Evaluating ensemble flood forecasts: a survey of online news media and scientific papers

Schalk Jan van Andel (1), Anouk van Stokkom (2), Gerben Dekker (3), and Anne Margot Leicher (4)

(1) IHE Delft, Delft, the Netherlands (s.vanandel@un-ihe.org), (2) ab.vanstokkom@gmail.com, (3) Rijkswaterstaat, Middelburg, the Netherlands (gjdekker@outlook.com), (4) Nelen & Schuurmans, Utrecht, the Netherlands (margot.leicher@nelen-schuurmans.nl)

When evaluating performance of ensemble streamflow forecasts, in scientific papers mostly verification metrics are reported, such as relative mean error, relative operating characteristic, and mean continuous ranked probability skill score. The research presented here analyses how such reported performances, compare with evaluation of one of the application areas: flood forecasting and early warning. The anecdotal evaluation concerns reviewing online news media for forecasts and warnings of floods and for reports of floods occurred. While realising there is often a complex pathway between forecast and published warning online, it is discussed whether both evaluations leave us with a similar sense of 'where do we stand today' with the performance of ensemble flood forecasting, or, if different, how to come to a combined or integrated viewpoint.