Alternative configurations of quantile regression for estimating predictive uncertainty in water level forecasts for the upper Severn River: a comparison

P. Lopez Lopez[1,2], J. S. Verkade[2,3,4], A. H. Weerts[2,5] and D. P. Solomatine[1,3]

[1] UNESCO–IHE Institute for Water Education, Delft, the Netherlands
[2] Deltares, Delft, the Netherlands
[3] Delft University of Technology, the Netherlands
[4] Ministry of Infrastructure and the Environment, Water Management Centre of the Netherlands, River Forecasting Service, Lelystad, the Netherlands
[5] Wageningen University and Research Centre, Wageningen, the Netherlands
Email: patricia.lopez@deltares.nl

Abstract: Forecasting may reduce but can never fully eliminate uncertainty about the future. Hydrological forecasts will always be subject to many sources of uncertainty, including those originating in the meteorological forecasts used as inputs to hydrological models (e.g. precipitation and temperature), and in the hydrological models themselves (e.g. model structure, model parameters and human influences). Informed decision-making may benefit from estimating the remaining uncertainties.

In the literature, various approaches to estimate predictive uncertainty have been presented. One of those is statistical post-processing. Estimating predictive uncertainty through statistical post-processing techniques comprises an analysis of past, "observed" predictive uncertainty to build a model of future predictive uncertainties. Several hydrologic post-processors have been described in the scientific literature, including the Hydrological Uncertainty Processor, the Bayesian Model Averaging, the Model Conditional Processor, and Quantile Regression. The present research focuses on the latter technique.

The study comprises an intercomparison of different configurations of a statistical post-processor that is used to estimate predictive hydrological uncertainty. It builds on earlier work by Weerts, Winsemius and Verkade (2011; hereafter referred to as WWV2011), who used the quantile regression technique to estimate predictive hydrological uncertainty using a deterministic water level forecast as a predictor. The various configurations are designed to address two issues with the WWV2011 implementation: (i) quantile crossing, which causes non-strictly rising cumulative predictive distributions, and (ii) the use of linear quantile models to describe joint distributions that may not be strictly linear. Thus, four configurations were built: (i) a "classical" quantile regression, (ii) a configuration that implements a non-crossing quantile technique, (iii) a configuration where quantile models are built in normal space after application of the normal quantile transformation (NQT) (similar to the implementation used by WWV2011), and (iv) a configuration that builds quantile model separately on separate domains of the predictor. Using each configuration, four reforecasting series of water levels at 14 stations in the upper Severn River were established. The quality of these four series was intercompared using a set of graphical and numerical verification metrics for a large number of sub-sets of available data, each representative for increasingly higher events.

Verification showed that, unconditionally, in terms of all skills and metrics, forecast quality is positive. However, the analysis also shows that forecast quality and skill decreases with increasing value of the event. Intercomparison showed that reliability and sharpness vary across configurations, but in none of the configurations do these two forecast quality aspects improve simultaneously. Further analysis shows that skills in terms of the brier skill score, mean continuous ranked probability skill score and relative operating characteristic score is very similar across the four configurations.

Keywords: *Uncertainty, hydrological forecasting, Quantile Regression, verification*