

Reviewer # 2

Title: A novel technique to improve the hydrological estimates at ungauged basins by swapping workspaces

Authors: M. U. Qamar, M. Azmat, M. Usman, D. Ganora, M. A. Shahid, F. Baig, S. Mushtaq

Note: The reviewer's comments are written in "blue" followed by our response in "black".

OVERALL EVALUATION

The manuscript focuses on regional streamflow regimes predictions in ungauged sites by using a dissimilarity-based method. Looking in a comprehensive way at the whole study, I regretfully have to inform the Authors that, in my opinion, the manuscript is unsuitable for publication in Hydrology and Earth System Science. My main concerns about the manuscript are listed below. I hope the Authors will find them useful should they decide in the near future to critically revise their study.

General comment: Given the topic and the analyses reported, I think it is misleading to classify this manuscript as "technical note". Although the Authors do not introduce any novel method, they apply hydrological tools and models through a somewhat novel procedure, which I suppose it can ascribed to a scientific paper, rather than a technical note. I see the technical note more like a document that reports further tests on a well-known procedure or method, with the final aim to be readily available for operational purposes. Instead, I think that the procedure presented might have a potential for predicting streamflow regimes in ungauged sites, as it relies on previous studies on dissimilarity-based techniques, published by one of the co-authors (Ganora et al., 2009). Nevertheless, this potential must be further investigated. For this reason, I think it is worth the effort to work on a better in-depth comparative study, with detailed results and comparisons with other models.

Reply: First of all, we thank reviewer for positive criticism. By and large, we agree with the reviewer. However, as the worthy reviewer himself/ herself discussed in the later part of the comment that the procedure relies on previous studies done in the context of dissimilarity-based hydrological prediction, published by one of the co-authors (Ganora et al., 2009). For this sole reason, we opted to go for submission as a technical note instead of a full fledge paper. The suggestion of reviewer about the comparison of our method with other established procedures is very useful. In the revised draft, we will briefly compare the performance of our method with other practiced methods.

Methodological comments:

Even though I have understood the idea behind the whole study, I really struggled with how the swapping applies in ungauged sites. In addition, the Authors use leave-one-out cross-validation strategy (see L 170-171 P5), however they never emphasize this. In my view, given the

operational purposes of the study, this should be reported better, e.g. added in the abstract and clearly stated within the body of the text.

Reply: We agree with the reviewer. We elaborate a step-wise procedure for the application of the proposed methodology on the ungauged basins;

1) Selection of Original and Swapped models by considering Δ and R_{adj}^2 .

2) Estimating Δ_{NN} and C_f values for each basin using OM and SM. Where, Δ_{NN} represents the error generated in predicting hydrological data of NNs of u_g in the cluster formed by u_g^{NN} and NNs of u_g^{NN} (using hydrological data although unknown for u_g but known for its NNs (u_g^{NN}) as well as NNs of u_g^{NN}). Whereas, C_f represents the coverage factor created by NNs of u_g in the descriptor space (using descriptor data known for NNs as well as u_g).

3) Prefer SM over OM iff $\Delta_{NN}^{OM} > \Delta_{NN}^{SM}$; and $C_f^{OM} < C_f^{SM}$.

Although these points are already inculcated in the text but we will elaborate these points more comprehensively in the revised draft.

I am not sure about the normalization applied to either the descriptors or the discharges (i.e. hydrological variables). Variables must be comparable from one site to another. This assumption is fundamental in regional analyses, in fact, in many cases reported in the literature (among all, refer to see Blöschl et al., 2013) some sort of standardization is always employed to the streamflows, e.g. using the mean annual flow (or monthly in this case) as reference values. I might be wrong, but I do not see this step in the manuscript except for a general statement at L89-90 P3, however other normalizations seem to be used the dissimilarity indices (see L140 P4), but none for the descriptors, correct?

Reply: We completely agree with the reviewer. Indeed, some sort of standardization is always employed to the streamflows, e.g. using the mean annual flow (or monthly in this case) as reference values to make variables must be comparable from one site to another. The descriptors and hydrological data in our work are normalized by using mean values at each basin.

I have found the mathematical notation really poor and misleading throughout the text, with some relevant inconsistencies (see e.g. how the subscripts in eq. 4 do not match with the definition in section 2).

Reply: This comment has also been raised by the reviewer 1. In revised draft, we will make sure to remove these inconsistencies.

Other comments:

The manuscript is not well structured and the writing is confusing, with many errors that sometimes make very difficult the understanding of the analysis.

Reply: We completely agree with the reviewer. Honestly speaking, we tried to make the article as short as possible to make it look like a technical note. However, with serious concerns raised over the structure we are compelled to restructure it. The revised draft will be restructured carrying more general contents.

Please, use English for reporting variable name (descriptors' names, see table 1). This is a minimum requirement for any manuscript, whether it is a technical note or a scientific paper.

Reply: The revised draft will have all descriptors notated in English.

Since the Authors are using leave-on-out cross-validation it is possible to draw scatterplot of empirical vs. predicted dissimilarity-indices, I strongly recommend using graphical tools rather than long tables (which do not really help understanding the results), or, alternatively, please prefer summary tables in the body of text with supplementary material for the complete reporting of the results.

Reply: We agree with the reviewer. We will add summary tables and scatterplots in the revised draft. We initially thought of adding the scatterplots. However, we ditched the idea as this will considerably increase the length of the manuscript owing to the fact that we will have to add multiple graphs between different variables.