Mapping the benefit of multi-variable model calibration for hydrological predictions

S. Pool, K. Fowler and M. Peel

Department of Infrastructure Engineering, Faculty of Engineering and Information Technology, University of Melbourne, Australia
Email: sandra.pool@unimelb.edu.au

Abstract: Hydrological predictions from rainfall-runoff models are of vital socioeconomic importance. The predictions are typically based on models that were previously calibrated against observed streamflow and are assumed to reasonably represent hydrological processes in a particular catchment. Here, we systematically test this assumption for 550 Australian catchments by evaluating simulated monthly actual evapotranspiration (ETa), soil wetness (SW), and total water storage (TWS) after calibrating the SIMHYD model against ten years of observed daily streamflow data. Model performance is quantified by a non-parametric version of the Kling-Gupta efficiency. We show that ETa, SW, and TWS are for many catchments poorly simulated with simulations being comparable to those of the uncalibrated SIMHYD. Mapping the results in space indicates two regional hotspots of low model performance. First, ETa is difficult to simulate along the south and southwest coast, where ETa and streamflow have a different seasonality. Second, predictions of SW and TWS are most challenging along the southeast coast, where SM and TWS are weakly linked to streamflow. Given these surprisingly low model performances for ETa, SW, and TWS, we hypothesize that their consideration in model calibration leads to more satisfying results. We, therefore, conducted an additional multi-variable calibration exercise in which ETa, SW, and TWS were used to complement the traditional calibration against streamflow. To assess the added value of a multi-variable calibration, we use the concept of benchmarks. We first defined a lower benchmark, which corresponds to the median performance of the uncalibrated SIMHYD (using 100’000 random parameter values) for the variable of interest (X_{uncal}). Second, we define an upper benchmark, which corresponds to the performance of SIMHYD calibrated solely to the variable of interest (X_{cal}). The difference between the upper and the lower benchmarks ($\Delta X = X_{cal} - X_{uncal}$) represents the benefit of calibration for that variable. We can now use ΔX as reference for comparing the performance of SIMHYD when calibrated against streamflow or streamflow plus the variable of interest. For example, the reproduction of ETa reached, on average, 16% of ΔX when calibrated against streamflow, which increased to 63% when calibrated against streamflow and ETa. Similarly, the average performance increased from -6% to 37% for SW, and 18% to 56% for TWS. Particularly high benefits of the multi-variable calibration are observed in the two hotspot regions, where the traditional model calibration had most difficulties in simulating a range of hydrological processes. However, these benefits typically come at the expense of a reduced model performance for streamflow (on average up to -12%). This trade-off between reduced streamflow performance and improved ETa, SW, and TWS predictions tends to get more pronounced with increasing catchment wetness. The limitations of our study include that the “observed data” for ETa, SW and TWS are each based upon remotely sensed information, and these data may themselves be subject to error. Despite these limitations, our results demonstrate the benefits and trade-offs related to an improved representation of simulated hydrological processes, which is likely to become more important under changing climatic conditions.

Keywords: *Model calibration, remote-sensing data, large-sample hydrology*