EGU22-8739
https://doi.org/10.5194/egusphere-egu22-8739
EGU General Assembly 2022

Accelerating large scale groundwater simulation with machine learning: modeling approaches and science implications for changing systems

Laura Condon[1], Andrew Bennett[1], Hoang Tran[2], Ben Horowitz[2], Elena Leonarduzzi[2], Peter Melchior[2], and Reed Maxwell[2]

[1]University of Arizona, Hydrology and Atmospheric Sciences, Tucson, United States of America (lecondon@arizona.edu)
[2]Princeton University, Princeton, United States of America

It is well established that groundwater is an important buffer to hydrologic systems; stabilizing water supplies across spatial scales and long time frames. However, groundwater surface water interactions are non-linear and can vary greatly based on climate and hydrogeologic setting. This challenge is exacerbated in changing systems where shifting land cover, extreme droughts and floods can significantly change groundwater storage, discharge and recharge dynamics. Observations of groundwater levels and the hydrogeologic properties that govern flow are sparse both in space and time. As a result, we rely heavily on physically based numerical models to help us understand this critical component of the hydrologic cycle. There are an increasing number of national to global scale groundwater models that take a variety of numerical approaches and simplify the system to varying degrees. One of the challenges we face is that the integrated models best suited to capture changing dynamics, are also by far the most computationally expensive. This creates a trade-off between the physical complexity we can represent and the size of the ensembles we can explore (another critical dimension in highly uncertain systems).

Here we will explore the potential for machine learning emulators to help accelerate solutions while maintaining physically rigorous solutions in changing systems. First, we present progress in the development of the next generation high resolution (1km2) ParFlow model of the contiguous US (ParFlow-CONUS). Next, we explore a range of machine learning architectures that have been developed to emulate the national model. Here we focus on solutions that can emulate the full 3D subsurface system as this allows for the most flexibility in hydrologic applications. Specifically, we explore 3D convolutional neural networks, recursive neural networks and LSTM approaches. For every approach we evaluate the fidelity with which the machine learning model can emulate the physics-based model. We focus specifically on performance for extreme hydrologic conditions both when the ML emulator is provided training data on these cases and when the ML model is applied to out of sample scenarios. One of the key strengths of physically based hydrologic models is their ability to represent scenarios that we haven't seen in the past. This can be a large challenge for purely data driven ML approaches which can provide erroneous results on conditions that fall outside historical behavior.