Hydrology and Earth System Sciences Discussions

Hydrol. Earth Syst. Sci. Discuss., referee comment RC2
https://doi.org/10.5194/hess-2022-213-RC2, 2022

Comment on hess-2022-213

Elena Toth (Referee)

Referee comment on "Hydrologic Interpretation of Machine Learning Models for 10-daily streamflow simulation in Climate sensitive Upper Indus Catchments" by Haris Mushtaq et al., Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2022-213-RC2, 2022

Dear Authors,

I do apologise for the delay in closing the discussion: unfortunately not all the referees who accepted the review uploaded their comments, despite several reminders, and I will therefore provide a comment myself, as Referee as well as Editor (on the other hand, you should have contributed to the discussion replying at least to the comments uploaded by Ref#1 more than 6 weeks ago, since you had all the time…).

I will also send separately to you the annotated manuscript with additional details of my revision.

REFEREE'S COMMENTS

The work presents an analysis of different data-driven methods for predicting streamflow in a set of catchments in the Upper Indus. Such a comparison is not innovative, but the study region is extremely interesting, also given the importance of snow-generated flow and the main novelty is the application of a novel method (SHAP) for understanding the role of the different meteorological predictors.

GENERAL COMMENTS:

-check all the references, both in the text (also amending format, according to journal guidelines) and in the final list (where some cited works are missing)

-revise the English that is still often not clear and editorial review on the English is needed

MAIN COMMENTS

Abstract: I agree with Ref#1's comment that you should better explain and justify why ML approaches should be particularly suitable for data-scarce catchments and why you think that "Given the catchment characteristics, there is an utmost need to develop machine learning models that are hydrologically robust". To do so, on the first point (and first line of the abstract), I would perhaps suggest replace "Data-scarce" with a phrase explaining that there is a recent abundance of meteorological data, but little information on the catchment properties and on the characterisation of detailed hydrological processes.

The **Introduction** should be re-organised to better present and separate the two main issues of your work, i.e. i) ML methods and their potential also for identifying the influence of the inputs and ii) the challenges in snow-driven areas like the Indus upper catchments

In particular:

ll 45-52 and 53-64 should go together with ll 24-35

ll 53-64 may be removed (not needed: you may limit the state-of-the-art to the models you use and the applications in your region (already included in the following)

ll 35-39 should go with together with ll. 19-24 (and ll. 19-24 that should be rephrased/shortened or even removed, since there is no need to highlight, in a general way. the importance of streamflow prediction, that is obvious and certainly not limited to the last decades…)

Study area: clarify which is the closure sections of each catchment (adding also the drainage area) and denote them always with the same name (not mixing name of river and name of section)

Data sets: add more information on the choice and justification in reference to previous studies in the region

Predictors: the main features are too many (5 multiplied by the number of elevation zones) and they are certainly strongly correlated (for the same feature between elevation zones, but also between PET and temperature for each zone): why not performing a Principal Component analysis (or other, possibly non-linear, methods), in order to get a smaller set of predictors, to simplify also the interpretation of the results?

Methodologies: the description of the SHAP method (see Eq. 1), that is the core of the work, is far from clear: try to explain better, possibly with the help of figures/flowchart?

Results: clarify over which period (training/calibration or testing/validation?) the goodness-of-fit indexes and the difference in observed/modelled signatures are calculated (text Section 5.1+ Table 4 + Figure 6)

Move section 5.3 together with 5.1 (and in 5.1 elaborate more on the results and comment on differences between the models in capturing signatures)

Section 5.2 is far from clear and needs to be thoroughly revised, starting from clarifying the meaning of the figures you present: first explain and discuss the Feature Importance Order you obtained for each basin and for each model. Then explain the meaning of the shape, colors and position on the plot. Finally present the interpretation for each basin (or couple of basin) in separate subsections. This section is the core of the paper, representing the novelty of your research and as it is in not understandable, so it is impossible to assess the value of the work.