Increasing the Reliability of Reliability Diagrams

J Bröcker & LA Smith

Weather and Forecasting, 22 (3), 651-661, Nov 2006

Abstract

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 and Winkler, 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 easies quantitative evaluation and comparison. Both presentations can easily be employed for any method of binning. Code to implement this approach is available at www.lsecats.org.