Geophysical Research Abstracts
Vol. 13, EGU2011-6832-1, 2011
EGU General Assembly 2011

Evaluation of Different Performance Criteria for Calibrating Conceptual Hydrological Models

Thomas Wöhling (1,2), Luis Samaniego (3), and Rohini Kumar (3)

(1) Water & Earth System Science Research Centre (WESS), University of Tübingen, Institute for Geoscience, Sigwartstr. 10, 72076 Tübingen, Germany, Email: thomas.woehling@uni-tuebingen.de, (2) Lincoln Ventures Ltd, Environmental Research Division, Hamilton, New Zealand, (3) Helmholtz-Centre for Environmental Research - UFZ, Computational Hydrosystems, Permoserstraße 15, 04318 Leipzig, Germany

Different criteria can be used for the calibration and evaluation of conceptual hydrological models with observed data. One of the most common performance criteria is the Nash-Sutcliffe efficiency (NSE, Nash and Sutcliffe 1970). However, Gupta et al. (2009) showed that there are systematic problems inherent with any optimization based on formulations related to the NSE. In this study, 53 transfer parameters of the distributed conceptual mHM model (Samaniego et al. 2010) for the upper part of the meso-scale Goldersbach catchment (36.84 km²) in the Neckar basin (Germany) are calibrated using nine years of observed daily runoff and the global search algorithm AMALGAM. The model performance is evaluated using the following criteria: the NSE of discharge (NSE_Q), the NSE of log-transformed discharge (NSE_{lnQ}), the sum-squared error of discharge (SSE_Q), the sum-squared error of monthly discharge sums (SSE_{MQ}), the King-Gupta efficiency (KGE_Q, Gupta et al. 2009), and three criteria obtained by a decomposition of the NSE. These criteria are the linear correlation coefficient (R), a measure for relative variability (α), and a bias component (β). The model calibration is posed in a multi-criteria context using three different performance criteria. NSE_Q, NSE_{lnQ}, and SSE_{MQ} are used in the first optimization run whereas R, α, and β are utilized in the second multi-criteria calibration run. In each of the two AMALGAM runs, the trade-off between the different performance criteria is determined by Pareto efficient parameter solutions. The extreme ends of the bi-criterion Pareto fronts and an equally weighted compromise solution of all three criteria are selected for further analysis. The results suggest that the 53 optimised transfer parameter values are very sensitive to the utilized performance criterion. The parameter set that yields the largest NSE_Q – value in the first run tends to slightly favour peak flows but exhibits a very good overall performance during the calibration period. On the other hand, the parameter set that yields the largest NSE_{lnQ} – value (a common criterion to give emphasis to low flows) results in relatively large volume balance errors and tends to underestimate variability. A very good overall performance was achieved with the compromise solution of the first optimization run that yielded better NSE_Q, NSE_{lnQ}, SSE_{MQ}, and SSE_Q – values than any of the solutions in the second AMALGAM run while maintaining R, α, and β – values close to their ideal values. The results demonstrate the advantages of multi-criterion calibration of conceptual hydrological models. To strengthen confidence in the generality of the results, a number of other catchments will be tested.