Behind the scenes of runoff performance

Tanja de Boer-Euser¹, Laurène Bouaziz¹,², Guillaume Thirel³, Lieke Melsen⁴, Joost Buitink⁴, Claudia Brauer⁴, Jan de Niël⁵, Sotirios Moustakas⁵, Patrick Willems⁵,⁶, Benjamin Grelier⁷, Gilles Drogues⁷, Fabrizio Fenicia⁸, Jiri Nosent⁹, Fernando Pereira⁹, Hubert Savenije¹, Albrecht Weerts³,⁴, and Markus Hrachowitz¹

¹Water Resources Section, Faculty of Civil Engineering and Geosciences, Delft University of Technology, P.O. Box 5048, 2600 GA Delft, the Netherlands
²Department Catchment and Urban Hydrology, Deltares, Boussinesqweg 1, 2629 HV Delft, the Netherlands
³Université Paris-Saclay, INRAE, UR HYCAR, 92160, Antony, France
⁴Hydrology and Quantitative Water Management Group, Department of Environmental Sciences, Wageningen University, Wageningen, the Netherlands
⁵Hydraulics division, Department of Civil Engineering, KU Leuven, Kasteelpark Arenberg 40, BE-3001 Leuven, Belgium
⁶Vrije Universiteit Brussel (VUB), Department of Hydrology and Hydraulic Engineering, Pleinlaan 2, 1050 Brussels, Belgium
⁷LOTERR – Centre de recherche en Géographie, Université de Lorraine, Ile du Saulcy, 57045 Metz Cedex 1, France
⁸Eawag, Uberlandstrasse 133, CH-8600 Dubendorf, Switzerland
⁹Flanders Hydraulics Research, Berchemlei 115, B-2140 Antwerp, Belgium

Hydrological models are valuable tools for short-term forecasting of river flows, long-term predictions for water resources management and to increase our understanding of the complex interactions of water storage and release processes at the catchment scale. Hydrological models provide relatively robust estimates of streamflow dynamics, as shown by the countless applications in many regions across the world. However, various model structures can lead to similar aggregated outputs, i.e. model equifinality. To provide reliable estimates, it is of critical importance that not only the aggregated response but also the internal behaviors are consistent with their real-world equivalents. In a previous international comparison study (de Boer-Euser et al., 2017), eight research groups followed the same protocol to calibrate their twelve models on streamflow for several catchments within the Meuse basin. In the current study, we hypothesize that these twelve process-based models with similar runoff performance have similar representations of internal states and fluxes. We test our hypothesis by comparing internal states and fluxes between models and we assess their plausibility using remotely-sensed products of actual evaporation, snow cover, soil moisture and total storage anomalies. Our results indicate that models with similar runoff performance represent internal states and fluxes differently. The dissimilarities in internal process representation imply that these models cannot all simultaneously be close to reality. Using remotely-sensed products, the plausibility of process representation could only be evaluated to some extent as many variables remain unknown, highlighting the need for more experimental research. The study further emphasizes the value of multi-model, multi-parameter studies to reveal to decision-makers the uncertainty inherent to the lack of evaluation data and the heterogeneous hydrological landscape.
References: