Please find below the comments of W. R. Berghuijs and our replies (preceded by "»").

W.R Berghuijs
wb14708@bristol.ac.uk
Received and published: 13 January 2017
I think the authors address an excellent point by stating that progress in scaling can benefit from utilising data wisely instead of focusing on modelling all the time. While reading the paper I came across a few things that potentially (/hopefully) help to improve the manuscript. This short comment is NOT intended as a full review of the paper.
Overall I enjoyed reading the paper, but I refrain from giving an explicit opinion on the suitability of the manuscript for HESS, because: (i) I am not asked to review the paper, (ii) I did not fully review all aspects of the paper, and (iii) one of the authors (Ross Woods) is my current PhD supervisor.

>>Thank you for your constructive comments

- While reading the paper I was expecting a clear definition of "the fourth paradigm". While the reader will eventually grasp your opinion on this, it seems that the paper can benefit by adding a clear explicit definition of the 4th paradigm early on in the manuscript (e.g. in the final part of the introduction or maybe even in the abstract).

>>Agreed…the revised version of the paper adds a definition of the 4th paradigm in the Introduction. The following definition seems appropriate, and perhaps this broader definition also addresses your next point below:

> The Fourth Paradigm is a concept that focuses on how science can be advanced by enabling full exploitation of data via new computational methods. The concept is based on the idea that computational science constitutes a new set of methods beyond empiricism, theory, and simulation, and is concerned with data discovery in the sense that researchers and scientists require tools, technologies, and platforms that seamlessly integrate into standard scientific methodologies and processes. By integrating these tools and technologies for research, we provide new opportunities for researchers and scientists to share and analyze data and thereby encourage new scientific discovery.

- Your definition, or at least emphasis, for the "fourth paradigm for hydrology" seems to be on systematic testing of hypotheses. This is narrower than the definition of the fourth paradigm as discussed by Hey et al. (2009) (which is something like "insights are wrested from vast troves of existing data"). In the latter definition, there is more emphasis on the data-driven discovery of new laws, rather than the focus on testing (existing) concepts. Do in interpret that correctly? If no: addressing the previous comment may resolve my misinterpretation. If yes: is it worth emphasising the difference between the definitions?

>>Thank you for helping us clarify this important point. Yes, the definition of the 4th paradigm is much broader than we are using here. However, the key point is that (following on the broad definition above) we need to seamlessly integrate computational methods and data into our scientific methodologies and processes. The point being that we have not adequately exploited the "vast troves of data" in testing existing theories and models. We hope this is clarified in the revised version.

- Connected to the previous point: (In my view), it is the combination of the 4 paradigms (empiricism, theory, modeling, systematic testing models/theories with data) that will lead to advances. Should the connection between the four paradigms not be discussed explicitly? Or is there no place for empiricism, new theories and model development in the future of scaling?

>>We do not mean that the 4th paradigm supplants the other 3, and similar points were raised by at least one reviewer. We do believe that perhaps the pendulum has swung a bit too far in the area of simulation such that we have lost our way in terms of what is actually knowable from the data and how to properly test hypotheses that might arise from theory or empiricism. For example, how do we know that processes being represented in a "hyper-resolution" model are adequate without using the data to provide an upper bound to describe the information available? Grounding the scientific method in information theory will help to reconcile this issue.

- Very little is said about past work that tried to systematically assess the validity of scaling hypotheses. Especially, since the paper is introduced at a "review" rather than an "opinion paper" I expected to read more about past efforts before you introduce the need for a fourth paradigm.

>>Yes, another reviewer also commented on the brevity of the paper—although it is really somewhere between an opinion paper and a review. We expand the discussion of previous work in the revised version.

- Can you summarize the vision of your paper in a Figure? I think the paper will be more appealing with such a figure

>>This is a good suggestion. One possibility is a variant of Figure 1 from Gupta et al., HESS, 2014. In this case the diagnostic signatures are based on the patterns in the data as a benchmark and compared with patterns in the model simulations based on a particular similarity concept or hypothesis.

Gupta, H. V., Perrin, C., Blöschl, G., Montanari, A., Kumar, R., Clark, M. and Andréassian, V.: Large-sample hydrology: A need to balance depth with breadth, Hydrol. Earth Syst. Sci., 18(2), 463–477, doi:10.5194/hess-18-463-2014, 2014.

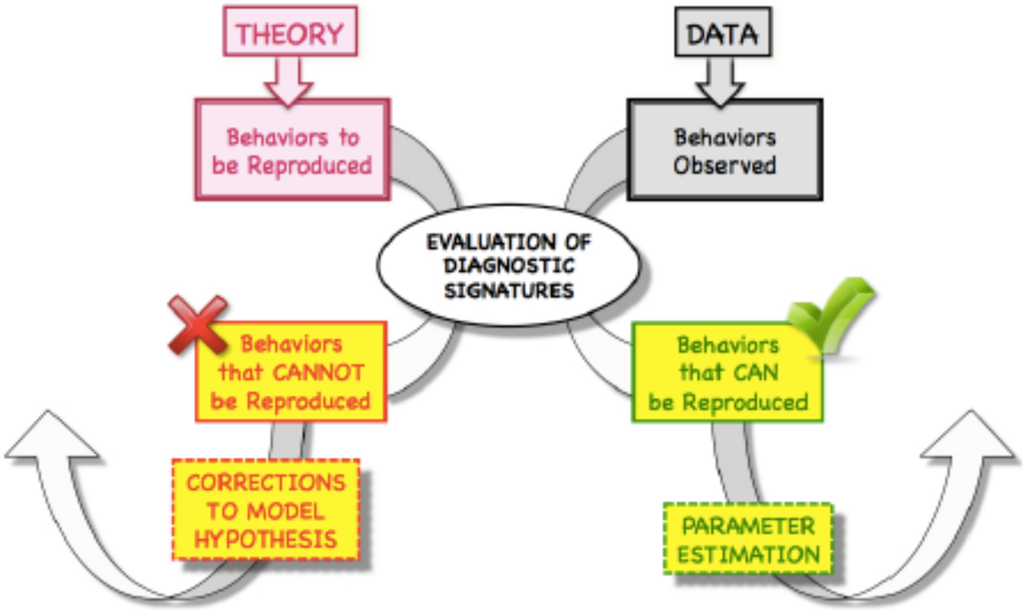


Fig. 1. How an approach based in evaluation of signature properties can be used to detect and diagnose model deficiencies and develop appropriate ways to improve the model/hypothesis (figure based on ideas presented in Gupta et al., 2008).

Technical comments:
Line 14: "larger/longer" suggests that scaling is limited to "upscaling". Why not change it to "other" so it refers to both upscaling and downscaling?
>>Agreed.

References

- In the text there is one citation of Albergel et al (2012). However, in in the reference list, there are two articles by Albergel et al (2012).

>>This is actually an erroneous repetition of the same article. One will be deleted.

-

- Berghuijs et al. (2014) is listed in the references, but not cited in the main text. - Köhli et al (2016) does not include the journal it is published in (WRR?)

>>Thank you for pointing out these oversights. We now include a citation for Berghuijs et al. (2014) in the main text, and the Köhli et al (2016) reference has been corrected to WRR.

Please find below the comments of S. Mylevaganam and our replies (preceded by "»").

S. Mylevaganam
sivarajah@abzwater.com
Received and published: 16 January 2017
Title: Scaling, Similarity, and the Fourth Paradigm for Hydrology
Authors: Christa D. Peters-Lidard, Martyn Clark, Luis Samaniego, Niko E. C. Verhoest, Tim van Emmerik,Remko Uijlenhoet, Kevin Achieng, Trenton E. Franz, Ross Woods
Journal: Hydrology and Earth System Sciences
Review:
In hydrologic science, a system is associated to a set of processes by admitting the fact that there exists no universal law that can identify all the processes associated with a system. Keeping this in mind, these processes are translated into governing equations to formulate a model that can help to resolve a problem of interest. Unfortunately, the inputs data to these models do not come at the same resolution (i.e., physical and temporal dimension). Therefore there exists a need to scale the inputs data to a common resolution that is best suited for the model and its discretization. Moreover, some of the inputs data may not be readily available. This leads to derive these inputs data by relating the characteristics of the system of interest to another system.
Though the growth of hydrologic science, from empiricism (1st paradigm), to theory (2nd paradigm), to computational simulation (3rd paradigm) has yielded important advances in understanding and predictive capabilities in hydrologic science, scaling (i.e., transfer of information across scales) and similarity (i.e., relating characteristics of one system to another system) remain one of the most challenging problems in hydrologic sciences. However, as underscored in the literature, there has been a dramatic increase in the type and density of hydrologic information that is becoming available at multiple scales, from point- to meso-scale and regional to global. Therefore, in this paper, the authors assert that it is time for the hydrologic sciences community to embrace the 4th paradigm of data-intensive science for a data-driven hypothesis testing framework for scaling and similarity.
Based on this review, the following comments are made:
1) The authors assert that it is time for hydrology to embrace a fourth paradigm of data-intensive science. In this paper, the authors too could have used some data to demonstrate the fourth paradigm of data-intensive science.

>> In this paper, we cite the Nearing et al., work as the primary example.

2) In this paper, the authors should clearly state the difference between "concepts" and "hypotheses".

>> Here they are used interchangeably, although to be precise a concept would need to be re-stated as a testable hypothesis in order to apply the scientific method.

3) As per the authors, the advances in data-intensive hydrologic science have laid the foundation for a data-driven hypothesis testing framework for scaling and similarity. For hydrologic sciences community, the concept of hypothesis testing is not new. This concept has been floating and researched for many years. Therefore, the authors need to elaborate more on hypothesis testing that are relevant for the purpose that is outlined in this paper. The paper should also mention the criteria used in a data-driven hypothesis testing framework.

>> Yes it is true that hypothesis testing is not new, and has been practiced in hydrology for many years. The new aspect promoted in this paper is that we need to seamlessly integrate computational methods and data into our scientific methodologies and processes. The point being that we have not adequately exploited the "vast troves of data" in testing existing theories and models. We hope this is clarified in the revised version.

4) Though the authors assert that it is time to embrace the 4th paradigm of data-intensive science, it is hard to understand the actual reasons that have motivated the authors to call upon the hydrologic sciences community to develop a focused effort towards adopting the fourth paradigm of data-intensive science. The authors' assertion gives an indication that, until now, the hydrologic sciences community is reluctant to use the existing data. If this is the case, instead of asking for a paradigm shift, it would be more appropriate to find out the reasons that harbor the hydrologic sciences community from using the existing data.

>> The key point that we would like to emphasize is that hydrologists have been focused primarily on integrated, external characteristics of catchments, such as streamflow, while the advent of higher-resolution remote sensing and other technologies now enables the routine observation of internal dynamics in catchments (soil moisture patterns, evapotranspiration, snowpack, etc.), so that we can now simultaneously test the validity of existing scaling concepts. These scaling concepts or hypotheses were typically developed and tested over only a few catchments, owing to limited data availability. Hence, the 4th paradigm will enable the exploitation of this rich array of hydrological variables. There have been several attempts at comparative hydrology across multiple catchments (e.g., Coopersmith et al., 2012; Berghuijs et al., 2014), but these are typically limited to "macroscale" signatures such as aridity index, runoff ratio, flow-duration curve, etc., rather than examining the internal dynamics and scaling in these watersheds.

5) From the reader's point of view, the problem of scaling and similarity is always going to be there in hydrologic science. The problem of scaling and similarity needs to be addressed when the data is acquired. From the reader's point of view, hydrologic science should not be pleased for having dramatic increase in the type and density of hydrologic information, should rather be disappointed for not being able to device a proper mechanism and methodology to acquire the data that can eliminate the problem of scaling and similarity.

>> As noted in the conclusions, we hope that the community will recognize the opportunity and devise a strategy to test existing hypotheses by combining multiple data types in an information theoretic framework so that we can design future experiments including key observations to reconcile this issue.

6) On page number five, “REH” is undefined.

>> Thank you for catching this. Actually it should be RH for “Representative Hillslope”. We include a definition of Representative Hillslope (RH) on page 3 along with the REA, REW definitions.

7) On page number two, as per the current version of the paper, heterogeneity or variability in hydrology manifests at multiple spatial scales, from local (O(1 m); e.g.,macropores) to hillslope (O(100 m); e.g., preferential flowpaths) to catchment (O(10 km); e.g., soils) and regional (O(1000 km); e.g., geology). Are these numerical values widely accepted by the hydrologic sciences community?

>> These scales come directly from Figure 6 in Blöschl & Sivapalan, HP 1995

Please find below the comments of M. Bierkens and our replies (preceded by "»").

M. Bierkens (Referee)
bierkens@geo.uu.nl
Received and published: 12 February 2017
In this synthesis paper the authors make a case to move from the 3th paradigm of computational science to the 4th paradigm of data-intensive science to overcome the challenges to represent hydrological processes in models over a range of scales. After a short review on scaling and similarity in hydrology they propose an existing framework by Nearing et al. (on information theory) for testing new data-driven scaling laws. They also provide an overview of data requirements to do this as well as what is demanded of a modelling framework to provide for this hypothesis testing (SUMMA being a likely candidate).
The paper is generally well written. I have three comments that I feel are important to address:

>>Thank you for your helpful comments

1) I get the idea behind the information theory to test the model/scaling hypotheses as described in Figure 2, but how it works does not become clear from the figure and this paper. I know, I could read the original paper, but I think a paper should be sufficiently self-contained to convey the message by reading it. So please add a bit more details on the method; especially on how to calculate the different information measures with a given model and scaling method.

>>More description and discussion will be added.

2) My second comment is the most important one, although it may result from my limited understanding of the Nearing et al. method. The authors take the information content of the observations as a benchmark to compare alternative hypotheses against. However, this may be fundamentally problematic if the data are not intensive enough to properly describe the reality you aim to describe. For instance, if I compare a complex distributed model with a lumped-conceptual model against discharge data only, there is no way I will do it better with my complex model, even though the purpose of my model would possibly be to model the internal states (i.e. runoff generation pro- cess or groundwater recharge). So the first question should be: what is the information of the hydrological process I aim to describe. Next, the question than is what data I would need to approximate the necessary information content.

>>As described in Nearing et al., there are three requirements for a benchmark: a particular dataset, a particular performance metric, and a particular reference value for that metric. The dataset must contain information about the true process you seek to model, so it would seem that the streamflow example you mention would require many interior streamflow points in addition to that at the outlet. To

calculate a benchmark, you first measure the information in the observations themselves (using the entropy metric). In the Nearing example, we used SCAN soil moisture data and FLUXNET for ET. To derive the reference value, you generate a synthetic representation of relationship between the forcing and/or parameters and the metric (using a machine learning such as sparse pseudo-input Gaussian Process Regressions, as in Nearing et al.,) that provides an "upper bound" of information contained in the model about the variable of interest. Then the information contained in a model that conforms to a given hypothesis may be compared to the information contained in this "upper bound", and differences in mutual information indicate losses of information due to model formulation.

3) This is generally an optimistic paper providing a way forward to arrive at multi-scale-consistent parameterizations. However, many of the data sources mentioned as needed to support this path forward are remotely sensed data. These will not provide any information on the subsoil. Also georadar or even airborne EM, albeit promising at the local scale, will not give us this information at larger scales. As long as we do not involve (hydro)geologistists and sedimentologists in large-extent high-resolution hydrogeological mapping, the subsoil will remain closed to us I am afraid.

>> The hydrogeophysical techniques described in Binley et al., 2015 (as referenced) do provide some hope for subsurface mapping, although we agree that more work is needed to provide large-scale mapping of the subsurface. We will acknowledge these challenges in the revised manuscript.

Small remark page 4, line 7: also refer to the (improved) ARNO scheme here, e.g. DuÌ´Lmenil, L. and Todini, E., 1992. A rainfall-runoff scheme for use in the Hamburg climate model. Advances in theoretical hydrology. J. P. o ÌA˛Kane, (Editor), Elsevier, 129-157.
Hagemann, S. and L.D. Gates (2003), Improving a subgrid runoff parame-
terization scheme for climate models by the use of high resolution data derived from satellite observations, Climate Dynamics 21, 349–359.

>> These references to the improved ARNO scheme have been added.

Please find below the comments of M. Sivapalan and our replies (preceded by ">>").

M. Sivapalan (Referee)
sivapala@illinois.edu
Received and published: 8 March 2017

This synthesis paper serves an important purpose. It not only helps to celebrate the legacy of Eric Wood (who has made valuable contributions to the topics of scaling and similarity), but also provides a long term (40 years) perspective about where we have come from, what we have learned and where we might go in the future in terms of theory development.

I liked much of what saw in the paper, I am sure there are critical comments that can be made on some details of the history of progress that the authors have provided. I hope these are picked up by other reviewers. In view of the historical nature of the article and a purported new vision offered by the authors, to be effective (and be different) I chose to focus only on the big picture, and decided to keep my comments at a philosophical level. My comments are not necessarily a criticism of the paper, but provide a broader perspective that, if the authors agree and choose to adopt, might bring about a more satisfactory closure to the paper. I am sorry that I use several of my own papers to buttress these arguments, and believe the same opinions are held by many others too.

>> Thank you for your constructive and substantive commentary.

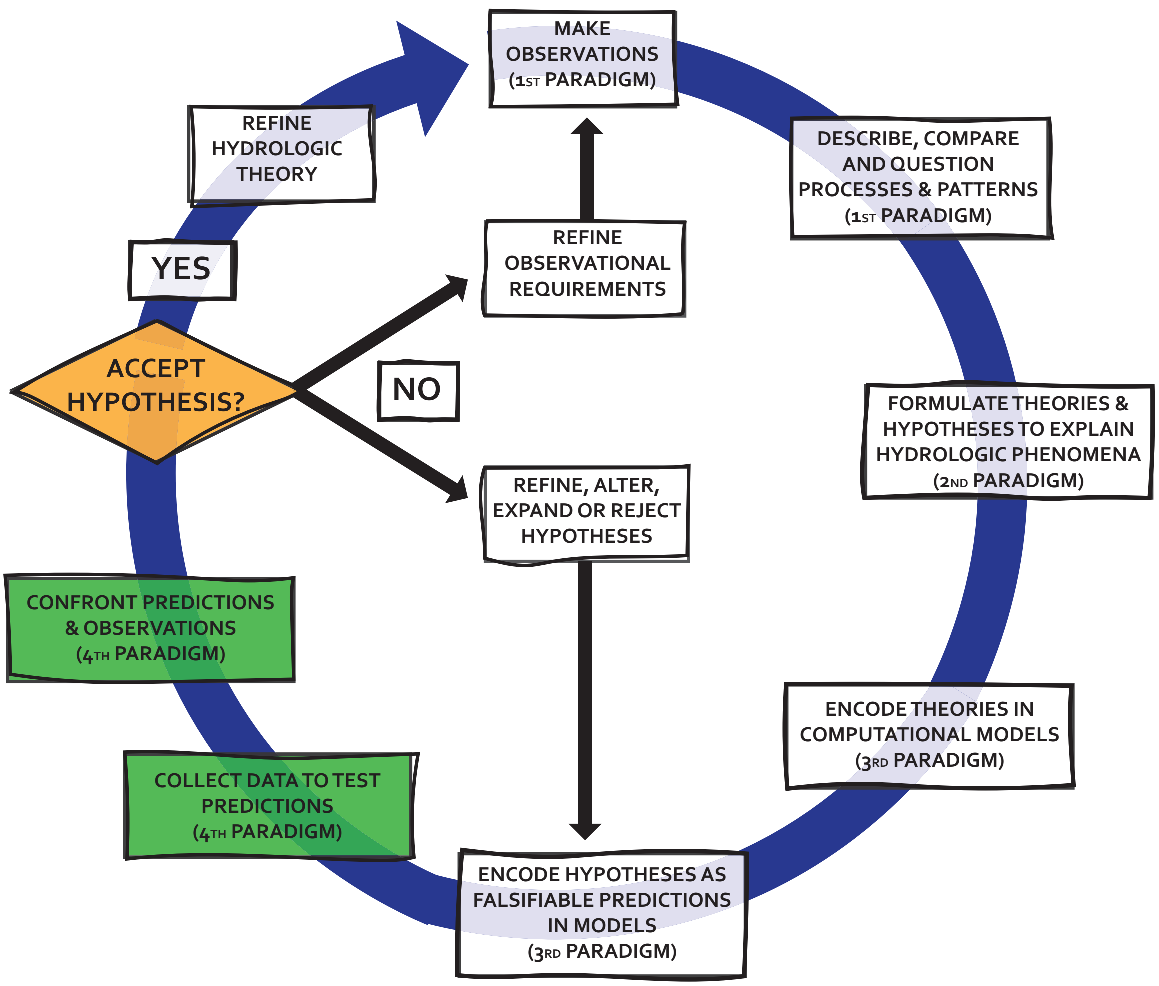
The main argument behind the paper is that even while we have made a lot of progress in the last 40 years on issues of scale and similarity, progress towards a universal theory of hydrology, has slowed down, and the current paradigm is unlikely to lead to further advances. The argument then is that we need a new paradigm (which the authors call the Fourth Paradigm). The fourth paradigm is supposed to be somewhat related to learning from data (and lots of it). The key statement in the paper in this context is this one:

"Fundamentally, these approaches conform to the third paradigm, in the sense that they take as given a set of conservation equations that govern behavior at the fundamental (patch, tile, grid, hillslope, or REW) scale. Testing both the scaling and closure assumptions as hypotheses using data would move hydrology towards the fourth paradigm."

This is confirmation to me that the authors continue to approach the problem within the constraints of the Newtonian framework or worldview, now supplemented by approaches fashionably borrowed from the information sciences currently in vogue. This may advance computational hydrology (I am sure it will), but I am afraid that it will not advance theory development, which was ostensibly the primary focus of the paper.

>> We view the use of information sciences to test existing hypotheses a necessary but not independent step in the scientific method, as illustrated in our new Figure 1

(also below), which was inspired by Figure 1 in Clark et al, 2016. The focus on using new information is not so much on how to generate new theory itself, but rather on how to test it properly. We see this as a significant contribution, but it is not complete in itself. Proper development of testable hypotheses coupled with 4th paradigm-enabled "gathering of data to test predictions" provides opportunities to refine, alter, expand or reject hypotheses, which in turn, can lead to more general theories.



From a medical doctor analogy, I am in agreement with the authors about the nature of the disease (the theory challenge), their diagnosis, and even the direction from where a cure might come from (data/information based learning). To my mind, much of what appears after the presentation of this viewpoint is a lot of hand waving, and does not convince me that it will lead to theory development of the kind they are hoping for. This gap in their logic or unfinished business is surprising, given that the nature of the cure has been evident for some time.

>>As noted above, we are saying "we can formulate a framework for testing hypotheses". This is not the same as generating the hypotheses, but it plays an important role in the process.

Of course, in the era of the "big data", one can understand the thinking that big data will be the panacea to solve all of our ills. I am sure there will be lots and lots of action (including lots of hits and misses), to keep lots of people busy (a veritable cottage industry dealing with lots and lots of noisy statistics and uncertainty analysis). Real progress will be limited unless the focus on data-based learning is guided by some kind of over- arching vision or theoretical framework. This is currently lacking in the paper – what I see is a blind faith that lots and lots of data will somehow bring about breakthroughs that we otherwise have not managed to obtain so far. It could, but only under certain conditions. In my mind, it is not data that produces theoretical breakthroughs, but the kinds of questions that you ask of the data (Sivapalan, 2009). The authors themselves cite Beven and Kirkby (1979) – TOPMODEL theory did not come from data mining, but from somebody sitting down, observing things and letting the imagination go wild. The same think can be said of Budyko (1974). The solution here is not more information theory, but more process hydrology, and plain hard science.

>>As noted in the expanded Introduction, "The Fourth Paradigm is a concept that focuses on how science can be advanced by enabling full exploitation of data via new computational methods." It seems that theories such as TOPMODEL can be more rigorously tested when we utilize large catchment databases (such as MOPEX) coupled with observations of topography, saturated area, streamflow, etc. Our focus is on testing hypotheses in the age of big data, but as shown in Figure 1, this could also lead to refinements in the hypotheses as well as data requirements.

The disappointment for me is that a theoretical framework (one I can mention confidently, others may also exist) to guide this kind of data analysis (i.e., the Fourth Paradigm) already exists. It started becoming articulated a decade ago (Sivapalan, 2005; McDonnell et al., 2007) and has gathered momentum since then, and has found expression as the Darwinian Approach in several papers (Thompson et al., 2011; Harman and Troch, 2014). A prelude to the kind of big-data based Newtonian-Darwinian synthesis that is relevant to this paper has already appeared in the PUB Synthesis Book (Blöschl et al., 2013). In fact, the PUB book carried out a synthesis of catchment scale predictions organized across scales, places and processes. The notion of scale and similarity was the foundation for the extrapolation across places found in the PUB Book.

Chapter 2 of the PUB Book carefully presented the theoretical basis for the synthesis, which was the notion that catchments are co-evolved complex systems. This means that one does not look at catchment as a physical object that provides the boundary conditions for the balance equations for water movement (as one does in a Newtonian approach, which is traditional), but as co-evolved "living" systems, with a long history of co-evolution. Patterns of landscape properties and processes are just a snapshot of a something that has been co-evolving, and one looks at the similarity, differences and scaling that one observes at a moment in time or at a point or area in space arise from multiple trajectories of the same co-evolutionary (land forming and life sustaining) processes, underpinned by the same organizing (if not well known yet) organizing principles. Chapter 12 of the PUB Book presented the outcomes of the synthesis, and discussed how work along these lines can lead to accumulation of knowledge, which is a prelude to new theories. There is much more that can be done along these lines, with new data that is coming on line, as the authors say.

>>As noted above, the principal focus of this work is how the 4th paradigm can assist in testing hypotheses. We believe that the 4th paradigm is consistent with a Darwinian approach. Consider Darwin's "Structure of Coral Reefs" as quoted in Harman and Troch, 2014:

"... In effect, what an immense addition to our knowledge of the laws of nature should we possess if a tithe of the facts dispersed in the Journals of observant travellers, in the Transactions of academies and learned societies, were collected together and judiciously arranged! From their very juxtaposition, plan, co-relation, and harmony, before unsuspected, would become instantly visible, or the causes of anomaly be rendered apparent; erroneous opinions would at once be detected; and new truths – satisfactory as such alone, or supplying corollaries of practical utility – be added to the mass of human knowledge. A better testimony to the justice of this remark can hardly be afforded than in the work before us."

This is precisely the issue that the 4th paradigm seeks to address—using advanced computational technologies to gather together data of different types, collected by different means, and knit them together in an information framework that enables the testing of different hypotheses.

My point is that the Fourth Paradigm will not be a new paradigm unless backed up a broader Earth science perspective, such as this co-evolutionary view. By the way, this is the same worldview that is behind the highly successful Critical Zone Observatories in the United States and also in Europe. So what I am saying is not a biased perspective to impose my own views, but is a widely held perspective in the Earth science communities. Of course in the era of big data and hyper-resolution modeling, one is tempted to believe more in the power of satellites and subsurface geophysics and the power of computers (and of techniques like data assimilation) to generate results that are satisfactory enough for predictions.

>>Agreed. If we agree that "information" is the unifying concept one of the key findings of the Nearing et al, example cited in the paper is that the current macroscale models used to predict soil moisture and evapotranspiration are losing information relative to that contained in the data. Through information theoretic metrics and machine learning designed to provide proper (i.e. asymptotically convergent) estimates of information, it is shown that for soil moisture the majority of information is being lost in the parameters rather than the physics themselves. Further, for evapotranspiration, it is shown that the input boundary conditions ("forcings") are the primary source of lost information. Hence, for these predictands a more fruitful path is to spend effort on properly characterizing soils and/or near-surface meteorology rather than on model physics. This is not a model calibration

exercise—rather it is a demonstration of information content in the model and in the observations, along with attribution of errors.

But if one seriously believes that improvements in theory will be needed for predictions, or can in the long term lead to better predictions (predictions for the right reasons), as I am sure the authors do believe, then there is no alternative but to seriously consider the new co-evolutionary worldview to generate new kinds of questions with which to interrogate the patterns that one finds in the data, test hypotheses about the underlying causes, and use a multitude of tests of hypotheses to move towards general theories. In the absence of such a vision, the combination of traditional Newtonian paradigm and the big data, in my opinion, is a massive exercise in model calibration, parameter regionalization and data assimilation, that will keep a lot of people busy, but will not advance fundamental theory.

>>As shown in new Figure 1, we believe that the 4th paradigm represents an enhancement to the scientific method for hydrology, not a replacement. We agree that co-evolution is a worthwhile avenue to investigate for hydrology, given its demonstrated relevance to other fields.

Big data can indeed help us generate new patterns (at a range of time and space scales) that trigger curiosity and imagination, and will lead to many more examples of simple theories such as the TOPMODEL theory. Indeed the availability of data from thousands of catchments around the world, such as MOPEX and other datasets, is already generating new non-Newtonian understanding through the mechanism of comparative hydrology, as the paper by Berghuijs et al. (2014) illustrates (for example).

In conclusion I enjoyed reading the paper, and indeed agree with the authors on what they are proposing, but believe that they should go to the next (and final) step and frame the problem from a co-evolutionary perspective. The co-evolutionary view is also very critical to frame the new prediction problems in a changing world in the new Anthropocene era (Sivapalan and Blöschl, 2015). They should present avenues, in the style of the natural history approach adopted by Charles Darwin (as described in Harman and Troch, 2016 and Thompson et al., 2012) to generate hypotheses from the data and methods (experiments, numerical simulations etc.) to test these hypotheses to develop new theories.

>> Agreed, Figure 1 now illustrates this.

Of course, this not anything new or unique to hydrology: it is indeed the scientific method, and for this reason I draw inspiration from Jacob Bronowski, and point to a quote from his TV series and book (Bronowski, 1956, p. 23) of the same name "The Ascent of Man":

"All science is the search for unity in hidden likenesses. . .The progress of science is the discovery at each step of a new order which gives unity to what had long seemed unlike. . . For order does not display itself of itself; if it can be said to be there at all, it is not there for the mere looking. . . order must be discovered and, in a deep sense, it must be created. What we see, as we see it, is mere disorder.”

References Cited

Sivapalan, M. (2005). Pattern, Process and Function: Elements of a New Unified Hydrologic Theory at the Catchment Scale. Contribution to: Encyclopaedia of Hydrologic Sciences, M. G. Anderson (Managing Editor), Chapter 13 (Vol. 1, Part 1), pp. 193-219, John Wiley & Sons.

McDonnell, J. J., M. Sivapalan, K. Vaché, S. Dunn, G. Grant, R. Haggerty, C. Hinz, R. P. Hooper, J. W. Kirchner, M. L. Roderick, J. Selker, and M. Weiler (2007). Moving beyond heterogeneity and process complexity: A new vision for watershed hydrology. Water Resources Research, Vol. 43, W07301, doi: 10.1029/2006WR005467.

Thompson, S. E., C. J. Harman, R. A. Schumer, J. S. Wilson, N. B. Basu, P. D. Brooks, S. D. Donner, M. A. Hassan, A. I. Packman, P. S. C. Rao, P. A. Troch and M. Sivapalan (2011). Patterns, puzzles and people: Implementing hydrologic synthesis. Hydrologi- cal Processes, Vol. 25, No. 20, pp. 3256–3266, doi: 10.1002/hyp.8234.

Blöschl, G., M. Sivapalan, T. Wagener, A. Viglione and H. H. G. Savenije, Editors (2013). RUNOFF PREDICTIONS IN UNGAUGED BASINS: A SYNTHESIS ACROSS PROCESSES, PLACES AND SCALES. Cambridge University Press, 500p.

Harman, C. and P. A. Troch (2014). What makes Darwinian hydrology "Darwinian"? Asking a different kind of question about landscapes. Hydrol. Earth Syst. Sci., 18, 417-433, 2014

Sivapalan, M. and G. Blöschl (2015). Time scale interactions and the coevo- lution of humans and water. Water Resources Research, 51(9), 6988–7022 doi:10.1002/2015WR017896.

Please find below the comments of R. Maxwell and our replies (preceded by “>>”).

R. Maxwell (Referee)
rmaxwell@mines.edu
Received and published: 12 April 2017
This manuscript makes the argument for a 4th paradigm, or data-intensive science, as an additional path towards understanding scaling relationships in hydrology. In general this manuscript is clearly written and is of a topic that should be of interest to the readers of HESS and is a good fit for the special issue. I have detailed comments organized topically below that I think the authors should consider in a revised manuscript. I recommend that this discussion article undergo revision prior to publication. Finally, I did find the article to be a bit too concise and while much can be written on the topic and brevity is appreciated, some sections felt like they would benefit from additional discussion (e.g. conclusions and recommendations).

>>Thank you for your constructive comments. Agreed. More description and discussion has been added on these points.

Types of scaling relationships. As the authors point out, scaling may be a function of organization or hydrologic inputs and responses. Functional relationships between hydrologic variables may also exist and these may be scale-independent (or often referred to as scale-invariant).

>>Agreed. This point is noted in the introduction.

Closure. Closure is still important and a formal closure of these systems need be embraced. All models invoke some form of closure at an REV scale and whether this is appropriately represented by the model resolution is an interesting question. Many of the scaling studies cited in this work point out that topography is fractal and this drives much of the scaling behavior seen in hydrology. This presents a special challenge to closure, similar to the atmospheric sciences, where some sub grid information is always needed no matter what the resolution.

>>Agreed. See additional discussion added to the introduction.

Model-data interplay. Increasing model resolution can be for multiple reasons, some of which are pointed out but a critical one not included is to ensure numerical convergence of the underlying PDE solution. This latter reason is in tandem with many of the others mentioned by the authors, datasets and scale of process. An important distinction is models that have parameterizations that represent sub grid processes may not improve with increased resolution, models that have a single column tile / subtile form will only increase as a function of finer / better data. I think the important point here is that data limitations affect all models, but differently. Integrated models with lateral flow of water in surface, subsurface systems that generate runoff directly will have a different spatial sensitivity to the resolution of the input data than more traditional land surface models with no lateral flow and a parametrized runoff generation, which will have a very different spatial sensitivity. The input data matters in all cases, but differently for different classes of model. Spatial resolution matters in all cases, but for some models increasing resolution will have diminishing returns or no effect past a certain point, while for other models these thresholds may not be seen or may be found at different spatial resolutions.

>> Agreed. The PDE/truncation issue noted in introduction and additional discussion added in Section 5 on modeling framework requirements.

Scaling of process. The sub grid parameterizations in e.g. VIC are an important step to understand how processes, such as runoff production, scale up with heterogeneous parameters. High resolution numerical studies indicate that this may depend upon process, excess infiltration doesn't appear to have an ergodic limit (e.g. Maxwell and Kollet AWR 2008), while excess saturation processes scale with the geometric of subsurface K (e.g. Meyerhoff and Maxwell, Hydrogeology 2011). This is but what I'm sure must be only one (self-serving) example of how model simulations like this can inform how parameterizations may be constructed, tested or scaled. In this type of example, the effects of heterogeneity may decrease with scale for runoff that involves groundwater, diminishing the impact of uncertainty in the framework presented by the authors, while they may not for runoff that is purely surface-flow and connection driven. I would imagine that other examples can be used to extend this idea to other processes, e.g. ET, land energy fluxes.

>> These are important points, and we have them along with citations in section 2.

Distinctions between input data, observations for model and process validation. The 4th paradigm, or intensive spatial temporal data, is important in hydrology. I think it is worth distinguishing between uses and types. Some of these data will be used as model input, some to compare or validate. Uncertainty and spatial scale / scope is important in these large (e.g. continental) datasets and although published after this article went into discussion, Christensen et al (WRR 2017), provide some novel approaches for improving the subsurface datasets. Additionally, I think it is important for the so-called 3rd and 4th paradigm to interact in that models can be useful hypotheses generation tools to better inform use of observations in new ways.

>> Agreed. We have included this distinction in section 4 data requirements, and have also included a reference to Christensen et al.

Please find below the comments of U. Lall and our replies (preceded by ">>").

My recommendation is to publish with minor revision. This recognizes that the paper is part of the set for the Wood symposium, and addresses a specific audience. My main reactions are:
1) I am lukewarm to the idea of the 4 paradigms that the authors mention. I am not sure that there is such a clear, sequential separation.

>>Based on comments from other reviewers, we have expanded the definition of the 4th paradigm and also included Figure 1 to help explain that they are all interconnected in the scientific method.

2) I am very sympathetic to the idea that data at multiple scales be used simultaneously in the context of setting up a model and exploring what constitutes similarity. The authors really touch this only towards the end, and do not really develop a mutual information based approach that they promise in the beginning of the paper

>>Based on comments from another reviewers, we have expanded some of this discussion, but we do also rely on the citations to present the background on the method.

3) The authors had me confused with their title – I expected that the paper would develop some notions of self similarity, fractals and emergent behavior across scales from the interactions across coupled hydrologic systems. This would have been an exciting idea for the fourth paradigm, I suppose, albeit not new. However, they are really talking about how to better parameterize surface hydrologic models in a multi scale context, and are developing the notion of similarity and homogeneity that Wood introduced, in parallel to the subsurface literature where such concepts were also being explored. This is perhaps a useful direction for the researchers involved in such an enterprise, and the references to VIC and recent improvements are helpful. Perhaps, I am the only one likely to be confused by the scaling and similarity notions expressed here versus the fractals and nonlinear dynamics literature, but it may be useful to draw the distinction early on

>>We agree that the scope is more narrowly focused on scaling and similarity in hydrology, with the main contribution being to use "big data" to test hypotheses. We have attempted to clarify this in the introduction.

4) I am quite averse to the whole bias correction game that seems endemic in our models nowadays. The one paragraph devoted to it seems to suggest that the authors do not think it is a great idea in the present context, but stop shy of actually trying to clarify that it is not a good thing to do. I would suggest that they make this a stronger statement and emphasize that ideally one needs to use the multiscale

data in a way that best leverages it and demonstrates the ability of the models to reproduce processes at the scales at which those data are available, without any bias correction. Where they talk about dynamics, it would be useful to discuss the reproduction of attributes of dynamics, such as the time rate of decorrelation using an information metric, and the mutual information across variables, space and time. Of course I realize that most of my suggestions reflect my idiosyncratic views and the authors may or may not agree with them

>>This is a good point. We were mostly acknowledging the issue without providing a clear statement of how to apply the multiscale data to address it. We now include statements about both multiscale data and attributes of dynamics.

>>Thank you for your constructive comments.

Scaling, Similarity, and the Fourth Paradigm for Hydrology

Christa D. Peters-Lidard[1], Martyn Clark[2], Luis Samaniego[3], Niko E. C. Verhoest[4], Tim van Emmerik[5], Remko Uijlenhoet[6], Kevin Achieng[7], Trenton E. Franz[8], Ross Woods[9]

[1]Earth Sciences Division, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA
[2]Research Applications Laboratory, National Center for Atmospheric Research, Boulder, CO 80301, USA
[3]UFZ-Helmholtz Centre for Environmental Research, Leipzig, 04318, Germany
[4]Laboratory of Hydrology and Water Management, Ghent University, Coupure links 653, B-9000 Ghent, Belgium
[5]Water Resources Section, Delft University of Technology, Delft, 2628 CN, The Netherlands
[6]Hydrology and Quantitative Water Management Group, Wageningen University, 6700 AA Wageningen, The Netherlands
[7]Department of Civil and Architectural Engineering, University of Wyoming, Laramie, WY 82071, USA
[8]School of Natural Resources, University of Nebraska-Lincoln, Lincoln, NE 68583, USA
[9]Department of Civil Engineering, University of Bristol, Bristol, BS8 1TR, UK

Correspondence to: Christa D. Peters-Lidard (christa.peters@nasa.gov)

Abstract. In this synthesis paper addressing hydrologic scaling and similarity, we posit that roadblocks in the search for universal laws of hydrology are hindered by our focus on computational simulation (the third paradigm), and assert that it is time for hydrology to embrace a fourth paradigm of data-intensive science. Advances in information-based hydrologic science, coupled with an explosion of hydrologic data and advances in parameter estimation and modelling, have laid the foundation for a data-driven framework for scrutinizing hydrological scaling and similarity hypotheses. We summarize important scaling and similarity concepts (hypotheses) that require testing, describe a mutual information framework for testing these hypotheses, describe boundary condition, state/flux, and parameter data requirements across scales to support testing these hypotheses, and discuss some challenges to overcome while pursuing the fourth hydrological paradigm. We call upon the hydrologic sciences community to develop a focused effort towards adopting the fourth paradigm and apply this to outstanding challenges in scaling and similarity.

Peters-Lidard, Chris…, 5/15/2017 4:14 PM
Deleted: review

Peters-Lidard, Chris…, 5/15/2017 4:15 PM
Deleted: of

1 Introduction

This synthesis paper is an outcome of the "Symposium in Honor of Eric Wood: Observations and Modeling across Scales", held June 2-3, 2016 in Princeton, New Jersey, USA. The focus of this contribution is the heterogeneity of hydrological processes, their organization, scaling and similarity, and the impact of the heterogeneity on water and energy states and fluxes (and vice versa). We argue here that the growth of hydrologic science, from empiricism (1st paradigm), via theory (2nd paradigm), to computational simulation (3rd paradigm) has yielded important advances in understanding and predictive capabilities – yet we argue that accelerating advances in hydrologic science will require us to embrace the 4th paradigm of

Uijlenhoet, Remko 6/5/2017 5:47 PM
Deleted: to

data-intensive science, to use emerging datasets to synthesize/scrutinize theories and models, and improve the data support for the mechanisms of Earth System change.

The Fourth Paradigm is a concept that focuses on how science can be advanced by enabling full exploitation of data via new computational methods. The concept is based on the idea that computational science constitutes a new set of methods beyond empiricism, theory, and simulation, and is concerned with data discovery in the sense that researchers and scientists require tools, technologies, and platforms that seamlessly integrate into standard scientific methodologies and processes. By integrating these tools and technologies for research, we provide new opportunities for researchers and scientists to share and analyze data and thereby encourage new scientific discovery. As shown in Figure 1, the scientific method applied to hydrology is not a linear process—rather, because hydrology is already in the 3rd paradigm, empiricism (the 1st paradigm) and theoretical development (the 2nd paradigm) both lead to new theories and hypotheses that are embodied in computational models. These hypotheses may not be rigorously tested with many datasets, either because the datasets have not been gathered into an effective, accessible platform, or because the datasets require additional processing and information theoretic techniques to apply them to the model predictions for hypothesis testing. Further, as noted by Pfister and Kirchner (2017), hypothesis testing with models is fraught with challenges that require not only consideration of the data required to test a given hypothesis, but also careful consideration of how to encode hypotheses as uniquely falsifiable predictions (Figure 1). Advances in data science now allow the 4th paradigm to inject "big data" into the scientific method using rigorous information theoretic methods without eliminating the other parts of the scientific method.

Our focus here on scaling and similarity directs attention to one of the most challenging problems in the hydrologic sciences. As defined by Blöschl and Sivapalan (1995), scale is a "characteristic length (or time) of process, observation, model" and scaling is a "transfer of information across scales" (see also Bierkens et al., 2000; Grayson and Blöschl, 2000). Functional relationships between hydrologic variables may also exist and these may be scale-independent (or scale-invariant). Similarity is present when characteristics of one system can be related to the corresponding characteristics of another system by a simple conversion factor, called the scale factor. We should note that the terms 'scaling' and 'similarity' used here are specific to the hydrology literature and distinct from the general notions of self-similarity, fractals, and emergent behavior in the nonlinear dynamics literature. Classic examples of similarity include the ratio of catchment areas (Willgoose et al., 1991; Smith, 1992), and the topographic index ln(a/tanβ) (Beven and Kirkby, 1979) that are used for relating flows of two catchments and relating the topographic slopes and contributing areas to water table depths, respectively. Another example is the hillslope Péclet number (Berne et al., 2005; Lyon and Troch, 2007). Heterogeneity or variability in hydrology manifests itself at multiple spatial scales (e.g., Seyfried and Wilcox, 1995; Blöschl and Sivapalan, 1995), from local (O(1 m); e.g., macropores) to hillslope (O(100 m); e.g., preferential flowpaths) to catchment (O(10 km); e.g., soils) and regional (O(1000 km); e.g., geology). Similarly, temporal variability is reflected on event, seasonal and decadal time scales (e.g., Woods,

Peters-Lidard, Christ…, 6/9/2017 9:04 AM
Deleted: , but

Uijlenhoet, Remko 5/9/2017 10:00 AM
Comment [1]: Would it make sense to add other similarity indices, such as the hillslope Péclet number (Berne et al., 2005; Lyon and Troch, 2007; both WRR) at this stage? I think it would...

Peters-Lidard, Chris…, 4/24/2017 8:47 AM
Deleted:),

2005). Understanding scaling and similarity requires understanding how the interactions among multiple processes across scales affect the (emergent) hydrologic behaviour at otherother space-time scales; such understanding underpins methods for computational simulation.

The scaling and similarity problem is nevertheless very difficult. As asserted by Dooge (1986), “within the physical sciences and the earth sciences there is and can be no universal model for water movement.” Despite numerous attempts at integrating local models across soils (e.g. Kim et al., 1997), hillslopes (Troch et al., 2015) and watersheds e.g., (Reggiani et al., 1998, 1999, 2000, 2001), universal laws in hydrology and the required closure relations remain elusive because the physics are likely scale-dependent (e.g. Bierkens, 19961996) and the data required to test these hypotheses are either not readily available or not easily synthesized, or, even worse, would never be observable (Beven, 2006). Further, computational advances have enabled so-called “hyper-resolution” or, using anan alternative term that is not necessarily equivalent, “hillslope-resolving” modellingmodelling (e.g. Chaney et al., 2016; Wood et al., 2011), but as noted in the discussion between Beven and Cloke (2012) and Wood et al. (2012), and later discussed in Beven et al. (2015), the ability to provide meaningful information from hillslope-resolving models is limited both by a lack of tested parameterizations at a given model scale as well as by lack of data for model evaluation (e.g. Melsen et al., 2016a).

In principalprincipleprinciple, moving to finer spatial and temporal resolutions may improve accuracy simply by reducing the truncation error in the numerical solution of the system of partial differential equations. In an analogy with fluid mechanics and the atmospheric sciences where “large eddy simulations” are designed to capture the most energetic motions and thereby reduce the sensitivity to turbulence closure, one might ask whether “hillslope-resolving” models might resolve the most energetic components (in an information theoretic/entropy sense) of the terrestrial water storage spectrum such that the closure problem may be simplified. As discussed in many of the studies cited above, topography is fractal and this, combined with scaling between the pedon and the hillslope, drives much of the scaling behavior seen in hydrology. Most of the apparent fractal nature in relation to hydrology has been demonstrated at the scale of river networks (e.g. Tarboton et al, 1988), so a hypothesis that could be tested with data following the 4th paradigm is to what extent resolving these river networks in models reduces the information loss. Further, proposed scaling relationships may be appropriate above a given scale, but as we move downward in scales from watershed to hillslope to local, these relationships may break down.

These current tactics in the hydrologic sciences are representative of the third paradigm of scientific investigation (Hey et al., 2009), characterized by applying computational science to simulate complex systems. The so-called third paradigm builds on the earlier first (empirical) and second (theoretical) paradigms. As discussed by Clark et al., (this issue), computational science approaches to modeling hydrologic systems have been discussed for decades. With the advent of high-resolution earth observing systems (McCabe et al., this issue), proximal sensing (Robinson et al., 2008), sensor networks (Xia et al., 2015), and advances in data-intensive hydrologic science (e.g., Nearing and Gupta, 2015), there is now an opportunity to

Uijlenhoet, Remko 6/5/2017 6:07 PM
Deleted: s

Peters-Lidard, Christ…, 5/5/2017 9:40 AM
Deleted: larger/longerlonger

Uijlenhoet, Remko 5/9/2017 9:53 AM
Deleted: (

Uijlenhoet, Remko 5/9/2017 9:53 AM
Deleted: ,

Martyn Clark 5/28/2017 10:02 AM
Deleted: e.g., (Beven, 2006)

Uijlenhoet, Remko 5/9/2017 9:58 AM
Deleted: 1994

Uijlenhoet, Remko 5/15/2017 4:13 PM
Deleted: 1994

Martyn Clark 5/28/2017 10:02 AM
Deleted: (

Peters-Lidard, Chris…, 5/15/2017 4:26 PM
Deleted: a more physical

Uijlenhoet, Remko 5/9/2017 10:03 AM
Comment [2]: I am not sure “hyper-resolution” and “hillslope-resolving” are necessarily the same.

Peters-Lidard, Chris…, 5/18/2017 1:40 PM
Deleted: modeling

Uijlenhoet, Remko 6/5/2017 6:09 PM
Deleted: , e.g.,

Peters-Lidard, Chris…, 6/9/2017 11:29 AM
Deleted:).

Uijlenhoet, Remko 5/9/2017 9:59 AM
Deleted: .

Martyn Clark 5/28/2017 10:03 AM
Deleted:

Martyn Clark 5/28/2017 10:03 AM
Deleted: alprincipal

Peters-Lidard, Chris…, 5/23/2017 1:38 PM
Deleted: Proposed

recast the hydrologic scaling problem into a data-driven hypothesis testing framework e.g., (Rakovec et al., 2016a). By embracing such a framework, hydrologic analysis can become explicitly "scale-aware" by testing specific parameterizations at a given model scale. Now is the time for a fourth paradigm in hydrologic science.

With this goal in mind, this paper addresses the following questions:

1. What are the key scaling and similarity concepts (hypotheses) that require testing?
2. What framework could we use to test these hypotheses?
3. What are the data requirements to test these hypotheses? and
4. What are the model requirements to test these hypotheses?

2 Scaling and similarity hypotheses

Most scaling work to date has built on the Representative Elementary Area (REA) concept (Wood et al., 1988), and extensions to the Representative Elementary Watersheds (REW) introduced by Reggiani et al. (1998, 1999, 2000, 2001) – the REA/REW concept seeks to define physically meaningful control volumes for which it is possible to obtain simpler descriptions of the rainfall-runoff process (i.e., simpler than those at the point scale). An alternative, but related, concept is the Representative Hillslope (RH; Troch et al., 2003; Berne et al., 2005; Hazenberg et al., 2015). The REA/REW approach is conceptually similar to Reynolds averaging, and relies on the fundamental assumption that the physics are known at the smallest scale considered (e.g. Miller and Miller, 1956). Critically, the fluxes at the boundaries of the model control volumes require parameterization (the so-called "closure" relations). These closure assumptions are typically ad-hoc, and include sub-grid probability distributions, scale-aware parameters, or new flux parameterizations. Fundamentally, these approaches conform to the third paradigm, in the sense that they take as given a set of conservation equations that govern behaviour at the fundamental (patch, tile, grid, hillslope, or REW) scale (Figure 22). Testing both the scaling and closure assumptions as hypotheses using data would move hydrology towards the fourth paradigm.

The examples above represent the classic "Newtonian" approach in hydrology, but the 4th paradigm advocated here is not specific to testing hypotheses derived from that approach, and as shown in Figure 1, represents an augmentation to the scientific method in hydrology. Foundational (Sivapalan, 2005; McDonnell et al., 2007) and more recent work (Thompson et al., 2011; Harman and Troch, 2014) on "Darwinian" hydrology has used scale and similarity concepts to synthesize catchments across scales, places and processes. As noted in McDonnell et al., (2007) there has been a call for a reconciliation of the Newtonian and Darwinian approaches, starting first in the ecology community (Harte, 2002), and we believe that

Uijlenhoet, Remko 5/9/2017 10:07 AM
Deleted: manuscript

Peters-Lidard, Chris…, 4/26/2017 9:25 AM
Deleted: representative

Peters-Lidard, Chris…, 4/26/2017 9:25 AM
Deleted: hillslope

Peters-Lidard, Chris…, 5/15/2017 4:31 PM
Deleted: 1

Uijlenhoet, Remko 6/5/2017 6:15 PM
Deleted:

moving to a 4th paradigm with the augmented scientific method depicted in Figure 1 will embody the wishes of Darwin from his "Structure of Coral Reefs" as quoted in Harman and Troch (2014):

"... In effect, what an immense addition to our knowledge of the laws of nature should we possess if a tithe of the facts dispersed in the Journals of observant travellers, in the Transactions of academies and learned societies, were collected together and judiciously arranged! From their very juxtaposition, plan, co-relation, and harmony, before unsuspected, would become instantly visible, or the causes of anomaly be rendered apparent; erroneous opinions would at once be detected; and new truths – satisfactory as such alone, or supplying corollaries of practical utility – be added to the mass of human knowledge. A better testimony to the justice of this remark can hardly be afforded than in the work before us."

An important avenue to advance hydrologic understanding and predictive capabilities is through attention on hypotheses of hydrologic scaling and similarity, i.e., different ways to relate processes and process interactions across spatial scales. One of the foundational works in hydrologic similarity is the topographic index (Beven and Kirkby, 1979) – the topographic index defines local areas of topographic convergence, and is used to relate the probability distribution of local water table fluctuations to catchment-average surface runoff and sub-surface flow. Building on this topographic similarity, this index was expanded to include soils and study runoff production (Sivapalan et al., 1987; Sivapalan et al., 1990), and further applied to examine scaling of evaporation (Famiglietti and Wood, 1994) and soil moisture (Wood, 1995; Peters-Lidard et al., 2001). Such controls of water table depth on runoff production and evapotranspiration at catchment scales represent just one hypothesis of similarity and scaling behaviour – an example alternative hypothesis, used in the VIC model (Liang et al., 1994), is the description of how sub-element variability in soil moisture affects the development of saturated areas in a catchment and the partitioning of precipitation into surface runoff and infiltration (Moore and Clarke, 1981; Dümenil and Todini, 1992; Wood et al., 1992; Hagemann and Gates, 2003). Other scaling hypotheses are used for other physical processes, for example, how small-scale variability in snow affects large-scale snow melt (Luce et al., 1999; Liston, 2004; Clark et al., 2011a), and how energy fluxes for individual leaves scale up to the vegetation canopy (de Pury and Farquar, 1997; Wang and Leuning, 1998).

The critical issue here is the interplay between the scale of the model elements and the choice of the closure relations: As computational resources permit higher resolution simulations across larger domains (Wood et al., 2011), more physical processes can be represented explicitly, and the closure relations must be tailored to fit the spatial scale of the model simulation. To some extent such "hyper-resolution" approaches abandon the quest for physically meaningful control volumes that characterizes the REA and REW concepts, and the representation of sub-element processes in fully 3-D simulation of watersheds (e.g., Kollet and Maxwell, 2008; Maxwell and Miller, 2005) is becoming less and less obvious, and perhaps less and less necessary. A key question now is whether "hyper-resolution" applications through explicit 3-D models, or (at least for some variables) with clustered 2-D simulations (e.g., the HydroBlocks of Chaney et al., 2016), provide

Uijlenhoet, Remko 5/9/2017 10:26 AM
Comment [3]: Would it make sense to add other similarity indices, such as the hillslope Péclet number (Berne et al., 2005; Lyon and Troch, 2007; both WRR) at this stage?

Peters-Lidard, Chris..., 4/24/2017 4:49 PM
Deleted:).

Peters-Lidard, Chris..., 4/24/2017 4:55 PM
Deleted:

Peters-Lidard, Chris..., 6/9/2017 11:29 AM
Deleted: .

Uijlenhoet, Remko 5/9/2017 10:27 AM
Deleted: .

reasonable representations of scaling and similarity. Considering infiltration excess and saturation excess runoff generation processes, high resolution numerical studies indicate that excess infiltration doesn’t appear to have an ergodic limit (e.g. Maxwell and Kollet (2008), while excess saturation processes scale with the geometric of subsurface saturated hydraulic conductivity (e.g. Meyerhoff and Maxwell, 2011). Similarly, one might imagine different scaling relations for evapotranspiration depending on the nature of controls due to radiation (topography), vegetation, and/or soil moisture (e.g., Rigden and Salvucci, 2015). For example, as recently shown by Maxwell and Condon (2016), the interplay of water table depths with rooting depths along a given hillslope exerts different controls on evaporation and transpiration, which links the water table dynamics with the land surface energy balance, even at continental scales. This finding is based on limited data, and would benefit from formal hypothesis testing in an information-based framework, as described in the next section.

3 A hypothesis testing framework for hydrologic scaling and similarity

As demand increases for hillslope-resolving or “hyper-resolution”modelling modelling (e.g., Beven et al., 2015; Beven and Cloke, 2012; Bierkens et al., 2015; Wood et al., 2011, 2012), the question arises as to whether the physics in our models, the parameters that are used in the models, and the input data (e.g., “forcings”) are adequate to support such endeavoursendeavours (e.g. Melsen et al., 2016b). Following from Nearing and Gupta (2015), we can formulate a framework for testing hypotheses based on measuring information provided by a model (e.g., parameterizations based on similarity concepts) as distinct from information provided to a model (e.g., forcing data or parameters). We should note that this is not hypothesis testing in the traditional sense, but rather a framework for scrutinizing hydrological scaling and similarity hypotheses with data. This concept was demonstrated by Nearing et al. (2016)), who evaluated the information loss due to forcing data, parameters, and physics in the North American Land Data Assimilation System (NLDAS) model ensemble. In this example, information was first measured using point data for soil moisture and evaporation, and compared to regressions that are kernel density estimators of the conditional probability densities and represent the upper bound of information available on a given variable from the forcing data alone and given the forcing data and parameters. As shown in Figure 2, we can measure the total information about a given variable z contained in observations (H(z), left bar), and then measure the information about that variable provided by a given model simulation (I(z; y^M), right bar). The intermediate bars represent losses of information due to forcing data (boundary conditions) and due to parameters.

If we take this example, and expand it to conceptualize a framework for hypothesis testing in hydrology, we can imagine multiple instances of H(z) computed at different spatial scales, as well as multiple instances of mutual information I(z, y^M), computed for models employing different representations of processes at that scale. One concrete example hypothesis described in the previous section is the use of TOPMODEL parameterizations for groundwater, versus representative hillslopes, versus “HydroBlocks” (Chaney et al., 2016) versus explicit 3-D modeling.

Peters-Lidard, Chris…, 5/23/2017 2:12 PM
Deleted:

Uijlenhoet, Remko 6/5/2017 6:22 PM
Deleted:

Peters-Lidard, Chris…, 6/9/2017 11:26 AM
Deleted: .

Uijlenhoet, Remko 5/9/2017 10:27 AM
Deleted: (

Uijlenhoet, Remko 5/9/2017 10:27 AM
Deleted: ,

Peters-Lidard, Chris…, 5/15/2017 4:29 PM
Comment [4]: Agreed, I have noted that the terms are not necessarily equivalent above, but I do think that the notion of hillslope-resolving (similar to convection-resolving or eddy-resolving) is more meaningful than “hyper-resolution”

Uijlenhoet, Remko 5/10/2017 1:43 PM
Comment [5]: I’m not sure all hyper-resolution modeling efforts are actually hillslope resolving. In addition, if we really go hyper-resolution, then our models should not only resolve hillslopes, but vegetation, urban areas, and the subsurface at that resolution as well.

Peters-Lidard, Chris…, 5/15/2017 4:28 PM
Deleted:

Peters-Lidard, Chris…, 6/9/2017 11:29 AM
Deleted: modeling

Peters-Lidard, Chris…, 5/18/2017 1:39 PM
Deleted: endeavors

Uijlenhoet, Remko 5/10/2017 9:54 AM
Deleted:

Peters-Lidard, Chris…, 6/9/2017 11:29 AM
Deleted: .

Uijlenhoet, Remko 5/10/2017 9:54 AM
Deleted: ,

Uijlenhoet, Remko 5/15/2017 4:13 PM
Deleted:)

Uijlenhoet, Remko 5/10/2017 11:10 AM
Deleted:

Uijlenhoet, Remko 5/10/2017 9:54 AM
Deleted:

Uijlenhoet, Remko 5/10/2017 9:55 AM
Deleted:

Critical to this exercise is the availability of forcing data, such as precipitation, radiation, humidity, temperature and wind speed, that have sufficient information content at the scale being evaluated such that it can adequately characterize the variable (e.g., soil moisture) or process (e.g., evapotranspiration; runoff) being studied (e.g. Berne et al., 2004). Similarly, the parameters provided to the model must also contain information about the variable or process being studied at a particular spatial and temporal scale. The Nearing and Gupta approach provides a framework for explicitly measuring the information available from observations, comparing that to information provided by a model, and attributing lost information to forcings, parameters and physics, and hence provides a rigorous method to test our physics assumptions by confronting them with observations. Clearly, this leads to requirements for data that can support such framework.

4 Data requirements

As shown in Figure 1, the 4th paradigm for hydrology is characterized by the rigorous application of large datasets towards testing hypotheses as encapsulated in models. The process of constructing models requires observations both as input data, and for model and process validation or hypotheses testing. A distinguishing characteristic of data for model and process validation will be that we are observing spatial and temporal patterns of fluxes and states represented in our modeling frameworkFor, for example, soil moisture, snow pack or evapotranspiration. As discussed by McCabe et al. (this issue), there has been a dramatic increase in the type and density of hydrologic information that is becoming available at multiple scales, from point- to meso-scale and regional to global. For example, the number of remote sensing missions dedicated to observing the water cycle, allows further development of (large scale) hydrological models and data assimilation frameworks for more accurate soil moisture, evaporation, and streamflow prediction. In particular, there are exciting developments in meso-scale (i.e. hillslope to catchment) observations, which are critical for testing hypotheses about scaling (REA, RH, REW) by connecting point measurements, hydrological models, and remote sensing observations. Examples include recent advances in cosmic ray neutron sensors (Franz et al., 2015; Köhli et al., 2016; Zreda et al., 2008), distributed temperature sensing (DTS; Steele-Dunne et al., 2010; Bense et al., 2016), ; Dong et al., 2016), soil moisture observations, the use of crowd-sourcing (De Vos et al., 2016) and microwave signal propagation from telecommunications towers for precipitation (Leijnse et al., 2007), to the rise in the use of unmanned autonomous vehicles to characterize the landscape at centimeter scale (Vivoni et al., 2014). These alternative data sources enhance our ability to observe, understand, and simulate the hydrological cycle. Advances in citizen science (Buytaert et al., 2014; Hut et al., 2016) and the use of so-called "soft" data for hydrological modeling (Van Emmerik et al., 2015; Seibert and McDonnell, 2002) show that even though these new data are collected on nontraditional spatiotemporal scales, they might give us new insights in how processes at different scales are coupled.

Uijlenhoet, Remko 5/10/2017 1:44 PM
Deleted:

Uijlenhoet, Remko 6/5/2017 6:37 PM
Deleted: .

Peters-Lidard, Chris…, 6/9/2017 11:26 AM
Deleted: For

Uijlenhoet, Remko 6/5/2017 6:37 PM
Deleted: F

Peters-Lidard, Chris…, 6/9/2017 11:29 AM
Deleted: REH

Peters-Lidard, Chris…, 6/9/2017 11:29 AM
Deleted:)

Uijlenhoet, Remko 5/10/2017 1:49 PM
Deleted: and

Uijlenhoet, Remko 5/15/2017 4:13 PM
Deleted:)

Uijlenhoet, Remko 6/5/2017 6:38 PM
Deleted: van

Advances in hydrogeophysical characterization of the subsurface (Binley et al., 2015), such as electrical methods, ground penetrating radar and gravimetry, offer non-invasive meso-scale information that can be used to provide parameters or to infer boundary conditions, states or fluxes. Recently, Christensen, et al. (2017) demonstrated that dense airborne electromagnetic data can be used to map hydrostratigraphic zones, which is an encouraging capability. Imaging the subsoil may be feasible at local scales, but it is a challenge at river basin or continental scales. Hence, we encourage more joint efforts in hydrogeophysical imaging for integrated characterization of the subsurface.

Combined, these observations may be used in a benchmarking exercise similar to Nearing et al. (2016). Synthesizing hydrogeophysical methods with point observations and laboratory/field techniques for estimating "effective" soil hydraulic functions/parameters is a challenging opportunity (e.g. Kim et al., 1997), but one which might be tractable using a data-driven hypothesis testing framework. These new data sources allow us to understand and apply scaling between data sources (point scale to remotely sensed data) and between model scales; and provide the critical data required to test alternative scaling hypotheses.

Beyond the new meso-scale observations, extensive catchment databases now exist to support hypothesis testing including the TERENO (Zacharias et al., 2011), MOPEX (Duan et al., 2006), CONUS benchmarking (Newman et al., 2015a), GRDC (http://www.bafg.de/GRDC/EN/01_GRDC/13_dtbse/database_node.html) and EURO-FRIEND databases (Stahl et al., 2010). Recent similarity studies (Sawicz et al., 2011) have systematically analyzed large numbers of catchments focusing on streamflow-oriented signatures such as the runoff coefficient, baseflow index and slope of the flow duration curve, and then have explored relationships between these signatures and model process time scales (Carrillo et al., 2011). Coopersmith et al. (2012) generalized this work with four nearly orthogonal signatures that included aridity, seasonality of rainfall, peak rainfall, and peak streamflow, and demonstrated that 77% of MOPEX catchments can be described by only six classes defined by combinations of the four signatures. Clearly there is information contained in these catchment databases about not just the coevolution of climate (forcing) and landscape properties (parameters), but also the physics of the catchment responses. Comparative hydrology (e.g., Kovács, 1984; Falkenmark and Chapman, 1989; Gupta et al, 2014) takes a first needed step in the direction of the fourth paradigm, and following the framework described above, we can explicitly quantify the mutual information in the signatures, parameters and forcings to help elucidate these connections beyond classification. One of the crucial factors that complicate scaling is the anthropogenic effect on catchments. Recent advances in modeling the co-evolution of the human-water system (see e.g. Troy et al., 2015; Ciullo et al., 2017) focused on identifying generic key processes and relations. Yet, it is unknown how these relate to systems on larger (and smaller) scales. To arrive at new understandings of scaling and similarities in human-influenced catchments, studying these issues from a socio-hydrological point of view should be an integrated part of the way forward (e.g. Van Loon et al., 2016).

Peters-Lidard, Chris…, 5/23/2017 2:59 PM
Deleted:

Uijlenhoet, Remko 5/10/2017 1:55 PM
Deleted:

Uijlenhoet, Remko 5/10/2017 2:01 PM
Deleted: (

Peters-Lidard, Chris…, 6/9/2017 11:29 AM
Deleted: .

Uijlenhoet, Remko 5/10/2017 2:01 PM
Deleted: ,

Peters-Lidard, Chris…, 6/9/2017 11:29 AM
Deleted:

Uijlenhoet, Remko 5/10/2017 2:02 PM
Deleted:

Uijlenhoet, Remko 5/15/2017 4:13 PM
Deleted: ,

Uijlenhoet, Remko 5/10/2017 2:03 PM
Deleted:

Uijlenhoet, Remko 6/5/2017 6:39 PM
Deleted:

Uijlenhoet, Remko 5/10/2017 2:04 PM
Deleted: ,

Uijlenhoet, Remko 5/10/2017 2:04 PM
Deleted: ,

Uijlenhoet, Remko 6/5/2017 6:42 PM
Comment [6]: See also recent work by Giuliano Di Baldassare et al.

5 Modeling framework requirements

Embracing the fourth paradigm in hydrology will face several challenges. First, it is necessary to implement/extend a hydrologic modelling framework with sufficient flexibility to evaluate competing hypotheses of similarity and scaling behavior (Clark et al., 2011b). One possible framework is the Structure for Unifying Multiple Modeling Alternatives (SUMMA), recently introduced by Clark et al. (2015), which has the capability to incorporate alternative spatial configurations and alternative flux parameterizations. Frameworks like SUMMA, which pursue the method of multiple working hypotheses, enable decomposing complex models into the individual decisions made as part of model development, and focusing attention on specific decisions (e.g., related to scaling and similarity) while keeping all other components of a model constant, hence enabling users to isolate and scrutinize specific hypotheses. One confounding issue is that models with parameterizations designed to represent sub grid processes may not add information in a manner proportional to increased information in the inputs, while models that have a single column tile / subtile form may show a more direct relationship between information in inputs and information in outputs. Similarly, integrated models with lateral flow of water in surface and subsurface systems that generate runoff directly will have a different spatial sensitivity to the resolution of the input data than more traditional land surface models with no lateral flow and a parameterized runoff generation. Hence, the modeling framework must be able to isolate the role that surface and subsurface connectivity play in processing information at different scales.

Uijlenhoet, Remko 6/5/2017 6:42 PM
Deleted:

Peters-Lidard, Christ…, 6/9/2017 9:07 AM
Deleted: , which will have a very different spatial sensitivity

A second challenge consists of understanding how to deal with different uncertainties/errors of different observational products and hydrologic models when comparing them for studying the scaling behavior. Several papers have highlighted the problem of different climatologies or sensitivities of remote sensing products (e.g. Albergel et al., 2012; Brocca et al., 2011), gridded meteorological products (Clark and Slater, 2006; Newman et al., 2015b), and streamflow observations (Di Baldassarre and Montanari, 2009; McMillan et al., 2010). A true correspondence of these remotely sensed variables with model results is often hampered, due to vertical mismatches in the soil column between the different products (Wilker et al., 2006), approximations in the structure of the hydrological model used, its parameterization and discretization, the initial conditions, and errors in forcing data (De Lannoy et al., 2007). Because of this, modeled variables often do not correspond well to observations; nevertheless, similar trends and dynamics between the different products are found (Koster et al., 2009).

Uijlenhoet, Remko 6/5/2017 6:45 PM
Deleted: di

In several data assimilation studies, the problem of differences in climatologies is resolved by bias-correcting the observations towards the model (e.g. Crow et al., 2005; Kumar et al., 2014; Lievens et al., 2015a, 2015b; Martens et al., 2016; Reichle and Koster, 2004; Sahoo et al., 2013; Verhoest et al., 2015). Yet, such (statistical) operations may not be appropriate for scaling studies. First of all, these methods only rescale the remotely sensed value, yet the uncertainties in these products need rescaling as well. Second, depending on the bias-corrections method used (ranging from only correcting for the first moment to full CDF matching) different scaling relations may be found. Ideally, multiscale data should be used in a way that best demonstrates the ability of the models to reproduce processes at the scales at which those data are available, particularly with respect to reproducing attributes of dynamics, such as the time rate of decorrelation using an information metric, and the mutual information across variables, space and time.

Testing hypotheses with multiple scale information also require assimilation/modeling frameworks that allow integrating data into models at their native resolution so that simulations and observations can be compared without the need of introducing ad-hoc downscaling/upscaling rules. One such framework has recently been proposed by Rakovec et al. (2016b). This framework uses the multiscale parameter regionalization (MPR) (Samaniego et al., 2010) technique to link the resolutions of the various data sources with the target modeling resolution and keeping a single set of model transfer parameters that are applicable to all scales. As a result, seamless, flux-matching simulations can be obtained. The MPR-based assimilation framework proposed by Rakovec et al. (2016b) is general and can be used within any land surface or hydrologic model. This framework was originally tested with mesoscale hydrological model (mHM) (Kumar et al., 2013; Samaniego et al., 2010) in order to test hypotheses related to model transferability across scale and locations as well as process description. This data assimilation approach is general and can be used—for example within the SUMMA (Clark et al., 2015) modeling framework—to test hypothesis related with the appropriate model complexity at a given scale. A model agnostic MPR system called MPR-flex has been recently applied to the Variable Infiltration Capacity (VIC) model to estimate seamless parameter and flux fields over CONUS (Mizukami, N., Clark, M., Newman, A., Wood, A., Gutmann, E., Nijssen, B., Samaniego, L. Rakovec,) under reviewunder review). This symbiosis of model parameterization (MPR-Flex) and simulation frameworks (e.g., SUMMA, mHM, etc.) is a very promising avenue to test scaling laws as well as the uncertainty decomposition described above. Finally, the issue of subjective modeling decisions (e.g. the choice of time step, spatial resolution, numerical scheme, study region, time period for calibration / validation, performance metrics, etc.) and associated uncertainties is an issue that requires further attention (e.g. Krueger et al., 2012).

6 Summary and Next Steps

In this paper we review advances in hydrologic scaling and similarity. Beginning with the challenge of Dooge (1986), we posit that roadblocks in the search for universal laws of hydrology are hindered by our third-paradigm approach, and assert that it is time for hydrology to embrace a fourth paradigm of data-intensive science. Building on other synthesis papers in this issue (Clark et al., McCabe et al.), advances in data-intensive hydrologic science (e.g., Nearing and Gupta, 2015) have laid the foundation for a data-driven hypothesis testing framework for scaling and similarity. To achieve this goal, we have (1) summarized important scaling and similarity concepts (hypotheses) that require testing; (2) described a mutual information framework for testing these hypotheses; (3) described boundary condition, state/flux, and parameter data

Peters-Lidard, Chris…, 6/9/2017 11:29 AM
Formatted: English (US)

Uijlenhoet, Remko 5/10/2017 2:14 PM
Deleted: (

Peters-Lidard, Chris…, 6/9/2017 11:29 AM
Deleted: .

Uijlenhoet, Remko 5/10/2017 2:14 PM
Deleted: ,

Uijlenhoet, Remko 5/10/2017 2:17 PM
Deleted: (

Peters-Lidard, Chris…, 6/9/2017 11:29 AM
Deleted: .

Uijlenhoet, Remko 5/10/2017 2:17 PM
Deleted: ,

Peters-Lidard, Chris…, 6/9/2017 11:29 AM
Deleted: hypothes

Peters-Lidard, Chris…, 4/24/2017 8:49 AM
Deleted: is

Martyn Clark 5/28/2017 10:09 AM
Deleted: in preparation,

Peters-Lidard, Chris…, 6/9/2017 11:29 AM
Deleted: 2016

Uijlenhoet, Remko 5/15/2017 4:13 PM
Deleted:)

Martyn Clark 5/28/2017 10:09 AM
Deleted: 2016

Martyn Clark 6/5/2017 9:34 AM
Deleted: 2016

Peters-Lidard, Chris…, 6/9/2017 11:29 AM
Deleted: model

Uijlenhoet, Remko 5/10/2017 2:18 PM
Deleted: l

Peters-Lidard, Chris…, 6/9/2017 11:29 AM
Deleted: ing

requirements across scales to support testing these hypotheses, and (4) discussed some challenges to overcome while pursuing the fourth hydrological paradigm.

Figure 1 illustrates the concept that embracing a 4th paradigm in hydrology where we enable a rigorous confrontation of our hypotheses embodied within our models with a range of data types across many locations and spatial-temporal scales. This paradigm represents a union and extension of previous scientific methods within a formal hypotheses driven framework. Models are a synthesis of all what we have learned (e.g., conservation equations; constitutive relationships for soil infiltration) and data, particularly through first paradigm examples like comparative hydrology, yields empirical relationships, signatures, fingerprints that helps lead to new understanding and theory (2nd paradigm). By coupling traditional (e.g., in situ) and new data sources (e.g., satellites) we can use the power of information theory and rigorous hypothesis testing to elucidate the causes for behaviours that may not be evident in the analysis of individual sites or catchments. In this sense, a move to the 4th paradigm means that we seek modelling-driven monitoring, and simultaneously, monitoring-driven modelling. The formal hypotheses driven framework will indicate where we have weak processes understanding because we cannot explain the data obtained at high resolution. In other cases, comprehensive integrated simulations and big-data relationships would allow the identification of where the measurement errors are too large (i.e. data has little information content, entropy) and point out what kind of sensors or new measurements/sensors are needed to improve our physical understanding. These are the feedback loops in Figure 1, and these represent two important paths to optimizing the use of models and data to enhance hydrologic science.

Peters-Lidard, Christ…, 6/9/2017 9:06 AM
Deleted: modelling

Peters-Lidard, Christ…, 6/9/2017 9:06 AM
Deleted:

As a next step, we propose a focused community effort to shape the development of the fourth paradigm for hydrology. To this end, a workshop following the publication of this special issue would be a good first step.

Uijlenhoet, Remko 6/5/2017 6:51 PM
Deleted:

Acknowledgements

We thank Wouter Berghuijs, Sivarajah Mylevaganam, Marc Bierkens, Murugesu Sivapalan, Reed Maxwell, and Upmanu Lall for their constructive comments on an earlier version of this manuscript.

References

Albergel, C., de Rosnay, P., Gruhier, C., Muñoz-Sabater, J., Hasenauer, S., Isaksen, L., Kerr, Y. and Wagner, W.: Evaluation of remotely sensed and modelled soil moisture products using global ground-based in situ observations, Remote Sens. Environ., 118, 215–226, doi:10.1016/j.rse.2011.11.017, 2012.

Peters-Lidard, Chris…, 4/21/2017 7:09 PM
Deleted: Albergel, C., P. de Rosnay, C. Gruhier, J. Muñoz-Sabater, S. Hasenauer, L. Isaksen, Y. Kerr, W. Wagner, Evaluation of remotely sensed and modelled soil moisture products using global ground-based in situ observations, Remote Sens. Environ., 118, 215-226, 2012.

Bense, V. F., Read, T., Bour, O., Le Borgne, T., Coleman, T., Krause, S., Chalari, A., Mondanos, M., Ciocca, F. and Selker, J. S.: Distributed Temperature Sensing as a downhole tool in hydrogeology, Water Resour. Res., 52(12), 9259–9273, doi:10.1002/2016WR018869, 2016.

Berghuijs, W. R., Sivapalan, M., Woods, R. A. and Savenije, H. H. G.: Patterns of similarity of seasonal water balances: A window into streamflow variability over a range of time scales, Water Resour. Res., 50(7), 5638–5661, doi:10.1002/2014WR015692, 2014.

Berne, A., Delrieu, G., Creutin, J. D. and Obled, C.: Temporal and spatial resolution of rainfall measurements required for urban hydrology, J. Hydrol., 299(3–4), 166–179, doi:10.1016/j.jhydrol.2004.08.002, 2004.

Berne, A., Uijlenhoet, R. and Troch, P. A.: Similarity analysis of subsurface flow response of hillslopes with complex geometry, Water Resour. Res., 41(9), W09410, doi:10.1029/2004WR003629, 2005.

Beven, K.: Searching for the Holy Grail of scientific hydrology: Qt=(S, R, Δt)A as closure, Hydrol. Earth Syst. Sci., 10(5), 609–618, doi:10.5194/hess-10-609-2006, 2006.

Beven, K., Cloke, H., Pappenberger, F., Lamb, R. and Hunter, N.: Hyperresolution information and hyperresolution ignorance in modelling the hydrology of the land surface, Sci. China-Earth Sci., 58(1), 25–35, doi:10.1007/s11430-014-5003-4, 2015.

Beven, K. J. and Cloke, H. L.: Comment on “Hyperresolution global land surface modeling: Meeting a grand challenge for monitoring Earth’s terrestrial water” by Eric F. Wood et al., Water Resour. Res., 48(1), W01801, doi:10.1029/2011WR010982, 2012.

Beven, K. J. and Kirkby, M. J.: A physically based, variable contributing area model of basin hydrology, Hydrol. Sci. Bull., 24(1), 43–69, doi:10.1080/02626667909491834, 1979.

Beven, K. J., Wood, E. F. and Sivapalan, M.: On hydrological heterogeneity — Catchment morphology and catchment response, J. Hydrol., 100(1), 353–375, doi:10.1016/0022-1694(88)90192-8, 1988.

Bierkens, M. F. P.: Modeling Hydraulic Conductivity of a Complex Confining Layer at Various Spatial Scales, Water Resour. Res., 32(8), 2369–2382, doi:10.1029/96WR01465, 1996.

Bierkens, M. F. P., Finke, P. A. and de Willigen, P.: Upscaling and Downscaling Methods for Environmental Research, Springer Netherlands, ISBN 978-0-7923-6339-2, 2000.

Bierkens, M. F. P., Bell, V. A., Burek, P., Chaney, N., Condon, L. E., David, C. H., de Roo, A., Döll, P., Drost, N., Famiglietti, J. S., Flörke, M., Gochis, D. J., Houser, P., Hut, R., Keune, J., Kollet, S., Maxwell, R. M., Reager, J. T., Samaniego, L., Sudicky, E., Sutanudjaja, E. H., van de Giesen, N., Winsemius, H. and Wood, E. F.: Hyper-resolution global hydrological modelling: What is next?: "Everywhere and locally relevant" M. F. P. Bierkens et al. Invited Commentary, Hydrol. Process., 29(2), 310–320, doi:10.1002/hyp.10391, 2015.

Binley, A., Hubbard, S. S., Huisman, J. A., Revil, A., Robinson, D. A., Singha, K. and Slater, L. D.: The emergence of hydrogeophysics for improved understanding of subsurface processes over multiple scales., Water Resour. Res., 51(6), 3837–3866, doi:10.1002/2015WR017016, 2015.

Blöschl, G, Grayson R.B., and Sivapalan M. On the representative elementary area (REA) concept and its utility for distributed rainfall-runoff modelling. Hydrological Processes 9: 313-330, 1995.

Blöschl, G. and Sivapalan, M.: Scale issues in hydrological modelling: A review, Hydrol. Process., 9(September 1994), 251–290, doi:10.1002/hyp.3360090305, 1995.

Brocca, L., Hasenauer, S., Lacava, T., Melone, F., Moramarco, T., Wagner, W., Dorigo, W., Matgen, P., Martínez-Fernández, J., Llorens, P., Latron, J., Martin, C. and Bittelli, M.: Soil moisture estimation through ASCAT and AMSR-E sensors: An intercomparison and validation study across Europe, Remote Sens. Environ., 115(12), 3390–3408, doi:10.1016/j.rse.2011.08.003, 2011.

Buytaert, W., Zulkafli, Z., Grainger, S., Acosta, L., Alemie, T. C., Bastiaensen, J., De Bièvre, B., Bhusal, J., Clark, J., Dewulf, A., Foggin, M., Hannah, D. M., Hergarten, C., Isaeva, A., Karpouzoglou, T., Pandeya, B., Paudel, D., Sharma, K., Steenhuis, T., Tilahun, S., Van Hecken, G. and Zhumanova, M.: Citizen science in hydrology and water resources: opportunities for knowledge generation, ecosystem service management, and sustainable development, Front. Earth Sci., 2(October), 1–21, doi:10.3389/feart.2014.00026, 2014.

Carrillo, G., Troch, P. A., Sivapalan, M., Wagener, T., Harman, C. and Sawicz, K.: Catchment classification: Hydrological analysis of catchment behavior through process-based modeling along a climate gradient, Hydrol. Earth Syst. Sci., 15(11), 3411–3430, doi:10.5194/hess-15-3411-2011, 2011.

Chaney, N. W., Metcalfe, P. and Wood, E. F.: HydroBlocks: a field-scale resolving land surface model for application over continental extents, Hydrol. Process., 30(20), 3543–3559, doi:10.1002/hyp.10891, 2016.

Christensen, N. K., Minsley, B. J. and Christensen, S.: Generation of 3-D hydrostratigraphic zones from dense airborne electromagnetic data to assess groundwater model prediction error, Water Resour. Res., 53(2), 1019–1038, doi:10.1002/2016WR019141, 2017.

Clark, M. P., Bierkens, M. F. P., Samaniego, L., Woods, R. A., Uijenhoet, R., Bennet, K. E., Pauwels, V. R. N., Cai, X., Wood, A. W., and Peters-Lidard, C. D.: The evolution of process-based hydrologic models: Historical challenges and the collective quest for physical realism, Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-693, in review, 2017.

Clark, M. P., Hendrikx, J., Slater, A. G., Kavetski, D., Anderson, B., Cullen, N. J., Kerr, T., Örn Hreinsson, E. and Woods, R. A.: Representing spatial variability of snow water equivalent in hydrologic and land-surface models: A review, Water Resour. Res., 47(7), W07539, doi:10.1029/2011WR010745, 2011a.

Clark, M. P., Kavetski, D. and Fenicia, F.: Pursuing the method of multiple working hypotheses for hydrological modeling, Water Resour. Res., 47(9), W09301, doi:10.1029/2010WR009827, 2011b.

Clark, M. P., Nijssen, B., Lundquist, J. D., Kavetski, D., Rupp, D. E., Woods, R. A., Freer, J. E., Gutmann, E. D., Wood, A. W., Brekke, L. D., Arnold, J. R., Gochis, D. J. and Rasmussen, R. M.: A unified approach for process-based hydrologic modeling: 1. Modeling concept, Water Resour. Res., 51(4), 2498–2514, doi:10.1002/2015WR017198, 2015.

Clark, M. P., Schaefli, B., Schymanski, S. J., Samaniego, L., Luce, C. H., Jackson, B. M., Freer, J. E., Arnold, J. R., Moore, R. D., Istanbulluoglu, E. and Ceola, S.: Improving the theoretical underpinnings of process-based hydrologic models, Water Resour. Res., 52(3), 2350–2365, doi:10.1002/2015WR017910, 2016.

Clark, M. P. and Slater, A. G.: Probabilistic Quantitative Precipitation Estimation in Complex Terrain, J. Hydrometeorol., 7(1), 3–22, doi:10.1175/JHM474.1, 2006.

Coopersmith, E., Yaeger, M. A., Ye, S., Cheng, L. and Sivapalan, M.: Exploring the physical controls of regional patterns of flow duration curves - Part 3: A catchment classification system based on regime curve indicators, Hydrol. Earth Syst. Sci., 16(11), 4467–4482, doi:10.5194/hess-16-4467-2012, 2012.

Crow, W. T., Koster, R. D., Reichle, R. H. and Sharif, H. O.: Relevance of time-varying and time-invariant retrieval error sources on the utility of spaceborne soil moisture products, Geophys. Res. Lett., 32(24), 1–5, doi:10.1029/2005GL024889, 2005.

Peters-Lidard, Chris…, 5/23/2017 3:54 PM

Deleted: Clark, M., Bierkens, M., Samaniego, L., Woods, R., Uijlenhoet, R., Bennett, K., Pauwels, V., Cai, X., Wood, A., and Peters-Lidard, C., The evolution of process-based hydrologic models: Historical challenges and the collective quest for physical realism, submitted to HESS, 2016.

Ciullo, A., Viglione, A., Castellarin, A., Crisci, M. and Di Baldassarre, G.: Socio-hydrological modelling of flood-risk dynamics: comparing the resilience of green and technological systems, Hydrol. Sci. J., 62(6), 880–891, doi:10.1080/02626667.2016.1273527, 2017.

De Lannoy, G. J. M., Houser, P. R., Pauwels, V. R. N. and Verhoest, N. E. C.: State and bias estimation for soil moisture profiles by an ensemble Kalman filter: Effect of assimilation depth and frequency, Water Resour. Res., 43(6), 1–15, doi:10.1029/2006WR005100, 2007.

De Pury, D. G. G. and Farquhar, G. D.: Simple scaling of photosynthesis from leaves to canopies without the errors of big-leaf models, Plant, Cell Environ., 20(5), 537–557, doi:10.1111/j.1365-3040.1997.00094.x, 1997.

Di Baldassarre, G. and Montanari, A.: Uncertainty in river discharge observations: a quantitative analysis, Hydrol. Earth Syst. Sci., 13(6), 913–921, doi:10.5194/hess-13-913-2009, 2009.

De Vos, L., Leijnse, H., Overeem, A. and Uijlenhoet, R.: The potential of urban rainfall monitoring with crowdsourced automatic weather stations in Amsterdam, Hydrol. Earth Syst. Sci, 21, 765–777777, doi:10.5194/hess-21-765-201720172017, 2017.

Dong, J., Steele-Dunne, S. C., Ochsner, T. E., Hatch, C. E., Sayde, C., Selker, J., Tyler, S., Cosh, M. H. and van de Giesen, N.: Mapping high-resolution soil moisture and properties using distributed temperature sensing data and an adaptive particle batch smoother, Water Resour. Res., 52(10), 7690–7710, doi:10.1002/2016WR019031, 2016.

Dooge, J. C. I.: Looking for hydrologic laws, Water Resour. Res., 22(9 S), 46S–58S, doi:10.1029/WR022i09Sp0046S, 1986.

Duan, Q., Schaake, J., Andréassian, V., Franks, S., Goteti, G., Gupta, H. V., Gusev, Y. M., Habets, F., Hall, A., Hay, L., Hogue, T., Huang, M., Leavesley, G., Liang, X., Nasonova, O. N., Noilhan, J., Oudin, L., Sorooshian, S., Wagener, T. and Wood, E. F.: Model Parameter Estimation Experiment (MOPEX): An overview of science strategy and major results from the second and third workshops, J. Hydrol., 320(1–2), 3–17, doi:10.1016/j.jhydrol.2005.07.031, 2006.

Dümenil, L. and Todini, E.: A rainfall-runoff scheme for use in the Hamburg climate model, in Advances in Theoretical Hydrology - A Tribute to James Dooge, pp. 129–157., 1992.

Van Emmerik, T., Mulder, G., Eilander, D., Piet, M. and Savenije, H.: Predicting the ungauged basin: model validation and realism assessment, Front. Earth Sci., 3(October), 1–11, doi:10.3389/feart.2015.00062, 2015.

Falkenmark, M. and Chapman, T. (Eds.): Comparative Hydrology, UNESCO, Paris, 1989.

Uijlenhoet, Remko 6/5/2017 6:27 PM
Deleted: de

Peters-Lidard, Chris…, 6/9/2017 11:29 AM
Deleted: .

Uijlenhoet, Remko 5/10/2017 1:57 PM
Deleted: Discuss.

Uijlenhoet, Remko 5/10/2017 1:57 PM
Deleted: 11–22

Uijlenhoet, Remko 5/10/2017 1:58 PM
Deleted: doi:10.5194/hess-2016-505

Peters-Lidard, Chris…, 6/9/2017 11:29 AM
Deleted: ,

Uijlenhoet, Remko 5/10/2017 1:59 PM
Deleted: 2016

Uijlenhoet, Remko 5/15/2017 4:13 PM
Deleted: 2016

Uijlenhoet, Remko 6/5/2017 6:27 PM
Deleted: van

Famiglietti, J. S. and Wood, E. F.: Multiscale modeling of spatially variable water and energy balance processes, Water Resour. Res., 30(11), 3061–3078, doi:10.1029/94WR01498, 1994.

Fan, Y. and Bras, R. L.: On the concept of a representative elementary area in catchment runoff, Hydrol. Process., 9(7), 821–832, doi:10.1002/hyp.3360090708, 1995.

Franz, T. E., Wang, T., Avery, W., Finkenbiner, C. and Brocca, L.: Combined analysis of soil moisture measurements from roving and fixed cosmic ray neutron probes for multiscale real-time monitoring, Geophys. Res. Lett., (May), 1–8, doi:10.1002/2015GL063963.Received, 2015.

Gao, H., Wood, E. F., Jackson, T. J., Drusch, M. and Bindlish, R.: Using TRMM/TMI to Retrieve Surface Soil Moisture over the Southern United States from 1998 to 2002, J. Hydrometeorol., 7, 23–38, doi:10.1175/JHM473.1, 2006.

Gupta, H. V., C. Perrin, G. Blöschl, A. Montanari, R. Kumar, M. Clark, and V. Andréassian (2014), Large-sample hydrology: A need to balance depth with breadth, Hydrol. Earth Syst. Sci., 18, 463–477, doi:10.5194/hess-18-463-2014.

Grayson, R. and Blöschl, G.: Spatial Patterns in Catchment Hydrology: Observations and Modelling, Cambridge University Press, ISBN 0-521-63316-8, 2001.

Hagemann, S. and Gates, L. D.: Improving a subgrid runoff parameterization scheme for climate models by the use of high resolution data derived from satellite observations, Clim. Dyn., 21(3–4), 349–359, doi:10.1007/s00382-003-0349-x, 2003.

Hazenberg, P., Fang, Y., Broxton, P., Gochis, D., Niu, G. Y., Pelletier, J. D., Troch, P. A. and Zeng, X.: A hybrid-3D hillslope hydrological model for use in Earth system models, Water Resour. Res., 51(10), 8218–8239, doi:10.1002/2014WR016842, 2015.

Hey, T., Tansley, S. and Tolle, K. M.: The fourth paradigm: data-intensive scientific discovery, Microsoft research Redmond, WA., 2009.

Hut, R., Tyler, S. and Van Emmerik, T.: Proof of concept: Temperature-sensing waders for environmental sciences, Geosci. Instrumentation, Methods Data Syst., 5(1), 45–51, doi:10.5194/gi-5-45-2016, 2016.

Kim, C. P., Stricker, J. N. M. and Feddes, R. A.: Impact of soil heterogeneity on the water budget of the unsaturated zone, Water Resour. Res., 33(5), 991–999, doi:10.1029/97WR00364, 1997.

Köhli, M., Schrön, M., Zreda, M., Schmidt, U., Dietrich, P. and Zacharias, S.: Footprint Characteristics Revised for Field-Scale Soil Moisture Monitoring with Cosmic-Ray Neutrons, 51(7), 5772-5790, doi:10.1002/2015WR017169, 2016.

Kollet, S. J. and Maxwell, R. M.: Capturing the influence of groundwater dynamics on land surface processes using an integrated, distributed watershed model, Water Resour. Res., 44(2), W02402, doi:10.1029/2007WR006004, 2008.

Koster, R. D., Guo, Z., Yang, R., Dirmeyer, P. A., Mitchell, K. and Puma, M. J.: On the nature of soil moisture in land surface models, J. Clim., 22(16), 4322–4335, doi:10.1175/2009JCLI2832.1, 2009.

Kovács, G.: Proposal to construct a coordinating matrix for comparative hydrology, Hydrol. Sci. J., 29, 435–443, doi:10.1080/02626668409490961, 1984.

Krueger, T., Page, T., Hubacek, K., Smith, L. and Hiscock, K.: The role of expert opinion in environmental modelling, Environ. Model. Softw., 36, 4–18, doi:10.1016/j.envsoft.2012.01.011, 2012.

Kumar, R., Samaniego, L. and Attinger, S.: Implications of distributed hydrologic model parameterization on water fluxes at multiple scales and locations, Water Resour. Res., 49(1), 360–379, doi:10.1029/2012WR012195, 2013.

Kumar, S. V., Peters-Lidard, C. D., Mocko, D., Reichle, R., Liu, Y., Arsenault, K. R., Xia, Y., Ek, M., Riggs, G., Livneh, B. and Cosh, M. H.: Assimilation of remotely sensed soil moisture and snow depth retrievals for drought estimation, J. Hydrometeorol., 15(6), 2446-2469, doi:10.1175/JHM-D-13-0132.1, 2014.

Leijnse, H., Uijlenhoet, R. and Stricker, J. N. M.: Rainfall measurement using radio links from cellular communication networks, Water Resour. Res., 43(3), W03201, doi:10.1029/2006WR005631, 2007.

Liang, X., Lettenmaier, D. P., Wood, E. F. and Burges, S. J.: A simple hydrologically based model of land surface water and energy fluxes for general circulation models, J. Geophys. Res., 99(D7), 14415, doi:10.1029/94JD00483, 1994.

Lievens, H., De Lannoy, G. J. M., Al Bitar, A., Drusch, M., Dumedah, G., Hendricks Franssen, H. J., Kerr, Y. H., Tomer, S. K., Martens, B., Merlin, O., Pan, M., Roundy, J. K., Vereecken, H., Walker, J. P., Wood, E. F., Verhoest, N. E. C. and Pauwels, V. R. N.: Assimilation of SMOS soil moisture and brightness temperature products into a land surface model, Remote Sens. Environ., 180, 292–304, doi:10.1016/j.rse.2015.10.033, 2015a.

Lievens, H., Tomer, S. K., Al Bitar, A., De Lannoy, G. J. M., Drusch, M., Dumedah, G., Hendricks Franssen, H. J., Kerr, Y. H., Martens, B., Pan, M., Roundy, J. K., Vereecken, H., Walker, J. P., Wood, E. F., Verhoest, N. E. C. and Pauwels, V. R.

Peters-Lidard, Chris…, 6/9/2017 11:29 AM
Deleted: Hydrol

Uijlenhoet, Remko 5/10/2017 2:29 PM
Deleted: og

N.: SMOS soil moisture assimilation for improved hydrologic simulation in the Murray Darling Basin, Australia, Remote Sens. Environ., 168, 146–162, doi:10.1016/j.rse.2015.06.025, 2015b.

Liston, G. E.: Representing Subgrid Snow Cover Heterogeneities in Regional and Global Models, J. Clim., 17(6), 1381–1397, doi:10.1175/1520-0442(2004)017<1381:RSSCHI>2.0.CO;2, 2004.

Luce, C. H., Tarboton, D. G. and Cooley, K. R.: Sub-grid parameterization of snow distribution for an energy and mass balance snow cover model, Hydrol. Process., 13(12–13), 1921–1933, doi:10.1002/(SICI)1099-1085(199909)13:12/13<1921::AID-HYP867>3.0.CO;2-S, 1999.

Lyon, S. W. and Troch, P. A.: Hillslope subsurface flow similarity: Real-world tests of the hillslope Péclet number, Water Resour. Res., 43(7), doi:10.1029/2006WR005323, 2007.

Martens, B., Miralles, D., Lievens, H., Fernández-Prieto, D. and Verhoest, N. E. C.: Improving terrestrial evaporation estimates over continental Australia through assimilation of SMOS soil moisture, Int. J. Appl. Earth Obs. Geoinf., 48, 146–162, doi:10.1016/j.jag.2015.09.012, 2016.

Maxwell, R. M. and Condon, L. E.: Connections between groundwater flow and transpiration partitioning, Science (80-.)., 353(6297), 377–380, doi:10.1126/science.aaf7891, 2016.

Maxwell, R. M. and Kollet, S. J.: Quantifying the effects of three-dimensional subsurface heterogeneity on Hortonian runoff processes using a coupled numerical, stochastic approach, Adv. Water Resour., 31(5), 807–817, doi:10.1016/j.advwatres.2008.01.020, 2008.

Maxwell, R. M. and Miller, N. L.: Development of a Coupled Land Surface and Groundwater Model, J. Hydrometeorol., 6(3), 233–247, doi:10.1175/JHM422.1, 2005.

McCabe, M. F., Rodell, M., Alsdorf, D. E., Miralles, D. G., Uijlenhoet, R., Wagner, W., Lucieer, A., Houborg, R., Verhoest, N. E. C., Franz, T. E., Shi, J., Gao, H., and Wood, E. F.: The Future of Earth Observation in Hydrology, Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2017-54, in review, 2017.

McDonnell, J. J., Sivapalan, M., Vaché, K., Dunn, S., Grant, G., Haggerty, R., Hinz, C., Hooper, R., Kirchner, J., Roderick, M. L., Selker, J. and Weiler, M.: Moving beyond heterogeneity and process complexity: A new vision for watershed hydrology, Water Resour. Res., 43(7), 1–6, doi:10.1029/2006WR005467, 2007.

Peters-Lidard, Chris…, 5/23/2017 3:52 PM
Deleted: McCabe, M., Rodell, M. et al., The future of earth observation in hydrology, Submitted to HESS, 2016.

McMillan, H., Freer, J., Pappenberger, F., Krueger, T. and Clark, M.: Impacts of uncertain river flow data on rainfall-runoff model calibration and discharge predictions, Hydrol. Process., 24(10), 1270–1284, doi:10.1002/hyp.7587, 2010.

Melsen, L. A., Teuling, A. J., Torfs, P. J. J. F., Uijlenhoet, R., Mizukami, N. and Clark, M. P.: HESS Opinions: The need for process-based evaluation of large-domain hyper-resolution models, Hydrol. Earth Syst. Sci., 20(3), 1069–1079, doi:10.5194/hess-20-1069-2016, 2016.

Melsen, L., Teuling, A., Torfs, P., Zappa, M., Mizukami, N., Clark, M. and Uijlenhoet, R.: Representation of spatial and temporal variability in large-domain hydrological models: case study for a mesoscale pre-Alpine basin, Hydrol. Earth Syst. Sci., 20(6), 2207–2226, doi:10.5194/hess-20-2207-2016, 2016.

Meyerhoff, S. B. and Maxwell, R. M.: Quantifying the effects of subsurface heterogeneity on hillslope runoff using a stochastic approach, Hydrogeol. J., 19(8), 1515–1530, doi:10.1007/s10040-011-0753-y, 2011.

Miller, E. E. and Miller, R. D.: Physical Theory for Capillary Flow Phenomena, J. Appl. Phys., 27(4), 324–332, doi:10.1063/1.1722370, 1956.

Mizukami, N., Clark, M., Newman, A., Wood, A., Gutmann, E., Nijssen, B., Samaniego, L. Rakovec, O.: Towards seamless large domain parameter estimation for hydrologic models, WaterWater Resources Research, under review, 20172017.

Moore, R. J. and Clarke, R. T.: A distribution function approach to rainfall runoff modeling, Water Resour. Res., 17(5), 1367–1382, doi:10.1029/WR017i005p01367, 1981.

Nearing, G. S. and Gupta, H. V.: The quantity and quality of information in hydrologic models, Water Resour. Res., 51(1), 524–538, doi:10.1002/2014WR015895, 2015.

Nearing, G. S., Mocko, D. M., Peters-Lidard, C. D., Kumar, S. V and Xia, Y.: Benchmarking NLDAS-2 Soil Moisture and Evapotranspiration to Separate Uncertainty Contributions, J. Hydrometeorol., 17(3), 745–759, doi:10.1175/JHM-D-15-0063.1, 2016.

Newman, A. J., Clark, M. P., Craig, J., Nijssen, B., Wood, A., Gutmann, E., Mizukami, N., Brekke, L. and Arnold, J. R.: Gridded Ensemble Precipitation and Temperature Estimates for the Contiguous United States, J. Hydrometeorol., 16(6), 2481–2500, doi:10.1175/JHM-D-15-0026.1, 2015a.

Newman, A. J., Clark, M. P., Sampson, K., Wood, A., Hay, L. E., Bock, A., Viger, R. J., Blodgett, D., Brekke, L., Arnold, J. R., Hopson, T. and Duan, Q.: Development of a large-sample watershed-scale hydrometeorological data set for the

Martyn Clark 5/28/2017 10:10 AM
Deleted: in

Peters-Lidard, Chris…, 6/9/2017 11:29 AM
Deleted: preparation, 2016

Martyn Clark 5/28/2017 10:10 AM
Deleted: preparation

Martyn Clark 5/28/2017 10:10 AM
Deleted: 6preparation, 2016

contiguous USA: data set characteristics and assessment of regional variability in hydrologic model performance, Hydrol. Earth Syst. Sci., 19(1), 209–223, doi:10.5194/hess-19-209-2015, 2015b.

Peters-Lidard, C. D., Pan, F. and Wood, E. F.: A re-examination of modeled and measured soil moisture spatial variability and its implications for land surface modeling, Adv. Water Resour., 24(9–10), 1069–1083, doi:10.1016/S0309-1708(01)00035-5, 2001.

Pfister, L. and Kirchner, J. W.: Debates-Hypothesis testing in hydrology: Theory and practice, Water Resour. Res., 53(3), 1792–1798, doi:10.1002/2016WR020116, 2017.

Rakovec, O., Kumar, R., Attinger, S. and Samaniego, L.: Improving the realism of hydrologic model functioning through multivariate parameter estimation, Water Resour. Res., 52(10), 7779–7792, doi:10.1002/2016WR019430, 2016a.

Rakovec, O., Kumar, R., Mai, J., Cuntz, M., Thober, S., Zink, M., Attinger, S., Schäfer, D., Schrön, M. and Samaniego, L.: Multiscale and multivariate evaluation of water fluxes and states over European river basinsbasins, J. Hydrometeorol., 17(1), 287–307, doi:10.1175/JHM-D-15-0054.1, 2016b.

Reggiani, P., Sivapalan, M. and Majid Hassanizadeh, S.: A unifying framework for watershed thermodynamics: balance equations for mass, momentum, energy and entropy, and the second law of thermodynamics, Adv. Water Resour., 22(4), 367–398, doi:10.1016/S0309-1708(98)00012-8, 1998.

Reggiani, P., Hassanizadeh, S. M., Sivapalan, M. and Gray, W. G.: A unifying framework for watershed thermodynamics: constitutive relationships, Adv. Water Resour., 23(1), 15–39, doi:10.1016/S0309-1708(99)00005-6, 1999.

Reggiani, P., Sivapalan, M. and Hassanizadeh, S. M.: Conservation equations governing hillslope responses: Exploring the physical basis of water balance, Water Resour. Res., 36(7), 1845–1863, doi:10.1029/2000WR900066, 2000.

Reggiani, P., Sivapalan, M., Hassanizadeh, S. M. and Gray, W. G.: Coupled equations for mass and momentum balance in a stream network: theoretical derivation and computational experiments, Proc. R. Soc. A-Mathematical Phys. Eng. Sci., 457(2005), 157–189, doi:10.1098/rspa.2000.0661, 2001.

Reichle, R. H. and Koster, R. D.: Bias reduction in short records of satellite soil moisture, Geophys. Res. Lett., 31(19), 2–5, doi:10.1029/2004GL020938, 2004.

Rigden, A. J. and Salvucci, G. D.: Evapotranspiration based on equilibrated relative humidity (ETRHEQ): Evaluation over the continental U.S., Water Resour. Res., 51(4), 2951–2973, doi:10.1002/2014WR016072, 2015.

Peters-Lidard, Chris…, 5/18/2017 1:33 PM
Deleted:

Uijlenhoet, Remko 5/9/2017 10:05 AM
Deleted: Multivariate

Uijlenhoet, Remko 5/9/2017 10:05 AM
Deleted: Evaluation

Uijlenhoet, Remko 5/9/2017 10:05 AM
Deleted: Water

Uijlenhoet, Remko 5/9/2017 10:05 AM
Deleted: Fluxes

Uijlenhoet, Remko 5/9/2017 10:05 AM
Deleted: States

Uijlenhoet, Remko 5/9/2017 10:05 AM
Deleted: River

Uijlenhoet, Remko 5/9/2017 10:06 AM
Deleted: Basins

Uijlenhoet, Remko 5/15/2017 4:13 PM
Deleted: Basins

Robinson, D. A., Binley, A., Crook, N., Day-Lewis, F. D., Ferré, T. P. A., Grauch, V. J. S., Knight, R., Knoll, M., Lakshmi, V., Miller, R., Nyquist, J., Pellerin, L., Singha, K. and Slater, L.: Advancing process-based watershed hydrological research using near-surface geophysics: a vision for, and review of, electrical and magnetic geophysical methods, Hydrol. Process., 22(18), 3604–3635, doi:10.1002/hyp.6963, 2008.

Sahoo, A. K., De Lannoy, G. J. M., Reichle, R. H. and Houser, P. R.: Assimilation and downscaling of satellite observed soil moisture over the Little River Experimental Watershed in Georgia, USA, Adv. Water Resour., 52, 19–33, doi:10.1016/j.advwatres.2012.08.007, 2013.

Samaniego, L., Kumar, R. and Attinger, S.: Multiscale parameter regionalization of a grid-based hydrologic model at the mesoscale, Water Resour. Res., 46(5), 1–25, doi:10.1029/2008WR007327, 2010.

Sawicz, K., Wagener, T., Sivapalan, M., Troch, P. A. and Carrillo, G.: Catchment classification: Empirical analysis of hydrologic similarity based on catchment function in the eastern USA, Hydrol. Earth Syst. Sci., 15(9), 2895–2911, doi:10.5194/hess-15-2895-2011, 2011.

Seibert, J. and McDonnell, J. J.: On the dialog between experimentalist and modeler in catchment hydrology: Use of soft data for multicriteria model calibration, Water Resour. Res., 38(11), 23, 1–14, doi:10.1029/2001WR000978, 2002.

Sivapalan, M.: Pattern, Process and Function: Elements of a Unified Theory of Hydrology at the Catchment Scale, in Encyclopedia of Hydrological Sciences, John Wiley & Sons, Ltd, Chichester, UK., 2005.

Sivapalan, M., Beven, K. and Wood, E. F.: On hydrologic similarity: 2. A scaled model of storm runoff production, Water Resour. Res., 23(12), 2266–2278, doi:10.1029/WR023i012p02266, 1987.

Sivapalan, M., Wood, E. F. and Beven, K. J.: On hydrologic similarity: 3. A dimensionless flood frequency model using a generalized geomorphologic unit hydrograph and partial area runoff generation, Water Resour. Res., 26(1), 43–58, doi:10.1029/WR026i001p00043, 1990.

Smith, J. A.: Representation of basin scale in flood peak distributions, Water Resour. Res., 28(11), 2993–2999, doi:10.1029/92WR01718, 1992.

Stahl, K., Hisdal, H., Hannaford, J., Tallaksen, L. M., Van Lanen, H. A. J., Sauquet, E., Demuth, S., Fendekova, M. and Jódar, J.: Streamflow trends in Europe: Evidence from a dataset of near-natural catchments, Hydrol. Earth Syst. Sci., 14(12), 2367–2382, doi:10.5194/hess-14-2367-2010, 2010.

Steele-Dunne, S. C., Rutten, M. M., Krzeminska, D. M., Hausner, M., Tyler, S. W., Selker, J., Bogaard, T. A. and van de Giesen, N. C.: Feasibility of soil moisture estimation using passive distributed temperature sensing, Water Resour. Res., 46(3), W03534, doi:10.1029/2009WR008272, 2010.

Tarboton, D. G., Bras, R. L. and Rodriguez-Iturbe, I.: The fractal nature of river networks, Water Resour. Res., 24(8), 1317–1322, doi:10.1029/WR024i008p01317, 1988.

Thompson, S. E., Harman, C. J., Schumer, R., Wilson, J. S., Basu, N. B., Brooks, P. D., Donner, S. D., Hassan, M. A., Packman, A. I., Rao, P. S. C., Troch, P. A. and Sivapalan, M.: Patterns, puzzles and people: Implementing hydrologic synthesis, Hydrol. Process., 25(20), 3256–3266, doi:10.1002/hyp.8234, 2011.

Troch, P. A, Paniconi, C. and Emiel van Loon, E.: Hillslope-storage Boussinesq model for subsurface flow and variable source areas along complex hillslopes: 1. Formulation and characteristic response, Water Resour. Res., 39(11), 1316, doi:10.1029/2002WR001728, 2003.

Troch, P. A., Lahmers, T., Meira, A., Mukherjee, R., Pedersen, J. W., Roy, T. and Valdés-Pineda, R.: Catchment coevolution: A useful framework for improving predictions of hydrological change?, Water Resour. Res., 51(7), 4903–4922, doi:10.1002/2015WR017032, 2015.

Troy, T. J., Pavao-Zuckerman, M. and Evans, T. P.: Debates-Perspectives on socio-hydrology: Socio-hydrologic modeling: Tradeoffs, hypothesis testing, and validation, Water Resour. Res., 51(6), 4806–4814, doi:10.1002/2015WR017046, 2015.

Van Loon, A. F., Gleeson, T., Clark, J., Van Dijk, A. I. J. M., Stahl, K., Hannaford, J., Di Baldassarre, G., Teuling, A. J., Tallaksen, L. M., Uijlenhoet, R., Hannah, D. M., Sheffield, J., Svoboda, M., Verbeiren, B., Wagener, T., Rangecroft, S., Wanders, N. and Van Lanen, H. A. J.: Drought in the Anthropocene, Nat. Geosci, 9(2), 89–91 [online] Available from: http://dx.doi.org/10.1038/ngeo2646, 2016.

Verhoest, N. E. C., Van den Berg, M. J., Martens, B., Lievens, H., Wood, E. F., Pan, M., Kerr, Y. H., Al Bitar, A., Tomer, S. K., Drusch, M., Vernieuwe, H., De Baets, B., Walker, J. P., Dumedah, G. and Pauwels, V. R. N.: Copula-based downscaling of coarse-scale soil moisture observations with implicit bias correction, IEEE Trans. Geosci. Remote Sens., 53(6), 3507–3521, doi:10.1109/TGRS.2014.2378913, 2015.

Vivoni, E. R., Rango, A., Anderson, C. A., Pierini, N. A., Schreiner-McGraw, A. P., Saripalli, S. and Laliberte, A. S.: Ecohydrology with unmanned aerial vehicles, Ecosphere, 5(10), art130, doi:10.1890/ES14-00217.1, 2014.

Wang, Y. P. and Leuning, R.: A two-leaf model for canopy conductance, photosynthesis and partitioning of available energy I: Model description and comparison with a multi-layered model, Agric. For. Meteorol., 91(1–2), 89–111, doi:10.1016/S0168-1923(98)00061-6, 1998.

Wilker, H., Drusch, M., Seuffert, G. and Simmer, C.: Effects of the Near-Surface Soil Moisture Profile on the Assimilation of L-band Microwave Brightness Temperature, J. Hydrometeorol., 7(3), 433–442, doi:10.1175/JHM498.1, 2006.

Willgoose, G., Bras, R. L. and Rodriguez-Iturbe, I.: A coupled channel network growth and hillslope evolution model: 2. Nondimensionalization and applications, Water Resour. Res., 27(7), 1685–1696, doi:10.1029/91WR00936, 1991.

Wood, E. F.: Scaling behaviour of hydrological fluxes and variables: Empirical studies using a hydrological model and remote sensing data, Hydrol. Process., 9(3–4), 331–346, doi:10.1002/hyp.3360090308, 1995.

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

Wood, E. F., Roundy, J. K., Troy, T. J., van Beek, L. P. H., Bierkens, M. F. P., Blyth, E., de Roo, A., Doell, P., Ek, M., Famiglietti, J., Gochis, D., van de Giesen, N., Houser, P., Jaffe, P. R., Kollet, S., Lehner, B., Lettenmaier, D. P., Peters-Lidard, C., Sivapalan, M., Sheffield, J., Wade, A. and Whitehead, P.: Hyperresolution global land surface modeling: Meeting a grand challenge for monitoring Earth's terrestrial water, Water Resour. Res., 47, W05301,doi:10.1029/2010WR010090, 2011.

Wood, E. F., Roundy, J. K., Troy, T. J., van Beek, R., Bierkens, M., Blyth, E., de Roo, A., Doell, P., Ek, M., Famiglietti, J., Gochis, D., van de Giesen, N., Houser, P., Jaffe, P., Kollet, S., Lehner, B., Lettenmaier, D. P., Peters-Lidard, C. D., Sivapalan, M., Sheffield, J., Wade, A. J. and Whitehead, P.: Reply to comment by Keith J. Beven and Hannah L. Cloke on "Hyperresolution global land surface modeling: Meeting a grand challenge for monitoring Earth's terrestrial water'', Water Resour. Res., 48, W01802, doi:10.1029/2011WR011202, 2012.

Wood, E. F., Sivapalan, M., Beven, K. and Band, L.: Effects of spatial variability and scale with implications to hydrologic modeling, J. Hydrol., 102(1–4), 29–47, doi:10.1016/0022-1694(88)90090-X, 1988.

Woods, R. A. "Hydrologic Concepts of Variability and Scale" in Encyclopedia of Hydrological Sciences. (1 ed., Vol. 1, pp. 23-40). John Wiley and Sons, Inc., 2005.

Xia, Y., Ek, M. B., Wu, Y., Ford, T. and Quiring, S. M.: Comparison of NLDAS-2 Simulated and NASMD Observed Daily Soil Moisture. Part I: Comparison and Analysis, J. Hydrometeorol., 16(5), 1962–1980, doi:10.1175/JHM-D-14-0096.1, 2015.

Zacharias, S., Bogena, H., Samaniego, L., Mauder, M., Fuß, R., Pütz, T., Frenzel, M., Schwank, M., Baessler, C., Butterbach-Bahl, K., Bens, O., Borg, E., Brauer, A., Dietrich, P., Hajnsek, I., Helle, G., Kiese, R., Kunstmann, H., Klotz, S., Munch, J. C., Papen, H., Priesack, E., Schmid, H. P., Steinbrecher, R., Rosenbaum, U., Teutsch, G. and Vereecken, H.: A Network of Terrestrial Environmental Observatories in Germany, Vadose Zone J., 10, 955–973, doi:10.2136/vzj2010.0139, 2011.

Zreda, M., Desilets, D., Ferré, T. P. A. and Scott, R. L.: Measuring soil moisture content non-invasively at intermediate spatial scale using cosmic-ray neutrons, Geophys. Res. Lett., 35(21), 1–5, doi:10.1029/2008GL035655, 2008.

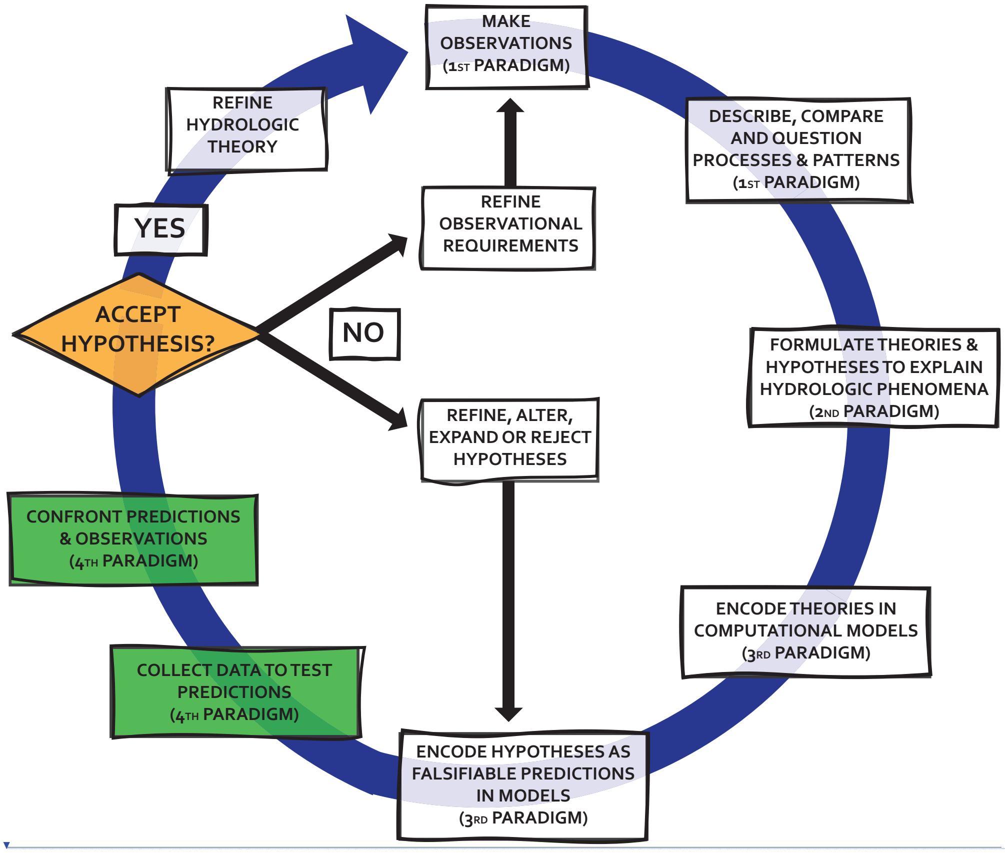
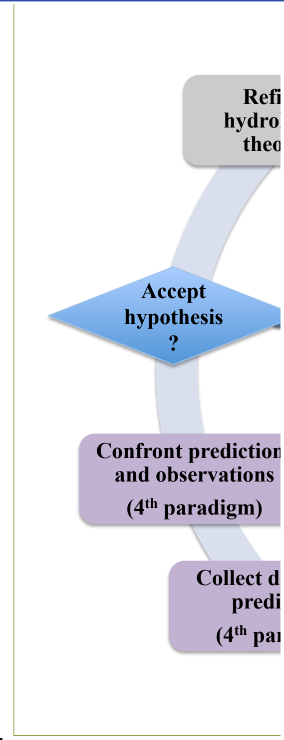


Figure 1: An illustration of the scientific method in hydrology, highlighting how each component of the method reflects the various paradigms of science. The 4th paradigm is characterized by advanced data collection and analysis, as noted in the green boxes. Based on Figure 1 in Clark et al., 2016.

Peters-Lidard, Chris…, 5/26/2017 1:16 PM

Deleted:



Peters-Lidard, Christ…, 6/9/2017 9:07 AM

Deleted: purple

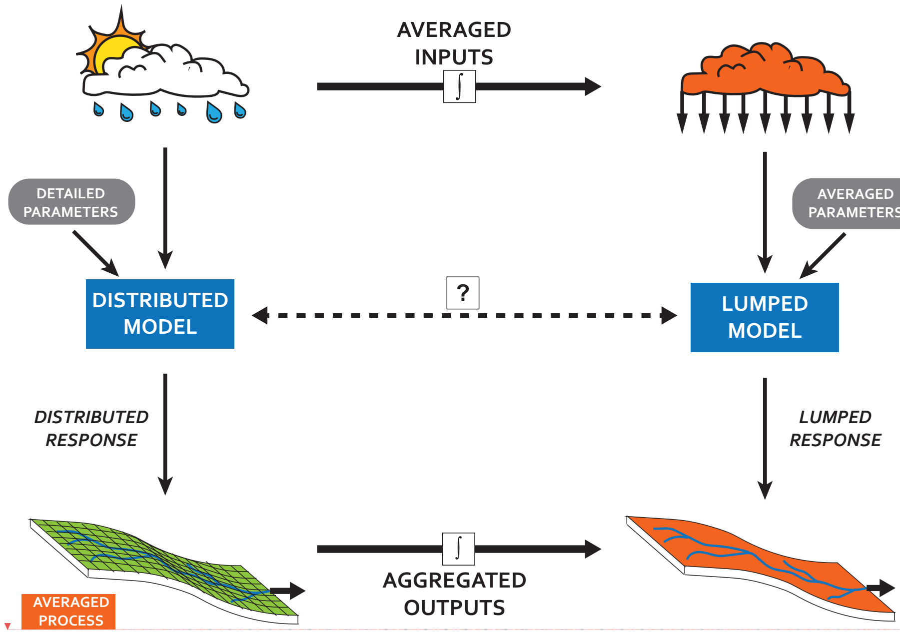


Figure 2: Aggregation and scaling schematic following Wood (1995).

Uijlenhoet, Remko 5/15/2017 4:13 PM
Deleted:

Peters-Lidard, Chris…, 5/15/2017 4:31 PM
Deleted: 1

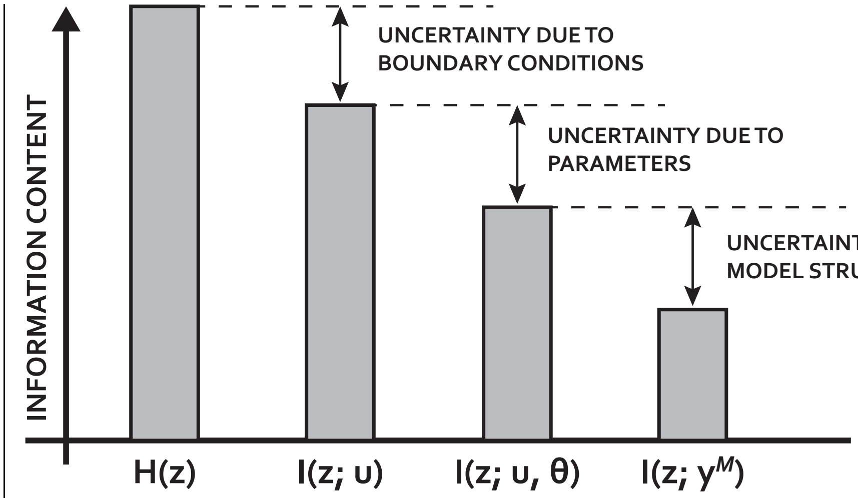


Figure 3: A conceptual diagram of uncertainty decomposition using Shannon information following Nearing et al., (2016). The term H(z) represents the total uncertainty (entropy) in the benchmark observations, and I(z; u) represents the amount of information about the benchmark observations that is available from the forcing data. Uncertainty due to forcing data is the difference between the total entropy and the information available in the forcing data. The information in the parameters plus forcing data is I(z; u), and I(z; u, θ)<I(z; u) because of errors in the parameters. The term I(z; y^M) is the total information available from the model, and I(z; y^M)<I(z; u, θ) because of model structural error.

Peters-Lidard, Chris…, 5/15/2017 4:31 PM
Deleted: Page Break

Peters-Lidard, Chris…, 5/15/2017 4:31 PM
Deleted: 2