

## General comments:

The manuscript by Lapillonne et al. addresses an important concern with a proposal for a uniform protocol for event analysis at gaging (automated monitoring) stations and represents a valuable contribution to the debris-flow community with the information provided in an event database from the French Réal monitoring stations. The authors show how the procedure can be applied to their own measurement series. They critically examine their own measurement data (quality, acquisition methods, positioning, etc.) and show the possibilities but also the limitations of the proposed methodology. The work also shows that additional information (metadata) and event descriptions can be very important to interpret measurement data correctly. The manuscript is basically well written in a good and understandable English and the procedure is described in a comprehensible way. It should therefore be of great interest to the community and fits well into egosphere. Nevertheless, various statements are not formulated precisely enough and require more detailed explanation or additions.

We are very grateful for your interesting and constructive review. Your insights are very valuable and the manuscript is improved by your suggestions and expert comments.

## Specific Comments:

12: what means “precise”? Can the authors specify how precise is precise enough... perhaps by giving a measure of uncertainty?

Thank you for the comment. In this case we mean that complete debris flow measurements, including flow height sensors, imagery, detection devices, ... are hard to install due to the harsh condition. This will be clarified as the use of the term ‘precise’ is not conveying the message correctly.

12: In general, please check how debris flow is written when used "adjectivally", i.e. in combination with other nouns. Either in this case always write a hyphen between debris and flow (debris-flow), e.g. debris-flow measurements or never use a hyphen. Both variants exist and personally, I prefer the hyphenated style.

Thank you, corrected!

15: use “observed” or “measured” instead of “monitored”.

Thank you, corrected!

51: The authors need to better derive why the characteristic values of individual surges are so important and differentiate between characteristic values of an entire debris-flow event vs. individual pulses or surges, which can sometimes differ greatly within an event.

Characteristics of surges are the focus of this paper and protocol because the numerical models getting interested in the physics of debris flow and impact studies focus on the surges. We will better explain this difference.

53: write “automated debris-flow monitoring station”

Thank you, corrected!

57: Not all can realize instrumenting multiple sites in a catchment area. This depends on the one hand on the location and on the other hand on the financial resources and manpower. Please specify that you mean the measuring points along a channel covering different channel sections and not stations at several torrents from sub-catchments. Perhaps one could introduce to differentiate that the entire measuring infrastructure in an area is called automatic debris-flow measuring station and the individual stations are called sub-stations or similar.

Thank you, that is a great observation. Vocabulary has been corrected but we would like to point out that having multiple sub stations is not in any way a requirement for this protocol. We just took advantage of our 'luck' concerning our instrumentation in this paper.

Fig. 1: Please complete the associated (or: derived) surge parameters also in the figure 1

Thank you, corrected!

65: It is not clear where the value comes from and how it is justified. If this is a minimum requirement, then conversely it should say less than or equal to 2 Hz. The changes in flow height can be very rapid and if the resolution is too low, you may miss the peak value. There is also a larger discrepancy between the different measurement methods used, including time delays in determining the value.

You are correct, there is an inconsistency here. 2Hz was initially chosen because of the refresh ability of the ultrasound sensors of our setup at that time, as well as the cut off frequency of our acquisition system. If you are interested in having more information on these sensors, a publication is currently being written on the sensors used in this station.

This has been modified to "flow stage measurements with representative frequency sufficient to describe accurately the flow front rise on the hydrograph".

66: It is essential to know the value before, during and after the passage of a surge as accurately as possible to minimize calculation errors.

Thank you, corrected! This has been added to the requirements for clarity.

67: please write "mean velocity" because here an average value of the propagation of the surge is calculated and not the instantaneous and local velocity

Thank you, corrected!

71: This vague statement about sufficiently close locations must be specified and substantiated (e.g. erosion, deposition, alteration of flow path, etc.)

Thank you, we will add a sentence to clarify what we mean by sufficiently.

"These measurements must be done at sufficiently close locations to reasonably assume that the measured flow stage is associated with the measured surge velocity. Between two sensors, there should be no major change in flow path, channel width and slope to ensure that the geomorphological processes are consistent along the interdistance."

81: write: "...is the maximum volume of the flow depth  $h$  [m]

Something was wrong in the sentence indeed, it has been corrected to is the maximum value of the flow depth  $h$  [m]!

83: provide a reference

Thank you, done, a reference to [2] will be added here.

95: use “flow height” instead of flow stage.

Thank you, corrected!

98: not clear, if raw data is used. Please specify. There are also alternative methods in which pulses are counted that exceed a specified output signal threshold (see: e.g. Abancó et al., 2012 - <https://doi.org/10.3390/s120404870>)

Thank you, raw data was used. The text will be corrected accordingly and the alternative method will be mentioned.

104: It is not made clear whether an and-link of flow height (please be consistent and use “flow height”) and geophone signal or an or-link is used to divide the debris-flow event into surges, and exactly when the separation occurs. Please specify more precisely.

Flow height has been correctly rewritten in the figures and in the text. An and-link is used. This part is reformulated to clarify.

119: see comment for line 71

This distance will be more precisely specified, similarly to our answer to L71.

120: How is the distance determined? Is the exact flow path measured (possibly a variable value) or is the direct visual distance meant? Please specify more precisely.

The distance is taken as the average flow path between the sensors i.e. the path of the main channel between the two sensors. This assumption of it corresponding to the flow path is not very precise indeed. This has been added in text.

134: Please note that debris flows often have a convex curved surface in the transverse profile. This must be taken into account when determining the flow height and the characteristic values derived from it (see: e.g. Jacquemart et al., 2017 - <https://doi.org/10.1007/s11069-017-2993-1>)

Yes, this has to be a case-by-case assumption, depending on the location of the debris flow. In our case, this was covered in Bel, 2017. This point will be added to the text.

145: please mention also Fig. 11 here

Thank you, corrected!

162: More out of curiosity: what is the advantage of placing a geophone downstream the flow height (stay consistent) sensor?

This is mainly to ensure that we have two zones to compute the velocity as the sensors are in a difficult environment they can fail relatively often, so multiplying the measurements make it safer.

168: The selection procedure should be described in more detail, e.g. by stating decision criteria and their weighting. Otherwise, the data reduction remains not transparent for others and could be applied incorrectly.

The selection is mainly based on a visual estimation of which sensor worked better : usually the sensor giving the highest values are chosen to avoid having a noise on the sensor that is of the same order of magnitude as the measured values. If one of the sensors failed (oddly high or oddly low values temporarily) then we will prefer the others. The influence of this choice has been tested and remains marginal.

174: What is meant by "...a clear appearance of the debris-flow behaviour...". Please specify more precisely.

This is in reference to section 2.1.2. We will change the text to make this point more clear for the reader.

181: Rainfall measurement is important but not of interest in this paper. This information could be omitted.

Thank you, we initially mentioned it for future user of our data, but this could be moved to supplemental information!

185: What are "significant events"? Please specify more precisely.

This is arbitrarily defined using the expertise on the station. Added to text.

Fig. 7a/b/c: For better readability and to avoid confusion, use a different color code for each legend on the right side. In the legend, a space is missing after c) and the word "literature" is not spelled with a double t (also occurs several times in the continuous text).

Thank you, corrected!

206: Turn over sentence to: "The surge volume was normalized by the catchment area to cross-compare measurements performed at different stations, but also to help transferring these results to other catchments."

Thank you, corrected!

210: literature

Thank you, corrected!

212: add "clear trend"

Thank you, corrected!

213: use "carful" instead of "prudent"

Thank you, corrected!

220: Please check whether the calculation for the Illgraben data uses the total catchment area (11.69 km<sup>2</sup>) or only the area of the active contributing sub-catchment (4.83 km<sup>2</sup>). The large difference could also be due to this. cf. Hürlimann et al., 2019 , p. 12

In this calculation, we used the total catchment area. The contributing area would not be the only contributor to the total volume of the flow. Thank you for the comment, we will add this precision to the text.

226: literature

Thank you, corrected!

237: please explain what you mean by “heavy surge”. Also not clear what you mean by “with the topography of this catchment...”. Please specify more precisely.

High in volume, that is corrected in the paper now. The reference to topography was unnecessary : high volumes are not as present because the hydrology of the catchment allows for sediment transfers to occur rather often (see Bel [1]) and the steep slopes lead to low yield criterion of the accumulated sediments.

243: what is the driver to flowing conditions then? Please specify more precisely.

Composition of the mixture changes the mobility of surges. We will specify more precisely in text.

255: Please provide an additional figure or a supplement to an existing figure with a longitudinal profile showing the position of stations S1, S2 and S3. Moreover, just out of curiosity, is the area regularly flown (with a drone) to calculate a sediment budget. Such data could provide valuable additional information.

The additional figure will be added to supplemental material. Unfortunately, we have not had the opportunity to regularly calculate sediment budget on this catchment. Such data would be very precious to our study and could be something to investigate in the future.

Fig. 8: Write (pictures from the authors). Add “note persons in pictures c) and d) for scale”

Thank you, corrected! Persons are now also lightly highlighted to better detect them in gray scale.

Fig. 9: If applicable, make the figure wider (extend the timeline) and connect related events with a line.

Thank you, corrected!

273: call them “sub-stations”

Thank you, corrected!

283: Here it becomes clear, among other things, why additional information and event descriptions are of great importance as a supplement to the pure measurement data at automatic monitoring stations.

Fully agree, this will be added to the text.

Fig 10: To make the figures easier to read, I have connected the points that belong together with a line. Could possibly be supplemented. In addition, the symbols could, for example, be displayed larger or smaller depending on the volume, which could allow additional statements to be made.

Thank you, corrected for the line, and grey scale by date was added, thanks to other comments. However, we think that changing the size of the symbols would make the graphs harder to read and might not be in our best interest. We will try and see if we can implement this in a useful and clear way.

Fig. 11: Note that the data basis of the empirical datasets are not based on measurements of individual surges (at the very most on the dominant surge of an entire debris-flow event)!

Thank you, this was more clearly added to the 4.4 section and in the title of the figure.

322: As was mentioned, data sets already exist. Accordingly, the contribution represents another and not the first data set, but it is based on the evaluation of surges. This still needs to be specified more precisely.

The wording was misleading, we meant a first dataset 1/for the Réal torrent & 2/ with this protocol for surge evaluation. This was corrected, thank you!

We thank you for the time spent on this very helpful review. The work will benefit greatly from your insights.

[1] C. Bel, Analysis of debris-flow occurrence in active catchments of the French Alps using monitoring stations, Ph.D. thesis, Université Grenoble Alpes, 2017

[2] Hungr, O. (2005) Classification and Terminology. In: Jakob, M. and Hungr, O., Eds., Debris Flow Hazards and Related Phenomena, Springer Verlag, Heidelberg.

[https://doi.org/10.1007/3-540-27129-5\\_2](https://doi.org/10.1007/3-540-27129-5_2)