Machine learning identifies characteristic “local” droplet size distributions in stratocumulus clouds

Nithin Allwayin[1] , Michael Larsen[2] ,Alexander Shaw[3] , Susanne Glienke[4] ,Shreya Joshi[1], Claudio Mazzoleni[1], and Raymond A. Shaw[1]

[1]Michigan Technological University

[2] College of Charleston

[3] Brigham Young University

[4] Pacific Northwest National Laboratory

2022 JOINT ARM USER FACILITY AND ASR PI MEETING

Background and Motivation

- Cloud droplet size distributions are often represented in global climate models (GCMs) and remote sensing algorithms as modified gamma distributions that capture the average behavior across a large spatial domain.
- Cloud processes, such as growth by condensation or collisions, on the other hand, span across a range of spatial scales down to the droplet scale.
- The effects of these “local’’ scale processes are not well represented in GCMs and constitute one of the largest uncertainties in model predictions.
- We ask the question whether the cloud drops at the smallest scales are indeed gamma like. In other words, if we sample a small region in the cloud do the size distributions resemble the grid averaged distribution values?
- Further, is there any similarity between the droplet populations at different points in a cloud?

Methods

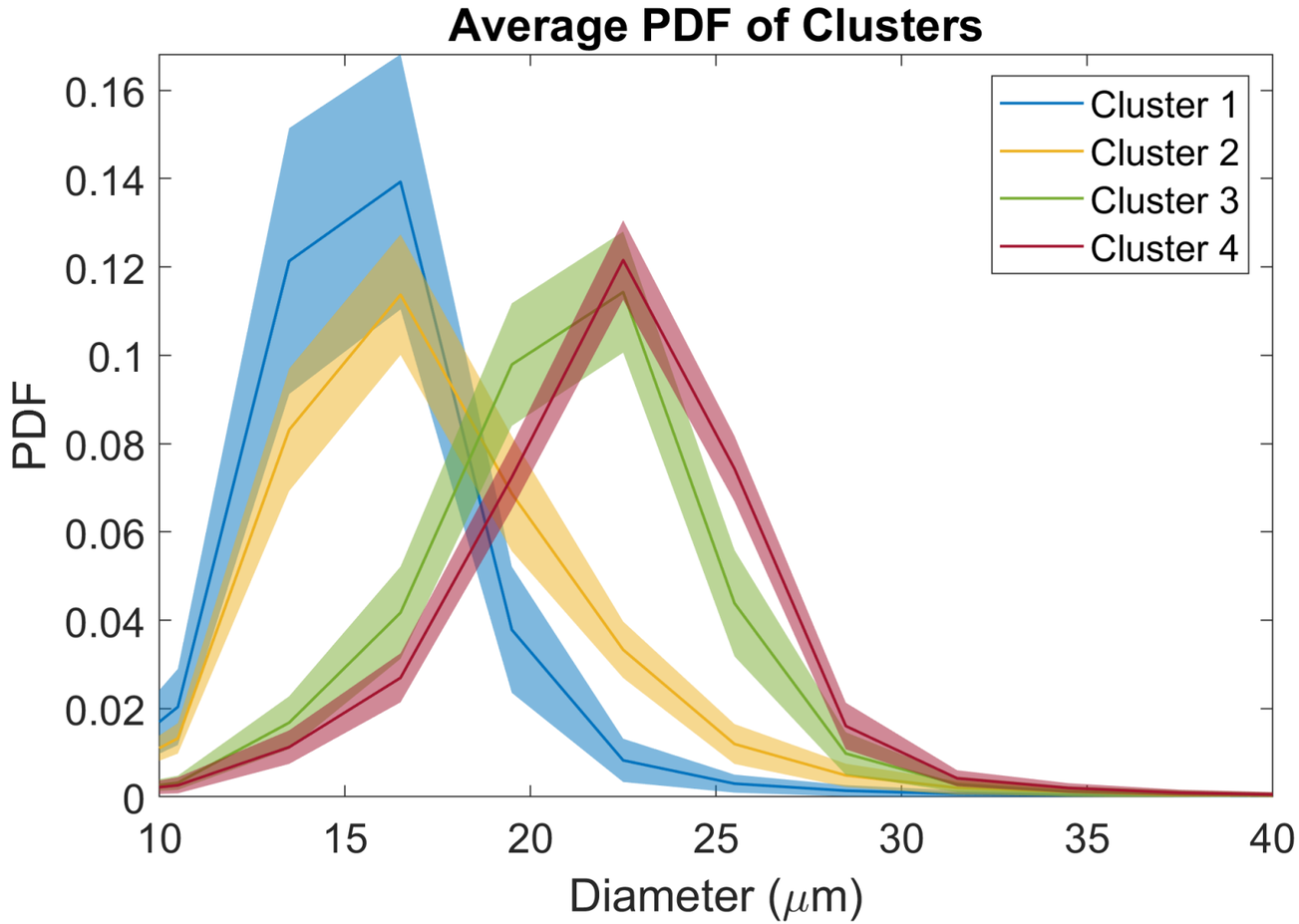
- We used hypothesis testing to evaluate this possibility to see if “local” samples of cloud have statistically similar size distributions.
- A semi-parametric algorithm was developed to search for and identify distinct sets of cloud droplet size distributions. Many cloud samples taken at different points in the cloud fall into one of these distribution types.
- The algorithm works by combining hypothesis testing and data clustering techniques. Kolmogorov–Smirnov test identifies the statistically similar size distributions which are then grouped by data clustering using a class of density based clustering algorithms(e.g. DBSCAN/OPTICS).
- The algorithm does not presume any specific distribution shape, avoids biases from binning and the results are determined in an relatively unsupervised fashion.
- In the end we obtain different sets of cloud segments having almost the same distribution of droplet populations. These form the characteristic distributions.

ASR
Atmospheric System Research

1
2
3
4
5
6

Results and Discussion

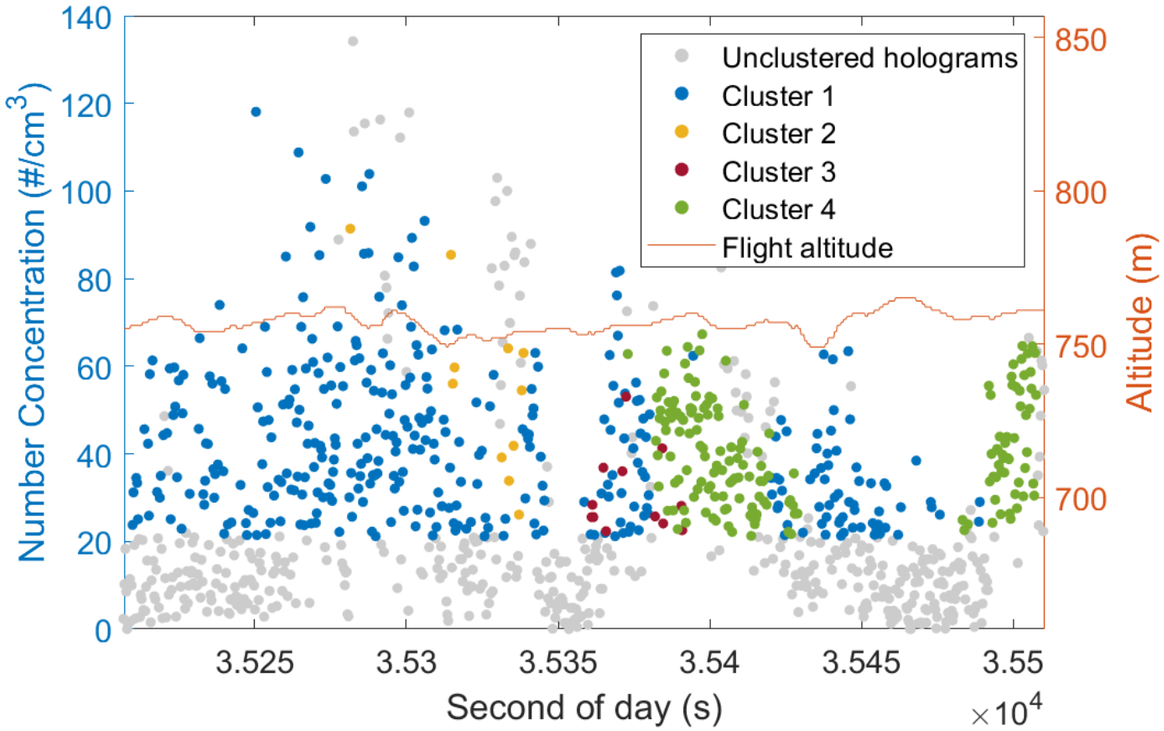
- Results from the holographic cloud probe (HOLODEC) transects from the Aerosol and Cloud Experiments in the Eastern North Atlantic (ACE-ENA) research flight on July 18, 2018 identifies these characteristic distributions



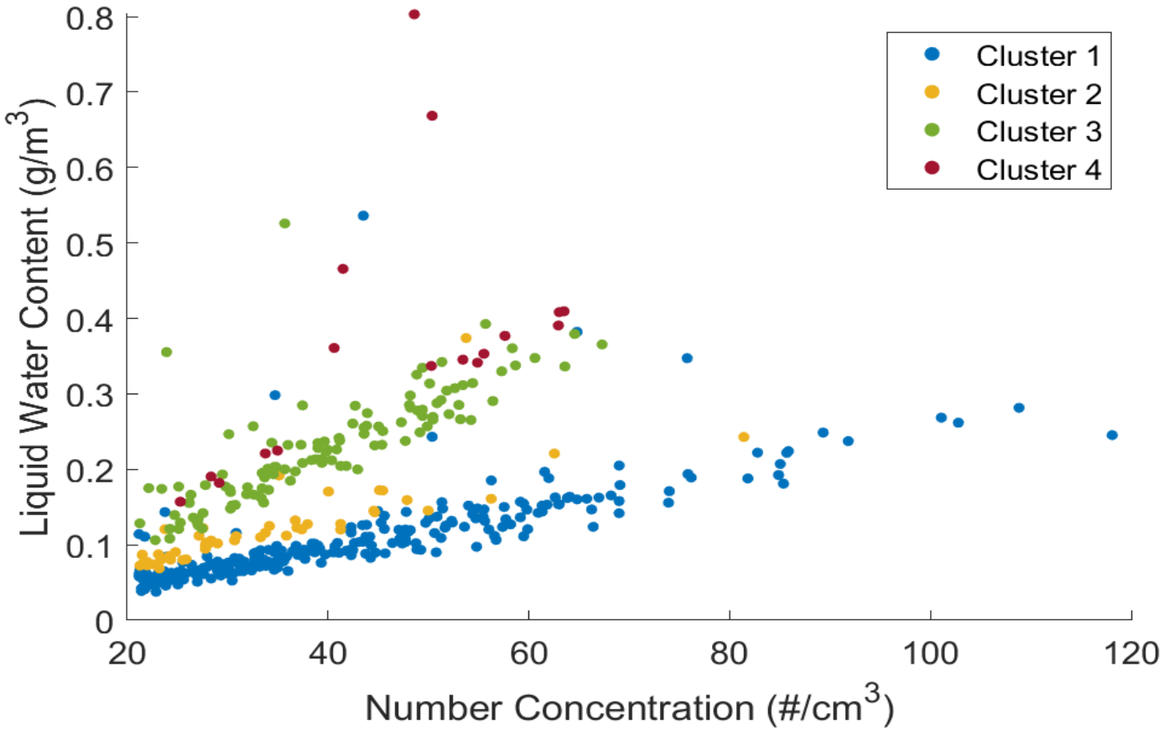
Distributions of the identified “characteristic” clusters. For this cloud segment, 4 characteristic distributions are identified. The shaded region gives one standard deviation.

ASR Results and Discussion

Atmospheric System Research



Clusters are ubiquitous and spatially persistent. Cloud regions with same distributions shapes are found kilometers apart in the same cloud. Here each point corresponds to a “local” sample volume with defined cloud size distribution. The time series is analogous to spatial position in the cloud



In this mixing-diagram-like plot, clusters tend to form lines. These lines represent similar size distribution shapes with varying levels of dilution. This may indicate signatures of individual inhomogeneous mixing events

Summary, References, and Acknowledgements

- We developed an algorithm that identifies a distinct set of cloud droplet size distributions that occur often within samples taken in a cloud and thus can represent the sub-grid scale variabilities in cloud representations.
- The algorithm is demonstrated via measurements from holographic cloud probe (HOLODEC) transects from the Aerosol and Cloud Experiments in the Eastern North Atlantic (ACE-ENA) field campaign and identifies characteristic cloud distributions for homogenous looking stratocumulus clouds.
- These clusters are ubiquitous, persistent and not necessarily spatially correlated.
- The nature of the identified clusters suggest inhomogeneous mixing

Allwayin, N., Larsen, M. L., Shaw, A. G., & Shaw, R. A. (2022). Automated Identification of Characteristic Droplet Size Distributions in Stratocumulus Clouds Utilizing a Data Clustering Algorithm, Artificial Intelligence for the Earth Systems, 1(3), e220003.