EMS Annual Meeting Abstracts
Vol. 16, EMS2019-172, 2019


# High-frequency ensemble gust predictions for surface sites and wind turbines using machine learning and data mining techniques

Irene Schicker, Petrina Papazek, Markus Dabernig, and Alexander Kann
ZAMG, DMM-VHMOD, Vienna, Austria (irene.schicker@zamg.ac.at)

Wind gusts and high wind speeds can be a thread to not only wildlife and nature but also to infrastructure, dispersion of air pollution and aviation. Furthermore, they need to be considered in wind energy and power grid management for construction, material, siting and maintenance of turbines and power lines.
Gusts and their measurements are an important information source on turbulence conditions at specific sites. Two main mechanisms can be described: wind shear (dominates in near-neutral conditions) and buoyancy (both in unstable conditions, under stable conditions buoyancy damps wind shear). Additionally, gusts can occur as convective gusts (downbursts), gravity waves in vicinity of higher mountains (e.g. Bora) and breaking gravity waves.
Accurate numerical weather prediction (NWP) forecast depend strongly on model skills in reproducing turbulence and the sub-grid scale processes. Except for rapid update cycle NWP models they often do not provide sub-hourly forecasts and are not available in time for the nowcasting (0 – max. 6 hours ahead). Methodologies based solely on observations can help here but bear the risk of not being useful beyond 2-hours ahead.
In a first step a high-frequency wind speed and gust nowcasting ensemble based on different machine learning methodologies, including an ensemble for every method, will be presented. It is based on boosting, random forest, linear regression and a feed forward neural network. As second step the feed forward network will be combined with the ZAMG interval Artifical Neural Network (ZiANN) for medium range forecasting.
To be able to provide a reasonable set of input features a data mining based input feature selection will be carried out. Therefore, the possibilities of crowd sourced and clustered data will be investigated, too, to increase our set of available training data.