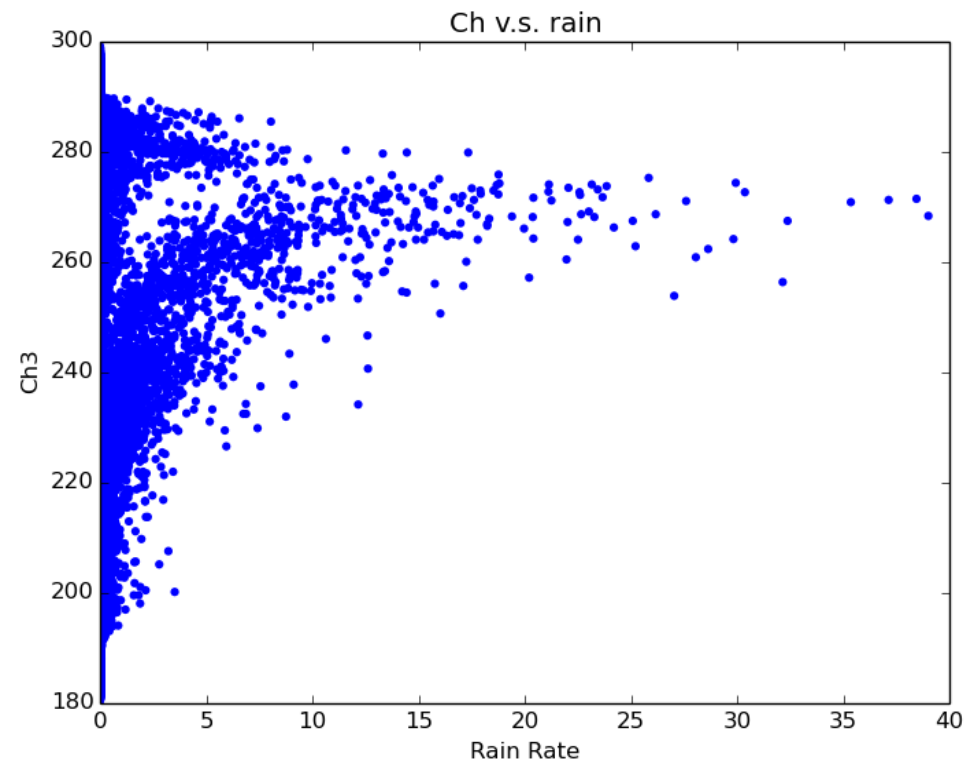
- Source: Tropical Rainfall Measuring Mission (TRMM)
 - http://trmm.gsfc.nasa.gov/data/quicklook/last_2_cal.html
- Huge!
 - 2.2 Tb, 16 times day for 20 years
- High dimensions!
 - 11 variables
- Noisy
 - Dominant signals are irrelevant to rain fall

Data

```
Loading sample data ...
```

```
Data has 67877 records, and 12 variables.
```

lat	lon	Ch1	Ch2	...	rain
-36.2	-116.3	178.7	107.7	...	0.0
-36.1	-116.3	179.2	108.6	...	0.1
-36.0	-116.3	177.9	107.9	...	0.2
-35.9	-116.4	175.8	105.8	...	0.0
-34.5	-116.4	182.4	120.4	...	0.0
-34.4	-116.4	182.9	117.3	...	0.1
-34.3	-116.4	183.0	115.3	...	0.1
-34.2	-116.3	180.4	115.6	...	0.0
-34.1	-116.3	180.8	115.7	...	0.0
-34.0	-116.3	183.9	115.4	...	0.0



K-means Clustering

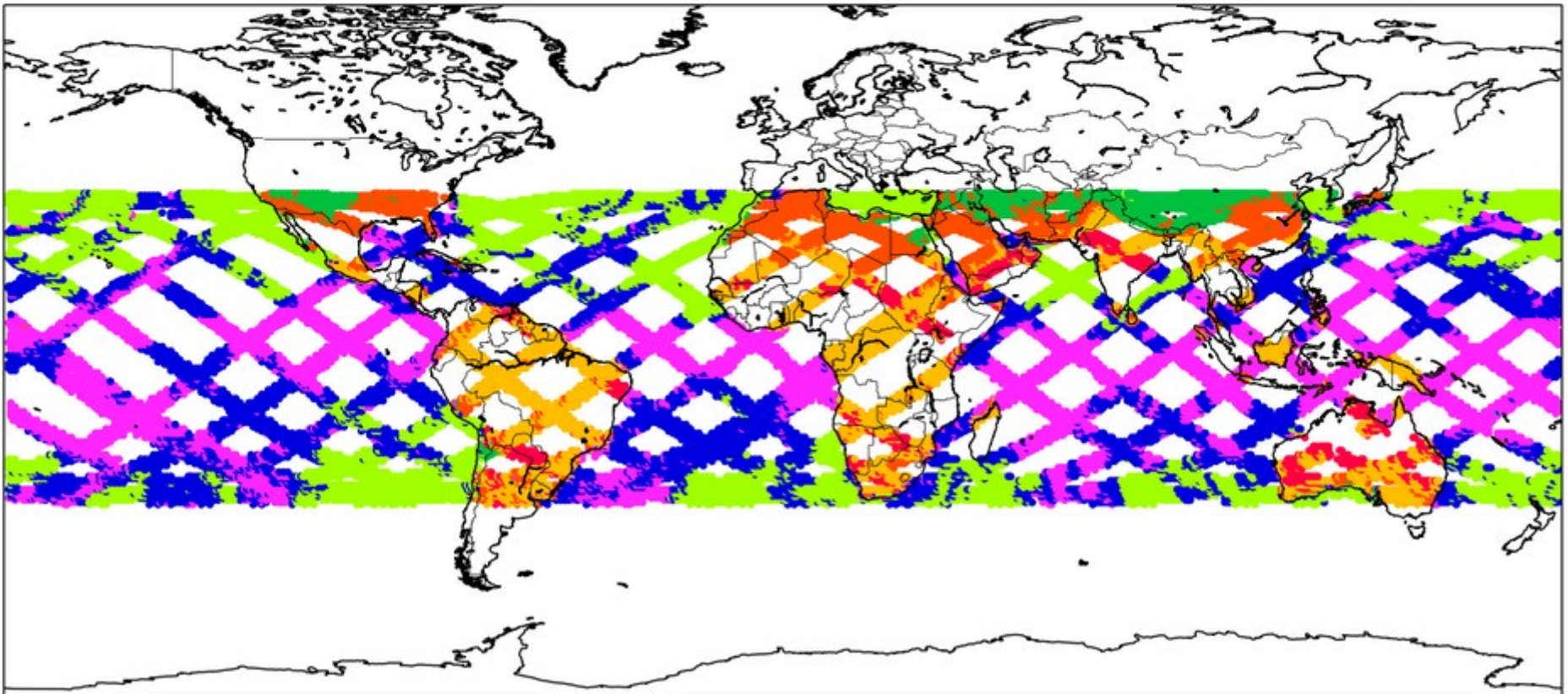
==Bin Data==

Bin data into 0.5 x 0.5 degree pixels for surface categorization.

Press any key to continue.

==K-means Clustering==

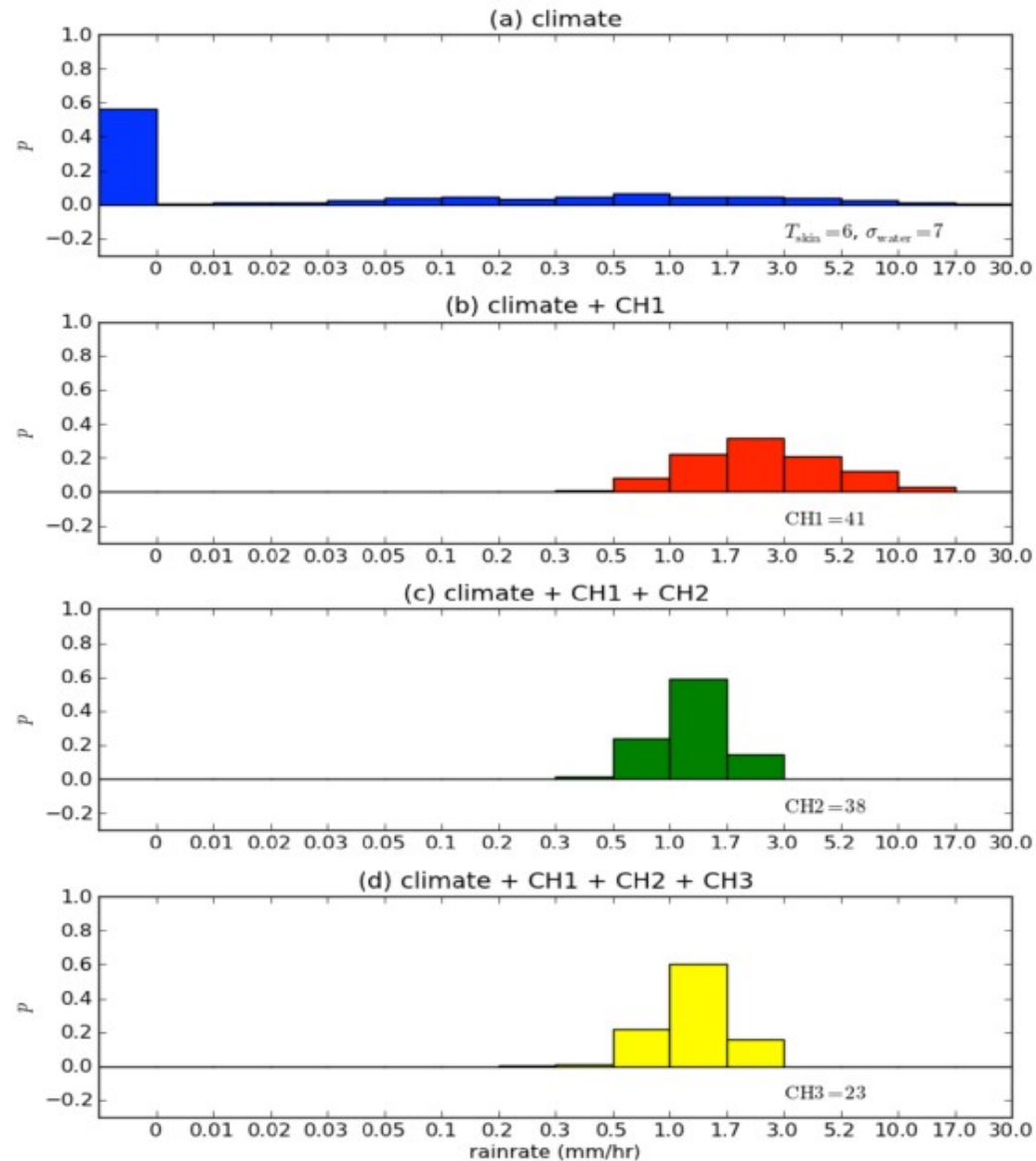
Number of Clusters: 7



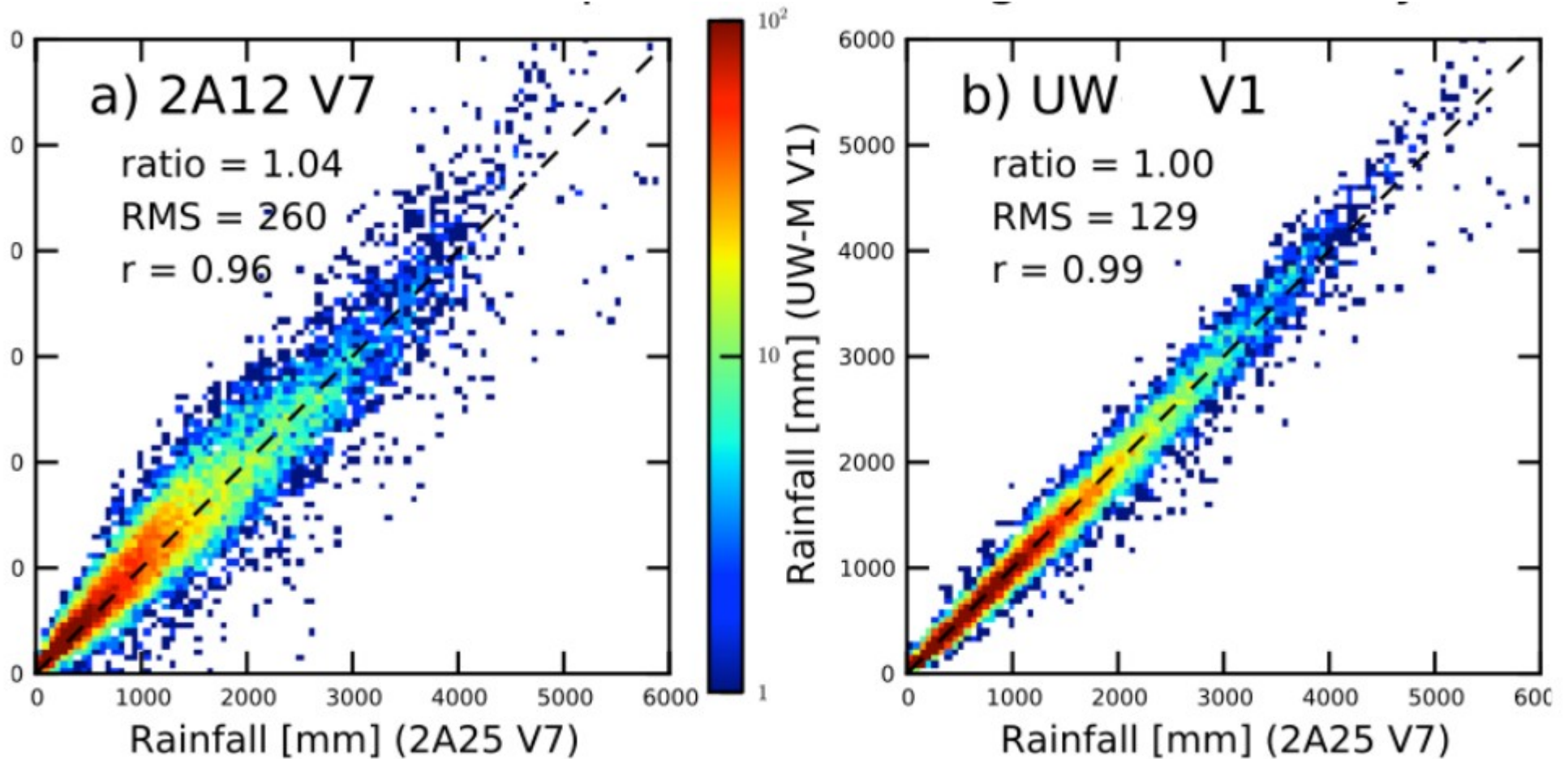
Dimension Reduction

- 9 Channels \Rightarrow 3 Pseudo Channels (PCs)
- Save computational power
(show PCs v.s. rain rate)

Naive Bayes Classifier



Result

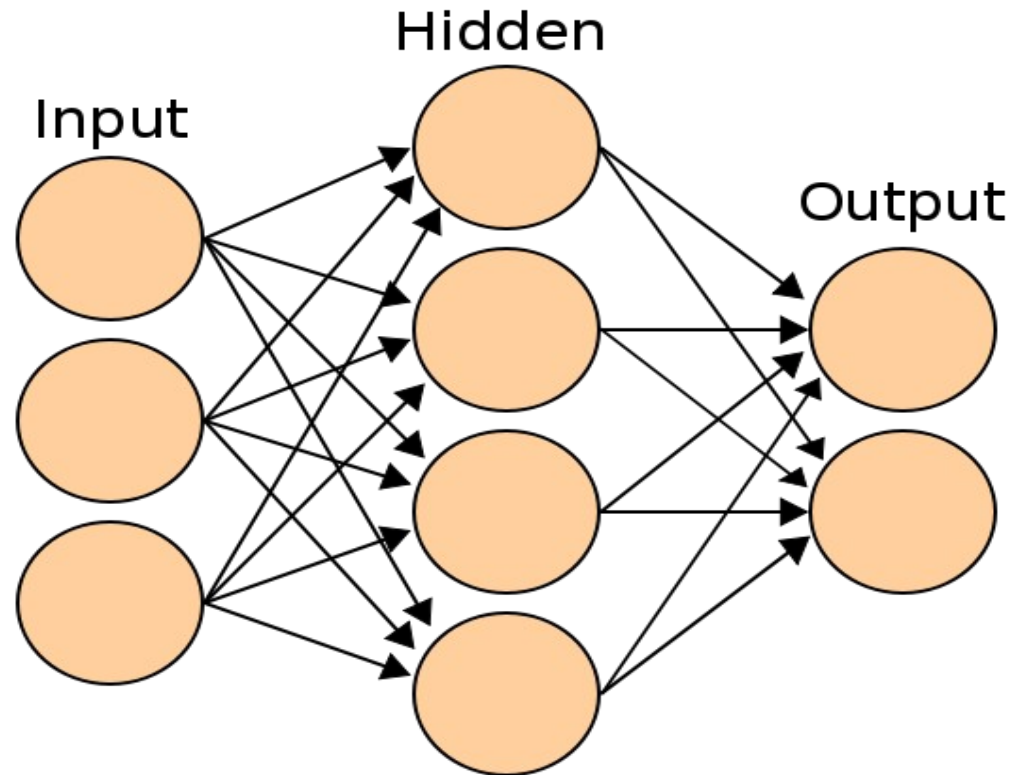


So far

- Machine Learning name:
 - Bayesian Based Algorithm
- Beats NASA previous algorithm by 3% (96% v.s. 99%)
- Currently used for Global Precipitation Measuring Mission Satellite
- Complicated (additional variables)
- Large computational power (land classification, large samples needed)

Neural Networks (NN)

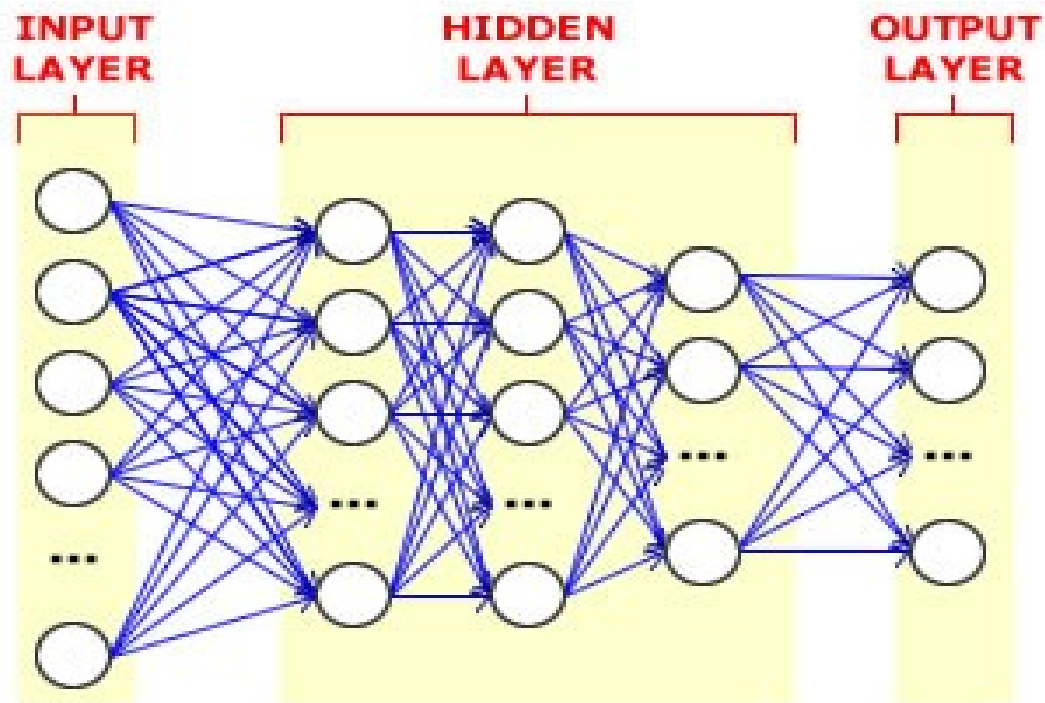
- Neurons: transmit patterns of data
 - Interconnected: Human brain
 - Receive input signals
 - Fire an output signal
- *show github



(Picture comes from "Artificial neural network" wikipedia)

Future Work

- Current NN: under-fitting
- Increase NN hidden layer number



A SIMPLE NEURAL NETWORK