The Cryosphere Discussions

The Cryosphere Discuss., referee comment RC1
https://doi.org/10.5194/tc-2022-34-RC1, 2022

Comment on tc-2022-34

Pascal Hagenmuller (Referee)

Referee comment on "A random forest model to assess snow instability from simulated snow stratigraphy" by Stephanie Mayer et al., The Cryosphere Discuss., https://doi.org/10.5194/tc-2022-34-RC1, 2022

Summary :

Snow instability data is essential for avalanche forecasting. The authors developed a random forest model to estimate snow instability from simulated snow stratigraphy. In details, the model estimates whether a weak layer - slab system simulated by a detailed snowpack model is unstable or stable according to a measured rutschblock test. To train the model, they used about 140 data points (rutschblock score + measured stratigraphy) in the region of Davos (Swiss) and they manually associated the observed weak layer - slab system (revealed by the stability test) to its equivalent in the modeled stratigraphy. The model evaluated on other points around Davos (300 points) and in other Swiss regions (120 points), shows a very good overall performance (accuracy 88% precision 96%, recall 92%). The model can then be indifferently applied to any weak layer - slab system (i.e. all the layers in a profile) to provide a vertical profile of the instability probability P_unstable. It appears that the weak layer revealed by the measured rutschblock test is in most cases the most unstable in the simulated snow profile. In addition, they computed the time evolution of P_unstable with simulated snow stratigraphy representative of the regions around Davos and compared it to the avalanche activity in this region.

Overall comment :

The authors tackle a very important problem for the snow and avalanche community: how to provide synthetic indicators relevant for avalanche forecasting from potential huge amont of simulated snow data. I am very impressed by the obtained results. Moreover, the paper is very well written and is comprehensive with a deep analysis of the model behavior and a detailed presentation of the data pre-processing which essential for

machine learning approaches. It might be sometimes difficult to catch the key results among all the results presented, but with a second read it becomes clear enough. However, the interpretation of the model explanatory variables should be qualified as the used feature importance metric is affected by correlation between the input variables and contains only partial information of the underlying « physics ». Overall, I suggest accepting this very good paper with only minor revision I have listed below.

Minor comments :

Abstract : add somewhere that the study domain is mainly around Davos and in Swiss.

L11 : give number of points in the validation data set

L14-16 : you provide the accuracy for discriminating the non avalanche / avalanche days. However, if the data is not balanced it is difficult to interpret. Use the same clear sentence as in l390-392.

L44 : the model MEPRA (Giraud, 1992) is one of the first model that tried to combine different metrics of snow instability into a synthetic index. Add historical reference in the text.

Fig. 1 : « virtual slope simulation » => « simulated snow profiles »

L83 : give reference of the rutschblock score from 1 to 7 or explain its meaning.

L89 + 105 : « RB tests failed adjacent to layer of persistent grain types ». I do not understand what is meant here. Do you mean: the weak layer revealed by the RB test is in 64% cases composed of FC, DH or SH ?

L90 and throughout the text : « to evaluate the model », it is not clear what is the model here. Indeed, the « basic » model predicts whether a weak layer - slab system is unstable or not. I understand that you applied your model more extensively to simulated snow profiles. But be more specific.

Fig. 2 : x-label and plot title appear in the same form which is confusing.

L214 : « similarity criteria » to be defined. Do you mean criteria 1-5 ?

L274-277: reword. Not clear to me. The use of the probability is not related to the fact that you want to apply the model to any layer of the profile ???

Fig. 8 : I do not understand the role of Fig.8b as the goal is here to see how the model works on intermediate instability classes.

L325-329: I did not understand your point here, could you be clearer to explain your point (L330-331).

L389-390: could you plot on Fig. 13 the avalanche and non-avalanche days as defined in this paper.

L402: « Figure c » => « Figure 15 c » ?

L425: « they were mostly developed to align complete profiles ». In practice, this is not true as a parameter of the model can be used to align only a sub part of the profile. In particular it is used to relax the assumption that the snow-ground interface must be matched. Besides, it is not a limit of the method since for the manually matching you also look below the weak layer for stratigraphy markers (eg. MF-crust). « these additional parameters are not included in the current available automated methods » It is implemented and shown in Viallon-Galinier et al. (2020). Actually your manual method seems to works fine enough and you do not necessarily need an automatic method. You might see the automated matching method as a further development to reduce the time spent to prepare the data but you do not need to say something wrong about the automated method limits.

Section 5.3 : all your analysis is based on the feature importance as computed by the scipy package. First, here, you do not give any information on the « sign » (> or <) of the important feature. For instance, it is not clear (and there is no info about that) whether it is high or low values of « mean density divide by mean grain size » that promote instability. To be added. Besides, the feature importance are somehow « shared » between correlated variables. For instance, viscous deformation might be correlated to the initiation criteria such as SK38 (stress over strength) which is itself correlated to strength, stress (and so importance shared …). Your comment about the absence of initiation criterion must therefore be qualified. Moreover, your comparison of your model score (6 parameters, training) to the « physical » model with only two parameters and no training

is unfair (L. 478).

Fig. 13 and 14 and Section 4.2.5 : the results at the regional scale are very interesting but never discussed in the paper. In particular, the model apparently failed (?) to detect clearly the big avalanche events (high AAI) at the regional scale. Add a discussion on the inherent difficulty to predcit high AAI from only slab stability indices (size, spatial distribution, natural release).

Reference :

Giraud, G., 1992. MEPRA an expert system for avalanche risk forecasting, in: International Snow Science Workshop. Breckenridge, Colorado, USA, pp. 97–104.

Viallon-Galinier, L., Hagenmuller, P., Lafaysse, M., 2020. Forcing and evaluating detailed snow cover models with stratigraphy observations. Cold Regions Science and Technology 180, 103163. https://doi.org/10.1016/j.coldregions.2020.103163

Pascal Hagenmuller