EGU23-15817, updated on 23 Apr 2024
https://doi.org/10.5194/egusphere-egu23-15817
EGU General Assembly 2023

Evaluating the generalization ability of a deep learning model trained to detect cloud-to-ground lightning on raw ERA5 data

Gregor Ehrensperger[1], Tobias Hell[1], Georg Johann Mayr[2], and Thorsten Simon[1,3]
[1]Department of Mathematics, University of Innsbruck, Innsbruck, Austria
[2]Department of Atmospheric and Cryospheric Sciences, University of Innsbruck, Austria
[3]Department of Statistics, University of Innsbruck, Innsbruck, Austria

Atmospheric conditions that are typical for lightning are commonly represented by proxies such as cloud top height, cloud ice flux, CAPE times precipitation, or the lightning potential index. While these proxies generally deliver reasonable results, they often need to be adapted for local conditions in order to perform well. This suggests that there is a need for more complex and holistic proxies. Recent research confirms that the use of machine learning (ML) approaches for describing lightning is promising.

In a previous study a deep learning model was trained on single spatiotemporal (30km x 30km x 1h) cells in the summer period of the years 2010--2018 and showed good results for the unseen test year 2019 within Austria. We now improve this model by using multiple neighboring vertical atmospheric columns to also address for horizontal moisture advection. Furthermore data of successive hours is used as input data to enable the model to capture the temporal development of atmospheric conditions such as the build-up and breakdown of convections.

In this work we focus on the summer months June to August and use data from parts of Central Europe. This spatial domain is thought to be representative for Continental Europe since it covers mountainous aswell as coastal regions. We take raw ERA5 parameters beyond the tropopause enriched with a small amount of meta data such as the day of the year and the hour of the day for training. The quality of the resulting paramaterized model is then evaluated on Continental Europe to examine the generalization ability.

Using parts of Central Europe to train the model, we evaluate its ability to generalize on unseen parts of Continental Europe using EUCLID data. Having a model that generalizes well is a building block for a retrospective analysis back into years where the structured recording of accurate lightning observations in a unified way was not established yet.