Ensemble verification: Old scores, new perspectives

Sabrina Wahl (1,2), Jan D. Keller (2,3), and Petra Friederichs (1)

(1) University of Bonn, Meteorological Institute, Bonn, Germany (wahl@uni-bonn.de), (2) Hans Ertel Center for Weather Research, Climate Monitoring Branch, (3) Deutscher Wetterdienst, Offenbach, Germany

The recently derived decomposition of the quantile score into uncertainty, resolution and reliability (Bentzien and Friederichs, 2014) can be generalized to a decomposition of the continuous ranked probability score (CRPS). This decomposition reveals detailed insights into ensemble performance in terms of calibration and information content. We utilize this framework to investigate the performance of raw ensemble forecasts, i.e. without further translation of the ensemble members into probabilistic predictions. Such an evaluation is of particular interest for developers and designers of ensemble prediction systems. We further present a novel graphical exploration of the forecast attributes.