Title: **National datasets and workflows for integrated hydrologic modeling across the US**

Laura E. Condon[1]*, Reed M Maxwell[2], Hoang Tran[2,3], Jun Zhang[1,4], Chen Yang[2], Jackson Swilley[2], Danielle Tijerina[2], Mary Michael O'Niell[5], Jennifer Steyaert[1]

[1] University of Arizona, Tucson, AZ;
[2]Princeton University, Princeton, NJ;
[3]Pacific Northwest National Laboratory, Richland, WA;
[4]Nanjing Normal University, Nanjing, China;
[5]Hydrological Sciences Laboratory, NASA Goddard Space Flight Center, Greenbelt, MD;

Contact: (lecondon@arizona.edu)

Project Lead Principal Investigator (PI): David Moulton

BER Program: ESS

Project: IDEAS-watersheds

Project Website: https://ideas-watersheds.github.io/

Project Abstract:

The ParFlow-CONUS simulation platform has been in ongoing development for a decade and spans input data, several decades' worth of simulations, and comprehensive analysis tools to make the input and output data publicly available. This work has resulted in many publications that have advanced both technical capabilities for large-scale simulation, as well as scientific understanding of continental-scale hydrologic interactions including how transpiration and groundwater are linked, how the US' legacy of groundwater pumping has impacted the rest of the hydrologic cycle and how climate change may impact our groundwater resources both regionally and locally). Work is now focused on version 2 of the model (CONUS2.0) which expands from version 1 to cover the entire drainage area of the contiguous US. Over the last year we have published several publications that detail new national dataset development (Zhang et al 2021; Steyaert et al 2022), model intercomparison with the US National Water Model (Tijerina et al 2021), and a detailed evaluation of long-term version 1.0 simulations (O'Neill et al 2021).

Key to the successes of the ParFlow-CONUS platform are interoperable workflows that facilitate data and model development. While the static inputs to the model are small, forcing datasets for CONUS2.0 are much larger (~50TB in size) and the simulation outputs are expected to be almost 1PB. These large datasets have not only taken years to develop but are too large for traditional analysis approaches, and required our team to develop custom reproducible workflows in tandem with the dataset creation. To manage these large data sets and ensure the community can access them in practical ways, we are leveraging a dedicated datacenter (the Princeton Hydrologic Data Center, or PHDC) was constructed as a platform to store and serve these community resources.

In addition, our team has built subsetting workflows which have been used to construct regional models of the Upper Colorado (e.g. Tran et al 2022) and the Delaware-Susquehanna basins and will be available for the ESS community to use in their work.

This presentation will highlight the data-workflow-integration for high resolution continental- and regional-scale community model development and recent CONUS 2 accomplishments. National-scale model input datasets combine multiple sources and involve rigorous testing over both the entire CONUS 2 domain and subsets of it. The development of these data and how they improve the CONUS 2 simulation results will be discussed.

References

O'Neill, M. M. F., D. T. Tijerina, L. E. Condon, and R. M. Maxwell. "Assessment of the ParFlow–CLM CONUS 1.0 Integrated Hydrologic Model: Evaluation of Hyper-Resolution Water Balance Components across the Contiguous United States." Geoscientific Model Development 14, no. 12 (2021): 7223–54. https://doi.org/10.5194/gmd-14-7223-2021.

Steyaert, J., L.E. Condon, S. Turner and N. Voisin, "ResOpsUS a dataset of historical reservoir operations in the contiguous United States", In Press, Scientific Data.

Tijerina, D., Condon, L., FitzGerald, K., Dugger, A., O'Neill, M. M., Sampson, K., Gochis, D.J., and Maxwell, R.M. Continental hydrologic intercomparison project, phase 1: A large-scale hydrologic model comparison over the continental United States. Water Resources Research, 57, e2020WR028931, doi:10.1029/2020WR028931, 2021. https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2020WR028931

Tran, H., Zhang, J., O'Neill, M.M., Ryken, A., Condon, L.E., and Maxwell, R.M. A hydrological simulation dataset of the Upper Colorado River Basin from 1983 to 2019. Sci Data 9, 16 (2022). https://doi.org/10.1038/s41597-022-01123-w

Zhang, J., L. E. Condon, H. Tran, and R. M. Maxwell. "A National Topographic Dataset for Hydrological Modeling over the Contiguous United States." Earth Syst. Sci. Data 13, no. 7 (July 8, 2021): 3263–79. https://doi.org/10.5194/essd-13-3263-2021.