Hydrol. Earth Syst. Sci. Discuss.,
doi:10.5194/hess-2016-546-AC3, 2016

Hydrology and Earth System Sciences
Open Access
Discussions

Interactive comment on "Streamflow characteristics from modelled runoff time series – importance of calibration criteria selection" *by* Sandra Pool et al.

Sandra Pool et al.

sandra.pool@geo.uzh.ch

Received and published: 7 December 2016

Dear Björn Guse,

Thank you for your efforts with our manuscript. We greatly appreciated the comments, which will help us to improve the manuscript. Please see below our detailed response to each of the comments.

Best regards,

Sandra Pool and Co-Authors

Major comments RC 3

Comment 1: In its current state, the article is in my opinion mainly related to hydrology (also by the selection of the journal). At several parts the authors emphasized its ecological importance (e.g. Page 1, Line 10, P.2, L.1-2, P.11, L.14-16). However, I do not really see a connection to ecology. Thus, either the article has to be fully focused on hydrology or it is required to emphasize its ecological relevance

Reply 1: We agree that our modelling approach for estimating SFCs is a typical hydrological application. However, the motivation for our study is strongly related to eco-hydrology because of its focus on the simulation of ecologically relevant SFCs of our study catchments (chapter 2.2). While many of our SFCs are also the focus of other ecohydrological studies, they are untypical for purely hydrological modelling studies where signatures such as segments of the FDC or runoff ratio are of interest. Given the importance of SFCs in ecology we find it important that the hydrological community addresses the issue how to compute these values for ecological studies and applications.

Comment 2: In this article, the Nash-Sutcliffe Efficiency is used as an example for a traditional calibration approach. However, in recent studies the use of several performance measures related to different parts of the hydrological system is recommended. In particular the use of typical performance metrics such as Reff or Kling-Gupta-Efficiency (KGE) in combination with signature measures is recommened (see eg. Van Werkhoven et al., 2009). Thus, I think that a comparison only with the Reff is not sufficient. I recommend to use two or three performance measures such as PBIAS or KGE to show that SFCs also outperforms a calibration approach based on NSE in combination with PBIAS (or other performance measures).

Reply 2: We agree that combined objective functions based on e.g. Nash-Sutcliffe efficiency (Reff) and volume error are widely used for runoff model calibration. We actually used such combined objective functions (that are not based on SFCs) in our previous study (Vis et al., 2015) to estimate the same SFCs of the same catchments as in the current study. The average SFCs error (percent error) in calibration was lowest for calibrations with Reff (error of -4.9 %), followed by calibrations with objective functions based on a) Reff, LogReff and volume error (error of -5.1 %), b) Reff and volume error (error of -5.6 %) and c) Reff, MARE, spearman rank correlation and volume error (error of -6.1 %). Some other combined objective functions were also tested, but resulted in clearly poorer SFC estimates than the ones listed above. We therefore decided to use Reff as a benchmark in the current study. We plan to add some information about the results of the preceding study in the introduction of our manuscript.

Vis, M., Knight, R., Pool, S., Wolfe, W., and Seibert, J.: Model calibration criteria for estimating ecological flow characteristics, Water, 7, 2358–2381, doi:10.3390/w7052358, 2015.

Comment 3: Moreover, it needs at least to be discussed how a calibration approach based on these SFCs is related to recent studies using hydrological signatures such as segments of the FDC (see Yilmaz et al., 2008, Pfannerstill et al., 2014).

Reply 3: Thank you for making us aware of the study from Yilmaz et al. (2008) which we will mention in the introduction in addition to the study of Pfannerstill et al. (2014). These two studies use hydrological signatures to better predict the overall hydrograph and thus their main goal slightly differs from ours. We plan to briefly compare our results with theirs in the discussion.

Comment 4: P. 1, L. 24-27: I do not fully agree with this statement. There are different ways of how to calibrate a discharge time series. Certainly there are studies which are focused on certain parts such as on high or low flows. However, other studies aim to represent the whole hydrological system at best without neglecting or emphasizing certain parts. In the way towards a good representation of the hydrological system, the latter one should be the general goal and a strong focus on certain parts of the hydrological system should be a specific case.

Reply 4: We agree that some general agreement often is the calibration goal. Our statements aims at the fact that even with such a general agreement, specific aspects

might be poorly simulated as we have shown in the preceding study (see Vis et al., 2015).

Comment 5: P.2, L. 21: Here, SFCs are defined "as equivalent to hydrological signatures“. In this case, I do not understand the use of SFC. It is stated before that hydrological signatures are a more common term in hydrology. It is certainly required to justify why you used the term SFC since I do not really see the strong relationship to ecology.

Reply 5: As described in the reply of comment 1, the selection of the SFCs is motivated by their ecological relevance in the study catchments. While both terms, hydrological signature and SFC, describe characteristics of the hydrograph, only the term SFCs makes a distinct connection to ecology. The term SFC has also been used for many years, whereas the term hydrological signature has been introduced more recently.

Comment 6: I think that the article would benefit from a more detailed interpretation of the results. For example on P. 7. L. 12-15; P.8, L.17-18: Can you explain these results or more specifically the behaviour of these SFCs?

Reply 6: We will discuss the behavior of the indicated exceptions in the revised manuscript.

Comment 7: It could be interesting to analyze the relationship/correlation between the SFCs.

Reply 7: We did a Spearman rank correlation test for the 13 SFCs using all 25 catchments. We did the correlation test for both modelling time periods and figure 1 (in this document) shows the average correlation value of the two time periods. The correlation values of the two time periods were similar.

Comment 8: The combination of different metrics might outperform in general single-metric approaches. However, the more metrics are included, the more a trade-off might occur and the equifinality problem arises. In this context, can you give a recommendation for a good number of required SFCs? In the best case, a systematic way of how to select the best SFCs will be provided. Even though when I expect that it is difficult to find a precise number, it is worth discussing this point.

Reply 8: Adding more criteria into a combined objective function usually rather decreases than increases the equifinality issue according to our experience and other studies, although you are right that also an increase because of compromise situations might be possible.

We systematically selected SFCs for a multi-objective function based on the criteria of information value and robustness. Our results indicate the importance of jointly evaluating both criteria for the selection of SFCs for a multi-objective function. However, we cannot give a minimum number of SFCs required for such an objective function, because this will depend on the type and combination of SFCs one is interested in. We could show that the four SFCs used for the multi-objective function preserved similar hydrograph characteristics as the Nash-Sutcliffe efficiency. How much value additional SFCs have for the representation of the hydrograph characteristics would have to be evaluated using many more than 13 SFCs and eventually using synthetic data to also see the effect of redundant information (see discussion chapter 4.3).

Comment 9: The figures 3-8 are very similar (at least visually). I think that the article would benefit from emphasizing the relevance of each figure. It is partly difficult to differentiate them. Maybe you can also thinking about reducing the number of figures to improve the overall message. For example, the figures 6 and 8 have almost the same figure caption. To summarize this point, it is easier to detect the whole message in the case of a more distinct presentation of the results. One example for this is the figure 9 which can be clearly distinguished from the other figures. These results are easier to understand.

Reply 9: We agree that some figures are similar and maybe hard to distinguish. We will try to change or reduce the most similar figures to present the results in a more

interesting or intuitive way in the revised manuscript.

Comment 10: P. 8, L.12-13: Could you specify how you can here differentiate between error dependence on time period or objective function?

Reply 10: We were interested in finding systematic patterns in the error magnitude (absolute value of normalized SFC error) of the SFCs. For some SFCs (e.g. TL1) the error magnitude could be considerably higher in one modelling time period than in the other one. We described the error of such SFCs as being time depended. Some other SFCs had clearly higher error magnitudes when calibrated on a certain objective function (e.g. MH10). The estimation accuracy of such SFCs was considered as being dependent on the objective function.

Comment 11: P.8, L.14-15: I do not understand this statement that the SFCs are neither related to flow components nor to flow conditions. Hydrological signatures (as an equivalent term) are known to be of special importance to explain the hydrological behaviour. Thus, what can we learn from using these SFCs in terms of the hydrological behavior in the catchment. And how is this related to the general idea of the hydrological signatures?

Reply 11: We are sorry for the confusion this statement evoked and will make the sentence more clear. We wanted to say that we could not relate the estimation accuracy (error magnitude, spread of error magnitude and dependency of error magnitude on modelling time period/ objective function (see comment 10)) of the analyzed SFCs to the flow components or flow conditions they belong to. E.g. we cannot say that all high-flow related SFCs had very low estimation accuracies.

Comment 12: P.8, L.26 to P. 9, L.21: I agree with this part which is clearly understandable, but certainly also not surprising. It is mostly existing knowledge of hydrological modelling. What can we learn here except of using several and different metrics. I recommend to shorten this part and emphasize the most important points from this study. In contrast, I really like to following passage (P. 9, L.21-31). Please also discuss the impact of a SFC-based calibration for the process representation. Can you state that the hydrological system is overall better represented by using several SFCs? Please discuss the benefit of optimizing one specific SFC. This leads to a modelled hydrograph which is able to represent a very specific condition but probably not the overall hydrograph. This implies that the part of the hydrological system which is not in the focus of this SFC is probably not adequately considered. This might be of particular relevance when using very specific SFCs such as MA26.

Reply 12: Thank you for this helpful comment. We will adapt the discussion part you mentioned focusing on your proposed aspect of how the calibration with SFCs affects process representations:

The benefit of optimizing one specific SFCs lies in the relatively accurate estimation of the respective SFC compared to a calibration with Reff or a multi-SFC objective function. Model calibration on one single SFC clearly emphasizes the hydrograph aspects of the selected SFC which can negatively affect other hydrograph aspects. This implies that calibrations with ISingle can lead to poor model performance for SFCs not included in the objective function (figure 4). E.g. calibrating on the frequency of moderate floods (FH6) leads to poor model efficiencies for baseflow (ML20) (figure 7b) which indicates that the representation of the main runoff processes can suffer by SFC-specific model calibration. From the fact that a calibration with Reff and a calibration with multiple SFCs lead to comparable SFC estimates we infer that the main hydrological processes of the catchments are similarly well represented with the two approaches. We assume that these two calibration criteria result in a better process representation than the calibration with a single SFC, because they outperform the calibration with ISingle for those SFCs not included in ISingle.

Minor comments RC 3:

P.1, L. 12: maybe "optimization“ instead of "minimization or maximization“.

Reply: We will do the replacement as suggested.

P.1., L.16: Are these over- and underestimations a general aspect of these SFCs or are they case-specific?

Reply: These results represent the general tendency of the model and the results from a specific objective function and/or modelling time period can deviate from these general conclusions (figure 9).

P.2, L.16-19 and L. 30: I recommend to include the study from Yilmaz et al. (2008).

Reply: Thank you for this suggestion. We will add the study of Yilmaz et al. (2008) to the mentioned part in the introduction.

P.2, L. 34: The meaning of "esoteric and subtle aspects of the flow regime“ is unclear.

Reply: Good point, we will adapt the sentence.

P.3, L. 19: Please think about renaming the section to "Methods and materials“, since a catchment is not a method.

Reply: We will change this title as suggested.

P.5, L.31: Why you have used 0.2 and 0.25 as weights?

Reply: IMulti consists of four SFCs that we wanted to weight equally which leads to weights of 0.25 for each SFCs. These four SFCs were combined with the Nash-Sutcliffe efficiency (IMulti_Reff) and each of the components was again assigned the same weight (0.2). We weighted all components of IMulti and IMulti_Reff equally because there was no convincing reason to weight one/some SFC(s) more than others.

P.6, L. 9: Could you specify "median parameter set“?

Reply: This sentence should say “. . ., the median model efficiency of each catchment was selected.” We will change this sentence.

P. 11, L.15: Could you specify "later application of simulated SFCs related to flow alteration – ecosystem change relationships“. This aspect was up to now neither in the focus of the article nor emphasized as an overall aim.

Reply: This statement refers to the first paragraph of the introduction where we motivate the need for accurate SFC estimates by giving the example of flow alteration – ecosystem change relationships used for sustainable flow management. We will adapt the sentence in the discussion to make this connection more clear.

Table 1: TA1: runoff with two f

Reply: We will add the missing f.

Table 1: Why you have named the SFCs FH6 and FH7 and not FH3 and FH7.

Reply: We agree that the abbreviation can be confusing, but decided to follow the abbreviations used in previous publications and in the EflowStats R-package that was used for the calculation of the SFCs. The same abbreviations are commonly used in many studies (see e.g. Olden and Poff, 2003).

Olden, J. D., Poff, N. L.: Redundancy and the choice of hydrologic indices for characterizing streamflow regimes, River Research and Applications, 19(2), 101-121, doi: 10.1002/rra.700, 2003.

Fig. 5: Can you explain the outlier TL1?

Reply: We actually already discussed the low estimation accuracy of TL1 in the discussion (last paragraph of chapter 4.1).

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-546, 2016.

	Mean-flow				Low-flow				High-flow				
	ma41	ma26	ra7	ta1	e85	fl2	tl1	ml20	mh10	dh13	dh16	fh6	fh7
ma41	-												
ma26	-0.529	-											
ra7	-0.449	0.843	-										
ta1	-0.844	0.653	0.632	-									
e85	0.803	-0.811	-0.801	-0.852	-								
fl2	0.213	-0.506	-0.697	-0.395	0.500	-							
tl1	-0.393	0.130	0.120	0.377	-0.165	-0.097	-						
ml20	0.546	-0.894	-0.962	-0.736	0.869	0.688	-0.148	-					
mh10	0.840	-0.425	-0.402	-0.725	0.689	0.198	-0.475	0.463	-				
dh13	-0.650	0.864	0.867	0.804	-0.918	-0.600	0.120	-0.945	-0.540	-			
dh16	0.513	-0.646	-0.722	-0.642	0.678	0.600	-0.319	0.750	0.545	-0.650	-		
fh6	-0.484	0.865	0.902	0.632	-0.796	-0.662	0.103	-0.934	-0.364	0.876	-0.677	-	
fh7	-0.614	0.884	0.895	0.762	-0.901	-0.609	0.112	-0.947	-0.503	0.966	-0.666	0.917	-

Fig. 1. Spearman rank correlation coefficients for the 13 SFCs.