



How good is good in hydrological modelling?

Jan Seibert (1), Marc Vis (1), Elizabeth Lewis (2), and Ilja van Meerveld (1)

(1) University of Zurich, Department of Geography, Zürich, Switzerland (jan.seibert@geo.uzh.ch), (2) Newcastle University, School of Civil Engineering and Geosciences

While obviously good performance for current conditions is not sufficient when using a hydrological model for an impact study, the capability of a model to reproduce streamflow for current conditions increases our confidence that a model might be a suitable representation of catchment functioning. However, when assessing the performance of a hydrological model, a question that can be raised is ‘how good is really good’? Despite several calls to use benchmarks, model performance in the scientific literature, conference presentations and discussions among hydrological modelers, is still often solely judged based on the value of some performance measure. For instance a model is rated as well-performing because the (Nash-Sutcliffe) model efficiency values are above 0.7. Based on our experiences with the application of hydrological models for largely varying catchments, we argue that such judgments on model performance can only be made if model performances are related to benchmarks that represent what could and should be expected. In this contribution we propose the use of upper and lower benchmarks and suggest possibilities for concrete benchmarks based on simulations using simple hydrological models which implicitly take observation uncertainties in both input and output time series into account. We evaluated how much upper and lower benchmarks vary for 300+ and 600+ catchments in the UK and the US respectively, how these benchmark performances are related to catchment characteristics and demonstrate the potential use of upper and lower benchmarks to evaluate model performance based on the example of the uncalibrated application of the SHE model to the 300+ catchments in the UK.