Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2020-543-AC1, 2021

Hydrology and Earth System Sciences
Open Access
Discussions

Interactive comment on "A new fractal-theory-based criterion for hydrological model calibration" *by* Zhixu Bai et al.

Zhixu Bai et al.

zhixu_bai@wzu.edu.cn

Received and published: 1 February 2021

Dear Editors and Reviewers: Thanks for your kind comments about our manuscript. Your comments are not only helpful but also inspiring. The comments provide new perspectives to understand the application of fractal theory in hydrological modeling. We have studied the reviewers' comments carefully and made responses in the following texts. We are looking forward for further advice from you.

Kind regards, Zhixu Bai, Yao Wu, Di Ma, Zixia Wang

Responses to the reviewers' comments:

Reviewer #1:

COMMENT 1 The metric E which the authors are using in their strategy is known to have a number of issues in its application for assessing "goodness-of-fits". Eventually, the need to modify E has been on the radar of hydrologist for decades. In other words, several variants of E exist to address the issues related to the use and interpretation of the original version from Nash and Sutcliffe (1970) which is still widely applied in hydrology. The question to answer is: why did the authors adopt the original version of E but not any of the existing variants?

RESPONSE 1 Thanks for your kind comment. As far as we know, although the different modification versions of E have been studied for decades, there are still no dominant dimensionless coefficients to measure the performance of hydrological models. When only one metric should be used with RD in our study (else the calibration and selection of parameter sets could be too complex to understand the effects of introducing RD), there are not many choices. We finally chose E rather than KGE or other variants of E because the pros and cons of E are more familiar for hydrologists, and this original version is still most often used in hydrological model calibration.

This concept will be included in Section 2.4 E-RD strategy of our final manuscript: "Another reason to choose E in our study is that the pros and cons of E are more familiar for hydrologists than other metrics, and this original version is still mostly often used in hydrological model calibration."

COMMENT 2

RD varies from zero to positive infinity (see line 155 of the discussion paper). However, E varies from negative infinity to zero. The point is that both E and RD are relative error measures. For relative error measure, we focus on the "standard" range in which values vary from zero and one with association to imperfect and perfect model, respectively. Therefore, how can a modeler interpret E and RD in a combined way yet the range of the values from each of these metrics is wider than the "standard" one?

RESPONSE 2

Thanks for your kind comment. RD varies from zero to positive infinity, but the RD value of a perfect model should be equal to 1 because the simulated streamflow series and the observed streamflow series have the same Hausdorff dimensions. We found that a small range of E near the best E in certain cases corresponds to a relatively large range of RD. Besides, there is always a set of parameters makes RD=1 and E close to the best E. Therefore, we applied a genetic algorithm to find individuals with smallest value of objectives. The flow chart is below (Figure 2 in our manuscript).

Figure 2: Flow chart of E-RD strategy. In the multi-objective optimization, we made some adjustments. The objectives used in the multi-objective optimization are 1-E and |1-RD| (see Line 255).

COMMENT 3

There is a possibility in modelling that the larger the number of calibration runs, the better the value of the objective function (especially if the parameter spaces are not small). However, the modeler needs to compute both E and RD in each calibration run as a requirement for the strategy being introduced. Thus, application of the introduced strategy brings about the problem of computation time required to reach optimum during calibration of a hydrological model. How can this problem be addressed to ensure application of the introduced strategy is not at the expense of calibration time (especially if the modeler is making use of long-term hydrological series)?

RESPONSE 3

We made an experiment to show the effects of introducing two objectives into an automatic calibration to the computation time.

We made an experiment to compare the runs needed for finding the best E (single-objective calibration) and the Pareto optimum of E and RD of HBV model used in this study. The calibration algorithm and parameters are the same with those in our original manuscript.

The results show that the multi-objective calibration took 2160 seconds to run 106 generations (63600 individuals) while the single-objective calibration took 1170 seconds to run 51 generations (31150 individuals).

Besides, to overcome the problem of computation time in multi-objective calibration of hydrological models, hydrologists have adopted several different types of methods. In our study, we have adopted NSGA II genetic algorithm and parallel computing technique to accelerate the calibration.

All in all, the introduction of a new strategy will increase the time required, and several methods were adopted. The calibration time has been controlled to a reasonable range in our study. When the E-RD strategy is used with distributed models, more techniques such as parameters' sensitivity analysis could be applied to reduce the number of parameters to be calibrated.

COMMENT 4

The best RD does not guarantee that E will be at its highest value. Furthermore, E reduces as the modeler searches for the best RD (see lines 330-331 of the discussion paper). This brings about (i) the issue of subjectivity in determining which values of E and RD should be used to select the set of optimal model parameters, (ii) the complication in dealing with the trade-off regarding the decision on which study objective should be preferred to others. To explain (ii), the authors need to note that a modeler may be aiming at reproducing extreme hydrological extremes especially peak high flows, and low flows. Applying the E-RD strategy means, the modeler should also aim at ensuring Ds and Do are the same or very close to one another. The question that the authors need to answer is: How can a modeler deal with the issues (i) and (ii) in application of the calibration strategy being introduced?

RESPONSE 4

Thanks for your kind comment. We would like to respond from two aspects. Firstly, in multi-objective calibration, the objectives, generally, cannot be at their highest values at the same time. And if they can, the introduction of multiple objectives becomes worthless because a single-objective calibration is able to achieve the same results. Secondly, we believe that RD could help modelers find the best fractality of simulated series. The improved performance (of low flows in our study) is the by-products of the improvement of fractality. We believe the issue (ii) proposed by the reviewer is not a drawback of our strategy because our strategy improves the simulation of low flows and has little effects on high flows. In other words, our strategy provides a better metric. Based on above, our answer is: a modeler may make gentle adjustments of our strategy to make it more suitable for his/her own cases. But the introduction of RD, by making the Hausdorff dimensions of simulated series and observed series closer, could improve the performance and the internal rationality (components of streamflow in our study) of hydrological models.

COMMENT 5

Sub-flows' separation procedure adopted for this study (incorporated in the tool named WESTPRO) makes use of a number of parameters. The authors never mentioned any values of such parameter in their discussion paper. Examples of such parameters (among others) include recession constants, and the filter parameter. At least two parameter values are required to extract base flow. Again, not less than two parameters are required to filter interflow. Thus, for each river flow time series one requires not less than four parameters to obtain the various sub-flows. The problem is that the choice of this parameters can be largely subjective (even if one takes into account his or her expert judgment in deciding on the parameter values to use for sub-flow filtering of a given streamflow). Moreover, sets of parameters required to separate subflows vary from one catchment to another. Finally, there are several methods available for separation of flows (what we also call the baseflow separation techniques). All these problems compound the challenge of using E-RD to judge model performance (or select which calibration run is the best). Furthermore, the overall problems that the authors need to

take into account, here, are with respect to the uncertainty (i) due to the choice of the baseflow separation technique (whether manual approach as the authors adopted or automated technique), (ii) the subjectivity of selecting which parameter values to use in filtering the observed and modeled streamflow. Here, the fact that the same set of parameter values are required to be applied to both observed and modeled streamflows should be considered basic and they need to go beyond it in addressing this comment. Finally, given the above background on sub-flow separation and unanswered question, statements made by the authors in the manuscript citing that the use of RD improves simulations of sub-flows remain claims (or are vague) unless they prove otherwise.

RESPONSE 5

Thanks for your kind comment. We put the parameters of WETSPRO here. And we are pleased to provide the value of parameters into the revised manuscript. Notably, the WETSPRO tool could separate the streamflow into fast flow and slow flow first, and then separate the fast flow into overland flow and interflow. In our study, only the first step is applied and only the first-step-related parameters of WETSPRO are listed in the table below. We selected the parameters by following the procedure. In WETSPRO's procedure, the parameters are selected one by one. For each parameter/step, there is a corresponding criterion. Thus, the separated streamflow components are relatively objective. Fig. R1-5 is an example of the objective procedure of selection. In this step, the user selects the w-parameter filter, which represents the case-specific average fraction of the quick flow volumes over the total flow volumes. According to the literature, the filtered baseflow should be close to the total streamflow in dry periods (Willems, 2009). The selection can be considered relatively objective.

Fig. R1-5 An example of the objective procedure of selection.

Table R1-5 Parameters of WETSPRO

The description of Table R1-5 is as follows: "Table R1-5 lists the parameters of WETSPRO in three cases. The recession constants are close to each other. The w-parameter filter, representing the case-specific average fraction of the quick flow volumes over the total flow volumes, shows the difference. The w-parameter filter of Dong catchment is 0.14, smaller than the other catchments, meaning that less proportion of total flow in Dong is baseflow, showing the catchment features of small area and high slope."

COMMENT 6

The authors attempted to relate optimal values of the model parameters to obtained RD's. In a number of cases (see, for instance, lines 436 and 505) the authors pointed that the selected model lacked capacity to simulate certain hydrological processes. The question to answer is: Why did the authors not take into account the uncertainty in their results due model selection? Models differ with respect to their structures (or underlying assumption and equations). It becomes imperative that the authors need to select at least two models and apply them to various catchments. In doing so, I suggest the authors focus on clear objectives of modeling so as to allow them comprehensively judge the influence of application of RD on the model results. Such objectives may include reproducing (i) extreme peak high flows, (ii) low flows, (iii) fractality in the observed streamflow. Furthermore, results on comparison of RD with model parameter should be put as supplementary material (if they cannot be discarded from the manuscript).

RESPONSE 6

Thanks for your kind comment. In our study, we analyzed the parameters' behavior when RD is taken into consideration instead. Besides, we trust that our performance of models is good enough for our cases. We would like to add the analysis about the objectives suggested by the reviewer. We agree that more objectives could make our study of E-RD strategy more comprehensive. Table 2 now becomes:

"The high-flow percentiles (Q_5) and low-flow percentiles (Q_75) are reasonable in three cases for all typical models. However, the high-flow percentiles and low-flow

percentiles of best RD models are still closest to the observation."

COMMENT 7

Instead of only selecting catchments from China, the authors need to take into account the influence from the differences in climatic conditions on the use of the E-RD strategy. This is because the difficulty in reproducing fractality in observed streamflow from catchments selected across various climatic regions may not be comparable. Furthermore, to guard against manipulations of model inputs, the catchments should be selected in such a way that their datasets for modeling should be from sources which readers can easily access. There are a number of catchments with complete information such as, hydro-meteorological data, which can be used for rainfall-runoff modelling. To mention, but one example, is the Rainfall-Runoff Library data which can be obtained via https://toolkit.ewater.org.au/Tools/RRL (accessed: 8th December, 2020).

RESPONSE 7

Thanks for your kind comment. We agree that the selection of catchments across various climatic regions leads to a more convincing result. However, in our manuscript, the three catchments are located in very different climatic regions (see Section 3.1). Dong is a small catchment with continental plateau climate. Xiang is a large catchment dominated by Dominated by subtropical monsoon climate. Jinhua is subject to Asian monsoon climate and effected by typhoon in summer. And we are glad to use open-source data and models in our following studies.

MINOR COMMENTS

RESPONSE

We'll make the corrections as suggested.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2020-543, 2020.

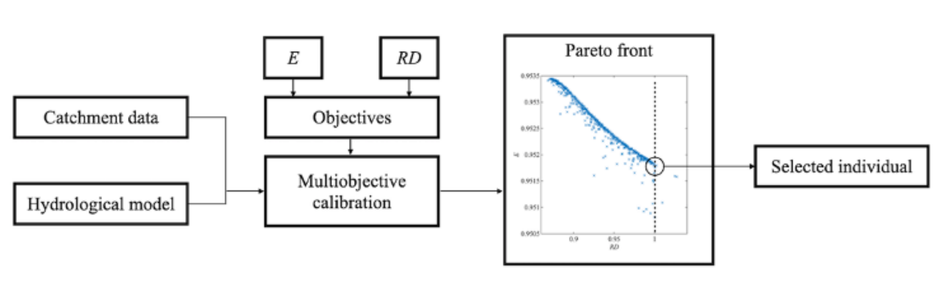

Fig. 1. Figure 2

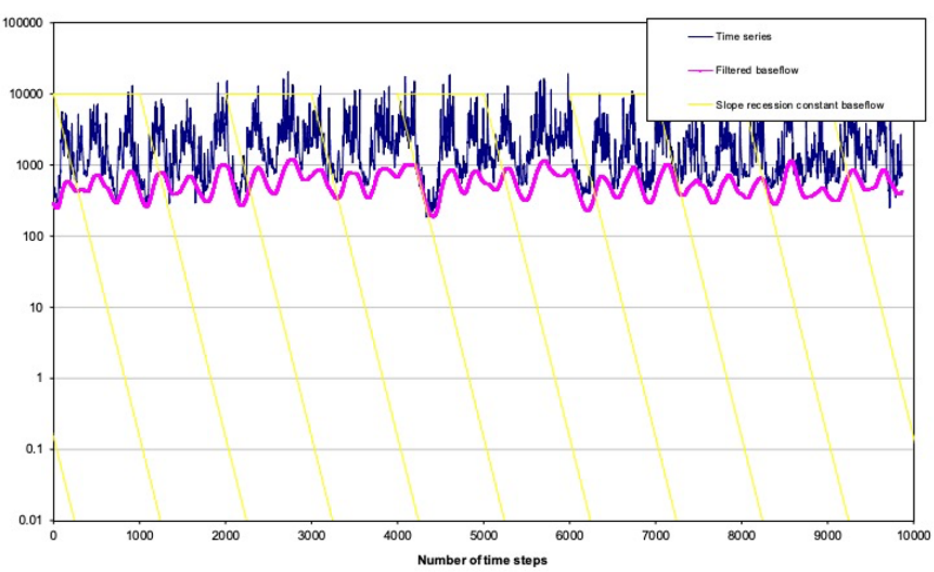

Fig. 2. Fig. R1-5

Parameter	Dong	Jinhua	Xiang
Recession constant (days)	90	80	90
w-parameter filter	0.14	0.43	0.38

Fig. 3. Table R1-5

		Observation	Best RD	Best E	Largest RD
Auto correlation	Dong	0.97	0.99	1.00	1.00
	Jinhua	0.76	0.76	0.76	0.75
	Xiang	0.94	0.95	0.94	0.94
Relative variance	Dong	0.53	0.56	0.58	0.57
	Jinhua	1.87	1.87	1.87	1.89
	Xiang	0.99	0.82	0.92	0.92
Maximum monthly flow (m^3/s)	Dong	1.54	1.40	1.42	1.39
	Jinhua	531.19	497.40	503.68	496.77
	Xiang	4210.01	3956.24	4027.68	4042.94
Minimum monthly flow (m^3/s)	Dong	0.44	0.30	0.27	0.26
	Jinhua	60.64	58.85	50.45	60.19
	Xiang	961.00	975.07	812.02	840.02
High flow percentiles (Q_5) (m^3/s)	Dong	1.93	1.44	1.49	1.38
	Jinhua	752.00	745.02	734.12	740.28
	Xiang	6048.50	5817.06	5586.92	5817.08
low flow percentiles (Q_{75}) (m^3/s)	Dong	0.50	0.39	0.40	0.38
	Jinhua	37.77	38.55	37.80	37.31
	Xiang	803.75	790.95	845.76	744.52

Fig. 4. Table 2