17 February 2023

This manuscript uses data from a controlled rockfall experiment to evaluate relationships between directly measured features of discrete rockfall impacts and the seismic energy generated by these impacts. A conventional approach is initially taken where linear scaling relations between seismic energy or amplitude and kinematic attributes of the impacts are explored. However, the authors also use a supervised machine learning regression to predict rockfall impact parameters (mass and velocity) using >100 features calculated from the seismic data. Labeled data examples (rockfall impacts with measured mass and velocity) are used to train a random forest regressor. Prediction errors of about 11% for velocity and 25% for mass are achieved. The authors examine the importance of each feature in the regression task and find that features associated with frequency-specific seismic energy are particularly influential, consistent with previous work. Operational potential and specific benefits of the machine learning workflow (versus other more conventional methods) are discussed.

Generation of a labeled dataset for supervised machine learning tasks is a significant challenge when working with exotic seismic sources such as rockfalls, as these events are not usually cataloged. The documentation of the rockfall experiment and how the kinematic features of the rock impacts are obtained will be of interest to ESurf readers. The benefits of a machine learning approach for linking seismic observations with desired mass movement parameters (e.g., avoiding the specification of a velocity model, event locations not required) are well-motivated. I appreciate that the authors make a distinction between larger events capable of producing long period seismic energy suitable for inversion versus the smaller events studied here which cannot be analyzed using the same techniques.

The text could benefit from some editing to reduce emphasis on the wavefield attenuation analysis and add more detail elsewhere. I don't think that the introduction and evaluation of the three simple attenuation models is relevant for the thrust of the paper. It is enough to share the "best fit" model and show (via Figs. 3 and 4) the limitations of this. Instead, the authors could further discuss the utility of this work in the context of monitoring. For example, the input to a trained model is presumably the triggered, windowed seismograms for each impact. How might those windowed seismograms be obtained? For the training dataset presented in this work they were manually picked. An application involving the training of a site-specific model for each monitoring location is introduced. Do the authors think that their existing random forest model could instead be generalized to new locations? If not, why not? Presumably this would not be possible for different source types (e.g., granular flow instead of rockfall). The model encompasses both the path/site effects as well as the source physics, though, so generalization to new geologic settings may itself be non-trivial.

In light of the above general comments, I feel this paper requires moderate revisions prior to publication. Please see the specific scientific and technical comments that follow, with line/section/figure numbers provided.

Scientific comments:

§2.5	A figure showing the binary impact time series compared with the seismic signal would be helpful to explain this somewhat unusual process. For example, the binary impact time series could be shown as an additional panel at the top of Fig. 2. This would help readers understand the time lag present as well as how the camera-obtained timings of impacts correspond with the seismic signals of the same.
165, 168	Eq. 3 is not a linear function of r — though such a linear function is what is plotted as a black line in Fig. 3. Please reconcile this — but also see my synoptic comment on the attenuation model analysis in the third paragraph above.
208–210	Please provide justification for including the standard deviations of these features, as this is not done in Provost et al. (2017) or Hibert et al. (2017c) — and some of the feature standard deviations are shown to be important in the subsequent feature importance analysis.
252	This is the first time the mass of the blocks is discussed. Since the random forest predicts mass and velocity for each impact, is the mass expected to be constant? In other words, what is the "real value" used for the mass error analysis? Is it the measured block mass? I assume the modulus of the velocity is taken from the kinematics reconstruction as explained in §2.4?
253	"real value" — see comment immediately above. Please be explicit about what these values are and how they are obtained.
338–346	I generally find this paragraph hard to follow. Is "frequency content" referring to a seismic parameter that was predicted like the impact force was in the study cited? Additionally, in the sentence starting on line 342 ("This may suggest…"), what is meant by "a change of the velocity" or "a change of the mass" (for the latter, see also my comment on line 252)? Is the "variability of the impacted forces" being linked with the measured seismic features implicitly in this argument? The final sentence of this paragraph is an important conclusion, but I think the preceding body of the paragraph needs to be rewritten for clarity with definitions in particular made more explicit.
389–390	The code used for this work is only partially shared in the linked Zenodo repository. The authors should make the code for the rest of the workflow (in particular, the setup/training/testing of the random forest) available in this repository.

Technical comments:

10	"and an *a priori* knowledge of the environment" — assumptions about the seismic propagation model are mentioned in the next phrase, so what is this referring to? Please be more specific, or remove.
13	"constrain" → "constraints on"
16, 18	Do not capitalize "world" or "earthquake" here.
24	"complete" seems an odd choice here, since of course there is always more information that can be obtained. Suggest "augment"
42	"will most of the time generate high-frequency seismic waves" — this energy is always present even if the long-period energy is not prominent. I think "will most of the time" could be removed here.
42–45	This is an important point that could be rephrased to avoid the false binary of "either small or catastrophic volumes" (what about events in between these two end members?) — suggest something like "...most mass wasting processes, including smaller-volume events, which…"
55	It isn't surprising that source physics affect the scaling laws, so I suggest that "sometimes even of" be removed here.
63	For clarity, I suggest noting that the wave propagation model is a challenge for the high-frequency case specifically.
86–94	I could not find information about the mass and dimensions of the blocks anywhere in the text. This paragraph is a logical place to insert that information.
87–89	Two methods were used to measure the block shapes. Which method was ultimately chosen?
167	Note erroneous p inserted before $A(r)$.
191	Remove "In other words" as what follows are examples (of applications of random forest classification and regression).
196	I feel the quotes around "predict" can be removed as this terminology is widely used in machine learning (e.g., scikit-learn syntax).
Fig. 1	The color scale used for the trajectories (absolute velocities) should be changed so that it doesn't conflict with the geophone colors and confuse the reader. Also, please include a legend for this color scale (a colorbar).
Fig. 2	Per my comment on §2.5 above, consider adding a panel to the top of this figure which shows the impact time series derived from the camera-based workflow, so that readers can see the correspondence between those impacts and the transients in the seismic waveform.
Fig. 3	See my comment on lines 165, 168 above. The black line does not correspond with Eq. 3. Please reconcile. Also, note that the subpanel letters are not currently referenced in the figure caption.

Fig. 4	Note that the rainbow color scale is now being used for rockfall launch specification instead of geophone specification, which could confuse readers. Consider using a different color scale.
Fig. 6	It appears that the standard deviation features are plotted with some transparency. Please be explicit about this in the legend or the caption to clarify for the reader.
Table A1	In the table header, "[mean(std)]" is confusing since it appears like a mathematical expression. Perhaps just write "number for standard deviation of feature given in parentheses" or similar.