Responses to Anonymous Referee #3

Dear Authors,

Thank you very much for your work. I think the work is interesting but I have some concerns.

We thank the reviewer for his/her suggestions as they will contribute to improve the manuscript. We have indicated in our responses those references that were not included in the initial version of the manuscript.

Introduction

- Introduction miss literature review: a) what about the intercomparison project MOPEX; b) what about prediction in ungauged catchments. Please, see key references, also for work that has been performed in Spain.

- Line 31: "across climates" I suggest you to have a look at Addor et al. (2018).
- Line: 40 "evaluation and benchmarking". I suggest you to have a look at Prieto et al. (2021, 2022).
- Line 43: parameter regionalization techniques, I recommend you to have a look at Almeida et al. (2016) and Prieto et al. (2019).

Addor, N., Nearing, G., Prieto, C., Newman, A. J., Le Vine, N., & Clark, M. P. (2018). A ranking of hydrological signatures based on their predictability in space. Water Resources Research, 54, 8792–8812. https://doi.org/10.1029/2018WR022606

Almeida, S., Le Vine, N., McIntyre, N., Wagener, T., and Buytaert, W. (2016). Accounting for dependencies in regionalized signatures for predictions in ungauged catchments, Hydrol. Earth Syst. Sci., 20, 887–901, https://doi.org/10.5194/hess-20-887-2016

Prieto, C., Le Vine, N., Kavetski, D., Fenicia, F., Scheidegger, A., & Vitolo, C. (2022). An exploration of Bayesian identification of dominant hydrological mechanisms in ungauged catchments. Water Resources Research, 58, e2021WR030705. https://doi.org/10.1029/2021WR030705

Prieto, C., Kavetski, D., Le Vine, N., Álvarez, C., & Medina, R. (2021). Identification of dominant hydrological mechanisms using Bayesian inference, multiple statistical hypothesis testing, and flexible models. Water Resources Research, 57, e2020WR028338. https://doi.org/10.1029/2020WR028338

Prieto, C., Le Vine, N., Kavetski, D., García, E., & Medina, R. (2019). Flow prediction in ungauged catchments using probabilistic random forests regionalization and new statistical adequacy tests. Water Resources Research, 55, 4364–4392. https://doi.org/10.1029/2018WR023254

We thank the reviewer for suggesting key references to improve the introduction. All of the references provided above will be incorporated into the introduction of the revised version of the manuscript. In addition, the MOPEX intercomparison project as well as Prediction in Ungauged Basins will be discussed from a large-sample hydrology perspective to improve the introduction.

- Line 64: "there is an increasing tendency towards aridity conditions": what is the difference for different catchments in Spain.

The increasing tendency towards aridity conditions manifests a similar pattern for the Spanish catchments and generally reveals a clear latitudinal gradient with greater aridity corresponding to the southern catchments. This will be specified in the revised version of the manuscript.

Study area and data

- Line 80: specify the northern districts (there are "many", eg, Aguas de Galicia, CHC, URA, ARA)

We will specify the Northern Districts in the text of the revised version of the manuscript. We would like to clarify that these districts were grouped under the term “Northern Districts” because the identifiers in the SAIH-ROEA dataset corresponding to all the northern catchments share the first digit.

- Also, in section 2 I recommend to provide the range of mean annual precipitation, mean annual flow, mean annual potential evapotranspiration and rainfall runoff coefficient across catchments and maybe per river basin district in the text. This is to guide the reader.

We thank the reviewer for pointing this out. We will provided the four ranges across catchments in the text to guide the reader in the revised version of the manuscript. Providing them per River Basin District can potentially confuse the reader as there are eight River Basin Districts and this would lead to a total of 32 values. The reader can always refer to Fig. 1 and Fig. 3 to visualize the hydrologic variability in space for the study catchments.

- Line 205: you are using SIMPA as benchmark which we know is a very simple model. Maybe, include the pros and cons or similarities and differences as most of the readers won't know what SIMPA is.

The importance of SIMPA for water resources management in Spain must be recognized as it constitutes a reference tool for water resources allocation at the national and at the basin scale. SIMPA is a semidistributed implementation of the lumped conceptual model of Témez (1977) and has evolved since its inception to include, among other features, a snow module (https://www.miteco.gob.es/content/dam/miteco/es/agua/temas/evaluacion-de-los-recursos-hidricos/cedex-informeerh2019_tcm30-518171.pdf), and recently a new hydrogeological module (http://hdl.handle.net/10261/335461). In our understanding, the simplicity of the original model of Témez has been already left behind, and SIMPA has been used in many previous studies for comparison purposes (e.g., Pellicer-Martínez and Martínez-Paz, 2018; Suárez-Almiñana et al., 2020; Yeste et al., 2020). We agree that describing the similarities and differences between VIC and SIMPA is going to be helpful

for the reader and beneficial for the manuscript. This description will be included in Section 3.3 in the revised version of the manuscript.

Témez, J.R., 1977. Modelo matemático de transformación “precipitación-aportación”. Asociación de Investigación Industrial Eléctrica (ASINEL). Madrid.

Discussion

- I miss 1) talking about uncertainty, 2) talking about model structure error, and 3) talking or comparing (which would go to the methods) with a more well established model, e.g. GR4F, even if SIMPA is used as benchmark. So that you would have SIMPA, VIC and GR4J.

We thank the reviewer for his/her suggestions to improve the discussion of the manuscript. As per the requirement of the Anonymous Referee #1, the value of the evaporation dataset in reducing equifinality and uncertainty will be addressed according to the results of the Monte Carlo experiment. The equifinality assessment has resulted in a new figure that will be introduced in Section 4.1 and that will be discussed in Section 5.1 in the revised version of the manuscript. Please refer to our response to the major suggestion from Anonymous Referee #1 for a thorough explanation of our new findings.

In relation to model structure errors, it is not possible to analyze them when only one model structure is used as in this work. The focus of this study is the VIC model and its application for the Spanish catchments after the previous experience using VIC in Yeste et al. (2020, 2021, 2023), given the suitability of VIC for large-sample and large-scale applications (e.g., Sepúlveda et al., 2022). The evaluation of model structure errors is better performed with frameworks such as FUSE (Clark et al., 2008) or SUMMA (Clark et al., 2015). The inherent limitations of using only one model structure will be recognized in the revised version of the manuscript and the use of FUSE and SUMMA will be identified as a potential future work.

Finally, as for the comparison of VIC with a more established model such as GR4J, to our knowledge there are neither lumped GR4J simulations available for the 189 study catchments nor gridded simulations available for the Spanish domain, whereas gridded SIMPA simulations are publicly available and regularly updated. Implementing GR4J ourselves would notably require an additional calibration effort and an intensive data processing step for all the 189 study catchments to subsequently compare it against VIC. Such an endeavor is unfortunately unfeasible at this stage of project implementation and thus it is out of the scope of this work. Please see also our response to the following question.

Clark, M. P., Slater, A. G., Rupp, D. E., Woods, R. A., Vrugt, J. A., Gupta, H. V., Wagener, T., & Hay, L. E. (2008). Framework for Understanding Structural Errors (FUSE): A modular framework to diagnose differences between hydrological models. Water Resources Research, 44(12), 1–14. https://doi.org/10.1029/2007wr006735

Clark, M. P., Nijssen, B., Lundquist, J. D., Kavetski, D., Rupp, D. E., Woods, R. A., Freer, J. E., Gutmann, E. D., Wood, A. W., Brekke, L. D., Arnold, J. R., Gochis, D. J., & Rasmussen, R. M. (2015). A unified approach for process-based hydrologic modeling: 1. Modeling

concept. Water Resources Research, 51(4), 2498–2514. https://doi.org/10.1002/2015WR017198

- I also miss to compare with other models and results that were run at daily time scale in Spain, e.g. look at URA

Several studies were already discussed in Section 5.2, in particular those involving the Duero River Basin (Morán-Tejeda et al., 2014; Yeste et al., 2020, 2023), Tajo (Pellicer-Martínez and Martínez-Paz, 2018; Pellicer-Martínez et al., 2021), Guadalquivir (Yeste et al., 2018), Segura (Pellicer-Martinez and Martínez-Paz, 2015; Pellicer-Martínez et al., 2015) and Júcar (Marcos-Garcia et al., 2017; Suárez-Almiñana et al., 2020). We are happy to also include the key references suggested by the reviewer at the beginning of the review report in relation to models that were run at daily time scale for the comparison with other models and results that were run at daily time scale in northern Spain.

Conclusions

- "The soil and routing parameters were reveled as the most important parameters". Could you add in which type of catchments were most and least important?

A similar question was posed by Anonymous Referee #1 and Ilhan Özgen-Xian. Both reviewers suggested to interpret the correlations in Fig. 6 and the importance of the different parameters from a hydrologic perspective. Please refer to our response to Anonymous Referee #1 for further details, as this topic has been extensively discussed there and the discussion on the mechanisms behind the correlations will be integrated in Section 5.1 in the revised version of the manuscript. As required by the reviewer in this question, we will also indicate for which catchments the routing parameters are more important in the conclusions of the revised version of the manuscript.

Once again, thank you very much for your work.

In case the editor asks for a revised version of the manuscript, I am very happy to serve as reviewer of the revised version

All the Best,

Reviewer