Interactive comment on “Data efficient Random Forest model for avalanche forecasting” by Manesh Chawla and Amreek Singh

Anonymous Referee #2

Received and published: 5 February 2020

(1) This paper predicts snow avalanche using the Random Forest model. The research area is young and the paper is interesting, however, some improvements should be considered before publication:

(2) Forecast or predict? Hazard or risk? Please be consistent with using the phrases.

(3) All full names should be presented in their first occurrences, for example, CROCUS in L40, SAFRAN in L44, and etc.

(4) Literature review is missing. As a research paper, this submission needs to critically assess work previously carried out in the scientific field. Although this has been done to a limited extent in the introduction, some key references are missed. For example, Choubin et al. (2019) predicted the snow avalanche hazard using machine learning methods.

(5) L101-102: I do not agree with using this sentence in the introduction. Transfer it to the conclusion/discussion or delete it. Instead, state the objectives clearly (preferably as listing (i) (ii), etc) at the end of the introduction.

(6) L 172: Figure 4 or Fig. 3?!

(7) What is 0830 hours?

(8) Table 2: It is mentioned that star represents parameters derived using training labels. What does that mean? Please clarify more.

(9) Table 3: Add reference for FAR, POD, and precision, too.

(10) How did you split the dataset into training and testing data? What ratio has used?

(11) The figures’ number must be checked.

(12) Location of the avalanches should be presented.

(13) Tables 5 and 6 are confusing. These metrics are calculated after modeling run, but I can not see the modeling conditions in each row of the tables. There is some missing information in these Tables.

(14) Table 7: Performance of the Random Forest model is lower. So, how did you suggest this model?
