EGU2020-6111
https://doi.org/10.5194/egusphere-egu2020-6111
EGU General Assembly 2020

Machine Learning is Central to the Future of Hydrological Modeling

Grey Nearing[1], Frederik Kratzert[2], Craig Pelissier[3,4,5], Daniel Klotz[2], Jonathan Frame[1], and Hoshin Gupta[6]
[1]University of Alabama, Department of Geological Sciences, Tuscaloosa, AL, USA (gsnearing@ua.edu)
[2]LIT AI Lab & Institute for Machine Learning, Johannes Kepler University Linz, Austria
[3]NASA Goddard Space Flight Center, NASA Center for Climate Simulation, Greenbelt, MD, USA
[4]University of Maryland Baltimore County, Department of Computer Science and Electrical Engineering, Baltimore, MD, USA
[5]Science Systems Applications Inc., Lanham, MD, USA
[6]University of Arizona, Department of Hydrology and Water Resources, Tucson, AZ, USA

This talk addresses aspects of three of the seven UPH themes: (i) time variability and change, (ii) space variability and scaling, and (iii) modeling methods.

During the community contribution phase of the 23 Unsolved Problems effort, one of the suggested questions was “Does Machine Learning have a real role in hydrological modeling?” The final UPH paper claimed that “Most hydrologists would probably agree that [extrapolating to changing conditions] will require a more process-based rather than calibration-based approach as calibrated conceptual models do not usually extrapolate well.” In this talk we will present a collection of recent experiments that demonstrate how catchment models based on deep learning can account for both temporal nonstationarity and spatial information transfer (e.g., from gauged to ungauged catchments), often achieving significantly superior predictive performance compared to other state-of-the-art (process-based) modeling strategies, while also providing interpretable results. This is due to the fact that deep learning can learn, exploit, and explain catchment and hydrologic similarity in ways and with accuracies that the community has not been able to achieve using traditional methods.

We argue that the results we have obtained motivate a path forward for hydrological modeling that centers around ‘physics-informed machine learning.’ Future model development might focus on building hybrid (AI + process-informed) models with three objectives: (i) integrating known catchment behaviors into models that are also able to learn directly from data, (ii) building explainable deep learning models that allow us to extract scientific insights, and (iii) building hybrid models that are also able to simulate unobserved or sparsely observed variables. We argue further that while the sentiments expressed in the UPH paper about process-based modeling are common, the community currently lacks an evidence-based understanding of where and when process-based understanding is important for future predictions, and that addressing this question in a meaningful way will require true hybrids between different modeling approaches.

We will conclude by providing two fundamentally novel examples of physics-informed machine

learning applied to catchment-scale and point-scale modeling: (i) conservation-constrained neural network architectures applied to rainfall-runoff processes, and (ii) integrating machine learning into existing process-based models to learn unmodeled hydrologic behaviors. We will show results from applying these strategies to the CAMELS dataset in a rainfall-runoff context, and also to FluxNet soil moisture data sets.