Geophysical Research Abstracts
Vol. 12, EGU2010-12998, 2010
EGU General Assembly 2010
© Author(s) 2010

What we can learn from probabilistic verification scores in the context of hydrological ensemble forecasting

Maik Renner (1) and Micha Werner (2)

(1) TU Dresden, Institute of Hydrology und Meteorology, Meteorology, Tharandt, Germany (maik.renner@mailbox.tu-dresden.de), (2) Deltares and UNESCO-IHE, Delft, The Netherlands

The acknowledgement that hydro-meteorological forecasts may be uncertain, has resulted in an increasing use of probabilistic forecasting techniques in operational forecasting systems, often through the use of ensemble forecasting. Rather than providing a single value forecast, these ensemble forecasts provide a probability of a future state, e.g. precipitation, water level or discharge. Although these probabilistic forecasts provide valuable information on the range of possible future states, the reliability, sharpness and resolution of these probability forecasts should be evaluated through a process of forecast verification.

In this paper we present results of an extensive verification of forecasts made using two meteorological ensemble forecast products; the global scale ECMWF-EPS and the dynamically downscaled COSMO-LEPS. These are used to provide input forcing to the operational forecasting system of the Rhine basin, used by both the German Federal Institute of Hydrology, and the Dutch Centre for Water Management for predicting water levels and discharges at key forecasting locations. Ensemble flow forecasts are generated by forcing a calibrated hydrological model (HBV), using either of the two meteorological ensemble products. Verification of a large set of hindcast runs shows that when compared to climatology, positive skill scores are found at all river gauges considered for lead times of up to 9 days. This shows that the medium-range flow forecasts obtained to be useful. However, the comparison between the low resolution ECMWF and the high resolution COSMO-LEPS model shows that the downscaled forecasts provide better representation of the variability. Higher skills are found across all catchment sizes, particularly for shorter lead time forecasts. The downscaling of the ensemble forecasts to a scale commensurate with the sub-basin scale in the hydrological model is thus recommended.

Additionally we demonstrate the use of the probabilistic verification scores established for (i) the forecasting system developer, and (ii) the operational forecaster. On the one hand the developer aims to compare and improve forecasts, and thus make decisions on data to be used in the system. Suitable measures are threshold based binary verification methods such as Reliability diagrams and the Brier score and multicategorical methods such as Rank histogram and the Ranked Probability Score. For the forecaster, on the other hand, verification statistics that can help judge the forecast at hand and provide guidance on the decision to be taken. For this situation we discuss the use of the Reliability diagram and the Relative Operating Characteristic.