Geophysical Research Abstracts
Vol. 14, EGU2012-11549-1, 2012
EGU General Assembly 2012

Comparing Goodness-of-fit Measures for Calibration of Models Focused on Extreme Events

M. Zambrano-Bigiarini (1) and A. Bellin (2)

(1) Water Resources Unit, Institute for Environment and Sustainability, Joint Research Centre, Ispra, Italy (mauricio.zambrano@jrc.ec.europa.eu), (2) Dept. of Civil and Env. Engineering, Università degli Studi di Trento, Trento, Italy (alberto.bellin@ing.unitn.it)

Hydrological models have been increasingly used during the last decades for flood forecasting and water resources management, among others. Therefore, a proper assessment of the reliability of hydrological simulations is of utmost importance to enhance societal confidence on model predictions. To date, several goodness-of-fit measures have been proposed to assess model performance and to measure the agreement between observations and simulated equivalents. Despite serious and well-known limitations, many single-objective goodness-of-fit measures are still of widespread use. As an example, the Nash-Sutcliffe efficiency (*NSE*) has been highly criticised as an inappropriate benchmark for comparing modelling results to observations, nonetheless, it is still one of the most common performance measures used by both practitioners and environmental scientists.

This work examines how different goodness-of-fit measures reported in literature behave when used within a single-objective optimisation procedure, both for the identification of the model's parameters and in the reproduction of high- and low-flow events. By doing so, several parameters of the semi-distributed Soil and Water Assessment Tool (SWAT) 2005 are calibrated, by using a novel global optimisation technique called Particle Swarm Optimization (PSO) and changing only the goodness-of-fit measure to be optimised. In particular, the Kling-Gupta efficiency (KGE), the *NSE*, the index of agreement (*d*), along with modified and relative versions of the last two measures are compared for a calibration of SWAT aiming at reproducing low flows. On the other hand, KGE, *NSE*, *d*, and the coefficient of persistence (*cp*) are compared for a calibration aiming at reproducing high flows. In addition, a relatively new goodness-of-fit measure is introduced, which, when used along with some ad-hoc weighting scheme, allows capturing model errors in the high or low spectrum of the analysed time series with little influence of errors in other portions of the signal.

Optimal parameter values presented a considerable variation depending on the objective function used in PSO. Discharge values obtained during calibration are used to compute Empirical Cumulative Distribution Functions (ECDFs) for three different quantiles representative of simulated low, medium and high flows. Simulated quantiles computed with the new goodness-of-fit function in combination with the ad-hoc weighting scheme were closer to their observed counterparts. Results provide quantitative guidance about the bias of the calibrated hydrological model to reproduce low and high flows when different well-known goodness-of-fit measures are used as objective function during calibration. The latter facilitates the elaboration of standards about which benchmarks to use when trying to represent different extreme events.