EMS Annual Meeting Abstracts
Vol. 18, EMS2021-490, 2021
https://doi.org/10.5194/ems2021-490
EMS Annual Meeting 2021

End-to-end verification of the ensemble precipitation-to-river-flow forecasting chain: how to maximise skill for the user and does the uncertainty propagate?

Marion Mittermaier[1], Seonaid Anderson[2], Ric Crocker[1], Steven Cole[2], Robert Moore[2], and Sebastian Cole[1]
[1]Met Office, Exeter, United Kingdom of Great Britain – England, Scotland, Wales (ken.mylne@metoffice.gov.uk)
[2]UK Centre for Ecology and Hydrology, Wallingford, UK

Forecasting the potential for flood-producing precipitation and any subsequent flooding is a challenging task; the process is highly non-linear and inherently uncertain. Acknowledging and accounting for the uncertainty in precipitation and flood forecasts has become increasingly important with the move to risk-based warning and guidance services which combine the likelihood of flooding with the potential impact on society and the environment.

A standard approach to accounting for uncertainty is to generate ensemble forecasts. Here the national Grid-to-Grid (G2G) model is coupled to a Best Medium Range (BMR) ensemble which consists of three models spanning different time horizons: an ensemble nowcast for the first 6h, which is blended with the short-range 2.2 km Met Office Global Regional Ensemble Prediction System (MOGREPS-UK) ensemble up to 36h and the ~20 km global MOGREPS-G up to day 6. The G2G model is driven by 15-minute accumulations on a 1 km grid.

16-months of precipitation and river flow ensemble forecasts have been processed to develop and assess a joint verification framework which can facilitate the evaluation of the end-to-end forecasting chain. Analysis concluded the following: (1) daily precipitation accumulations provide the best guidance in terms of rain volume for hydrological impacts. One reason may be because it removes the impact of timing errors at the sub-daily scale. However, sub-daily precipitation can be more closely related to river flow on an ensemble member-by-member basis. (2) Observation uncertainty is important. The same forecasts verified against three different observed precipitation sources (raingauge, radar or merged) can provide markedly different results and interpretations. G2G river flow performance can also be affected, when driven by these datasets rather than forecasts. (3) The change in precipitation-intensity with model is evident and has an impact on downstream modelling and verification. (4) The period used for ensemble verification should be at least two years. The 16-month test period was sufficient for generating enough precipitation threshold-exceedances for the 95th percentile: but insufficient for higher thresholds and for river flow thresholds above half the median annual maximum flood at sub-regional scales. (5) A new method of presenting Time-Window Probabilities (TWPs) has been developed for precipitation thresholds that are hydrologically relevant. Verification of these shows that probabilities are larger, and more reliable so that users can have greater confidence in them. (6) Overall precipitation forecast skill was far more uniform than for river-flow, primarily because the atmosphere is a continuum whilst catchments are finite and subject to external, non-atmospheric factors including antecedent moisture. (7) Though G2G can be sensitive to precipitation outliers, the precipitation

ensemble is generally under-spread and spread does not appear to amplify or propagate to enhance the river flow ensemble spread, so spread is reduced rather than increased in the downstream application.