Geophysical Research Abstracts
Vol. 21, EGU2019-5932, 2019
EGU General Assembly 2019

Improving the Reservoir Module of a Process-based Distributed Hydrological Model with Machine Learning: The Case of Unoptimized Dam Reservoirs

Andrew Cutler (1), Jenny Hagen (2,3), Patricia Trambauer (2), Albrecht Weerts (2,4), Pablo Suarez (5,6), and Dimitri Solomatine (7)

(1) Boston University, College of Engineering, Department of Electrical Engineering, 8 St. Mary's street, Boston, MA 02215, USA., (2) UNESCO-IHE Institute for Water Education, Department for Water Science and Engineering, Westvest 7, 2611 AX Delft, Netherlands., (3) Deltares, Inland Water Systems Division, Boussinesqweg 1, 2629 HV Delft, Netherlands., (4) Wageningen University and Research, Hydrology and Quantitative Water Management group, Department of Environmental Sciences, Lumen, Droevendaalsesteeg 3a, 6708 PB Wageningen, Netherlands., (5) Red Cross Red Crescent Climate Centre, Anna van Saksenlaan 50, 2593 HT Den Haag, Netherlands., (6) University College London, Department of Science, Technology, Engineering and Public Policy, 36-37 Fitzroy Square, London W1T 6EY, UK., (7) Delft University of Technology, Water Resources Section, Building 23, Stevinweg 1, 2628 CN, Delft, Netherlands.

With increasing data availability globally, opportunities are unlocked for sophisticated hydrological modelling in poorly gauged basins. For a case-study on flood forecasting in Togo, West-Africa, the process-based distributed hydrological model, wflow_sbm, was set up, calibrated and forced using globally available data. While wflow_sbm successfully improved the model currently predicting inflows to a dam reservoir upstream of flood-prone villages, the Nash-Sutcliffe Efficiency of outflow forecasts dropped below zero beyond two days lead time. An exquisite challenge in predicting outflows with a model like wflow_sbm is provoked when human decision-making – as opposed to reservoir optimization – governs dam releases; thus, a mismatch results between the linear model structure and the non-linear underlying mechanisms. However, machine learning techniques can be effective means of capturing highly non-linear relationships. Therefore, machine learning techniques of increasing complexity were trained to predict dam releases with up to ten days lead-time. First, an autoregressive random forest was trained, using the 50 preceding daily outflows. Extending the input variable selection to include satellite precipitation estimates, potential evapotranspiration calculated with de Bruin Equation and reanalysis temperature estimates, a three-dimensional convolutional neural network was trained. The outflow predictions were substantially improved with both techniques, but distinct advantages were also identified; the random forest – being simpler and more robust – provided better predictions at higher lead times, while the convolutional neural network obtained the highest forecast skill in terms of Kling-Gupta Efficiency at lower lead times. However, at lead times extending beyond the period of high autocorrelation, the CNN will have distinct advantages relying on forecastable variables like precipitation and temperature. The results demonstrate improvement of the wflow_sbm reservoir module in case of dam release prediction for unoptimized reservoirs, and point to the potential of exploiting machine learning techniques as complementary model structures for flood prediction where highly non-linear relationships are prevalent.