Contact: tkurihana@uchicago.edu

A35C-1635

Cloud Clustering Applied to MODIS Calibrated Radiances Over 2003 via Rotationally Invariant Autoencoder

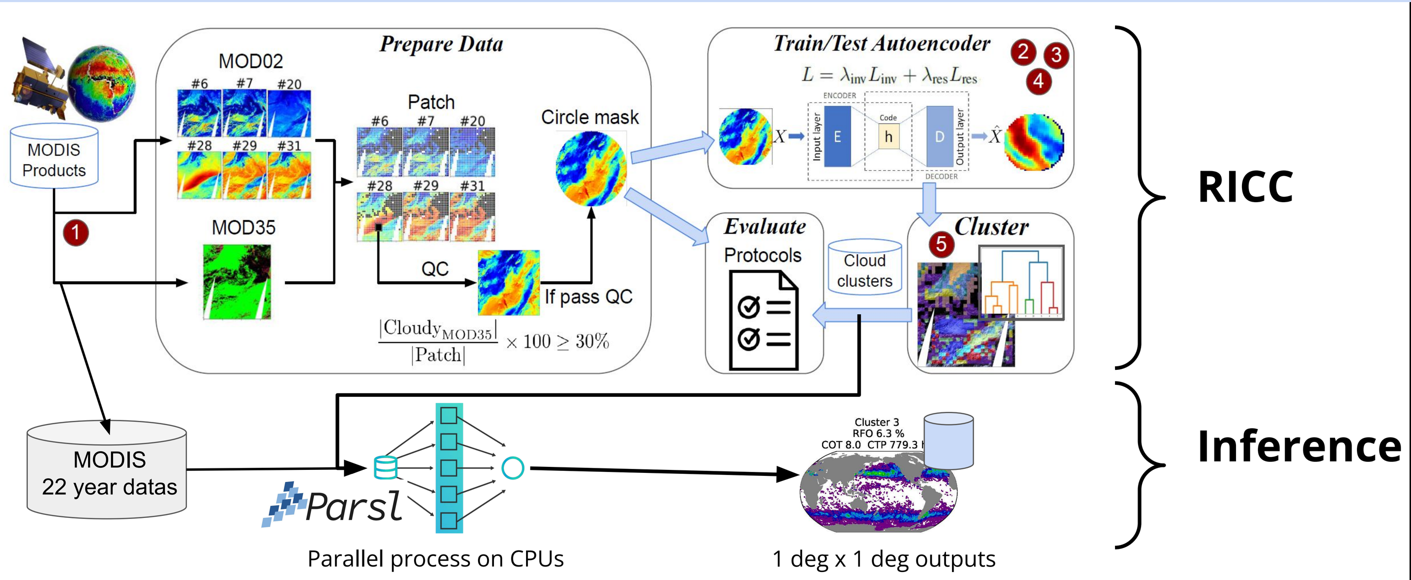
Takuya Kurihana[1,4], Elisabeth Moyer[3,4], Rebecca Willett[1,2], Davis Gilton[5] and Ian Foster[1,6]

1.Department of Computer Science, University of Chicago; 2. Department of Statistics, University of Chicago; 3. Department of Geoscience, University of Chicago; 4. Center for Robust Decision Making on Climate and Energy Policy (RDCEP); 5. Department of Electrical Engineering and Computer Science, University of Wisconsin Madison; 6. Argonne National Laboratory

Rotation-Invariant Cloud Clustering

- Automation of classification and clustering of cloud (*Cloud Classification*) gains an insight of cloud impact on the Earth's energy balance
- Unsupervised learning approach can address
  - Limited number of categories in training sets
  - Free from artificial definitions
  - Rarely available use of a large number of perfectly labeled datasets
- Kurihana et al. (2021)**[1]** developed Rotation-Invariant Cloud Clustering (RICC) for MODIS satellite imagery
- NASA archives 21-year of MODIS imagery from TERRA and AQUA = approximately ~700TB of data!
- Apply RICC to ocean image of a year 2003 and run inference over the entire January 2003 data.

Workflow



RICC workflow comprises a data download from NASA LAADS and four principle steps for the implementation. **Prepare Data** step subdivides each large MODIS swath into smaller unit (*patches*); **Train/Test Autoencoder** step produces a lower-dimensional representation; **Cluster** step applies agglomerative hierarchical clustering; **Evaluate** step applies protocol to evaluate the clusters.
Inference workflow applies cluster centroids to assign cluster labels to the rest of 22 year sof MODIS dataset.

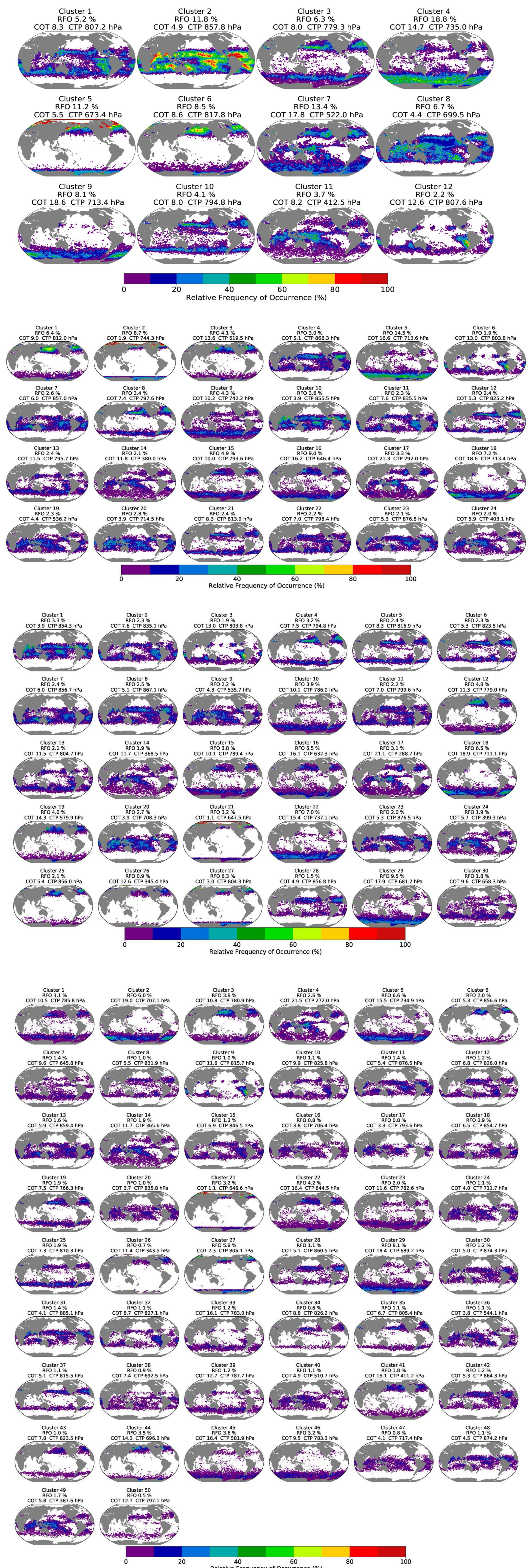
Use of MODIS Satellite Data

Product	Description	Band
MOD02	S/w infrared (1.230-1.250 μm)	5
	S/w infrared (1.628-1.652 μm)	6
	S/w infrared (2.105–2.155 μm)	7
	L/w thermal infrared (3.660–3.840 μm)	20
	L/w thermal infrared (7.175–7.475 μm)	28
	L/w thermal infrared (8.400–8.700 μm)	29
	L/w thermal infrared (10.780–11.280 μm)	31
MOD35	Cloud mask	
MOD06	Cloud optical thickness	
	Cloud top pressure	
	Cloud phase infrared	
	Cloud effective radius	

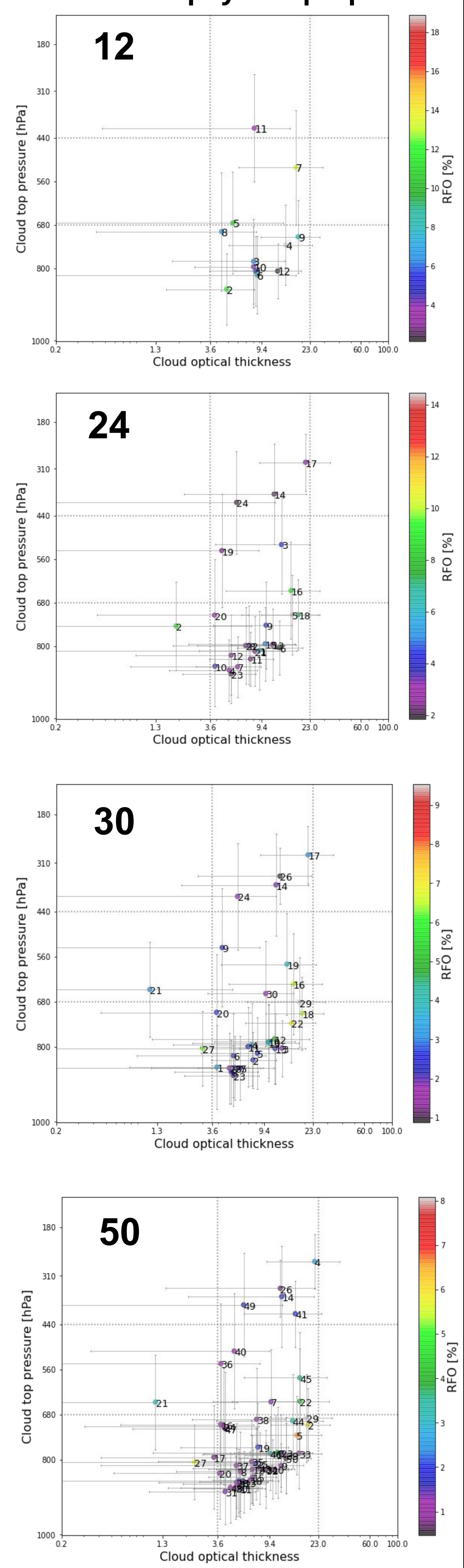
MOD02: Use for train and test data; **MOD35**: Detect cloud pixels; **MOD06**: Evaluate physical association

Scalable RICC Autoencoder

Relative frequency of occurrence



Mean & Std physical properties



Total cloud fraction for Jan. 2003

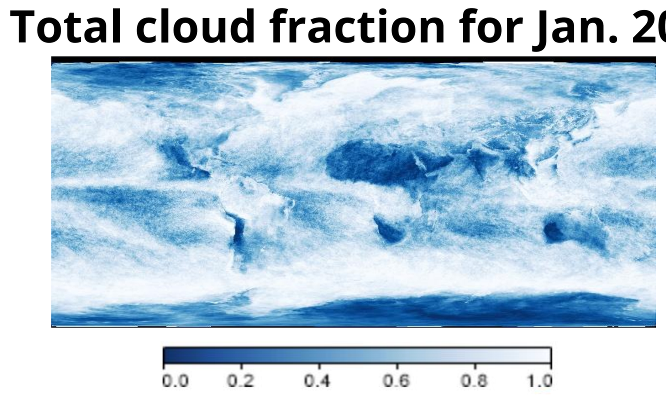
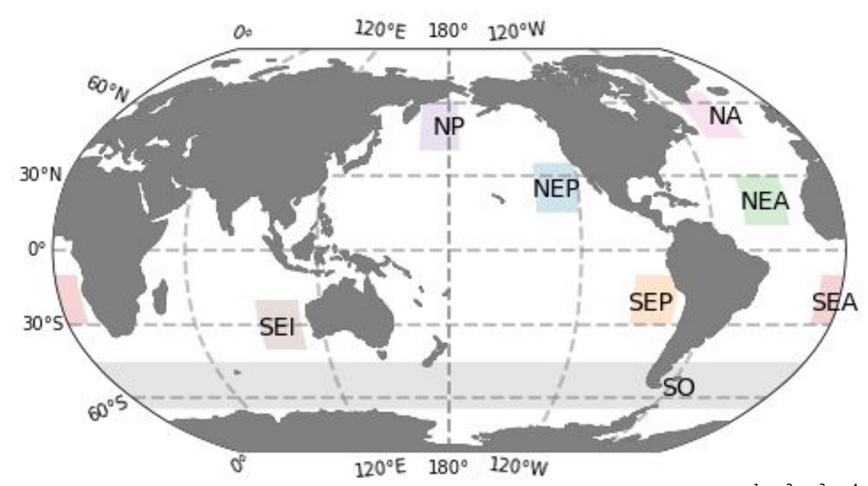
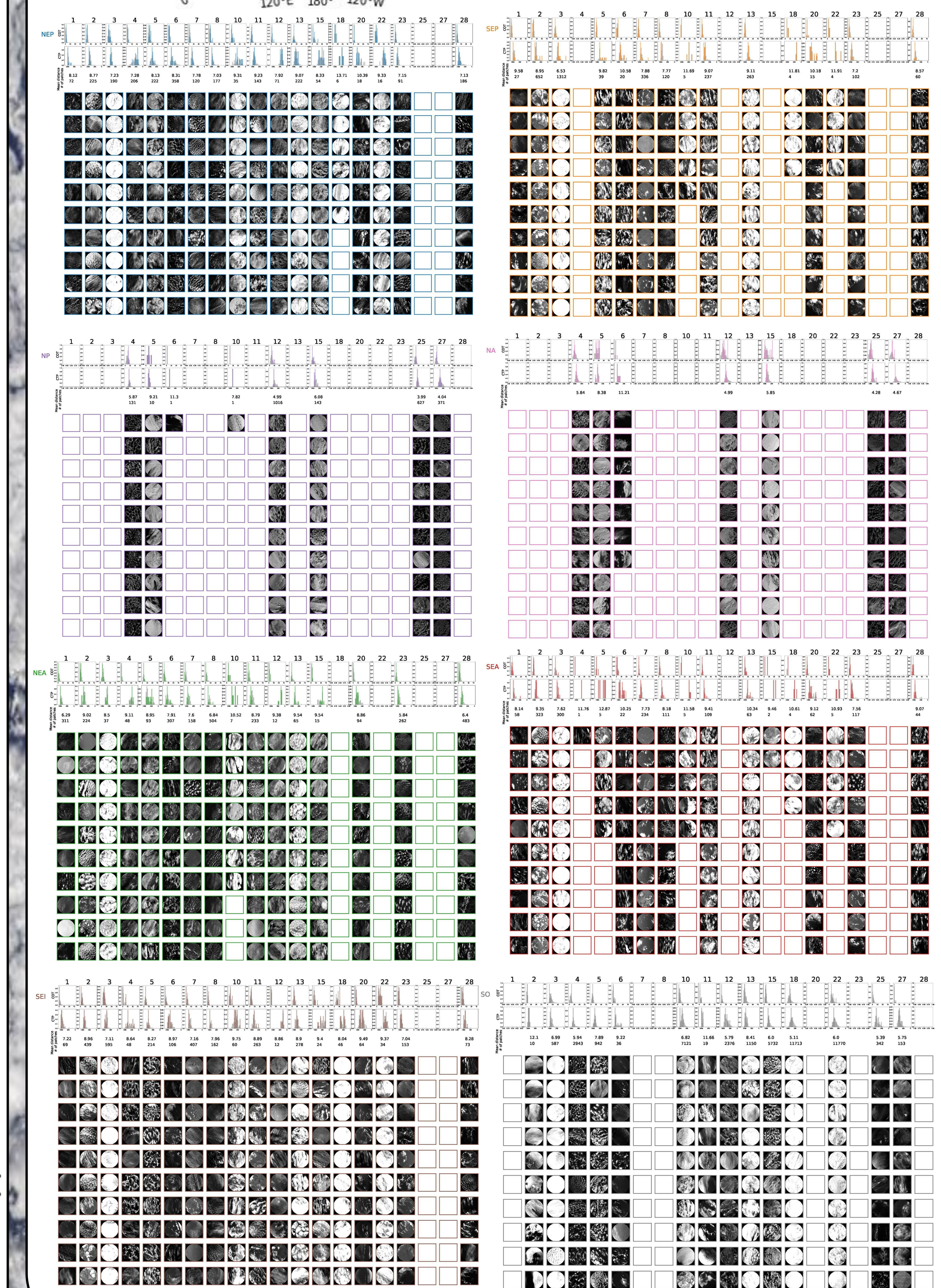


Figure acquired from NASA Earth Observation

Regional low cloud images



Display 10 cloud images closest to their centroid in cluster on the subset of ocean regions.



Results & Discussions

- Cloud clusters generated from RICC separate satellite imagery into diverse clusters in terms of their physical properties and patterns, indicating that our algorithm identifies the rich information of cloud patterns and textures in satellite observation.
- Example cloud images show coherent texture of clouds within the cluster.
- Next attempt is to determine the optimal number of clusters.

[1] RICC paper is available in IEEE TGRS

T. Kurihana, E. Moyer, R. Willett, D. Gilton and I. Foster, "Data-Driven Cloud Clustering via a Rotationally Invariant Autoencoder," in IEEE Transactions on Geoscience and Remote Sensing, doi: 10.1109/TGRS.2021.3098008.