Think multidimensionally: How well do we sample the parameter space when searching for optimal parameter sets in hydrological modelling?

Hydrological models are used for water management decisions and to predict the effects of climate or land use change on water resources. The parameters used to calculate the different water fluxes in a hydrological model usually need to be calibrated. This is typically done by maximizing the agreement between the observed and simulated discharge (i.e., optimizing model fit) and thus finding well-performing parameter sets.

To consider the fact that there might be different, similarly suitable parameter sets, we often use Monte Carlo approaches with millions of model runs (starting with randomly selected parameter sets). However, these approaches are very demanding computationally. Therefore, we often use the genetic algorithm and Powell optimization (GAP; Seibert, 2000) as implemented in the HBV-light software (Seibert & Vis, 2012) to search the parameter space for suitable parameter sets. With the genetic algorithm, the evolutionary development ("survival of the fittest") is mimicked: From a generation of parameter sets, the ones that perform better have a higher chance to be promoted as "parents" for the next generation. Thus, the algorithm tends to find better performing parameter sets with every generation and moves towards local maxima in the parameter space. However, it remains an open question how well the genetic algorithm is suitable to identify parameter sets in the entire parameter space, i.e., if the whole parameter space is covered by the search algorithm. After the use of the genetic algorithm, the Powell optimization allows for a local optimization by efficiently fine-tuning a parameter set to reach the local optimum. The exact benefit from running the Powell optimization after the genetic algorithm remains to be investigated.

Even though these questions are relevant for all research projects that make use of the GAP optimization, no systematic investigation has been conducted regarding these questions yet. Thus, for all research projects using the GAP optimization in the future, answers to these methodological questions will provide opportunities to improve the modelling strategies.

To be able to make general statements, we suggest an investigation of these processes in a variety of catchments, i.e., to use a large-sample data set, such as the CAMELS data set (Addor et al., 2017) containing discharge data and catchment information for over 600 unregulated but gauged catchments in the United States.

The thesis requires an interest in hydrologic modelling and (developing) programming skills (e.g., R or Python). The scope of the Master's thesis can be adjusted to the student's interest. For more information, contact Franziska Clerc-Schwarzenbach (Y25-L74): franziska.schwarzenbach@geo.uzh.ch

Reading suggestions

Addor, N., Newman, A. J., Mizukami, N., & Clark, M. P. (2017). The CAMELS data set: Catchment attributes and meteorology for large-sample studies. *Hydrology and Earth System Sciences*, *21*(10), 5293–5313. https://doi.org/10.5194/hess-21-5293-2017

Seibert, J. (2000). Multi-criteria calibration of a conceptual runoff model using a genetic algorithm. *Hydrology and Earth System Sciences*, *4*(2), 215–224. https://doi.org/10.5194/hess-4-215-2000

Seibert, J., & Vis, M. J. P. (2012). Teaching hydrological modeling with a user-friendly catchment-runoff-model software package. *Hydrology and Earth System Sciences*, *16*(9), 3315–3325. https://doi.org/10.5194/hess-16-3315-2012