Hydrol. Earth Syst. Sci. Discuss., 11, C3658–C3674, 2014
www.hydrol-earth-syst-sci-discuss.net/11/C3658/2014/

Hydrology and
Earth System
Sciences
Discussions
Open Access

Interactive comment on "Hydrological hysteresis in catchments and its value for assessing process consistency in conceptual models" *by* O. Fovet et al.

O. Fovet et al.

ophelie.fovet@rennes.inra.fr

Received and published: 8 September 2014

"Most parts of the manuscript are well written and structured. Its scientific contribution will fit well into Hydrology and Earth System Sciences after some revision have been performed. Part from some more elaborations about the Hysteresis Index and some necessary shortening of subsection 3.1 I have two major comments".

We appreciate the reviewer's positive assessment of the manuscript. In the following, we provide further precisions about the discussion points highlighted by the reviewer and suggest practical modifications to integrate these comments.

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

Major comments: -“Already in the methodology the authors refer to another study (Hrachowicz et al., 2014, in revision at WRR) that is not available for the reader. In particular the reference to hydrological signatures that are not explained in the text or shown in the figures through the entire text made some of the interpretations and conclusions hardly understandable.”

Reply: We agree that the definition of the signatures was missing so we propose to provide additional information about this work in a new table (cf. New Table 4 at the end of this reply) for clarifying the objective functions used for calibrating the models, the hydrological signatures used for assessing them; and as new version of Figure 3 (provided at the end of this reply). It has to be noticed that the manuscript submitted to WRR is now accepted and available for further details on the previous work (doi: 10.1002/2014WR015484).

-“In the description of the models and their parameters (which is partly referring to the above-mentioned study) the authors choose one final parameter set for each of the four models based on a weighted performance measure that only uses discharge observations. However, many preceding studies showed that models with more than 4-6 parameters face problems of over-parameterization when they only use discharge for calibration (Jakeman and Hornberger, 1993; Wheater et al., 1986). The low spread of weighted efficiencies/Euclidean distances in Fig3 in the manuscript might disprove that but the distributions shown there are re-shaped (with an exponent of 10) and might appear much more uniform in their original distribution. Since the model simulations are a substantial part of the interpretations and second part of the manuscript the authors need to provide some more information about the reliability of their models and the chosen parameters”

Reply: We fully agree with the reviewer that, if insufficiently constrained, models with elevated numbers of parameters are subject to increased parameter uncertainty and associated predictive uncertainty, as many parameter combinations will merely provide a mathematically suitable fit while essentially misrepresenting the internal dynamics

Full Screen / Esc
Printer-friendly Version
Interactive Discussion
Discussion Paper

of the system as pointed out by many previous studies (e.g. Beven, 2006; Kirchner, 2006; Gupta et al., 2008; Andreassian et al., 2012). To reduce that problem while in the same time allowing for higher process complexity, we chose a double strategy for model calibration/selection in the Hrachowitz et al. (2014) manuscript: 1) multiple objective calibration based on 4 calibration objectives, which in the past has been shown to be in itself already a valuable tool for identifying parameter sets that would otherwise be kept as feasible if only 1 calibration objective (e.g. Nash-Sutcliffe Efficiency) was used (e.g. Gupta et al., 1998; Seibert and McDonnell, 2002; Vrugt et al., 2003; Fenicia et al., 2007; see also a recent review paper on the topic by Efstradiadis and Koutsoyiannis, 2010), highlighting the different information content of objective functions based on different catchment signatures (Euser et al., 2013). 2) to further increase the confidence that the selected parameter sets actually reproduce the observed system dynamics to a certain extent, we complemented the multi-objective calibration strategy with the use of expert-knowledge and data driven parameter and process constraints to ensure that the selected parameter sets are in themselves consistent (e.g. the unsaturated storage capacity in wetlands needs to be lower in wetlands than on hillslopes) and that they reproduce system dynamics that do not contradict what we know about the system (e.g. unreasonably high/low long-term average base flow contributions or actual evaporation as estimated from the Budyko relationship) within certain limits of acceptability (e.g. Winsemius et al., 2009; Gharari et al., 2013; Gao et al., 2014). Applying these constraints a wide range of mathematically feasible parameter sets, violating these constraints, can be discarded significantly reducing the ill-posed nature of the problem. This is also reflected in our Figure 3: not only does the calibration performance and its spread improve with the progression of M1-M4, more importantly, the performance and its spread during VALIDATION also improves, indicating improved predictive power of the model, which in turn points toward potentially improved process representation.

We will clarify this in the revised manuscript. Please also note, that the performance measurements used, i.e. the Euclidean distances, are not weighted, rather the Euclidean distances themselves (together with an exponent of 10) were used as informal likelihood measurements (following the concept of GLUE) to construct uncertainty bounds around the modelled variables in the Hrachowitz et al. (2014) manuscript. Thus, the weights do not affect the actual performance assessment for the manuscript under review. We will therefore remove any reference to it in the revised version.

Detailed comments:

1. p. 5565, line 20. “Please elaborate the link between the scale problem of lumped and (semi) distributed models and the storage behavior in a bit more detail”.

Reply: In order to clarify our statement we suggest adding the following precisions: “A time-series of groundwater table level from a single piezometer is not representative of the behaviour of the groundwater, even at the hillslope scale; therefore it is difficult to link it with either a reservoir volume simulated by a lumped model or an average water table level of a grid point simulated by a fully distributed model.”

2. p. 5667, line 7. “Please also mention studies that used water quality data for model assessment, e.g A Hartmann, T Wagener, A Rimmer, J Lange, H Brielmann, M Weiler Water Resources Research 49 (6), 3345-3358 or Hartmann, A., Weiler, M., Wagener, T., Lange, J., Kralik, M., Humer, F., Mizyed, N., Rimmer, A., Barberá, J. A., Andreo, B., Butscher, C. and Huggenberger, Hydrol. Earth Syst. Sci., 17(8), 3305–3321” .

Reply: The use of water quality data is indeed another example of multi-calibration studies, in the early version of the manuscript we cited only quantitative examples but we suggest adding the following references as examples of the use of tracer data in multi-data modelling approaches:

“chloride concentrations (Hrachowitz et al., 2011), atmospheric tracers (Molenat et al., 2013) or nitrates and sulfate concentrations (Hartmann et al, 2013 a), and water isotope as δ18O (Hartmann et al., 2013 b)”.

3. p. 5668, line 7. “PET + drainage are smaller than precipitation - please elaborate. In addition, please mention also that there is a strong seasonal behavior. Otherwise the

Full Screen / Esc
Printer-friendly Version
Interactive Discussion
Discussion Paper

definition of the HI would be hard to understand (see comments below)."

Reply: There is indeed a high water deficit in the annual budget of the catchment almost each year (PET+Q<R). It is true for the neighbour catchments and nested catchments too (not presented in the paper but also part of the ORE AgrHys). This deficit is likely to be due to underflows below the outlet, as it was suggested by previous studies (e.g. Ruiz et al. 2002). We propose to add the following sentences in the study site section to notify these points to the reader more explicitly:

"Mean annual rainfall over the period 1992-2012 is 1113 mm (+/-20%) and mean annual Penman potential evapotranspiration (PET) is 700 mm (+/- 4%). Mean annual drainage is 360 mm (+/- 60%) at the outlet. There is a high water deficit in the annual budget almost each year due to underflows below the outlet (Ruiz et al., 2002). The catchment is laying under granite called leucogranodiorite of Plomelin, which upper part is weathered on 1 to more than 20 m deep. Soils are mainly sandy loam with an upper horizon rich in organic matter, depths are comprised between 40 and 90 cm. Soils are well drained except in the bottomlands which represent 7% of the total area. Agriculture dominates the land use with 86% of the total area. The base flow index is about 80 to 90%, thus the hillslope aquifer is the main contributor to stream (Molenat et al., 2008; Ruiz et al., 2002). Both stream flow and shallow groundwater tables exhibit a strong annual seasonality in this catchment (Fig. 2)".

And to illustrate the strong annual seasonality we suggest adding the hydrographs in new version of Figure 3.

4. p. 5669, line 9. "I think you mean 'groundwater storage dynamics' and not groundwater storage', which would be related to volumes rather than dynamics"

Reply: Yes the word "dynamic" has to be added: as explained in the introduction (and related to comment 1) the groundwater level and soil moisture measurements are more representative of the dynamics of the storages rather than of the volumes themselves.

5. p. 5669, l. 23-24. “Rephrase”.

Reply: We propose the following reformulation of the sentence: “The two profiles are located on the upslope and downslope parts of the hillslope respectively. Thus, we assumed that averaging their normalized values will allow us to build a proxy for the dynamics of the unsaturated zone storage on the whole hillslope”.

6. p. 5670, lines 8-12, “This is not clear. Please provide some more detail why an index has to be used.”

Reply: To make it clearer we propose to reformulate the sentence as “For storage-discharge hysteresis at the annual scale, this approach is not sufficient as the same type of hysteretic loop is likely to happen for almost all the years because despite of stream flow inter annual variations, the seasonality (with a high flow period during winter) is the same for all years. Moreover a preliminary cross correlation analysis revealed that storage and stream flow are strongly correlated”

7. p. 5670-5671 Eq. (1): “How often can you calculate this difference within one year? Is HI their mean? (by reading through the proceeding chapters it appears that there is a strong seasonality in discharge and these values might only be passed once a year, but this is not clear at this stage of the manuscript)“, “define Qmid before eq. 1 and mention how often it occurs within one hydrological year.“

Reply: Indeed, the strong seasonal cycle observed on the studied catchment allowed us to compute a HI based on a Qmid value which is taken only 2 times per year : during the recharge period and during the recession period. Actually, high and low stream flow values are more likely to occur several times a year in this catchment than medium values. Moreover, we smoothed the data using a 7-day moving average to remove highest Q values due to rapid storms. We suggest explaining and presenting the presence of this seasonal cycle in the case study section (cf. Reply to comment 3 and illustration in new Figure 3), and adding some precisions when defining Qmid in this section:

"In this paper, as the hydrological variables exhibit a strong annual uni-modal cycle, we calculated the hysteresis index (HI) each year as the difference between water storages at the dates of mid-point discharge [...] In order to reduce the impact of the quick variations of discharge or groundwater level due to individual storm events, we smoothed the time series using 7-day moving averages. The strong seasonal discharge cycle led to identify two occurrences of Qmid per year only: during the recharge period (tR) and during the recession period (tr) , while high and low stream flow values are taken several times per year as explained by Lawler et al. (2006).

8. p. 5671, lines 2-4. "Parts of subsection 2.4 appear like a literature review that could also be part of the introduction".

Reply: We would like to keep the literature related to the hysteresis descriptions in this section as it is really specific to this methodological point and we do not see how it could fit in the general introduction

9. p. 5671, lines 15. "The authors should also add some more detail about their reasons to exactly choose this hysteresis index. Considering only 2 points instead of shape/rotation/etc. might omit some convolutions but it also gives the impression that a lot of information is ignored and misinterpretations might be possible, too"

Reply: We agree that HI does not integrate a full description of the hysteresis and does not pretend to do so. It is only an index which gives already 2 types of information as explained in the paragraph, and allows a quantitative comparison between simulations and observations. Classification methods for storm events are indeed based on the rotational pattern (clockwise/anticlockwise), curvature (shape), and trend or rotational angle. Rotation pattern is given by the sign of HI. Curvature is defined from concave to convex, the trend is generally used to identify on concentrations-discharge hysteresis if the solutes are diluted (negative trend) or concentrated (positive trend) during a storm. Our hystereses are always concave with a positive trend. This pattern similar among years tends to support the hysteresis as a signature of internal catchment behaviour.

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

Note also that all our interpretations are driven by the full observations and not only on the HI values. But we agree that if the objective was to calibrate models, more analysis would be required to identify the best index (or a combination of several ones) to lose as little information as possible.

10. p. 5672 l. 7, “Just considering water balance it appears obvious that M1 will not work (see my comment at the study site description)”.

Reply: We fully agree with the reviewer, yet, although many catchment worldwide exhibit similar water balance deficits due to deep infiltration losses/underflow (e.g. Le Moine et al., 2007), only a small minority of models actually caters for this process. M1 has been used in the Hrachowitz et al. (2014) manuscript as a starting point and benchmark model that resembles many frequently used models (e.g. HBV, Sacramento, FLEX, etc.) The use of such an overly simplistic model structure allowed us not only to illustrate that it cannot sufficiently well reproduce the hysteretic behaviour, but it also helped in model diagnosis to see where and how the model fails to reproduce the catchment behaviour. Thus, M1 was included merely for instructive purposes to demonstrate how wrong modelling can go with frequently used, yet unsuitable model architectures.

11. p. 5673, line 1, “This is only three objective functions”.

Reply: There are four objective functions: the Nash Sutcliffe Efficiency Criterion applied to the stream flow (1), to the logarithm of stream flow (2), and to the flow duration curve (3), and finally the Volumetric Efficiency criterion applied to stream flow (4). In order to clarify this point we suggest presenting the four criteria with the hydrological signatures in the additional table (cf. New Table 4 at the end of this reply).

12. p. 5673, “this is really high did Freer et al also use the Same value ? Is it really necessary to introduce p at all?”

Reply: It is indeed a high value and Freer et al. (1996) even tested the effect of using an exponent of 30. To further answer the question: no it is not absolutely necessary to introduce an exponent for the informal likelihood measure. However, an exponent, in particular a high one, can serve two purposes: (1) it reduces the width of the uncertainty interval by giving relatively little weights to poor model realizations, thereby addressing the frequently raised criticism of GLUE that it overestimates uncertainty (e.g. Mantovan and Todini, 2006; Stedinger et al., 2008) and (2) it reduces the sensitivity of uncertainty interval to the subjective choice of behavioural models, which is a second point frequently criticised in GLUE (e.g. Montanari, 2005). Please note that the GLUE and the informal likelihood weights (+exponent) do not affect the actual performance assessment for the manuscript under review. We will therefore remove any reference to it in the revised version.

13. p. 5674, line 1. , p. 5680, line 23 and p. 5681, line 25 , p. 5682, line 26, "Signatures are not defined anywhere and not shown".

Reply: We suggest adding the signatures used in the modelling work presented in Hrachowitz et al., (2014) in an additional table (see new Table 4) and the performances of each model on these signatures.

14. p.5674 lines 3-16. "This whole paragraph presents results that should be moved to the results section. Furthermore the authors explain differences among the model by state variables and fluxes that are not shown in Figure 3. It is not clear, which part of these results was done by Hrachowitz et al. 2014. "

Reply: These results are a contribution from a previous work (Hrachowitz et al., 2014) therefore they are presented here in this material and method section as they are considered as previous results/knowledge for the present study. We agree with the reviewer that this has to be clarified (see also reply to reviewer 1). To make this point clearer we suggest an explicit reformulation in the introduction of this paragraph.

(i)to add a mention to this previous work at the end of the introduction :

“ (...) ii) to which degree a suite of conceptual rainfall-runoff models with increasing complexity, which were calibrated and evaluated for this catchment in a previous work, using a flexible modelling framework (Hrachowitz et al., 2014), can reproduce the observed storage-discharge hysteresis (...)”

(ii)to explain this choice to the reader at the beginning of the corresponding section

“In a previous work, a range of 11 conceptual models were calibrated and evaluated for the Kerrien catchment in a stepwise development using a flexible modelling framework (see Hrachowitz et al., 2014). This section aims at summarizing the results of this previous study that are used as a basis for the present work.”

We also suggest providing additional information about this work as, especially the results of model calibrations in a new version of Figure 3.

Moreover the manuscript submitted to WRR is now accepted and available for further details on the previous work (doi: 10.1002/2014WR015484).

15. p.5674, line 26., p. 5683 line 20, “It is very confusing to refer to the results of another study. It is also critical to go on only with the "best" parameter set, which might be very similar to other parameter sets in the sample if p was not applied on the original likelihoods. Is there any proof that the selected parameter set were sufficiently identifiable, ie. that there is no equifinality ? “

Reply: The study of Hrachowitz et al. (2014) aimed at proposing a stepwise modelling approach where increasing model complexity (and increasing model number of parameters) was always associated with an increase of model constraints (parameter or architecture constraints) and always motivated by the need of reducing the predictive uncertainties, and the difference between calibration and evaluation period uncertainties (so called model consistency) rather than increasing model performance in the calibration. This approach limits the equifinality which may appear when increasing the model complexity (see also reply to major comment #2). To provide the reader more

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

information about this point, we suggest adding some details from the previous work in the new version of Figure 3 with illustrations of model performances on the objective functions, for both calibration and independent evaluation period.

16. p.5676, line 12. “It would be very helpful to show schematic figures /conceptual models that visualise these interpretations “.

Reply: We thank the reviewer for this suggestion, an additional Figure with a scheme of the interpreted mechanisms is proposed as a new Figure 9 (see at the end of this reply).

17. p. 5676, “the subsection above is also about observations in hill slope - please choose other header”.

Reply: Indeed the previous section is already dealing with hillslope observations so we suggest modified the 2 titles as “3.1.1 Observations in hillslope and riparian zones: saturated storage vs. Flow” and “3.1.2 Observations in hillslope: saturated and unsaturated storages vs. Flow”.

18. p.5677, section 3.1.3. “There is quite a lot of interpretation subsections 3.1.1 to 3.1.3. which overloads the manuscript combined with the proceeding modeling. I recommend to shorten this part (Subsections 3.1.1 to 3.1.3) significantly and providing conceptual drawings of the interpreted system behavior instead“.

Reply: We agree that the addition of a conceptual scheme would be useful. The proposed new Figure 9 would provide the required conceptual drawing. However, we propose to maintain the text, also as it has been appreciated by reviewer 1.

19. p.5680, line 24, “As mentioned before, this is already obvious only by considering water balance...”

Reply: we agree with the reviewer on the fact that model M1 can be expected as non consistent regarding our knowledge of the catchment (see reply to comment 10). However it is interesting to see that the hysteresis comparison shows clearly and immediately what is inconsistent and how the model behaved to compensate the error: “model M1, the overestimation of the hillslope saturated storage was partially compensated by the underestimation of the hillslope unsaturated storage”.

20. p. 5681 line 28, “this can only be answered when the model parameters are evaluated for their sensitivity and parameter interactions. With the present information equifinality in calibration could also be a very probable explanation”

Reply: We included a sensitivity analysis of HI on the basis of the parameter sets retained as feasible (see below and also replies to the other reviewers). These results have been integrated at the end of section 3.2.1 and in Figure 11b. About the equifinality see also reply to comment 15.

“The hysteresis index sensitivity to parameter uncertainty increases with the number of parameters from M1 to M2 and then stays in the same range from M2 to M4 (Figure 11b). This analyse confirms the importance of considering the Hysteresis Indices both between saturated and unsaturated storage (HSS and HUS) to avoid accepting an inadequate model architecture. For example, considering only the performance on the HSS(Q) relationship could lead to accept model M1 while its performance on HUS is lower and it is not able to reproduce the Riparian compartment hysteresis. For readability purposes, Figure 11b illustrates this sensitivity for the different HI in the year of 2011-2012 only but similar behaviour is observed every year. It showed that best behavioural parameters sets (bbp) lead to modelled HI values closer to the observed values than average modelled HI values. Using an additional calibration criterion related to the hysteresis could reduce the sensitivity of HI to parameter uncertainty and lead to narrow range of feasible parameter sets.”

21. p. 5686 line 4-8. “Large part of this paragraph is referring to Hrachowitz et al. 2014, which was not part of this study”.

Reply: The previous work of Hrachowitz et al. (2014), analyzed the model ability to reproduce the classical hydrological signatures but the comparison with the performance on the hysteretic signature is actually a result of the present work. We propose to reformulate in order to avoid any ambiguity:

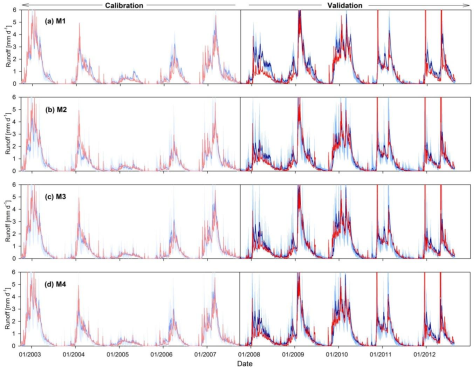
“The tested models were characterized by an increasing degree of complexity and also an increasing consistency, as shown in a previous study using classical hydrologic signatures. In this study, we showed that if all of them simulated a hysteretic relationship between storages and discharge, their ability to reproduce hysteresis index also increased with model complexity. In addition, we suggest that if classical hydrological signatures help to assess model consistency, the hysteretic signatures help also to identify quickly when the models give “right answer for wrong reasons” and can be used as a descriptor of the internal catchment functioning.”

Please also note the supplement to this comment:
http://www.hydrol-earth-syst-sci-discuss.net/11/C3658/2014/hessd-11-C3658-2014-supplement.pdf

---

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 11, 5663, 2014.

Figure 3. a. Observed (red line) and modelled runoff for model set-ups (a) M1, (b) M2, (c) M3 and (d) M4 in calibration and independent evaluation (validation) periods. Modelled runoff shown as most balanced solution (dark blue line) and the 5/95th uncertainty bounds (light blue shaded area). Adapted from Hrachowitz et al. (2014).



b. Overall model performance for all model set-ups (M1-M4) expressed as Euclidean Distance from the "perfect model" computed from all calibration objectives and signatures with respect to calibration and validation periods. Triangles represent the optimal solution, i.e. the solution obtained from the parameter set with the lowest Euclidean Distance during calibration. Box plots represent the Euclidean Distance for the complete sets of all feasible solutions (the dots indicate 5/95th percentiles, the whiskers 10/90th percentiles and the horizontal central line the median). Adapted from Hrachowitz et al. (2014).

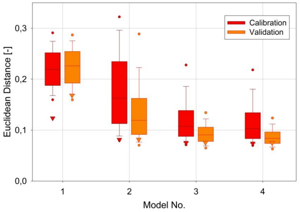


Fig. 1. New Figure 3

Figure 9: Conceptual scheme of successive mechanisms which explain the annual hysteresis between storages and stream flows. HUS: Hillslope unsaturated storage, HSS: hillslope saturated storage, RUS: riparien unsaturated storage, RSS: riparian saturated storage, Q: stream flow, bold characters indicate varying components, grey arrows indicate if the component is increasing or decreasing, black arrows indicate the water flow paths.

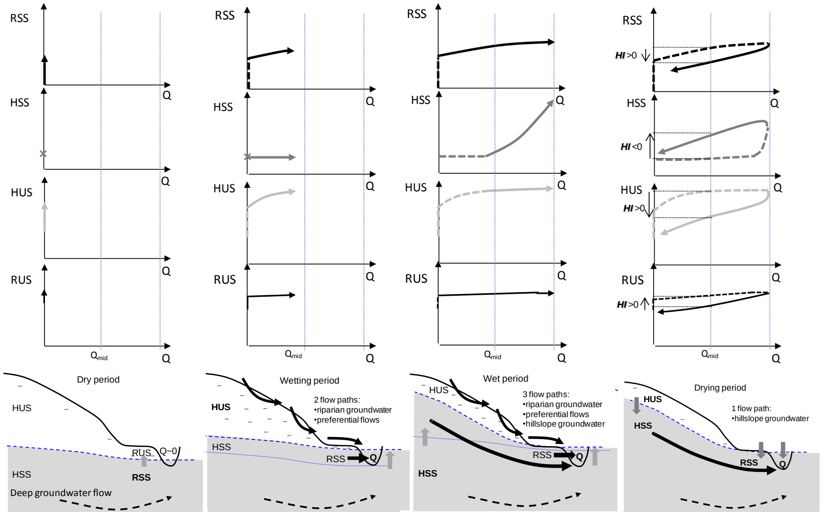


Fig. 2. New Figure 9

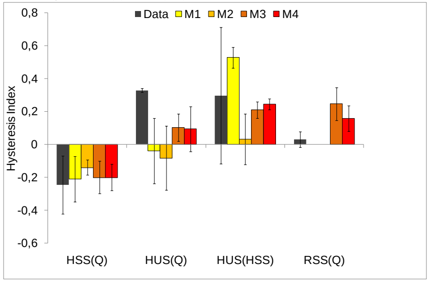
Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

Figure 11. a. Mean annual hysteresis Indices observed and simulated with the 4 models M1 to M4, for Hillslope saturated storage vs. discharge HSS(Q), Hillslope unsaturated storage vs. discharge HUS(Q), Hillslope unsaturated storage vs. Hillslope saturated storage HUS(HSS), and Riparian saturated storage vs. discharge RSS(Q). RSS is simulated only in models M3 and M4. Error bars show the standard deviation for the 10 years for HSS(Q) and RSS(Q), and the values for the two available years for HUS(Q) and HUS(HSS).



b. Sensitivity of Hysteresis Index values to parameter uncertainty for the year 2011-2012. Mx bbp indicates the value for best behavioural parameter sets, the circles, triangles, squares ,and diamonds indicate the mean HI value for the all the behavioural parameter sets, and the corresponding bars its range of variation.

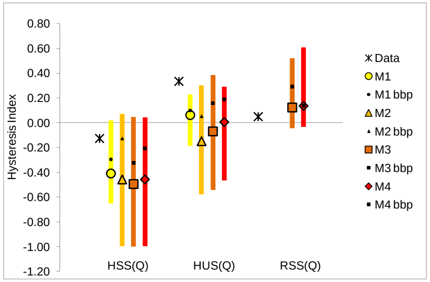


Fig. 3. New Figure 11

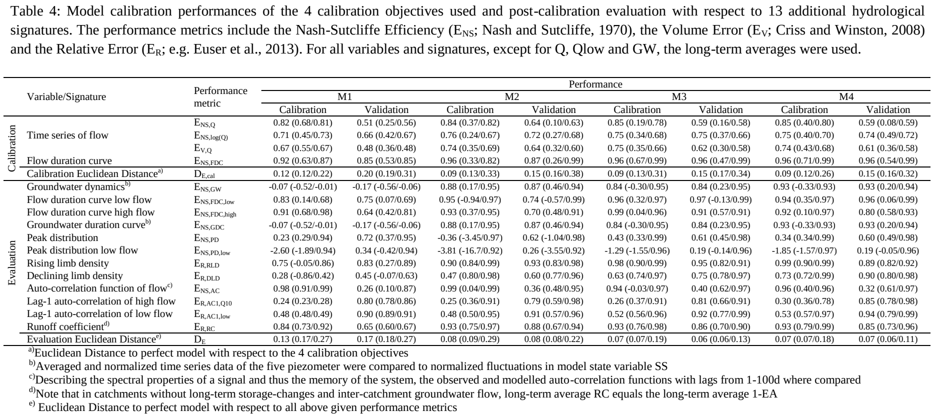


Table 4: Model calibration performances of the 4 calibration objectives used and post-calibration evaluation with respect to 13 additional hydrological signatures. The performance metrics include the Nash-Sutcliffe Efficiency (E_{NS}; Nash and Sutcliffe, 1970), the Volume Error (E_V; Criss and Winston, 2008) and the Relative Error (E_R; e.g. Euser et al., 2013). For all variables and signatures, except for Q, Qlow and GW, the long-term averages were used.

	Variable/Signature	Performance metric	Performance							
			M1		M2		M3		M4	
			Calibration	Validation	Calibration	Validation	Calibration	Validation	Calibration	Validation
Calibration	Time series of flow	$E_{NS,Q}$	0.82 (0.68/0.81)	0.51 (0.25/0.56)	0.84 (0.37/0.82)	0.64 (0.10/0.63)	0.85 (0.19/0.78)	0.59 (0.16/0.58)	0.85 (0.40/0.80)	0.59 (0.08/0.59)
		$E_{NS,log(Q)}$	0.71 (0.45/0.73)	0.66 (0.42/0.67)	0.76 (0.24/0.67)	0.72 (0.27/0.68)	0.75 (0.34/0.68)	0.75 (0.37/0.66)	0.75 (0.40/0.70)	0.74 (0.49/0.72)
		$E_{V,Q}$	0.67 (0.55/0.67)	0.48 (0.36/0.48)	0.74 (0.35/0.69)	0.64 (0.32/0.60)	0.75 (0.35/0.66)	0.62 (0.30/0.58)	0.74 (0.43/0.68)	0.61 (0.36/0.58)
	Flow duration curve	$E_{NS,FDC}$	0.92 (0.63/0.87)	0.85 (0.53/0.85)	0.96 (0.33/0.82)	0.87 (0.26/0.99)	0.96 (0.67/0.99)	0.96 (0.47/0.99)	0.96 (0.71/0.99)	0.96 (0.54/0.99)
	Calibration Euclidean Distance[a)]	$D_{E,cal}$	0.12 (0.12/0.22)	0.20 (0.19/0.31)	0.09 (0.13/0.33)	0.15 (0.16/0.38)	0.09 (0.13/0.31)	0.15 (0.17/0.34)	0.09 (0.12/0.26)	0.15 (0.16/0.32)
Evaluation	Groundwater dynamics[b)]	$E_{NS,GW}$	-0.07 (-0.52/-0.01)	-0.17 (-0.56/-0.06)	0.88 (0.17/0.95)	0.87 (0.46/0.94)	0.84 (-0.30/0.95)	0.84 (0.23/0.95)	0.93 (-0.33/0.93)	0.93 (0.20/0.94)
	Flow duration curve low flow	$E_{NS,FDC,low}$	0.83 (0.14/0.68)	0.75 (0.07/0.69)	0.95 (-0.94/0.97)	0.74 (-0.57/0.99)	0.96 (0.32/0.97)	0.97 (-0.13/0.99)	0.94 (0.35/0.97)	0.96 (0.06/0.99)
	Flow duration curve high flow	$E_{NS,FDC,high}$	0.91 (0.68/0.98)	0.64 (0.42/0.81)	0.93 (0.37/0.95)	0.70 (0.48/0.91)	0.99 (0.04/0.96)	0.91 (0.57/0.91)	0.92 (0.10/0.97)	0.80 (0.58/0.93)
	Groundwater duration curve[b)]	$E_{NS,GDC}$	-0.07 (-0.52/-0.01)	-0.17 (-0.56/-0.06)	0.88 (0.17/0.95)	0.87 (0.46/0.94)	0.84 (-0.30/0.95)	0.84 (0.23/0.95)	0.93 (-0.33/0.93)	0.93 (0.20/0.94)
	Peak distribution	$E_{NS,PD}$	0.23 (0.29/0.94)	0.72 (0.37/0.95)	-0.36 (-3.45/0.97)	0.62 (-1.04/0.98)	0.43 (0.33/0.99)	0.61 (0.45/0.98)	0.34 (0.34/0.99)	0.60 (0.49/0.98)
	Peak distribution low flow	$E_{NS,PD,low}$	-2.60 (-1.89/0.94)	0.34 (-0.42/0.94)	-3.81 (-16.7/0.92)	0.26 (-3.55/0.92)	-1.29 (-1.55/0.96)	0.19 (-0.14/0.96)	-1.85 (-1.57/0.97)	0.19 (-0.05/0.96)
	Rising limb density	$E_{R,RLD}$	0.75 (-0.05/0.86)	0.83 (0.27/0.89)	0.90 (0.84/0.99)	0.93 (0.83/0.98)	0.98 (0.90/0.99)	0.95 (0.82/0.91)	0.99 (0.90/0.99)	0.89 (0.82/0.92)
	Declining limb density	$E_{R,DLD}$	0.28 (-0.86/0.42)	0.45 (-0.07/0.63)	0.47 (0.80/0.98)	0.60 (0.77/0.96)	0.63 (0.74/0.97)	0.75 (0.78/0.97)	0.73 (0.72/0.99)	0.90 (0.80/0.98)
	Auto-correlation function of flow[c)]	$E_{NS,AC}$	0.98 (0.91/0.99)	0.26 (0.10/0.87)	0.99 (0.04/0.99)	0.36 (0.48/0.95)	0.94 (-0.03/0.97)	0.40 (0.62/0.97)	0.96 (0.40/0.96)	0.32 (0.61/0.97)
	Lag-1 auto-correlation of high flow	$E_{R,AC1,Q10}$	0.24 (0.23/0.28)	0.80 (0.78/0.86)	0.25 (0.36/0.91)	0.79 (0.59/0.98)	0.26 (0.37/0.91)	0.81 (0.66/0.91)	0.30 (0.36/0.78)	0.85 (0.78/0.98)
	Lag-1 auto-correlation of low flow	$E_{R,AC1,low}$	0.48 (0.48/0.49)	0.90 (0.89/0.91)	0.48 (0.50/0.95)	0.91 (0.57/0.96)	0.52 (0.56/0.96)	0.92 (0.77/0.99)	0.53 (0.57/0.97)	0.94 (0.79/0.99)
	Runoff coefficient[d)]	$E_{R,RC}$	0.84 (0.73/0.92)	0.65 (0.60/0.67)	0.93 (0.75/0.97)	0.88 (0.67/0.94)	0.93 (0.76/0.98)	0.86 (0.70/0.90)	0.93 (0.79/0.99)	0.85 (0.73/0.96)
	Evaluation Euclidean Distance[e)]	D_E	0.13 (0.17/0.27)	0.17 (0.18/0.27)	0.08 (0.09/0.29)	0.08 (0.08/0.22)	0.07 (0.07/0.19)	0.06 (0.06/0.13)	0.07 (0.07/0.18)	0.07 (0.06/0.11)

[a)] Euclidean Distance to perfect model with respect to the 4 calibration objectives

[b)] Averaged and normalized time series data of the five piezometer were compared to normalized fluctuations in model state variable SS

[c)] Describing the spectral properties of a signal and thus the memory of the system, the observed and modelled auto-correlation functions with lags from 1-100d where compared

[d)] Note that in catchments without long-term storage-changes and inter-catchment groundwater flow, long-term average RC equals the long-term average 1-EA

[e)] Euclidean Distance to perfect model with respect to all above given performance metrics

Fig. 4. Table 4

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper