

An Approach for Open and Reproducible Hydrological Modeling using Sciunit and HydroShare

Choi, YoungDon ; Goodall, Jonathan ; Ahmad, Raza ; Malik, Tanu ; Tarboton, David

It is widely acknowledged that the reproducibility of published computational results is critical to advancing science. Creating reproducible computational workflows, however, is burdensome and requires significant work to share the complete package that efficiently encapsulates all required data and software. Computational hydrology is one field that has seen rapid advancements through fast-evolving technologies for supporting increasingly complex computational hydrologic modeling and analysis. This growing model complexity, along with rapidly evolving underlying software technologies, makes the options and approaches for achieving computational reproducibility extremely challenging to settle. We argue that the technologies needed to achieve open and reproducible hydrological modeling can be grouped into three general categories: 1) data (and metadata) sharing, 2) containerizing computational environments, and 3) capturing and executing modeling workflows. While a growing set of science gateways and virtual research environments have been created to support one or more of these technologies to improve reproducibility, the integration and interoperability across all three needs are still lacking, making end-to-end systems still out of reach. The objective of this research is to advance such an end-to-end solution that can support open and reproducible hydrological modeling that effectively integrates data sharing, containerization, and workflow execution environments. Our approach emphasizes 1) well-documented modeling objects shared with meaningful metadata through the HydroShare open repository, 2) version control with efficient containerization using the Sciunit software, and 3) immutable, but flexible, computational environments to use newly developing software packages. A key to this work is advancing Sciunit, a tool for easily containerizing, sharing, and tracking deterministic computational applications, to minimally containerize reproducible hydrologic modeling workflow objects into the same container with version control capabilities. We present how to add new model input and modeling dependencies into the Sciunit container for flexibility and how to create Docker images through Sciunit containers for compatibility with popular containerization tools. In this presentation, we will emphasize both the underlying technological developments made possible through this research along with a user-centric case study showing the application of the technology from a hydrologic modeler's perspective.

Publication: vEGU21, the 23rd EGU General Assembly, held online 19-30 April, 2021, id.EGU21-13763
Pub Date: April 2021
DOI: 10.5194/egusphere-egu21-13763 
Bibcode: 2021EGUGA..2313763C

 Feedback/Corrections?

View page in SciX