Geophysical Research Abstracts
Vol. 19, EGU2017-6719, 2017
EGU General Assembly 2017

Towards Reproducibility in Computational Hydrology

Christopher Hutton (1), Thorsten Wagener (1), Jim Freer (2), Dawei Han (1), Chris Duffy (3), and Berit Arheimer (4)

(1) Department of Civil Engineering, University of Bristol, Bristol, UK, (2) School of Geographical Sciences, University of Bristol, Bristol, UK, (3) Department of Civil Engineering, Pennsylvania State University, State College, Pennsylvania, USA, (4) Swedish Meteorological and Hydrological Institute, Norrköping, Sweden

Reproducibility is a foundational principle in scientific research. The ability to independently re-run an experiment helps to verify the legitimacy of individual findings, and evolve (or reject) hypotheses and models of how environmental systems function, and move them from specific circumstances to more general theory. Yet in computational hydrology (and in environmental science more widely) the code and data that produces published results are not regularly made available, and even if they are made available, there remains a multitude of generally unreported choices that an individual scientist may have made that impact the study result. This situation strongly inhibits the ability of our community to reproduce and verify previous findings, as all the information and boundary conditions required to set up a computational experiment simply cannot be reported in an article's text alone.

In Hutton et al 2016 [1], we argue that a cultural change is required in the computational hydrological community, in order to advance and make more robust the process of knowledge creation and hypothesis testing. We need to adopt common standards and infrastructures to: (1) make code readable and re-useable; (2) create well-documented workflows that combine re-useable code together with data to enable published scientific findings to be reproduced; (3) make code and workflows available, easy to find, and easy to interpret, using code and code metadata repositories. To create change we argue for improved graduate training in these areas.

In this talk we reflect on our progress in achieving reproducible, open science in computational hydrology, which are relevant to the broader computational geoscience community. In particular, we draw on our experience in the Switch-On (EU funded) virtual water science laboratory (http://www.switch-on-vwsl.eu/participate/), which is an open platform for collaboration in hydrological experiments (e.g. [2]). While we use computational hydrology as the example application area, we believe that our conclusions are of value to the wider environmental and geoscience community as far as the use of code and models for scientific advancement is concerned.

References:
[1] Hutton, C., T. Wagener, J. Freer, D. Han, C. Duffy, and B. Arheimer (2016), Most computational hydrology is not reproducible, so is it really science?, Water Resour. Res., 52, 7548–7555, doi:10.1002/2016WR019285.
[2] Ceola, S., et al. (2015), Virtual laboratories: New opportunities for collaborative water science, Hydrol. Earth Syst. Sci. Discuss., 11(12), 13443–13478, doi:10.5194/hessd-11-13443-2014.