

Interactive comment on “Hydrologically Informed Machine Learning for Rainfall-Runoff Modelling: Towards Distributed Modelling” by Herath Mudiyansele Viraj Vidura Herath et al.

Anonymous Referee #1

Received and published: 11 November 2020

The paper addresses the problem of coupling traditional hydrological models with machine learning techniques. The problem addressed is relevant and a current topic of interest in the hydrological community and this paper potentially represents a valuable contribution.

However, I see some deficiencies in the paper organization and in the clarity of the content that make difficult to fully exploit the potential of the work. Therefore I believe that a major revision, mainly addressed to the form rather than to the content, is needed.

GENERAL COMMENT

There are many extremely long paragraphs (e.g. lines 70 to 90) that express multiple

C1

concepts. I would make paragraphs shorter, creating a paragraph per concept. This should help the readability of the paper.

There are several statements without a reference justifying them.

OVERALL STRUCTURE

I do not like the structure of the paper and the amount of content given to each paragraph. Your main message is to present MIKA-SHA but this is left to section 4, which is barely 1 page out of 42. If that was the main concept of the paper, I would give it more space.

I like the material in sections 1 to 3 (included) but they are basically a mix of introduction and a “methods” section. I would move some content from section 2 and 3 to the introduction, which can be divided in subsections. A possible structure of the introduction can be:

- Quick introduction on hydrological modelling and TGDS
- On hydrological models
 - * Physics-based models vs. conceptual models vs. data science models (mix of 2.1 and 2.2)
 - * Focusing on conceptual models, difference between fix and flexible structure (some part of 2.2)
 - * Lumped vs. distributed (2.3)
- On ML models
 - * Some generalities (3)
 - * ANN (3.1, maybe reduce)
 - * GP (3.2)

- Physics informed ML (3.3)

I would then add a “methods” part where you write about what of sections 2 and 3 you actually use: SUPERFLEX, FUSE, GP. I would also add to this the metrics that you later use in section 5.

Section 5 and 6 are about the case study and I think this should represent a minor part of the paper (which objective is to present how MIKA-SHA works). To this end, I would more or less keep the same content and put it in a single section divided in:

- Presentation of the case study
- Settings
- Results that you get
- Meaningful discussion, potentially showing that MIKA-SHA works.

I would therefore move some aspects of this section elsewhere:

- Metrics to a “methods” section
- Further explanation of MIKA-SHA functioning to the section that presents the model
- General implications on the goodness of the approach to a general “discussion and conclusion” section

PAPER CONTENT

Keeping in mind that I do not have a deep knowledge of GP, I find quite difficult to understand what MIKA-SHA actually does and my difficulty can be motivated by the following reasons:

- Use of jargon from GP, that may be not common in the hydrological community (e.g. model induction vs. model selection)
- Assuming good familiarity of the reader with GP

C3

- Assuming that the reader knows ML-RR-MI: you do not have to re-write here that paper but at least explain here the concepts that are necessary. I find it difficult to follow something that says that MIKA-SHA is basically ML-RR-MI plus something else.

DATA AND CODE AVAILABILITY

I do not know if HESS forces the sharing of the source code but I believe that, potentially, MIKA-SHA can be a valuable tool for the hydrological community and, therefore, I invite you to make it publically available.

FURTHER COMMENTS

See the attached PDF with the comments.

Please also note the supplement to this comment:

<https://hess.copernicus.org/preprints/hess-2020-487/hess-2020-487-RC1-supplement.pdf>

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., <https://doi.org/10.5194/hess-2020-487>, 2020.

C4