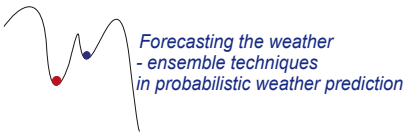
EMS Annual Meeting Abstracts
Vol. 8, EMS2011-183, 2011
11th EMS / 10th ECAM


# Calibrating the ECMWF EPS 2m Temperature and 10m Wind Speed for Austrian Stations

S. Radanovics and A. Kann

Central Institute for Meteorology and Geodynamics, Model Applications, Vienna, Austria (sabine.radanovics@zamg.ac.at)

In operational forecasting the ensemble predictions are used mainly to assess the model uncertainty in a qualitative way. Quantitative applications are limited, because they are uncalibrated. The ensemble mean, as well as the distribution, have systematic errors that can be adjusted by appropriate statistical calibration, also known as ensemble model output statistic. The calibrated point forecasts can be used for quantitative probability forecasts.
About 290 Austrian station locations are chosen for calibration. ECMWF EPS forecasts are interpolated to these station locations and an ensemble model output statistic for each station and each available leadtime for the parameters 2m temperature and 10m wind speed is performed.
For 2m temperature the non-homogeneous Gaussian regression (NGR), minimising the CRPS, is used for calibration. Experiments with different lengths of training data are performed. Previous findings suggested that between 30 and 50 days of training data are sufficient for calibrating the 2m temperature, which is confirmed by the present study. No significant improvement is obtained by using more training data and the verification results do not depend much on the season. In order to have still enough training data for higher leadtimes (up to 15 days), the last 50 days are used in the operational application.
For 10m wind speed a comparison between two calibration methods, cut-off NGR and logistic regression, is performed. 120 days of training data are used for the logistic regression, where the ensemble mean is taken as a predictor. Additional predictors like ensemble spread or wind direction are tested, but did not improve the results. More training data did not improve the results either. For cut-off NGR, 120 days of training data were taken into account, but only the 30 most similar days (with the smallest absolute difference in the ensemble mean) were actually used. This selection gives better results than taking the whole 120 days training data set or just taking the last 30 days. Selecting training days with similar wind direction does not improve the results. With both methods a reduction of the Brier score up to 50% can be seen from the verification results. For low thresholds like 1 m/s the logistic regression performs slightly better, for higher thresholds like 5 m/s the performance of the two methods is nearly equal. Since the logistic regression performs better, this method is used in the operational application.