Geophysical Research Abstracts
Vol. 14, EGU2012-10827-1, 2012
EGU General Assembly 2012

Misleading estimates of forecast quality: quantifying skill with sequential forecasts

A. S. Jarman (1) and L. A. Smith (1,2)

(1) Centre for the Analysis of Time Series, London School of Economics, London, United Kingdom (a.s.jarman@lse.ac.uk),
(2) Oxford Centre for Industrial and Applied Mathematics, University of Oxford (lenny@maths.ox.ac.uk)

Quantifying the skill in probability forecasts is complicated by the fact that the behaviour of many physical systems changes slowly over long timescales. This can lead to unreliable evaluation of the skill and the value of a forecast system. Demonstrating robust out-of-sample skill requires a sufficient sample of probability forecasts and corresponding verifications. Larger samples of evaluations result in more precise confidence intervals but what often appears to be overlooked (D. Wilks. QJRMS, 136(653), pp2109-2118, 2010) is the effect of serial correlation in a forecast/outcome archive on skill score statistics. Linear autocorrelation in a forecast time series can contribute to variance inflation in the sampling distribution, and thus result in falsely narrowed confidence intervals. The investigation is applied to the ignorance score in a broader range of systems where sample variance inflation effects have been observed. Sampling variances can sometimes be deflated in a linearly autocorrelated time series. In fact, it is not merely linear correlation but the lack of independence of consecutive forecasts which can lead to a misleading estimate of skill. This is demonstrated using a chaotic time series which lacks independence yet has no linear autocorrelation in the forecast/observation time series. These results support the need for sample size corrections to avoid overconfidence in forecast skill but also indicate that a forecast user should be aware of the implications of any serial correlation for statistical inference with skill scores.