

Interactive comment on “Deep learning applied to glacier evolution modelling” by Jordi Bolibar et al.

Anonymous Referee #1

Received and published: 28 August 2019

Deep learning applied to glacier evolution modelling: Manuscript review

The research presented in this manuscript shows promising results in the application of an ANN model used for surface mass balance modelling. The manuscript is, for the most part, well organized. The manuscript can be greatly improved by increasing clarity and specificity throughout. I hope that my comments are helpful to the authors in this effort. No single one of my comments identifies a major flaw with the manuscript; rather, there are many small changes that I believe can be made to improve the quality of the paper. I have organized my comments in sequential order by section, preceded by one general note.

General Note:

The difference between “machine learning” and “deep learning” is not clearly defined in the literature, but a 6-layer ANN is likely at the very tip of what may constitute “deep

C1

learning”. Considering that, as you note, deep learning is not a common tool among the glaciological community, it would be good to provide further context as to what an ANN is (a type of model, which is relatively simple in the deep learning world as compared to, say, a convolutional neural network or long short-term memory network) versus what deep learning is. I believe that this is required especially because the title refers to deep learning broadly, not a deep ANN specifically, and it should be made clear that there is much more to deep learning than ANNs.

Sequential Notes:

Page 1, Line 22: What does “individual glaciers at regional scale” mean? Do you mean to say, “individual glaciers within the same region”?

1 Introduction: Page 1, Line 25: “. . . being climate proxies which can clearly depict the evolution of climate for the global audience”; remove “clearly” – “if the evolution of climate was clear for the global audience, then why is there so much disagreement among the global audience?”

Page 1, Line 26: “For the coming decades. . .”; I believe this should be “In the coming decades. . .”

Page 1, Line 28: “The reduction in ice volume may produce an array of consequences which requires to be properly predicted.” This sentence, and the following, is vague. What consequences are you talking about? Be explicit.

Page 2, Line 2: “For any glacier model. . .”; Saying “any” makes this sentence too broad and not necessarily true. Be explicit for the classes/types/purposes of models which require SMB and glacier dynamics (e.g. “SMB and glacier dynamics both need to be modelled to understand glacier evolution on regional and sub-regional scales. Models of varying complexity exist for both processes.”)

Page 2, Line 18: “. . . these different approaches strongly depend on available data. . .”; Change to “. . . the use of these different approaches strongly depend on available

C2

data...” since it the model usage, not the model itself, which depends on what data one has.

Page 2, Line 21: “... relationships remain stationary.”; Change to “... relationships remain stationary in time.”

Page 2, Line 34: “... the glaciological community has remained quite oblivious to these advances...”; Oblivious is a strongly negative word to use here, and it is a disservice to insult your readers.

Page 6, Line 10: “... relevant predictors must be selected, performing a sensitivity study...”; Change to “... relevant predictors must be selected, so we perform a sensitivity study...”

Page 6, Line 14 and Equation 1: Is there a reference for this “effective way of expanding the training dataset”?

Page 7, last sentence: Here you describe the types of cross validations available in ALPGM. Which did you use?

Page 8, Line 6: “... (2) the optimizer: the method for...”; change to “...(2) the optimizer, which is the method for...”

Page 8, Line 6: “...(3) its (possibly nonlinear) activation functions...”; When are activation functions linear?

Page 8, Line 10: “... allowing to train deep neural networks...”; change to “allowing the training of deep neural networks...”

Page 8, Line 11: “...ANNs are best suited when the quality of predictions prevails over the interpretability of the model.” This is vague, and does not help readers know when ANNs are ‘best suited’. How are either of these things quantified?

Page 10, Line 21: “...should be long enough to be representative of the glacier evolution...”; How long is ‘representative’? Representative of what? How does one

C3

know this?

Page 11, Line 3: Refer to Figure 4 here.

Page 11, Line 5: Is there a reference to this study?

Page 11, Line 9: “... using remote sensing based on changes in glacier volume and the snow line altitude is used...”; Remove second “is used”

Page 12, Figure 4: Axes should be labelled.

Page 13, Line 2: Cite RGI (check here for reference: <https://www.glims.org/RGI/>)

Page 13, Line 12: Qualifications here are vague (e.g. “quite satisfactorily”, “good overall”, “certain altitudinal ranges”). Give quantitative measures of “goodness”, and refer to specific parts of Figure S2 that demonstrate what you’re talking about.

Page 13, Line 26: This sentence can be improved by maintaining consistency across clause structure. You use “we verb” statements (e.g. we go through, we assess, and we show) for all clauses except for “the building of the machine learning SMB models”.

Page 14, Line 25 (and paragraph): You discuss that you dynamically calculate the accumulation/ablation periods based on the CPDD, and that you keep constant periods to account for winter and summer snowfalls. Later, you use ‘transition months’ as predictors – are these predictors kept constant, or dynamically calculated? Are results improved when the transition months are dynamically computed? I ask because I would expect that what constitutes a ‘transition month’ may change in the future. Or do you think that this approach, applied to more variables, then forces the model to depend too much on CPDD when the CPDD is not the only variable involved in melt?

Page 15, Equations 2 and 3: Are input variables normalized? If so, how?

Page 15, Line 20: When you say ‘linear machine learning’, are you referring to the linear regression methods? Be consistent in how you refer to your methods.

C4

Page 15, Line 20: How did you choose the function f ?

Page 15, Equation 25: You create linear models using the predictors shown here. You then create nonlinear models using only the predictors in Equations 2 and 3. Then, you compare the results of these models and conclude that the nonlinear model is better because of the nonlinear nature of the model; however, how do you know that the improved performance is not simply due to using a different set of predictor variables? Your argument would be more convincing if you first showed that the linear model performance improved when you change predictor variables from the standard case (those only in Equations 2 and 3) to the combination case (Equation 4), and then showed that a nonlinear model using variables from the standard case outperformed even this improved linear model.

Page 15, Equation 5: Is there a missing '('? This equation ends with ')_g,y'.

Page 16, Line 1: It is not clear to me why there are 50 predictors, when there are 33 coefficients in Equation 4.

Page 16, Line 2: Can you please be more explicit about what this matrix is, and to what matrix equation it is input into?

Page 16, Line 8: "...the annual CPDD as well as the winter and summer snowfall appear as significant predictors as well as several monthly mean temperatures and snowfall values. . ."; Change to "...the annual CPDD, winter and summer snowfall, and several monthly mean temperature and snowfall were found to be significant at $p < ?$..."

Page 20, Figure 8: This is a challenging plot to interpret (in my opinion). Can you plot deep learning bias vs lasso bias and deep learning MAE vs Lasso MAE? That may more clearly demonstrate the points you make, and may reveal structure. The points in the scatter plots could be coloured by region (Ecrins, Vanoise, Mont Blanc) if there are regional patterns. If this approach is not useful or helpful, then the current figure will suffice.

C5

Page 20, Line 15: "This implies, that. . ."; remove the comma

Page 22, Line 7: "...using Leave-Some-Glaciers-and-Years-Out (LSYGO)"; the abbreviation should be LSGYO, or the full phrase should be Leave-Some-Years-and-Glaciers-Out, for consistency.

Page 25, Lines 4-5: "The greater the dropout, the more we will constrain the learning of the ANN so the higher the generalization will be, until a certain point." This sentence is not clear. What does it mean to "constrain the learning"? Why is there a "certain point", and what happens beyond that point? This could be made more explicit.

Page 25, Line 7: Why is it that slower convergence leads to better generalization? Is this always true?

Page 25, Line 8: "... that best suits a certain dataset and model." How does one define "best"?

Page 25, Line 17: "Despite it has been shown. . ."; Change to "Although it has been shown" or "Despite the fact that it has been shown"

Page 25, Line 28: "The results were quite astonishing. . ."; If the results are astonishing, then this result warrants further emphasis in the paper. The methods used to come to this conclusion should be brought up in Section 3, and further discussion is warranted in Section 4. It is worth a figure to communicate these results.

Page 26, Lines 5-16: This paragraph is speculative, but is presented with a high degree of confidence. Phrases such as "unprecedented efficiency" and "excellent" are used without supporting evidence. Much of the discussion is implied; for example: "An interesting way of expanding a dataset would be to use a deep learning approach to fill the data gaps." It is unclear how this would be done. "Such an approach would be an excellent way of obtaining more SMB data in remote glacierized regions such as the Andes or the Himalayas." This is not known or demonstrated by the rest of the paper. I would recommend either removing this paragraph entirely or severely limiting

C6

its scope.

Page 26, Lines 18-22: This paragraph is speculative. It does not follow from the results presented in the paper, and is more of a justification for using deep learning in glaciology than it is an item of discussion in the context of the preceding research. These final two paragraphs do a disservice to the rest of the paper; prior to this, the organization had nice flow, and the first two paragraphs in Section 4.3 were both interesting and directly relevant.

Page 27, Lines 4-12: This paragraph is quite vague and does not explicitly follow from the research. For example: "It might still be too early for the development of such models in certain regions" is a vague statement. Conclusions should follow directly and explicitly from the work and should not reach beyond the scope of the research. The first paragraph in Section 5 is much better.

Supplementary Figures: In the SMB_lasso_ANN_no_weights_SMB_simulations.pdf file, y-axes are missing units.

Figures 6, 7, 9, and 11: Please increase font size, especially of axis labels.

Interactive comment on The Cryosphere Discuss., <https://doi.org/10.5194/tc-2019-163>, 2019.