Response Letter

Report #1

I thank the authors for their thoughtful and thorough response to my criticism. The revised version of the manuscript satisfactorily addresses my principal comments that resulted in improvements to the manuscript in comparison with the previous version.

Reply: Thank you very much for your valuable suggestions and encouragement.

I suggest softening some statements:

L. 89-90 "…the concept has not been used in hydrology under natural conditions". => "…the concept has rarely been used in hydrology under natural conditions"

Reply: Thank you for the suggestion. The sentence has been revised accordingly in Line 91.

L 121 "…no studies investigated the dynamics of the catchment water storage" => … there are not many studies investigated the dynamics of the catchment water storage"

Reply: Thank you for the suggestion. The sentence has been revised accordingly in Line 122-123.

Some papers as an example

Birkel, C.and Tetzlaff, D. (2011). Modelling catchment-scale water storage dynamics: Reconciling dynamic storage with tracer-inferred passive storage. Hydrological Processes. 25. 3924 - 3936. 10.1002/hyp.8201.

Chaffaut, Q. et al (2022) New insights on water storage dynamics in a mountainous catchment from superconducting gravimetry, Geophysical Journal International, Volume 228, Issue 1, January 2022, Pages 432–446, https://doi.org/10.1093/gji/ggab328

Huang, CC. and Yeh, HF. (2022) Evaluation of seasonal catchment dynamic storage components using an analytical streamflow duration curve model. Sustain Environ Res 32, 49. https://doi.org/10.1186/s42834-022-00161-8

Oswald, C et al. (2011). Water storage dynamics and runoff response of a boreal Shield headwater catchment. Hydrological Processes. 25. 3042 - 3060. 10.1002/hyp.8036.

L 501-504

"In conclusion, the system dynamics approach provides a hierarchical view to understand endogenous linkage structure of a hydrological system, and better reproduces the slow hydrological processes at interannual to decadal scales compared to conventional hydrological models" =>

"In conclusion, the system dynamics approach provides a hierarchical view to understand endogenous linkage structure of a hydrological system, and has the potential to better reproduce the slow hydrological processes at interannual to decadal scales compared to conventional hydrological models

I recommend the manuscript for publication after these technical revisions.

Reply: Thank you for the suggestion. The sentence has been revised accordingly in Line 532-535. Your recommendation is greatly appreciated.

Report #2

The authors have gone to some effort to revise their manuscript. However, they have opted merely to discuss the issues raised rather than address them. Given I earlier wrote that the issues raised were serious, I find it disappointing that a more substantive response has not been attempted. I do not feel favourable towards publication unless significant extra analysis is undertaken, as discussed below.

Reply: First, we would like to express our appreciation for your valuable and constructive comments and suggestion. We apologize for any oversight over concerns and comments you have shared with us. We will try our best to address your concerns in this revision.

CORE ISSUE

The main issue I have with the paper is that the authors imply via both words and tone (particularly in the abstract and short summary) that they have improved upon existing methods, yet they have not undertaken a comparison with existing methods - in fact, they have not even demonstrated how to quantify the performance they are interested in. When I pointed out the poor performance on an annual timestep, the defence was that the method is focussed on the multi-annual and multi-decadal timescales. Fine - but the onus is on the authors to find a way to directly quantify this, which has not been attempted. One option could be to separate out the shorter term fluctuations by focus on the residuals around the annual rainfall runoff relationship, which would remove much of the short term fluctuation. Another option is to separate off the low frequency component using, say, empirical mode decomposition or similar. This could be applied in similar fashion to both simulated and observed data, thus isolating the interesting component of the signal and allowing focus on whether it is correctly simulated. The specific technique doesn't matter, the point is that such things are possible, and no attempt has been made. Whatever technique is chosen, it then needs to be applied consistently to allow comparison between their method and a traditional rainfall-runoff model or models.

So, in summary, I would consider the following two things to be a minimum requirement for publication:

1. Quantification of performance on multi-annual to multi-decadal timescales since this is the timescale the authors claim to be focussing on; and
2. Application of (1) to allow a comparison of performance between a traditional rainfall-runoff model and the proposed method.

Reply: Thank you for the comments. We did comparison between our model and a rainfall-runoff model using the empirical mode decomposition method at both short-term and long-term scales as suggested. As expected, the system dynamics approach outperforms current approaches at long-term scales. Please see the text in line 450-458 in discussion 4.2 and Figure S13 for the revision.

"By adopting the combined structure of a vegetation reinforcing feedback and a soil water-vegetation balancing feedback, our model significantly outperforms traditional rainfall-runoff models and large models at the long-term scale, while demonstrating marginally inferior performance compared to them at the short-term scale. On one hand, a comparison between our model and SIMHYD, a widely used rainfall-runoff model in Australia (Chiew et al., 2002), was conducted in Fuping catchment as an example. Monthly rainfall and potential evapotranspiration were used as inputs to SIMHYD model, and monthly results were aggregated to annual values. Then empirical model decomposition method was applied using Matlab to quantify the performance of both models on short-term and long-term scales (Fig. S13). Results showed that SIMHYD model only surpassed our model in short-term Q simulation, but performed poorly in short-term ΔS simulation and long-term AET, Q, and ΔS simulation."

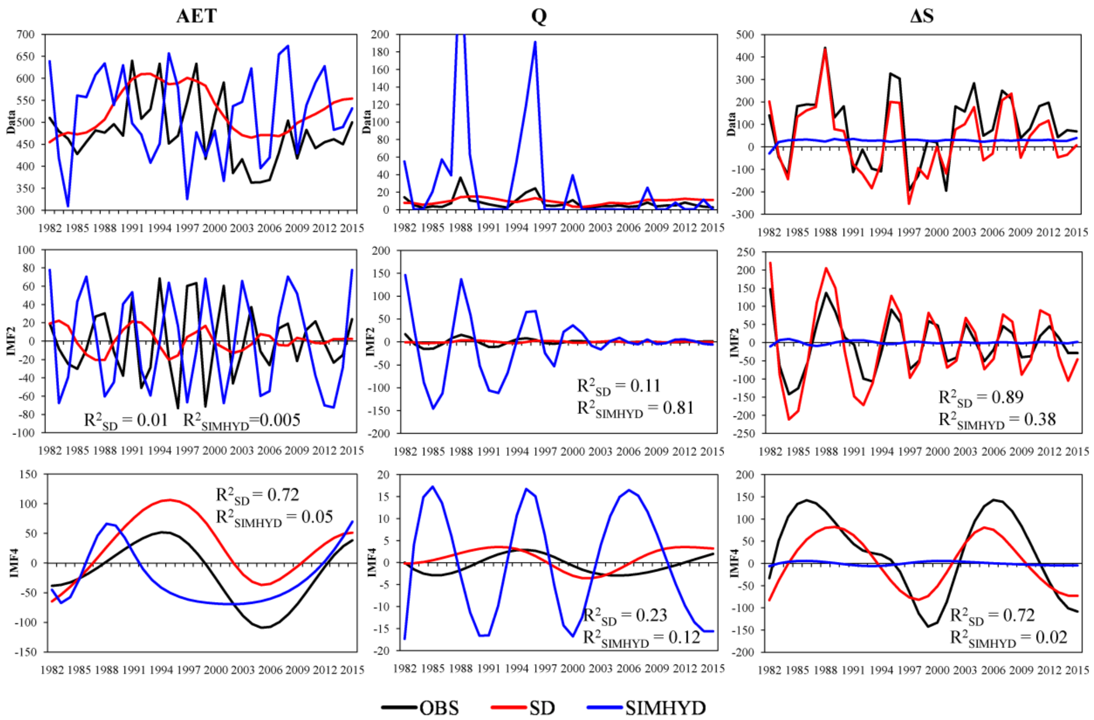

Figure S13 Comparison between the proposed system dynamics model and traditional rainfall-runoff model SIMHYD on simulating short- and long-term hydrological dynamics in Fuping catchment in China as example. The comparison is conducted by the empirical model decomposition method. OBS is observations, SD is system dynamics model. Here the results of SD model were obtained with varied ESMS (expected soil moisture stock, see eq. 6 in the main text), without VEG and GP which mimic human activity parameters for, respectively, vegetation-related activities such as reforestation, and groundwater pumpage. IMF is the intrinsic mode functions which can be used to extract different resolutions from the data without the use of fixed functions or filters.

OTHER COMMENTS

In the case of a decision to proceed with publication, the following comments might help to improve the manuscript:

1. AET as input. In response to my suggestion that AET needed to be included as an input, the authors argued that because of the arid nature of the case study, it is not needed. I sense that arguing about this further is not helpful, so I merely request that the manuscript:
- acknowledge that the vast majority of cases will need AET as an input. This needs to be done in the methods and the limitations section of the manuscript, at least.
- confirm that this is possible to include as an input - ie. that the method would allow this.
- write a paragraph aimed at potential users of the method, commenting on how easy or difficult it might be to incorporate AET as an input. Any instructions on how to do this would be welcome, but I recognise that it might be difficult to do this breifly.
Reply: Thank you for the comment. We have integrated vegetation parameter as input in our model, implicitly equating it to the input of AET due to the inherently close correlation between these two factors. A discussion has been added in line 475-482 in discussion 4.2 for the revision.

"On the other hand, the parameter values should be allowed to vary over time rather than being fixed to improve model's performance at the long-term scale. For instance, the fixed utilization ratio of soil water (C2 in Eq. 8) implies that AET solely depends on vegetation coverage and soil bound water, but not on plant species. However, the dominant plant species in the study area have shifted from herbs to shrubs due to the implementation of afforestation projects (Liu et al., 2011), resulting in significant changes in the utilization ratio of soil water (Liu et al., 2014). By incorporating variable parameters, the model's performance can be enhanced, as the input of high-quality vegetation data, including vegetation coverage and plant species over time, is actually equivalent to the input of AET based on the intrinsically close links between the two factors."

2. Given the authors' defence of the poor performance is that their method is focussed on longer timescales, the authors need to discuss options for how to augment their method such that it can also match shorter timescales, otherwise the concept of using it for projection (ie. Figure 7) adds little value. I feel this would fit well with their existing discussion in Section 4.2.
Reply: Thank you for the comment. A discussion has been added in line 465-475 in discussion 4.2 for the revision.

"In the future, our model can be improved from at least two aspects. On one hand, the mechanism of low hierarchy in the hydrological system can be integrated into the endogenous linking structure to enhance the model's performance at the short-term scale. As discussed earlier, the interaction between rainfall density and sink-filling/macropores controls the hydrological behaviors at intra-annual scale. Thus, high-temporal-resolution rainfall data and detailed soil properties of the topsoil layer are needed to determine the initiation and duration of short-term runoff events. In addition, to link the two hierarchies, intra- and inter-annual together, soil macropores/sinks should be modelled as dynamic features that evolve in response to vegetation changes. Previous studies showed that the change of plant productivity affects the input of plant products (above- and belowground litter), causing changes in fractions of particulate organic carbon (Shi et al., 2024), subsequently affecting soil water-holding capacity and storage-discharge relationship. Furthermore, as topsoil texture becomes finer or coarser, water infiltrating into the deep soil changes, which in turn affects vegetation physiology and structure across scales (Wankmüller et al., 2024)."

3. The manuscript is not consistent in its portrayal of the method. Late in the manuscript (Section 4.2) the method is described as "a 'toy' model, also known as a minimal or exploratory model". The implication is that it is a first step and is intended to explore the possibilities of alternative methods, which seems appropriate. It also implies the approach has strong limitations. This contrasts sharply with the abstract and short summary, which seem to be saying that this is a method that can replace conventional hydrological models because it "excels at explaining pattern of slow hydrological behaviours" and the case study application "successfully captured slow hydrological behaviours". I strongly suggest that these matters be harmonised, preferably by changing the wording of the abstract and short summary and adding phrases such as "toy model" and "exploratory analysis". The abstract should also contain a clear statement of limitations.

Reply: Thank you for the comment. More explanations have been added in Abstract (line 37-38) and Short summary (line 46-47) for the revision.

"The system dynamics model is in its early stage with applications primarily confined to water-stressed regions and long-term scales."

"In spite of the simplicity, it holds potential to integrate hydrological behaviors across scales."

4. Following on from the above, the use of the method to provide projections (Fig. 7) is questionable. It is never explained why projections are provided in the first place. What use are projections produced by a "toy" model that is acknowledged to be exploratory? I suggest these matters be clarified.

Reply: Thank you for the comment. The projection serves as a form of model comparison aimed at validating our model's ability for long-term simulations, as initially we didn't conduct a comparative analysis between our model and a rainfall-runoff model. We believe that this comparison is not only more rigorous but also offers additional information about future scenarios. A sentence has been added in line 395-397 for the revision.

"These findings further emphasized the system dynamics' structure is capable of producing long-term hydrological behaviors. Conversely, while adept at capturing short-term fluctuations, process-based models fall short in simulating long-term hydrological behaviors due to lack of the structure."

5. The wording of the title is not correct English. Two acceptable alternatives are:

- Can system dynamics explain long-term hydrological behaviours? The role of endogenous linking structure
- System dynamics explains long-term hydrological behaviours: the role of endogenous linking structure.

Given the issues raised above, it is more appropriate to phrase as a question - thus, I recommend the first one.
Reply: Thank you for the comment. The title has been changed accordingly.

6. Line 449-453. A correlation of X on a daily or monthly timestep has a different meaning to a correlation of X on an annual timestep. Thus, when comparing performance among methods, please state the timesteps on which the performance is calculated.
Reply: Thank you for the comment. The timesteps have been added in line 459-462 for the revision.

"Results showed that all models performed well in mean annual Q simulation but struggled with the simulation of inter-annual Q change, with median correlation coefficients (*r*) close to 0 for GCMs, and around 0.6 for GHMs and LSMs (r-square values ranging from 0.5-0.6) (Zhou et al., 2012; Hou et al., 2023)."

7. Line 467. The comments about an "attractor" seem to assume there is only one attractor. Some studies suggest there can be more than one - particularly Peterson et al. https://doi.org/10.1126/science.abd5085. Given the Peterson study was, like this study, framed around multi-annual behaviour in a relatively dry region, I suggest that the implications for the present manuscript be briefly discussed.
Reply: Thank you for the comment. A sentence has been added in line 497-498 for the revision.

"It is noteworthy that the ESMS is changeable in our study, as the factors that determine ESMS have evolved in tandem with the climate change, corresponding to multiple attractors."