Geophysical Research Abstracts
Vol. 15, EGU2013-9466, 2013
EGU General Assembly 2013

Use of different sampling schemes in machine learning-based prediction of hydrological models' uncertainty

Nagendra Kayastha (1), Dimitri Solomatine (1,2), Durga Lal Shrestha (3), Ann van Griensven (1,4)

(1) UNESCO-IHE Institute for Water Education, Hydroinformarmatics & Knowledge Management, Delft, Netherlands (d.solomatine@unesco-ihe.org, +31-(0)15-2122921), (2) Water Resources Section, Delft University of Technology, Netherlands, (3) CSIRO Land and Water, Highett, Australia, (4) Vrij University of Brussels, Belgium

In recent years, a lot of attention in the hydrologic literature is given to model parameter uncertainty analysis. The robustness estimation of uncertainty depends on the efficiency of sampling method used to generate the best fit responses (outputs) and on ease of use. This paper aims to investigate: (1) how sampling strategies effect the uncertainty estimations of hydrological models, (2) how to use this information in machine learning predictors of models uncertainty.

Sampling of parameters may employ various algorithms. We compared seven different algorithms namely, Monte Carlo (MC) simulation, generalized likelihood uncertainty estimation (GLUE), Markov chain Monte Carlo (MCMC), shuffled complex evolution metropolis algorithm (SCEMUA), differential evolution adaptive metropolis (DREAM), partical swarm optimization (PSO) and adaptive cluster covering (ACCO) [1]. These methods were applied to estimate uncertainty of streamflow simulation using conceptual model HBV and Semi-distributed hydrological model SWAT. Nzoia catchment in West Kenya is considered as the case study. The results are compared and analysed based on the shape of the posterior distribution of parameters, uncertainty results on model outputs.

The MLUE method [2] uses results of Monte Carlo sampling (or any other sampling shceme) to build a machine learning (regression) model U able to predict uncertainty (quantiles of pdf) of a hydrological model H outputs. Inputs to these models are specially identified representative variables (past events precipitation and flows). The trained machine learning models are then employed to predict the model output uncertainty which is specific for the new input data. The problem here is that different sampling algorithms result in different data sets used to train such a model U, which leads to several models (and there is no clear evidence which model is the best since there is no basis for comparison). A solution could be to form a committee of all models U and to sue a dynamic averaging scheme to generate the final output.

[1] N. Kayastha, D. L. Shrestha and D. P. Solomatine. Experiments with several methods of parameter uncertainty estimation in hydrological modeling. Proc. 9th Intern. Conf. on Hydroinformatics, Tianjin, China, September 2010.

[2] D. L. Shrestha, N. Kayastha, and D. P. Solomatine. A novel approach to parameter uncertainty analysis of hydrological models using neural networks. Hydrol. Earth Syst. Sci., 13, 1235–1248, 2009.