The Cryosphere Discussions

The Cryosphere Discuss., referee comment RC2
https://doi.org/10.5194/tc-2022-15-RC2, 2022

Comment on tc-2022-15

Hannah Vickers (Referee)

Referee comment on "Snow Avalanche Frequency Estimation (SAFE): 32 years of monitoring remote avalanche depositional zones in high mountains of Afghanistan" by Arnaud Caiserman et al., The Cryosphere Discuss., https://doi.org/10.5194/tc-2022-15-RC2, 2022

This paper "Snow Avalanche Frequency Estimation (SAFE): 32 years of remote hazard monitoring in Afghanistan" attempts to produce inventories of avalanche debris using Landsat optical satellite imagery in late spring when snow, bare ground and water are easily distinguishable. The concept of using a long time series of remote sensing data to identify hotspots of avalanche deposition zones and trends in their spatial occurrence is good, but there are many pitfalls with the overall implementation and communication of the work which reduces the impact.

1. The paper requires some major restructuring of the content, starting with the introduction. Throughout the paper I found that information was in the wrong order and/or wrong section. Results were presented already in section 2 (eg. Table 2) and discussions were being made in the results section. This makes the work difficult to follow, even with the flow chart provided. Moreover figures are wrongly labeled (Fig. 10) and have unsatisfactory captions or text to explain what is being shown or how they were produced, color scales are not constant making figures hard to compare (Fig. 6-8)

2. It seems to me that the authors are basically identifying late season snow patches in valley bottoms close to rivers which they are assuming to be avalanche deposits. This is made quite straightforward by the fact that the regions of interest are snow-free and snow is easily distinguishable by higher NDSI in the Landsat images compared with bare ground or water. This just reduces the problem to a simple thresholding and classification of image pixels into 3 classes, and I fail to see what is state-of-the-art in this approach. Moreover the authors have employed MODIS data to identify the snowline in order to select the dates and regions which are snow-free. MODIS has poorer spatial resolution than Landsat, so why not just use the Landsat data to identify the snowline? I can't see any value in using MODIS vs. Landsat for this purpose.

3. Throughout the paper the authors emphasise that the approach is based on Landsat

data and the use of the google earth engine because it should be used in areas where internet connection is poor. However, they also highlight that the main end-users of such a dataset are stakeholders and decision makers. Are these stakeholders and decision makers likely to be located in remote mountain villages or the main cities (where internet connection is presumably good)? Are local villagers in these mountain environments really likely to be making use of this dataset? I find it hard to believe that knowing where a large avalanche deposit has occurred several months prior to its detection is likely to be of interest to these people.

4. As pointed out by reviewer 1 the classification of avalanche size seems quite arbitrary and does not have much meaning when it is being detected late in the season after it has already partly melted out. It would be more meaningful to show for example a histogram of the avalanche size to show what is being detected rather than applying some random size classification to the detected deposits.

5. Inconsistent terminology. Avalanche debris/deposits are referred to as "snow packages", "snow patches", "avalanche depositional" in the paper. The authors should use the correct term and use it throughout.

6. Poor validation. In section 3 the authors state that over the 32 years of data analysed they identified around 810,000 avalanche deposits using their dataset. However for the calculation of POD and PPV as shown in Table 2 they have ony used 158 deposits observed using Google Earth images. Moreover they do not describe how the validation data were identified (was this done visually or was there some other algorithm used to detect them in these images?). Overall this does not come across as a satisfactory validation dataset with which to evaluate their detections.