Geophysical Research Abstracts, Vol. 9, 07461, 2007
SRef-ID: 1607-7962/gra/EGU2007-A-07461

Contrasting methods of ensemble interpretation

R. Binter (1), J. Broecker (1), J. Penzer (2), L. A. Smith (1) (3)

(1) Centre for the Analysis of Time Series, London School of Economics, Houghton St., London WC2A 2AE, UK, (2) Department of Statistics, London School of Economics, Houghton St., London WC2A 2AE, UK, (3) Mathematical Institute, University of Oxford, 24-29 St Giles', Oxford OX1 3LB, UK

When constructing a probability forecast from a Monte Carlo sample of ensemble members, one usually attempts to construct a continuous probability density function or assign probabilities to each of a finite set of "bins". To reach that objective, one is faced with several choices including (i) to fit the parameters of a candidate distribution (say, a Gaussian) to the ensemble members, (ii) to apply a kernel to each ensemble member or (iii) to interpret the joint distribution of ensemble members as a whole. The performance of kernel dressing is compared with that of distribution fitting, restricting attention to the case where the ensemble members and the verification are drawn from the same generating distribution while noting that, nothing guaranties this generating distribution will have the form of the candidate distribution. Both Gaussian and non-Gaussian generating distributions are considered in terms of their ignorance scores (Good, 1952, Broecker and Smith, 2007). For small ensemble sizes fitting a Gaussian distribution can be inferior to kernel dressing even if the generating distribution is in fact Gaussian. We consider the conjecture that, for small ensembles, non-Gaussian candidate distributions can out-perform fitting a Gaussian even when the generating distribution is in fact Gaussian (Jewson and Penzer, 2004). In the Gaussian case as the ensemble size increases, relative performance between kernel dressing and Gaussian fitting becomes comparable; in cases with even slight deviations from Gaussianity, however, the Gaussian fit is again found to be inferior. These claims are quantified, and it is suggested that they result from the occasional, but expected, appearance of low probability draws in the forecast-verification archive. These results carry over to the case of forecasting discrete probabilities when the number of bins is large. The case of only a 'few' bins differs, whether the target bins be of equal probability (terciles, quartiles, etc.) or equal width; both referred to hereafter as 'few-tiles'. For

various candidate distributions it is demonstrated that, in the case of ’few-tiles’ for the generating distributions considered, kernel dressing again systematically outperforms the other methods investigated.

Broecker,J., Smith, L.,A. Scoring Probabilistic Forecasts: The Importance of Being Proper. Weather and Forecasting (in print), 2007.

Good, I.J. Rational decisions. Journal of the Royal Statistical Society, 14, 107-114, 1952.

Jewson, S., Penzer, J. Optimal Year Ahead Forecasting of Temperature in the Presence of a Linear Trend, and the Pricing of Weather Derivatives. Available at SSRN: http://ssrn.com/abstract=563943, 2004.