



The potential of data driven approaches for quantifying hydrological extremes

Sandra Margrit Hauswirth¹, Marc Bierkens^{1,2}, Vincent Beijk³, and Niko Wanders¹

¹Utrecht University, Geosciences, Physical Geography, Netherlands

²Deltares, Utrecht, The Netherlands

³Rijkswaterstaat, Water, Verkeer en Leefomgeving, Utrecht, The Netherlands

Recent droughts have shown that national water systems are facing increasing challenges over the last few years. As such, the Netherlands has seen increasing needs to adapt their water management in order to improve their preparedness for current and future drought events. Ideally, the necessary information needed for operational water management decisions should be readily available ahead in time and/or computed in a flexible and efficient way to ensure the various management actions. In this study we show that in addition to the physically based hydrological models, the upcoming and promising trend of incorporating machine learning (ML) in hydrology can increase the information produced to support national and operational water management.

To investigate the potential of ML for this case, we assessed 5 different ML methods to predict the following hydrological variables relevant for water management at a national scale: timeseries of discharge, groundwater levels, surface water levels and surface water temperatures. We developed a unified workflow for all the methods and variables of interest. As inputs, we only used a limited set of hydro-meteorological variables and general water management policies that are readily available on a daily basis and that can be used when the ML methods are used in seasonal forecasting mode.

We show that all methods have a good performance, with a normalized RMSE ranging between 0.0 and 0.4, and Random Forest outperforming other methods. This performance remains stable for low flows, where we observe that complex ML methods outperform simpler algorithms. The addition of water management in the ML routine increases overall performance, although limited. Finally, we observe that locations further upstream show a better performance due to the limited water management influence and close proximity to input observations.

Our study shows that ML has potential in predicting different hydrological variables at various locations at a national scale with only a simple input data set of 5 meteorological and hydrological variables. We additionally were able to capture and incorporate water management information in our analysis, creating a base for future experiments where a combination of seasonal forecasting and scenario analysis might reveal ML-informed mitigation strategies. As such, our approach may improve the preparedness of the national water system of the Netherlands for future drought

events.