Dear Reviewer,

The authors are very thankful to your comments. After we submitted a first response to your comments several weeks ago addressing them point by point eliciting significant modifications to the manuscript. All responses to your comments, as well as modifications that we will make to the manuscript are written in red in the text below your comments. We believe that this revised manuscript is an improved version that clarifies and addresses most all of your comments. Based on your review, we redirected the terminology and the concept of our paper from *snow avalanches mapping* to *avalanche deposit zones* mapping. Moreover, the classification of deposits is now based on surface area instead of the confusing term ‘size’. The revised version also includes the comparison between SAFE outputs and the dataset you provided in Switzerland; the authors are very thankful for this. We believe that this comparison shows interesting results, where datasets are comparable (deposit zones). This comparison also provided a better view on SAFE applicability in other high mountains regions of the world.

1.The state of the art is incomplete. Several publications, very relevant for this topic have to be considered and discussed. In particular the optical mapping with SPOT6 over the Swiss Alps is important (Bühler et al., 2019). But also, several mappings with Sentinel-1 are missing (Leinss et al., 2020; Karas et al., 2021; Vickers et al., 2016). Considering hazard indication mapping new developments allow for applications over very large areas (Maggioni and Gruber, 2003; Barbolini et al., 2011; Bühler et al., 2022) and was even already conducted in Afghanistan (Bühler et al., 2018). Therefore, the introduction has to be overworked including the relevant publications.

Many of these references have now been added to the literature review:

- Lines 61-62 “Vickers et al. (2016) conducted one of the first studies utilizing Sentinel-1 products to detect avalanches debris by developing an unsupervised classification.”
- Lines 64-66 “Using TerraSAR-X and Sentinel-1 products, Leinss et al. (2020) mapped avalanches, demonstrating the potential of radar products in snow hazard detection.”
- Lines 67-68 “Moreover, a recent study also used SAR products to detect avalanches and demonstrated both the potential and limitations of radar products due to orientation and orbit of the images (Karas et al., 2021).”

In reference to other papers using DEM and GIS technics, we have added the following:

- Lines 67-72: “In addition to optical, radar or Lidar data, other studies used Digital Elevation Models (DEMs) and topographic parameters to determine the influence of terrain on avalanches in Switzerland (Maggioni and Gruber, 2009). Other studies incorporated other parameters such as morphology and vegetation to define potential avalanche zones and ran the Avalanche Flow and Run-out Algorithm to automatically detect potentially affected regions by avalanches (Barbolini et al., 2011). Moreover, the combination of snow measurements (depth) and high resolution DEMs have proved useful in snow hazard detection (Bühler et al., 2018a).”

For the optical state of the art, we have added the following:

- Lines 97-99: “Across a wide area (12,500 km^2) individual snow avalanches were manually digitized using high resolution SPOT-6 images (Bühler et al., 2019).

Moreover, we have added some of these references in Table 4 (line 536) when we compare SAFE accuracy with accuracies of other methods. We thank you for those references that enhance our paper.

2. It is essential to clearly communicate what can be expected from the presented approach in terms of accuracy and reliability. First of all, only very large avalanche debris can be mapped. Throughout the paper the authors should use this term and not the term avalanche to be clear. An avalanche consists of a release, a transition and a deposition zone. Only the deposition zone

can be partially mapped. There are several problems for example if the avalanche debris is covered by soil / rock or wood (The NDSI is reduced and the deposit is not mapped as avalanche). There is now information on how many avalanches deposited onto one mapped deposit. Typically, this happens several times a year. In the river basins there is often complex terrain with a lot of cast shadow leading to missed avalanche debris. All these uncertainties lead to a very limited reliability of the presented approach. Therefore, it is not eligible to draw all the statistics from the mapped debris as the authors do in the results. These statistics are strongly biased and not reliable. Applying them for hazard mapping or the planning of mitigation measures could be very dangerous.

We thank you for your comment related to the terminology used in our paper. We have modified snow avalanches to avalanche depositional zones or deposits throughout the manuscript (in red in the text). Moreover, based on your comment, we have decided to change the title as follows: “Snow Avalanche Frequency Estimation (SAFE): 32 years of monitoring remote avalanche depositional zones in Afghanistan”. We believe that this title better reflects what SAFE does.

As for avalanche structure, it was already explicitly articulated that SAFE only maps deposition zones: “These zones are indeed detectable by delineating the depositional zones of the avalanches (not their release or transition zones);” lines 175-176.

As for snow avalanche deposits that may have been missed using NDSI, we agree that SAFE can omit snow avalanche deposits as already acknowledged in line 268: “Another source of error arises when SAFE cannot detect avalanches depositional zones due to a dark color on the snow surface associated with surface debris or a debris flow on top of the avalanche.” However, because of the advantage of our long-term data base, if SAFE misses an event in one year, the model systematically looks at each pixel in every year – in our case, 32 times (i.e., 32 years of data). Thus, frequently impacted areas will be identified even if events in a few years are missed due to shadows or debris flows. Thus, as we answered in our previous response to this review, we disagree that our statistics are ‘not reliable’.

Regarding the comment that the application of this model ‘could be very dangerous’, as we wrote in the first response to your comment, we completely disagree with this value judgement. SAFE is one model, if not the first attempt, to map areas at avalanche risk across large scales and on over long time periods in this remote, high mountain region. As specifically mentioned in the paper, we looked at avalanche deposits on foot slopes where human settlements are most vulnerable to snow avalanches. SAFE does not examine mid- to upper-slope terrain because in mountainous Central Asia, as well as proximate mountain regions, these upper slope areas are not occupied by humans or their activities during winter. Only foot slopes represent an area at risk; high mountain winter recreational activities – e.g., skiing, and other winter tourism activities – are virtually non-existent in this vast mountain region. What matters in our region, such as Badakhshan, is the location and frequency of avalanche deposits on villages, roads and, to some extent, streams. SAFE represents a very needed model in this region because it can be freely replicated and used with minimal resources to determine which villages and roads are at frequent risk. This is the advantage of our long-term (albeit less accurate than SPOT) database. In our work here on the border of Afghanistan, we frequently experience road blockages and the isolation of villages for several days because of avalanche deposits. This issue is a higher priority in Central Asia and the surrounding mountain regions than detailed mapping of avalanches on upper slopes.

3. To assess the mentioned uncertainties and potential biases we recommend to test the algorithm with the most complete and accurate avalanche dataset mapped with SPOT6 imagery over the swiss Alps in 2018 and 2019 (Bühler et al., 2019; Hafner and Bühler, 2019) https://www.envidat.ch/dataset/spot6-avalanche-outlines-24-january-2018;https://www.envidat.ch/dataset/spot6-avalanche-outlines-16-january-2019. This exercise could bring clarity into very important questions and help to assess the potential of the presented approach.

First of all, the authors would like to thank the reviewer for allowing us access to this very interesting dataset in Switzerland. Hence, we have conducted the comparison between the outputs of SAFE and SPOT-6 outlining method. A totally new paragraph, plus a table, have been added to the manuscript. Based on the comment by Reviewer 2, we have decided to move the validation sections (including SAFE/SPOT comparison) to the beginning of the Results section - sub-sections 3.1 and 3.2. While we believe that the two datasets (SAFE/SPOT) are quite different as explained in the following paragraph, this comparison provides some interesting results. The comparison actually helped us to be more specific about the best use of SAFE. SAFE is more applicable in high mountains (Himalaya, Tien Shen, Hindu Kush, Karakoram, Andes…) where avalanches depositional zones remain longer compared to lower elevation mountains. And, interestingly, these mountain ranges are mostly within impoverished regions. Hence, from the title new version in comment 2, we have modified again the paper title as follows: "Snow Avalanche Frequency Estimation (SAFE): 32 years of monitoring remote avalanche depositional zones in high mountains of Afghanistan". Related to this, we have added two sentences, one in the Introduction and one in the Conclusion, noting that SAFE is more applicable for high mountain regions:

- lines 113-115 "SAFE is applicable in any high mountains of the world, such as Tien Shen, Himalaya, Hindu Kush, Karakoram or Andes, but not restricted to these, where snow avalanches deposits can be detected every year by satellite images for a long time before completely melting."
- and lines 553-555 "Moreover, the application of SAFE in Afghanistan, compared to its application in Switzerland, showed that the script can be applied worldwide, especially in high mountains (above 4000 m) since deposit zones are still detectable in late spring at those elevations."

Moreover, this comparison more precisely elucidates the goal of SAFE, which is mapping avalanche deposits, rather than the continuum from initiation to runout.

The results of the comparison between SAFE/SPOT are presented as follows lines 304-341:

"3.2 SAFE outputs compared with outlined avalanches using SPOT-6 images

As a potential method of strengthening our testing of SAFE, outputs of our model were compared with a method that applied a more precise and expensive remote sensing product in Switzerland in 2018 (Bühler et al., 2019). The Swiss area encompassed 12,500 km² where more than 18,000 snow avalanches were manually digitized using very high-resolution products SPOT-6 images (in January 2018). While our dataset is quite different from the Swiss data, the objective of this comparison was to assess how many snow avalanche deposits SAFE could detect compared to the approach using SPOT-6 (Table 3).

Table 3. Comparison of snow avalanches deposits zones between SAFE outputs (April to June 2018) and manual digitization using high-resolution SPOT-6 images in Switzerland in January 2018*

Method	Number of snow polygons	Area of snow polygons (m²)
SPOT digitization	7574	362,187,741
SAFE detection	9948	494,454,599
Overlapping SPOT-SAFE	2194	223,907,868

*SPOT data based on (Bühler et al., 2019)

Importantly, not all avalanches manually digitized on SPOT-6 images were comparable to SAFE results. To make this comparison more consistent, we clipped the outlined avalanches with the valley bottom mask used in SAFE. Following this modification, the SPOT-6 digitization process identified 7574 avalanches deposits in valley bottoms compared with 9948 by SAFE. Overlapping these two datasets, we found that both approaches detected 2194 deposit zones in common. Much of this discrepancy is due to the timing of SAFE images, which examine deposits that remain in late spring and early summer, whereas SPOT digitization covered only January. The larger number of snow deposits detected by SAFE occur during late season snow avalanches that impact valleys. This suggests that SAFE could not detect all January snow deposits because many of those already melted by the time of SAFE detection (early April to late June in the Swiss case). In addition, optical image quality strongly depends on cloud cover that may cause avalanches to be obstructed. For instance, we could not compare the 2019 SPOT-6 derived dataset in eastern Switzerland (Hafner et al., 2021) due to cloudy images at the end of winter and early spring because these snow avalanches had already melted, implying that SAFE is more suitable for high mountain areas (>4000 m) where snow deposits remain longer in valleys, thus inflicting greater damages and obstructions. Using LANDSAT images, SAFE somewhat circumvents this problem of cloud cover by assessing many years of data (in our case 32 y). However, SAFE does not distinguish individual events and considers overlapping snow deposits as one, in contrast to SPOT-6 which distinguishes these as discrete events. This, in addition to the different methods and spatial resolution difference between SAFE and SPOT, explains the somewhat low number of overlapping snow deposits between SAFE and SPOT. Moreover, the SPOT digitization procedure found a total avalanche area of 362,187,741 m² in January, while SAFE detected 494,454,599 m² of deposits at the end of the avalanche season, including 223,907,868 m² in common. The area detected by SAFE is naturally larger than SPOT-6 since SAFE maps all detectable deposits at the end of the winter. Moreover, SAFE did not detect the small avalanches of January that rapidly melted after they occurred. The polygons extracted by SAFE using Landsat images are obviously coarser than those outlined with SPOT-6 images, which partly explains the low number of overlapping snow deposit zones, but a much more comparable detected area (62%) between the two methods. Much of the discrepancy is related to SAFE's inability to detect individual events and missing deposits that rapidly melt (mostly from the early winter snow avalanches), as well as the very different resolution of these products."

4. The snow avalanche size classification is totally flawed with respect to reality/ methodology. According to the definition of the EAWS (https://www.avalanches.org/standards/avalanche-size/#largeavalanche) size is mostly defined by volume, runout-length and destruction potential: so basically only avalanches larger than size 3 (large to extremely large) have potential to even reach those places where they are later detected with enough snow for it to remain until summer, Additionally, as the authors state they cannot separate single events, a size classification with the same classes as assigned to whole avalanches shortly after their release is nonsensical also as the area covered in gullies usually means a lot more volume than one would think. This makes methodologically no sense as well as everything derived from this (whether as category or as size).

We understand and appreciate your comment about avalanche size classification. Indeed, we unfortunately do not have any data on avalanche volume, and we admit that our terminology 'avalanches size' is confusing. We however believe that the classification of the cumulative avalanche deposits could be relevant information to highlight and rank the most vulnerable areas. Valleys with large 'size' events represent more vulnerable areas impacted by repetitive avalanches than 'small size' events. And if one valley bottom is affected by only one, but a very large avalanche, SAFE will still be able to identify it as a single large event, since SAFE maps the cumulated avalanche deposits. In addition, we noticed that the avalanches mapped using SPOT-6 images in Switzerland (https://www.envidat.ch/dataset/spot6-avalanche-outlines-24-january-2018;https://www.envidat.ch/dataset/spot6-avalanche-outlines-16-january-2019) were actually also classified by size (m²), so the point about the 'flawed classification' raised by the reviewer is confusing and somewhat contradictory compared to this Swiss data. To address this issue, we have changed 'size' to 'surface area' in the manuscript wherever needed (text and Figure 11).

5. It is not clear why only Landsat is used. Sentinel-2 imagery would also be a big help for the presented approach (even though only available from 2015). What about the potential of other systems such as PLANET? This should be discussed.

As noted in our initial response to this comment, we agree that it would indeed be interesting to run SAFE with other products such as S2 or other products from PLANET. This could be an area to explore in a future paper. However, to address the objective of this study (i.e., a long-term assessment of hazards in valley bottoms related to avalanches), SAFE uses Landsat archives for two obvious reasons: (1) these data are open access, which suits the economic context of local research institutes and practitioner in this greater region who cannot afford expensive images such as SPOT; and (2) as noted, the objective of this paper was to look at a long-term – 32 years – period of avalanches deposits to assess frequencies of potential damages, and only Landsat archives can achieve this.