

# P-VALUES AS ANALYTICAL TOOLS IN PROBABILISTIC FORECAST ASSESSMENTS

Aline de H. N. Maia, Holger Meinke, Sarah Lennox, Roger Stone

**ABSTRACT** - Much has been written about 'quality' of probabilistic forecasts. Often, providers and users of such forecasts are unclear about what 'quality' entails, leading to confusion and misinformation. Here we try to provide some guidance and suggest a general approach to communicate aspects of probabilistic forecast quality related to discriminatory ability ( $DA$ ) and skill ( $S$ ). In our opinion, these two components of forecast quality should be considered independently.  $DA$  represents the additional knowledge about future states arising from some forecast system ( $FS$ ) over and above the total variability of the prognostic variable while  $S$  quantifies changes in the agreement between observed and predicted values when using a specific  $FS$  instead of a  $FS$  based on "climatology" only. The major concerns are: generally poor distinction between  $DA$  and  $S$ ; inappropriate use of significance testing to quantify  $DA$  and use of  $DA$  and  $S$  measures that do not account for the series lengths and/or number of classes of the  $FS$ . To address all of these issues, we propose the use of p-values derived from non-parametric tests as direct measures of  $DA$  and  $S$ . We illustrate this approach by quantifying  $DA$  and  $S$  of the Southern Oscillation Index applied to forecasting rainfall across Australia.