Developing observational methods to drive future hydrological science: can we make a start as a community?

Received: 4 October 2019 | Accepted: 4 October 2019
DOI: 10.1002/hyp.13622

INVITED COMMENTARY

Developing observational methods to drive future hydrological science: Can we make a start as a community?

Keith Beven[1] | Anita Asadullah[2] | Paul Bates[3] | Eleanor Blyth[4] | Nick Chappell[1] | Stewart Child[5] | Hannah Cloke[6] | Simon Dadson[4,7] | Nick Everard[2] | Hayley J. Fowler[8] | Jim Freer[3] | David M. Hannah[9] | Kate Heppell[10] | Joseph Holden[11] | Rob Lamb[12] | Huw Lewis[13] | Gerald Morgan[14] | Louise Parry[15] | Thorsten Wagener[16]

[1]Lancaster Environment Centre, Lancaster University, Lancaster, UK
[2]UK Environment Agency, Bristol, UK
[3]School of Geography, Bristol University, Bristol, UK
[4]Centre for Ecology and Hydrology, Wallingford, Wallingford, UK
[5]Hydro-Logic Services (International) Ltd., Reading, Reading, UK
[6]Department of Geography, University of Reading, Reading, UK
[7]School of Geography and the Environment, University of Oxford, Oxford, UK
[8]School of Engineering, Newcastle University, Newcastle upon Tyne, UK
[9]School of Geography, Earth & Environmental Sciences, University of Birmingham, Birmingham, UK
[10]School of Geography, Queen Mary University of London, London, UK
[11]water@leeds, School of Geography, University of Leeds, Leeds, UK
[12]JBA Trust, Skipton, Skipton, UK
[13]MetOffice, Exeter, Exeter, UK
[14]Edenvale Young Associates Ltd., Bristol, Bristol, UK
[15]Arup, Bristol, Bristol, UK
[16]Department of Civil Engineering, Bristol University, Bristol, UK

Correspondence
Keith Beven, Lancaster Environment Centre, Lancaster University, Lancaster, UK.
Email: k.beven@lancaster.ac.uk

Hydrology is still, and for good reasons, an inexact science (for a recent discussion, see Beven, 2019a), even if evolving hydrological understanding has provided a basis for improved water management for at least the last three millennia. The limitations of that understanding have, however, become much more apparent and important in the last century as the pressures of increasing populations, and the anthropogenic impacts on catchment forcing and responses, have intensified (see Abbott et al., 2019; Montanari et al., 2013; Sivapalan, Savenije, & Blöschl, 2012; Srinivasan et al., 2017; Wagener et al., 2010; Wilby, 2019). At the same time, the sophistication of hydrological analyses and models has been developing rapidly, often driven more by the availability of computational power and geographical data sets than any real increases in understanding of hydrological processes. This sophistication has created an illusion of real progress, but a case can be made that we are still rather muddling along, limited by the significant uncertainties in hydrological observations, knowledge of catchment characteristics, and related gaps in conceptual understanding, particularly of the subsurface. These knowledge gaps are illustrated by the fact that for many catchments, we cannot close the water balance without significant uncertainty (e.g., Beven, 2019a;

Schaller & Fan, 2009), uncertainty that is often neglected in evaluating models for practical applications. This lack of water balance closure can also result from a lack of information about the influence of water management on the water balance. We have seen improvements since the first crude U.K. water balance estimates of John Dalton (1791), but there remain important uncertainties in the estimates of every term in the water balance equation: precipitation inputs (especially snow); discharge, evapotranspiration and other outputs; and storages in the system.

The above issues are reflected in the discussions that have produced the 23 unsolved problems in hydrology (Blöschl et al., 2019) and the British Hydrological Society Working Group on the Future of Hydrological Science (to which all of the co-authors have contributed). The aim of these two initiatives has been to stimulate hydrological research by identifying future strategic priorities. Here, we will focus on those areas pertaining to improving the understanding and representation of hydrological processes. Many of the unsolved problems refer to the nature and controls of future *hydrological change*, which surely requires a fundamental understanding of present-day hydrological processes and also of the human impacts on those processes (e.g., Abbott et al., 2019).

It could be considered that our *perceptual* understanding of hydrological processes is actually quite good (see, e.g., the outline in Beven, 2012), though, as in all the sciences, we still expect that understanding to improve over time. Examples of that improvement include recent work on the connectivity on hillslopes (e.g., Bracken et al., 2013; Emanuel, Hazen, McGlynn, & Jencso, 2014; Jencso & McGlynn, 2011) and the isotope studies that reveal differences in soil water and vegetation storages (McDonnell, 2014; Sprenger, Llorens, Cayuela, Gallart, & Latron, 2019). The difficulty comes in translating perceptual understanding, often gained in local experimental situations, into practical *quantitative* analyses of flows, storages, and water quality variables across a range of useful and appropriate time and space scales for a given purpose (see, e.g., the discussions in Beven, 2006; Beven & Germann, 2013; Ward & Packman, 2019). Quantitative analyses will require a model (even if it is only the water balance equation), and it is clear that the quantitative representation of hydrological processes in models is lacking in rigour because of the difficulty of testing models as hypotheses when the observational data are uncertain, at an inappropriate scale, or too sparse (e.g., Beven, 2019b; Beven & Lane, 2019). That is one reason why we have so many hydrological models. Current observational data are not adequate to reject many of our models (though see Hollaway et al., 2018, for an example of the rejection of the rather widely used SWAT model).

To do better hydrology, we really need data streams for water fluxes, water storages, and water quality and catchment properties that will provide better inputs for hydrological predictions and support better hypothesis testing in improving hydrological science. That means better observational methods for all of the terms in the water balance equation as well as the tracer and quality observations required for a better understanding of residence time and transit time distributions and storage exchanges (see, e.g., Harman, 2015, 2019; Rinaldo et al., 2015; Remondi, Kirchner, Burlando, & Fatichi, 2018). Our current perceptual model allows for preferential flows, hot spots, hot moments, and other complexities in both surface and subsurface responses to forcing; most hydrological models do not include these and those that do have not been adequately tested as hypotheses. Scale is important here, since we do not fully understand how these small space-scale and timescale processes might integrate up to larger scales. What is clear is that such localized processes of recharge and run-off generation can be significant in affecting larger scale responses (e.g., Hartmann, Gleeson, & Wada, & Wagener, 2017; Fan et al., 2019; Ward & Packman, 2019).

But if we need to improve observations of all the water balance components, where to begin? This can, at least in part, be tested using simulations to establish what type of observational improvements might be more worthwhile for different purposes, for which different levels of uncertainty might be admissible. Within a Bayesian statistical framework, optimizing observational improvements can be explored using a form of pre-posterior prior analysis, where simulations are used to test the value of assuming new uncertain observations, or new types of observations, to be available. Such a framework has been used before in hydrology, for example, to assess where to place an additional observation well in assessing a groundwater model (see, e.g., Ben-Zvi, Berkowitz, & Kesler, 1988; Freeze, James, Massmann, Sperling, & Smith, 1992; Kollat, Reed, & Maxwell, 2011). In the remote sensing field, Observing System Simulation Experiments are similarly used to provide synthetic data sets for testing the utility of proposed missions (e.g., Durand et al., 2008; Biancamaria et al., 2011; in the case of the Surface Water and Ocean Topography [SWOT] satellite, still to be launched). The answers might not necessarily be simple. Bashford, Beven, and Young (2002), for example, looked at this type of observation gap problem from a slightly different perspective. In many parts of the world, including parts of the United Kingdom, evapotranspiration, rather than discharge, is the dominant output term in the water balance. Using simulations at a 30-m pixel scale, they produced a 1-km^2 scale evapotranspiration flux, which they assumed to be observed by remote sensing with different degrees of error. Using that spatial information, they explored what complexity of process model might be supported if such sensor signals could be made available. The outcome turned out to be much simpler than the representation of evapotranspiration in most hydrological models. This implies that both flux observations with low uncertainty and other types of information (e.g., internal states) would be required to support rigorous hypothesis testing to differentiate between model structures that reflect the complexity of processes in the environment. There will, inevitably, be a strong interaction between the development of model theory and the observational support available. The task then is to try to ensure that the right sort of data are collected for the purposes at hand, whether that be testing model structures or testing applied hydrological predictions.

As an example, one interesting possibility would be the development of a method for observing discharge in arbitrary channel cross sections but with sufficient accuracy to be able to identify spatial differences across the channel network. This spatial mapping is possible using tracers (see, e.g., Huff, O'Neill, Emanuel, Elwood, & Newbold, 1982; Genereux, Hemond, & Mulholland, 1993; Kelleher et al., 2013), at least under the assumption that the tracer is conservative (and with the permission of the relevant regulatory agencies). Such data collection might produce potentially significant improvements in understanding of the role of geology and topography in controlling hydrological processes (as also suggested by the inter-comparisons of catchment isotope responses in Tetzlaff, Seibert, & Soulsby, 2009; Birkel et al., 2018). It might also allow much stronger testing of distributed model predictions in ways not previously possible, subject to having adequate observations of the input forcing. Despite the large and ongoing investment in rainfall radar methods and better rain gauges, a major limitation on model testing is knowing just what the inputs to a catchment area are within complex terrain (e.g., Beven, 2019a; McMillan, Krueger, & Freer, 2012; Yatheendradas et al., 2008). If we did have better methods for estimating catchment inputs and discharges, we would be able to make much more rigorous hypothesis tests given information about storages and residence/transit times.

Simulations will only go so far in deciding what type of measurements should be prioritized. That is because what is produced by a simulation depends on the structural assumptions of the model that produced it, and we have a mismatch between the complexity of the perceptual model of the relevant processes and the relative simplicity of current model structures. This mismatch will particularly be the case with estimates of storage and residence times that are strongly dependent on the assumptions underlying any simulation. Therefore, it would be worth combining the pre-posterior prior approach with some direct monitoring where intensive effort is made to capture all elements of the water balance accurately. There are some existing examples of intensive water balance monitoring in experimental systems such as the artificial hillslopes in Biosphere II in Arizona (Gevaert, Teuling, Uijlenhoet, & Troch, 2014; Hopp et al., 2009; Scudeler et al., 2016) or Hydrohill in China (Gu et al., 2010) where considerable effort is made to capture fluxes and storages. Intensive measurement of fluxes and storage is, however, difficult at larger scales in more "natural" catchments, as the long history of research on experimental catchments in different countries and climatic regions testifies. The International Hydrological Decade (1964–1975) generated a large number of "experimental or representative basins" globally (Robinson & Whitehead, 1993). Some, such as those at Plynlimon in Wales, are still being monitored and provide a strong case for the continuation of routine monitoring in a time of changing hydrological responses. Recent initiatives such as the TERENO basins in Germany (Bogena, 2016; Bogena et al., 2018), the Heihe basin in China (Li et al., 2013), and the CZO basins in the United States have seen considerable investment. Nevertheless, because of the limitations of current measurement technologies, and the lack of control over boundary conditions, it is not clear that any such experimental basin has the information available to critically test perceptual understanding and model formulations.

From an experimental viewpoint, we can consider new data collection of fluxes, storages, and catchment properties as an exercise in constraining uncertainty (Beven et al., 2018). In developing what is known about an existing experimental catchment, or in collecting data from a new study catchment (with a particular purpose in mind), we need to determine what types of information would be most useful in constraining the uncertainties in the understanding and prediction of the catchment responses necessary for that purpose, whether that be testing models as hypotheses or some decision for water management. Such an assessment would include making the most of information we might be able to bring from studies elsewhere (e.g., Evaristo & McDonnell, 2017), as well as information gained from direct observations, remote sensing, intensive field campaigns, or other strategies. The issue has been addressed in the context of the prediction of flow in ungauged basins (e.g., Blöschl, Sivapalan, Savenije, Wagener, & Viglione, 2013) but not in terms of considering the requirements for new observational techniques that might serve to improve hydrological science.

The latter purpose implies a need for better observational technologies and network designs to support hypothesis testing in real catchments of interest that go beyond current monitoring capabilities. This technological mission is necessarily long term because it does not seem that significant improvements to existing methods are yet on the horizon. There have been some improvements in radar and microwave rainfall estimates (Diederich, Ryzhkov, Simmer, Zhang, & Trömel, 2015; Rico-Ramirez, Liguori, & Schellart, 2015); eddy correlation and remote sensing estimates of evapotranspiration (Franssen, Stöckli, Lehner, Rotenberg, & Seneviratne, 2010; Maes, Gentine, Verhoest, & Gonzalez Miralles, 2019); gravity anomaly estimates of storage (Güntner et al., 2017; Huang et al., 2019; Richey et al., 2015), acoustic Doppler measurements of discharge (Farina, Alvisi, & Franchini, 2017; Moore, Jamieson, Rainville, Rennie, & Mueller, 2016), and "citizen science" methods of getting more spatially distributed observations (e.g., Le Coz et al., 2016; Paul et al., 2018; Starkey et al., 2017). Significant epistemic uncertainties and some unmeasured states remain for all of these technologies. For some variables, the uncertainties might be reduced; for others, it might be necessary to seek new methods. There also remain important questions to be resolved about just how to test models as hypotheses when there are important epistemic uncertainties in the observational data, but certainly, a good starting point would be to reduce those uncertainties as far as technologically possible. This is likely to require some radically new approaches to provide a step-change improvement, given the limitations of existing observational techniques.

For such long-term aims, we might draw an analogy with defining a new satellite system for Earth Observation, such as the SWOT mission (e.g., Biancamaria et al., 2009; Biancamaria, Lettenmaier, & Pavelsky, 2016). First, we need to define a functional requirement and then a technical specification and provide a justification for funding, including simulations of the difference the sensor would make, before any satellite-based sensor can be designed, built, and successfully launched. SWOT was listed as a potential mission in NASA's Decadal Plan of 2007; it will hopefully be launched in 2021. In the meantime, SWOT work has generated a large number of papers about how the data will contribute to improving estimates of the global water balance, flood discharges and inundation from larger rivers, surface storage in lakes, and the calibration of hydrological models (e.g., Biancamaria et al., 2009; Lee et al., 2010; Pedinotti, Boone, Ricci, Biancamaria, & Mognard, 2014; Yoon, Beighley, Lee, Pavelsky, & Allen, 2015).

As hydrological science moves into the future, it seems essential to improve observational methods in testing process representations and thereby gaining improved understanding. The British Hydrological Society Working Group suggested a number of long-term needs for improved observational methods (to download the full report, including suggestions on shorter term needs and model and theoretical developments, go to http://www.hydrology.org.uk/bhs-working-group-future.php):

- discharge measurements sufficiently accurate to calculate incremental discharges downstream;
- catchment precipitation inputs to much higher accuracies for better characterization of catchment water balance and forecasting purposes;
- total subsurface storage at scales useful for defining some "process response unit";
- better characterization of dynamic storages in different layers; and
- better characterization of controls on fluxes of water and solutes in different layers (including hot spots/hot moments/preferential flows/non-homogenous turbulence/ …) in relation to soil hydrological functioning and land management.

A combination of such field observations and model testing might be one way of combatting the general decline of field hydrology relative to modelling (e.g., Burt & McDonnell, 2015). In doing so, however, we need to be ambitious: to start to evaluate just where the biggest advances might be made for the purposes of both hydrological science and applied hydrology. Initially, this would have to make use of the type of prior simulations suggested earlier, testing how different levels and types of observation might make a difference to hypothesis testing and hydrological practice. These combinations should lead, as a community effort, to defining and commissioning new technologies and would, we believe, lead to significant gains for hydrological science. There is, of course, the question of who would pay for those new technologies to be developed and made available, which also depends on issues of who might invest and who benefits, but the important point is that we should make a start on deciding what should be prioritized, even if the process might be long term.

ORCID

Keith Beven https://orcid.org/0000-0001-7465-3934
Paul Bates https://orcid.org/0000-0001-9192-9963
Eleanor Blyth https://orcid.org/0000-0002-5052-238X
Nick Chappell https://orcid.org/0000-0001-6683-951X
Hannah Cloke https://orcid.org/0000-0002-1472-868X
Simon Dadson https://orcid.org/0000-0002-6144-4639
Nick Everard https://orcid.org/0000-0002-9936-9902
Hayley J. Fowler https://orcid.org/0000-0001-8848-3606
Jim Freer https://orcid.org/0000-0001-6388-7890
David M. Hannah https://orcid.org/0000-0003-1714-1240
Kate Heppell https://orcid.org/0000-0001-6028-1359
Joseph Holden https://orcid.org/0000-0002-1108-4831
Rob Lamb https://orcid.org/0000-0002-9593-621X
Thorsten Wagener https://orcid.org/0000-0003-3881-5849

REFERENCES

Abbott, B. W., Bishop, K., Zarnetske, J. P., Minaudo, C., Chapin, F. S., Krause, S., … Plont, S. (2019). Human domination of the global water cycle absent from depictions and perceptions. *Nature Geoscience*, *12*, 533–540. https://doi.org/10.1038/s41561-019-0374-y

Bashford, K. E., Beven, K. J., & Young, P. C. (2002). Observational data and scale dependent parameterisations: Explorations using a virtual hydrological reality. *Hydrological Processes*, *16*(2), 293–312. https://doi.org/10.1002/hyp.339

Ben-Zvi, M., Berkowitz, B., & Kesler, S. (1988). Pre-posterior analysis as a tool for data evaluation: Application to aquifer contamination. *Water Resources Management*, *2*(1), 11–20. https://doi.org/10.1007/BF00421927

Beven, K. (2019a). Towards a methodology for testing models as hypotheses in the inexact sciences. *Proceedings of the Royal Society A*, *475*(2224), 20180862. https://doi.org/10.1098/rspa.2018.0862

Beven, K. J. (2006). The Holy Grail of scientific hydrology: as closure. *Hydrology and Earth Systems Science*, *10*, 609–618. https://doi.org/10.5194/hess-10-609-2006

Beven, K. J. (2012). *Rainfall–runoff modelling: The primer* (2nd ed.). Wiley Blackwell: Chichester, UK.

Beven, K. J. (2019b). How to make advances in hydrological modelling. *Hydrology Research*, in press. https://doi.org/10.2166/nh.2019.134

Beven, K. J., & Germann, P. F. (2013). Macropores and water flow in soils revisited. *Water Resources Research*, *49*(6), 3071–3092 https://doi.org/10.1002/wrcr.20156

Beven, K. J., Aspinall, W. P., Bates, P. D., Borgomeo, E., Goda, K., Hall, J. W., … Watson, M. (2018). Epistemic uncertainties and natural hazard risk assessment—Part 2: What should constitute good practice? *Natural Hazards and Earth System Sciences*, *18*(10), 2769–2783. https://doi.org/10.5194/nhess-18-2769-2018

Beven, K. J., & Lane, S. (2019). Invalidation of models and fitness-for-purpose: A rejectionist approach, Chapter 6. In C. Beisbart, & N. J. Saam (Eds.), *Computer simulation validation—Fundamental concepts, methodological frameworks, and philosophical perspectives* (pp. 145–171). Cham: Springer.

Biancamaria, S., Andreadis, K. M., Durand, M., Clark, E. A., Rodriguez, E., Mognard, N. M., … Oudin, Y. (2009). Preliminary characterization of SWOT hydrology error budget and global capabilities. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *3*(1), 6–19.

Biancamaria, S., Durand, M., Andreadis, K. M., Bates, P. D., Boone, A., Mognard, N. M., … Clark, E. A. (2011). Assimilation of virtual wide swath altimetry to improve Arctic river modeling. *Remote Sensing of Environment*, *115*(2), 373–381. https://doi.org/10.1016/j.rse.2010.09.008

Biancamaria, S., Lettenmaier, D. P., & Pavelsky, T. M. (2016). The SWOT mission and its capabilities for land hydrology. In *Remote Sensing and Water Resources* (pp. 117–147). Cham: Springer. https://doi.org/10.1007/978-3-319-32449-4_6

Birkel, C., Helliwell, R., Thornton, B., Gibbs, S., Cooper, P., Soulsby, C., … Midwood, A. J. (2018). Characterization of surface water isotope spatial patterns of Scotland. *Journal of Geochemical Exploration*, *194*, 71–80. https://doi.org/10.1016/j.gexplo.2018.07.011

Blöschl, G., Bierkens, M. F., Chambel, A., Cudennec, C., Destouni, G., Fiori, A., … Stumpp, C. (2019). Twenty-three unsolved problems in hydrology (UPH)—A community perspective. *Hydrological Sciences Journal*, *64*, 1141–1158. https://doi.org/10.1080/02626667.2019.1620507

Blöschl, G., Sivapalan, M., Savenije, H., Wagener, T., & Viglione, A. (Eds.) (2013). *Runoff prediction in ungauged basins: Synthesis across processes, places and scales*. Cambridge: Cambridge University Press.

Bogena, H. R. (2016). TERENO: German network of terrestrial environmental observatories. *Journal of large-scale research facilities JLSRF*, *2*, 2, A52.. https://doi.org//10.17815/jlsrf-2-98

Bogena, H. R., Montzka, C., Huisman, J. A., Graf, A., Schmidt, M., Stockinger, M., … Lücke, A. (2018). The TERENO-Rur hydrological observatory: A multiscale multi-compartment research platform for the advancement of hydrological science. *Vadose Zone Journal*, *17*(1). https://doi.org/10.2136/vzj2018.03.0055

Bracken, L. J., Wainwright, J., Ali, G. A., Tetzlaff, D., Smith, M. W., Reaney, S. M., & Roy, A. G. (2013). Concepts of hydrological connectivity: Research approaches, pathways and future agendas. *Earth-Science Reviews*, *119*, 17–34. https://doi.org/10.1016/j.earscirev.2013.02.001

Burt, T. P., & McDonnell, J. J. (2015). Whither field hydrology? The need for discovery science and outrageous hydrological hypotheses. *Water Resources Research*, *51*(8), 5919–5928. https://doi.org/10.1002/2014WR016839

Diederich, M., Ryzhkov, A., Simmer, C., Zhang, P., & Trömel, S. (2015). Use of specific attenuation for rainfall measurement at X-band radar wavelengths. Part I: Radar calibration and partial beam blockage estimation. *Journal of Hydrometeorology*, *16*(2), 487–502. https://doi.org/10.1175/JHM-D-14-0066.1

Durand, M., Andreadis, K. M., Alsdorf, D. E., Lettenmaier, D. P., Moller, D., & Wilson, M. (2008). Estimation of bathymetric depth and slope from data assimilation of swath altimetry into a hydrodynamic model. *Geophysical Research Letters*, *35*(20), L20401. https://doi.org/10.1029/2008GL034150

Emanuel, R. E., Hazen, A. G., McGlynn, B. L., & Jencso, K. G. (2014). Vegetation and topographic influences on the connectivity of shallow groundwater between hillslopes and streams. *Ecohydrology*, *7*(2), 887–895. https://doi.org/10.1002/eco.1409

Evaristo, J., & McDonnell, J. J. (2017). A role for meta-analysis in hydrology. *Hydrological Processes*, *31*(20), 3588–3591. https://doi.org/10.1002/hyp.11253

Fan, Y., Clark, M., Lawrence, D. M., Swenson, S., Band, L. E., Brantley, S. L., … Yamazaki, D. (2019). Hillslope hydrology in global change research and Earth system modeling. *Water Resources Research*, *55*, 1737–1772. https://doi.org/10.1029/2018WR023903

Farina, G., Alvisi, S., & Franchini, M. (2017). Estimating discharge in rivers through the combined use of dimensionless isovels and point velocity measurements. *Hydrology Research*, *48*(3), 616–633. https://doi.org/10.2166/nh.2017.029

Franssen, H. H., Stöckli, R., Lehner, I., Rotenberg, E., & Seneviratne, S. I. (2010). Energy balance closure of eddy-covariance data: A multisite analysis for European FLUXNET stations. *Agricultural and Forest Meteorology*, *150*(12), 1553–1567. https://doi.org/10.1016/j.agrformet.2010.08.005

Freeze, R. A., James, B., Massmann, J., Sperling, T., & Smith, L. (1992). Hydrogeological decision analysis: 4. The concept of data worth and its use in the development of site investigation strategies. *Groundwater*, *30*(4), 574–588. https://doi.org/10.1111/j.1745-6584.1992.tb01534.x

Genereux, D. P., Hemond, H. F., & Mulholland, P. J. (1993). Spatial and temporal variability in streamflow generation on the West Fork of Walker Branch Watershed. *Journal of Hydrology*, *142*(1-4), 137–166. https://doi.org/10.1016/0022-1694(93)90009-X

Gevaert, A. I., Teuling, A. J., Uijlenhoet, R., & Troch, P. A. (2014). Hillslope experiment demonstrates role of convergence during two-step saturation. *Hydrology and Earth System Sciences Discussions*, *11*, 2211–2232. https://doi.org/10.5194/hessd-11-2211-2014

Gu, W. Z., Shang, M. T., Zhai, S. Y., Lu, J. J., Frentress, J., McDonnell, J., & Kendall, C. (2010). Rainfall–runoff paradox from a natural experimental catchment. *Advances in Water Science*, *21*(4), 471–478.

Güntner, A., Reich, M., Mikolaj, M., Creutzfeldt, B., Schroeder, S., & Wziontek, H. (2017). Landscape-scale water balance monitoring with an iGrav superconducting gravimeter in a field enclosure. *Hydrology and Earth System Sciences*, *21*(6), 3167–3182. https://doi.org/10.5194/hess-21-3167-2017

Harman, C. J. (2015). Time-variable transit time distributions and transport: Theory and application to storage-dependent transport of chloride in a watershed. *Water Resources Research*, *51*(1), 1–30. https://doi.org/10.1002/2014WR015707

Harman, C. J. (2019). Age-ranked storage–discharge relations: A unified description of spatially lumped flow and water age in hydrologic systems. *Water Resources Research*, *55*, 7143–7165. https://doi.org/10.1029/2017WR022304

Hartmann, A., Gleeson, T., & Wada & Wagener, T. (2017). Enhanced recharge rates by altered recharge sensitivity to climate variability through subsurface heterogeneity. *PNAS*, *114*(11), 2842–2847. https://doi.org/10.1073/pnas.1614941114

Hollaway, M. J., Beven, K. J., Benskin, C. M. W. H., Collins, A. L., Evans, R., Falloon, P. D., … Haygarth, P. M. (2018). Evaluating a processed based water quality model on a UK headwater catchment: What can we learn from a 'limits of acceptability' uncertainty framework? *Journal of Hydrology*, *558*, 607–624. https://doi.org/10.1016/j.jhydrol.2018.01.063

Hopp, L., Harman, C., Desilets, S. L. E., Graham, C. B., McDonnell, J. J., & Troch, P. A. (2009). Hillslope hydrology under glass: Confronting fundamental questions of soil–water–biota co-evolution at Biosphere 2. *Hydrology and Earth System Sciences*, *13*(11), 2105–2118. https://doi.org/10.5194/hess-13-2105-2009

Huang, Z., Yeh, P. J. F., Pan, Y., Jiao, J. J., Gong, H., Li, X., … Zheng, L. (2019). Detection of large-scale groundwater storage variability over the karstic regions in Southwest China. *Journal of Hydrology*, *569*, 409–422. https://doi.org/10.1016/j.jhydrol.2018.11.071

Huff, D. D., O'Neill, R. V., Emanuel, W. R., Elwood, J. W., & Newbold, J. D. (1982). Flow variability and hillslope hydrology. *Earth Surface Processes and Landforms*, *7*, 91–94. https://doi.org/10.1002/esp.3290070112

Jencso, K. G., & McGlynn, B. L. (2011). Hierarchical controls on runoff generation: Topographically driven hydrologic connectivity, geology, and vegetation. *Water Resources Research*, *47*(11), W11527. https://doi.org/10.1029/2011WR010666

Kelleher, C., Wagener, T., McGlynn, B., Ward, A. S., Gooseff, M. N., & Payn, R. A. (2013). Identifiability of transient storage model parameters along a mountain stream. *Water Resources Research*, *49*(9), 5290–5306. https://doi.org/10.1002/wrcr.20413

Kollat, J. B., Reed, P. M., & Maxwell, R. M. (2011). Many-objective groundwatermonitoring network design using bias-aware ensemble Kalman filtering, evolutionary optimization, and visual analytics. *Water Resources Research*, *42*(2), W02529. https://doi.org/10.1029/2010WR009194

Le Coz, J., Patalano, A., Collins, D., Guillén, N. F., García, C. M., Smart, G. M., … Braud, I. (2016). Crowdsourced data for flood hydrology: Feedback from recent citizen science projects in Argentina, France and New Zealand. *Journal of Hydrology*, *541*, 766–777. https://doi.org/10.1016/j.jhydrol.2016.07.036

Lee, H., Durand, M., Jung, H. C., Alsdorf, D., Shum, C. K., & Sheng, Y. (2010). Characterization of surface water storage changes in Arctic lakes using simulated SWOT measurements. *International Journal of Remote Sensing*, *31*(14), 3931–3953. https://doi.org/10.1080/01431161.2010.483494

Li, X., Cheng, G., Liu, S., Xiao, Q., Ma, M., Jin, R., … Wen, J. (2013). Heihe watershed allied telemetry experimental research (HiWATER):

Scientific objectives and experimental design. *Bulletin of the American Meteorological Society*, *94*(8), 1145–1160.

Maes, W., Gentine, P., Verhoest, N., & Gonzalez Miralles, D. (2019). Potential evaporation at eddy-covariance sites across the globe. *Hydrology and Earth System Sciences*, *23*, 925–948. https://doi.org/10.5194/hess-23-925-2019

McDonnell, J. J. (2014). The two water worlds hypothesis: Ecohydrological separation of water between streams and trees? *Wiley Interdisciplinary Reviews Water*, *1*, 323–329. https://doi.org/10.1002/wat2.1027

McMillan, H., Krueger, T., & Freer, J. (2012). Benchmarking observational uncertainties for hydrology: Rainfall, river discharge and water quality. *Hydrological Processes*, *26*, 4078–4111. https://doi.org/10.1002/hyp.9384

Montanari, A., Young, G., Savenije, H. H. G., Hughes, D., Wagener, T., Ren, L. L., … Belyaev, V. (2013). "Panta Rhei—Everything flows": Change in hydrology and society—The IAHS scientific decade 2013–2022. *Hydrological Sciences Journal*, *58*(6), 1256–1275. https://doi.org/10.1080/02626667.2013.809088

Moore, S. A., Jamieson, E. C., Rainville, F., Rennie, C. D., & Mueller, D. S. (2016). Monte Carlo approach for uncertainty analysis of acoustic Doppler current profiler discharge measurement by moving boat. *Journal of Hydraulic Engineering*, *143*(3), 04016088.

Paul, J. D., Buytaert, W., Allen, S., Ballesteros-Cánovas, J. A., Bhusal, J., Cieslik, K., … Dewulf, A. (2018). Citizen science for hydrological risk reduction and resilience building. *Wiley Interdisciplinary Reviews Water*, *5*(1), e1262. https://doi.org/10.1002/wat2.1262

Pedinotti, V., Boone, A., Ricci, S., Biancamaria, S., & Mognard, N. (2014). Assimilation of satellite data to optimize large-scale hydrological model parameters: A case study for the SWOT mission. *Hydrology and Earth System Sciences*, *18*(11), 4485–4507. https://doi.org/10.5194/hess-18-4485-2014

Remondi, F., Kirchner, J. W., Burlando, P., & Fatichi, S. (2018). Water flux tracking with a distributed hydrological model to quantify controls on the spatio-temporal variability of transit time distributions. *Water Resources Research*, *54*(4), 3081–3099. https://doi.org/10.1002/2017WR021689

Richey, A. S., Thomas, B. F., Lo, M. H., Famiglietti, J. S., Swenson, S., & Rodell, M. (2015). Uncertainty in global groundwater storage estimates in a Total Groundwater Stress framework. *Water Resources Research*, *51*(7), 5198–5216. https://doi.org/10.1002/2015WR017351

Rico-Ramirez, M. A., Liguori, S., & Schellart, A. N. A. (2015). Quantifying radar–rainfall uncertainties in urban drainage flow modelling. *Journal of Hydrology*, *528*, 17–28. https://doi.org/10.1016/j.jhydrol.2015.05.057

Rinaldo, A., Benettin, P., Harman, C. J., Hrachowitz, M., McGuire, K. J., Van Der Velde, Y., … Botter, G. (2015). Storage selection functions: A coherent framework for quantifying how catchments store and release water and solutes. *Water Resources Research*, *51*(6), 4840–4847. https://doi.org/10.1002/2015WR017273

Robinson, M., & Whitehead, P. G. (1993). A review of experimental and representative basin studies. *Methods of Hydrological Basin Comparison*, *120*, 1–12.

Schaller, M., & Fan, Y. (2009). River basins as groundwater exporters and importers: Implications for water cycle and climate modeling. *Journal of Geophysical Research-Atmospheres*, *114*, D04103. https://doi.org/10.1029/2008JD010636

Scudeler, C., Pangle, L., Pasetto, D., Niu, G. Y., Volkmann, T., Paniconi, C., … Troch, P. (2016). Multiresponse modeling of variably saturated flow and isotope tracer transport for a hillslope experiment at the Landscape Evolution Observatory. *Hydrology and Earth System Sciences*, *20*(10), 4061–4078. https://doi.org/10.5194/hess-20-4061-2016

Sivapalan, M., Savenije, H. H., & Blöschl, G. (2012). Socio-hydrology: A new science of people and water. *Hydrological Processes*, *26*, 1270–1276. https://doi.org/10.1002/hyp.8426

Sprenger, M., Llorens, P., Cayuela, C., Gallart, F., & Latron, J. (2019). Mechanisms of consistently disjunct soil water pools over (pore) space and time. *Hydrology and Earth System Sciences*, *23*, 2751–2762. https://doi.org/10.5194/hess-23-2751-2019

Srinivasan, V., Sanderson, M., Garcia, M., Konar, M., Blöschl, G., & Sivapalan, M. (2017). Prediction in a socio-hydrological world. *Hydrological Sciences Journal*, *62*(3), 338–345.

Starkey, E., Parkin, G., Birkinshaw, S., Large, A., Quinn, P., & Gibson, C. (2017). Demonstrating the value of community-based ('citizen science') observations for catchment modelling and characterisation. *Journal of Hydrology*, *548*, 801–817. https://doi.org/10.1016/j.jhydrol.2017.03.019

Tetzlaff, D., Seibert, J., & Soulsby, C. (2009). Inter-catchment comparison to assess the influence of topography and soils on catchment transit times in a geomorphic province; the Cairngorm mountains, Scotland. *Hydrological Processes*, *23*(13), 1874–1886. https://doi.org/10.1002/hyp.7318

Wagener, T., Sivapalan, M., Troch, P. A., McGlynn, B. L., Harman, C. J., Gupta, H. V., … Wilson, J. S. (2010). The future of hydrology: An evolving science for a changing world. *Water Resources Research*, *46*, W05301. https://doi.org/10.1029/2009WR008906

Ward, A. S., & Packman, A. I. (2019). Advancing our predictive understanding of river corridor exchange. *WIREs Water*, *6*, e1327. https://doi.org/10.1002/wat2.1327

Wilby, R. (2019). A hydrology research agenda fit for the 2030s. *Hydrology Research*, submitted, https://doi.org/10.2166/nh.2019.100

Yatheendradas, S., Wagener, T., Gupta, H., Unkrich, C., Goodrich, D., Schaffner, M., & Stewart, A. (2008). Understanding uncertainty in distributed flash flood forecasting for semiarid regions. *Water Resources Research*, *44*, W05S19. https://doi.org/10.1029/2007WR005940

Yoon, Y., Beighley, E., Lee, H., Pavelsky, T., & Allen, G. (2015). Estimating flood discharges in reservoir-regulated river basins by integrating synthetic SWOT satellite observations and hydrologic modeling. *Journal of Hydrologic Engineering*, *21*(4), 05015030.

How to cite this article: Beven K, Asadullah A, Bates P, et al. Developing observational methods to drive future hydrological science: Can we make a start as a community? *Hydrological Processes*. 2019;1–6. https://doi.org/10.1002/hyp.13622