



Jun 25th, 9:00 AM - 10:20 AM

Managing Big Data Output from the U.S. National Water Model

Zhiyu Li

Brigham Young University, zyli2004@gmail.com

Daniel P. Ames

Brigham Young University, dan.ames@byu.edu

E. James Nelson

Brigham Young University, jimn@byu.edu

Michael Stealey

Renaissance Computing Institute

Follow this and additional works at: <https://scholarsarchive.byu.edu/iemssconference>

Li, Zhiyu; Ames, Daniel P.; Nelson, E. James; and Stealey, Michael, "Managing Big Data Output from the U.S. National Water Model" (2018). *International Congress on Environmental Modelling and Software*. 40. <https://scholarsarchive.byu.edu/iemssconference/2018/Stream-B/40>

This Oral Presentation (in session) is brought to you for free and open access by the Civil and Environmental Engineering at BYU ScholarsArchive. It has been accepted for inclusion in International Congress on Environmental Modelling and Software by an authorized administrator of BYU ScholarsArchive. For more information, please contact scholarsarchive@byu.edu, ellen_amatangelo@byu.edu.

Managing Big Data Output from the U.S. National Water Model

Zhiyu Li^a, Daniel P. Ames^a, E. James Nelson^a, Michael Stealey^b

*a Civil and Environmental Engineering, Brigham Young University, Provo, Utah, USA
(zhiyu.li@byu.edu, dan.ames@byu.edu, jimn@byu.edu)*

b Renaissance Computing Institute, Chapel Hill, North Carolina, USA (stealey@renci.org)

Abstract: The National Water Model (NWM) is a newly designed national-scale hydrologic model that provides streamflow forecasts for the entire continental United States. Since released in August 2016, its daily outputs are about 400GB worth of NetCDF files, which has proven challenging in terms of data storage, mobility, and accessibility. As a complement to the 48-hour storage of model forecast data on the NOAA NOMADS server, the CUAHSI HydroShare project has been archiving the NWM outputs -- extending storage duration up to 40 days. We have added new capabilities to HydroShare to meet the requirement for long-term storage of regional NWM data to support research or applications such as replication and validation, cross-model comparison, and historical data analysis (e.g. hurricane or flooding events). We will present the design and implementation of a GIS-based subsetting tool package that enables users to subset NWM data for archival and analysis purposes. The subsetted data contains model outputs that geographically fall in a user-specified watershed polygon, and the resulting files follow the original NWM NetCDF file conventions being compatible with existing NWM tools. This tool package has been integrated by the latest NWM Forecast Viewer Tethys App which is part of the HydroShare web app environment. All of its key functions have been exposed through a simple user interface and also through machine accessible application programmer interfaces (APIs).

Keywords: National Water Model; Subsetting; NetCDF; HydroShare; Tethys