Geophysical Research Abstracts
Vol. 20, EGU2018-538, 2018
EGU General Assembly 2018

Accuracy and frequency of crowd-sourced data with respect to their use in hydrological models

Simon Etter, Barbara Strobl, Ilja van Meerveld, and Jan Seibert
University of Zurich, Faculty of Science, Department of Geography, Switzerland (simon.etter@geo.uzh.ch)

Areas with the greatest scarcity in hydrological data are typically also those that suffer the greatest impacts from adverse hydrological conditions. These areas can potentially benefit greatly from simple hydrological data collection methods, such as those obtained by citizen scientists. Streamflow can be estimated by anyone without specialized equipment, but these estimates are potentially too inaccurate to be useful for hydrological model calibration. In the CrowdWater project, we assess the requirements of streamflow data in terms of frequency and accuracy for their use in the calibration of bucket-type hydrological models. We asked passersby to estimate the streamflow for different rivers in Switzerland. A comparison of the people's estimates with the measured streamflow data suggested a lognormal error distribution. We applied this error distribution to a year of measured hourly streamflow data to obtain synthetic datasets, from which we then took a limited number of data points to represent the temporal characteristics of crowdsourced data. We used these synthetic data and the HBV model to test the effects of the number of data points, their distribution throughout the year and the observation errors on hydrological model calibration.
The results indicate that the crowdsourced estimates of streamflow are too inaccurate to significantly improve model calibration compared to the use of random parameter sets. However, if training could reduce the standard deviation of the error distribution by a factor of two, these data would be useful for hydrological model calibration, but the effect depends on the number or crowd-sourced streamflow estimates and their distribution throughout the year. These results highlight the potential value of training citizen scientists in order to reduce the uncertainty in their streamflow estimates. Where training is not an option, water levels might be a more reliable parameter to estimate and use in hydrological model calibration. Above all, this study shows the value of using synthetic datasets based on observed data to assess the value of the data collected by citizen scientists prior to implementing a new citizen science project.