EGU23-2701, updated on 18 Apr 2024
https://doi.org/10.5194/egusphere-egu23-2701
EGU General Assembly 2023
© Author(s) 2024. This work is distributed under
the Creative Commons Attribution 4.0 License.

Investigating properties of statistical tests for comparing predictive performance with application to probabilistic weather forecasting

Annette Möller[1] and Friederike Grupe[2]
[1]Bielefeld University, Faculty of Business Administration and Economics, Bielefeld, Germany (annette.moeller@uni-bielefeld.de)
[2]KPMG Hamburg, Germany

This work investigates several statistical tests in the context of probabilistic weather forecasting and ensemble postprocessing. The tests are commonly used for comparing predictive performance of e.g. two statistical postprocessing models.

In the first part of the analysis a case study is conducted on temperature data consisting of observations and ensemble forecasts. The tests are applied to compare the performance of two probabilistic temperature forecasts at different stations, for different lead times, investigating several standard verification metrics to measure prediction performance. The analysis shows that the tests generally behave consistently in the context of temperature forecasts. However, for certain scenarios some tests might be be preferred over the others. In general, the combination of the original Diebold-Mariano test with the continuous ranked probability score (CRPS) to assess forecast accuracy leads to the most consistent and reliable results.

The second part of the analysis uses simulated data to investigate the general behaviour of the tests in different postprocessing scenarios as well as their size and power properties. Again, the original Diebold-Mariano test appears to perform most reliably and shows no noticeable inconsistent behaviour for different simulation settings.