Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2020-566-RC2, 2021

Hydrology and Earth System Sciences
Discussions

Interactive comment on "Using an ensemble of artificial neural networks to convert snow depth to snow water equivalent over Canada" *by* Konstantin Franz Fotios Ntokas et al.

Anonymous Referee #2

Received and published: 6 January 2021

This manuscript describes the application of machine learning techniques, specifically an ensemble of multilayer perceptrons, to estimate the hydrological variable Snow Water Equivalent (SWE). As described by the authors, SWE is a crucial variable which is difficult to directly measure at scale. The paper subject advances the application of ML techniques for SWE estimation by application of ensemble methods, discerning model applicability over climatic regions. and demonstrating advantages over empirical methods.

This research builds on the Ordy et al (2020) study by extending the geographical scope, using direct estimation of SWE and introducing snow classes in MLP train-

Printer-friendly version

Discussion paper

ing. There is sufficient novelty in this paper for publication, however it should first be strengthened in clear justification for decisions, interpretation of results and conclusion.

The manuscript leaves out some essential elements from Ordy et al, including descriptions of evaluation metrics and why they are selected. The addition of explanatory variables such as Snow Density from ERA5 lack explanation and background. Following strong results reporting sections, the discussion finding and conclusions of the manuscript require additional reflection on the limitations of the study and the context.

Overall, a strong effort worthy of publication on the basis of some revision and structural improvement. Some coherence is missing in the experimental design, in the inclusion of variables, the applicability of the study and the conclusions drawn from it. These require revision, hope the comments that follow can be of help.

Pg 1, ln 15: "Using a greater number of MLP parameters could lead to further improvements" It is somewhat self-evident that increased parameterization of an MLP model could potentially produce better results, can this statement be focused to the study specific outcomes?

pg 2, 50: This description of the application of physics-based models for SWE estimation is a bit too simplistic here, given ERA5 snow density as used later as an explanatory variable. The iSnobal mentioned is a coupled energy and mass-balance model that requires a great deal of meteorological data derive accumulating snow density and in turn modelled SWE. Please provide some further description on the advanced requirements these approaches and limitations, beyond only computational cost.

pg 2, ln 54: Consider including Snauffer et al, 2018. https://doi.org/10.5194/tc-12-891-2018.

pg 3, ln 76: The second hypotheses seems too broad. "in-depth testing". Please be more specific as to the methodology to be tested.

pg 3, ln 78: "The entire area of Canada." Is this an overreach given the the limited

density of measurements across much of Canada? "Applicability in a broader context". Be more specific in the what this broader context is.

pg 3, ln 78: This last sentence seems out of place to close the paragraph. Moving one sentence earlier would improve the paragraph.

pg 3, ln 94: Is MSE the definitive objective function for regression problems? Better likely to phrase as "commonly used"

pg 5, ln 131: "The algorithms RMSProp and AdaDelta produce good results". Please elaborate in this statement, or tie in better with the following two sentences.

pg 5, ln 155: Avoid starting sentences with "Because". This is a general comment also through the manuscript. Would recommend re-writing this initial sentence, breaking into sections.

pg 6, ln 171: Can references be provided for some of these conclusions? The linkage of snow depth only to precipitation requires some basis. There are a lot of varied physical processes for snow accumulation between tundra and taiga eco-zones.

pg 8, ln 222: I don't follow the second sentence, though can be my ignorance. Consider a clearer explanation if including.

pg 12, Figure 3e. It is possible to rescale the number of records for each site? It is not very descriptive. Is the maximum records up above 3000?

pg 13, Ordy et al had recommended the inclusion of additional explanatory variables from meteorology, such as wind or solar radiation. This study has included ERA5 daily averaged snow density data. This output from a physically based model is included without description of its generation, assessment of the quality or the relevance or applicability of this data source. Although the variable is kept as least important for the conversion model, and minimal impact on the ignorance score, it is kept in complete assessment. The manuscript should include some rationale for the inclusion of this model output, and why it was chosen. Are the assumptions in producing the snow

density relevant? How does this data perform compared to available measurements?

Pg 15, Table 2, To clarify, on pg 4, ln 97 it is mentioned that modifying the order of the input data is recommended. Is that done in this study (Shuffling data before each epoch)?

Pg 17, Figure 5. MBE should be introduced before used. The use of error metrics is not entirely clear (MBE, MAE, RMSE) compared with clearly rational and description for other scores. Clearer rationale and explanation would help. Is there a reason the RMSE is shown compared to the objective function MSE? RMSE can be a more comparable error metric for SWE, but this is not explained.

Pg 20, line 449: This appears a notable and relevant finding (what explanatory variables are ultimately useful) that can be better articulated in study findings.

Pg 27, Figure 13: Consistency would be useful for interpretation between RB and MBE. Can see in the following paragraph why RB is substituted for MBE, but would like to see this graph included.

Pg 27, line 555: The conclusions drawn in this section appear to have a relatively weak causal or testable links. Ranging from regression model structure, to physical processes to reference model performance, several comments seem quite speculative. For example, the tundra region has poor performance, but would be subject to may be similar topographic controls of the prairies. It would seem better to reflect on what information can truly be derived from these results, or at least address that there are many contributing factors that are not represented by this method.

Pg 28: line 570: This opening sentence for the Conclusions section should be more descriptive and engaging in the content of the study.

Pg 29, ln 591: What is the additional geophysical information beyond snow class from Sturm et al.? If this refers the discretization by elevation class, this should be elaborated on in the rest of the document to include in conclusions.

Pg 29, ln 596: These statements are quite generalized, and should be refined. What variables should be added and what information content due they bring? What information is missing that could be provided by other sources and why are they not now included?

Pg 29, general: What are the limitations of the study? What is it's applicability?

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2020-566, 2020.

Printer-friendly version

Discussion paper