

Comparison between clustering algorithms for the detection of cloud regimes

An objective way to classify clouds into „cloud regimes“ based on analysis of satellite cloud property datasets is by automatic clustering of joint optical depth – cloud top pressure histograms (e.g. Jakob and Tselioudis, 2003). Using as reference available results obtained by applying the „k-means clustering“ for the cloud classification into regimes, we suggest here to repeat the study applying „hierarchical clustering“, a more advanced machine learning method.

As the two clustering algorithms operate in different ways, a comparison between them would help to understand the uncertainty in the definition of cloud regimes, since the k-means clustering represents a quite common choice in the community.

In this thesis, the student will:

- work with existing python scripts for the clustering and adjust them for hierarchical clustering
- apply the hierarchical clustering to one or several cloud datasets (CLARA-A2, Cloud_cci)
- create global maps and histograms as output of the clustering
- interpret the results.

Supervisors: Prof. Dr. Corinna Hoose, M. Sc. Olimpia Bruno

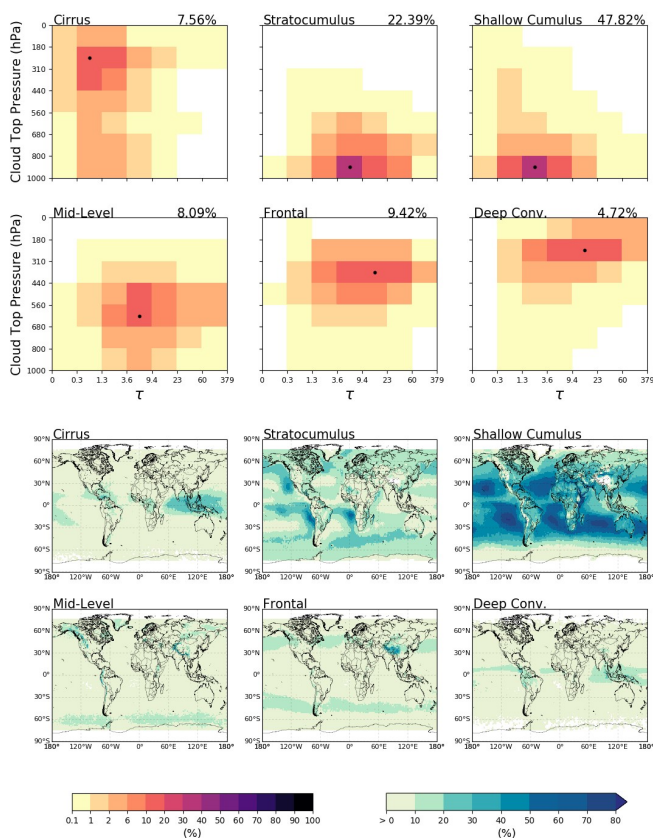


Figure: Cloud regimes represented by cluster centroids (upper panels) and their geographical distribution (lower panels) obtained by k-means clustering of cloud properties from the Cloud_cci v2.0 dataset.

Reference

Jakob, C., and G. Tselioudis, 2003: Objective identification of cloud regimes in the Tropical Western Pacific. *Geophys. Res. Lett.*, **30**, no. 21, 2082, doi:10.1029/2003GL018367.