Title: Towards the systematic reconnaissance of seismic signals from glaciers and ice sheets - Part A: Event detection for cryoseismology
Author(s): Rebecca B. Latto et al.
MS No.: egusphere-2023-1340
MS type: Research article

Please access the discussion at:
https://egusphere.copernicus.org/preprints/2023/egusphere-2023-1340/#discussion

---

Below, the R1 comments are copied in grey. Author Comments continue in blue.

The authors present an exciting study that describes a “recognisance” algorithm for detecting seismicity from a range of seismic signals. The method is novel and the concept of a recognisance algorithm that is sensitive to a range of seismic signals will likely be of much use to the community.

Many thanks for this very positive appraisal – we confirm that the emphasis in this work is capturing the range of seismic signals (Part A) in a reconnaissance of the seismic events (and likely following analysis, as in Part B), and event-like noise that is present in a glacier environment. (A minor confirmation, we carry out ‘reconnaissance’ information capture, whereas ‘recognisance' may have been inserted by a spell-checker, and is a different/legal term).

However, although I would like to see this work published, I think the wording of the manuscript needs some work before it is ready to accept for publication. Firstly, the novelty of the algorithm needs to be clarified. This is probably a minor point, but at the moment the method appears to already be published (Turner et al., 2021), but the authors do not make this explicitly clear throughout the paper. I’d like to see the original methods paper properly acknowledged where the method is first introduced, and the tone of the paper changed to reflect that the method is applied to cryoseismology here rather than introduced as a new method.

We’re happy to add a clarification as follows: The Turner et al., 2021 citation refers to a (properly documented, code-reviewed, software library) handling framework for waveforms as a pipeline for automated analysis (i.e. machine learning). It was developed by the same research team, and the full reference is here:

Turner, R.J., Latto, R.B. and Reading, A.M., 2021. An ObsPy Library for Event Detection and Seismic Attribute Calculation: Preparing Waveforms for Automated Analysis. *Journal of Open Research Software*, 9(1), p.29.DOI: https://doi.org/10.5334/jors.365

This software pipeline has multiple options for the event detection algorithm (including the standard STA/LTA). The current submission ‘Part A’ is, accordingly, the correct reference for the ‘multi-STA/LTA’ algorithm as a novel approach, and we confirm that it was designed primarily for cryoseismology (although it could well be useful for other environmental seismology studies also). Going forward, other users could also select the ‘multi-‘ algorithm (for example) if they use the Turner et al. 2021 software pipeline.

Secondly, I remain to be convinced by the concept that such a deliberately broad method can outperform a more specific method suited to one task (e.g. basal icequake detection). Perhaps it can, but I see no evidence in the paper to back up the claim made in the conclusions that the method presented can provide a “near-comprehensive event catalogue”.

We agree that we need to define what we mean by 'near-comprehensive' (wide range of event and event-like noise types). Should a research task be focussed on a specific event type, another algorithm may well be more appropriate as this would avoid having to handle other event types in the same catalogue. The currently presented algorithm aims to capture a wide range of seismic signals, prior to semi-automated ongoing analysis. In fact, basal stick-slip events -are- very well captured by the multi-STA/LTA algorithm and we are happy to add a comment on this point.

To conclude, I think this work will be a valuable contribution to the field and I don't find any issues with the results themselves. However, the text needs to be somewhat revised to tone down the claims made. Assuming the authors are happy to do this, then I would be very happy to see this paper published.
We apologise if we seemed to be overstating any point. Hopefully the amendments (as per above and comments below) provide suitable background re: the novelty and scope of the current study, and intended future contribution of the reconnaissance workflow, in the context of other options).

Comments:

The introduction does seem to be written in rather a bold way (e.g. "Since STA/LTA and correlation-type algorithms have enjoyed only limited success when applied to environmental seismology"). Both STA-LTA and cross-correlation methods are effective when used in certain ways, as the authors elude to and indeed that is the premise of this paper. I'd suggest toning down the limitations of these algorithms a little, since actually there has been much success in using these algorithms carefully, as part of broader methods. Indeed, I cannot think of any passive seismology method (other than manual detection) that is not at least somewhat built upon one of these two foundations. I think it would be useful to outline the scope of the study from the very beginning of the introduction, to clearly clarify the scope of the work to the reader before making the perhaps bolder claims.
We are happy to reword as suggested. This will be partly covered in response to Reviewer 1 (see our general statement in response to that reviewer).

From the introductory text, I was expecting Section 2.2 to be its own section. However, I can see the logic behind having it as a subsection of the overall algorithm method. Regardless of the structure, I think it would be useful to provide more details on the analysis of the algorithm testing results presented in Fig. 2,3. There is a lot of information held in those two figures, but they are not yet adequately described in the text. In other words, I was excited to read that the algorithm would first be tested on synthetic data, but was then left a little disappointed that the key findings are buried in the figures without much explanation of what they show.
This information is included in the journal and electronic supplements to avoid the main text. We are happy to provide better pointers to this material in Section S2 in the Supplement.

L183-187: Perhaps it is common to only use the vertical component. However, best practice for any body wave data should be to use both vertical and horizontal components. However, mixing the three components via the Eucllidean norm might cause one to loose information about whether a phase is a P or S phase. While information loss is not inherently a problem for detecting events, incorrectly identifying S-wave phase arrivals as P-waves would result in false detections, since they are likely from the same event. I'm not suggesting the authors should revisit this component of their method, but I think they should be clear about the limitations, especially the possibility of false triggering. I'd suggest it would be beneficial to also discuss somewhere how their method might be developed further to act on the vertical and horizontal phases separately. If the authors would like any pointers to literature describing how to use

the vertical and horizontal component information for P and S wave association in detail, including how a firn layer can affect such results, then this paper and references therein provide further details (https://doi.org/10.5194/egusphere-2023-657). (No requirement to cite this work, just a potentially useful paper that covers the point raised).
This is a good point, and we are happy to add a few sentences to the discussion section in this regard. We're fine with citing the suggested paper (at the time of writing, it is in the final stages of review), and might perhaps also add a further citation that certainly supports being able to detect/work with S waves (https://doi.org/10.1093/gji/ggw150), and also following work with the same lead author (not included in the previous suggested reference list).

L189-190: Some clarification on the novelty of the method is required. Until this point in the paper, I was under the impression that the authors were presenting a new algorithm for event detection. However, from a glance at Turner et al. (2021), it looks like the value of this work is more in showing how the cited algorithm is implemented? Could this be clarified, and if the implementation is originally from Turner et al. (2021), then that paper be clearly referred to in Section 2.
See above clarification re: the Turner et al. (2021) software library. The first impression is correct.

L200-202: Perhaps too technical a question, but more for my interest: Are the signals instrument-response corrected before being passed through the algorithm? If not, then "energy" might be better referred to as "an approximation of the energy", since different frequencies might exhibit different amplitude responses dictated by the instrument transfer function (response). Normally I wouldn't raise this point, but since the authors are attempting to describe as broad an event detection algorithm as possible, frequency response could become important in certain instances. Maybe at least worth making readers aware of this point.
We are happy to add a comment as suggested.

Figure 5: It would be nice to see more detail in each of the time-series. I cannot dicern any differences from the plots in this figure. Maybe for each, the authors could include an inset figure zooming in on any differences in the first arrival, perhaps plotting all three signals over one another? Otherwise I'm left questioning what difference the multi-STA/LTA algorithm makes. In summary, I imagine it is better, but the figure does not currently communicate that.
We intend that the reader will focus on the difference between the purple, light orange and light blue overlays, which are clearly different (assuming the .pdf is showing OK on the screen or printout). The waveforms should be the same.

L276-277: Great aim. However, this sentence is very much left hanging. I'd like to see some text introducing how the authors will justify how the work presented here meets that aim in the remaining discussion.
This should be better supported following the clarification made in response to Reviewer 1 (general statement). Also, we're happy to expand on this point in the 'Applications' section.

Section 4.1: The biggest limitation of this work to meet its aim is that such a broad search algorithm does likely not perform as well for certain event detection scenarios, compared to more specific methods. For example, a coalescence/stacking based migration algorithm will, by design, be better for filtering out false triggers caused by noise. Therefore, taking basal icequakes as an example, the authors method is unlikely to outperform a migration method. This does not detract from the method presented by the authors, since a general recognisance method is definitely very useful in many contexts. However, it is definitely a limitation that needs to be mentioned, because it really does present a barrier for "a consistent approach to the generation … of event catalogues" (L276-277).

Again, this should be better supported following clarification made in response to Reviewer 1 (general statement). They key point is whether the event detection intends to capture a wide range of event or event-like noise types (as in the current study), or whether the nature of the event in question is previously known.

L411-412: This statement is definitely too bold. I am not convinced that one can ever develop algorithms that produce “near-comprehensive” event catalogues. However, more specifically, a general recognisance algorithm, which this is presented as, is unlikely to outperform a specific algorithm for a specific purpose. The algorithm presented here is definitely a valuable contribution to the field, but I see no evidence in the manuscript that it can produce “near-comprehensive” catalogues of seismicity.
We are happy to amend the text to be consistent with previously addressed comments, as in the above point, and also following a careful definition of ‘near-comprehensive’ (wide range of event types), or a wording change.

Minor comments:

There should be spaces between units (e.g ma^-1 should be m a^-1, otherwise it is technically milli years).
Corrected accordingly.

L19: “exceptional” – rather emotive language. Consider removing.
Happy to remove this word.

L23-24: I think the definitions of the two event detection types are a little narrow. I’d view STA/LTA algorithms as just one subclass of any algorithm that searches for a peak in energy with particular frequency content (related to the length of the STA window). These can then be used to detect phase-arrivals, or be used in combination with more sophisticated phase associators, or coalescence-based algorithms to improve detection. Perhaps worth somehow very briefly mentioning this (as an STA/LTA algorithm on its own is not very useful/causes lots of false triggers).
Happy to follow this suggestion.

L23-24: Further to the point above, one could consider array-based methods as a common earthquake detection type too? Not sure if this is worth mentioning, but picking phase arrivals on individual channels is only one class of earthquake detection. If the authors do decide to mention this, then examples of some relevant papers are included in the introduction to this paper: https://doi.org/10.5194/egusphere-2023-657. (sorry to refer to this paper again – just easiest way to point to the relevant literature cited within it).
We’re very familiar with array analysis, and are happy to include the suggested reference, and also other earlier relevant citations not in the suggested reference in support of this point more generally, e.g. https://doi.org/10.1029/2018JB015526 and book chapter ‘Beamforming and polarization analysis’, Gal and Reading (2021) in Seismic Ambient Noise (Eds Nakata, Gualtieri and Fichtner).

L43-44: I’d say that cross-correlation algorithms are inherently also similarly prone to missed detections since they are typically based on the assumption that similar events occur within a catalogue.
Happy to amend wording accordingly.

L72: Is QuakeMigrate a spectral-based method?
Apologies, this occurred inadvertently, when making some of the background text more succinct.

QuakeMigrate uses migration based techniques that use coherency and waveform stacking in detection. We'll fix the wording.

L122: "event" – would it be perhaps better to refer to it as an "event phase arrival"?
Happy to amend wording accordingly.

Figure 1: Isnt the data in Figure 1b also plotted on Figure 1c? Not sure this section is therefore required. Similarly, Figure 1a could also be removed as all the information is contained in (d).
We were asked to include this figure by the Editor, as a step by step illustration of the STA/LTA concept. Hence, we prefer to keep the separate figures. We're happy to follow further editor guidance, perhaps combining panels a and b as a compromise.

L170: The minimum distance rather than the maximum distance is probably the more relevant number.
Happy to amend word choice from maximum to "up to 600 km" (L170-171) to follow the exact statement in the quoted reference (Wiens et al., 2016).

L199: That reference is not to software documentation, but to a paper describing the software. Pointing to any software should really be done via a software repository DOI in the Acknowledgements. However, I think the authors are actually refering to the paper here, and so should simply remove the word "documentation".
Happy to amend wording accordingly.

L236-237: I think this sentence doesn't make sense with the word "event" at the end. Apologies if I misread, but would it be possible to reword if indeed it doesn't make sense?
Happy to amend wording accordingly.

Figure 5: Add subplot labels, then refer to accordingly in the text (rather than left, right etc). That way the UTC time stamps can be remove from the text. Also, the caption is too long. Consider shortening.
We are happy to follow this suggestion.