Natural Hazards and Earth System Sciences Discussions

Nat. Hazards Earth Syst. Sci. Discuss., referee comment RC1
https://doi.org/10.5194/nhess-2020-411-RC1, 2021

Comment on nhess-2020-411

Anonymous Referee #1

Referee comment on "Debris-flow velocity and volume estimations based on seismic data" by Andreas Schimmel et al., Nat. Hazards Earth Syst. Sci. Discuss., https://doi.org/10.5194/nhess-2020-411-RC1, 2021

The manuscript presents a study on key debris flow characteristics (velocity and volume) from three study sites. The flow characteristics are estimated using a predominantly seismic approach. The goal of the study is timely, relevant and matches the scope of the journal well. From my perspective, the study design is well suited and the gathered data is substantial, providing proper substance to pursue most of the discussed topics. That said, the manuscript also shows a few drawbacks that need to be addressed before it has reached a state that renders it eligible for publication. I believe all these points can easily be addressed with the data that is inherent to the study.

In the scope section (lines 40-42) the reader gets the impression that the approach would not require any site-specific calibration efforts. This is not true. Evidently, the autors record independent data (flow stage at all four sites, and velocity by a Doppler Radar sensor at one site) and use that (table 2, fig. 8) to convert the integrals of squared seismic velocity to volume (integral of flow stage throughout an event), the corresponding fit function is presented as Eq 1. So, there is a need for calibration.

Now, it is a matter of discussion whether the authors have demonstrated that this fit function is a universal law to reach suitable results or not. My visual impression of the content of fig. 8 is: No, this is not the case. If one would fit the data of the three sites (coloured symbols) individually, the resulting regression coefficients would be quite different, and that difference could and should actually be tested quantitatively. Hence, when propagating the impact of these supposedly three different regression coefficients to the results shown in fig. 9, I would assume that we would see for each subset of dots quite different results of predicted volumes. In any way, to make things short here: if the authors wish to claim there is a universal law to relate seismic energy integrals to total flow volumes then they have to prove this hypothesis in a proper way.

It remained unclear to me, which stations the authors used to calculate the integral of squared ground velocity, one sensor (which one) or all available? Was an average value used (and can the scatter be quantified), or have the amplitudes been scaled by distance to source?

The authors raise the claim that they can deliver a metric like average flow velocity based on seismic sensors. This is only true for that stretch of the channel that lies between two sensors, some tens of metres, and thus close to an "extended point measurement" considering the total channel lengths under investigation. That information is kind of implicit but should be brought up explicitly in the abstract and other appropriate places, because seismic sensors can also be used to study the average velocity of a debris flow as it propagates down the channel (Walter et al., 2017), which is a quite different type of information.

Following the authors approach to study flow volume and velocity using seismic sensors, raises the question why this approach would really be superiour to other, more classic measurements. Sure, the other measurements require infrastructure built to host the required sensors (Doppler systems and flow stage meters) while seismometers can be installed relatively easy adjacent to the channel – a point that the authors correcty elaborate on. But, at least for the calibration work to be able to relate seismic signals to volume, some independent measurements need to be gathered, as well, no? Thus, I suggest to authors spend a few words to frame this topic a bit: advantage of seismic approach in the light of efforts for calibration work.

Line 7, "methods was", change to "methods were"

Line 12 (and further cases), the terms magnitude and volume are used alternatingly. And it is not clear to me how especially the term magnitude is defined. As I read, volume is defined as the time integral of stage height. But what about magnitude?

Line 27-28, the study of Manconi on rockslides (and also the work of Perez-Guillen of snow avalnches) does not really match with the context of the introduction (and actually the scope of the manuscript as a whole). Either provide a more elaborated overview on more of the existing approaches to relate seismic signal properties to material volume or leave this part out. This issue links to another point, see my comment regarding lines 189-198, which describe such additional (but by far not all relevant) approaches to turn seismic metrics into volume and other kinetic process attributes.

Line 33-34, this reads like magnitude is volume times velocity. Is that so? If yes, this should be mentioned explicitly and also it should be discussed to give more substance, I see there is a reference to Coviello et al. (2019), but a few more words would be really helpful, here. At a first glance the product has the unit m^4/s, doest this make sense?

Line 37-38, this is a fair point and I fully agree. However, actually the results of this study show exactly this point: there is no universal "approach" to scale seismic signals to flow properties, without any site-specific calibration. I suggest this part should be revised to not raise the implication that this study solves this issue. I do not see the point that no other studies have yet presented a universal simple method. There are numerous other studies that have used seismic sensors to investigate debris flows and have come up with methods to reveal key flow properties, including those studies cited by the authors, and especially previous studies by the author team (see line 34).

Line 38, again I see some ambiguity in the wording, here. The presented study does not at all use seismic amplitude data only, to provide an estimate of flow velocity and volume: figures 5-7 and 8-9 basically show independent data, and the data from table 2 is used to fit a regression model, which ultimately allows relating seismic data to total volumes. So in essence, this study also uses additional data as many other studies did and which is logical because as the authors mention, the seismic signal properties depend on a lot of site specific paramaters. Please revise this section to be clear. This includes especially line 40-42, which raises the claim to overcome site specific calibration needs.

Fig.1, this figure is of very limited use, providing no relevant information, not even a scale bar. Either remove it or expand its content, for example by combining it with figs 2-4. Regarding these latter figures, just as a hint, I would double check if using Google Earth shreen shots is is agreement with the CC-BY license of the journal. It would any way be better to provide proper topographic maps instead if the perspective views, unless they reveal content that a proper map would not be able to deliver.

Line 52, "G1 and G2 … marked with yellow circle … 75 m", these information bits do not add up. Yellow circles are around G2, G3 and G4, 75 m are between G2 and G3. Please clarify. In addition, I think it is not clear at all why the velocity was only estimated between two geophones when a nice linear array of four sensors is present that can be exploited. Imagine the increased depth of information and robustness if you would use four sensors, i.e., six possible combinations of velocity estimates! Why did you limit your study so drastically? Here would be an excellent chance to estimate the robustness of your velocity estimation approach, and also at the Cancia site (station pairs 1-2, 2-3, 1-3) this would be possible.

Line 55, "which reliable detects", first correct wording to "reliably" and second, I feel more detail is needed, here. What gives rise to that reliability? Can you quantify that based on the referenced work, e.g., ratio of correct versus incorrect detections, or a confusion matrix? What is this "specially designed detection algorithm" and how is it related to the STA-LTA algorithm mentioned above? I know there are other articles about this system, which I actually really see as an asset to the field of seismic hazard detection, but a few lines of explaining text would be great in this context here, as well.

Line 62, "two stations for testing the warning system", this section does not make sense.

Do these two stations belong to the system, or is MAMODIS an extra device/setup? In the former case, how can the system be tested by the system, in the latter case, where is the system in the map shown in fig. 3? Please clarify.

Line 81, I think I understand there are two approaches to get flow velocity from two seismic stations. It is not so clear from the wording, that you actually applied these two methods independently. Can you please revise the text to make this obvious to the reader? I only got this information when I looked at the legend of fig. 5 and then trying to move myself backwards through the manuscript to find the indication of these two methods. I see you mentioned a reference for the amplitude modelling approach but a few lines of explaining text would be very helpful to understand the context without needing to search for the referenced article.

Line 84, "mean surge velocity", you need to be specific here, this is only valid for that stretch of the channel/flow between the the two sensors, not the entire flow as it propagates down the channel.

Line 85, "manually analysed", my first impression was that the study pursues an automatic detection and characterisation of debris flows. How does this match up? Can you clarify, ideally at the first introduction of the idea of an automatic system. More importantly, based on which criteria did you identify comparable peaks? Was that just a subjective eye-spotting approach? What would the uncertainties be that arise here?

Line 89, the phrasing of the window size definition is somewhat unclear to me. How is the "number of samples equal to the distance" defined? Number of samples relates to the temporal domain and distance to the spatial domain. The linking factor would be velocity – which is not known beforehand. Please clarify because it seems like this selection of the window size appears to be a very sensitive parameter.

Line 95, if cross correlation is performed twice, what are the two pairs for time series that are used? Or do you mean within a fixed time window you do a cross-correlation of amplitudes (actually you should rather call this envelopes, because you only have positive values) in a sliding sub-window? If so, why only twice and based on which sub-window size and overlap? This information is not really clear, and I suggest to simply add more detail, here.

Line 101, for which stations does this energy estimate hold? Again, the source energy requires application of a bit more calculus and information about ground parameters, e.g as done by Le Roy et al. (2019, JGR). And thus, the information about which station or ideally which stations is essential. In addition to this, you mention to unit Joule, but to get from m^2/s^2 to kg m $/s^2$ there is a bit more necessary. In other words, just squaring the signal envelope will not give you seismic energy on a short track.

Line 103, this links to my above comment. “estimation of seismic energy” is misleading here, at best you get a rough proxy of seismic energy following $v^2 \sim E$, but you need to estimate the scaling factor that turns this relation to a real function to estimate the energy from seismic amplitude values. Agian, my suggestion is to either reword the text (relax the tough claim on energy estimate) or follow a similar approach as Le Roy et al. did.

Line 107, can you really justify that especially in the near field, where it is far from easy to understand the wave field, it is legitimate to ignore any attenuation of the signal? If not, consider reworking the text to be less “confident” and glossing over this topic.

Line 111-114, you can significantly shorten/consolidate this part. The first sentence repeats things we know from previous sections, The second sentence would make more sense in the second paragraph.

Line 117, Defining peak discharge as all periods above a threshold of 3.5 m does not make sense. Peak discharge is the one value of maximum discharge, not several values. I think you mean local maxima in the amplitude time series, right?

Line 121, this paragraph should be connected to the one above. The same holds for the decription of the third study site.

Line 123, be specific and replace “several” by the actual number, ideally illustrated also in the figure, for example by small numbers denoting the selected surges.

Line 124-127, can you please be more specific and elaborate regarding these results (range of values, number of surges, signal-to-noise ratio, range of cross-correlation values, and so on)? In the methods you mention a lot of things that you did but here we see only a very shortend presentation of the results of these methods. Specifically, the results of the two velocity estimate approaches (cross-correlation and amplitude modelling) should be presented in a more elaborate form. Also, avoid interpretations of your results already in this chapter but move them to the discussion section.

More to the above point, I am missing any presentation of the Doppler velocimeter results as well as of the other non-seismic instruments you use to calibrate the seismic signals. All we get is the condensed version in table 2.

Table 2, where do these numbers come from? From the gauge measurement devices? Please specify and if from independent measurements, these results deserve a bit more content than just their appearance in the table.

Line 138, a) please also give uncertainties on the fit coefficients and b) – related to my general comments – these fits should be performed also for each site individually so that the reader can judge how justified a global fit approach is.

Figure 9, what are the implications of the 20 % error range with respect to the point that only the blue dots fall roughly into it while the majority of the other coloured dots do not? How much % scatter would you need to catch all dots? Or 95 % of the dots? What do the +/- 2 sigma lines depict exactly and how does this metric relate to the data you show?

Line 143, add at the end of the sentence something like "between two closely spaced seismic sensors".

Line 145-148, this is a very broad and arm-waiving statement with little crisp information. Either include more specific (and thus justified) content or leavt it out.

Line 154-155, I think you can and should be more specific here. You can for example quantify the ratio of channel distance to station distance (20/90) as a metric to better define the term "significant difference".

Line 156-160, Ideally, you would test explicitly by signal aggregation and inspection of the impact of different sampling frequencies on the rsults. This could be easily done and I encourage the authors to do so, in order to be able to replace some of the "may"'s by justifiable hard results.

Line 165-174, this section is not really helpful in the discussion. The information given there is material I would rather expect in the introduction part, giving an overview of possibilities to measure debris flow height and/or velocity and therefore in the end justifying the usefulness of seismic sensors. Here, you have little results to raise a discussion about other potential approaches. The discussion should be based on your study's findings.

Line 176-188, I would welcome also a bit more discussion on the actual downsides of the seismic approach. It is good to underline its strengths, but there are also obvious weaknesses that deserve a discussion.

Line 178, that "variance" is indeed due to the multitude of site specific parameters, parameters that must and can be accounted for by a calibration of the seismic data.

Line 189-190, can you explain how/why the velocity would affect the frequency spectrum? This does not seem intuitive for me.

The conclusion is a weak one. It merely repeats what has been discussed before, rather than putting the findings into a wider context. Can you reach out a bit more and touch this wider impact? What is this study relevant for? What are the great assets? Which fundamental research gap/questions has been tackled? What is possible to engage with, now that the technique is there to seismically estimate important debris flow parameters?