Hydrol. Earth Syst. Sci. Discuss., 9, C1414–C1418, 2012
www.hydrol-earth-syst-sci-discuss.net/9/C1414/2012/


Hydrology and Earth System Sciences Discussions



# *Interactive comment on* "Moving beyond traditional model calibration or how to better identify realistic model parameters: sub-period calibration" *by* S. Gharari et al.

**S. Gharari et al.**

shervangharari@yahoo.com

Received and published: 11 May 2012

The authors would like to thank the anonymous referee for his/her constructive comments which will help us to improve our discussion paper (DP) further and make it transparent for the readers.

**1- Page 1889, l. 6. What is the difference between multi-objective and multi-criteria? In general, the multi-objective calibration problem may be defined with respect to (i) multiple variables, (ii) multiple response modes (e.g. high flow**

**and low flow), and (iii) multiple sites (in case of distributed model output).**

We agree with the other reviewer that the terms multi-criteria and multi-objective should be better defined. As we wrote in our reply to the anonymous referee 1 what the reviewer refers to as "multivariate" is referred to in the DP as "multi-criteria". In other words, multi-criteria calibration is defined here as calibration with respect to different criteria, such as stream flow, groundwater fluctuations or tracer response. On the other hand, "multi-objective" calibration is defined here as calibration with respect to multiple objective functions of a given criterion, such as RMSE and NSE of stream flow.

**2- Page 1890, l. 6. How is distance defined? In the application example Euclidian distance of the objective functions is used. However, this distance measure will be sensitive to the magnitudes and units of the objective functions. Some transformation is needed to obtain a consistent distance measure (e.g., Madsen, 2000).**

We are aware of the work of Madsen (2000) and we agree that the way to calculate the distance to the Pareto front involves critical assumptions. We clarified in the revised paper that the distance can be calculated in different ways, also using normalized values or any other measures. However in this study for simplicity we decided to use Euclidian distance, which is sensitive to the different scales of the objective functions. It is also possible that the distance be presented in different direction for example the horizontal and vertical distance to Pareto front which is independent of the scale of each objective function and increase the optimization dimension of second step of SuPer calibration by factor of 2; however due to the fact that those distances are correlated (when one approaches zero the other also approaches zero) we decided to use a single distance.

We would like to emphasize that the number and length of sub-periods, number and type of objective function(s), number and type of calibration criteria, the actual optimization algorithm as well as the distance metric for computing "distance to Pareto front" is entirely up to the discretion of the modeler.



HESSD

9, C1414–C1418, 2012



**3- Page 1894, l. 5-10. The objective functions do not explicitly define high flow and low flow performance. Performance measures for flows above a "high flow threshold" and below a "low flow threshold" would be more consistent.**

What we meant by high flow and low flow is the calibration focus. By using RMSE and RMSE of log of the flow the emphasis of calibration changes from fitting the peak flow to mimicking the low flow. We will elaborate our intention of using this metrics more precisely in the DP.

**4- Page 1897, l. 15. What is "normalized cumulative frequency"? I expect you mean cumulative distribution function (cdf).**

That is true we mean cumulative distribution function (CDF). We will change this sentence to make it clearer.

**5- Page 1898, l. 10-12. I think the results are quite similar with the multi-year calibration results. Please elaborate.**

We will make this explanation clearer, actually what we intended to express is that range and behavior of parameter sets inferred from SuPer calibration is wider and less identifiable compared to other panels which used traditional calibration, i.e. reducing false positive and false negative (type I and II error) by widening the feasible parameter range and by not excluding parameterizations that could actually be feasible ones.

**6- Page 1898, l. 14-16. I don't think the results strongly support this statement. Obviously parameter estimates will be very different when based only on one year of data. Larger records are needed to investigate the performance of the proposed framework compared with the traditional calibration approach.**

We agree with this statement; there is less heterogeneity between longer time series. That is why it became a common practice for hydrological community to calibrate and



HESSD

9, C1414–C1418, 2012



validate on longer time series because the model tends to behave similarly in less heterogeneous time series (Seibert, 2003). This is the entire issue that we want to draw the attention of hydrological community to: By calibrating to longer time periods, the characteristics of certain sub-periods are averaged out and the best average parameterization is estimated. By using Sub-Period calibration, it is possible to keep and consider the different characteristics of different time periods in the final overall parameterization.

As we mentioned in the reply to the anonymous referee 1, we will incorporate evaluation of parameter sets selected by different combination of Super calibration and traditional calibration in completely different time period with different time length to make it clear for the readers.

**7- Page 1899, l. 24 - Page 1900, l. 5. I don't think this is a particular feature of the proposed method. A traditional calibration approach will also put less emphasis on few storms that are not well represented in the precipitation record in the optimization.**

After some consideration, we have to agree with the reviewer. Sentence removed.

**8- Figure 1. Not clear what this figure illustrates.**

The picture illustrates the concept of SuPer calibration. The first step of SuPer calibration is to calibrate each sub-period separately and find the best performance of each sub-period. The second step is to sample the parameter space and determine the parameters which perform closest to the best performance of sub-periods. We will elaborate it more in manuscript and caption.

**9- Figure 7a. Partly same information as in Figure 4.**

We agree with this statement. We preferred to repeat the entire result with different



HESSD

9, C1414–C1418, 2012



combination and calibration strategy for readers to compare.

**10- Figure 10. Not clear what this figure illustrates and is not discussed in the text.**

The figure is explained in page 1898 line 18-20. It illustrates the selection of best parameter set from Pareto members based on subjective thresholds. Mostly the balanced points were selected (the closest point or area to the origin). However SuPer calibration has the capability to select other parts of Pareto front, which may be different from the closest to the origin.

**11- The written English should be improved.**

We will improve the use of English language in the final version of the paper.

Once again the authors would like to thank anonymous referee 2 for his/her constructive comment on our work. We hope that having addressed these comments in our work meets the demands of the referee and improves the transparency of the paper.

**References**

Madsen, H.: Automatic calibration of a conceptual rainfall–runoff model using multiple objectives, Journal of Hydrology, 235(3-4), 276–288, 10.1016/S0022-1694(00)00279-1, http://www.sciencedirect.com/science/article/pii/S0022169400002791, 2000.

Seibert, J.: Reliability of model predictions outside calibration conditions, Nordic Hydrology, 34(5), 477–492, 2003.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 9, 1885, 2012.