Earth **Surface**
Dynamics
Discussions

Open Access

EGU

Interactive comment on "Systematic Identification of External Influences in Multi-Year Micro-Seismic Recordings Using Convolutional Neural Networks" *by* Matthias Meyer et al.

Matthias Meyer et al.

matthias.meyer@tik.ee.ethz.ch

Received and published: 16 November 2018

article [utf8]inputenc

RC: Referee Comment
AC: Author Comment

1   Response to RC1

AC: We would like to thank the anonymous reviewer for the extensive review and the valuable feedback. We will incorporate the feedback and address all comments in the following. For a preliminary revised manuscript highlighting all the modifications please refer to the response to the editor.

RC: This paper addresses the issue of accurate attribution of seismic events to the correct source in long-term/large (micro-)seismic datasets. This paper has the potential to form a helpful methodological contribution to the geomorphic literature, and the overall result is promising. However, I do not believe the paper is ready for publication in its current format. Whilst there is some interesting information presented here, the focus, clarity and structure of the paper require further work.

AC: We restructured the manuscript to make it more precise in terms of terminology and the geoscientific field we consider. Moreover, we have condensed the information to make it more accessible.

RC: General points The language is often vague, with loose use of specific terminology. For example, in the abstract, the authors mention that 'Successful analysis depends strongly on the capability to cope with such external influences'. What do they mean by 'successful analysis' and 'coping' with these influences?

AC: We acknowledge the loose use of specific terminology in the initial submission. In the revised manuscript we use a more precise terminology and avoided misleading formulations, such as 'successful analysis' mentioned in your comment.

RC: Similarly, the authors mention 'correct slope characterisation' in the next sentence. What does this mean? It suggests consideration of the structural/strength/geometric properties and/or damage condition of the slopes. It is not clear which the authors are addressing, and why. Linked to this, Fig. 5 suggests that the focus of the paper is on rockfalls, which again is different to 'slope characterisation'. In short, what is the geomorphic nature of the seismic activity the authors are considering?
RC: Links to geomorphic processes are implicit at best, and largely absent. For example, what exactly are you trying to monitor? Rockfall occurrence? Ground cracking and associated micro-seismic signal? This isn't clear. There is also a stark lack of reference to appropriate literature (e.g. page 2, lines 10 – 16).

AC: In contrast to the submitted manuscript, we have precisely defined the application context in the revised manuscript. The context is geophysical analysis on micro-seismic signals in general and event-based analysis in particular. The specific analysis performed using event-based methods depends on the use case and is not the focus of our study. For example the case study we use to demonstrate our method focuses on rock fracturing. We have also added additional literature to define the different geophysical applications and state more precisely which application we focus on: (Hardy, 2003), (Michlmayr et al., 2012), (Gischig et al., 2015), (Burjánek et al., 2012), (Weber et al., 2018).

RC: The final sentence in the abstract is also rather obvious and can be made without the detailed assessment presented in the paper. Indeed, this type of source characterization is commonly done (and done well) by geomorphologists (see e.g. the work of Adam Young on coastal microseismic monitoring). The most interesting part here is the ability to distinguish between sources of microseismic activity in large/long-term monitoring datasets, and this needs to be more clearly presented.

AC: You are correct that the final sentence "Due to these findings we argue that a systematic identification of external influences, like presented in this paper, is a prerequisite for a qualitative analysis." is rather obvious. Indeed, we want to show in this paper how the source characterization can be done systematically for a large and long-term monitoring experiment. Consequently, we have updated the statement to be more precise: "Due to these findings we argue that a systematic identification of external influences using a semi-automated approach and machine learning techniques as presented in this paper is a prerequisite for the qualitative and quantitative analysis of long-term monitoring experiments."

RC: The Introduction repeats the same points multiple times in subtly different ways – this section could be condensed considerably.
RC: Section 2 again repeats much of what we have already been told in the introduction.
RC: Sections 2.4.2 –3.4 contain some ostensibly important methodological steps, but again much of these sections feels descriptive, lacks an appropriate justification and a logical structure to follow the workflow and the choices made.
RC: The aim of the paper is not clear and the authors present instead a bullet-point list

of study conclusions. What is the focus here and what is novel?

AC: The introduction has been condensed significantly. Moreover, due to the reorganization of the first sections the reading flow has been improved. Whereas the initial manuscript had a potentially confusing structure, the updated manuscript follows a clear structure of

- Introduction
- Concept of the classification method
- Case Study
- Manual Data Assessment
- Classifier Selection and Training
- Automatic Classification
- Evaluation
- Discussion
- Conclusions

The aims of the paper are now highlighted at the end of the introduction with precise statements about the contributions and novelty.

RC: The methods section is again repetitive, justifying the need for, and broad benefits of, the approach, rather than stating concisely how it works. Much of the information

here is not clear. For example, Page 6, Lines 16 -21 - there is no specific detail about how tasks are undertaken and how 'a good set of classifiers' is objectively specified. Much of the methods section lacks detail and feels very descriptive and subjective; many of the choices made are not fully/objectively demonstrated.

AC: We have condensed the methods section (now called "Manual Data Assessment") significantly and concisely described the steps required for manual data assessment. A detailed description to objectively demonstrate our choices is given in each subsection. Additionally, to avoid the "descriptive feel" we have added specific examples to the classifier selection description (see p. 6 l. 24 - 31). In this way it remains clear which steps are to be taken on a high, methodology level while having short specific details on e.g. 'how a good set of classifiers' is specified (which is later defined in detail in the respective subsection).

RC: For example, Fig. 6 does not clearly demonstrate the wind speed threshold required for a 'visible influence' on tremor amplitude. Important definitions do not appear in logical places (e.g. tremor amplitude is defined after it has been used in the text).

AC: For better understanding we have indicated the wind speed threshold in the respective figure. Moreover, we have rephrased the section about the tremor amplitude such that the explanation is directly available to the reader (see Page 9, Lines 1 - 5).

RC: The key aspect of the event trigger threshold by STA/LTA is not appropriately addressed; I would like to see more critique of the application of the method in this setting. Is it too sensitive and/or appropriate given the plots in Fig. 5?

AC: We discussed the STA/LTA characteristics throughout the submitted manuscript. In the revised manuscript we added an additional paragraph to discuss the STA/LTA settings in the context of Fig.5, including the effect of the threshold (see Page 9, Lines 12 - 17). We hope that thereby the application of the method in this setting are described in greater detail.

RC: How is the accuracy of event attribution assessed, other than by ruling out moun-taineers etc. and process of elimination (page 9, lines 8 – 10 suggests this is the case)?

AC: We assume that the term event in this question relates to geophysical events for example the rockfalls discussed on page 9, lines 8 – 10 in the initial submission. In this case the accuracy of the attribution is related to the accuracy of our sources, which are incidents reported and logged by local observers, for example during maintenance of the monitoring setup. In addition, some rockfalls can also be seen by analyzing image sequences. We relate the characteristics of the micro-seismic signal to the timestamps of a rockfall report containing beginning and end timestamp. We do not use additional information/knowledge about a characteristic signal pattern which makes a rockfall identifiable only with the micro-seismic signal. Since we take only verified events into account the accuracy of these events is rather high but it also means that we probably missed to annotate rockfalls occurrences during the two years. As a result, we did not use our classifier to automatically annotate rockfall occurrences since the dataset is not accurate enough to train a rockfall classifier. Additionally, to evaluate the accuracy

of event attribution even more we introduce a new evaluation which investigates how false labels affect the classifier performance. This evaluation is presented and discussed in the section "Classifier Evaluation".

RC: The level of assumed knowledge about neural network is also rather high.

AC: We have added a new subsection "Convolutional Neural Networks" (starting Page 9, Line 7) to explain the concept of convolutional neural networks and we recommend additional literature for the interested reader

RC: I am not convinced by the 'statistical analysis' presented in Table 1 – this seems rather weak and limited in terms of the depth of data analysis.

AC: We have restructured the statistical evaluation and improved the analysis. First, methods and evaluation are strictly separated into their individual section. Second, the "Statistical Evaluation" section now comprises all results from the initial submission and new results we have added as a consequence of your comment. The results presented and discussed in depth are (i) statistics for the manually annotated test set, (ii) statistics for the automatically annotated set for the year 2017, (iii) a plot which illustrates the distribution of STA/LTA events over time.

RC: The results section draws out the key argument that the authors wish to make, but I would like to see more assessment of the data presented in Fig. 5, even at the basic level, including the duration and frequency range/spectral density of different seismic sources. Can this information be used in a simpler manner to draw the same conclusions?

AC: In the new section "Feature Extraction" we have addressed the before mentioned comment. We make an assessment of the different source characteristics and how these can be used to classify/distinguish event sources. Moreover, we discuss the pro/cons of classifying based on manually extracted features versus classifying on learned features.

RC: How sensitive are the patterns shown by the graphs to the colour scale of the spectral density information?

AC: The color scale is only a visualization guideline and has been set to the same range for all subplots to maintain comparability between the subplots. The visibility of the patterns would change with a different scale but please note that this visualization is only used for illustration in the paper. The input to the convolutional neural network is not using a color-coded representation but uses the raw spectrogram matrix.

RC: The discussion section is underdeveloped, lacks grounding and critique in the context of related literature and does not address the geomorphic significance of the

approach addressed.

AC: We have addressed the fact that the discussion section is underdeveloped and expanded by adding two new subsections ("Feature Extraction" and "Overfitting") and extending the existing subsections. Moreover, we added more literature, such as (Walter et al., 2008), (Kuyuk et al., 2011), (Eibl et al., 2017), (Fei-Fei et al., 2006). Additionally, we have included more information in the evaluation section to support discussion about the possible advantages/disadvantages of our approach in regards to the geophysical application of event-based analysis. We show how our method can be used in our case study to extend event-based geophysical analysis. However, we find that a more detailed assessment and discussion about the results of our case study (in regards to geophysical significance) is out of scope of the submitted manuscript.

RC: How does the constrained uncertainty of the approach considered compare to other sources of uncertainty, such as seismometer tilt (indeed, which component of the seismometer is being used, and why? – again see the work of Adam Young)

AC: We have updated the manuscript to include which components of the seismometer are used (all three components, see Page 9, Lines 1-2). However, since we are not performing a device characterization study the analysis of tilt impact is not in the scope of our work.

RC: [How does the constrained uncertainty of the approach considered compare to

other sources of uncertainty, such as] rock slope resonance and site effects (see e.g. Burjanek et al, 2012, 2017 GJI)? Some of the claims made about trade-off between time and accuracy feel poorly considered and require a more robust demonstration. There is also no discussion of the representativity of the case studies provided and how changes in the nature of the rock mass may affect the accuracy/source attribution of the seismic readings through time (e.g. resonance effects on duration and frequency as a rock mass degrades).

AC: We accounted for changes in the nature of the rock by using a test set from a different year. We assume that the test set is representative for upcoming, never-seen-before data (see Discussion section "Overfitting" of the revised manuscript). Since the classifier shows good performance on the test set, the changes in nature of the rock are not assumed to have a significant impact on the performance of our classifier. This assessment is of course only valid in the scope of our dataset and an interesting future work would be to investigate how specific resonance or side effects have an impact on the accuracy of the classifier.

RC: Specific comments (not exhaustive)

AC: The following reviewer comments have been acknowledged and corrected but do not need a dedicated answer in our opinion.

RC: There are many uses of e.g. in the manuscript – remove these and replace with 'such as' or 'for example' as appropriate.
RC: Brackets for citations are not always used correctly. Please check and amend.

RC: Page 6 Line 28 - has monitored? Tense is not correct.
RC: [Page 7] Line 32 – sentence beginning 'Additionally' is not clear.
RC: Page 10 Line 5 – do not use comma splices (re: therefore)

AC: The passages the following comments refer to have been removed in the revised manuscript.

RC: [Page 2] Line 3 – HAS been demonstrated [Page 2] Line 5 – micro-seismic RECORDS?
RC: [Page 2] Line 6 – biased assessments of what? [Page 2] Lines 3 – 4 and 5 – 9 – very repetitive.
RC: [Page 2] Line 22 – its (not it's)
RC: [Page 2] Lines 18 – 19 – Do you mean the accurate attribution of seismic events?

AC: The following specific comments are replied to individually.

RC: Fig. 5 - what are the red/purple circles? Are these triggered microseismic events? This isn't clear in the figure or the caption.

AC: The red circles indicate the timestamps of the STA/LTA triggers we use for the paper. We have updated the caption accordingly.

RC: Page 1 Line 19 – check terminology. Rockfalls are a type of landslides (see e.g. Varnes, 1978, and subsequent iterations of this work).

AC: We acknowledge a loose usage of terminology and a possible sources of misunderstanding given our formulation and thus we have rewritten the passage mentioned. However, we would like to highlight that in Varnes, 1978 and in the subsequent work it is recommended that the term slope movement is used instead of landslides to avoid confusion. Consequently, you are right that rockfalls are a type of slope movement.

RC: Page 2 Line 1 – what is the difference between acoustic emission and micro-seismic emission? Clarify.

AC: The difference is the frequency range in which the emission is detected. The particular sentence related to this question was rewritten in the revised manuscript. Now, acoustic emission is not mentioned anymore to avoid confusion since the focus of the study is only on micro-seismic emission.

RC: [Page 2] Line 9 – expand on scaling issues.
RC: [Page 7] Line 8 – expand on 'scaling issues' – this is unclear.

AC: By scaling issues we mean that it is for example unfeasible to manually analyze and annotate continuous micro-seismic recordings of many years. We have reformulate a similar statement to "... manual methods suffer from their inability to scale to increasing data volumes ..."

RC: [Page 2] Line 21 – what is the significance of footsteps?

AC: The paragraph has been changed in the revised manuscript. It is now made clear that the STA/LTA event detectors can be used to register external influences, such as footsteps, but "cannot reliably discriminate geophysical seismic activity from external (unwanted) influence factors"

RC: Page 3 Line 16 – F1 is not defined at this point. Indeed, much of the terminology in this section (e.g. ensemble classifier) is not clearly defined.

AC: We have addressed this point by defining F1 before it is used in the text (except for the abstract where the value is required to make a statement about the performance of our method). Additionally, it is made clear than an ensemble classifier consists of multiple classifiers.

RC: Page 7 Lines 6/7 – clarify 'sampling rate' – how was the sampling done? Or are you referring to the data transmission interval?

AC: The sampling is done by performing a measurement every two minutes with the respective sensor and then transmitting that measurement via the wireless sensor

network to the server. We have adjusted the phrasing to make this aspect clear. To answer your question: In our case the sampling rate equals the inverse data transmission interval.

RC: Lines 26 – 29 – the distinction between acoustic events and seismic events is confusing, seems a little arbitrary and lacks reference to the literature; some of the terms do not follow some conventions in e.g. laboratory monitoring of acoustic emissions; this is important for a contribution to the geophysical literature. These definitions and distinctions also come too late in the manuscript, since these terms are used earlier.

AC: We acknowledge that given the geophysical background a consistent usage of terms is required. We have updated our definitions such that an event is defined as a trigger from a STA/LTA event detector. An event can have sources of geophysical or non-geophysical nature. In our study we apply a systematic method to identify non-geophysical sources in order to take them into account when analyzing geophysical sources. The definitions are now defined at the beginning of the paper and are consistently used throughout the manuscript

RC: Table 2 – this needs a lot more detail – what is this showing?

AC: The table caption has been updated to explain the table content, the structure of the neural network, its layers, strides and output channels.

RC: [Page 6] Line 19 – triaxial or three-axis?

AC: We use the word three-component in the revised manuscript

RC: [Page 7] Line 11 – The microseismic records considered in the case study were affected?

AC: It has been updated to "The recordings of the case study were affected..."

RC: [Page 7] Line 24 – is 'sounds' the correct word here?

AC: The passage the comment refers to has been removed in the revised manuscript.

RC: Table 1 – Reword the caption. It is not the case that none of the other categories 'apply'. Rather, it is where you have not been able to classify the signal as one of the three categories discussed.

AC: It has been updated to "when none of the other categories could have been identified".

RC: *Page 9 Line 9 – Figure 5(e) does not show an example of a rockfall. It shows an example of the seismic signature of a rockfall event.*

AC: It has been updated to "Shown are seismic signatures of ..."

C17