Geophysical Research Abstracts
Vol. 18, EGU2016-11352, 2016
EGU General Assembly 2016

The impact of objective function selection on the influence of individual data points

David Wright, Mark Thyer, Seth Westra, and David McInerney
School of Civil, Environmental and Mining Engineering, University of Adelaide

Across the field of hydrology practitioners apply a range of objective functions which are selected based upon the intended model application and suitability of the objective function assumptions to the data in question. Despite most objective functions providing fundamentally different calibration results there are currently limited methods for comparison of alternatives. Influence diagnostics quantify the impact of individual data points on model performance, parameters and predictions. The goal of this study is to use compare four commonly applied objective functions in hydrology using influence diagnostics to provide insights on how objective function selection changes the influence of individual data points on model calibration. The specific aims are to: 1) explore the impact on magnitude of influence of objective functions, 2) investigate similarities between influential points identified by objective functions and, 3) categorise flows that are influential under objective functions.

We use case-deletion influence diagnostics to examine four objective functions: Standard Least Squares (SLS), Weighted Least Squares (WLS), Log transformed flows (LOG) and the Kling-Gupta Efficiency (KGE). We apply these objective functions to six scenarios: two conceptual hydrological models (GR4J and IHACRES) across three catchment case studies with varying runoff coefficients (0.14 to 0.57). We quantify influence using the case-deletion relative change in flow metrics: mean flow prediction, maximum flow prediction, and the 10th percentile low flow prediction.

The results show that when using objective functions SLS and KGE influential data points have larger magnitude influence (maximum of 10% change in the flow metrics across all data points for both objective functions) than heteroscedastic WLS and LOG (WLS maximum of 8% and LOG maximum of 6% change in the flow metrics). SLS and KGE identify similar influential points (75% of the most influential points are common to both objective functions) whereas other objective function combinations identify less similar points (ranging from 10 to 20% most influential points in common). We show that in general, the most influential flows for SLS and KGE tend to be high flows, whereas for WLS and LOG the most influential flows tend to be low flows. This highlights that the choice of objective function plays an important role in determining the data points that are most critical in model calibration. Unexpectedly there were a small number of cases across the six scenarios, where the opposite was true, highlighting the need to examine influence diagnostics during the model calibration process to ensure that the calibration is robust to highly influential points.

In a field where the norm is increasing theoretical complexity with limited tools to examine the robustness of calibration, influence diagnostics provide important insights. This work has shed light on four objective functions that are commonly applied across the hydrological discipline and shown how the choice of objective function can produce fundamentally different model predictions leading to substantial changes in predictions of extreme and base flows.