Geophysical Research Abstracts
Vol. 17, EGU2015-8503, 2015
EGU General Assembly 2015
© Author(s) 2015. CC Attribution 3.0 License.

Identifying influential data points in hydrological model calibration and their impact on streamflow predictions

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

Highly influential data points are those that have a disproportionately large impact on model performance, parameters and predictions. However, in current hydrological modelling practice the relative influence of individual data points on hydrological model calibration is not commonly evaluated. This presentation illustrates and evaluates several influence diagnostics tools that hydrological modellers can use to assess the relative influence of data. The feasibility and importance of including influence detection diagnostics as a standard tool in hydrological model calibration is discussed.

Two classes of influence diagnostics are evaluated: (1) computationally demanding numerical "case deletion" diagnostics; and (2) computationally efficient analytical diagnostics, based on Cook's distance. These diagnostics are compared against hydrologically orientated diagnostics that describe changes in the model parameters (measured through the Mahalanobis distance), performance (objective function displacement) and predictions (mean and maximum streamflow).

These influence diagnostics are applied to two case studies: a stage/discharge rating curve model, and a conceptual rainfall-runoff model (GR4J). Removing a single data point from the calibration resulted in differences to mean flow predictions of up to 6% for the rating curve model, and differences to mean and maximum flow predictions of up to 10% and 17%, respectively, for the hydrological model. When using the Nash-Sutcliffe efficiency in calibration, the computationally cheaper Cook's distance metrics produce similar results to the case-deletion metrics at a fraction of the computational cost. However, Cooks distance is adapted from linear regression with inherit assumptions on the data and is therefore less flexible than case deletion.

Influential point detection diagnostics show great potential to improve current hydrological modelling practices by identifying highly influential data points. The findings of this study establish the feasibility and importance of including influential point detection diagnostics as a standard tool in hydrological model calibration. They provide the hydrologist with important information on whether model calibration is susceptible to a small number of highly influent data points. This enables the hydrologist to make a more informed decision of whether to (1) remove/retain the calibration data; (2) adjust the calibration strategy and/or hydrological model to reduce the susceptibility of model predictions to a small number of influential observations.