Geophysical Research Abstracts
Vol. 17, EGU2015-6517, 2015
EGU General Assembly 2015

Evaluation of ensemble forecast uncertainty using a new proper score: application to medium-range and seasonal forecasts

Hannah Christensen (1), Irene Moroz (2), and Tim Palmer (1)

(1) Atmospheric, Oceanic and Planetary Physics, University of Oxford, Oxford, United Kingdom (h.m.christensen@atm.ox.ac.uk), (2) Oxford Centre for Industrial and Applied Mathematics, University of Oxford, UK

Forecast verification is important across scientific disciplines as it provides a framework for evaluating the performance of a forecasting system. In the atmospheric sciences, probabilistic skill scores are often used for verification as they provide a way of unambiguously ranking the performance of different probabilistic forecasts. In order to be useful, a skill score must be proper — it must encourage honesty in the forecaster, and reward forecasts which are reliable and which have good resolution. A new score, the Error-spread Score (ES), is proposed which is particularly suitable for evaluation of ensemble forecasts. It is formulated with respect to the moments of the forecast. The ES is confirmed to be a proper score, and is therefore sensitive to both resolution and reliability. The ES is tested on forecasts made using the Lorenz '96 system, and found to be useful for summarising the skill of the forecasts. The European Centre for Medium-Range Weather Forecasts (ECMWF) ensemble prediction system (EPS) is evaluated using the ES. Its performance is compared to a perfect statistical probabilistic forecast — the ECMWF high resolution deterministic forecast dressed with the observed error distribution. This generates a forecast that is perfectly reliable if considered over all time, but which does not vary from day to day with the predictability of the atmospheric flow. The ES distinguishes between the dynamically reliable EPS forecasts and the statically reliable dressed deterministic forecasts. Other skill scores are tested and found to be comparatively insensitive to this desirable forecast quality. The ES is used to evaluate seasonal range ensemble forecasts made with the ECMWF System 4. The ensemble forecasts are found to be skilful when compared with climatological or persistence forecasts, though this skill is dependent on region and time of year.