The manuscript discusses the role of expert knowledge in hydrological modelling, contrasts several perspectives described as "modeller" and "experimentalist", and provides a simple case study illustrating the discussion.

I generally enjoyed reading the manuscript, it is written in an easy-to-follow style, and does not oversell its results. The case study is presented in a fairly succinct way. Overall I believe the authors arguments are valid and of interest to the community, even if they are not necessarily new. However there are several important instances with questionable gaps in the logic and some critical parts of the method description seem to be missing.

Reply: We would like to thank the anonymous reviewer for his comments and suggestions. We acknowledge that some of our arguments might not be seen as new and may already have been enunciated by other authors. However, we found exclusively review papers (e.g. Clark et al., 2017) or opinion papers (e.g. Hrachowitz and Clark, 2017) on the topic. Our contribution is therefore intended to be one of the first hands-on exercise on the coupling of bottom-up and top-down thinking, with a specific regard to DRP-based modelling approaches. In the following, we replied to each point raised by the reviewer and indicated how we will implement the suggestions in the revised manuscript.

An immediate question I had when reading the manuscript is that the term "expert knowledge" does not appear to be defined / explained. Should it be contrasted to some other kind of "general" knowledge (or "alternative" knowledge?). As a result it is unclear what type of information/data do the authors consider to be "expert knowledge", and which they do not, and the discussion that follows then becomes a bit too vague. For example, is the classification into top-down vs bottom-up considered expert knowledge? I would have argued this would nowdays represent "general" knowledge rather than "expert" knowledge, and so forth.

Reply: With the term “expert knowledge” we refer to someone’s acquaintance with hydrological sciences as a result of study and experience. In this context, the term “expert” should be understood as “relating to a person who has knowledge in a particular field” (WordReference Random House Learner's Dictionary of American English © 2017) and it should not be contrasted to other kinds of knowledge. The expression “expert knowledge” was already used by several authors: Gharari et al. (2014), Hrachowitz et al. (2014), and Safavi et al. (2015) used it in the title, whereas e.g. Antonetti et al. (2016), Bahremand (2016), Hellebrand et al. (2011), Hümann and Müller (2013), Nijzink et al. (2016), Peschke et al. (1999), Sivakumar (2004) and Smith et al. (2013) refer to it in their manuscript. In the revised manuscript, we will therefore keep the expression as it is for consistency.

As acquaintance is uncountable by definition, information to be considered as “expert knowledge” cannot be listed. Of course, the expert knowledge of a hydrologist with 30 years of experience cannot be compared with that of a PhD student. In the same way, expert knowledge of a modeller will differ from that of an experimentalist. What we tried to do in this study is to compare *different ways* of implementing expert knowledge in hydrological conceptual models, following either a top-down or a bottom-up thinking. In the revised manuscript we will therefore add a definition of „expert knowledge“ at the beginning of the introductory section and will distinguish in a clearer way between (i) expert knowledge per se and (ii) (top-down or bottom-up) strategies for the use of expert knowledge in conceptual models.

The case study description should also more clearly categorise and describe the types of "expert knowledge" used. Some of the expert knowledge seems to be just a particular data set (which is in principle accessible to "non-expert" modellers also). Perhaps I missed something here, in which case a clarification in the manuscript is still needed to prevent such confusion for other readers.

Reply: As mentioned in the previous reply, we would not refer to *types* of expert knowledge but rather to *ways* of using it during the three steps of the process-based modelling step (i.e. process mapping, process parameterisation, and parameter allocation). With regard to process mapping, the mapping approach of Schmocker-Fackel et al. (2007) and Margreth et al. (2010) is the one involving the largest amount of expert knowledge, which is used for manually mapping the small test areas (pg. 5, line 13) and setting-up the mapping algorithm. In contrast, the two other mapping approaches used for the study - Gharari et al.'s (2011) approach based on topographical information only, and Müller et al.'s (2009) approach based on DTM, geological and land use maps - barely rely on expert knowledge, which was used by the corresponding authors exclusively for verifying the classification criteria. One of the research questions addressed in the manuscript is therefore whether the use of more expert knowledge during the mapping phase improves hydrological simulations given unavoidable uncertainties in the forcing data (pg. 4, line 8).

We will try to clarify this in the revised manuscript as requested by the reviewer.

The manuscript makes an overly general statement is when suggesting "modellers, or 'dry' hydrologists" tend to develop theories at the catchment scale. Surely this would not apply to hydrologists that use "physically-based" models, which are classic bottom-up approaches. And the contrast to "experimentalists, or 'wet' hydrologists" is itself questionable - surely some of them work at the small scale and others extrapolate to the larger scale. Some additional supporting references here are needed. Or maybe clarify that "modellers" here mean "conceptual" modellers (?)

Reply: To our knowledge, the two appellations *modellers* and *experimentalists* were introduced by Seibert and McDonnell (2002) in their study on the use of soft data during model building and calibration. At the end of the paragraph in question we therefore cited their work. Some modellers will certainly develop physically-based models, as well as some experimentalists will focus on the regionalisation of small scale outcomes. In the revised manuscript we will soften the statement in question and add supporting references.

The ANOVA analysis used to estimate dominant source of uncertainty is a worthwhile undertaking, but the description here lacks essential detail. For example, how are the individual terms in equation 6 calculated? Clearly some assumptions need to be made here, eg, about the errors in the rainfall inputs (eg, Renard et al WRR2011), about the errors in the maps, and so on - how are the assumptions made and how are they checked in this study? Otherwise the conclusions of this analysis would not really be justified.

Reply: With the analysis of variance (ANOVA) it is possible to quantify the relative impact of source of uncertainty, named factors, on a given response variable, in our case the Kling-Gupta-Efficiency, based on the assumption that the output variance can describe the uncertainty of an environmental system (von Storch and Zwiers, 1999). In addition, the approach assumes independence between the different levels (e.g., between the different process maps). This assumption is not fulfilled, as e.g. three datasets use the same forcing (i.e. Combiprecip), and the two model setups are based on the same runoff generation module (i.e. RGM-PRO).

The method is clearly less complex than the Bayesian total error analysis (BATEA) method suggested by the reviewer, and others, but has advantages in terms of robustness (Tang et al., 2007), and is easy to implement thanks to the R package *stats* (version 3.2.2; R Core Team) where a function for the ANOVA is built in.

For the calculation of the individual terms in equation 6 we refer to equations from 18 to 22 of Bosshard et al. (2013). However, we believe that a longer description of the method would distract the reader from the objective of the study. It should be also noticed that the length and the amount of information contained in the paragraph where ANOVA is described, is comparable to that of other studies relying on ANOVA for quantifying uncertainty sources (e.g. Addor et al., 2014; Finger et al., 2012; Köplin et al., 2013; Rössler et al., 2012).

In the revised manuscript we will therefore add:

- A reference to a comprehensive description of the ANOVA method, e.g. von Storch and Zwiers (1999)
- A sentence about the assumption of independency between different levels of a given factor.
- A sentence on how to calculate each individual term of equation (6) and refer to Bosshard et al. (2013) for further details.

Pages 12 and 16: "Model realism" ... Of course every scientist and engineer would want their model to be "realistic", and there have been many opinion papers on this in hydrology advocating improving the "realism", whatever that means beyond ultimately matching some performance criteria. In this manuscript this issue is raised on Page 12, but then after a fairly basic modelling attempt produced results deemed "unrealistic" (page 13, lines 10-20), any improvement was deemed "beyond the scope". So I am not sure the manuscript and its case study in their current form can convincingly call on other modellers to achieve such "realism".

Reply: Here we disagree with the reviewers' definition of the expression "improving the realism". In our view, it is not exclusively a question of matching performance criteria, but more an attempt of letting the model behave according to the hydrologist perception of how it should behave. The crucial point here is the perception of which hydrologist should we refer to. For instance, Gharari et al. (2014b) claimed that "using expert knowledge to increase realism in environmental system models can dramatically reduce the need for calibration". In a parallel study, they define expert knowledge as "the *modeler's* perception of catchment behaviour and characteristics" (Gharari et al., 2014a). According to the two hydrologists types described in this study, the reviewer will agree that the *experimentalist's* perception would be more detailed than the modeller's one.

As we exclusively compared existing methods for process mapping, process parameterisation and parameter allocation, any improvement of one of those method was actually "beyond the scope" of the study. However, we disagree with the reviewer's reconstruction of our reasoning in section 4.1, as, for instance, the expression "beyond the scope"(pg. 13 line 20) belongs to the next section of the manuscript (4.2), which is not about model realism. The line of reasoning of section 4.1. is therefore following:

1. A modeller might conclude that using one of the two top-down mapping approaches (MU09 and GH11) used here may be the best choice in terms of cost-benefit, as the results obtained with them were, on average, only slightly lower than those obtained with the process maps with higher involvement of expert knowledge (SF07).
2. This conclusion is however not acceptable from an experimentalist point of view. The results may seem acceptable at the gauging stations, but the local representation of the DRP mapped would most likely differ from that expected by an experimentalist, as shown by Antonetti et al. (2016).
3. Therefore, we advocate that the hydrological community should aspire to develop models able to reproduce processes in a "realistic" way (i.e. in agreement with the *experimentalists'* expectation) at least at the sub-catchment or, even better, at the hillslope scale. At these scales, the added value of the process maps with higher involvement of expert knowledge should get highlighted.
4. As runoff measurements at this scales were not available on the investigated catchments, this issue remains a speculation, and will be verified during future research.

In the revised manuscript we will make clearer the call for the use of experimentalist's rather than modeller's perception for verification purposes. In addition, we will clearly state that our considerations about the scales, at which model realism should be achieved, will be verified during future research.

Other

1. Page 2 Line 5: What does the "It" refer to? "Expert knowledge"? Or "linking observations and laws"? The latter would make far more sense, but is not apparent from the way the sentence is written.

Reply: "It" is referred to "linking observations and laws". In the revised manuscript we will rephrase this sentence.

2. Page 2 Line 24 "Wet and dry" - this usage here is confusing and unnecessary. Taken literally "dry hydrologists" might be mistaken for hydrologists working in dry (arid/semiarid) catchments. I think it is fun to refer to experimentalists and modellers as wet and dry, but I think once the point is made to continue using these terms is not necessary and can lead to confusion.

Reply: Agreed. In the revised manuscript we will change "wet and dry" with their corresponding appellations "experimentalists and modellers".

3. Page 3 line 19 "too simplistic" - I think this requires some statement of the purpose - too simplistic to achieve what?

Reply: Agreed. In the revised manuscript we will modify the sentence as follows: "Top-down models and parameterisations may be too simplistic to depict the spatial variability of runoff processes within a given catchment and, therefore, require calibration […]"

4. Pages 3-4: Fenicia et al WRR2016 is an example where bottom-up ("distributed" and scaled up) and "top-down" (conceptual) approaches (as per definitions given earlier) are combined, applied and several hypotheses about process representations and hydrological controls were tested.

Reply: We want to thank the anonymous reviewer for the suggested reference, which will be integrated in the revised manuscript.

5. Page 5 line 25: This text correctly refers to the distinction between top-down and bottom-up approaches not being sharp - this aspect was overlooked in the earlier pages when discussing these two approaches, which lead to some too-strong statements being made.

Reply: Agreed. In the revised manuscript we will bring this aspect forward in the introduction.

6. Page 6 Line 25: "parameter allocation" sounds a bit awkward? Wouldn't "estimation" be a better choice here, covering all options (though measurement, calibration, etc)?

Reply: This issue was already raised by a reviewer in a previous publication (Antonetti et al., 2017), and the use of the expression "parameter allocation" was suggested by the other reviewer. We would therefore like to keep the expression "parameter allocation" throughout the manuscript, for consistency with Antonetti et al. (2017). In his recent opinion paper, Bahremand (2016) defined "parameter allocation" as "a logic-based specification", which completely fits in the context of the present study.

7. Page 6 line 31: give a reference for this "so-called hydrological approach", and explain how it is used in this work.

Reply: In the revised manuscript we will add the reference Dyck and Peschke (1995) but the explanation of how the hydrological approach is used in this work is already given in pg 7 line 5-9; there, we will explicitly refer to the hydrological approach.

8. Page 7 line 22: "storage parameter" might read better here than "storage constant"?

Page: The reviewer might be right maintaining that "parameter" would be better than "constant" as a worth choice. However, if the reviewer agrees, we would like to keep the expression "storage constant" for consistency with Antonetti et al. (2016, 2017).

9. Eqn 1: numbering each of these 2 equations individually would be helpful.

Reply: Agreed.

10. Page 8 line 20: what is a "modelling chain combination"? please briefly define/explain what this means.

Reply: The term in question is borrowed from the HEPEX (Hydrologic Ensemble Prediction EXperiment) community (https://hepex.irstea.fr/), where it is usually used to indicate a cascade of modelling elements, such as forcing data, hydrological models, pre- or postprocessing methods etc. (see for instance Bosshard et al., 2013). In this work, a single modelling chain consists of a given dataset of forcing data, a DRP map, and a model setup.

In the revised manuscript we will add a brief explanation.

11. Page 9 eqn 5 - clarify that "KGE applied to runoff time series" or similar, for clarity

Reply: Agreed. In the revised manuscript we will modify the sentence as follows: "*Runoff* simulations were evaluated with the Kling Gupta Efficiency (KGE; Gupta et al., 2009).

12. Page 9 line 18-23: Please define these P-factors and R-factors by equations, at least to the same standard as the KGE in eqn 1. Then all metrics become properly documented and easily reproducible without guessing how exactly they are calculated.

Reply: Agreed. In the revised manuscript we will report equations for P- and R-factors.

13. Section 4.5 It is good that the study is noting its limitations. I would also suggest the overall calibration approach used here is quite simplistic - it is quite possible a more detailed application would've produced quite different results. Worth also noting the use of a single case study makes it impossible to know if similar conclusions would hold elsewhere. This does not invalidate the results, just should be noted as one of the limitations.

Reply: Here we are not really sure which calibration criteria the reviewer is referring to, as both top-down and bottom-up model setups were applied uncalibrated on the study catchments.

The fact that a single case study makes it impossible to know if similar conclusions would hold elsewhere was already pointed out in pg. 16 lines 4-5: "Finally, to generalise the findings of this study, the number of catchments and events investigated should be increased considerably."

14. Section 5 Conclusions - to make sense of this, need to list the actual "expert knowledge" used here. This relates to a major concerned raised earlier that this concept is not sufficiently clearly delineated here.

Reply: In the revised manuscript we will address this topic according to the reply to the major concern.

15. Fig 1 - I found this figure rather confusing - somehow physically based models became closer to experimentalists approaches than conceptual models? I would not have said that, these kinds of models are certainly not developed by "field" "wet" hydrologists ... I think they do not fit within this "planar" dichotomous figure, they almost form another dimension of their own.

Reply: Agreed. We will re-design Fig 1 in the revised manuscript.

16. Fig 4 - worth providing references to several earlier studies where such "distributed model structures" were used.

Reply: Agreed.

17. Fig 6 - what are the benchmark options for the process maps, parameterizations, and parameter values?

Reply: There is no benchmark for process maps and model setup, as they are all depicted in Fig. 8 and Fig. 10. The reason why that forcing data was used as benchmark is reported on pg.9 line 11-12.

18. Figure 7 - not sure I am seeing in the figure how you arrived at the last sentence in the caption.

In the revised manuscript we will modify the last sentence of the caption as follows: “The SF07 map reproduced best the peak runoff with the bottom-up setup, whereas the GH11 map outperformed the other maps with the top-down setup”.

References

Addor, N., Rössler, O., Köplin, N., Huss, M., Weingartner, R. and Seibert, J.: Robust changes and sources of uncertainty in the projected hydrological regimes of Swiss catchments, Water Resour. Res., 50(10), 7541–7562, doi:10.1002/2014WR015549, 2014.

Antonetti, M., Buss, R., Scherrer, S., Margreth, M. and Zappa, M.: Mapping dominant runoff processes: an evaluation of different approaches using similarity measures and synthetic runoff simulations, Hydrol. Earth Syst. Sci., 20(7), 2929–2945, doi:10.5194/hess-20-2929-2016, 2016.

Antonetti, M., Scherrer, S., Kienzler, P. M., Margreth, M. and Zappa, M.: Process-based Hydrological Modelling: The Potential of a Bottom-Up Approach for Runoff Predictions in Ungauged Catchments, Hydrol. Process., doi:10.1002/hyp.11232, 2017.

Bahremand, A.: HESS Opinions: Advocating process modeling and de-emphasizing parameter estimation, Hydrol. Earth Syst. Sci., 20(4), 1433–1445, doi:10.5194/hess-20-1433-2016, 2016.

Bosshard, T., Carambia, M., Goergen, K., Kotlarski, S., Krahe, P., Zappa, M. and Schär, C.: Quantifying uncertainty sources in an ensemble of hydrological climate-impact projections, Water Resour. Res., 49(3), 1523–1536, doi:10.1029/2011WR011533, 2013.

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., 1–14, doi:10.5194/hess-2016-693, 2017.

Dyck, S. and Peschke, G.: Die Abflusskonzentration im Gewässernetz, in Grundlagen der Hydrologie, p. 536, Verlag für Bauwesen, Berlin., 1995.

Finger, D., Heinrich, G., Gobiet, A. and Bauder, A.: Projections of future water resources and their uncertainty in a glacierized catchment in the Swiss Alps and the subsequent effects on hydropower production during the 21st century, Water Resour. Res., 48(2), doi:10.1029/2011WR010733, 2012.

Gharari, S., Hrachowitz, M., Fenicia, F. and Savenije, H. H. G.: Hydrological landscape classification: Investigating the performance of HAND based landscape classifications in a central European meso-scale catchment, Hydrol. Earth Syst. Sci., 15(11), 3275–3291, doi:10.5194/hess-15-3275-2011, 2011.

Gharari, S., Shafiei, M., Hrachowitz, M., Kumar, R., Fenicia, F., Gupta, H. V. and Savenije, H. H. G.: A constraint-based search algorithm for parameter identification of environmental models, Hydrol. Earth Syst. Sci., 18(12), 4861–4870, doi:10.5194/hess-18-4861-2014, 2014a.

Gharari, S., Hrachowitz, M., Fenicia, F., Gao, H. and Savenije, H. H. G.: Using expert knowledge to increase realism in environmental system models can dramatically reduce the need for calibration, Hydrol. Earth Syst. Sci., 18(12), 4839–4859, doi:10.5194/hess-18-4839-2014, 2014b.

Gupta, H. V., Kling, H., Yilmaz, K. K. and Martinez, G. F.: Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling, J. Hydrol., 377(1–2), 80–91, doi:10.1016/j.jhydrol.2009.08.003, 2009.

Hellebrand, H., Müller, C., Matgen, P., Fenicia, F. and Savenije, H.: A process proof test for model concepts: Modelling the meso-scale, Phys. Chem. Earth, 36(1–4), 42–53, doi:10.1016/j.pce.2010.07.019, 2011.

Hrachowitz, M. and Clark, M.: HESS Opinions: The complementary merits of top-down and bottom-up modelling philosophies in hydrology, Hydrol. Earth Syst. Sci. Discuss., 1–22, doi:10.5194/hess-2017-36, 2017.

Hrachowitz, M., Fovet, O., Ruiz, L., Euser, T., Gharari, S., Nijzink, R., Freer, J., Savenije, H. H. G. and Gascuel-Odoux, C.: Process consistency in models: The importance of system signatures, expert knowledge, and process complexity, Water Resour. Res., 50(9), 7445–7469, doi:10.1002/2014WR015484, 2014.

Hümann, M. and Müller, C.: Improving the GIS-DRP Approach by Means of DelineatingRunoff Characteristics with New Discharge Relevant Parameters, ISPRS Int. J. Geo-Information, 2, 27–49, doi:10.3390/ijgi2010027, 2013.

Köplin, N., Schädler, B., Viviroli, D. and Weingartner, R.: The importance of glacier and forest change in hydrological climate-impact studies, Hydrol. Earth Syst. Sci., 17(2), 619–635, doi:10.5194/hess-17-619-2013, 2013.

Margreth, M., Naef, F. and Scherrer, S.: Weiterentwicklung der Abflussprozesskarte Zürich in den Waldgebieten, Zurich., 2010.

Müller, C., Hellebrand, H., Seeger, M. and Schobel, S.: Identification and regionalization of dominant runoff processes – a GIS-based and a statistical approach, Hydrol. Earth Syst. Sci., 13(6), 779–792, doi:10.5194/hess-13-779-2009, 2009.

Nijzink, R. C., Samaniego, L., Mai, J., Kumar, R., Thober, S., Zink, M., Schaefer, D., Savenije, H. H. G. and Hrachowitz, M.: The importance of topography-controlled sub-grid process heterogeneity and semi-quantitative prior constraints in distributed hydrological models, Hydrol. Earth Syst. Sci., 20(3), 1151–1176, doi:10.5194/hess-20-1151-2016, 2016.

Peschke, G., Etzenberg, C., TöpferJ, Zimmermann, S. and Müller, G.: Runoff generation regionalisation: analysis and a possible approach to a solution, in Proceedings of a conference on Regionalisation in hydrology held at Braunschweig, pp. 147–156., 1999.

Rössler, O., Diekkrüger, B. and Löffler, J.: Potential drought stress in a Swiss mountain catchmentEnsemble forecasting of high mountain soil moisture reveals a drastic decrease, despite major uncertainties, Water Resour. Res., 48(4), doi:10.1029/2011WR011188, 2012.

Safavi, H. R., Golmohammadi, M. H. and Sandoval-Solis, S.: Expert knowledge based modeling for integrated water resources planning and management in the Zayandehrud River Basin, J. Hydrol., 528, 773–789, doi:10.1016/j.jhydrol.2015.07.014, 2015.

Schmocker-Fackel, P., Naef, F. and Scherrer, S.: Identifying runoff processes on the plot and catchment scale, Hydrol. Earth Syst. Sci., 11(2), 891–906, doi:10.5194/hess-11-891-2007, 2007.

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.

Sivakumar, B.: Dominant processes concept in hydrology: Moving forward, Hydrol. Process., 18(12), 2349–2353, doi:10.1002/hyp.5606, 2004.

Smith, T., Marshall, L., McGlynn, B. and Jencso, K.: Using field data to inform and evaluate a new model of catchment hydrologic connectivity, Water Resour. Res., 49(10), 6834–6846, doi:10.1002/wrcr.20546, 2013.

von Storch, H. and Zwiers, F. W.: Statistical Analysis in Climate Research, Cambridge University Press., 1999.

Tang, Y., Reed, P., Wagener, T. and van Werkhoven, K.: Comparing sensitivity analysis methods to advance lumped watershed model identification and evaluation, Hydrol. Earth Syst. Sci., 11(2), 793–817, doi:10.5194/hess-11-793-2007, 2007.