4–8 November 2019, Kraków, Poland
ECSS2019-165

A novel multi-sensor nowcast algorithm of thunderstorm intensity using machine learning

Ulrich Hamann (1), Joel Zeder (1,2), Hélène Barras (1,3,4), Lorenzo Clementi (1), Loris Foresti (1), Alessandro M. Hering (1), Shruti Nath (1), Daniele Nerini (1,2), Luca Nisi (1), Marco Sassi (1), and Urs Germann ()

(1) MeteoSwiss, Locarno, Switzerland (Ulrich.Hamann@meteoswiss.ch), (2) Institute for Atmospheric and Climate Science, ETH, Zurich, Switzerland, (3) Institute of Geography and Oeschger Centre for Climate Change Research, University of Bern, Bern, Switzerland, (4) Mobiliar Lab for Natural Risks, University of Bern, Bern, Switzerland

Today's warning procedures for severe convective storms are often based on nowcasting methods in which, typically, the thunderstorm position is extrapolated with its current motion, while the storm severity is kept constant (Lagrangian persistence assumption). In this presentation, we examine the potential of machine learning methods for nowcasting the evolution of thunderstorms (Zeder et al, 2018), e.g. expressed as heuristic severity rank (as defined by TRT algorithm of Hering et al, 2008) or as lightning and hail activity. We train the machine learning algorithm with a comprehensive database of thunderstorm properties in Switzerland and its surroundings for the summer 2018, consisting of satellite, radar and lightning observations, NWP model data, topography, time, and thunderstorm position and movement, in total more than 70 variables. We only consider thunderstorm severities relevant for operational warnings, resulting in around 10'000 snapshots of single cells in total. For each snapshot, the cell's history of the last 45 min in 5 min time steps was evaluated by deriving the thunderstorm displacement using the open-source library pysteps (https://pysteps.github.io/, Pulkkinen et al, 2018). At each location all predictors were analysed in a local surrounding to avoid sensitivities to particular cell boundary definitions as well as cell splitting and merging. Within this surrounding, several statistics were computed for the predictors, which include the minimum, maximum, standard deviation, and several percentiles. All possible combinations of variables, time steps, and statistics provided a very large number of possible predictors. Hence, we applied shallow decision trees (XGBoost, Chen and Guestrin, 2016) as a feature selection algorithm to estimate the importance of the predictors. For instance, when predicting the TRT rank, radar observations were ranked as most important input data for lead times of 5 min and 10 min, followed by information from the NWP model, satellite and lightning observations. For nowcasts with a longer lead time, the NWP model is ranked as the most important data source, followed by radar as second. Satellite and lightning observations remain at the third and fourth position for our predictor selection. In its current version for nowcasting of near-future thunderstorm intensity changes, shallow trees were found to perform best, using a total of 750 input features. In the end, we validate the algorithm and compare the quality of the thunderstorm severity nowcast to the assumption of Lagrangian persistence.