



Assessment of classical performance measures and signature indices from Flow Duration Curves for model evaluation.

Rita Ley (1), Hugo Hellebrand (1), Markus C. Casper (1), and Fabrizio Fenicia (2)

(1) University of Trier, Germany, (2) Swiss Federal Institute of Aquatic Science and Technology (Eawag), Switzerland

The result of model evaluation is strongly influenced by the choice of the used performance measures. There exist a large variety of performance measures, each with its strengths and weaknesses. Although all of them represent the ability of a hydrological model to reproduce observed stream flow, it is unclear which one is most appropriate for specific applications. The objective of this study is to investigate which performance measure is best suited to find a best performing model structure for a single basin out of multiple model structures. We compare the usability of a new performance measure, the Standardized Signature Index Sum, with several classical statistical performance measures and hydrological performance measures like the Root Mean Square Error or the Nash and Sutcliffe Efficiency. In contrast to the classical and hydrological performance measures, the Standardized Signature Index Sum is based on the comparison of observed and simulated Flow Duration Curves (FDCs). It combines the performance for different parts of the FDC to one measure considering the whole FDC and therefore the whole hydrograph.

For this purpose 12 model structures were generated using the SUPERFLEX modeling framework and applied to 53 meso-scale basins in Rhineland Palatinate (Germany). For all calibrated models based on the 12 model structures and 53 basins, we calculate several performance measures and compare their usability to identify a best performing model structure for each basin.

In many cases the classical performance measures and the hydrological performance measures assigned similar values to seemingly different hydrographs simulated with different model structures. Therefore, these measures are not well suited for model comparison. The proposed Standardized Signature Index Sum is more effective in revealing differences between model results. Furthermore, it provides information in which part of the hydrograph and how a model fails. The Signature Index Sum allows for a more distinctive identification of a best performing model for individual basins than the tested classical and hydrological performance measures.