Geophysical Research Abstracts
Vol. 21, EGU2019-11238, 2019
EGU General Assembly 2019

Classification of Atmospheric Meteorological Events by a supervised machine learning

Anton Sokolov (1), Egor Dmitriev (2), Hervé Delbarre (1), and Cyril Gengembre (1)

(1) University of Littoral Cote d'Opale, Laboratory for Physico-Chemistry of the Atmosphere, Dunkerque, France (anton.sokolov@univ-littoral.fr), (2) Institute of Numerical Mathematics, Russian Academy of Sciences, 8 ul. Gubkina, 119333, Moscow, Russia

The dispersion of pollutants and the weather forecast are closely connected with atmospheric meteorological phenomena like breezes, fogs and storms [1]. The detection and classification of these events are currently realized by qualified specialists in atmospheric physics based on analysis of in-situ and remote sensing data.

We propose to classify these meteorological phenomena by means of the supervised machine learning (SML). Four algorithms were applied for meteorological events detection: the k-nearest neighbors, the multiclass support vector machine with the error correcting output codes, the quadratic discriminant analysis and the Parzen-Rosenblatt window-based classifier.

Dunkerque region, in the North of France, was chosen as the area of study. Ultrasound anemometer coupled with a meteostation and a Doppler LIDAR datasets were obtained and analysed during a one-year long measurement campaign. These measurements were investigated and classified by a meteorologist. Four classes of meteorological events were studied: sea breezes, fogs, storms and other events. These data served to train algorithms. We optimized the sequence of predictors with the regularized greedy algorithm of stepwise forward selection to overcome the curse of dimensionality problem.

Four classifiers were compared; it was shown that, for in-situ anemometer data, these SML algorithms allow the correct detection of more than 80% of each type of events.

On the other hand, we failed to get a considerable improvement by taking into account the wind profile obtained by the LIDAR. Informative predictors differ for each algorithm, but the absolute wind speed, the North – South wind component and the solar radiation are the most often selected.

In perspective, we are planning to apply the developed SML algorithms to atmospheric modelling data.

1. Sokolov et al. Cluster analysis of atmospheric dynamics and pollution transport in a coastal area. Boundary-Layer Meteorology, 2016. DOI: 10.1007/s10546-016-0174-5