Hydrology and Earth System Sciences Discussions

Hydrol. Earth Syst. Sci. Discuss., author comment AC1
https://doi.org/10.5194/hess-2021-514-AC1, 2022

Reply on RC1

Yan Liu et al.

Author comment on "Pitfalls and a feasible solution for using KGE as an informal likelihood function in MCMC methods: $DREAM_{(ZS)}$ as an example" by Yan Liu et al., Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2021-514-AC1, 2022

We thank the reviewer for the comments that help us to improve our manuscript. In the following text, **the reviewer comments are shown in regular font, and our point-by-point replies are shown in *italic***. Upon revision we will make the following major changes to the manuscript and Supplementary Material:

- ***Introduce one case study to show that the adapted KGE approach works while the formal likelihood fails to highlight the advantages and the need to use the adapted KGE;***
- ***Provide the analysis with the known analytical solution of posterior in one of the case studies and compare results derived using adapted KGE to it;***
- ***Compare the performance of the adapted KGE with the GLUE framework and another formal likelihood function.***

The authors proposed an informal likelihood function based on KGE (with modifications), and demonstrated its performance against a formal likelihood function based on RMSE in DREAM_ZS with three cases. There are several key questions that were not clearly answered.

1. Why should one use the KGE-based informal likelihood function? Why Gamma distribution? It seems that it is not advantageous over the formal likelihood function in the three case studies. It would be essential to design a case where the formal likelihood function would fail while the KGE-based one still works. Simply introducing a new metric (without solving challenging problems) has no significance.

Response: The motivation of proposing this adapted KGE is that KGE is widely used as the performance measure in hydrological studies and also used as objectives for calibrations. However, we have seen some flaws using the original KGE in MCMC-type calibrations. Gamma distribution is an easily applicable distribution function and can solve the two problems for using the original KGE: (i) ensure the monotonically increase of probability density even with negative KGE values, and (ii) achieve a proper nonlinearity of performance increase due to the increase in KGE. They can lead to an efficient and proper chain evolution. Another reason is that among other functions we tried, the Gamma distribution is better since it does not introduce more parameters to calibrate and maintains the good performance compared to the formal likelihood function. In our case studies 2 and 3 the adapted KGE even has a higher general performance, the mean KGE *of the evaluation, and a smaller bias overestimation of low flows than the formal likelihood function. We will discuss more on the use of Gamma distribution function in the revised version of the manuscript.*

It is a very good idea to include one case study where our adapted KGE works while the formal likelihood fails. This will also highlight the need to use KGE-based informal likelihood function. In the revision, we will include the above mentioned case study and discuss more why the KGE-based informal likelihood function should be used.

2. No theoretical analysis has been provided. At least one case where analytical form of posterior is available should be considered to verify whether the new likelihood can obtain the right answer.

Response: In Case study 1, the true model parameters are known by setting. We compared the performance between the formal and our adapted KGE approach. In revision, we will include one analysis with the known analytical solution of posterior in one of our case studies and compare our results with it.

3. The numbers of unknown parameters are generally small. A case with more than 20 unknown parameters (>100 would be better) is suggested to demonstrate its performance in more challenging settings.

Response: Our approach was developed based on lumped or semi-distributed hydrological models, where the number of model parameters is mostly smaller than 20 to which $DREAM_{(SZ)}$ *is usually applied (Liu et al., 2021; Shafii et al., 2014; Vrugt et al., 2008, 2009). Some other new likelihood measures are also usually tested with simple analytical models or models with similar complexity as ours (Knoben et al., 2019; Schwemmle et al., 2020).*

4. Comparison with other informal likelihood functions (NSE, GLUE, etc.) is lacking.

Response: In revision, we will add the comparison of our approach and GLUE using NSE as the objective in one of our case study. Additionally, we will compare our approach with another formal likelihood functions, such as using the log transformation of model errors as suggested in the major comment 1 by Reviewer #2.

Minor comments

1. Lines 47-48: confused about what is N about.

Response: N is the variable symbol that was used as a parameter. We will make it clearer in revision.

2. Lines 57-60: The proposal should not affect the shape of posterior if the chain is sufficiently long.

Response: We agree that if the chain is long enough, the 'true' shape of posterior can be explored. However, in practice one needs to consider efficiency due to the computational cost. This means a limited number of realizations will be performed. Using the original KGE, the differentiation of very good (e.g. KGE=0.8) and good (e.g. KGE=0.6) in the standard MCMC is small. This will lead to a very fast convergence (indicating by the diagnostic index), which means using the limited realizations and its converged chains will result in a very flat posterior distribution, i.e. the exploration of the shape of posterior is largely affected.

3. Line 82: if the types of observations are different and with different magnitudes, how to calculate the ED metric?

Response: Since the adapted KGE is informal, we can combine multiple KGEs with each KGE for one type of observations (such as the weighted sum). The ED metric will be 1 subtracts the combined KGE. The combination of KGE will be based on the importance of each type of information defined by the user. It will be like using multi-objectives.

4. There is no need to include results of KGE_ori, as they are obviously wrong.

Response: We wanted to show problems exist when using KGE_ori. In revision, we will minimize using results of KGE_ori and put the comparison into supplement.

5. Figures 6 (h-g), curves of KGE_ori and formal are quite different, why? A synthetic case with similar settings is needed to check which one failed to capture the truth.

Response: Curves of KGE_ori and formal are quite different because KGE_ori cannot well explore the posterior. The differences between KGE_formal and KGE_gamma are most probably due to the interactions between model parameters. We will check the auto-correlations of model parameters and analyze other factors to express the reasons in revision.

6. Line 364: capable to->capable of

Response: We will change it.

7. What is equation of the likelihood function based on RMSE? There are also many forms of formal likelihood function (e.g., Table B1 in J.A. Vrugt / Environmental Modelling & Software 75 (2016) 273e316)

Response: We will include the equation in revision. It is the first, "lik=11", in Table B1.

Reference:

Knoben, W. J. M., Freer, J. E. and Woods, R. A.: Technical note: Inherent benchmark or not? Comparing Nash-Sutcliffe and Kling-Gupta efficiency scores, Hydrol. Earth Syst. Sci., 23(10), 4323–4331, doi:10.5194/hess-23-4323-2019, 2019.

Liu, Y., Wagener, T. and Hartmann, A.: Assessing Streamflow Sensitivity to Precipitation Variability in Karst□Influenced Catchments With Unclosed Water Balances, Water Resour. Res., 57(1), doi:10.1029/2020WR028598, 2021.

Schwemmle, R., Demand, D. and Weiler, M.: Technical note: Diagnostic efficiency – specific evaluation of model performance, Hydrol. Earth Syst. Sci. Discuss., (2008), 1–15, doi:10.5194/hess-2020-237, 2020.

Shafii, M., Tolson, B. and Matott, L. S.: Uncertainty-based multi-criteria calibration of rainfall-runoff models: A comparative study, Stoch. Environ. Res. Risk Assess., 28(6), 1493–1510, doi:10.1007/s00477-014-0855-x, 2014.

Vrugt, J. A., ter Braak, C. J. F., Clark, M. P., Hyman, J. M. and Robinson, B. A.: Treatment of input uncertainty in hydrologic modeling: Doing hydrology backward with Markov chain Monte Carlo simulation, Water Resour. Res., 44(12), 1–15, doi:10.1029/2007wr006720, 2008.

Vrugt, J. A., ter Braak, C. J. F., Gupta, H. V. and Robinson, B. A.: Equifinality of formal (DREAM) and informal (GLUE) Bayesian approaches in hydrologic modeling?, Stoch. Environ. Res. Risk Assess., 23(7), 1011–1026, doi:10.1007/s00477-008-0274-y, 2009.