Hydrol. Earth Syst. Sci. Discuss., 11, C3630–C3644, 2014
www.hydrol-earth-syst-sci-discuss.net/11/C3630/2014/

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

"This paper is very well written and surely of interest for the hydrology community. I agree with the Authors in that the purpose of hydrology is not to maximise performance measures but to correctly understand/reproduce what happens (in this case, what are the catchment internal dynamics). This is valid for practical purposes too, since models that can correctly capture the processes going on are expected to be more reliable in predicting the catchment response in conditions non observed in the past. I am definitively supporting for the publication of the paper in HESS. I have some specific comments below, but since they mostly involve additional discussion, from my side the resulting revision should be minor.”

We thank the reviewer for the positive assessment of our article and his/her comments and suggestions which helped us to make our manuscript clearer for the reader and to extend the discussion. Below we reply to each of the specific comments.

Specific comments:

1. “The analysis is done on only one (very small) catchment, while from the title I would have expected more examples”

We agree on the ambiguity in the title using plural form of “catchments” therefore we propose to reformulate as “Hydrological hysteresis and its value for assessing consistency in catchment conceptual models”.

2. p. 5669, section 2.3. “is the normalisation of the storage/saturation values using the minimum and maximum observed values a robust choice? How much does it depend on the record length? How sensitive are the hysteresis indices to this choice? The Authors should add one sentence here to justify that this choice is robust and/or that it has no effect on the results of the study.”

Reply: The hysteresis index is defined as a difference between recharge and recession storage values which correspond to either a normalized groundwater level or normalized soil moisture. As highlighted in the following comment (3) the catchment is characterized by a strong annual cycle with clear recharge and discharge periods. At the inter-annual scale, while stream flow varies in a quite large range due to rainfall variability, the groundwater levels and the soil moisture values are varying within a narrower range of values from one year to the other. The normalization of storage values aims at making the comparison between measurement points (upslope/downslope for the piezometers and depth for the soil moisture sensors) more readable. Using the maximum and the minimum values is a simple solution for normalizing because it is difficult to estimate the actual storage capacity of both unsaturated and saturated zones. In the Hysteresis Index as it is defined here, the normalization is equivalent to dividing the difference between recharge and recession non-normalized values by the maximal amplitude over the records: Denoting Z the groundwater depth (always negative) and θ the soil moisture (always positive) HI could be written as:

HI=(Z(t_R)-Z(t_r))/(Min(Z)- Max(Z))

Or

HI=(θ(t_R)-θ(t_r))/(Max(θ)-Min(θ))

According to the respective ranges of variation of Z and θ, the denominator is always a positive real. Therefore the normalization does not affect the sign of HI. The value of the denominator increases with the amplitude of groundwater level variations or soil moisture variations in the record, thus HI values are likely to decrease when the amplitude of variations increases. However it does not affect our results because: (i)The normalization would tend to increase the absolute values of HI computed from the dowslope piezometers where groundwater fluctuations are lower than in the upland piezometers but we still observed that HI absolute value tends to decrease from upslope to downslope areas so the normalization does not erase this trend. (ii)When HI is used to compare the model to the observations the normalization is done on the same period. As suggested by the reviewer we propose to add a sentence in order to explain this point at the end of section 2.4:

“The normalization of the observed variables related to the storages (here either groundwater level or soil moisture) has no effect on the sign of HI, the HI values are being only divided by the maximal amplitude observed in the storage during the whole period. Therefore, as long as the normalization is applied for the whole period (for all years and for both measurements and simulations), it does not affect the interpretation related to absolute values of HI. “

3. p.5670, Eq. (1). “This definition for the hysteresis index is used by the Authors at

the annual scale. This makes sense in this work because the storage dynamics have an annual period (see Fig. 2). Do the Author expect this to be the case in general? I would think that in other catchments there could be more cycles in one year or even a non-periodic behaviour (in arid climates). However Eq. (1) would still be valid but at the event (rather than annual) scale. If so, a sentence could be added here as a guidance for researchers willing to use the same index in different hydrological settings."

Reply: We agree with the reviewer that the HI index as defined here is useful for characterizing periodic behaviours. Here, the index led to annual values because of the strong seasonal cycle occuring in the studied catchment. The same index could be used similarly for flood events if they are more or less 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 hillslope storage is likely to be less empty at the beginning of the second recharge period than at the beginning of the first one. In snow-melt driven catchments, 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. In order to help the identification of the limits of our HI, we propose to add the following precision before explaining how HI is computed:

"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"

4. p. 5671 Eq. (2). "Related to the previous comment, is the choice of Qmid robust? This is because I expect Qmax to be very variable from year to year and maybe related to short term rainfall response (flood event)."

Reply: 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 in the study site (cf. Figure 4). Even if stream flow is varying considerably among years, these variations are not so abrupt. Moreover forgot to mention that we have used smoothed time series using 7-day moving average. To clarify this important point we added the following explanation in section2.4:

"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)."

However we agree that the relevance of Qmid will depend on the shape of the loop. 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 can be either a concentration or the turbidity in hysteresis studies). If the stream flow maximal values flatten the loop, a better metric of the hysteresis width could be the difference of stream flow values between recharge and recession for the annual mid-point storage value (cf. schematic representations in Figure A)

5. p.5671, lines 15-17. "What do clockwise and anticlockwise hysteresis loops mean from a process point of view?"

Reply: at this stage of the manuscript (Material and method section) the processes are not further developed as the interpretation of the underlying processes is discussed in the results and discussion section, but the direction of the hysteresis indicates immediately which variable (storage or flow) is reacting first to rainfall. We propose the

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

following additional explanations related to HI information:

“HI is a proxy for the importance of lag time response between variations in catchment storages (unsaturated and saturated) and stream discharge, its sign indicates if storage reacts before or after the stream flow.”

6. p.5671, line 26. “how does this work differ from Hrachowitz et al. (2014)? That paper is under review in WRR and has a title which could be the title of this manuscript, although more general. A sentence should be added in the introduction (and maybe also here) to clarify what are the different contributions of the two papers.” Reply: The contribution of the previous work from Hrachowtiz et al., used as a basis of the present work, is detailed in sections 2.5 and 2.6 of the Material and Method section as it is 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 :

“ (...) 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 Figure 3 has been revised to provide the hydrographs on both calibration and evaluation periods (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).

7. p. 5674, line 18. “in Hrachowitz et al. (2014) more model structures were considered while here just four of them are analysed. What is the rationale for the choice of these four?”

Reply: The selection of only 4 of the 11 models from the previous work of Hrachowitz et al. (2014) has been motivated by the fact they correspond to the main different model architectures. The other models are rather constrained differently. We agree that this rationale should be explained and we propose to add this precision at the beginning of the section 2.5:

“Four of these 11 models (noted M1 to M4) were selected for the present work, as they correspond to the sequence of model architectures that provide the most significant performance improvements.”

8. p. 5677, section 3.1.3. “I like this section a lot. Just a suggestion: a figure/schematic that illustrates the mechanisms leading to opposite directions of the hysteresis loops in the hillslope and riparian zone (hypothesis 3) would be very useful (here or later).”

Reply: We thank the reviewer for this suggestion: we propose to add this new figure for illustrating our interpretations in terms of mechanisms (see new Figure 9 at the end of this reply).

9. p. 5679, sections 3.1.4, 3.1.5. “Maybe also the sensitivity to Qmid could be explored. Do the results change if the second annual peak is chosen as Qmax?” Reply: This would be interesting but as explained in reply to comment 4 we are working on smoothed data in order to eliminate the highest Q values due to rapid storm. However, we agree that its sensitivity analysis would be needed if the index was to be used e.g. to calibrate the models. For this purpose, we suggest that a range of indices rather

than a unique one should be used to fully describe the hysteresis. Therefore, the sensitivity analysis would be worth when all this range of indices will be defined. What we propose to add from the revision process is an investigation of the sensitivity of HI values to the parameter uncertainties (see below and also comments of the other reviewers)

10. p. 5683, lines 25-27. “The Authors state that “...a model able to represent the internal catchment behaviour will generate a wrong discharge value but consistent with the storage value and will be rejected in traditional calibration procedures”. This is a very valid point. If the Authors could show that this actually happens in the study they made, that would be great. The model parametrisations chosen for the analysis are optimal in maximising the performance measure Eq. (5) (page 5674, lines 23-26). It would have been very interesting to find out whether non-optimal parametrisations result in better modelling of the hysteresis.”

Reply: 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). (see end of section 3.2.1 and Figure 11b). It seems that the best parametrisation provides also the best hysteresis modelling.

“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."

11. p. 5685, line 25. "please recollect what are the four periods mentioned here."

Reply: We proposed to summarize the four periods in the conclusion as following:

"Four periods have been identified along the hydrological year: 1) first, at the end of the dry period, rainfall starts to refill unsaturated storages; 2) in the wetting period, riparian unsaturated storage is filled and the saturated storage starts to supply the stream while hillslope unsaturated storage is still being replenished; 3) during the wet period, unsaturated storage in the hillslope is also filled and the saturated hillslope storage also feeds the stream. Finally when rainfall declines, flow from the riparian groundwater recedes and during the recession period, the stream discharge is sustained only by hillslope groundwater."

12. Figures 7 and 8: "just a suggestion: the years could be associated to the points in the graphs (e.g., "03-04", "10-11") so that the relationship with the other figures can be seen explicitly."

Reply: We agree. A new version of Figures 7 and 8 is proposed at the end of this reply.

13. Figure 9: "please indicate the direction of the loops".

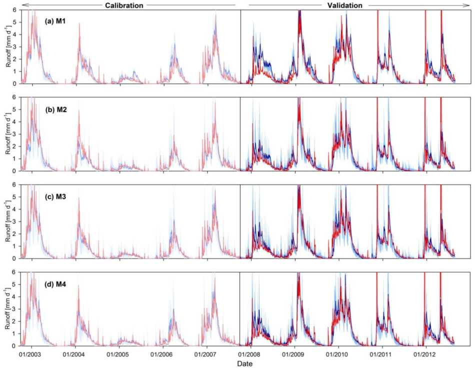
Reply: The direction of the loops has been added on Figure 10 (previously Figure 9) at the end of this reply.

Please also note the supplement to this comment:
http://www.hydrol-earth-syst-sci-discuss.net/11/C3630/2014/hessd-11-C3630-2014-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).

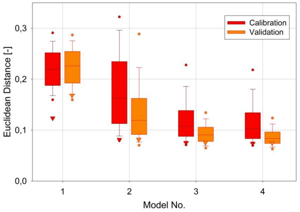
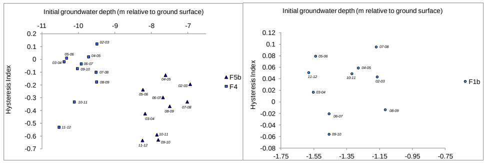


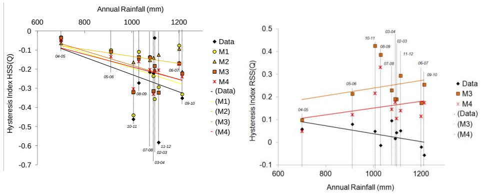
Fig. 1. new Figure 3

Figure 7. Year to year variations, for the 10 monitoring years, of the hysteresis indices a) HSS-F5b(Q) and HSS-F4(Q) (HI) versus the initial groundwater table level depth in the corresponding hillslope piezometer (F5b or F4) and b) HSS-F1b(Q) versus the initial groundwater table level depth in the piezometer in the riparian area (F1b). Hydrological years are labelled in italic.



a b

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. Thin gray lines mark the hydrological years labelled in italic.



a b

Fig. 2. new Figures 7 and 8

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.

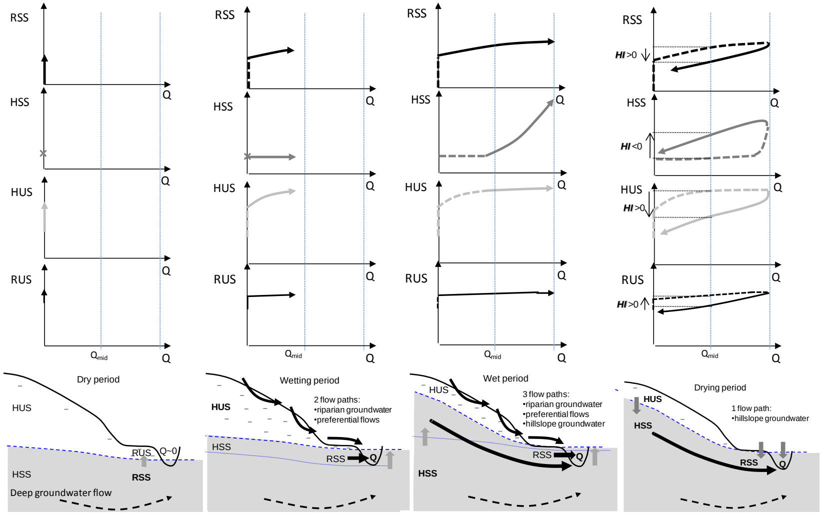


Fig. 3. new Figure 9

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

Figure 10. Observed and simulated annual hysteresis between stream flow (Q) and (a, b) Saturated Storage in the hillslope HSS (for observed, HSS is the average of F5b and F4) and (c, d) Saturated Storage in the riparian area RSS (for simulated, only M3 and M4 represent the riparian area), for the water years (a, c) 2003-2004 (wet year, calibration period) and (b, d) 2007-2008 (dry year, validation period).

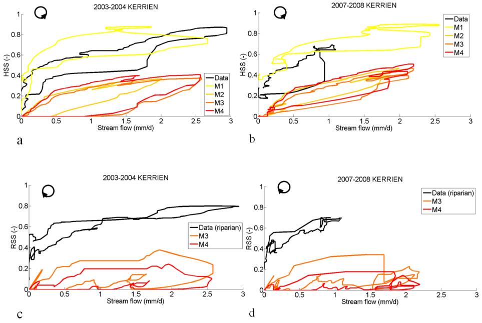


Fig. 4. new Figure 10

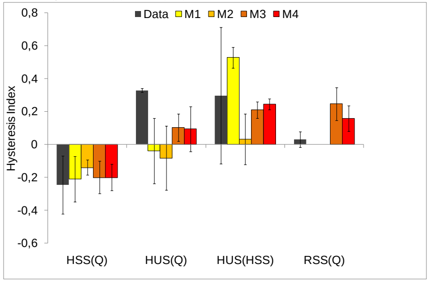
Full Screen / Esc

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.

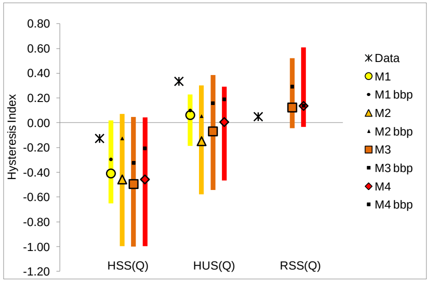


Fig. 5. new Figure 11

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

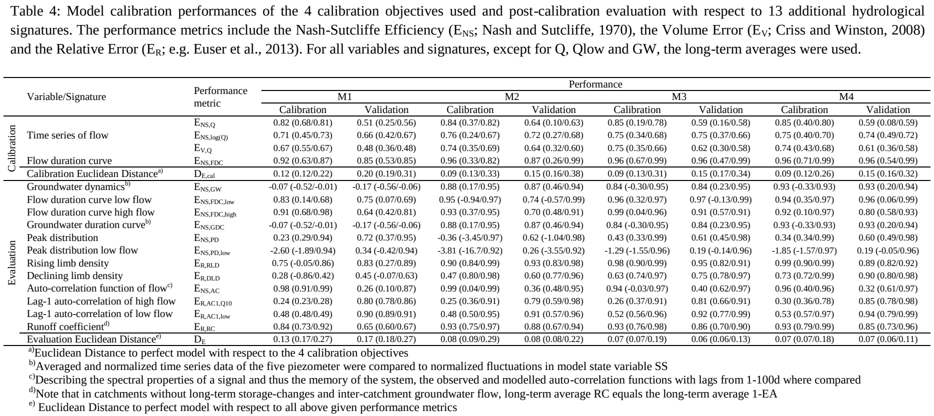


Table 4: Model calibration performances of the 4 calibration objectives used and post-calibration evaluation with respect to 13 additional hydrological signatures. The performance metrics include the Nash-Sutcliffe Efficiency (E_{NS}; Nash and Sutcliffe, 1970), the Volume Error (E_V; Criss and Winston, 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. 6. Table 4