Geophysical Research Abstracts
Vol. 21, EGU2019-11185, 2019
EGU General Assembly 2019

A data-centric Bayesian learning framework and the outlook for a deep-learning-powered avenue for hydrologic knowledge discovery

Chaopeng Shen, Wen-Ping Tsai, and Kuai Fang

Pennsylvania State University, Civil and Environmental Engineering, University Park, United States (cshen@engr.psu.edu)

Machine learning methods can serve as highly efficient and more objective generators for data-consistent hypotheses that explain how hydrologic systems function. However, they alone cannot differentiate between causal and associative relationships. The hypotheses, however, can be further assessed by the physics and causal relationships encoded in process-based models in a data-centric Bayesian learning framework, as opposed to the typical model-centric one. As a demonstration case, we show how a classification tree algorithm generated competing hypotheses explaining the gradients in storage-streamflow correlation relationships found in basins in the eastern continental US. With the help of numerical experiments using a process-based model, we then rejected soil texture and terrain slope and validated large soil thickness as the factor that causes high correlations between streamflow and storage. Based on our work in soil moisture modeling with deep learning and a trans-disciplinary literature review, we then provide an outlook of how this framework could be migrated to deep-learning-based effort, opening up a new research avenue.