AGU PUBLICATIONS

Water Resources Research

COMMENTARY

10.1002/2016WR019285

Key Points:

- Articles that rely on computational work do not provide sufficient information to allow published scientific findings to be reproduced
- We argue for open reuseable code, data, and formal workflows, allowing published findings to be verified
- Reproducible computational hydrology will provide a more robust foundation for scientific advancement and policy support

Correspondence to:
C. Hutton,
chutton294@gmail.com

Citation:
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.

Received 30 MAY 2016
Accepted 23 SEP 2016
Accepted article online 28 SEP 2016
Published online 19 OCT 2016

Most computational hydrology is not reproducible, so is it really science?

Christopher Hutton[1], Thorsten Wagener[1,2], Jim Freer[2,3], Dawei Han[1], Chris Duffy[4], and Berit Arheimer[5]

[1]Department of Civil Engineering, University of Bristol, Bristol, UK, [2]Cabot Institute, Royal Fort House, University of Bristol, BS8 1UJ, Bristol, UK, [3]School of Geographical Sciences, University of Bristol, Bristol, UK, [4]Department of Civil Engineering, Pennsylvania State University, State College, Pennsylvania, USA, [5]Swedish Meteorological and Hydrological Institute, Norrköping, Sweden

Abstract Reproducibility is a foundational principle in scientific research. Yet in computational hydrology the code and data that actually produces published results are not regularly made available, inhibiting the ability of the community to reproduce and verify previous findings. In order to overcome this problem we recommend that reuseable code and formal workflows, which unambiguously reproduce published scientific results, are made available for the community alongside data, so that we can verify previous findings, and build directly from previous work. In cases where reproducing large-scale hydrologic studies is computationally very expensive and time-consuming, new processes are required to ensure scientific rigor. Such changes will strongly improve the transparency of hydrological research, and thus provide a more credible foundation for scientific advancement and policy support.

1. Introduction

Upon observing order of magnitude differences in Darcy-Weisbach Friction Factors estimated from hillslope surface properties in two previous studies [*Weltz et al*., 1992; *Abrahams et al*., 1994], *Parsons et al*. [1994] conducted additional experiments to identify factors controlling hillslope overland flow in semiarid environments, and identified that the experimental setup was the main factor controlling the difference between the previous experimental results. While exact reproducibility is impossible in open hydrological systems, attempting to reproduce the main scientific finding within an acceptable margin of error is a core principle of scientific research [*Popper*, 1959]. As illustrated, independent observation helps to verify the legitimacy of individual findings. In turn, this helps us to build upon sound observations so that we can evolve hypotheses (and models) of how catchments function [*McGlynn et al*., 2002], and move them from specific circumstances to more general theory [*Wagener et al*., 2007].

As in *Parsons et al*. [1994] attempts at reproducibility have failed in a number of disciplines, leading to increased focus on the topic in the broader scientific literature [*Begley and Ellis*, 2012; *Prinz et al*., 2011; *Ioannidis et al*., 2001; *Nosek*, 2012]. Such failures have occurred not just because of differences in experimental setup, but because of scientific misconduct [*Yong*, 2012; *Collins and Tabak*, 2014; *Fang et al*., 2012], poor application of statistics to achieve apparent significant results [*Ioannidis*, 2005; *Hutton*, 2014], and importantly, insufficient reporting of methodologies and data quality in journals to enable reproducibility to be assessed by the community. An oft-cited underlying reason for such failures is the present reward system in scientific publication, which prioritizes the publication of innovative, and seemingly statistically significant results over the publication of both null results [*Franco et al*., 2014; *Jennions and Møller*, 2002; cf *Freer et al*., 2003], and reproduced experiments. Such a system provides few incentives to adopt open science practices that support and enable verification [*Nosek et al*., 2015].

The prominence of computational research across scientific disciplines—from big data analysis in genomic research to computational modeling in climate science—has brought increased focus on the reproducibility issue. This is because the full code and workflow used to produce published scientific findings is typically not made available, thus inhibiting attempts to verify the provenance of published results [*Buckheit and Donoho*, 1995; *Mesirov*, 2010]. Given the extent to which this lack of transparency is considered a problem for reproducibility more broadly in the scientific literature [*Donoho et al*., 2009], to what extent is reproducibility, or a lack thereof, also a problem in computational hydrology? Computational analysis has grown rapidly in hydrology over the past 30 years, transforming the process of scientific discovery. While code is most obviously used for hydrological modeling [e.g., *Clark et al*., 2008; *Wrede et al*., 2014; *Duan et al*., 2006], some form of code is used to produce the vast majority of hydrological research papers, from data processing and quality analysis [*Teegavarapu*, 2009; *Mcmillan et al*., 2012; *Coxon et al*., 2015], regionalization and large-scale statistical analysis across catchments [*Blöschl et al*., 2013; *Berghuijs et al*., 2016], all the way to figure preparation. However, as in other disciplines, the full code that produces presented results is typically not made available alongside the publication, which inhibits attempts to reproduce published findings.

In order to advance scientific progress in hydrology, reproducibility is required in computational hydrology for several key reasons. First, the reliability of scientific computer code is often unclear. From our own experience, it is often very difficult to spot errors unless they manifest themselves in very obvious errors in model outputs. Thus, code needs to be transparent to allow the legitimacy of published results to be verified. Second, the complexity of many hydrologic models and data analysis codes used today makes it simply infeasible to report all settings that can be adjusted (e.g., initial conditions and parameters) in publications—a point recognized recently in a joint editorial published in five hydrology journals [*Blöschl et al*., 2014]. Transparency across hydrology is especially important given research builds on previous research. For example, being able to evaluate how "tidied up" data sets have been created by explicitly showing all of the assumptions made will lead to benefits in interpreting where and why subsequent models that are built upon such data sets fail. Finally, the complexity and diversity of catchment systems means that we need to be able to reproduce exact methodologies applied in specific settings more broadly across a range of catchment environments, so that we can robustly evaluate competing hypotheses of hydrologic behavior across scales and locations [*Clark et al*., 2016]. Our current inability to achieve this hinders both the ability of the broader community to learn from, and build on, previous work, and importantly, verify previous findings. So what material should be provided, and therefore what is required to reproduce computational hydrology?

The necessary information that leads to and therefore documents the provenance of the final research paper has been termed the research compendium [*Gentleman and Lang*, 2007]. In the context of computational hydrology, this includes the original data used, all analysis/modeling code, and the workflow that ties together the code and data to produce the published results. Although these components are not routinely published alongside journal articles, current practices in hydrology do facilitate reproducibility to varying extents. For example, initiatives are relatively well developed in hydrology for opening up and sharing data from individual catchments and cross-catchment data sets [*McKee and Druliner*, 1998; *Renard et al*., 2008; *Kirby et al*., 1991; *Newman et al*., 2015; *Duan et al*., 2006], including (quite recently) the development of infrastructures and standards for sharing open water data [*Emmett et al*., 2014; *Leonard and Duffy*, 2013; *Tarboton et al*., 2009; *Tarboton et al*., 2014]. In addition, different code packages have been made available by developers. Prominent examples include the hydrologic models such as Topmodel [*Beven and Kirkby*, 1979], VIC [*Wood et al*., 1992], FUSE [*Clark et al*., 2008], HYPE [*Lindström et al*., 2010], open-source groundwater models including MODFLOW [*Harbaugh*, 2005] and PFLOTRAN, and codes linked to modeling, including optimization/uncertainty algorithms such as SCE [*Duan et al*., 1993], SCEM [*Vrugt et al*., 2003] or GLUE [*Beven and Binley*, 1992]. By being made open, such code has helped spread new ideas and concepts to advance hydrology, and made reproducing each-others' work easier. However, while sharing data and code are important first steps, sharing alone does not provide the critical detail on implementation contained within a workflow that is required to reproduce published results.

2. Towards Reproducible Computational Hydrology

We argue that in order to advance and make more robust the process of knowledge creation and hypothesis testing within the computational hydrological community, we need to adopt common standards and infrastructures to: (1) make code readable and reuseable; (2) create well-documented workflows that combine reuseable code together with data to enable published scientific findings to be reproduced; (3) make code and workflows available and easy to find through the use of code repositories and creation of code metadata; (4) use unique persistent identifiers (e.g., DOIs) to reference reuseable code and workflows, thereby clearly showing the provenance of published scientific findings (Figure 1).

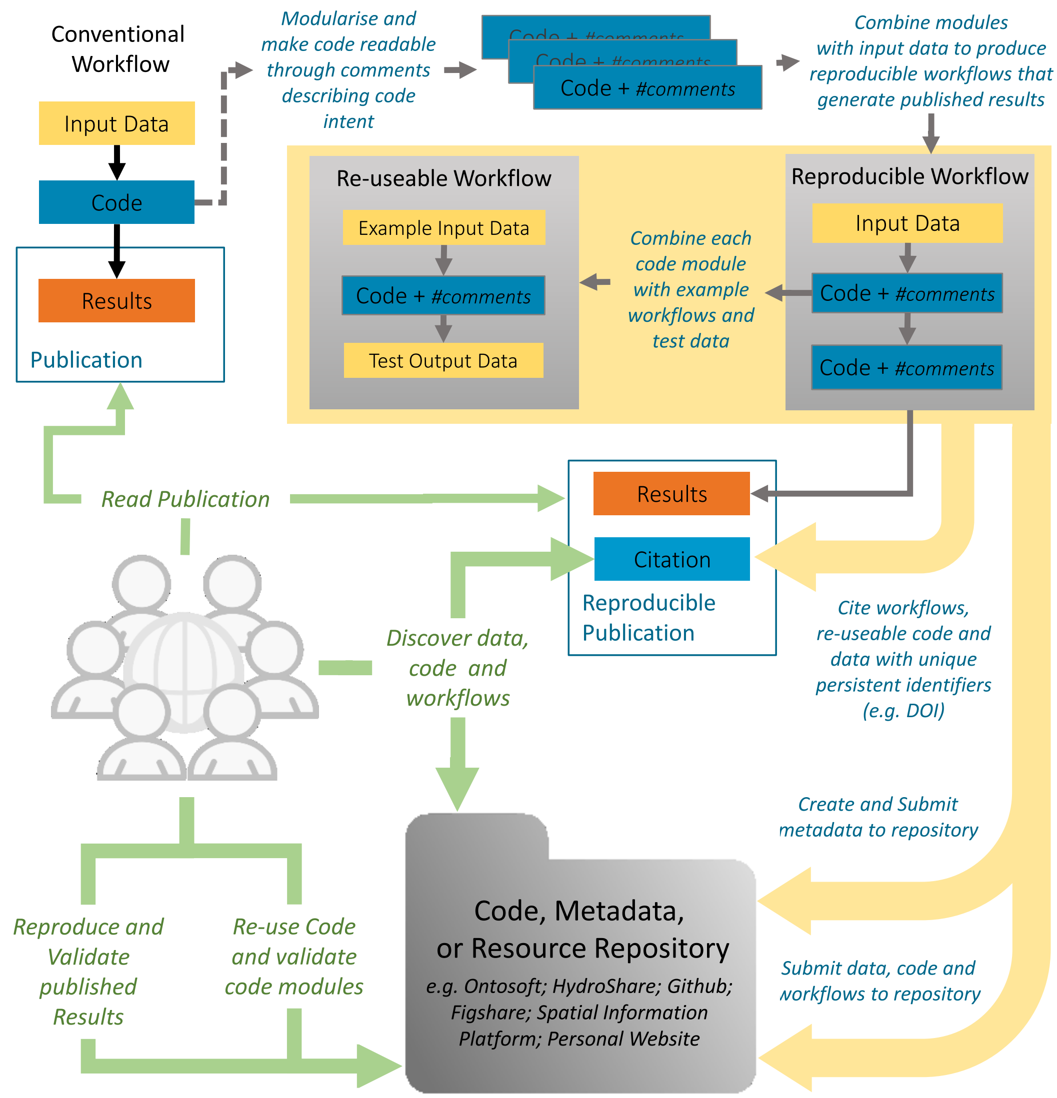

Figure 1. Schematic figure of steps required leading to reproducible and reuseable hydrological publications.

The first step toward more open, reproducible science is to adopt common standards that facilitate code readability and reuse. As most researchers in hydrology are scientists first, programmers second, setting high standards for code reuse may be counterproductive to broad adoption of reproducible practices. Yet long, poorly documented scripts are not reuseable, and certainly difficult to reproduce if their ability to do the intended job cannot be verified. As a minimum standard, we therefore recommend that code should come with an example workflow, as commonly adopted [e.g., *Pianosi et al*., 2015], and where possible, also packaged with input and output data to provide a means to ensure correct implementation of a method prior to application. Implementing code correctly however is not enough to make it reuseable; sufficient information is required to understand what the code does, and to be reproducible, whether it does this correctly. Therefore, code should be modularized into functions and classes that may be reuseable by the wider community, with comments that do not repeat the code, but explain at a higher level of abstraction what individual blocks within modular code are trying to do [*McConnell*, 2004]. Such readable code allows the broader community to verify code intent.

The second key requirement to reproduce published scientific results is a well-documented workflow, or protocol that combines reuseable code together with data to enable published scientific findings to be reproduced. Such workflows may take the form of code scripts themselves [e.g., *Ceola et al*., 2015; *Pianosi et al*., 2015], or when multiple programming environments/research partners are involved, schematic workflows that illustrate how individual scripts and intermediary results lead to the generation of the final, published paper. Regardless of the specific structure, or software/workflow management system used, we argue that the key requirement of such a workflow is that it clearly specifies all potential degrees of freedom, and therefore unambiguously ties together the component reuseable code and data to document the provenance of the published scientific results. For example, *Ceola et al*. [2015] identified the importance of a well-documented protocol to ensure correct execution, and avoid ambiguity in the interpretation of results, when five research groups attempted to reproduce the same hydrological model calibration experiment.

Third, code and code metadata need to be made open and available to allow others to reuse and reproduce scientific results. Numerous code and resource repositories exist to facilitate sharing of research outputs, such as Github, Zenodo, Figshare, the EU SWITCH-ON Virtual Water-Science Laboratory (www.water-switch-on.eu), and the US CUAHSI initiative Hydroshare, specifically designed for sharing hydrological data and models to serve the hydrological community [*Horsburgh et al*., 2015; *Tarboton et al*., 2014]. The development of metadata standards for water data is a key factor that has allowed data to be found, correctly interpreted and reused by the broader community [*Maidment*, 2008; *Taylor*, 2012]. In the same vein, we argue that in order to facilitate first the discovery, and second the reuse of disparate hydrological code across the web, the development and adoption of similar metadata standards are required. *Gil et al*. [2015], for example, have developed OntoSoft for the geoscience community; a metadata repository and ontology to describe software metadata. The development of code metadata, and consistent use of such a repository, while more challenging than development of metadata standards for data, will greatly facilitate the process of code identification and reuse, and through broad community engagement, lead the way toward the development of more formal ontologies for specific components of hydrological software, which will greatly improve model interoperability [see *Elag and Goodall*, 2013].

Finally, we recommend that reuseable code and reproducible code (workflows) need to be cited in research papers using unique persistent identifiers (e.g., DOIs) to clearly link published results to the code used to generate them, thereby documenting their provenance [*Horsburgh et al*., 2015]. Such DOIs should be specific to the exact code version used in generating the results. Appropriate citation in methodologies and results sections of papers will allow others to both reuse code and reproduce experimental results. While code may be included as supporting information in research articles, persistent links to repositories provide an open access approach that exploits existing infrastructures specifically designed for sharing research outputs. Furthermore, such an approach demands little from publishers other than adopting standards for code citation.

3. Changing the Research Culture

Making one's code reuseable in the first instance, then reproducible, undoubtedly requires extra effort. This is notwithstanding the effort to reproduce someone else's work, with little reward in the current system of publication to reproduce, and therefore validate, either positively or negatively, a prior result. Thus, it is a perfectly valid question to ask: why go to the effort? Within the current system of academic reward through citation [*Koutsoyiannis et al*., 2016], making code available and reuseable reduces the barriers to the adoption of developed methods, which as considered above, is more likely to lead to further citation and greater impact in the community. Furthermore, making code reuseable is beneficial for our own work efficiency [*Donoho et al*., 2009]. Across hydrology, much duplicated code is likely to be written for common tasks that are not deemed worthy of publication. However, if open, reuseable practices are adopted by the broader community to make all code open and citable, this would reduce the amount of individual code to be written, and lead to improved efficiency at a community level. In addition, this would allow researchers to gain credit for all of their research outputs, not simply the final publication. The key reason we recommend making code reuseable, however, is that this would allow a process of natural selection to occur at the community level, where freely chosen code that is assessed to be most fit-for-puropse through reuse and unittesting can form the individual building blocks of larger "off-spring" scripts/workflows. Verification of these individual code building blocks, potentially by many users in the community, means assessing the reproducibilty and provenance of derived results becomes much easier.

As has guided our recommendations we make above, there is wide recognition that gradual steps are required to change a deeply engrained research culture that does not currently require reproducibility [*Bailey et al.*, 2016; *Peng*, 2011; *Koutsoyiannis et al.*, 2016]. A key step to change this culture is to ensure that computational science training (e.g., http://software-carpentry.org) is properly embedded within hydrological science curriculums, so that future generations of hydrologists have the skills to build readable, version controlled and unit-tested software [*McConnell*, 2004], allowing them to engage more fully in an open scientific community by reproducing and reusing each other's research outputs. Thus, instead of seeing the need to make their work reproducibile as an inconvenient after-thought, it will be an integral part of their research process. Engaging with advances in the related disciplines of computational science and hydroinformatics through such training will help ensure future hydrologists, and in turn the science they produce, benefits from modern computational methods. To facilitate this training, Data and Modeling Driven Cybereducation (DMDC) methods [*Merwade and Ruddell*, 2012], and educational web-based tools [e.g., *Wagener and McIntyre*, 2007; *Habib et al.*, 2012], need to come to the forefront and ultimately form part of a holistic approach to hydrology education that considers future challenges and opportunities for hydrologists [*Sanchez et al.*, 2016].

Journals and funding bodies clearly have a role to play in facilitating the change to more open science. Some publishers and hydrological journals are revising their policies to encourage authors to make data and computer codes available to readers [*Blöschl et al.*, 2014], notably *Vadose Zone Journal* with the launch of a reproducible research program, which will verify that code is technically sound and can be used to reproduce the key results of the paper [*Skaggs et al.*, 2015]. AGU Publications also encourages references to data and software to find source material, facilitating transparency and recognition [*Hanson and Van Der Hilst*, 2014]. Other journals go further. *Science*, for example, states that all codes used in creation and analysis of data must be available to readers [*Sciencemag.org*, 2016]. *Nosek et al.* [2015] have developed guidelines to facilitate gradual adoption of open practices by journals. Funding guidelines for science funding bodies in the U.S. (NSF) and UK (NERC) have moved toward more open science practices, and both require that data and other research materials are made open [*NERC*, 2016; *NSF*, 2016]. NERCs open data policy, for example, is designed to "support the integrity, transparency, and openness of the research it supports." However, despite the intent, these guidelines currently fall short of software sharing, which is only encouraged by the NSF. Finally, changes such as the replacement of the "Publications" section in the NSF biosketch format for grant applications with a "Products" section to recongize other research outputs like software provides important additional incentives for open science practice.

While reproducibility is more achievable in smaller-scale studies, there are key technical challenges to address in making computational workflows in hydrology reproducible as the scale of application increases in terms of modeling domain, data, and computational requirements, large legacy codes authored by large, diverse scientific groups, and large user communities. Modeling large domains with complex models, or many catchments with complex algorithms are increasingly common [e.g., *Kollat et al.*, 2012; *Pechlivanidis and Arheimer*, 2015], yet such studies are computationally demanding, and one cannot currently expect these to be reproduced given the resources it would require, in particular by reviewers. We therefore need to improve our ability to reproduce larger-scale studies, and when not possible, identify formal processes that nonetheless ensure that such studies are scientifically verifiable.

Ongoing research in hydroinformatics is attempting to tackle these reproducibility issues, including development of workflows for large-scale data processing [*Essawy et al.*, 2016; *Billah et al.*, 2016], and the work undertaken over the past decade to develop the open source model RAPID [*David et al.*, 2016]. In addition, formal processes like benchmark comparison tests [e.g., *Maxwell et al.*, 2014] may help to provide confidence in key complex codes that are difficult to transfer between research groups. Other scientific communties have moved toward sharing complex codes between many research groups, including modelling projects in meteorology (HIRLAM) and oceanography (NEMO), which is beneficial for code development. The idea to establish such a community model has been discussed in hydrological sciences [*Weiler and Beven*, 2015]. Improved training in computational science, and open science practices considered above, will help in building large and inter-operable model codes across research groups, which can help in providing independent verification of model components.

In a competitive research climate, funding bodies in the UK and Europe are increasingly emphasising the importance of impact generated from science spending. Coupled with events such as the droughts in California, and persistent flooding in the UK over recent years, this change in emphasis highlights the increasing role that hydrological scientists have to play in informing public policy and public understanding of hydrological risks. The need for openness and transparency in scientific research was highlighted by the so-called *climategate scandal*, because of the potential loss of trust in climate scientists that resulted [*Leiserowitz et al*., 2012]. Thus, to play a credible role in informing public policy, trust in the hydrological science community is essential, and is built on transparency. Transparent, reproducible computational hydrology will then provide a solid foundation to address the more difficult problem of inference and reproducibility in open systems to forward scientific understanding; progress in which requires both innovation and verification.

4. Conclusions

Reproducibility is a foundational principle in scientific research. Yet in hydrology, the code and data that actually produces published results are not regularly made available, which strongly inhibits reproducibility. This situation hinders both the ability of the broader community to learn from, and build on, previous work, and importantly, verify previous findings. To help move toward reproducible computational hydrology, we recommend the following:

1. code needs to be made readable and reuseable for the community;
2. workflows that tie together data and reuseable code need to be created to document, unambiguously, the full provenance of published scientific results;
3. reuseable code and workflows need to be made available and easy to find through consistent use of repositories and creation of code metadata;
4. reuseable and reproducible code needs to be cited in publications using unique persistent identifiers (e.g., DOIs) to clearly show the provenance of published scientific findings; and
5. new procedures need to be developed that ensure scientific rigor in circumstances where reproducing large-scale studies is computationally very expensive and time consuming.

Making code reuseable is more likely to lead to citation and reuse of an individual's work, which provides an incentive within the current publication system that can be built upon to move toward reproducibility, and gain efficiencies across the hydrology community to advance scientific understanding across catchments. Ultimately however, a collective will is required across the community to adequately address the larger technical, scientific, and cultural challenges that need to be solved, including real buy-in from journals and funding bodies, and training of young scientists to adopt reproducible practices. To allow hydrology to play a credible role in informing public policy, trust in the hydrological science community is essential, and is built on the transparency that will result. Our view is that reproducible computational hydrology will provide this transparency.

Acknowledgments
This study was performed within the EU FP7-funded project SWITCH-ON (grant agreement 603587), which explores the potential of Open Data and promote Open Science to become a vehicle for new forms of collaborative research in hydrology. The authors would like to thank Glenn Hammond and Jon Goodall for their insightful comments and suggestions that have contributed to improve this paper. No data were used in producing this manuscript.

References

Abrahams, A. D., A. J. Parsons, and J. Wainwright (1994), Resistance to overland-flow on semiarid grassland and shrubland hillslopes, Walnut Gulch, Southern Arizona, *J. Hydrol.*, *156*(1-4), 431–446.

Bailey, D. H., J. M. Borwein, and V. Stodden (2016), Facilitating reproducibility in scientific computing: Principles and practice, in *Reproducibility: Principles, Problems, Practices*, edited by H. Atmanspacher and S. Maasen, pp. 205–232, John Wiley and Sons, N. J.

Begley, C. G., and L. M. Ellis (2012), Drug development: Raise standards for preclinical cancer research, *Nature*, *483*, 531–533.

Berghuijs, W. R., R. A. Woods, C. J. Hutton, and M. Sivapalan (2016), Dominant flood generating mechanisms across the United States, *Geophys. Res. Lett.*, *43*, 4382–4390, doi:10.1002/2016GL068070.

Beven, K., and A. M. Binley (1992), The future of distributed models: Model calibration and uncertainty estimation, *Hydrol. Processes*, *6*, 279–298.

Beven, K., and M. J. Kirkby (1979), A physically based variable contributing area model of basin hydrology, *Hydrol. Sci. Bull.*, *24*(1), 43–69.

Billah, M. M., J. L. Goodall, U. Narayan, B. T. Essawy, V. Lakshmi, A. Rajasekar, and R. W. Moore (2016), Using a data grid to automate data preparation pipelines required for regional-scale hydrologic *modelling, Environ. Model. Software*, *78*, 31–39.

Blöschl, G., M. Sivapalan, T. Wagener, A. Viglione, and H. Savenije (2013), *Runoff Prediction in Ungauged Basins: Synthesis Across Processes, Places and Scales*, Cambridge University Press, Cambridge, U. K.

Blöschl, G., A. Bardossy, D. Koutsoyiannis, Z. Kundzewicz, I. Littlewood, A. Montanari, and H. Savenije (2014), On the future of journal publications in hydrology, *Water Resour*. Bull., *50*, 2795–2797, doi:10.1002/2014WR015613.

Buckheit, J., and D. Donoho (1995), WaveLab and Reproducible Research, in *Wavelets and Statistics*, vol. 103, pp. 55–81, Springer, N. Y.

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.
Clark, M. P., A. G. Slater, D. E. Rupp, R. A. Woods, J. A. Vrugt, H. V. Gupta, T. Wagener, and L. E. Hay (2008), Framework for Understanding Structural Errors (FUSE): A modular framework to diagnose differences between hydrological models, *Water Resour. Res.*, *44*, W00B02, doi:10.1029/2007WR006735.
Clark, M. P., et al. (2016), Improving the theoretical underpinnings of process-based hydrologic models, *Water Resour. Res.*, *52*, 2350–2365, doi:10.1002/2015WR017910.
Collins, F. S., and L. A. Tabak (2014), Policy: NIH plans to enhance reproducibility, *Nature*, *505*, 612–613, doi:10.1038/505612a.
Coxon, G., J. Freer, I. K. Westerberg, T. Wagener, R. Woods, and P. J. Smith (2015), A novel framework for discharge uncertainty quantification applied to 500 UK gauging stations, *Water Resour. Res.*, *51*, 5531–5546, doi:10.1002/2014WR016532.
David, C. H., J. S. Famiglietti, Z.-L. Yang, F. Habets, and D. R. Maidment (2016), A decade of RAPID—Reflections on the development of an open source geoscience code, *Earth Space Sci.*, *3*, 226–244, doi:10.1002/2015EA000142.
Donoho, D. L., A. Maleki, M. Shahram, I. U. Rahman, and V. Stodden (2009), Reproducible research in computational harmonic analysis, *Comput. Sci. Eng.*, *11*, 8–18, doi:10.1109/MCSE.2009.15.
Duan, Q., et al. (2006), Model Parameter Estimation Experiment (MOPEX): An overview of science strategy and major results from the second and third workshops, *J. Hydrol.*, *320*, 3–17.
Duan, Q. Y., V. K. Gupta, and S. Sorooshian (1993), Shuffled complex evolution approach for effective and efficient global minimization, *J. Optim. Theory Appl.*, *76*(3), 501–521, doi:10.1007/BF00939380.
Elag, M., and J. L. Goodall (2013), An ontology for component based models of water resource systems, *Water Resour. Res.*, *49*, 5077–5091, doi:10.1002/wrcr.20401
Emmett, B., et al. (2014), Heads in the clouds: Innovation in data and model dissemination, *Int. Innovation*, *141*, 82–85.
Essawy, B. T., J. L. Goodall, H. Xu, A. Rajasekar, J. D. Myers, T. A. Kugler, M. M. Billah, M. C. Whitton, and R. W. Moore (2016), Server_side workflow execution using data grid technology for reproducible analyses of data-intensive hydrologic systems, *Earth Space Sci.*, *3*(4), 63–175.
Fang, F. C., R. G. Steen, and A. Casadevall (2012), Misconduct accounts for the majority of retracted scientific publications, Proc. *Natl. Acad. Sci. U. S. A.*, *109*, 17,028–17,033, doi:10.1073/pnas.1212247109.
Franco, A., N. Malhotra, and G. Simonovits (2014), Publication bias in the social sciences: Unlocking the file drawer, *Science*, *345*(6203), 1502–1505, doi:10.1126/science.1255484.
Freer, J., K. Beven, and N. E. Peters (2003), Multivariate seasonal period model rejection within the generalised likelihood uncertainty estimation procedure, in *Calibration of Watershed Models, Water Sci. Appl. Ser.*, edited by Q. Duan et al., pp. 69–88, AGU, Washington, D. C.
Gentleman, R., and T. Lang (2007), Statistical Analyses and Reproducible Research, *J. Comp. Graph. Stat.*, *16*, 1–23.
Gil, Y., V. Ratnakar, and D. Garijo (2015), OntoSoft: Capturing scientific software metadata, in *Proceedings of the 8th ACM International Conference on Knowledge Capture*, pp. 1–4, Palisades, ACM, N. Y.
Habib, E., Y. Ma, D. Williams, H. O. Sharif, and F. Hossain (2012), HydroViz: Design and evaluation of a web-based tool for improving hydrology education, *Hydrol. Earth Syst. Sci.*, *16*, 3767–3781.
Hanson, B., and R. Van Der Hilst (2014), AGU's data policy: History and context, *Eos Trans. AGU*, *95*(37), 337, doi:10.1002/2014EO370008.
Harbaugh, A. W. (2005), *MODFLOW-2005, the US Geological Survey Modular Ground-Water Model: The Ground-Water Flow Process*, pp. 6-A16, U.S. Dep. of the Interior, U.S. Geol. Surv., Reston, Va.
Horsburgh, J. S., M. M. Morsy, A. M. Castronova, J. L. Goodall, T. Gan, H. Yi, M. J. Stealey, and D. G. Tarboton (2015), Hydroshare: Sharing diverse environmental data types and models as social objects with application to the hydrology domain, *J. Am. Water Resour. Assoc.*, *52*(4), 873–889, doi:10.1111/1752-1688.12363.
Hutton, C. J. (2014), How significant ($p < 0.05$) is geomorphic research?, *Earth Surf. Process. Landforms*, *39*(11), 1559–1562, doi:10.1002/esp.3618.
Ioannidis, J. P., E. E. Ntzani, T. A. Trikalinos, and D. G. Contopoulos-Ioannidis (2001), Replication validity of genetic association studies, *Nat. Genetics*, *29*, 306–309, doi:10.1038/ng749.
Ioannidis, J. P. A. (2005), Why most published research findings are false, *PLoS Med.*, *2*, 0696–0701, doi:10.1371/journal.pmed.0020124.
Jennions, M. D., and A. P. Møller (2002), Relationships fade with time: A meta-analysis of temporal trends in publication in ecology and evolution, *Proc. Biol. Sci.*, *269*, 43–48, doi:10.1098/rspb.2001.1832.
Kirby, C., M. D. Newson, and K. Gilman (1991), Plynlimon research: The first two decades, *IH Rep. 109*, Institute of Hydrology, Wallingford, U. K.
Kollat, J. B., P. M. Reed, and T. Wagener (2012), When are multiobjective calibration trade-offs in hydrologic models meaningful?, *Water Resour. Res.*, *48*, W03520, doi:10.1029/2011WR011534.
Koutsoyiannis, D., G. Blöschl, A. Bardossy, C. Cudennec, D. Hughes, A. Montanari, I. Neuweiler, and H. Savenije (2016), Joint editorial—Fostering innovation and improving impact assessment for journal publications in hydrology, *Hydrol. Sci. J.*, *61*(7), 1170–1173.
Leiserowitz, A. A., E. W. Maibach, C. Roser-Renouf, N. Smith, and E. Dawson (2012), Climategate, public opinion, and the loss of trust, *Am. Behavioral Sci.*, *57*(6), 818–837, doi:10.1177/0002764212458272.
Leonard, L., and C. J. Duffy (2013), Essential terrestrial variable data workflows for distributed water resources modeling, *Environ. Model. Software*, *50*, 85–96, doi:10.1016/j.envsoft.2013.09.003.
Lindström, G., C. Pers, J. Rosberg, J. Strömqvist, and B. Arheimer (2010), Development and testing of the HYPE (Hydrological Predictions for the Environment) water quality model for different spatial scales, Hydrol. *Res.*, *41*, 295–319, doi:10.2166/nh.2010.007.
Maidment, D. (2008), Bringing water data together, *J. Water Resour. Plann. Manage.*, *134*(2), 95–96.
Maxwell, R. M., et al. (2014), Surface-subsurface model intercomparison: A first set of benchmark results to diagnose integrated hydrology and feedbacks, *Water Resour. Res.*, *50*, 1531–1549, doi:10.1002/2013WR013725.
McConnell, S. (2004), *Code Complete*, 2nd ed., 897 pp., Interfaces, Providence, R. I.
McGlynn, B. L., J. J. McDonnel, and D. D. Brammer (2002), A review of the evolving perceptual model of hillslope flowpaths at the Maimai catchments, New Zealand, *J. Hydrol.*, *257*, 1–26, doi:10.1016/S0022-1694(01)00559-5.
McKee, A., and P. Druliner (1998), *H.J. Andrews Experimental Forest Brochure*, U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station.
Mcmillan, H., T. Krueger, and J. Freer (2012), Benchmarking observational uncertainties for hydrology: Rainfall, river discharge and water quality, *Hydrol. Processes*, *26*, 4078–4111, doi:10.1002/hyp.9384.
Merwade, V., and B. L. Ruddell (2012), Moving university hydrology education forward with community-based geoinformatics, data and modeling resources, *Hydrol. Earth Syst. Sci.*, *16*, 2393–2404.
Mesirov, J. P. (2010), Accessible reproducible research, *Science*, *327*, 415–416.
NERC (2016), Nerc Data Policy. [Available at http://www.nerc.ac.uk/research/sites/data/policy/data-policy/.]

Newman, A. J., et al. (2015), Development of a large-sample watershed-scale hydrometeorological data set for the contiguous USA: Data set characteristics and assessment of regional variability in hydrologic model performance, *Hydrol. Earth Syst. Sci.*, *19*, 209–223, doi:10.5194/hess-19-209-2015.

Nosek, B. A. (2012), An open, large-scale, collaborative effort to estimate the reproducibility of psychological science, *Perspect. Psychol. Sci.*, *7*(6), 657–660, doi:10.1177/1745691612462588.

Nosek, B. A., et al. (2015), Promoting an open research culture, *Science*, *348*(6242), 1422–1425, doi:10.1126/science.aab2374.

NSF (2016), National Science Foundation: Grant General Conditions. [Available at http://www.nsf.gov/pubs/policydocs/gc1/jan16.pdf.]

Parsons, A. J., A. D. Abrahams, and J. Wainwright (1994), On determining resistance to interrill overland-flow, *Water Resour. Res.*, *30*(12), 3515–3521.

Pechlivanidis, I. G., and B. Arheimer (2015), Large-scale hydrological modelling by using modified PUB recommendations: The India-HYPE case, *Hydrol. Earth Syst. Sci.*, *19*, 4559–4579, doi:10.5194/hess-19-4559-2015.

Peng, R. D. (2011), Reproducible research in computational science, *Science*, *334*, 1226–1227, doi:10.1126/science.1213847.

Pianosi, F., F. Sarrazin, and T. Wagener (2015), A Matlab toolbox for global sensitivity analysis, *Environ. Model. Software*, *70*, 80–85, doi:10.1016/j.envsoft.2015.04.009.

Popper, K. R. (1959), *The Logic of Scientific Discovery (Routledge Classics)*, Routledge, N. Y.

Prinz, F., T. Schlange, and K. Asadullah (2011), Believe it or not: How much can we rely on published data on potential drug targets?, *Nat. Rev. Drug Discovery*, *10*, 712, doi:10.1038/nrd3439-c1.

Renard, K. G., M. H. Nichols, D. A. Woolhiser, and H. B. Osborn (2008), A brief background on the U.S. Department of Agriculture—Agricultural Research Service Walnut Gulch Experimental Watershed, *Water Resour. Res.*, *44*, W05S02, doi:10.1029/2006WR005691.

Sanchez, C. A., B. L. Ruddell, R. Schiesser, and V. Merwade (2016), Enhancing the T-shaped learning profile when teaching hydrology using data, modeling, and visualization activities, *Hydrol. Earth Syst. Sci.*, *20*, 1289–1299, doi:10.5194/hess-20-1289-2016.

Sciencemag.org (2016), Science: Editorial Policies. [Available at http://www.sciencemag.org/authors/science-editorial-policies.]

Skaggs, T. H., M. H. Young, and J. A. Vrugt (2015), Reproducible research in vadose zone sciences, Vadose Zone *J.*, *14*(10), doi:10.2136/vzj2015.06.0088.

Tarboton, D. G., J. S. Horsburgh, D. R. Maidment, T. Whiteaker, I. Zaslavsky, M. Piasecki, J. Goodall, D. Valentine, and T. Whitenack, (2009), Development of a community hydrologic information system, in *18th World IMACS Congress and MODSIM09 International Congress on Modelling and Simulation*, pp. 988–994, Modell. and Simul. Soc. of Aust. and N. Z.

Tarboton, D. G., et al. (2014), HydroShare: Advancing collaboration through hydrologic data and model sharing, in *Proceedings of the 7th International Congress on Environmental Modelling and Software*, pp. 978–988, Int. Environ. Modell. and Software Soc., San Diego, Calif.

Teegavarapu, R. S. V. (2009), Estimation of missing precipitation records integrating surface interpolation techniques and spatio-temporal association rules, *J. Hydroinf.*, *11*, 133–146, doi:10.2166/hydro.2009.009.

Valentine, D., P. Taylor, and I. Zaslavsky (2012), WaterML, an information standard for the exchange of in-situ hydrological observations, Abstract presented at EGU General Assembly, EGU, Vienna.

Vrugt, J. A., H. V. Gupta, W. Bouten, and S. Sorooshian (2003), A Shuffled Complex Evolution Metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters, *Water Resour. Res.*, *39*(8), 1201, doi:10.1029/2002WR001642.

Wagener, T., and N. McIntyre (2007), Tools for teaching hydrological and environmental modeling, *Comput. Educ. J.*, *XVII*(3), 16–26.

Wagener, T., M. Sivapalan, P. Troch, and R. Woods (2007), Catchment classification and hydrologic similarity, *Geogr. Compass*, *1*, 1–31, doi:10.1111/j.1749-8198.2007.00039.x.

Weiler, M., and K. Beven (2015), Do we need a Community Hydrological Model?, *Water Resour. Res.*, *51*, 7777–7784, doi:10.1002/2014WR016731.

Weltz, M. A., A. A.B., and L. J. Lane (1992), Hydraulic roughness coefficients for native rangelands, *J. Irrig. Drain. Eng.*, *118*, 776–790.

Wood, E. F., D. P. Lettenmaier, and V. G. Zartarian (1992), A land-surface hydrology parameterization with subgrid variability for general circulation models, *J. Geophys. Res.*, *97*(D3), 2717–2728, doi:10.1029/91JD01786.

Wrede, S., F. Fenicia, N. Martínez-Carreras, J. Juilleret, C. Hissler, A. Krein, H. H. G. Savenije, S. Uhlenbrook, D. Kavetski, and L. Pfister (2014), Towards more systematic perceptual model development: A case study using 3 Luxembourgish catchments, *Hydrol. Processes.*, *29*, 2731–2750, doi:10.1002/hyp.10393.

Yong, E. (2012), Replication studies: Bad copy, *Nature*, *485*, 298–300, doi:10.1038/485298a.