EGU22-4479
https://doi.org/10.5194/egusphere-egu22-4479
EGU General Assembly 2022

Extratropical high-wind feature identification using a probabilistic random forest

Lea Eisenstein[1], Benedikt Schulz[2], Peter Knippertz[1], and Joaquim G. Pinto[1]
[1]Karlsruhe Institute of Technology, Institute of Meteorology and Climate Research, Department Troposphere Research, Karlsruhe, Germany (lea.eisenstein@kit.edu)
[2]Karlsruhe Institute of Technology, Institute for Stochastic, Karlsruhe, Germany

Strong winds associated with extratropical cyclones are one of the most dangerous natural hazards in Europe. These high winds are mostly connected with four mesoscale dynamical features: the warm (conveyor belt) jet (WJ), the cold (conveyor belt) jet (CJ), (post) cold-frontal convective features (CFC) and the sting jet (SJ). While all four have high wind gust speeds in common, the timing, location and some further characteristics typically differ and hence likely also the forecast errors occurring in association with them.

Here we present an objective identification approach for the four features listed above, based on a probabilistic random forest using each feature's most important characteristics in wind, rainfall, pressure and temperature evolution. The main motivations for this are to generate a climatology for Central Europe, to analyse forecast errors specific to individual features, and to ultimately improve forecasts of high wind events through feature-dependent statistical post-processing. To achieve this, we strive to identify the features in irregularly spaced surface observations and in gridded analyses and forecasts in a consistent way.

To train the probabilistic random forest, we subjectively identify the four storm features – as well as high cold sector winds – in ten winterstorm cases between 2017 and 2020 in both hourly surface observations and high-resolution reanalyses of the German COSMO model over Europe, using an interactive data analysis and visualisation tool. Results show that mean sea-level pressure (tendency), potential temperature, precipitation amount and wind direction are most important for the distinction between the features. From the random forest we get probabilities of each feature occurring at the single stations, which can be interpolated into areal information using kriging. While the observational data are limited to surface measurements, the gridded data includes further useful parameters and the possibility to consider vertical structures.

The results show a good identification of CJ, CFC and WJ, while a distinction between SJ and CJ is difficult using surface observations alone, such that the two features are considered together at this stage. A climatology is currently being compiled for both surface observations and the reanalyses over a period of around 20 years using the respective trained probabilistic random forests and further for high-resolution COSMO ensemble forecasts, for which we want to analyse forecast errors and develop feature-dependent postprocessing procedures.