Machine learning methods for the classification of extreme rainfall and hail events

Reinhard Teschl, Barbara Süsser-Rechberger, and Helmut Paulitsch

Graz University of Technology, Institute of Microwave and Photonic Engineering, Graz, Austria (reinhard.teschl@tugraz.at)

In this study, an analysis of a meteorological data set with machine learning tools is presented. The aim was to identify characteristic patterns in different sources of remote sensing data that are associated with hazards like extreme rainfall and hail.

The data set originates from a project that was started in 2007 with the goal to document and mitigate hail events in the province of Styria, Austria. It consists of three dimensional weather radar data from a C-band Doppler radar, cloud top temperature information from infrared channels of a weather satellite, as well as the height of the $0°$ C isotherm from the forecast of the national weather service. The 3D radar dataset has a spatial resolution of 1 km x 1 km x 1 km, up to a height of 16 km above mean sea level, and a temporal resolution of 5 minutes. The infrared satellite image resolution is about 3 km x 3 km, the images are updated every 30 minutes. The study area has approx. 16,000 square kilometers.

So far, different criteria for the occurrence of hail (and its discrimination from heavy rain) have been found and are documented in the literature. When applying these criteria to our data and contrasting them with damage reports from an insurance company, a need for adaption was identified. Here we are using supervised learning paradigms to find tailored relationships for the study area, validated by a sub-dataset that was not involved in the training process.