Response to Review Comments regarding the manuscript "Large-Sample Hydrology: A Need to Balance Depth with Breadth" by HV Gupta, C Perrin, R Kumar, G Blöschl, M Clark, A Montanari and V Andréassian.

October 18, 2013

A. General Response: We are gratified that all of the reviewers found the manuscript worthy of publication, with no less than 25 individual remarks that it is a "*worthy contribution to the literature*" (B Jackson), "*well written and clear … recommend publication*" (H McMillan), a "*timely reminder … well presented and argued … significant contribution that will be widely read and cited*" (Anonymous), "w*ell written and may have high impact*" (D Wang), "*interesting and overall well-written*" (S Patil), and "th*e authors' plea is encouraging raising several issues worthy of attention*" (E Coopersmith). We thank the reviewers for their very constructive and thought-provoking comments, which will certainly help us to sharpen the message and improve the paper.

We are also encouraged by the fact that all of the reviewers felt that the concept of using large samples for hydrological investigation is a 'no-brainer'. While increasingly more work *is being done* in this regard (as illustrated by that of the reviewers), such work is still limited and by only a relatively sample of the community at large. We hope that this paper will help to encourage a greater focus on, and more intensive pursuit of, such studies on a regular basis.

We have provided our response to the Editors summary comments in a separate document. Based on that, and your detailed suggestions, our revision will certainly attempt to (a) Clarify the relationship to Panta Rhei, (b) Clarify what is new relative to the pub book, (c) Clarify the relation with comparative hydrology and include some reference to recent studies, (d) Better frame large-sample hydrology regionalization vs understanding, (e) Clarify that modeling and understanding are tightly coupled processes, and (d) Improve the sections dealing with information (and problems in collecting/reporting/synthesizing) various kinds of hydro-meteorological data.

B. Specific Responses to Constructive Comments provided by the Reviewers: *[BJ = B Jackson, HM = H McMillan, AN = Anonymous, DW = D Wang, EC = E Coopersmith, SP = S Patil, MO = M Ostrowski]*

Given the large number of reviewer comments (seven), in the following we have grouped and responded to similar comments to avoid repetition, and maintain a manageable length ☺.

(B1) Comments Regarding - Repetition/Length:

[BJ]: Is sometimes a *little repetitive* and could do with shortening / some reorganization to draw out the key points more effectively

[HM]: The authors provide a review of previous studies in Section 2.2 (some of which is repeated in Section 4.1; *perhaps these could be combined*).

[AN]: I found the manuscript repetitive in places and recommend the authors remove repeated material and concepts.
[EC]: The subsections discussing hydrologic processes and the need for classification are *somewhat redundant and again, omit numerous recent classification papers*.
[EC]: Section 5.4 also seems a bit redundant, but section 5.5 is a very important point that requires substantial attention.
[EC]: As an opinion paper (and not a research article), the length seems a bit unwieldy, especially as numerous arguments therein are either revisited, expanded upon at length or repeated outright. Considering that there are likely to be several papers in the years to come that cite a paper like this one (as it does attempt to encapsulate the highlights of many decades of hydrologic work), *it is important that this work incorporate the comparative hydrological work of the past five-to-ten years*. I would advocate that this work, ideally, *should be revised to 50-75% of its current length*, address some of the newer papers that focus on signatures and/or hydrologic classification, and avoid repeating arguments (however persuasive they may be!).

[EC]: Between pg. 9152, line 16 and pg. 9153, line 10, the authors discuss Linsley (1982) and Klemes (1986). *These arguments largely rehash what is stated when these two papers are cited within the introduction*. Unlike a standard research paper in which an idea is mentioned briefly in an introductory section, then expounded upon as quantitative analyses are performed, this paper is an opinion piece (and a somewhat long one at that). *Passages like the one described above are somewhat redundant and should be shortened or removed*.

Response: Four of the seven reviewers remarked that the manuscript could be shortened somewhat by removal of 'repetition', and that an attempt should be made to '*draw out the key points more effectively*'. We will certainly take this into account during the revision process. One thing we can do is to give less detail in the introduction so as to reduce repetition. Another thing that will help is to place the information on data sets in sections 2.2 & 4.1 into a table (giving the name of the dataset, URL, reference, comment, etc.), which would avoid repetition in the text, as suggested by the reviewer (HM).

However, we disagree about the redundancy of section 5.4, or the need to discuss hydrological processes and the need for classification. While the reviewers may be especially familiar with these concepts, the broader readership needs to be kept in mind. Further, we suspect that *the impression of length may come partly from repetitions in the text; we will* try to limit them so that the manuscript becomes easier to read. Meanwhile, we agree about the need to better clarify the important recent contributions being made by the field of comparative hydrology.

(B2) Comments Regarding Figures:

[HM]: I don't think the figures are very informative, apart from fig. 4.
***[AN]*:** I found that figures 2,3,4 and 5 added very little to the argument and were not discussed by the authors. I recommend removing these – they add little

insight to the argument.

Response: Two of the seven reviewers were not impressed by the figures. We note that the addition of figures was based on a request from the editor. Figures 1 and 2 have not been published elsewhere and can be better integrated. We agree that Figures 3 and 5 can be omitted. The reviewers are mixed on Figure 4 – our own view is that it helps to clarify the points made in the text and should be retained. Further, to support the data related arguments, we are considering including a new figure consisting of a map illustrating locations of basins used in studies cited in the appendix.

(B3) Comments Regarding Content, Scope and Language:

[BJ]: *Over-stresses the point that many hydrologists are concerned with "depth"* (i.e. understanding certain sites in detail), particularly given that the paper recognizes it is important that some studies continue down this line. There is already a rapidly increasing number of studies and researchers concerning themselves with the "breadth" issue, as also noted within the reviews by S Patil and H McMillan. *The importance of this paper is more to help this already emerging community by highlighting the broad context and the varied ways in which these studies may contribute to a broad range of hydrological challenges, as well as ways forward in establishing data protocols and calling for wider availability of good quality datasets*.

[SP]: I found some of the views expressed in section 1.2 "The context of current practice" to be a bit outdated. The main argument of the authors in this section is that catchment hydrology is still focused on "depth" (studying more and more about a handful of places) rather than "breadth" (including large number of places in a single study). *I'm not sure if this accurately represents the current state of the field*. *Many of us from the younger generation of hydrologists have fully embraced the message that large sample hydrology research is extremely valuable, partly thanks to the successful messaging of the PUB initiative* [Carrillo et al., 2011; Sawicz et al., 2011; Cheng et al., 2012; Coopersmith et al., 2012; Drummond et al., 2012; Kelleher et al., 2012; Patil and Stieglitz, 2012a; Patil and Stieglitz, 2012b; Yaeger et al., 2012; Ye et al., 2012; González-Pinzón et al., 2013; Safeeq et al., 2013]. *Note that the large sample hydrology papers cited in the previous sentence are: (1) in addition to the 84 papers that the authors have mentioned in their supplementary material, (2) they have been published within the last three years and by an early career scientist as first author, and (3) they are not restricted to rainfall-runoff modeling alone*. Therefore, large sample hydrology research is already being conducted in a big way since the last few years. *The question then becomes, how can the new IAHS initiative help channelize these diverse efforts towards a greater "decadal" goal? The authors might want to consider adding some discussion in this section that gives credit to the current (albeit diverse) efforts at large sample hydrology and also clarifies their larger goal*.

Response: While we agree that a key contribution of this paper is to encourage and

promote increasing work on large sample studies, and the varied ways these can contribute to a broad range of hydrological challenges, we disagree that there is not a need to stress the fact that hydrologists tend to be concerned more with depth than with breadth. Certainly there are increasing numbers of studies (more recently) concerned with the '*breadth*' issue, but these represent a rather small sub-set of the community at large where, given the engineering roots of our discipline, '*place-based*' science is still the norm. Arguably, PUB and the new Panta Rhei initiatives have and are helping hydrology mature into a true scientific discipline, but much progress still needs to be made. We will certainly improve citation to the recent and diverse literature mentioned by the reviewers. *More important however, is the need to better address how the new IAHS initiative can help channelize these diverse efforts towards a greater "decadal" goal*. Thanks for this excellent suggestion.

[BJ]: *Found it distracting when bits of history/quotes from researchers were inserted in the sections other than section 2*, as it broke up the discussion of the scientific issues. Often these quotes added little to the paper, with their content already having been clearly made earlier in the manuscript. I would suggest going through all quotes and historical mentions in the manuscript; where the authors feel they are important they could be moved to the narrative in section 2, where not the researchers could be cited/briefly mentioned. For example, I found the introduction opening with a quote from a paper, and then jumping to a more standard introduction without a direct flow through to comment on the quote strange/jarring. I don't think the paper would lose anything from the quote being removed; the content of the quote is fairly accepted; the issue you are raising in this paper is that such testing is not properly pursued. If the quote is retained I suggest you make that point directly following the quote.

[MO]: Scientific hydrology is not connected with myths, mysteries or sagas. Expressions like '*holy grail*' as stated in this manuscript or in other papers are not very helpful …

Response: Thanks for these suggestions. However, we suppose these are largely issues of style. We ourselves like the use of quotes in which the authors have expressed certain ideas very well and clearly. Certainly, we will bear this suggestion in mind as we revise the manuscript. Meanwhile, we ourselves do find it useful to think in terms of lofty goals, unattainable though they might be ☺.

[SP]: This paper primarily focuses on rainfall-runoff modeling issues (model structure, uncertainty, etc.) and how large sample hydrology can help resolve them. However, I believe that *the scope of large sample hydrology research is (and has to be) larger than rainfall-runoff modeling alone*, especially if we are to make any meaningful progress towards the "Holy Grail" of comprehensive hydrologic process understanding. Results from some recent studies suggest that synthesis of different types of data, such as numerous stream tracer experiments [Drummond et al., 2012; González-Pinzón et al., 2013] or flood events [Gaál et al., 2012], can provide equally valuable insights into catchment processes as any large sample modeling study. Most importantly, *the diversity of efforts in large sample hydrology research increases the likelihood of identifying gaps in our*

current theories, models, and measurement techniques. I think the authors have a great opportunity to call for increasing the diversity of efforts in large sample hydrology (beyond modeling issues) in the closing sections of the paper.

Response: Thanks for this comment. We agree that the scope of LSH is larger than rainfall-runoff modeling alone, and that the increasing diversity in LSH studies will increase the likelihood of identifying gaps in our current theories, models and measurement techniques. Given that the authorship of this manuscript is dominated mainly by modelers, it is perhaps inevitable that our 'opinion' will tend to have a modeling slant. Note however that, as expressed in our response to the Editor, we view modeling and understanding as tightly coupled processes, arguably impossible to address independently.

Thanks for the suggestion to include a call for increasing the diversity of efforts in large sample hydrology (beyond modeling issues) in the closing sections of the paper. We will certainly do so.

Note however that, while we will certainly enhance citation to the studies pointed out by the reviewers, we do not feel it is relevant here to extend the appendix to studies that do not focus on rainfall-runoff modeling. This paper is not intended as a comprehensive review, and we will leave that effort to some enterprising young researcher. Instead, we can reword slightly to present the table as an example dealing with modeling applications within hydrology, and will acknowledge in the text that other approaches such flood studies have also contributed to LSH. Certainly we will also augment the table to include the rainfall-runoff modeling studies cited by the reviewers that we had missed.

[SP]: *The call for classification made in this paper is not new*. However, the authors did not acknowledge that a lot of work has already been conducted over the last few years in the fields of catchment, stream, and landscape classification (all of which are highly relevant to large sample hydrology research)[Wolock et al., 2004; Kennard et al., 2010; Mücher et al., 2010; Poff et al., 2010; Ley et al., 2011; Reidy Liermann et al., 2011; Sawicz et al., 2011; Coopersmith et al., 2012; Olden et al., 2012; Wigington et al., 2012]. *It would be much more informative to the readers if the authors mention (and briefly summarize if possible) the classification-related research that already exists, and then discuss their views on what is still missing, whether previous studies have been wrong (or incomplete) in their assumptions, and what further steps should be taken to make progress in classification*. *Simply describing how a classification system will provide insight into a lot of unresolved issues is not enough* in my opinion. That has already been mentioned in older opinion articles [McDonnell and Woods, 2004; Blöschl, 2006; McDonnell et al., 2007].

Response: Thanks for pointing that out. We agree that it would be informative to briefly summarize the classification-related research that already exists, discuss what is still missing, and what further steps should be taken to make progress in classification. Note however, this is not primarily a paper on classification – as pointed out other authors have addressed that issue (see also the 2013 PUB book).

We will clarify this issue (misunderstanding) while revising our paper.

[MO]: A modified more suitable title could be: *Integrated-sample hydrology: a need to combine depth with breadth*. The present title builds an artificial barrier between large sample and individual detailed studies, which does and should not exist.

Response: Thanks for the suggestion to change the title. However, we prefer our existing title. Certainly we will strengthen the text in regards to the need to increase diversity in data types. However, we will leave the paper on "*Integrated Sample Hydrology*" for someone else to write ☺.

(B4) Comments Regarding Historical Perspective:

[HM]: *It would be worth discussing here some of the more recent initiatives*, such as Environmental Observatories in various guises: e.g. TERENO (http://teodoor.icg.kfa- juelich.de/overview-en?set_language=en), EVO-UK (http://www.evo-uk.org/), S – HYPE (http://www.smhi.se/en/Research/Research-departments/Hydrology/hype-in-sweden-s-hype-1.7891).

[SP]: In Section 2.2 (P9155), the authors mention MOPEX and DMIP as examples of large sample catchment datasets in the US. I think it would also be worth mentioning here the tremendous efforts of the USGS in developing large catchment datasets for research; first the HCDN database (1659 catchments)[Slack et al., 1993] developed in the nineties and its recent update called GAGES (6785 catchments)[Falcone et al., 2010].

[EC]: Pg. 9149, line 19 argues, "*what is needed is to begin taking advantage of extensive datasets...*" However, given the multiple works published in HESS in the fall of 2012 (Cheng et al., Ye et al., Coopersmith et al., and Yaeger et al.) attempting to draw broader insights from 400+ diverse catchments within the MOPEX database, this comment seems to ignore the work of numerous younger hydrologists over the past few years.

[DW]: In recent years, some *comparative hydrology studies* based on MOPEX catchments may be included in the revision as pointed by other reviewers.

[AN]: By focusing on LSH in the context of hydrological models, the authors missed the significant literature related to Comparative Hydrology. Although they refer to Falkenmark and Chapman (1989), the reference is toward the end of the manuscript and not in any way central to their overall argument. In my opinion the argument presented for LSH is very similar to the purpose of CH [*more detail provided in the comment*]. Since the authors seek to '*promote a potentially important theme for the upcoming IAHS scientific decade …',* this manuscript is likely to frame this theme and be a significant contribution that will be widely read and cited. Therefore it is incumbent upon the authors to revise the manuscript to fully reflect the history of this topic and to acknowledge the contributions of CH to their argument. To this end, it would benefit from

broadening the discussion of LSH beyond hydrological modeling [*more detail provided in the comment*].

Response: Thanks for pointing this out. We will certainly mention these initiatives and studies in our revision, and improve the discussion in regards to Comparative Hydrology. Such efforts are indeed essential to ongoing research. Various countries are already contributing in this regard (UK, Australia, France, etc.), at least in part due to the efforts and activities of PUB.

[EC]: Section 2.2, a pseudo-literature review, *either requires substantial additions, or more likely, a shortening and relocation towards the introduction*. It lists a variety of modeling studies from the 1980s, then mentions the PUB initiative, but *skips the various classification studies that occur between them*. Mentioning MOPEX, DMIP, and NLDAS as potential national-scale data sources is important in terms of their facilitation of large-scale studies (though they are biased perhaps in terms of their focus upon those regions of the globe that are more affluent and more densely populated (North America, Western Europe, etc).

Response: As mentioned earlier, we do not propose to conduct an exhaustive review, but rather to provide references to a subset of LSH studies dealing with rainfall-runoff modeling. We will revise the texts to make this clear. Note, however, that we do point the reader to a recent comprehensive PUB synthesis report by *Blöschl et al., 2013* for more details (see 1st paragraph, section 2.2) Further, we will change the title of sub-section 2.2 to be more in line with the content.

[EC]: The authors rightly describe the prevailing trends to eschew "data fitting" (pg. 9156, line 18) in favor of understanding developed by signatures. However, as mentioned previously, the works published in HESS this past fall do exactly this, focusing on the flow duration curve (Cheng et al), the annual regime curve (Ye et al; Coopersmith et al), and the hydrologic relationship between the two (Coopersmith et al; Yaeger et al).

Response: Yes, this trend is certainly true. However, as discussed by Gupta et al (2008) and others (see also 2013 PUB report), flow duration curves and annual regime curves are just a small step in the direction of informative diagnostics, and certainly not a recent invention. There is a clear need to think more creatively regarding the issue of diagnostics in hydrological analysis, so as to better bridge the gap between data and models, and to bring insight and understanding to bear on model evaluation.

[MO]: The paper does not describe the state of the art and practice in hydrology sufficiently well. *The authors ignore the considerable amount of relevant gray literature* and the *even larger amount of catchment data sets from applied sciences and operational practice*, which should at least be considered in a discussion paper, only to mention the NOAA Science plan, 2007. The problem of intentious information loss due to the ignorance of gray literature was recently described in detail by Uhlemann et al (2013). The statistics presented and the conclusions drawn in that paper seem to be representative for the modeling context also. *The dominant self-restriction of scientific hydrology to commercial*

literature leads to a loss of about sixty to seventy percent of the relevant information. Thus, it seems rather questionable whether a concept based on such a limited information base can be relevant for the future of hydrologic research, if it is also aiming at positive practical impacts. However, it can be expected that open access and other initiatives will change the transparency of discussions significantly and might support the idea of large sample hydrology significantly.

Response: Thanks for this excellent suggestion. We will certainly include mention of the vast body of grey literature.

(B5) Comments Regarding Data Issues and Meta-Data:

[BJ]: Takes a very model-centric approach. But *I would argue that progress will be faster if most "large sample studies" also include some focus on how data availability/requirements vary between regions, and where and how the models are limited by data*. *A major issue* that will impede progress in refining models, better understanding process, and improving our predictions/reducing uncertainty *is the disparity of data (amount, type, quality, resolution, etc) from region to region*. I do realize section 4 is partly a call for standardization of data, but there will still be regional issues, and issues with interpreting historical datasets even when/if appropriate protocols do become fairly widely taken up.

[SP]: I wonder if we have a *problem of "latitude sampling"* in catchment hydrology. For whatever socioeconomic and historical reasons, majority of the developed world (OECD countries) is located in the temperate latitudes. Not surprisingly, most of the large sample data also comes from the developed countries (as evidenced by the table shown in the supplementary material of this paper). Although a lot of climatic and physiographic variability exists within and among these countries, *can we really move towards a comprehensive hydrologic understanding without having much more large sample data from the tropical regions*? *It would be helpful if the authors can add some perspective on this issue in the paper*.

Response: These comments are very pertinent. Certainly there is a long way to go before datasets will answer all of the needs and regional specificities. In the revision we will include the point that even with standardization, there will be regional issues, and issues with interpreting historical datasets, and that progress can be enhanced if LSH studies include some discussion of data availability/requirements specific to the various regions, and where and how the models are limited by data. As mentioned below, IAHS could play a role in this, and the need to maintain/develop observational networks could be stressed.

[HM]: I was not completely convinced by the arguments in Section 3.2 that large samples help to reduce the impact of data errors. When comparing or combining large numbers of different datasets, *the need for characterization of the uncertainty in the different datasets becomes particularly strong*. Different data from different parts of the world may have been collected using very different techniques and with *very different quality control*. Without uncertainty

information, *comparisons of processes or model performances between different locations may be incorrect or biased*. I would be interested in the authors' comments on *to what extent it is ever going to be possible to combine very diverse data in terms of temporal/spatial scales, uncertainties and collection methods*.

Response: This is an excellent point. While the point about impact of large samples on data errors presumes the tendency of random variation in error types across catchments, this will have to be demonstrated by actual investigation. Meanwhile we can certainly stress the increased need for metadata when using (archiving) large samples with diverse origins. Regarding the question, we think that the use of modeling as a tool (and, of course, comparative studies) provides at least one way to '*combine very diverse data in terms of temporal/spatial scales, uncertainties and collection methods*'; i.e., by diagnostically improving models via LSH, we gain a repository of knowledge that can potentially cut across data diversity of the kind mentioned.

[AN]: Another useful contribution would be a summary of freely available large samples of catchment data that could form the basis of LSH (perhaps as an appendix). This summary should not be based solely on data sets related to hydrological modeling [*more details provided in the comment*].

[DW]: As discussed in section 5.1, there are barriers to the free exchange of hydrological data due to economic and legal issues with the public institutions who compiling and managing the data. *Individual researchers and local agencies might be able to complement, besides federal institutions, the data sharing tasks*. For example, a public website can be created and managed by IAHS or others so that individual researchers can upload and share their catchment data. The challenges for this include the required variables (such as precipitation, runoff etc.) and the time period of observations, the consistence of the data format, meta-data with the same format and others. If the data requirement is daily precipitation, runoff, temperature and potential evaporation like MOPEX catchments, a lot of catchments around the world meet this data requirement.

Response: Thanks for these suggestions. As we are sure the reviewers are aware, the compilation of data sets is easier said than done. That said, one of our colleagues has expressed interest in attempting to gather a '*super-sample*' of such data sets, and we will certainly give that effort all of our support.

Meanwhile, it is true that an association such as IAHS could play a key role in disseminating information regarding data sets, perhaps by creating a hydrological data portal with description of data sets and links, in a similar way to existing model inventories.

[EC]: In the section on practical challenges, the need for data availability is well described, *though perhaps some redundancy occurs in the reintroduction of MOPEX and DMIP*. One might point towards NASA's upcoming SMAP (remotely sensed soil moisture) or GPCP (global precipitation) missions as potential sources of widely available data for hydrologists in the years to come.

[EC]: The issues associated with soil texture and topography are appropriately

mentioned, though perhaps the potential solutions should be discussed in terms of national soil databases (SCAN, SoilWeb from the USDA, etc) and elevation datasets becoming increasingly ubiquitous through LiDAR data, etc.

[EC]: With regards to the specifics of the subsection associated with perspectives, the first section discusses barriers to sharing data. *This is a real concern, but seems more aligned with the previous section's discussion of practical challenges*.

Response: Thanks for these suggestions. As proposed above, we will introduce a table to provide this kind of synthetic information while reducing repetition/redundancy. The issue of discussing soils databases seems a bit beyond the scope of this manuscript. Regarding the overlap between subsections, we can move some of the information in section 5.1 to section 4.1, so that section 5.1 focuses more on 'perspectives' ☺.

[MO]: *If a large sample should be built up, much more information must be contained in their documentation than supplied e.g. by MOPEX*. As one example not the runoff data, but the water levels and related stage discharge relationships including upstream and downstream boundary conditions must become available. *The detail of documentation must be agreed upon à priori*.

[MO]: *Meta Data:* The elementary rule in my teaching of hydrological modeling approaches was: *Never model a catchment which you have not seen personally during an intensive field visit*! *To obey this rule would be hardly possible for a large sample*. However, digital means are available today to distribute far wider information such as online maps, photographs and videos, as partly available for the DMIP-II, but higher density and resolution should be provided. *The paper could be more specific, even being an opinion paper*. *E.g. the DMIP-II results (Smith et al, 2012) and the Strategic Science Plan NOAA, 2007 contain clear information and define directions for future scientific and development strategies; it seems that they were insufficiently discussed*.

Response: Thanks for these suggestions. Yes, this issue of metadata should be more clearly discussed.

[MO]: *In contrast to the regulations of most commercial publishers including HESS the data sets (raw and modified data) do rarely become publicly available (MOPEX and DMIP are positive exceptions, with some restrictions) nor are they stored in infinite repositories. In this respect, commercial publishing is quite contradictive. This is addressed in the manuscript but clearer proposals for the improvement of the situation are missing*.

[MO]: A further key aspect of the manuscript is the supplement containing a selection of commercially published large sample studies. *The data resolution is either one day or one hour. Daily data, however, is completely insufficient to analyze hydrological processes in a physically meaningful way*. Daily data is just sufficient to estimate reliable water balances. If a specific model contains physically meaningful parameters and equations, they degenerate to empirical functions and indices with increasing time step. Again this is well known in

practice (obviously less in science) and the general time resolution e.g. in Germany for hydrological modeling is one hour or smaller (30-15 minutes, in urbanized catchments 5 minutes). *A basic rule applied is that the time step must be shorter than one third of the time of concentration of the smallest sub-catchment*. The problem was demonstrated and explained by Ostrowski et al., 2010. Only for very short time steps and HRU resolution in the range of several hectares realistic and physically meaningful simulations and parameters can be expected.

Response: Data sufficiency and access are endemic problems in science. We do not have the 'golden' answer here, and concerted efforts must be made across our discipline to make gradual progress. One helpful thing is that some journals now offer sections to publish data sets and/or studies with the corresponding data online (e.g., WRR). Regarding sampling frequency, there is no reason why we cannot do good science and learn a lot about catchment hydrology even with daily data. Of course, a major reason for the widespread use of daily and hourly time steps in modeling studies is data availability. Obviously, we agree that the time step should be chosen in agreement with the objectives of the study and the time dynamics of processes under investigation.

[MO]: A final and severe concern is the assessment of data quality. Andréssian et al, 2009 try to convince the community of the high quality of the data in large catchment samples without any proof. With experience in science and teaching, water administration and practice, I come to the conclusion that this is simply not true; it is wishful thinking! Applying a model we have good reasons to mistrust any incoming data and to check it thoroughly. This is not at all meant to criticize organizations supplying this data. They do the best work possible under given resources limitations, but standard methods have their limits. Frequently, data assimilation and their control in practice take more than 60% of project duration; this is expensive. And: The model applied is an important if not the dominant tool itself to test the data for homogeneity and consistency. *Again, the manuscript addresses the point but should say more about the required procedures for data quality assurance*.

Response: The experience of at least two of our co-authors (Andréassian and Perrin) is that the percentage of catchments with serious data issues in their large sample for France is surprisingly small, and that it does not generally impact the conclusion of the studies due to the errors being drowned within a larger sample of somewhat "good" data (but this is gray literature in French!). Obviously, we do not suggest that compiling and using large samples absolves us of the need to attend to data quality. It is true that data access and quality control can take as much as 60-90% of the time and effort – this has generally been true ever since we can remember, and the compilation of large samples only aggravates the problem. We do not think we can say much more about quality assurance procedures than has already been said exhaustively by others.

(B6) Comments Regarding Modeling Issues:

[HM] *Global hydrological models get a very brief mention, but they may soon be a major user of large hydrological datasets*. There is an interesting discussion in Widen–Nilsson et al (2007) on the data challenges involved in applying and validating a global water–balance model. *Some global runoff products are already available*, e.g. Global Runoff Data Centre (http:/grdc.bafg.de; and an application by Fekete et al (2002)).

Response: Thanks for this comment. We can include reference to this interesting study. During the revision, we will certainly mention the effort of GRDC repository system in a new table.

(B7) Comments Regarding Expanding the Discussion:

[DW]: The two sections, '*Linking large sample studies with process hydrology*' (Section 5.2) & '*Trading depth of analysis for sample size*' (Section 5.6) are related and very important parts of this paper. *These two sections may deserve more substantial discussions*.

"Balancing depth with breadth" is one of important points advocated in this paper. Every catchment is unique like individual person; but catchments have some common characteristics like human beings. Individual detailed catchment analysis is helpful for us to understand the physical process of "tree"; large sample hydrology is helpful for us to see the "forest". The understanding of common characteristics will benefit to improve our understanding on the behavior of local catchments; detailed study on experimental catchments can be used to understand the common characteristics in the processes level.

Then, the challenging *question is how to balance (or integrate) large sample study and specific catchment study to improve hydrologic understanding*. *What are the strength and weakness of "large-sample" approach and "a limited number of heavily instrumented catchments", respectively*? How to reconcile or integrate these two approaches to improve understanding of catchment behaviors? These two approaches can mutually benefit to each other.

In the study of individual catchment, observations at the finer spatial and temporal scales are available and more state and flux variables are observed such as soil moisture and groundwater table. While in the large sample study, there may be limited variables, which are available in all the sample catchments. *Once a hypothesis has been tested in a heavily instrumented catchment, the hypothesis can be generalized and tested in other catchments through large sample study*. The challenge is that the required observations to test the hypothesis in these catchments are usually not available. *Some thoughts and discussions along this line (how to integrate or reconcile the two approaches) may be highly appreciated by readers*.

[DW]: Trading space for time (Section 5.5). This definitely is one of the benefits from large sample hydrology. The underlying assumption for this method is that the catchment system including parameters will co-evolve with the changed

climate in a certain path. Is there only one possible state for the catchment evolving to or are there multiple possible states? *If it is the latter case, the uncertainty associated with the traded parameters may need to be quantified or other constraints besides climate need to be considered for projecting catchment changes*. *More references can be cited such as Sivapalan et al., (2011, Functional model of water balance variability at the catchment scale, 1: evidence of hydrologic similarity and space-time symmetry).*

[EC]: The perspective provided in the introduction argues (pg. 9149, lines 23-24) that "*the context of much current hydrological practice is a focus on depth rather than breadth*" continuing to discuss the development of generalized models where parameters are specified as a function of location. In turn, this leads to the "uniqueness of place" reference from Beven (2000) and the challenges associated with generalization. *This perspective seems to caution against the idea that one conceptual framework can be applied at a diversity of locations, though this is, to some extent, the "holy grail" of which the authors speak*. The authors rightly note that this challenge is increased when additional features are incorporated in the name of realism, *but fail to mention that it is likely that the "holy grail," were it ever to be found, is likely to be a relatively simple model of the type they warn against (one we try to apply everywhere)*.

Response: These comments make the following suggestions:

(a) To increase discussion regarding sections 5.2 and 5.6 ('*Linking large sample studies with process hydrology*' & '*Trading depth of analysis for sample size*') and discuss strengths and weakness (and how to reconcile) large sample approaches with limited number of heavily instrumented catchments [DW]. As pointed out by [DW], once a hypothesis has been tested in a heavily instrumented catchment, it can be generalized and tested on larger samples. Yes, the challenge of data availability exists. We are not sure what we can really add to this.

(b) To mention the benefits of trading space for time [DW]. As commented by the reviewer, this has already been mentioned in Section 5.5, and we are not sure what more to add to comments already made by other authors such as Sivapalan et al., (2011). Certainly we will endeavor to discuss how LSH can contribute to the Panta Rhei initiative.

(c) That we caution against the idea of applying one conceptual framework everywhere [EC]. We are not sure that we agree that arriving at a single "relatively simple model" is really the "holy grail" of hydrology, or that we are suggesting here. There appear to be diverse points of view on this issue, even among the authors. However, we do suggest that one possible strategy is to attempt to do this (mentioned in the last two paragraphs of section 5.2) as a vehicle for learning – one can only learn by making assumptions and then testing them. More likely, in keeping with ideas of catchment classification, various sub-classes of dominant process models will hold sway. However, each must be tested thoroughly against large

samples, and the reasons for differences between them (and when and where they can be applied) must be well understood.

(B8) Comments Regarding Recommendations:

> ***[HM]: Recommendations for Enabling LSH -*** It would be *good to hear from the authors their recommendations for enabling large sample hydrology in the future*. *Should we/they be thinking big, to initiate a global database of catchment hydrology information*? For example, in New Zealand we have a database which assigns every river reach over a given catchment area a unique identifier, and allows us to associate further information with the reach – should we be aiming for something similar on a global scale, which holds all our climate and hydrometric data and which could be queried by any hydrologist? Should we be trying to combine some of the existing datasets that the authors discuss?

Response: Thanks. We will make this clearer in the revision. Certainly we believe that a global database of catchment hydrology information is a worthy goal, and both IAHS and WMO could play leading roles in progress towards this. Of course, this will take time and require that numerous political and practical problems be solved – such who will host the database, bear the expense, and set the protocols. This must clearly be an international effort, and Panta Rhei can help promote it. In the meantime, as mentioned above, one of our colleagues is eager to attempt to compile a '*super-sample'* of such data sets, and we will appreciate assistance in this matter.

(B9) Specific Comments:

Response: Thanks to all the reviewers for the quite numerous specific comments. These include the fact that a catchment classification system need not be simple [HM], missing citations [AN], problems of ill-posed ness in modeling [MO], connections to Panta Rhei [MO], building transdisciplinary links [MO], and accounting for human influences [MO]. We will carefully go through these and absorb them as relevant. In regards to transdisciplinary studies, several of the authors are heavily engaged in such efforts, particularly to bridge across the social science and engineering disciplines. The need to build complementary data bases of human influences will certainly become apparent as Panta Rhei progresses.