Geophysical Research Abstracts
Vol. 21, EGU2019-12364, 2019
EGU General Assembly 2019
© Author(s) 2019. CC Attribution 4.0 license.

Avoiding the 'forecaster's dilemma': testing a new ensemble flood forecasting system for Australia

James Bennett (1,2), David Robertson (1), and Prasantha Hapuarachchi (3)

(1) CSIRO Land & Water, Clayton, Australia (James.Bennett@csiro.au), (2) Institute for Marine and Antarctic Studies, University of Tasmania, Sandy Bay, Australia, (3) Bureau of Meteorology, Docklands, Australia (prasantha.hapuarachchi@bom.gov.au)

Streamflow forecasting services in Australia have historically relied on deterministic forecasting systems. A substantial research effort has been directed at upgrading the Bureau of Meteorology's 7-day streamflow forecasting service from deterministic to ensemble prediction (e.g., Bennett et al., 2014; Shrestha et al. 2015; Li et al. 2017, Kabir et al. 2018). The ensemble quantifies uncertainty in both rainfall and streamflow predictions, and substantially improves the quality of deterministic forecasts. The ensemble forecasts are expected to be skilful enough to meaningfully supplement the Bureau's existing deterministic flood forecasting models.

To assess the value of ensemble forecasts for floods, forecast evaluation must be restricted to extreme events. Lerch et al. (2017) identified the 'forecaster's dilemma' facing those who restrict the evaluation of forecasts to a set of historical disasters. The 'dilemma' is that this approach rewards frequent prediction of disasters, leading to a large number of false alarms. One way of avoiding the forecaster's dilemma is to evaluate instances where an extreme event is forecast (irrespective of whether a corresponding extreme occurred in the historical record).

We describe a method for assessing the value of ensemble flood predictions that avoids the forecaster's dilemma. First, a threshold is identified that would normally result in a flood watch or warning being issued. If a forecast exceeds this threshold, it is added to the pool of forecasts to be evaluated. These forecasts are then evaluated with a range of probabilistic measures of forecast quality. We show that this method is generally useful for assessing retrospective predictions of floods, but can result in counter-intuitive results in cases where ensemble distributions are not reliable. We discuss the inherent difficulty of assessing ensemble forecasts – which can only be robustly assessed over large numbers of forecasts – for extreme events, and make a number of practical recommendations to overcome these difficulties.

References

Bennett JC, Robertson DE, Shrestha DL, Wang QJ, Enever D, Hapuarachchi P, Tuteja NK. 2014. A system for continuous hydrological ensemble forecasting (SCHEF) to lead times of 9 days. Journal of Hydrology 519: 2832-2846. DOI: 10.1016/j.jhydrol.2014.08.010.

Kabir A, Hasan MM, Hapuarachchi HAP, Zhang XS, Liyanage J, Gamage N, Laugesen R, Plastow K, MacDonald A, Bari MA, Tuteja NK, Robertson DE, Shrestha DL, Bennett JC. 2018. Evaluation of multi-model rainfall forecasts for the national 7-day ensemble streamflow forecasting service, Hydrology & Water Resources Symposium, Melbourne, 3-6 December 2018.

Lerch S, Thorarinsdottir TL, Ravazzolo F, Gneiting T. 2017. Forecaster's Dilemma: Extreme Events and Forecast Evaluation. Statist. Sci. 32: 106-127. DOI: 10.1214/16-sts588.

Li M, Wang QJ, Robertson DE, Bennett JC. 2017. Improved error modelling for streamflow forecasting at hourly time steps by splitting hydrographs into rising and falling limbs. Journal of Hydrology 555: 586-599. DOI: 10.1016/j.jhydrol.2017.10.057.

Shrestha DL, Robertson DE, Bennett JC, Wang QJ. 2015. Improving precipitation forecasts by generating ensembles through postprocessing. Monthly Weather Review 143: 3642-3663. DOI: 10.1175/mwr-d-14-00329.1.