Geophysical Research Abstracts
Vol. 14, EGU2012-9729, 2012
EGU General Assembly 2012

Do quantitative decadal forecasts from GCMs provide decision relevant skill?

E B Suckling and L A Smith
Centre for the Analysis of Time Series, London School of Economics, United Kingdom (e.suckling@lse.ac.uk)

It is widely held that only physics-based simulation models can capture the dynamics required to provide decision-relevant probabilistic climate predictions. This fact in itself provides no evidence that predictions from today's GCMs are fit for purpose. Empirical (data-based) models are employed to make probability forecasts on decadal timescales, where it is argued that these 'physics free' forecasts provide a quantitative 'zero skill' target for the evaluation of forecasts based on more complicated models. It is demonstrated that these zero skill models are competitive with GCMs on decadal scales for probability forecasts evaluated over the last 50 years. Complications of statistical interpretation due to the 'hindcast' nature of this experiment, and the likely relevance of arguments that the lack of hindcast skill is irrelevant as the signal will soon 'come out of the noise' are discussed. A lack of decision relevant quantiative skill does not bring the science-based insights of anthropogenic warming into doubt, but it does call for a clear quantification of limits, as a function of lead time, for spatial and temporal scales on which decisions based on such model output are expected to prove maladaptive. Failing to do so may risk the credibility of science in support of policy in the long term.

The performance amongst a collection of simulation models is evaluated, having transformed ensembles of point forecasts into probability distributions through the kernel dressing procedure [1], according to a selection of proper skill scores [2] and contrasted with purely data-based empirical models. Data-based models are unlikely to yield realistic forecasts for future climate change if the Earth system moves away from the conditions observed in the past, upon which the models are constructed; in this sense the empirical model defines zero skill. When should a decision relevant simulation model be expected to significantly outperform such empirical models?

Probability forecasts up to ten years ahead (decadal forecasts) are considered, both on global and regional spatial scales for surface air temperature. Such decadal forecasts are not only important in terms of providing information on the impacts of near-term climate change, but also from the perspective of climate model validation, as hindcast experiments and a sufficient database of historical observations allow standard forecast verification methods to be used. Simulation models from the ENSEMBLES hindcast experiment [3] are evaluated and contrasted with static forecasts of the observed climatology, persistence forecasts and against simple statistical models, called dynamic climatology (DC). It is argued that DC is a more apropriate benchmark in the case of a non-stationary climate. It is found that the ENSEMBLES models do not demonstrate a significant increase in skill relative to the empirical models even at global scales over any lead time up to a decade ahead.

It is suggested that the contsruction and co-evaluation with the data-based models become a regular component of the reporting of large simulation model forecasts. The methodology presented may easily be adapted to other forecasting experiments and is expected to influence the design of future experiments. The inclusion of comparisons with dynamic climatology and other data-based approaches provide important information to both scientists and decision makers on which aspects of state-of-the-art simulation forecasts are likely to be fit for purpose.

[1] J. Bröcker and L. A. Smith. From ensemble forecasts to predictive distributions, *Tellus A*, **60(4)**, 663-678 (2007).
[2] J. Bröcker and L. A. Smith. Scoring probabilistic forecasts: The importance of being proper, *Weather and Forecasting*, **22**, 382-388 (2006).
[3] F. J. Doblas-Reyes, A. Weisheimer, T. N. Palmer, J. M. Murphy and D. Smith. Forecast quality asessment of the ENSEMBLES seasonal-to-decadal stream 2 hindcasts, *ECMWF Technical Memorandum*, **621** (2010).