AGU Fall Meeting 2019, A53H-03, 12/13/2019

Cloud Classification with Unsupervised Deep Learning

Takuya Kurihana, Ian Foster, Rebecca Willett, Sydney Jenkins, Kathryn Koenig, Ruby Werman and Elisabeth Moyer

University of Chicago, Department of Computer Science
Center for Robust Decision-making on Climate and Energy Policy

Why we need cloud classification?

- Cloud plays a dominant role in the Earth's radiation budget
- Several decades satellite observations are available for study of cloud dynamics and feedback
- Large datasets have not yet been fully utilized

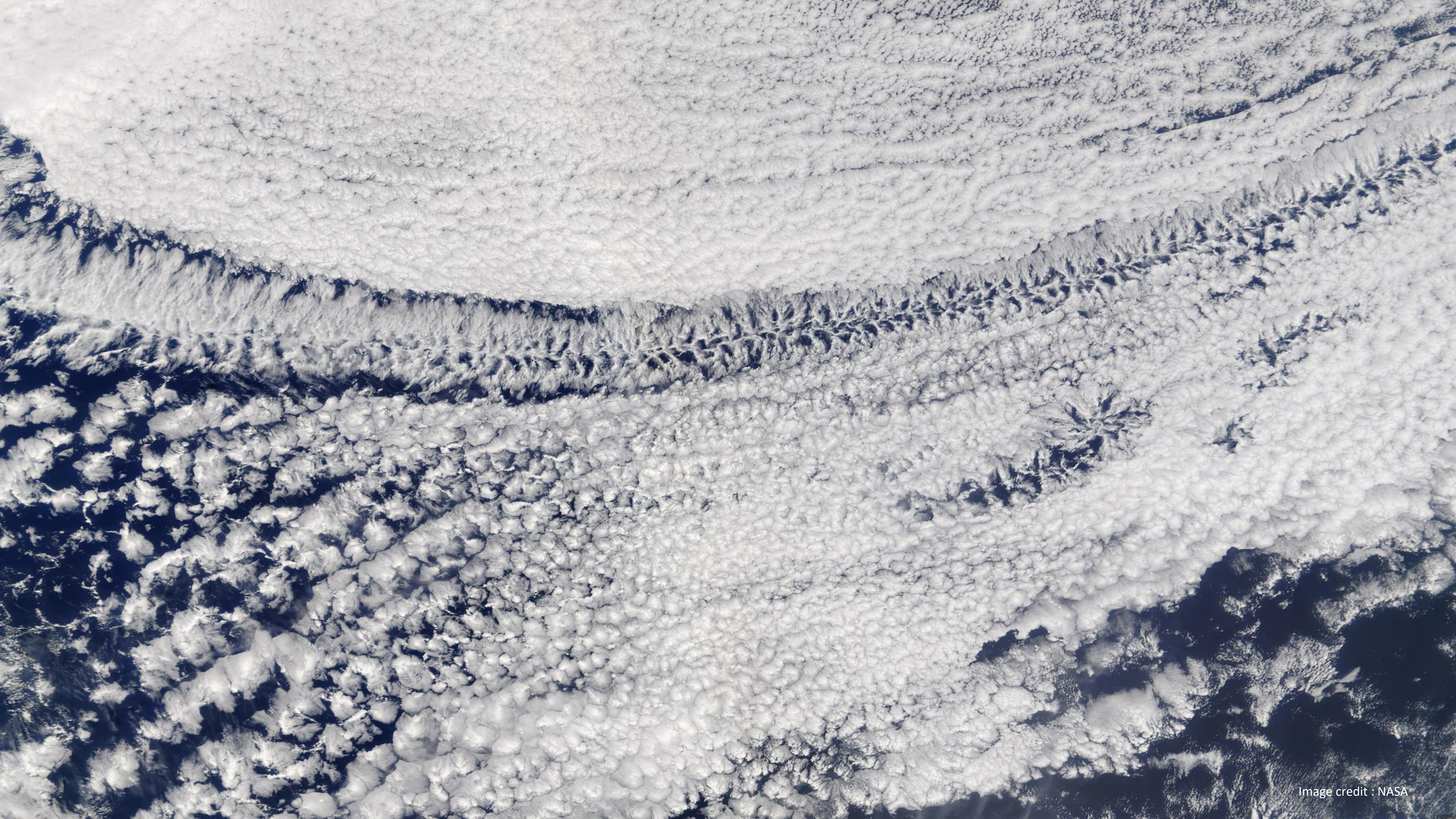


Image credit : NASA

International Satellite Cloud Climatology Project (ISCCP)

ISCCP Cloud definition

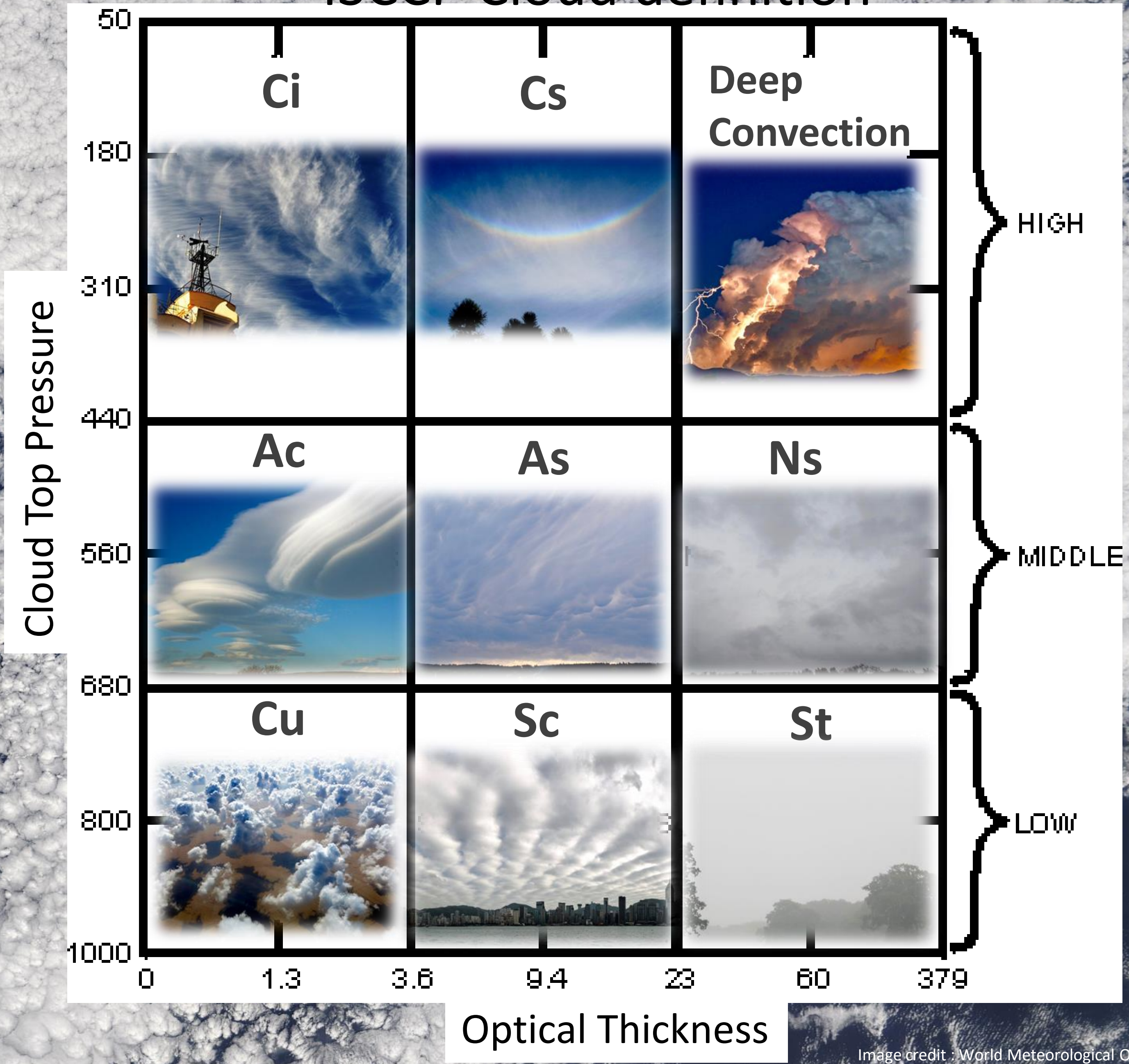


Image credit : World Meteorological Organization
International Satellite Cloud Climatology Project

International Satellite Cloud Climatology Project (ISCCP)

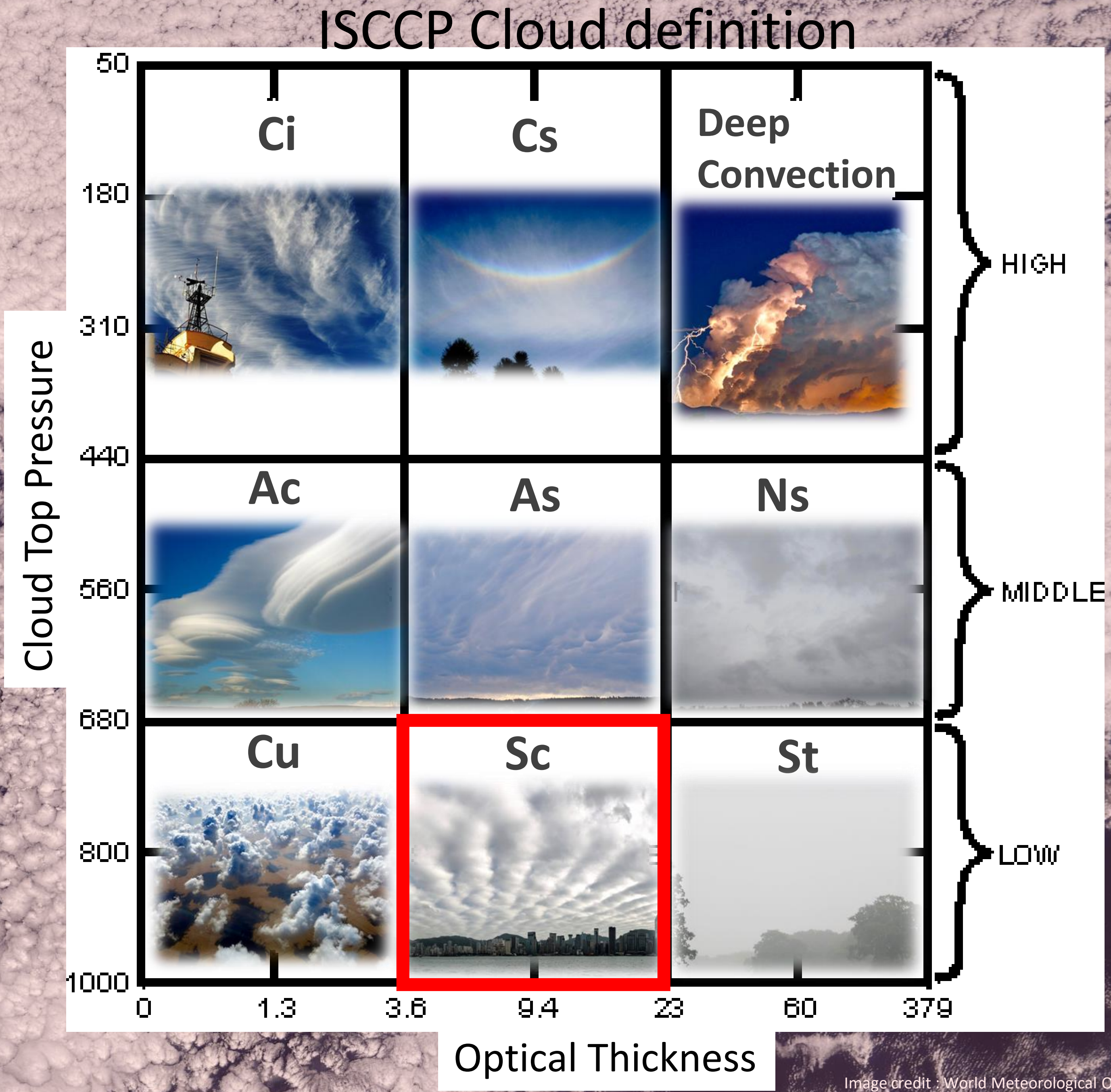
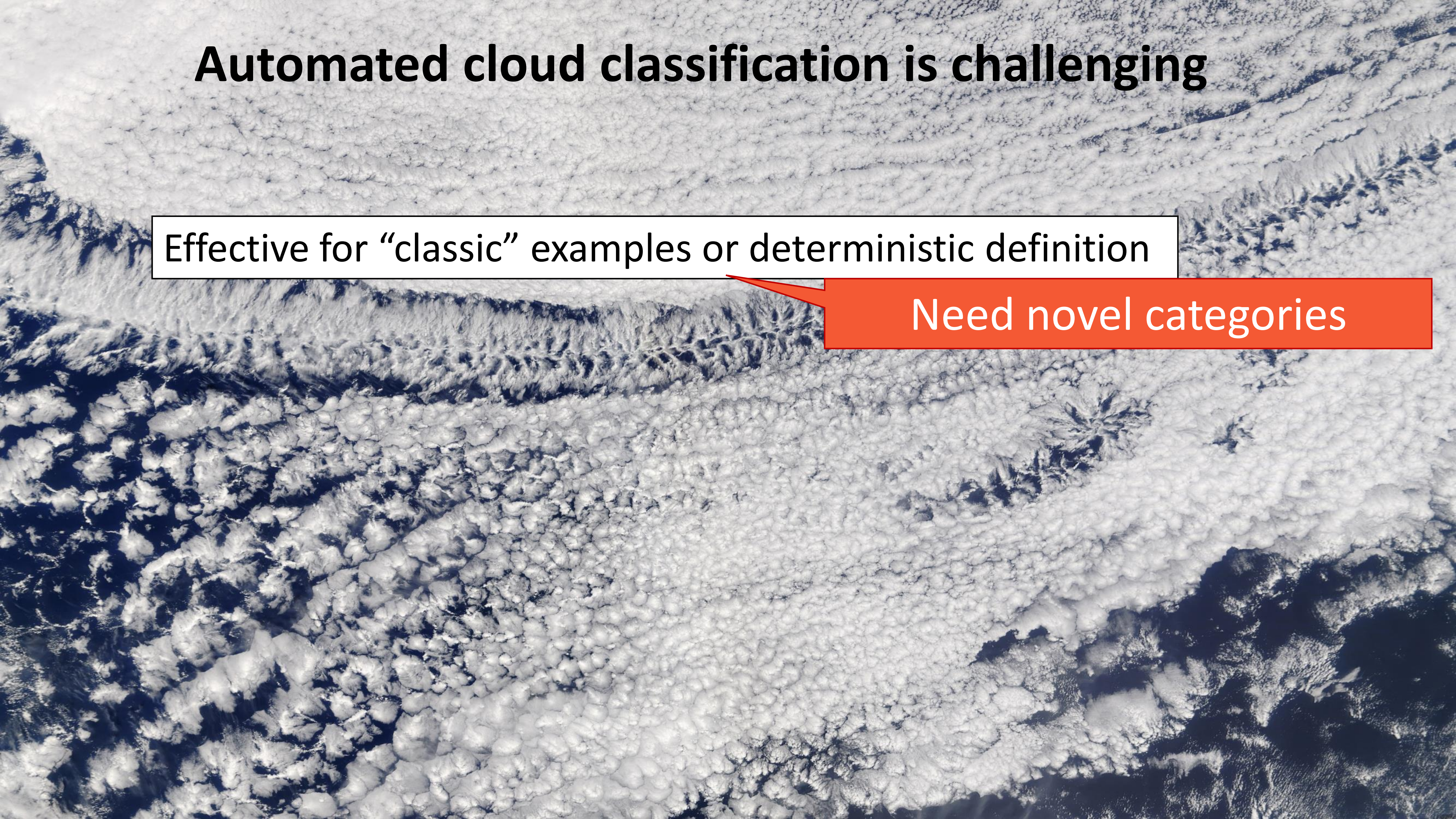


Image credit : World Meteorological Organization
International Satellite Cloud Climatology Project

Automated cloud classification is challenging

Automated cloud classification is challenging



Effective for “classic” examples or deterministic definition

Need novel categories

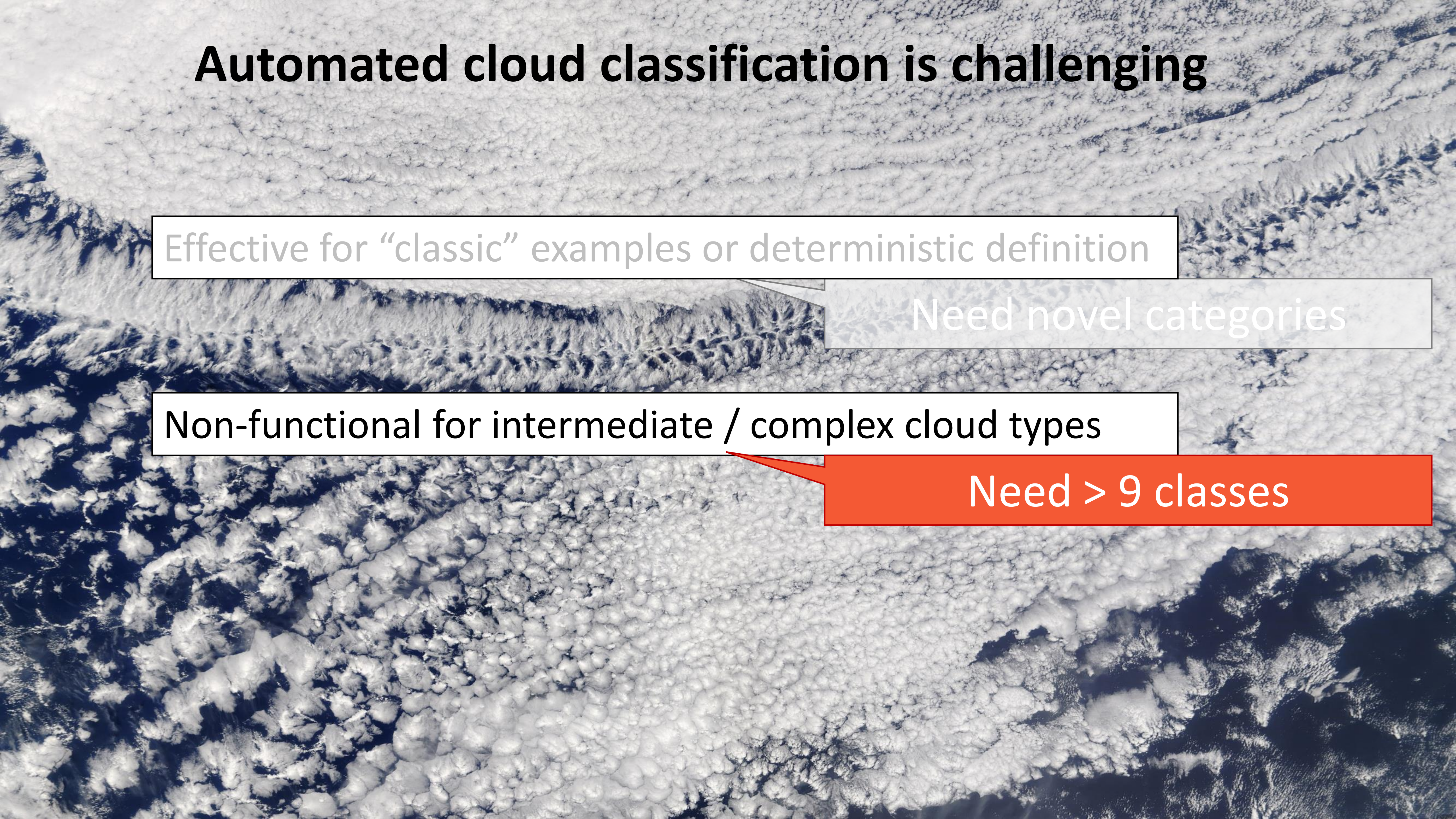
Automated cloud classification is challenging

Effective for "classic" examples or deterministic definition

Need novel categories

Non-functional for intermediate / complex cloud types

Need > 9 classes



Automated cloud classification is challenging

Effective for “classic” examples or deterministic definition

Need novel categories

Non-functional for intermediate / complex cloud types

Need > 9 classes

Mean cloud properties do not capture relevant physics

Use spatial information

Automated cloud classification is challenging

Effective for “classic” examples or deterministic definition

Need novel categories

Non-functional for intermediate / complex cloud types

Need > 9 classes

Mean cloud properties do not capture relevant physics

Use spatial information

Unsupervised learning enables novel and detailed classification

Prototype of unsupervised learning framework

Define	A) Define criteria for successful cloud classification
Process	B) Choose and process satellite data
Reduce	C) Reduce dimension with autoencoder
Cluster	D) Cluster resulting latent representation
Evaluate	E) Evaluate cloud classes against criteria

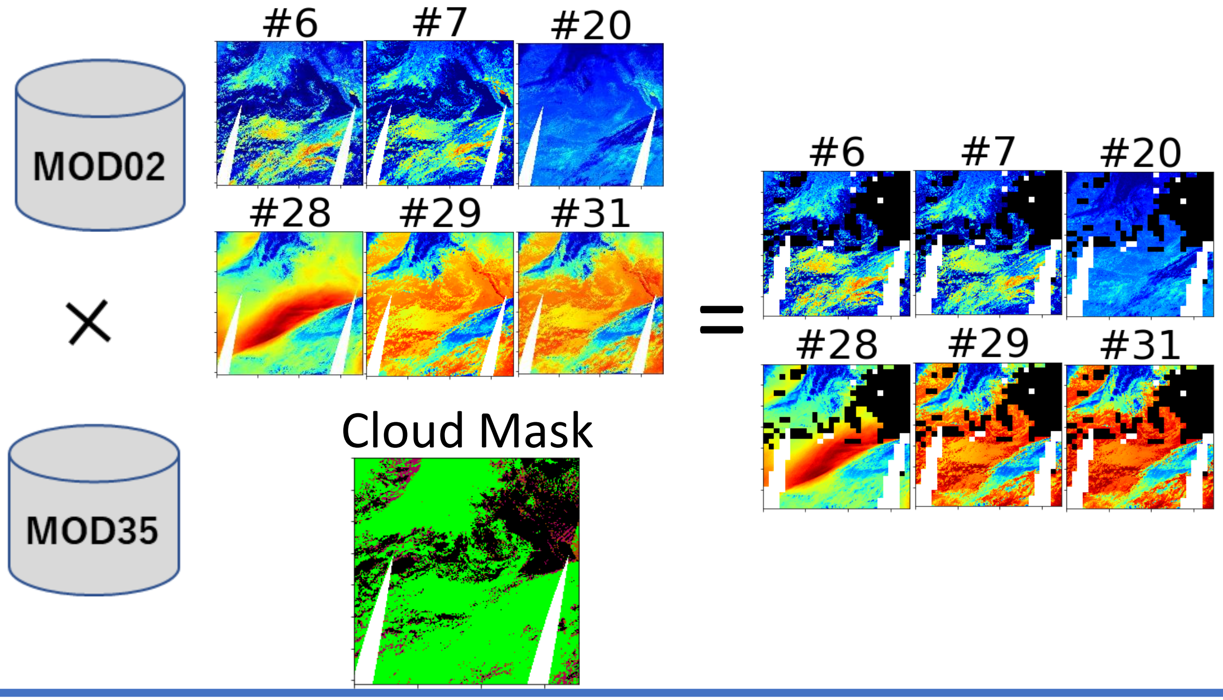
A) Define criteria for successful cloud classification

1. Physical reasonableness
2. Spatial coherence
3. Separable clusters
4. Rotational invariance
5. Cluster stability

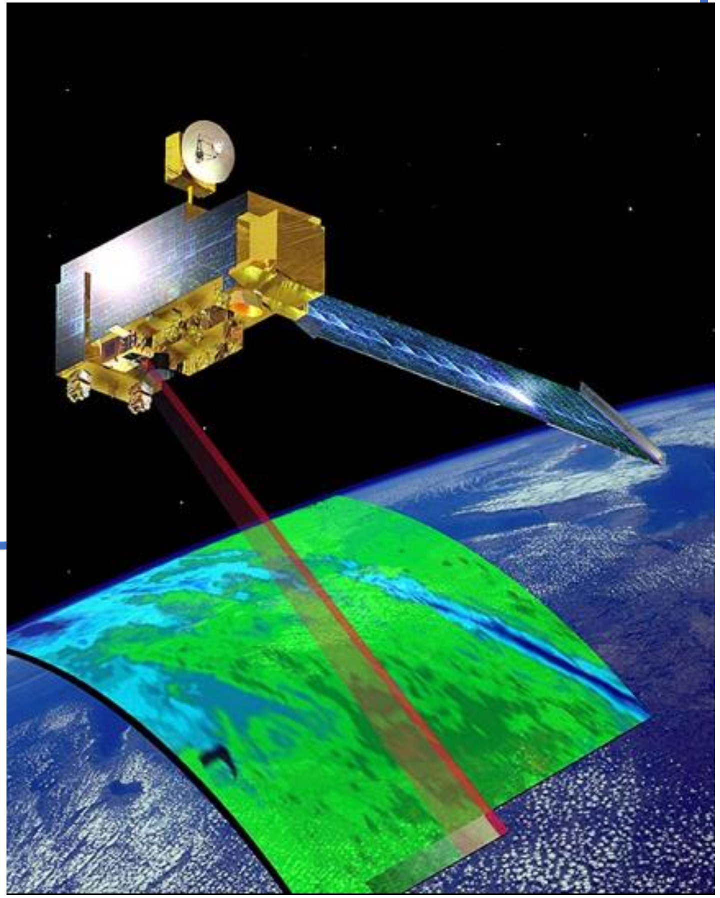
A) Define criteria for successful cloud classification

1. Physical reasonableness
2. Spatial coherence
3. Separable clusters
4. Rotational invariance
5. Cluster stability

B) Choose and process satellite data

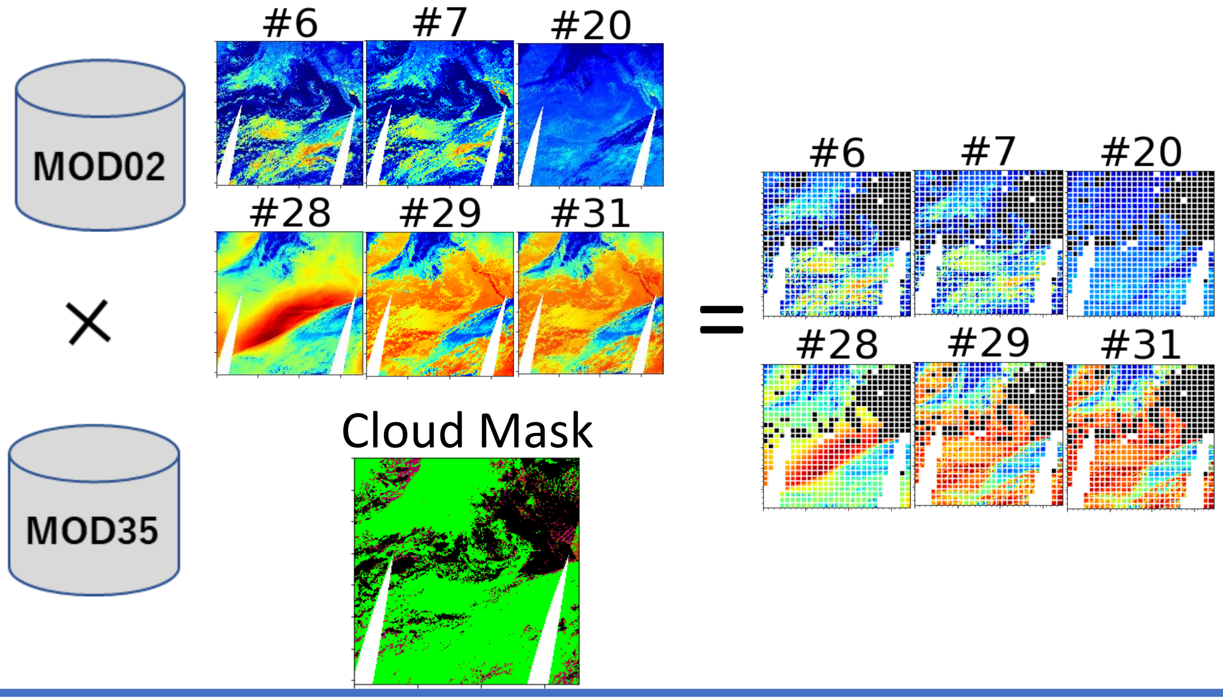


Moderate Resolution Imaging Spectroradiometer (MODIS)

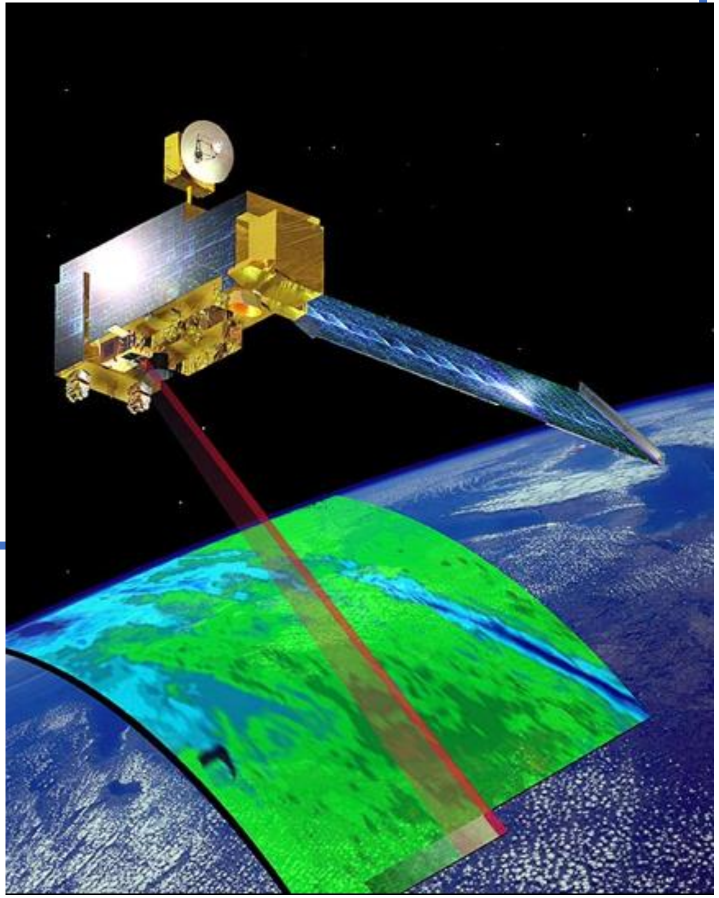


- MODIS satellite data, IR and near-IR radiances
- Use 6 radiance bands (MOD02)
- Cloud mask to identify cloud areas (MOD35)

B) Choose and process satellite data

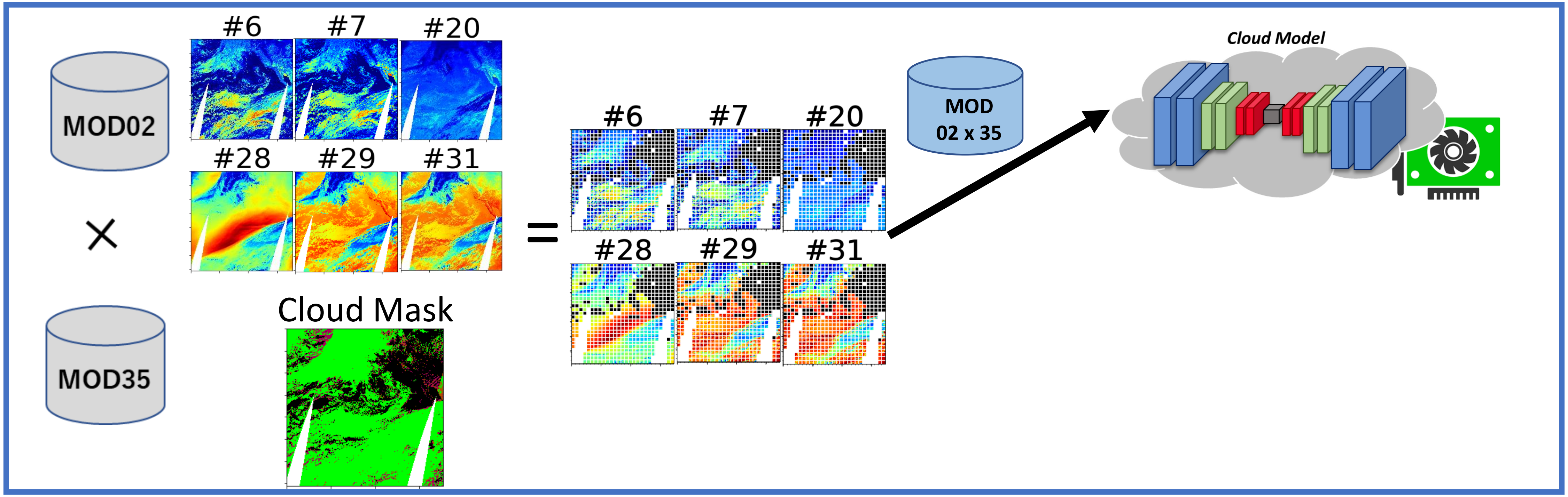


Moderate Resolution Imaging Spectroradiometer (MODIS)



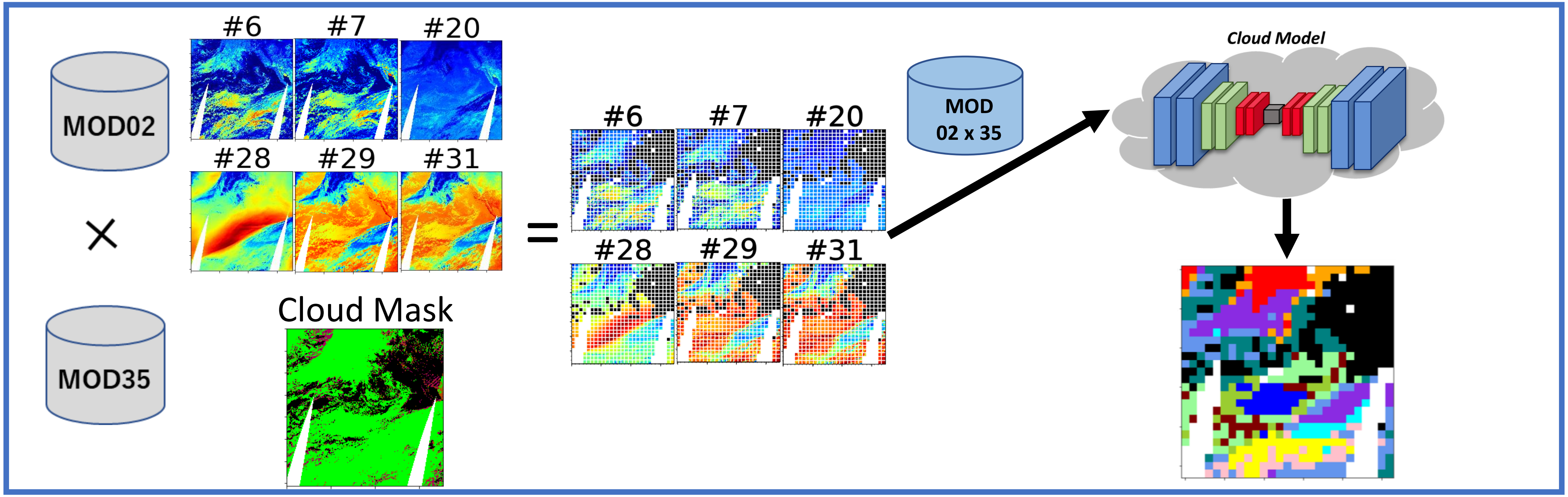
- 128 pixel x 128 pixel x selected 6 bands
- QC : Patches containing more than 30% of cloud pixels

C) Reduce dimension with autoencoder



- Train autoencoder
- Extract dimensional reduction information

D) Cluster resulting latent representation

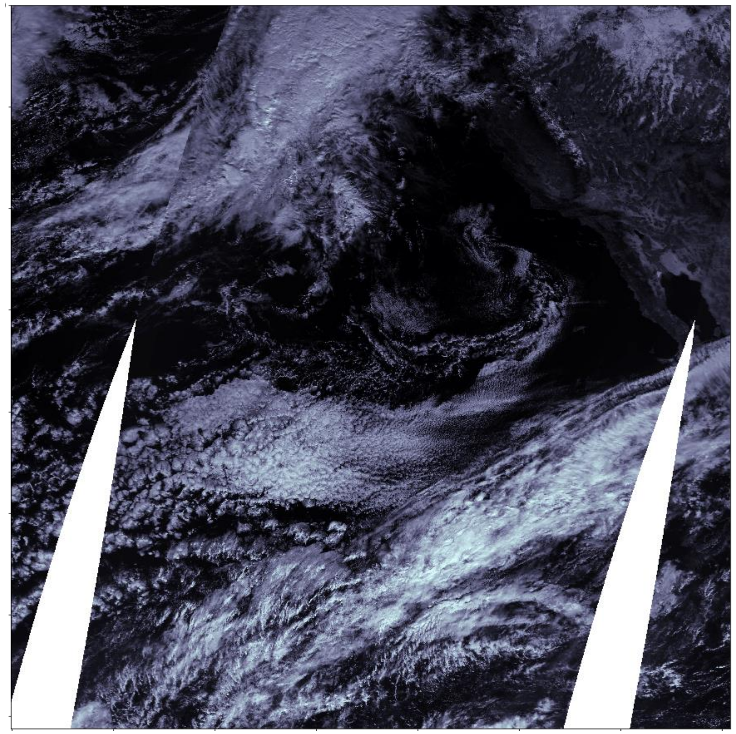


- Hierarchical Agglomerative Clustering (HAC)
  - Merge data as tree structure, minimizing variance among merging classes
  - Cluster based on Euclidian distance

E) Evaluate cloud classes against criteria

Example image:
coast of California, Dec. 1, 2015

12 clusters

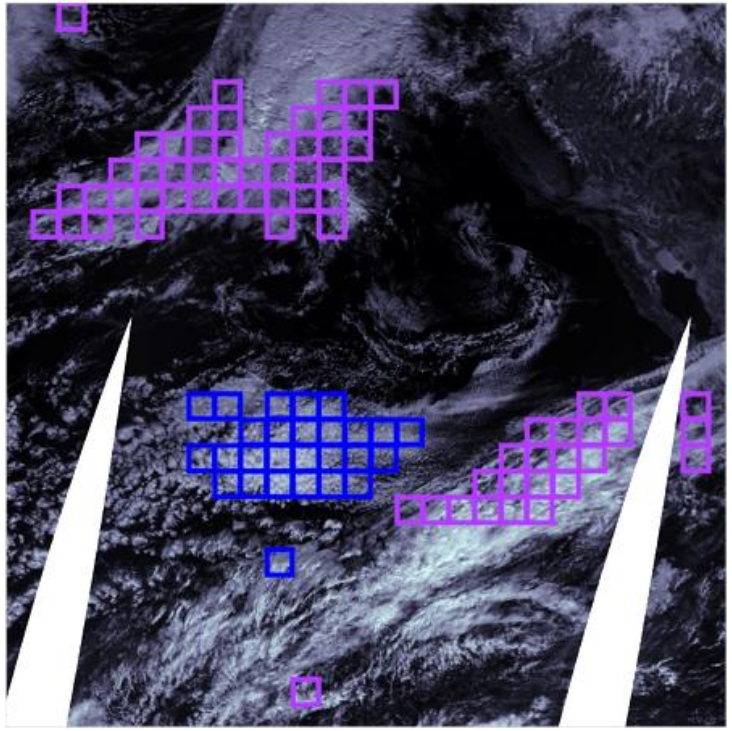


E) Evaluate cloud classes against criteria

Example image:
 coast of California, Dec. 1, 2015

12 clusters

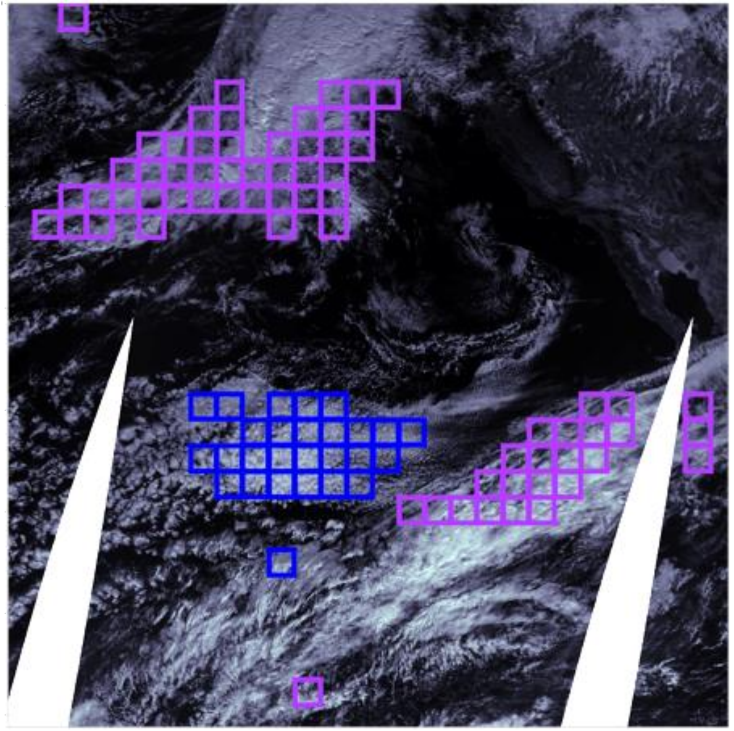
Raw visible image with clusters #0 & #2

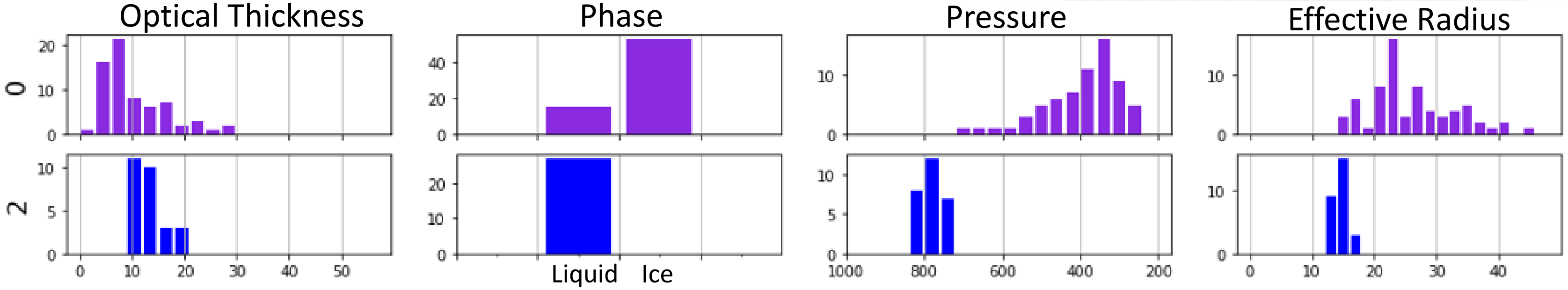


E) Evaluate cloud classes against criteria

Test – physical reasonableness

Raw visible image with clusters #0 & #2

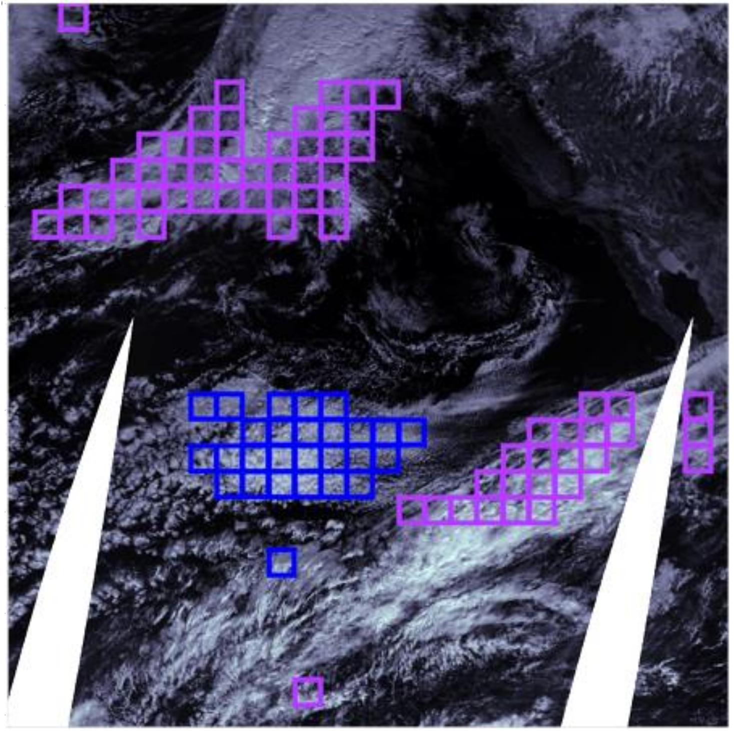




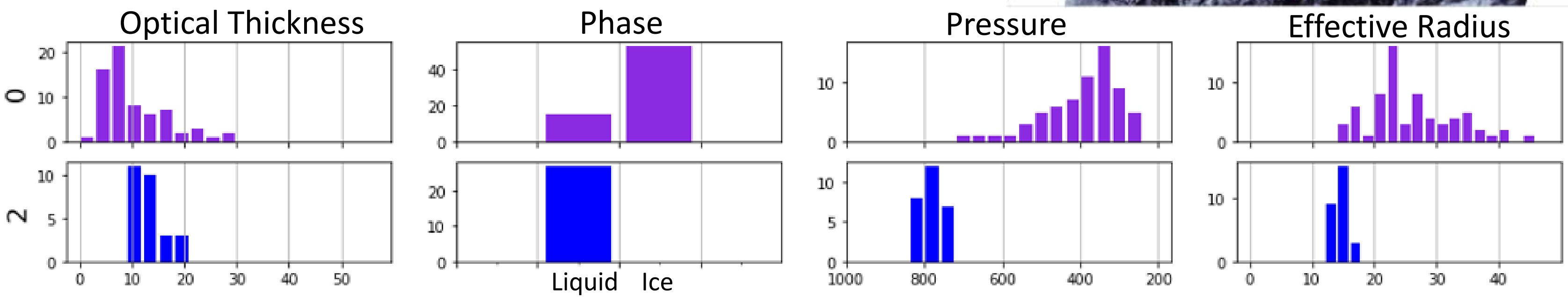
E) Evaluate cloud classes against criteria

Test – physical reasonableness

Raw visible image with clusters #0 & #2



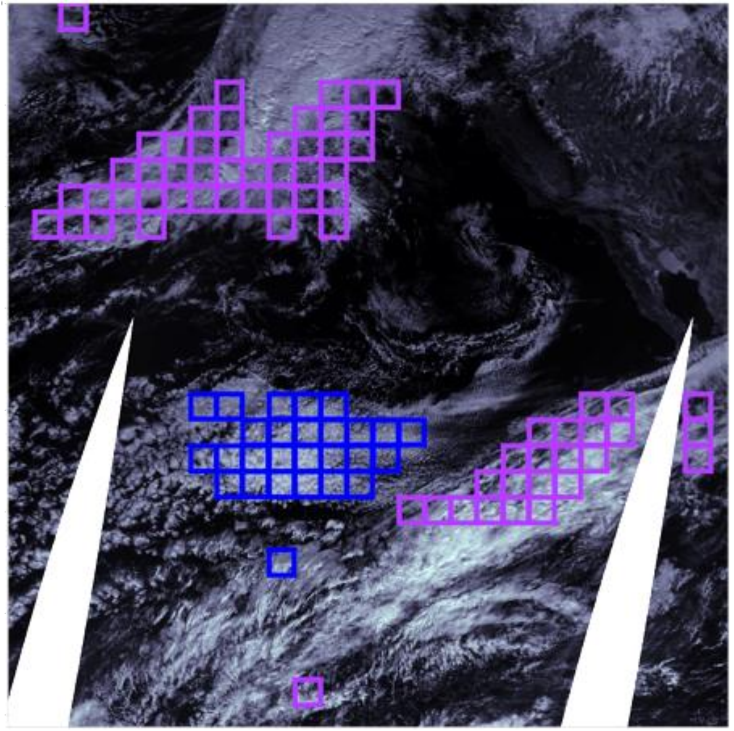
Outflow Cirrus

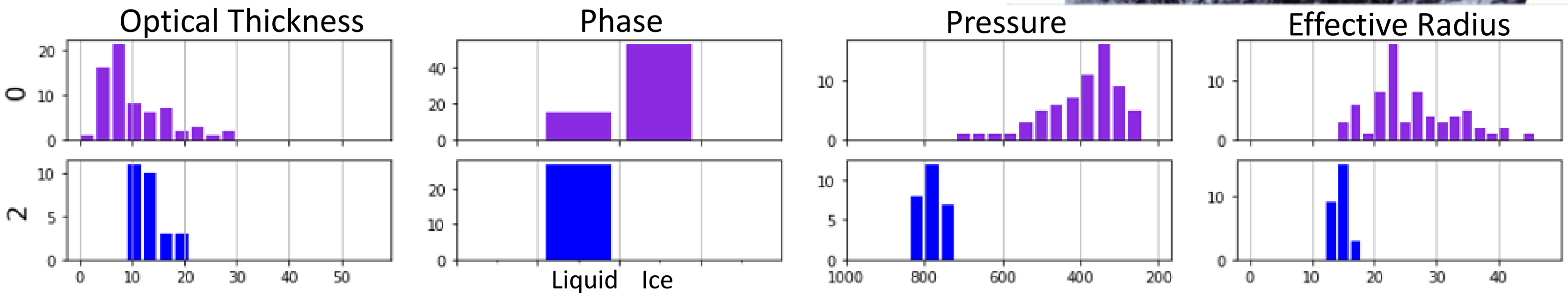


E) Evaluate cloud classes against criteria

Test – physical reasonableness

Raw visible image with clusters #0 & #2

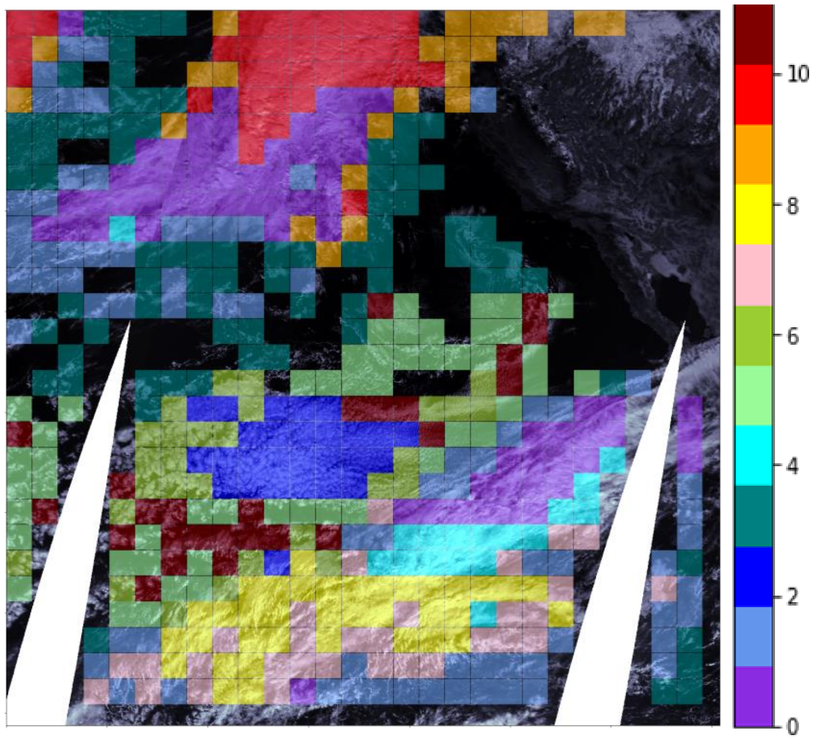




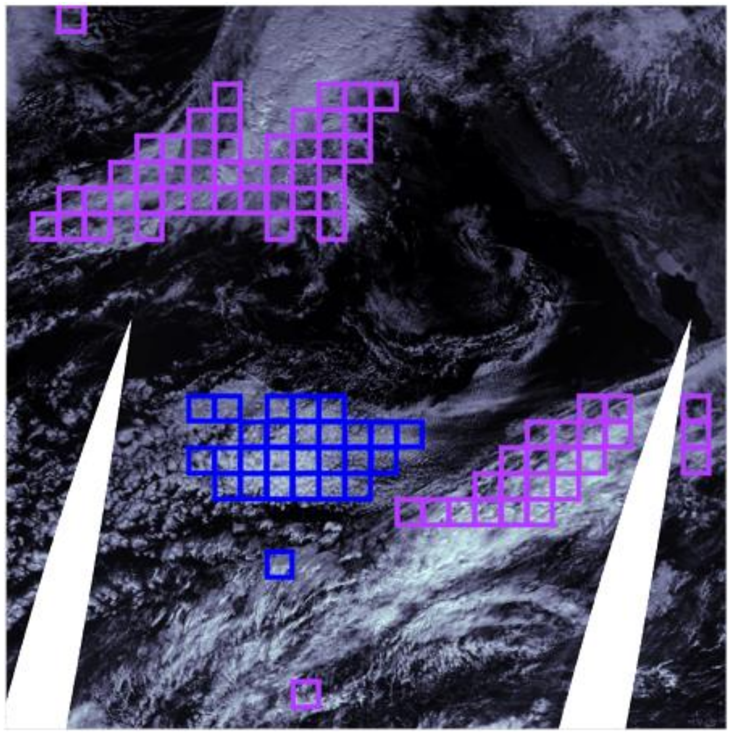
Low
stratocumulus

E) Evaluate cloud classes against criteria

All 12 Clusters

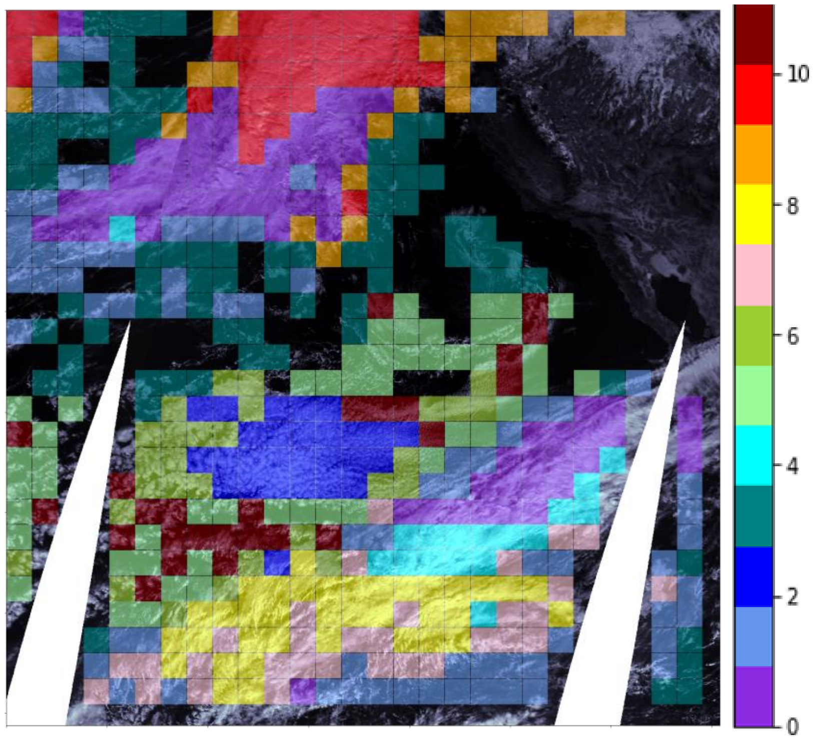


Raw visible image with clusters #0 & #2

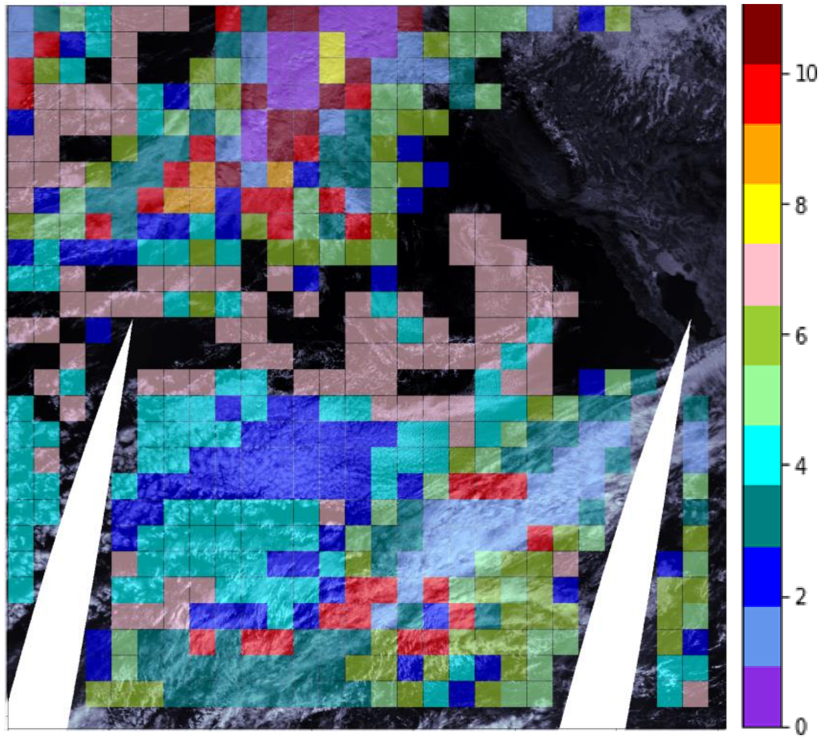


E) Evaluate cloud classes against criteria

All 12 Clusters



Clusters based on patch-mean physics values

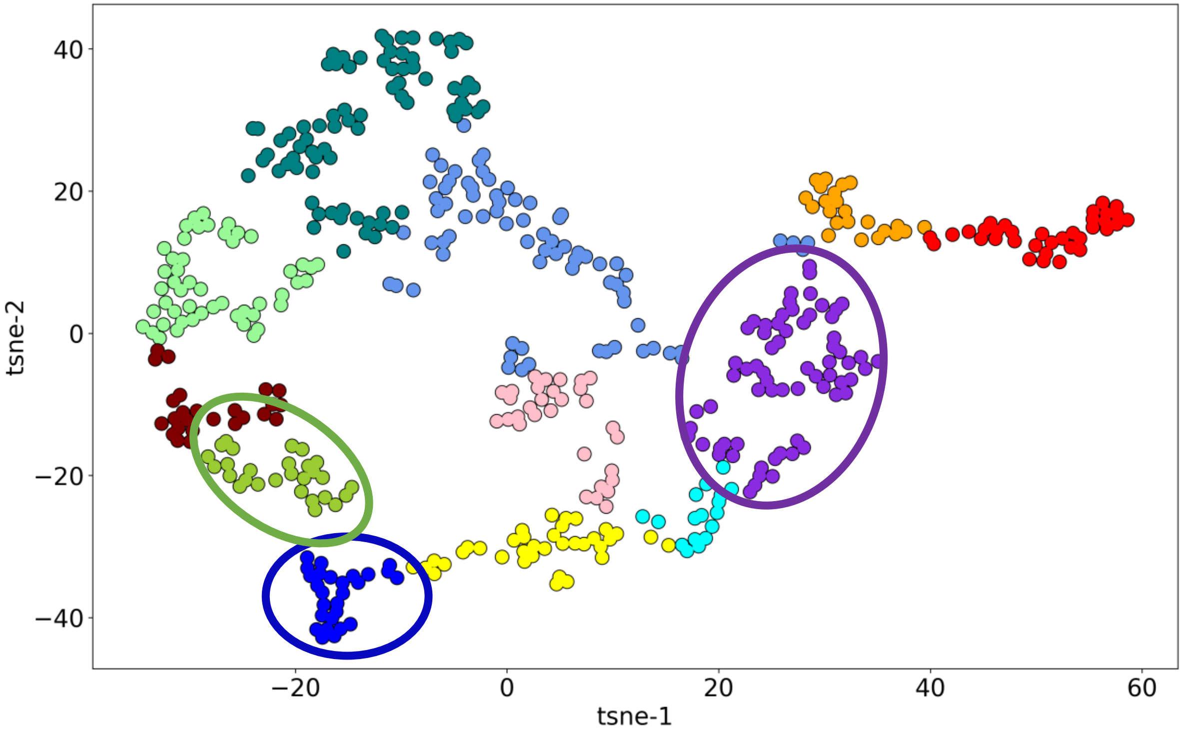


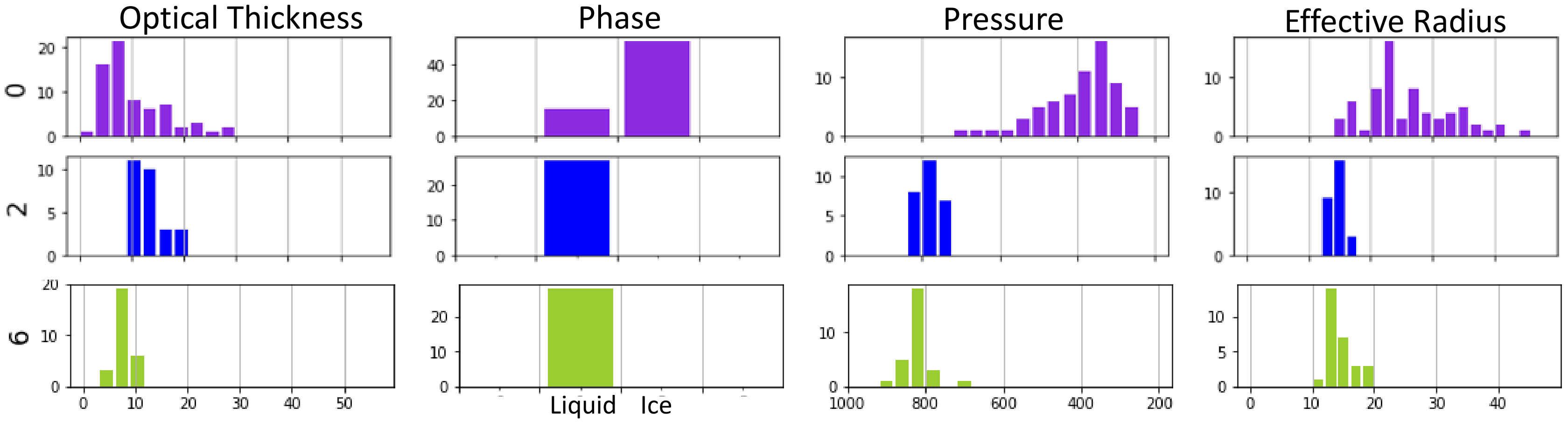
Test – spatial coherence

Clustering with autoencoder produces more spatial contiguity even than clustering using mean physics values

E) Evaluate cloud classes against criteria

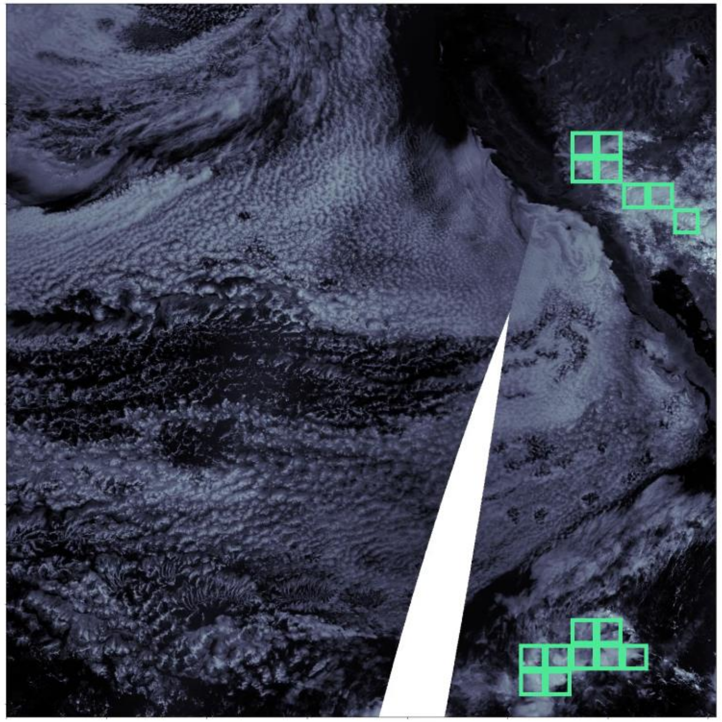
Test – separable clusters



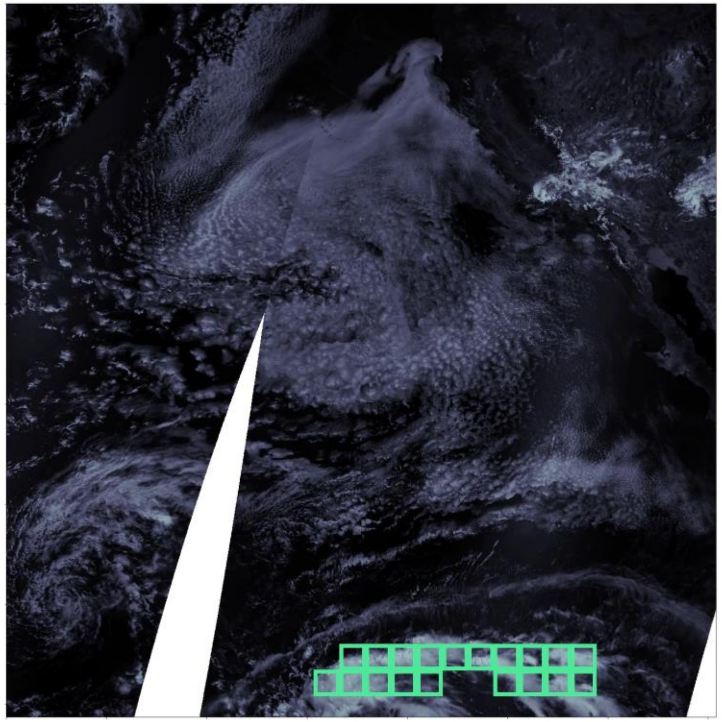


Examples: spatial-temporal evolution

Clusters #6 2012



Clusters #6 2015



2- and 3-week summer timeseries, CA coast, 14 clusters

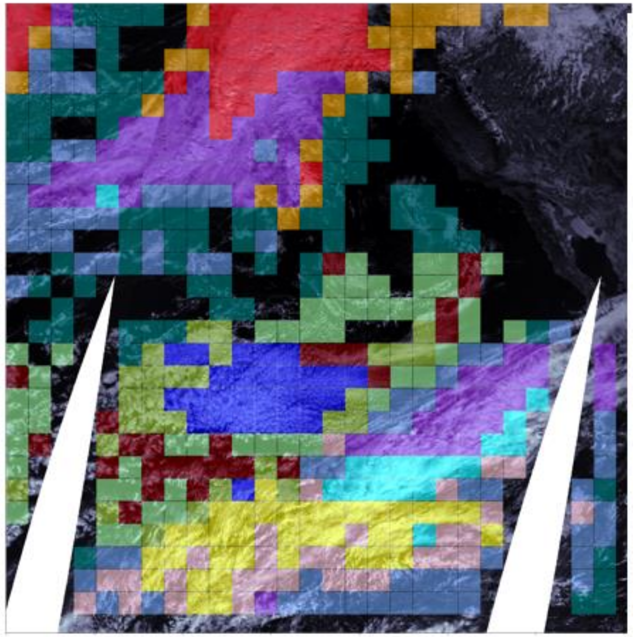
coherent evolution in time of cloud class shows advection

Conclusions & future work

- Prototyped a fully unsupervised cloud classification framework
- Produced physically reasonable clusters coherent in space and time
- Demonstrated promise of unsupervised deep learning

- Next steps
    - Rotational invariance
    - Cluster stability

Kurihana. T, et al 2019, Proceedings of the 9th International Workshop on Climate Informatics

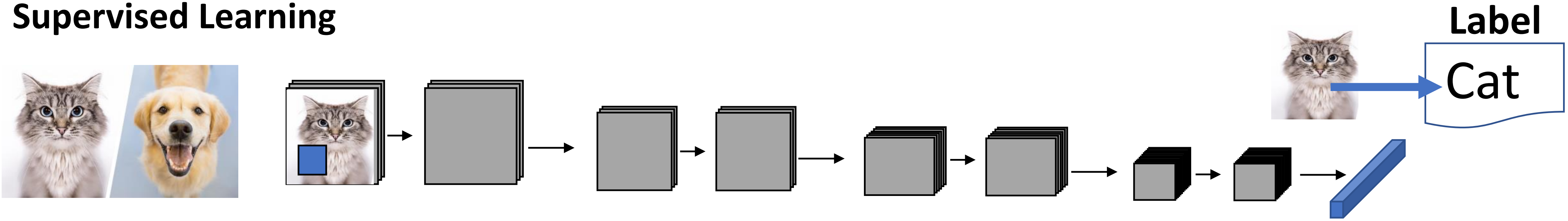


Contact: **tkurihana@uchicago.edu**

Supplement slides

Convolutional residual autoencoder

Supervised Learning



Unsupervised Learning