Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2018-168-AC1, 2018

Hydrology and Earth System Sciences
Discussions

Interactive comment on "HESS Opinions: Deep learning as a promising avenue toward knowledge discovery in water sciences" *by* Chaopeng Shen et al.

Chaopeng Shen et al.

cshen@engr.psu.edu

Received and published: 4 June 2018

Dear Sivarajah Mylevaganam,

Thanks for the online interaction discussion. Please see attached response to your comment.

"1) Even though the word "hydrology" is meant for science ("logy") of water ("hydro"), in the current version of the manuscript, the boundary between water sciences and hydrological/hydrologic sciences is not visible. As per the title of the manuscript, the manuscript is about deep learning in water sciences. However, the content of the

Printer-friendly version

Discussion paper

manuscript is merely constrained to hydrological sciences (e.g., see P-1 LN-22)."

After some collective deliberation, we indeed plan to change the title to hydrology rather than water sciences. Thanks.

"2) What is meant by "HESS" opinions? Are the opinions echoed in the manuscript represented by the journal office? Do the authors represent the journal office? Do the authors represent the editorial board? What is the expected outcome of the review process of this manuscript? If the null hypothesis is that the HESS opinions are always published regardless of the review process, what should be the appropriate alternate hypothesis that needs to be tested for a given significance level? Are the available data sufficient to conduct this hypothesis testing?"

HESS Opinions papers do not represent journal opinions. The reviewer is referred to some previous HESS Opinion papers. https://www.hydrol-earth-syst-sci.net/special_issue62.html https://www.hydrol-earth-syst-sci.net/18/2615/2014/hess-18-2615-2014.pdf https://www.hydrol-earth-syst-sci.net/13/157/2009/hess-13-157-2009.pdf https://www.hydrol-earth-syst-sci.net/20/3739/2016/hess-20-3739-2016.pdf Per HESS website https://www.hydrology-and-earth-system-sciences.net/about/manuscript_types.html ïĆğ Opinion articles discuss a topical aspect of hydrology. These articles are not peer reviewed in the traditional sense, but they are discussed openly in HESSD so as to stimulate an open debate among peers on new ideas, views, or perceptions in hydrology. Opinion articles will be published under the heading "HESS Opinions" and are handled by one of the executive editors. Opinion articles are generally invited, but authors with ideas for an opinion paper are encouraged to contact an executive editor. The manuscript title should start with "HESS Opinions:".

"3) As per the authors, deep learning, which has gained widespread attention since 2012(see P-2 LN-1), is a suite of tools centering on artificial neural networks. Is there a specific reason for the authors to prefer the year of 2012? In my opinion, the fusion

of information theory and deep learning in hydrological sciences was well rooted even before 2012. Therefore, an appropriate reference is needed to support the authors' statement."

2012 is when DL first won a major competition and started to garner attention, including the Merck competition and the ImageNet. Although there have always been researchers in this area, the winning of these two competitions added big fuels to the fire. Yes this point will be explained more in the revised version.

"4) The titles of some of the subsections are not acceptable at a significance level of 5%. For example, the tile of subsection 2.1 is with more data, opportunities arise. What is the HESS opinion on writing titles for sections/subsections?"

It is not certain what was referred to here. This sections are named as summaries of the section. We'd appreciate if more clarity is given on what is wrong with this subsection title. With or without this comment, this subsection title may be revised as part of the whole revision endeavor

"5) As per the authors, compared to classical DL problems, hydrology has a unique set of challenges that are research opportunities for DL (See P-10 LN-16). In the subsequent sentence, the authors state that DL research has not cover these questions extensively. What are those questions? I think, the paragraph (P-10 LN-16) needs to be re-written."

"These questions" means the points that immediately follow this paragraph. Yes, we will revise the wording to avoid this confusion.

"6) As per the authors, DL models have already been used as surrogate models for PBMs, but many novel ways that couple the two (i.e., PBMs and DLs) should be investigated (see P-11 LN-18). In my opinion, this has already been investigated in one of the PBMS (SWAT) in hydrology."

There have been limited studies that integrate ML with PBM. One of such examples

is the following reference, which is co-authored by one of the co-authors of the current paper. It is also perhaps what was referred to here. Mekonnen, B. A., Nazemi, A., Mazurek, K. A., Elshorbagy, A., & Putz, G. (2015). Hybrid modelling approach to prairie hydrology: fusing data-driven and process-based hydrological models. Hydrological Sciences Journal, 60(9), 1473–1489. https://doi.org/10.1080/02626667.2014.935778 In the above paper, SWAT is used to represent the runoff generation from contributing areas and the ANN model is used to represent the nonlinear overflow generation from the non-contributing areas. However, this is only one way such coupling is done. There are many other ways in many different applications, such as partially replace the functions for each other, use PBM as inputs to DL, use PBM as constraints for DL, etc. We will include this paper as a citation, but there are so many other things to do in this regard. We will add some examples of this in the revision.

"7) As per the authors, the evidence is mounting that when given "enough data", DL can provide the "unique ability" to automatically extract features, sometimes "better than human experts" do(see P-4 LN-28). Subsequent to this statement, the authors provide few bulleted points. What are meant with those bulleted points? Are the bulleted points meant to show that when given "enough data", DL can provide the "unique ability" to automatically extract features, sometimes "better than human experts" do?"

Indeed here we can use a sentence to transition more naturally. We will add the following sentence: 'The performance gain by DL can be witnessed by an increasing number of competition wins by DL-based models and adoption in the mainstream information technology industry.'

"8) Should the abbreviation ML (P-5 LN-33) be introduced in one of the previous pages (see P-3 LN-6)?"

We will remove this abbreviation as it is not needed. Thanks for pointing it out.

"9) On P-3 LN-20, except for satellite-based data products of precipitation, references are given for all other large available datasets (e.g., soil moisture, evapotranspiration,

and streamflows) mentioned in the manuscript. Is there a specific reason for not citing a research paper for satellite-based data products of precipitation?"

The reason is there were many precipitation products. However, we will add some references there, too.

"10) With the emerging datasets, DL models can be built and trained to learn features, organizational patterns and relationships and predicts outputs given new input instances (P-3 LN-28).However, the authors are not advocating a whole transition to DL as some of the problems, specifically the problems with just not enough data to train DL-based models, could be best tackled by specifically designed earlier-generation models. I think, it would be more appropriate to show an example (may be in hydrology) of how to use DL models and how to use specifically designed earlier-generation models to avoid transition to DL."

Yes this is a good suggestion. We can add an example in the revision.

"11) Considering the number of authors listed in the manuscript and the quantity of the work carried out in the manuscript, I think, it becomes vital to list each author's contribution in the manuscript."

This would not be necessary. We again refer the reviewer to earlier HESS opinion papers. None of those papers did this.

"12) What is meant by citizen "scientists"? What is the minimum required qualification? Does the definition of citizen scientist vary spatially and temporally? Minor Comments P-3 LN-19: should it be Srinivasan, 2013 P-4 LN-5 to P-4 LN-9: The language needs to be checked"

Citizen scientists mean volunteers who are willing to spend their time to help provide relevant measurements and data. Good point that we should not assume this term is generally known or accepted. More clarification will be added.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2018-168, 2018.

Printer-friendly version

Discussion paper