

Increasing the Reliability of Reliability Diagrams

J. Broecker (1), L. A. Smith (1)

(1) Centre for the Analysis of Time Series, London School of Economics, United Kingdom,
(cats@lse.ac.uk)

The reliability diagram is a common diagnostic graph used to summarise and evaluate probabilistic forecasts. Its strengths lie in the ease with which it is produced and the transparency of its definition. While visually appealing, major long noted shortcomings lie in the difficulty of interpreting the graph visually; for the most part, ambiguities arise from variation in the distribution of forecast probabilities and from various binning procedures (Murphy 1977, Smith 1997). A resampling method for assigning consistency bars to the observed frequencies is introduced, which allows immediate visual evaluation as to just how likely the observed relative frequencies are under the assumption that the predicted probabilities are reliable. Further, an alternative presentation of the same information on probability paper eases quantitative evaluation and comparison. Both presentations can easily be employed for any method of binning. Code to implement this approach is available at <http://www.lse.ac.uk/collections/cats/>

Murphy 1977: A. H. Murphy and R. L. Winkler. Reliability of subjective probability forecasts of precipitation and temperature. *Applied Statistics*, 26 (1): 41–47, 1977.

Smith 1997: L. A. Smith. The maintenance of uncertainty. *Proceedings of the International School of Physics Enrico Fermi*, 133: 177–246, 1997.

Broecker 2007: J. Broecker and L. A. Smith. Increasing the Reliability of Reliability Diagrams. *Weather and Forecasting* (in press). 2007. Preprint: www.lse.ac.uk/collections/cats/papersPDFs/JP_IncreasingReliabilityDiagrams_2006.pdf