Reviewer #4 (Marc Bierkens)

Comment:

I started to read this opinion paper with great anticipation because I think there is a desperate need for joining top-down and bottom-up approaches to arrive at solid hydrological theories. The paper is generally well written and starts out with a promising small review about the nature of bottom-up and top-down approaches.

Reply:

We highly appreciate this positive assessment.

Comment:

However, after reading the part thereafter, I have to admit I started to become a bit disappointed. The reason for this is that the second part of the paper becomes quite unbalanced and reads as an apologia for top-down modelling. What I miss is a section “Modelling myths or not” for bottom-up approaches. For example, statements as “Bottom up models are over-parameterized” can be elaborated on. After that I would have liked to have a section to sketch a way forward to marry both approaches taking account of their complementarities. Shortening the “Modelling myths or not” to make room for similar sections on bottom-up approaches would make the paper much more balanced and interesting.

Reply:

Reflecting also the points raised by Reviewer #1 and after re-analyzing the manuscript from the perspective of all reviewers, we fully agree that it comes across more like a defence of top-down models rather than the intended balanced evaluation of the two modelling strategies. We will accordingly re-structure and re-formulate the relevant sections in the revision by adding perspectives towards the bottom-up approach. We will also put more emphasis on how to take the best out of both approaches.

Having said that, and given that also the other reviewers noted that the paper should be less a defence of conceptual models, we would also like to stress one, potentially not irrelevant point: “Physically-based” models may largely benefit from a semantic-psychological bias. The term “physical-based” inherently implies that they are “correct” descriptions of real world-systems, which further implies that all other models are not “physical” and thus less “correct”. From this perspective, we believe that any type of comparison between model strategies will to some extend necessarily come across as a defence of conceptual models, i.e. explanations of why they can be as meaningful representations of reality as physically based models. In other words, already the term “physically-based” puts models in the (often not really justified) position of benchmarks other models have to be compared to, even if they are not necessarily “better” descriptions of reality.

Comment:

First, the authors underpin the statement that “At the macroscale, which in the realm of organized complexity is frequently characterized by the emergence of relatively simple functional relationships. . . that integrate typically unobservable natural heterogeneity over the model domain”, with a comparison with to statistical physics (e.g. gas laws). However, there is a big difference between an ideal gas and a hydrological system related to the assumption of ergodicity. In that context, this assumption loosely means that at all times all microstates are present when averaging over the volume. This assumption is valid for an ideal gas but not necessarily the case for hydrologic systems.

Reply:

We fully agree with the reviewer that there is no full correspondence between the two systems. We rather understand it as an analogy, i.e. a partial similarity, of the systems. In our understanding, the essential difference between the two systems is that a volume of an ideal gas is random and complex for it to be considered in the realm of unorganized complexity, where the microstates of large enough samples can be meaningfully characterized on the macroscale based on their statistical properties (ergodic system). In contrast, hydrologic systems are characterized by lower degrees of randomness and complexity and thereby fall in the realm of organized complexity, where systems cannot be fully described by statistics alone. We believe, that organisation in catchments is manifest in the structure of the hydrological response, which in turn is caused by the varying connectivity of processes acting on distinct time scales. In other words, depending on the wetness history and the “memory” of the system, any combination of these (statistically different) processes can be active at a given time. In spite of this overall structure (or organization), we think that the at least some of the individual processes may well approach the definition of ergodic processes (of course given the full knowledge of input, output and boundary conditions). An example may be the groundwater dynamics at the catchment scale: during low flow periods, the drainage of the “deep” ground water is the only processes sustaining stream flow (and due to the depth of the groundwater table only negligible evaporation is occurring) in many catchments world-wide. At the catchment-scale this emerges as exponential recession characteristics and thus suggests simple linear storage-discharge relationships. We think that it is not unreasonable to assume that a random samples of this process will reflect the statistical moments of the full process, which is the fundamental definition of an ergodic process. However, due to the reorganization of the manuscript we will remove the gas law analogy in the revised manuscript.

Comment:

Second, I feel that a problem with the way top-down megascopic hydrological laws are derived (also in comparative hydrology) is that often only (signatures) of the output variables are used to assess the form of the Q = F(S) relationship. This can only be done if a certain form (often a power function) is assumed a priori. I think that to really assess the form of these relationships one needs to jointly measure the state (groundwater storage, soil moisture, snow water equivalent) and the output variables (discharge, evaporation). Very rarely these state observations are used or available in catchments used in comparative hydrology. So we should get away from the fixation with hydrographs only and start measuring states. To add to this: energy conservation is often added by checking if the found megascopic laws follow Budyko's hypothesis. This is only a weak check on energy conservation, because it only checks for very long times and doesn't guarantee energy conservation at any given time.

Reply:

We also wholeheartedly agree with this comment and do not state otherwise in the manuscript. We would also take this point a step further and argue that the problem does not only apply to top-down models. Bottom-up models, based on extrapolations of anecdotal observations, are not unlikely to suffer from similar problems. Remote sensing products do have the potential to allow for real progress here (e.g. GRACE). Another point that is currently not fully exploited is the information content in spatial patterns. We think that systematically forcing (semi-)distributed models to produce good correlations with observed spatial pattern of, for example, soil moisture or snow cover will prove highly valuable to test models.

We also agree that the Budyko framework provides a models test, albeit a very weak one. However, the actual observed runoff coefficient may hold more information, as it cannot only be applied over the long-term, but models can also be trained to reproduce annual or even seasonal sequences of observed runoff coefficients. Doing this will strengthen the test on energy conservation (albeit not fully solving the problem, of course). We will clarify this in the manuscript.

Comment:

Third, once megascopic laws have been derived empirically, these laws' physical basis should be strengthened by also deriving them from upscaling from smaller-scale mechanics. A well-known example is Darcy's law. It was first established empirically - note that this was done by both observing states (heads or actually the head gradient) and fluxes. Later (much later), it was shown that it could be derived from the Navier-Stokes equations (by 1. neglecting quadratic inertia terms: laminar flow -> Stokes equations; 2. volume averaging by homogenization; 3. noting that drag forces are much larger than viscous forces). Obviously, heterogeneities in hillslopes and catchments are more complex than pore-scale heterogeneities in a REV. This makes simple homogenization not likely a suitable approach. However, hyper-resolution (cm-scale) modelling using simulated heterogeneities (including macropres etc) with 3D PDE-based models (e.g. Parflow, Hydrogeopshere, Cathy) and upscaling the results may be a way to derive megascopic laws from first principles.

Reply:

We agree with this. That was also the motivation behind the statement: "While top-down models approach the problem from a macroscale physical understanding, bottom-up models emphasize the microscale perspective. An ideal model would, almost needless to say, provide an equally good representation of both aspects." (p.13,l.19-21). We will clarify this make it more explicit in the revised manuscript.