

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

Anonymous Referee #2

Received and published: 5 January 2021

General comments

This paper introduces a new calibration criterion based on fractal theory and combines this criterion with a traditional criterion in a multi-objective setting for hydrological model calibration. The hydrological model HBV is applied to three catchments in China with different surface areas and calibrated using the two criteria and the multi-objective calibration algorithm. The additional value of the fractality-based criterion is evaluated in a general way and for different flow components, and relations between the newly-developed criterion and parameters are analysed.

Overall, the paper is reasonable written and presents interesting insights in the use and value of the fractality-based criterion. The authors generally use informative figures to illustrate their results. Several issues need attention such as the structure of the first

C1

three sections, the different lengths of the calibration periods of the three catchments, the analysis of results only for the calibration period (no temporal validation) and the comparison of the observed and simulated flow components. These and other specific comments can be found below. The English spelling, grammar and style can be improved; several examples and other technical corrections can be found below as well.

Specific comments

1. L20-39: The introduction starts with a paragraph about fractality without describing the (hydrological) context. Fractality is used as an additional criterion for model calibration in this paper to obtain calibrated models which perform well for better reasons than when only using traditional criteria based on for instance squared residuals. Hence, a fractality-based criterion is introduced and evaluated as a tool for a more robust model calibration. Therefore, it would be more logical to start the introduction with a description of the pros and cons of existing calibration criteria followed by the introduction of fractal theory as an additional evaluation framework for hydrological models. The last two sentences of sub-section 2.3 (L157-159) typically form (part of) the research gap and are a natural link to the research objective.

2. L23: Terms like ‘self-affinity’, ‘periodicity’, ‘long-term memory’ and ‘irregularity’ are listed without any explanation. In different contexts these terms might have different meanings. What is the meaning of these terms in this study and which of these terms are quantified by/ included in the ratio of fractal dimensions introduced in this paper?

3. L82: The structure of sections 2 and 3 can be improved. Part of the discussion of the traditional criteria and their cons (and pros) and the fractal dimensions and related indices (sub-sections 2.1 and 2.2) can be included in the introduction (section 1). The description of the fractality-based criterion used in this study (sub-section 2.3) and the calibration strategy (sub-section 2.4) can be merged with section 3. As a consequence, section 2 will disappear.

C2

4. L190-206: The authors use data from three catchments with different sizes and different data periods (and lengths of time series). In particular the data period for the Dong catchment is short (4 years) compared to the other two catchments. What is the influence of these differences in data periods on the results? Does it explain the relatively poor performance of the HBV model for the Dong catchment compared to the other two catchments, particularly for fast and slow flow? Did the authors test their framework with equal data periods for the three catchments (i.e. for Jinhua and Xiang also 4-year time series)? This would be a useful test to isolate the influence of difference lengths of data periods.
5. L211-247: The description of the HBV model is somewhat messy and not complete. For instance, actual evapotranspiration is not described, the order of the fluxes is not logical and the description of the parameters is not consistent with the literature. Since this model has been very frequently used and described in the literature, the authors are advised to reduce the description to a small general paragraph and refer for more details to the literature.
6. L257-263: Although the authors mention that most settings of the calibration algorithm are default ones, it is not completely clear what the meaning of these numbers is and why mostly default settings have been used. Moreover, which 14 HBV parameters need to be calibrated? This is a large number of parameters making the calibration cumbersome. Why not firstly carrying out a sensitivity analysis to select the most dominant parameters?
7. L268: Although the authors compared different observed and simulated signatures (and separated flow components, see next comment), a validation in time and/ or space has not been carried out. It would be very interesting to see how well the HBV model performs for another time period in the three catchments. This would enable a more independent and robust evaluation of the E-RD strategy proposed by the authors.
8. L283-285: The authors compare observed and simulated fast flow and baseflow.

C3

Observed components have been obtained using the WETSPRO tool. However, it is unclear how well the division into streamflow components is done by this tool (also mentioned by the authors in sub-section 4.3). It might well be that observed and simulated components describe a (totally) different flow mechanism. For the Dong catchment, this results in a poor performance for streamflow components and for the other two catchments in a good performance. What is the principle used by WETSPRO to separate flow components and to what extent is that principle related to the concepts of the HBV model? This needs more discussion by the authors.

9. L358: Section 4.2: the selection of parameters for further analysis is not completely clear and straightforward. The authors mention a threshold for the distance correlation (is this a correlation value of a squared correlation value), but they do not consequently apply this threshold. Furthermore, the (unexpected) high correlation between RD and the degree-day factor for the (mainly) rainfed Xiang catchment needs more discussion.
10. L413-416: Would it be possible to relate the different parameter values for different catchments to differences in characteristics of these catchments (e.g. slope, soil types, size)? See also e.g. lines 451-454.

Technical corrections

1. L26 and elsewhere: “studies” instead of “literatures”.
2. L26: “indices” instead of “indexes”.
3. L38: “. . . the hydrological model shall be able to”; what is meant by this sentence? How does it relate to the previous sentence?
4. L41: “hydrological” instead of “hydrology”.
5. L50: “interest” instead of “interests”.
6. L57: What kind of loss function is referred to here?
7. L90-92: What is meant with this sentence? Do you have an example of a situation

C4

where the individual data points are well simulated but physical behaviour of the model (for a particular catchment) is not realistic?

8. L110-111: What do the authors mean with 'dependent significances theoretically'?
9. L120: What is meant with 'classical criterion controlling water budget'? A calibration criterion which compares observed and simulated water balances?
10. L130: What is the meaning of the 'delta' symbol?
11. L132: This part of the sentence is not clear and is an example which needs reformulation.
12. L209: It would be better to use the same colours for the three DEMs.
13. L265: What is the E-RD calibration strategy? Probably here the explanation from sub-section 2.4 can be used.
14. L266: "corresponding" instead of "correspondent".
15. L276: Do you have a reference for 'distance correlation'?
16. L295: "1.0" instead of " $1+2.8 \times 10^{-12}$ ".
17. L298: What is the meaning of 'significant' here?
18. L298-299: And when RD is smaller than 1? In this study RD often is lower than 1 and sometime only slightly higher than 1. Why is this the case?
19. L314: Xiang catchment has a different (wrong?) x-axis.
20. L318: Are the best RD, best E and largest RD selected from the Pareto front?
21. L329-330: This sentence is not clear.
22. L359: "front" instead of "frontier".
23. L377-378: The ranges of the parameters in Table 3 need units. In addition, the

C5

ranges for parameter BETA are very small and all below 1. How do these ranges compare with recommended ranges from the literature?

24. L386: The first part of this line is not correct; the correlation for the Dong catchment is smaller than 0.74.
25. L391-392: Please include a unit for KF in Fig. 8 and switch the axes; i.e. RD is the dependent variable and the parameters are the independent variables. This also applies to other figures with RD as a function of parameter values.
26. L409-410: What is the meaning of the letters A, B, D, E, G and H?

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

C6