Reply to reviewer 2

We thank the reviewer for the constructive comments. Below we reply in detail.

1.) I'd like the author to better address the possible technical limitations of their methods, in particular the field deployment and the near-real-time application of the classification methods based on two stages.
We adapted the Discussions section and added a more detailed description of the limitations of our method (P. 22 L.5).

Although we were able to identify one major avalanche activity period in the winter season 2016-2017, the method presented here has its limitations. Based on the sensors used for the automatic monitoring, we identified avalanches within a range of 2 – 3 km. However, by using more sensitive sensors, e.g. seismological broadband stations, the detection range of avalanches can be increased, even up to 30 km for very large avalanches (run-out distance > 2 km) (Hammer et al. 2017). However, it is difficult to deploy such sensors in mountains terrain, since these stations require existing infrastructure (e.g. electricity, storage room in a hut), which is typically not available at remote locations. In addition, the last post-processing step requires a second array. Hence low energy systems with less sensitivity proved to be the best solution. Furthermore, the limited power supply at the field sites also prevents performing first processing steps directly at the field sites and hence limits the possibility of near real-time analysis. However, it is possible to overcome this problem by designing special hardware for this particular task.

2.) The network geometry has a strong impact on the success rate of this latter criterion, could you add some details on that?
Unfortunately, the resolution of our seismic array was limited due to the instrumentation used. Lacroix et. al 2012, e.g., used a seismic array with larger distances between the sensors resulting in a better resolution for seismic waves. They could use beam-forming methods to calculate the source direction. We assume that with a larger interstation distance the resolution can be improved. With a larger array, a new comparison between beam-forming and MUSIC would be required.

3.) I have the impression that this second stage can be surely useful to recognize earthquakes but it probably needs a calibration for anthropic sources.
It depends on the type of anthropogenic source. Helicopters, for example, have a very characteristic spectrogram and Doppler effects can be observed at all times. For airplanes, this is a little bit more complicated. Due to the different types of airplanes (turbine or propeller) and the different flight altitudes, signals vary. But most of the time, a dominant frequency is visible as well as Doppler effects. However, the airplanes, which were considered as avalanches by our algorithm, did not have a dominant frequency. We expect that these airplanes were flying at a larger distance to the seismic array and due to the topography signals were naturally filtered.
We gained this knowledge over the past years in a long learning step. We agree that calibrations for anthropogenic sources may provide more reliable classification results.

4.) In addition, technical limitations in such extreme environments like high Alpine areas (e.g., data transmission) can be a possible trivial but concrete limitation for a real time application.

That is correct. The biggest problem we have is the computational power at the field sites. At the moment, we rely on Raspberry Pis which are mainly used for data storage and data transfer purposes. Data are transmitted and then mainly processed at our institute. A first improvement would be, to establish a fast and stable wireless link to the arrays to provide good data transfer. Second, better hardware with a low energy consumption would also be a great advantage. However, hardware capable to perform our calculations near real-time cannot run solely on solar power and batteries.

5.) The application of the proposed methodology on another dataset gathered on another test site would be of great interest for the reader. For instance, is it possible to run the methods the other way round, testing it on the other array currently used for the second classification step?
Of course it would be possible to perform the classification task the other way round. However, due to some technical issues, we only recorded with two sensors at the second field site. As we showed in a previous study (Heck et. al 2018a), best classification results with the HMMs are obtained by only considering events classified as an avalanche by at least five sensors. We could perform the classification with only two sensors, however, we expect too many false alarms. Furthermore, it is impossible to determine the source direction of the signals based on only two vertical geophones. Hence, the localization step could not be used to confirm or neglect detected events as avalanches.

6.) Visual observations are used as validation, could the authors add some information about that? Which are the observation sources? How is compiled the avalanche catalog? If available, an image of one reference event could be also useful to show the test site.
We have installed several automatic cameras at the field sites, still, we could only determine the exact release time for two avalanches. For the remaining events, especially for the main avalanche activity period in March, we narrowed the release times down to a 24-hour window. This was due to the fact, that most avalanches released during snow storms or at night when the visibility was bad.
In addition to the cameras, we relied on the avalanche data base compiled by the avalanche warning service in Davos. They monitor the avalanche activity for the region of Davos (~ 175 km^2) and also use information from voluntary observers.
We included a picture shortly after the avalanche period in March taken during a field survey.

7.) Figure 2, it would be useful to add a map with terrain information (slope, morphology, etc).
Since the interstation distance between the sensors is two small, it is nearly impossible to show additional terrain information in these particular figures. Hence we added two additional figures, each showing additional information of the field sites also including the location of all instrumentation, at a lower scale (P.6).