EGU22-7355
https://doi.org/10.5194/egusphere-egu22-7355
EGU General Assembly 2022

Combining cloud properties and synoptic observations to predict cloud base height using Machine Learning

Julien Lenhardt[1], Johannes Quaas[1], and Dino Sejdinovic[2]
[1]Leipziger Institut für Meteorologie, Universität Leipzig, Leipzig, Germany (julien.lenhardt@uni-leipzig.de)
[2]Department of Statistics, University of Oxford, Oxford, UK

Cloud base height (CBH) is an important geometric parameter of a cloud and shapes its radiative properties. The CBH is also further of practical interest in the aviation community regarding pilot visibility and aircraft icing hazards. While the cloud-top height has been successfully derived from passive imaging radiometers on satellites during recent years, the derivation of the CBH remains a more difficult challenge with these same retrievals.

In our study we combine surface observations and passive satellite remote-sensing retrievals to create a database of CBH labels and cloud properties to ultimately train a machine learning model predicting CBH. The labels come from the global marine meteorological observations dataset (UK Met Office, 2006) which consists of near-global synoptic observations made on sea. This data set provides information about CBH, cloud type, cloud cover and other meteorological surface quantities with CBH being the main interest here. The features based upon which the machine learning model is trained consist in different cloud-top and cloud optical properties (Level 2 products MOD06/MYD06 from the MODIS sensor) extracted on a 127km x 127km grid around the synoptic observation point. To study the large diversity in cloud scenes, an auto-encoder architecture is chosen. The regression task is then carried out in the modelled latent space which is output by the encoder part of the model. To account for the spatial relationships in our input data the model architecture is based on Convolutional Neural Networks. We define a study domain in the Atlantic ocean, around the equator. The combination of information from below and over the cloud could allow us to build a robust model to predict CBH and then extend predictions to regions where surface measurements are not available.