

Dear editors,
Dear reviewers,

We highly appreciate that the referees have taken the time to review the revised manuscript once again. Their reviews and suggestions have contributed immensely to improving the manuscript and the content. For this, we want to thank them sincerely.

Kind regards,
Andri Simeon

Table of Contents

REFEREE 3#:	2
GENERAL COMMENTS	2
REFEREE 4#:	3
GENERAL COMMENTS	3
SPECIFIC COMMENTS.....	3
REFEREE 6#:	5
GENERAL COMMENTS	5

Referee 3#:

General Comments

I found the manuscript easier to read than the previous version (but probably because I already spent a lot of time on the previous version).

Most questions have been addressed, thanks.

But it is a bit frustrating having to wait for "future work" to test several ways to improve the method (including network and polarisation features, infrasound sensor, longer time windows...).

Thank you very much for your helpful feedback and for taking the time to review our manuscript carefully. We truly appreciate your suggestions regarding possible improvements to the method. At this stage, however, the paper is finalized and already large. We agree that exploring aspects such as network and polarisation features, infrasound sensors, and longer time windows would be highly valuable, and we will address these promising directions for future research.

Just one new question. Did you only apply your model to time periods when experts manually picked and classified events, or did you (ou could you?) apply your model on the full dataset, including summer periods? This could be interesting to further estimate the rate of false avalanche detections (eg, when there is no snow) and to compare the rate of avalanche activity with the estimated avalanche hazard.

Yes, so far we have only tested the models on the manually picked events. A large-scale transfer of these models to continuous data will certainly be considered for future work. However, we expect this transfer to be non-trivial and will likely involve further model development and tuning.

Due to restrictions from landowners, the sensor array can only be deployed during winter. Therefore, the dataset does not contain days in summer.

Can you explain in the zenodo archive what is shown in the catalog labels.csv: what is the label? av_score? eq_score?

We added a description to the repository to explain the columns in the labels.csv file.

By curiosity, is the seismic data also available from the FDSN?

Currently, it is only available on the Zenodo repository. Since the sensors are not running in summer, and the test site is not permanent and bound to the project duration we did not consider FDSN.

Referee 4#:

General Comments

This manuscript develops autoencoder derived seismic attributes and engineered seismic attributes as features in a random forest classification detection of snow avalanches. The results suggest that the autoencoder derived attributes perform as well as the engineered seismic attributes for event detection. The avalanche detection method is reported to be potentially used as an operational, near real-time avalanche detection system, though the relatively high number of false alarms requires further improvement. The work is presented with respect to the previous studies employing machine learning for seismic event detection while highlighting their significant and novel contribution of unsupervised feature extraction. I found the paper to be informative and complete in analysis and have been satisfied by the authors' responses to the initial review which improved the methodological development and comprehensibility of the work.

We are very grateful for your careful review and are happy that we followed the previous reviews satisfactorily.

Specific Comments

Line 6: "Therefore, we compiled a dataset of seismograms recorded with an array of five seismometers..." This sentence seems a bit disconnected from the main ideas presented. Is this statement intended to link back to "Monitoring snow avalanche activity is essential for operational avalanche forecasting..." or "Still, automatically distinguishing avalanche signals from other sources in seismic data remains challenging." I think the intent could be clarified by replacing "Therefore" with a descriptive intro to sentence like, "Because of the inherent complexity of interpreting signals travelling within the subsurface, we utilised an array of five seismometers..." This example expresses the importance of having an array of seismometers.

This sentence was intended to refer back to the challenge of distinguishing an avalanche from a noise signal, for which we needed a comprehensive dataset of avalanche and noise samples. Following your suggestion and to make it clearer, we changed the sentence to: «To study and interpret the variety of these signals, ...»

Line 104 – 106: Feels like a run-on sentence. Consider the revision, "Additionally, the site was equipped with a Doppler radar and three automatic cameras to obtain independent validation data, including accurate release times and information on the type and size of avalanches, provided favorable weather conditions."

Thank you for the suggestion. We changed this sentence accordingly.

Line 109- 110: "The cameras automatically photographed all surrounding slopes every 30 minutes (Fig. 1)." Consider including one sentence detailing how the photographs were utilized. Manually inspected as a corroboratory inspection of radar or other data

source detections, or automatically reviewed as an independent method? This is mentioned on lines 113-114, but not explicitly.

To clarify this, we added: «..., which we manually inspected to identify days with avalanche activity and verify avalanche events of the detection systems.»

Line 132: What is exactly meant by ground velocity? The derivative of the seismic displacement? Just a bit more detail on this method would be helpful.

We changed this sentence to: «..., we transformed the units of the raw recordings, i.e. counts, to meters per second (ground motion)». Seismometers detect ground vibrations and convert them into an electrical signal. The data logger then digitizes this signal and stores it as counts. To convert raw counts into physical units, the recorded counts are divided by the seismometer's sensitivity factor, yielding ground velocity in meters per second.

Line 232: "As activation function -> As an activation function"

Line 282: "Therefore, we used the three train folds -> Therefore, we used the three training folds"

Line 306: Similarly "train" -> "training"

Line 307: "weigh" -> "weight"

Line 465 "Tough" -> "Though"

We adapted all of the above five suggestions accordingly.

One thing to note: Between the Engineered Feature and AE feature avalanche detection, which had comparable recall values, were the same avalanches detected? Could the implementation of both SAE and engineered feature detection further increase the detection capability. For if they detected the same amount of avalanches, but different ones, perhaps this increases the overall detection. This analysis is explicitly missing, but could provide additional insights to the differences of the detection methods. Perhaps this is something you have already investigated, but did not note in the manuscript. Figure 13 touches on this conceptually, but it is still hard to discern if the false positive detections stem from different events or not.

This is indeed an interesting question. Unfortunately, we observe that the presented models strongly agree on the separate avalanches, meaning they detect and miss the same avalanches. Therefore, combining the models would lead to similar or even worse results. The reason for the agreement of the models is instead found in the respective signals. We qualitatively observed that true detections stem from strong avalanche signals, i.e. relatively large avalanches or avalanches flowing in the proximity of the sensor array. In contrast, weaker avalanche signals are often missed by all methods, which is what we had expected.

With Regards,
Tate Meehan

Referee 6#:

General Comments

I am accepting the manuscript "as is". The updated manuscript addresses all my concerns brought forward in my last review. Specifically, my ranking of the scientific quality and presentation quality has increased significantly. The scientific significance and reproducibility were already present during the last iteration (and still are).

I want to highlight the great effort the authors made to rework a large part of the manuscript, fixing code issues, updating figures, and even rerunning the experiments on a larger scale. All major and smaller suggestions have been taken into account and have led to either an adaptation of the manuscript or have been backed up by the authors with detailed explanations. Algorithm design choices were taken carefully, and are now communicated more clearly to the audience.

I thank the authors for their patience with the review process and for contributing their work to the scientific community.

We want to thank you for the careful past review. It has helped us immensely in improving the manuscript. We are happy that we could meet the previous concerns.