

Dear Reviewer 1

Thank you for your valuable feedback on our manuscript. We truly appreciate the time and effort that you have dedicated to the review process, and your interest in this work. Please see below, in **blue**, the original review comments, and in **black**, our responses to the comments and how we will implement the changes in the revised version.

Araki et al. (2026) presented a very interesting paper, which incorporates new methods, new interpretations and further develops the LSH field. The authors move from a more standard just prediction approach to one more focused on processes, and this is a very interesting path to go. Although I truly believe that the contribution should be accepted for publication at the journal eventually, it still needs some minor revisions before it is in full shape.

### **Major comments:**

1. Did the authors filter out basins from the sample based on time-series quality (apart from the time series length?). I think that this would be a very important step to make clear for readers and to base the interpretations on.

Yes, we filtered out basins based on time-series quality (L170–177 in the first submission), e.g., those with long periods of missing data, uncertain topographic boundaries, and, for some signatures, snow-dominated basins. We have applied this quality controlling criteria to both datasets used for signature calculations, as well as datasets used for random forest training. The flow chart below will be added to Text S1 to clarify the point. Furthermore, we will move the L170-177 sentences to “Section 3.1 Calculating hydrologic signatures” so that the filtering criteria are clearer.

Additionally, CAMELS-US, HYSETS, and USGS GAGES-II datasets also underwent filtering processes, such as removing watershed with limited streamflow data continuity (CAMELS, HYSETS, GAGES-II), dubious topographic boundaries (HYSETS, GAGES-II), peakflow impacted by regulation (HYSETS), on management channels (e.g., canals, flumes) (GAGES-II).

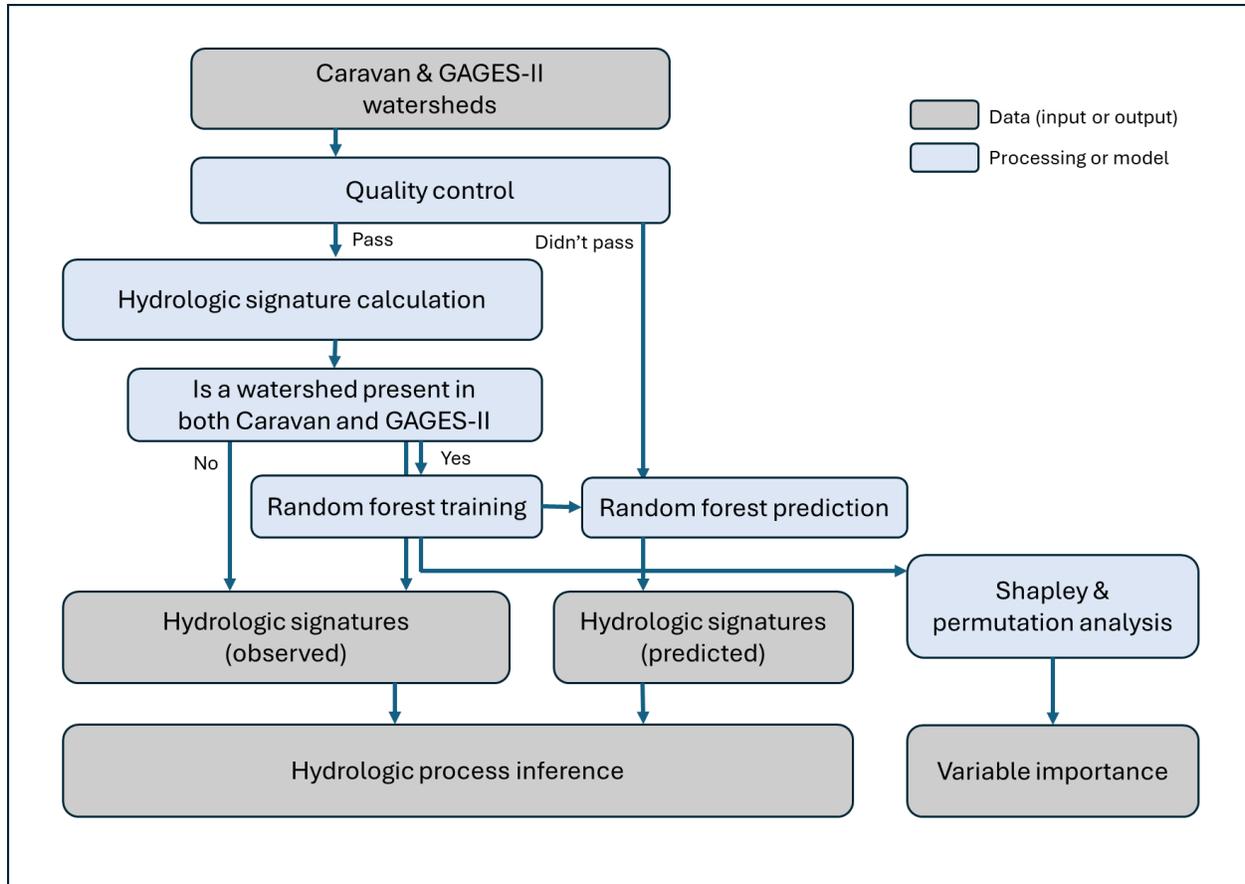
**GAGES-II:** James A. Falcone, Daren M. Carlisle, David M. Wolock, and Michael R. Meador. 2010. GAGES: A stream gage database for evaluating natural and altered flow conditions in the conterminous United States. *Ecology* 91:621.

<https://esapubs.org/Archive/ecol/E091/045/metadata.htm>

**HYSETS:** Arsenault, R., Brissette, F., Martel, J.-L., Troin, M., Lévesque, G., Davidson-Chaput, J., et al. (2020). A comprehensive, multisource database for hydrometeorological modeling of 14,425 North American watersheds. *Scientific Data*, 7(1), 243.

<https://doi.org/10.1038/s41597-020-00583-2>

**CAMELS-US:** Addor, N., Newman, A. J., Mizukami, N., & Clark, M. P. (2017). The CAMELS data set: catchment attributes and meteorology for large-sample studies. *Hydrology and Earth System Sciences*, 21(10), 5293–5313. <https://doi.org/10.5194/hess-21-5293-2017>



(Moved here from minor comments for relevancy:) L181: Could you please clarify what are these quality standards that you mean here?

This means the filtering criteria. Only signatures calculated from quality-controlled streamflow datasets are used for random forest training (with an additional criteria about subsetting). We will rephrase the sentence as: *“Only the signatures calculated from quality-controlled streamflow records (as described in Section 3.1) were used for training.”*

2. Also, to my understanding, the authors did compile the signatures as long-term averages for all time series. Although I fully acknowledge that this is valid and I see no problem at all in the methodological choice, I think it would be interesting to highlight and briefly discuss this in the paper. I know by experience that many catchments are under change, and perhaps a long term average could mask some of those? Again, no need to change anything, just discussion in my opinion.

Thank you for pointing that out, and we agree with your comment. We are currently working on the long-term trend of hydrologic signatures. We will add a discussion of this point in “Section 5.3 Limitations and future work” as:

*“Hydrological signatures in this study are long-term averages of the multi-year streamflow dynamics, which may not fully capture temporal variability in watershed processes. Future studies should account for long-term hydroclimatic changes (Hobeichi et al., 2022);*

*Gudmundsson et al., 2025*), as well as inter-annual variability (*Vogel et al., 1994*) and seasonal variations in watershed function (*Payn et al., 2012; Gomi et al., 2008*). ”

3. I am not fully used to the use of sections in the introduction. Although I think they are not a problem, could the authors check about the requirements from HESS? I would choose not to use, but again, this is a taste matter (in case HESS has nothing against it).

HESS guideline does not explicitly prohibit having subsections <https://www.hydrology-and-earth-system-sciences.net/submission.html>, however, we agree that we haven't seen many HESS papers with subsections in the introduction. We will leave it up to the proofing staff.

4. After reading I got a bit confuse. Did you classify the dominant based only on the 4 processes? Could you elaborate a bit better on that?

We classified whether each of the 6 process types is dominant or not, based on a combination of 2 signatures each (shown in the Figure 3 maps). Then, in Figure 2, we created a simplified overview map of dominant processes by overlaying the maps of 4 of these process types (see reasoning under next comment). We simplified by combining the two overland flow types, and omitting the seasonal variability. For example, “baseflow” in Figure 2 corresponds to the same map as Figure 3a (Baseflow), but only shows watersheds with the highest signature quantiles, indicating that the process is dominant. Dominant processes are not mutually exclusive, so some watersheds may show multiple dominant processes.

(Moved here from minor comments for relevancy:) Figure 2: why is the "seasonal variability" not present there?

We only show 4 processes in Figure 2, because overlaying 6 layers makes it congested and hard to read. If the readers are interested in remaining process dominance, they may refer to Figure 3.

The 4 selected processes are particularly important for hydrologic understanding. Baseflow sustains discharge across seasons and reflects groundwater connectivity, while overland flow drives stormflow and influences flood peaks. Storage governs buffering and recession dynamics, and water balance losses through evapotranspiration and deep percolation determine how much precipitation is converted to streamflow. Together, these processes span the continuum from slow to fast hydrologic response and integrate both vertical and lateral fluxes, making them essential for hydrologic theory and modeling. Notably, conceptual hydrologic models that focus on key functions, such as HBV (*Bergström, 1992*) and *Kirchner et al. (2009)*, are built around these 4 primary processes. We will include this passage in a Method subsection, Section 3.3 Interpretation of hydrologic signatures as process descriptors.

Of the remaining two processes, overland flow type is particularly relevant for hydrologic management, as noted in the introduction: “*Optimal management strategies, including the design of grey and green infrastructure, differ depending on which processes dominate hydrological response (Oswald et al., 2023; Thompson et al., 2020)*,” (note: these references concerns specifically the overland flow type). Additionally, understanding seasonal variability in dominant processes is also important (*Berghuijs, 2014*).

Kirchner, J. W. (2009). Catchments as simple dynamical systems: Catchment characterization, rainfall-runoff modeling, and doing hydrology backward. *Water Resources Research*, 45(2). <https://doi.org/10.1029/2008WR006912>

Bergström, S.: The HBV Model: Its Structure and Applications, Swedish Meteorological and Hydrological Institute (SMHI), Hydrology, Norrköping, 35 pp., 1992.

Berghuijs, W. R., Sivapalan, M., Woods, R. A., & Savenije, H. H. G. (2014). Patterns of similarity of seasonal water balances: A window into streamflow variability over a range of time scales. *Water Resources Research*, 50(7), 5638–5661. <https://doi.org/10.1002/2014WR015692>

(Moved here from minor comments for relevancy:) L192: Does it mean that we can get more than one dominant process per signature?

It is possible that the dynamics captured by one signature are confounded by other processes. We will add a discussion on this topic to “Section 5.3 Limitations and future work” as:

*“Another complication is that hydrologic signatures are often confounded by multiple processes (McMillan et al., 2022, 2023), whether driven by natural flow dynamics or impaired by human activities. For example, water abstraction by reservoirs reduces downstream flow variability and increases water balance deficits (Salwey et al., 2022; Veldkamp et al., 2016), but changes in vegetation or climate could induce similar effects. Disentangling these impacts remains challenging without testing narrower hypotheses about watershed function, incorporating expert knowledge, or having detailed information about human interventions. In this study, we partially mitigated this issue by using multiple signatures to characterize processes, and by representing human alteration through population density, which showed strong explanatory power for the signatures. Nevertheless, considerable effort is still needed to isolate the combined impacts of multiple processes, as well as the effects of urban development and agricultural practices on flow dynamics (Grantham et al., 2022) for improving the large-scale application of signatures.”*

(Moved here from minor comments for relevancy:) Figure 2: what does "unclassified" mean? I thought it was when no process dominates, but I see gray lines with the white lines. I am a bit confused here. Could you clarify it in more detail in the text?

“Unclassified” means the watershed is included in our dataset, but the estimated watershed processes were neither baseflow- nor overland-flow-dominant (i.e. neither the baseflow nor overland flow process had both signatures in the upper quantile). We colored these grey to distinguish from white areas not included in our dataset. We will add this to the Figure 2 caption.

### **Minor comments:**

Section 2.1: could you clarify better what are the number of gauged versus ungauged basins in the study?

Thank you for your comment. We will add this information to the beginning of “Section 3 Method”, as it involves explanation beyond the datasets. As stated in the conclusion currently: *“We analyzed 14,146 gauged U.S. watersheds; our map of processes was based on observational data from 10,261 gauged sites and extended using random forest predictions to an additional 3,885 watersheds.”*

Section 2.2: what was the motivation of merging Caravan and GAGES, and not using only GAGES (or vice versa) for example? Perhaps I missed it!

The purpose of merging Caravan and GAGES was to increase the sample size as much as possible. Approximately half of the watersheds in the Caravan and GAGES datasets overlap, while the other half do not. As stated in the introduction, in previous studies “[...], smaller regional sample sizes may limit prediction accuracy if datasets only provide tens of watersheds per region (Willard et al., 2024).” therefore, “[...], our work assessed 14,146 U.S. watersheds and was trained on 10,261 watersheds, nearly ten times more sample watersheds than previous studies; we leverage the Caravan and GAGES-II—the most extensive open-source large-sample datasets currently available (Falcone, 2011; Kratzert et al., 2023).”.

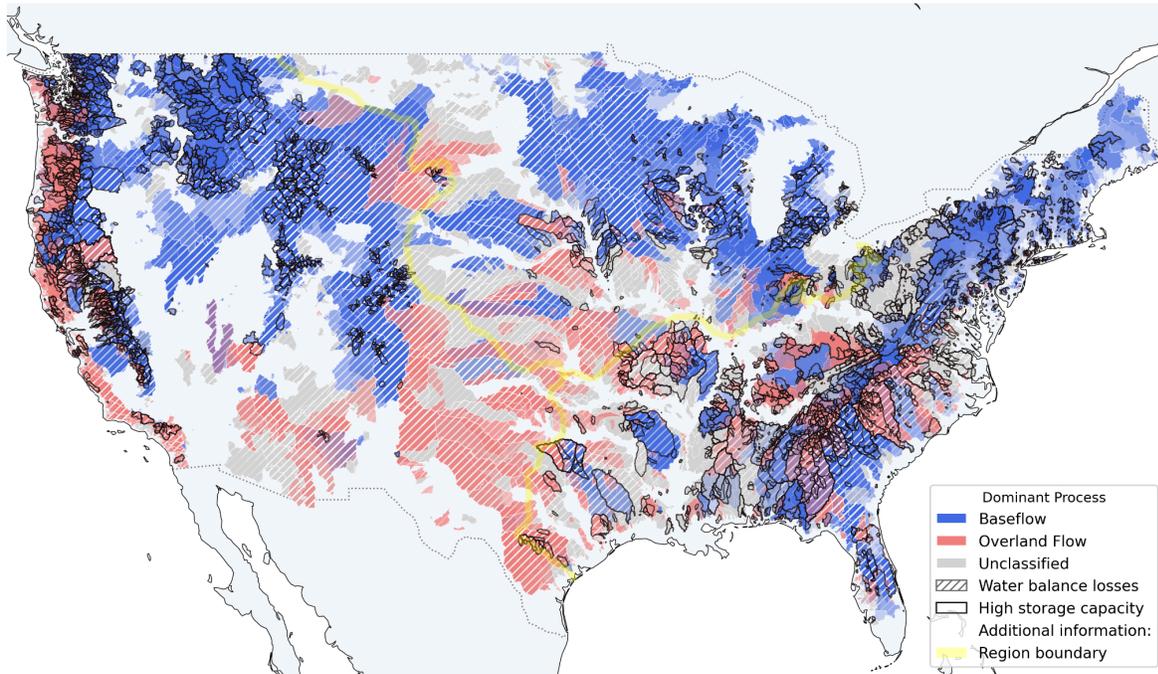
We will add a following statement to Section 2.2: *“The purpose of merging Caravan and GAGES-II dataset is to maximize the sample size of watersheds and better capture regional hydrologic variability (see Table S1 and an associated Venn diagram of watershed coverage across datasets). ”*

L188: I found it confuse here. Did you not use 4 signatures per process? Could you elaborate better on that?

The statement in the paper is correct that we used 2 signatures per process [not 4]. These are shown on the bivariate plots.

Figure 2: Would it make sense to add the three regions here? I am reading the text, and I found it a bit of a lot of going back and forward to try to match the regions and boundaries... In case it does not damage this (by the way very pretty) figure, could you add?

Thank you for your suggestion, we will try adding the three regions to the figure 5 as follows.



4.3 L351: Sorry for being a bit peaky, but could you perhaps jsut add that your hypothesis is that the RF and shapely can be interpreted as controls? Just that readers are aware that this is an hypothesis. The correlations, predictions, performances could always be just mathematical (although I believe that do reflect process, as you do!).

Thank you for pointing that out, we agree that it is really important to distinguish between statistical correlation and real causations. We will rephrase the sentence as: *In this section, we interpret the random forest models to estimate which aspects of climate and landscape are most important in controlling hydrologic processes in different regions of the U.S. We hypothesize that variable importance statistics from Shapley and permutation analysis reflect the relative importance of hydrologic process drivers.*