Increasing the Skill of Probabilistic Forecasts: Understanding Performance Improvements from Model-Error Representations

Judith Berner, Kathryn Fossell, Soyoung Ha, Joshua Hacker and Chris Snyder
NCAR

Four model-error schemes for probabilistic forecasts over the contiguous United States with the WRF-ARW mesoscale ensemble system are evaluated in regard to performance. Including a model-error representation leads to significant increases in forecast skill near the surface as measured by the Brier score. Combining multiple model-error schemes results in the best-performing ensemble systems, indicating that current model error is still too complex to be represented by a single scheme alone.
To understand the reasons for the improved performance, it is examined whether model-error representations increase skill merely by increasing the reliability and reducing the bias—which could also be achieved by postprocessing - or if they have additional benefits.
Removing the bias results overall in the largest skill improvement. Forecasts with model-error schemes continue to have better skill than without, indicating that their benefit goes beyond bias reduction. Decomposing the Brier score into its components reveals that, in addition to the spread-sensitive reliability, the resolution component is significantly improved. This indicates that the benefits of including a model-error representation go beyond increasing reliability. This is further substantiated when all forecasts are calibrated to have similar spread. The calibrated ensembles with model-error schemes consistently outperform the calibrated control ensemble.
Including a model-error representation remains beneficial even if the ensemble systems are calibrated and/ or debiased. This suggests that the merits of model-error representations go beyond increasing spread and removing the mean error and can account for certain aspects of structural model uncertainty.