Geophysical Research Abstracts
Vol. 13, EGU2011-13640, 2011
EGU General Assembly 2011

A comparison between raw ensemble output, (modified) Bayesian model averaging and extended logistic regression using ECMWF ensemble precipitation reforecasts

Maurice Schmeits and Kees Kok
KNMI, Weather Research and Development, De Bilt, Netherlands (schmeits@knmi.nl)

Using a 20-year ECMWF ensemble reforecast data set of total precipitation and a 20-year data set of a dense precipitation observation network in the Netherlands, a comparison is made between the raw ensemble output, Bayesian model averaging (BMA) and extended logistic regression (LR). A previous study has indicated BMA and conventional LR to be successful in calibrating multi-model ensemble forecasts of precipitation for a single forecast projection. However, a more elaborate comparison between these methods has not yet been made. This study compares the raw ensemble output, BMA and extended LR for single-model ensemble reforecasts of precipitation, namely from the ECMWF ensemble prediction system (EPS). The raw EPS output turns out to be generally well calibrated up to 6 forecast days, if compared to the area-mean 24-h precipitation sum. Surprisingly, BMA is less skillful than the raw EPS output from forecast day 3 on. This is due to the bias correction in BMA, which applies model output statistics to individual ensemble members. As a result, the spread of the bias-corrected ensemble members is decreased, especially for the longer forecast projections. Here an additive bias correction is applied instead and the equation for the probability of precipitation in BMA is also changed. These modifications to BMA are referred to as "modified BMA" and lead to a significant improvement in the skill of BMA for the longer projections. If the area-maximum 24-h precipitation sum is used as a predictand, both modified BMA and extended LR improve the raw EPS output significantly for the first 5 forecast days. However, the difference in skill between modified BMA and extended LR does not seem to be statistically significant. Yet, extended LR might be preferred, because incorporating predictors that are different from the predictand is straightforward, in contrast to BMA.