Hydrol. Earth Syst. Sci. Discuss., 11, C3675–C3691, 2014
www.hydrol-earth-syst-sci-discuss.net/11/C3675/2014/

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

"The manuscript is a relevant and important step towards diagnostic analysis of hydrologic models, with specific contribution through testing internal hydrologic process representation via comparing observed and simulated hysteretic patterns that exist between storage and streamflow discharge in the selected watershed. I think the topic is of interest to the HESS readership and the manuscript is well written, well-structured and the use of language is generally good. However, part of the analysis is described at an abstract level with frequent citation to a recent unpublished manuscript (Hrachowitz et al., 2014) by the authors. Therefore a clear distinction between the contributions by

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

these two manuscripts should be made in the manuscript.

Reply: We thank the reviewer for the positive assessment of our article. As highlighted also by the other reviewers the distinction with the results from the previous study of Hrachowitz et al. should be made clearer, this can be done by the addition of explicit sentences at the end of the introduction and when the results of this previous work are presented, i.e. in the “materials and methods” section, as they are previous knowledge for the present work. Also we suggest to provide additional information about these previous results 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 in a new version of Figure 3 for illustrating the results obtained from this previous study (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).

“I also suggest below a few cases where explicit discussion of the analysis should be provided. Overall, my assessment for the manuscript is minor revisions. The manuscript could be published after the authors address the comments listed in “Main Comments” and “Minor Comments” sections listed below.”

Reply: We also thank the reviewer for his/her comments and suggestions, our corresponding discussions and suggested modifications to take these comments into account are detailed below.

Main comments:

1.“The authors state that 86% of the study area is dominated by agriculture (Section 2.1). A discussion on the source of irrigation water (groundwater?) and how the agricultural use could affect the hysteretic patterns should be provided in the manuscript. Similar discussion related to percentage of snow and its possible impact on the results should be provided.”

Reply: Due to the temperate oceanic climate, there is no snow cover on the study catchment which is located only 10 km from the mouth of the Odet River in the Atlantic Ocean. This climate is indeed characterized by rainy and mild winters (with minimal temperature around 5.9 Celsius deg) and relatively wet and cool summers. The average total rainfall in the summer is more than 200 mm, and the numbers of day with rain per year is around 150 days in average. Due partly to this abundance of regular rainfall, agriculture in Brittany represents only 4 % of water uptakes. Irrigation is used in only 2% of the agricultural area and mainly for vegetable cropping. Irrigation is absent in the study catchment. We added in section 2.1:

“Agriculture dominates the land use with 86% of the total area covered by grassland, maize and wheat, none of them irrigated.”

2.“Hysteresis Indices (Section 2.4): The description of the mid-point discharge is not clear. In addition the manuscript lacks a hydrograph, which many hydrologists would very much like to see in a manuscript related to hydrologic processes and models. Therefore the authors should provide a representative hydrograph of the watershed with clear description of the mid-point discharge values on the hydrograph. Below the hydrograph a time series of water levels/moisture levels should be provided again indicating the selected points used in calculating the Hysteresis index. A figure as described above is very important for understanding the HI concept used in the study. My main concern is that the HI concept followed in the manuscript is only specific to the selected watershed. I also think that mid-point discharge could occur multiple times during recharge and recession periods, therefore which time to select should be clearly described in the manuscript”.

Reply: (See also reply to reviewer 1 and 2) In order to clarify the mi-point values we proposed to add a new Figure (Figure 9 the end this reply) that helps to identify that within the seasonal pattern observed on the studied catchment, the mid-point discharge value is taken only twice per year during the recharge and the recession periods respectively; and to add the observed and simulated hydrographs in new version of Figure

3. Regarding the observed hysteresis over the 10 years, the choice of Qmid succeeded in catching the difference of the saturated storage states in recharge and recession more or less in the middle of these periods. This behaviour is not specific to the studied catchment, and is also supported by previous studies e.g. Lawler et al., (2006) who "argue that computing HI by using mid-point discharge usually allows avoiding the small convolutions which are frequently observed at both ends of the hysteretic loop." (p.5670, lines 17-19).

Reply: We agree that HI does not integrate a full description of the hysteresis and does not pretend to do so. However it gives already 2 types of information as explained in the paragraph and allows a quantitative comparison between simulations and observations. It is relevant for studying annual pattern with strong seasonal cycle. It could be used similarly for flood event if it is uni-modal. For multiple-seasonality, e.g. if there are 2 recharge periods in the year, the hysteresis is likely to exhibit a double loop and 2 indices may be relevant to describe each of them. In particular one can imagine that 2 successive loops may have different directions (so different signs of HI) due to the successive activation of different flow paths and the fact that storages are likely to be less empty at the beginning of the second recharge than at the beginning of the first one. In snow-melt driven catchment, the hysteresis relative to the snow cover storage should be taken into account too (as a third storage). In arid catchment where the groundwater recession can occur during several years (see e.g. Ruiz et al., 2010), it would be more relevant to compare these relationships among the identified pluri-annual cycles composed by at least both a recharge and a recession rather than at the annual scale. As cited in the corresponding section 2.4, some authors prefer to describe the loop width using the extreme values of the Y variable (X variable is always stream flow, Y is storage in our case but often either a concentration or the turbidity in hysteresis studies).

In order to help the identification of the limits of our HI we propose to add the following precisions:

“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 the two phases of the hydrological year [...]“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 occurences 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).”

3. “Model calibration and Evaluation (Section 2.6): It seems that the whole section is taken from Hrachowitz et al. (2014). This should be stated right at the beginning of the section and specific contributions should be clarified. Overall, model calibration and evaluation steps need further explanations and clarifications in the current manuscript. First the selected likelihood measure is not mathematically correct and needs further discussion on the validity and specifically selection of parameter p=10. The authors should include a figure showing only a selected single 2-D representation of the 4-D objective function space to show the projected pareto surface together with the uncertainty intervals. The readers will then be able to understand the calibration and uncertainty analysis procedures with above information. Also, authors state that 13 signatures were used for evaluation, however there is no description of these signatures (perhaps only four is given at an abstract level; Page 5673-Line 22). The signatures used should be explicitly stated”.

Reply: Results from the previous work in Hrachowtiz et al. (2014) are used as a basis of the present work. They are presented in sections 2.5 and 2.6 of the Material and Method section as they are considered as previous results/knowledge. To clarify this, as suggested by the reviewers, we propose:

(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.”

(iii) and 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 in a new version of Figure 3 (provided at the end of this reply).

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

(iv) In spite of ongoing discussions of the most suitable technique to assess uncertainty and criticisms of GLUE for using formally incorrect likelihoods (e.g. Beven, 2006, 2008; Mantovan and Todini, 2006; Stedinger et al., 2008; Montanari et al., 2005; Clark et al., 2012; Hrachowitz et al., 2013a), the use of GLUE with its informal likelihood measures, as used here, has proven valuable in many studies in the past. Fully acknowledging the limitations of the approach, we would, however, also argue that other approaches using formally correct likelihoods suffer from other limitations. To further answer the comment, using likelihood measures that are unweighted or that have relatively small weighting exponents result in a relatively high sensitivity of uncertainty intervals to the choice of the threshold. In this study, we used a relatively high exponent p = 10 (which is still lower as p=30, as tested by Freer et al., 1996) to weight the informal likelihood measure (here: DE), so as to give higher weights to models with good performance and penalize models with poor performance, in an attempt to reduce the impact of subjectivity of the choice of feasible parameters in the results. High values of p significantly reduce the sensitivity of the uncertainty intervals of the modelled variables to changes in the subjectively chosen thresholds. In other words, the value of the threshold becomes irrelevant with high p as all performance thresholds will produce effectively identical uncertainty intervals. However, we believe that a more detailed description of the calibration and uncertainty assessment strategy used here (in particular as it is in detail given in the referenced manuscript) is somewhat out of the scope of this manuscript and will distract the reader from the actual story.

4.“A sensitivity analysis investigating the sensitivity of the hysteretic pattern simulation to the model parameters will significantly improve the manuscript. Currently it is not clear whether the improvements are solely due to the increase in the number of model parameters, or rightly due to addition of new conceptual component to the model as the complexity is increased”.

Reply: The previous study aimed at proposing a stepwise modelling approach where increasing model complexity (and increasing model number of parameters) is 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. To provide the reader more information about this point, we suggest adding the relevant results from the previous work in the new version of Figure 3 with illustrations of model performances on the objective functions and on the signatures, for both calibration and independent evaluation period. We agree with the reviewer and therefore we included a sensitivity analysis of HI on basis of the parameter sets retained as feasible (see also replies to the other reviewers). According to this analysis, when looking only at one of the hysteretic relationships such as the Hillslope saturated zone-flow, the increase of parameter number does not help to improve the hysteresis modelling, but taking into account the hysteresis between all components simultaneously, the increase of complexity allows for an improved overall simulation.

“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.“

5.“Low flow signatures vs. hysteresis patterns: Overall it can be seen that (e.g. Page 5680, Lines 22-23) hysteresis patterns are associated with the low flow signatures as expected. Although authors state briefly, an analysis showing the correlation between the low flow signatures to the hysteretic patterns should be provided. Perhaps a figure could be added showing the low flow signature performance vs. hysteretic index performance of different model structures. Currently low flow signatures are not analyzed independently in the manuscript to investigate their link to the hysteresis signature”.

Reply: at this stage are HI performances are not measured the problem is that we already know that HI is not sufficient to fully describe the hysteresis (as discussed in the 3 reviews in comments related to HI) and therefore was used as an assessment quantitative index but not as a calibration criterion. As discussed in section 3.3, perspectives from this work are a second step where a set of indices would be used to build a process-based objective functions usable for calibration. From our point of view a lot of work is still required to achieve such a set of indices, which would be the focus of a future work.

6.“Sensitivity of HI to annual rainfall (Section 3.1.5): The annual hysteretic patterns are sensitive to the timing of rainfall however the sensitivity to the annual rainfall is tested. My concern is that the HI will be sensitive to when the rainfall occurred (recharge period, recession period etc.) but less on the total annual rainfall. A discussion is needed”.

Reply: This is an interesting point raised here. Indeed, annual rainfall is only a proxy for annual recharge. In our case, it seems that in fact when the annual amount of rainfall increases it is related to an increase of precipitations mostly in the wet season i.e. the high flow period. Therefore, when the amount of annual rainfall increases, the recharge is almost unchanged while the recession is delayed. If the beginning of the recession is delayed, groundwater levels (saturated zone storage) will be still high when stream flow reaches the mid-point value.

7.“The authors presented the degree of hysteretic pattern mismatch between models with different complexities (Figures 9 and 10). A modeler will immediately wonder why the authors did not re-calibrate their models to improve the hysteretic patterns? Calibration using hysteretic patterns could provide additional information on the validity of the model structures and help to understand trade-offs in matching flow based vs storage based signatures. This comment is also linked to Comment 6 which is related to Sensitivity Analysis”

Reply: As in the previous study, Hrachowitz et al. (2014) did not use the signatures in the calibration procedure but rather use them as a post-calibration diagnostic tool for complementary evaluation of the model performances. We kept here the same approach, assuming that hysteresis pattern would provide other information about the model behavior than classical hydrological criteria. Proper calibration on the hysteretic signatures will require a full set of indices to really constraint the hysteretic pattern (see also reply to comment 5) and will be considered in further work.

8. “Conclusion (Section 4): P5686, L4-7: “They were previously calibrated using classical objective functions and assessed using classical hydrological signatures, and their overall performance at reproducing hysteretic signatures was consistent with their overall performance at reproducing the classical signatures. The analysis of the simulated hysteresis signatures helps to identify why the model may give a right answer for wrong reasons” The above statements by the authors are undermining their work and conflicting. According to the first sentence, one can conclude that the classical hydrologic signatures provided the same information as the hysteretic signatures with respect to the overall performance of the model (both say right model or wrong model). The second sentence is then conflicting since classical signatures are already capable of identifying right model from wrong model (right model for right reason). The authors should provide more in depth discussion about their contribution”.

Reply: The first statement is that the general improvement of model performance with increasing complexity observed on the classical hydrological signatures is consistent with a general improvement of the hysteretic signatures too, e.g. from all signatures point of view M4 is more consistent than M1. The second sentence states that the hysteretic signatures might help to identify why a model is wrong when it is assessed wrong by both groups of signatures: e.g. in model M1, performance is decreasing for low flows. It is visible on the low flow signatures and on the hysteresis in dry (low-flows) years. Looking at the HI values, one can see that the model M1 systematically under estimate the unsaturated hillslope storage. In other words, the model calibration tends to put all the water in the saturated storage for reaching the high flow values quickly, therefore in low flow there is not enough water stored in the unsaturated storage and dynamic is wrong. To summarize in other words we could say that classical and hysteretic signatures allow to identify when a model is wrong (or right for wrong reasons) and that additionally, hysteretic signatures allow to identify why the model is wrong (or

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

for which wrong reasons) in a quick and easy way.

We propose to avoid ambiguity the following reformulation: “The tested models were characterized by an increasing degree of complexity and also a 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 and why the models give “right answer for wrong reasons” and can be used as a descriptor of the internal catchment functioning.

9.“The authors should be more careful about the use of plural nouns, e.g. P5684, L23: soils moisture, P5683, L18: parameters sets etc. Please check throughout the manuscript as many more exist.”

Reply: all the manuscript has been reviewed with a particular attention to this point.

Minor comments

-Figure 4: “An explanation on how the recharge and discharge periods are represented on Figure 4 should be included”

Reply: Recharge and recession periods are both indicated in Figure 4 (see new version of Figure 4 below).

-Figure 8 a. “The markers overlap and hence some of them are hidden behind. Improve the marker representation”

Reply: A new version of Figure 8 is proposed at the end of this reply

-p. 5679, line 14: “There is no information related to unsaturated zone in Table 4. Page 5680, Line 24: Figure 3 does not have sub figures a and c”

Reply: The information related to the unsaturated zone is the initial storage of this zone denoted Initial HUS.

-p. 5684, line 6-7: “Clarify the sentence”

Reply: We propose to reformulate as “We argue that rather than increasing the number of constraints or objective functions to satisfy, an alternative could be to use some objective functions based on a combination of different variables as stream flow and the groundwater level, soil moisture, or stream concentrations.”

-p. 5676, line 22. “Clarify the sentence”

Reply: We propose to reformulate as “First, stream flow was close or equal to zero and was almost exclusively sustained by drainage of the saturated storage, while the unsaturated zone exhibited a significant storage deficit and only minor fluctuations due to transpiration and small summer rain events (dry period)”

“Page 5682, Line 16: Replace “discharge” with “recharge”.” “Page 5684, Line 11: Replace “though” with “through”.” “Page 5684, Line 2: “realisms constraints” choose one.” “Page 5683, Line 28: Remove “objectives”.” “Page 5674, Line 13: Replace “lower that” with “lower than that”.” “Page 5679, Lines 1-3: Typos with regard to citations.” All the spelling mistakes and typos regarding to citations raised by the reviewers have been corrected.

Please also note the supplement to this comment:
http://www.hydrol-earth-syst-sci-discuss.net/11/C3675/2014/hessd-11-C3675-2014-supplement.pdf

---

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

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

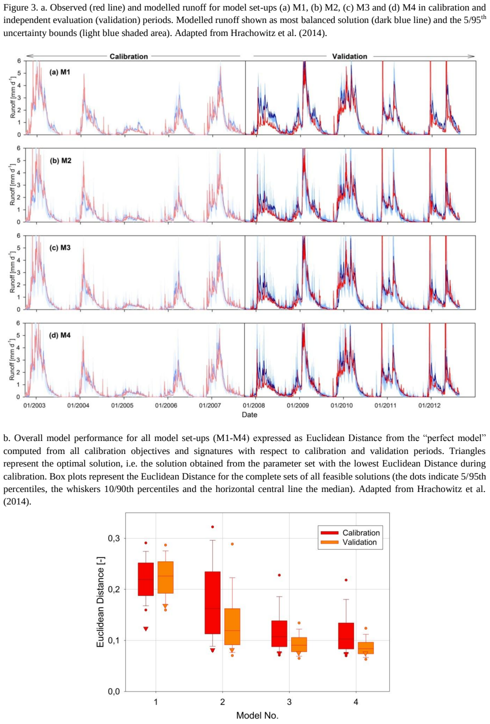


Figure 3. a. Observed (red line) and modelled runoff for model set-ups (a) M1, (b) M2, (c) M3 and (d) M4 in calibration and independent evaluation (validation) periods. Modelled runoff shown as most balanced solution (dark blue line) and the 5/95th uncertainty bounds (light blue shaded area). Adapted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

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

Figure 4. Examples of annual hysteretic loops for saturated zone storage vs. stream flow which are clockwise on the riparian zone (a, b) and anticlockwise on the hillslope (c, d) for the wet year 2003-2004 (a, c) and the dry year 2007-2008 (b, d).

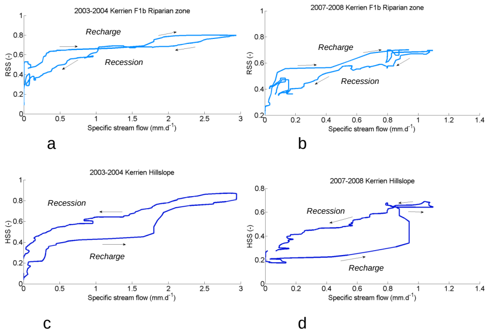


Figure 8. Variations of observed and simulated hysteresis Index versus annual rainfall for the 10 monitored water years for (a) Hillslope Saturated Storage versus discharge HSS(Q), (b) Riparian Saturated Storage vs. discharge RSS(Q). Solid lines indicate the linear regressions. Hydrological years are labelled in italic.

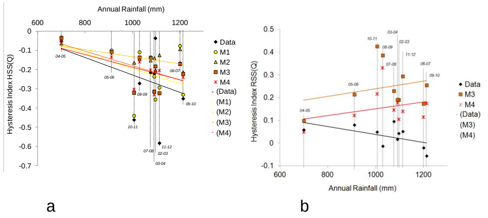


Fig. 2. New Figures 4 and 8

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

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.

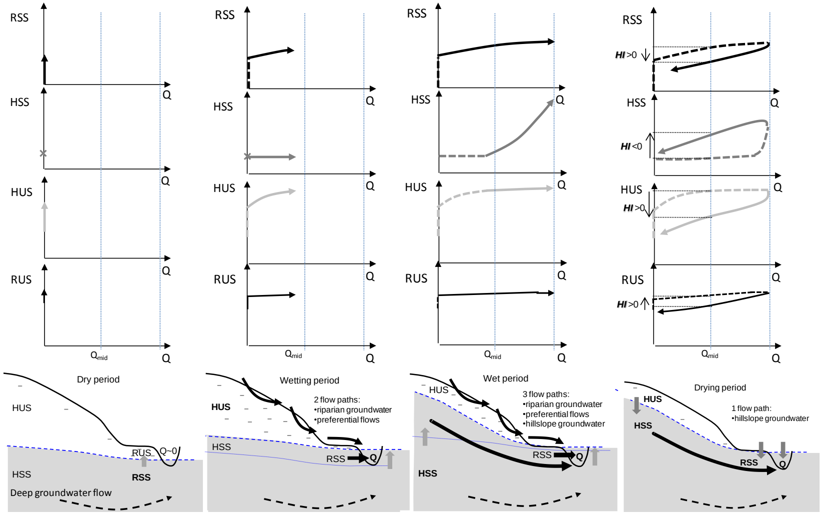
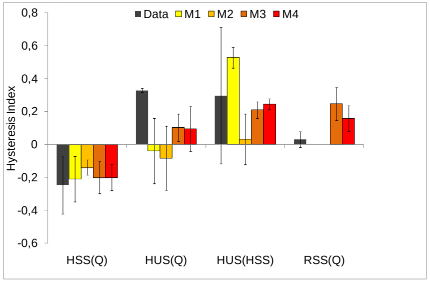


Fig. 3. New Figure 9

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.

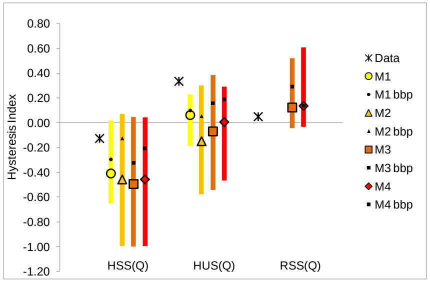


Fig. 4. New Figure 11

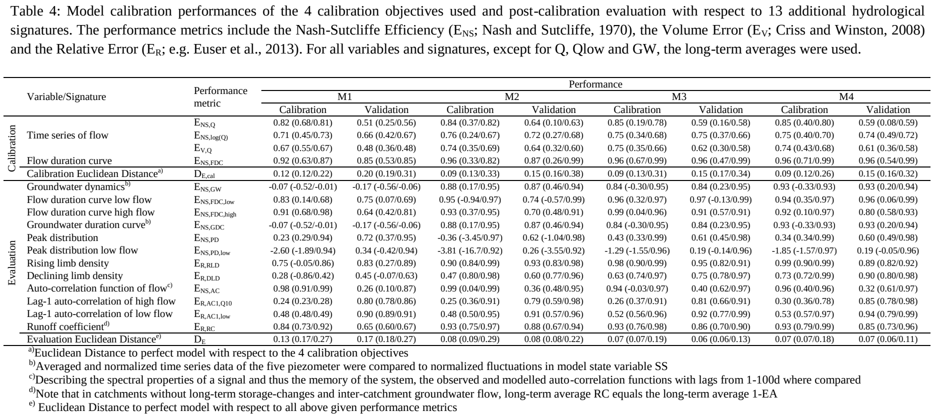


Table 4: Model calibration performances of the 4 calibration objectives used and post-calibration evaluation with respect to 13 additional hydrological signatures. The performance metrics include the Nash-Sutcliffe Efficiency (E_{NS}; Nash and Sutcliffe, 1970), the Volume Error (E_V; Criss and Winston, 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. 5. Table 4