Response to referee #1

We thank the reviewer for the time taken to review our manuscript and for the useful comments. We respond to each of the individual review comments below:

Referee text is black.

Response text is blue.

Changes are red.

General comments:

I think that the reasoning of the specific experimental setup should be explained.

The first sentence of 2.4 Experimental setup has been changed to:

The experimental setup was designed regarding the following requirements: (i) an exact and rapid triggering mechanism, (ii) multiple repetitions with identical boundary conditions due to homogeneous and dry material and (iii) flexibility for future studies.

Upscaling considerations to the larger scale, real world, landslides should be discussed including scaling analysis, as is expected from physical experiments.

Scaling of laboratory experiments is a complex topic. Various corresponding articles exist for different research objectives (e.g. hydraulic experiments: Heller, 2011; landslide and debris flow experiments: Iverson, 2015; gravity-scaled and unscaled debris flow experiments: Turnbull et al., 2015; granular slide experiments: Kesseler et al., 2020). We agree that estimating scaling effects and treating them appropriately is essential to transfer experimental findings to real world scenarios. In the present study, however, these aspects were not included, because the overall applicability of the probe for experimental application was investigated, no physical properties. These aspects will be the research objective of future studies with the Smartstone.

The following paragraphs will be added as new subsection to the discussion between Sect. 4.2 and Sect. 4.3:

Scaling

A scaling of the recording ranges will be necessary if the Smartstone method is adapted to other experimental scales or velocities. Additionally, the scaling of temporal persistence of movements has to be respected as the Nyquist frequency to observe the motion without undersampling changes with the rate of movement changes (Yang et al, 2009). This means that a small pebble in a fast-moving landslide will show more abrupt changes in its velocity, trajectory and mode than a large block in a slow landslide. Thus, the ranges of the sensors and the sampling frequency have to be adjusted depending on the landslide velocity *and* the particle size. Several aspects concerning the sensor recording range for different experimental applications have to be considered:

Acceleration range: The expected acceleration depends on the velocity of the landslide, as the strongest peaks occur during nonelastic collisions of moving particles with stationary boundaries, e.g. bedrock. Thus, the range needs to be increased (by choosing a different accelerometer chip in the Smartstone) with velocity. However, to choose a gratuitously large range to avoid clipping is counterproductive, as the quantisation error will also increase, as there is only a limited number of steps within the range. A deliberated balance needs to be chosen, e.g. by performing preliminary tests.

Gyroscope range: The rotational velocity depends on the movement of the landslide but also on the size of particles. For instance, if the mass moves with 1 m s^{-1}, a single rolling pebble with 30 mm diameter will show a rotational velocity of 3820 °s^{-1} (pebble circumference 942 mm, thus 10.6 rotations per second). Thus, the expected range can be calculated using the shortest circumference of the Smartstone's host particle and the expected landslide velocity. Again, choosing a gratuitously large range to avoid clipping will increase the quantisation error.

I suggest to send the manuscript to grammatical editing.

We will engage English copy-editing services of Copernicus.

Also some typos should be cleaned from the manuscript.

We corrected the typos mentioned by both referees. See also previous comment.

Repetitions between figure captions and main text – unneeded redundancy.

Some readers may only read captions; others will read the text and have afterwards a look at the figures. We belief, the explanation of figures should deal with both cases. Probably most readers will not do both.

Add a bit more about the implication of this method in the Abstract.

The implication of recording motion data of clasts embedded in moving artificial landslides was clarified in the abstract:

Compared to other observation methods Smartstone probes allow for the quantification of internal movement characteristics and, consequently, a motion sampling in landslide experiments.

Specific comments:

L11 - mention the size\type of pebbles is more interesting than the entire mass.

We added the size and type:

Using the Smartstone probe, the motion of single clasts (gravel size, d_{50} of 42 mm) within approx. 520 kg of a uniformly graded pebble material was observed in a laboratory experiment.

L79 - what is the former version? – maybe I missed.„

The former prototype version is now mentioned earlier in the text. The reference is Gronz et al. (2016).

L124 - needed?

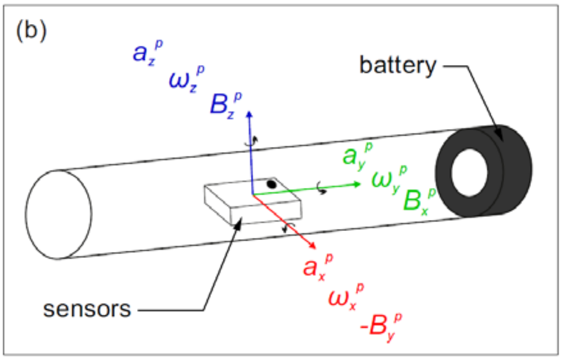
This information demonstrates that the Smartstone probe can be adapted to different research objectives.

L105 - The sampling rate (100 Hz) is low. Most probably, it cannot record the sharp impulse during a collision with another rigid body. Miss recording such impulse can lead to significant errors in the velocity and position, calculated by integrating the recorded acceleration. The authors should address this issue, provide an estimate of a typical collision duration in their experiment, and show that it is longer than 1/sampling rate. In case the above condition is not fulfilled in the experiment, the authors should explain the implications. A short discussion on the sampling rate in a broader context of a real landslide can be illuminating for the reader.

During the data analysis we did not notice any issues resulting from the sampling rate. Generally, the needed frequency does not depend on the duration of accelerations but on the uniformity of the movement. The sensor does not miss single strong acceleration peaks shorter than the sampling frequency due to its principle of operation: Inside the chip is a small movable piece shaped like a comb (for each axis). It is positioned inside a stationary second comb. Lateral movement due to accelerations changes the distance between the two combs, resulting in an electric capacity change, which is measured to derive acceleration. If a short, abrupt acceleration occurs, the movable comb will be displaced in any case, even if the acceleration does not last as long as the sampling period, due to its inertia. The chip integrates inherently during the sampling period. However, if the sensor is moved several times forward and backwards within a sampling period, the sensor will miss this information, as the movement's frequency is higher than the sampling frequency. Thus, the correct movement cannot be observed. Again, preliminary tests facilitate choosing the correct sampling frequency.

L133 - It appears that the system of coordinate of the ACC is not following the convention of the "Right-hand rule". Can the authors comment on that. In any case this is an important information for the reader.

Thanks for spotting the error in the figure. The figure has been modified. Of course, the axes follow the convention. Fig. 1 (b) has been modified:



L137 - 1. The term "higher-order" may not be the best choice as the position of a body/point is always relative. 2. The "real" axis system is not suitable for comparison of different probes; the "flume" system suit this purpose; this is why the authors used it for the graph in Figure 6 that compares the movements of different probes.

We modified the description in 2.2 beginning at L135 including the previous comment as follows:

Therefore, its x- and y-axis are also rotated. Following the right-hand rule, positive rotational directions are indicated by small curved black arrows.

To compare relative movement characteristics like distance or velocity of different probes, the inner data / coordinate system *p* must be transformed into an outer reference system *rel* (Fig. 1, c). The simplest way to do this is the construction of a reference system using the probe's starting position as coordinate origin. The system is defined by the sensor's inner coordinate system of the first timestep rotated so that z-axis follows gravity. The axes of this relative outer coordinate system are donated with x^{rel}, y^{rel} and z^{rel}. After the motion has started, the probe's inner orientation will change while the outer reference system keeps its axes configuration. Consequently, within this reference system, it is possible to calculate the probe's orientation and the covered distance in each timestep. In Fig. 1 (c) for instance, the probe has changed its orientation significantly compared to its starting position while moving along the assumed trajectory.

However, the different probe-specific outer coordinate systems must be transformed into the same local reference system to compare different probes' trajectories. For the present study, this local reference system is oriented towards the experimental flume (see section 2.4). Following the former conventions, the axes were donated as x^f, y^f and z^f. Note that the axes orientations of the outer (rel) and the local reference system (f) may not be identical, except of z^{rel} and z^f, as they follow gravity.

In different applications, where a global positioning is required, reference points of the outer coordinate systems must be known in the global system to determine the absolute probe position in the global system.

L223 - Can you clarify 0.0 g means?

Within the gravitational field of the earth, zero g can only be achieved if and only if an object moves in pure free fall without aerodynamic drag. In this case, the resultant acceleration vector of motion and the acceleration vector due to earth's gravitational acceleration exhibit the same magnitude. These vectors, however, point in the opposite direction compensating each other. This results in a measured acceleration magnitude of 0.0 g at a measuring device, such as the Smartstone probe.

L225 - The whole section (3.1) is very long and tedious. The formulas are trivial, in any case, the authors do not use the projections of g in the discussion.

We expect different subsets of so-called trivial knowledge at different readers. Thus, this section is an introduction on how this kind of data has to be read and interpreted. In discussions with colleagues it emerged that it is often not clear what is displayed within the graphs (e.g. Fig. 3). To our knowledge, there is no publication that describes spatiotemporal motion data adequately in a geoscientific context. Therefore, we decided to give a detailed description in this article.

L432 - the end of the sentence is missing.

Corrected:

It is visible that pebble 1 and 2 were embedded into the material, whereas pebble 3 and 4 were placed at the surface of the material (see Fig. 2).

L437-442 - The most probable reason for the wrong trajectory of pebble 3 is miss recording of collision with another pebble due to the slow sampling rate of the used IMU.

As explained above (comment L105), missing an acceleration peak is not possible due to the design of the acceleration sensor (as long as the motion frequency does not exceed the sampling frequency, which is given in this case). Therefore, this cannot be an explanation for the erroneous inclination of the trajectory. In addition, an overestimation of the vertical trajectory component occurs right from the beginning of the motion. The peak would have to be occurred before the motion starts, which obviously is not possible (not least because pebble 3 was placed at the surface of the material).

L446-447 This statement does not fit the description of the behavior of one body, out of many, in a multi-body system where collisions between bodies redistribute the energy of the system in a random way.

This is correct and we had to rethink this explanation. New text:

Okura et al. (2000) observed that blocks positioned at the front were also placed in the frontal deposition zone. In our experiment, the top pebbles travelled the longest distance. Regarding the high-speed video, the explanation is given by the tilted gate: The pebbles positioned on the top start their movement both downwards and to the right (from the camera perspective), thus not transferring energy to material formerly placed underneath in the storage box. The higher the pebbles are placed, the bigger is their overhang, resulting in less material vertically underneath. Compared to a vertical gate in Okura et al. (2000), less energy dissipation occurs.

L472-473 - Too much details.

Deleted.

L510 - In the present study, the probe monitored movements over a short period of ~ 2 sec. A brief discussion regarding the expected error in retrieving the trajectory over more extended periods can help to assess the type and scale of landslides that can be monitored in this way.

We agree that the duration of recorded motion is a critical aspect of this technique. Because the estimation of the displacement components requires double-integration, absolute errors will increase with time – depending on the motion. Therefore, longer durations will generally result in less accurate absolute results. The relative error will remain stable for certain kinds of movements. However, a general extrapolation of the expected error is not trivial, as it depends on the ratio of noise, quantisation error etc. to the magnitude of the true accelerations.

<u>Figures:</u>

Figure 3 - change the axis title to the same side.

Modified:

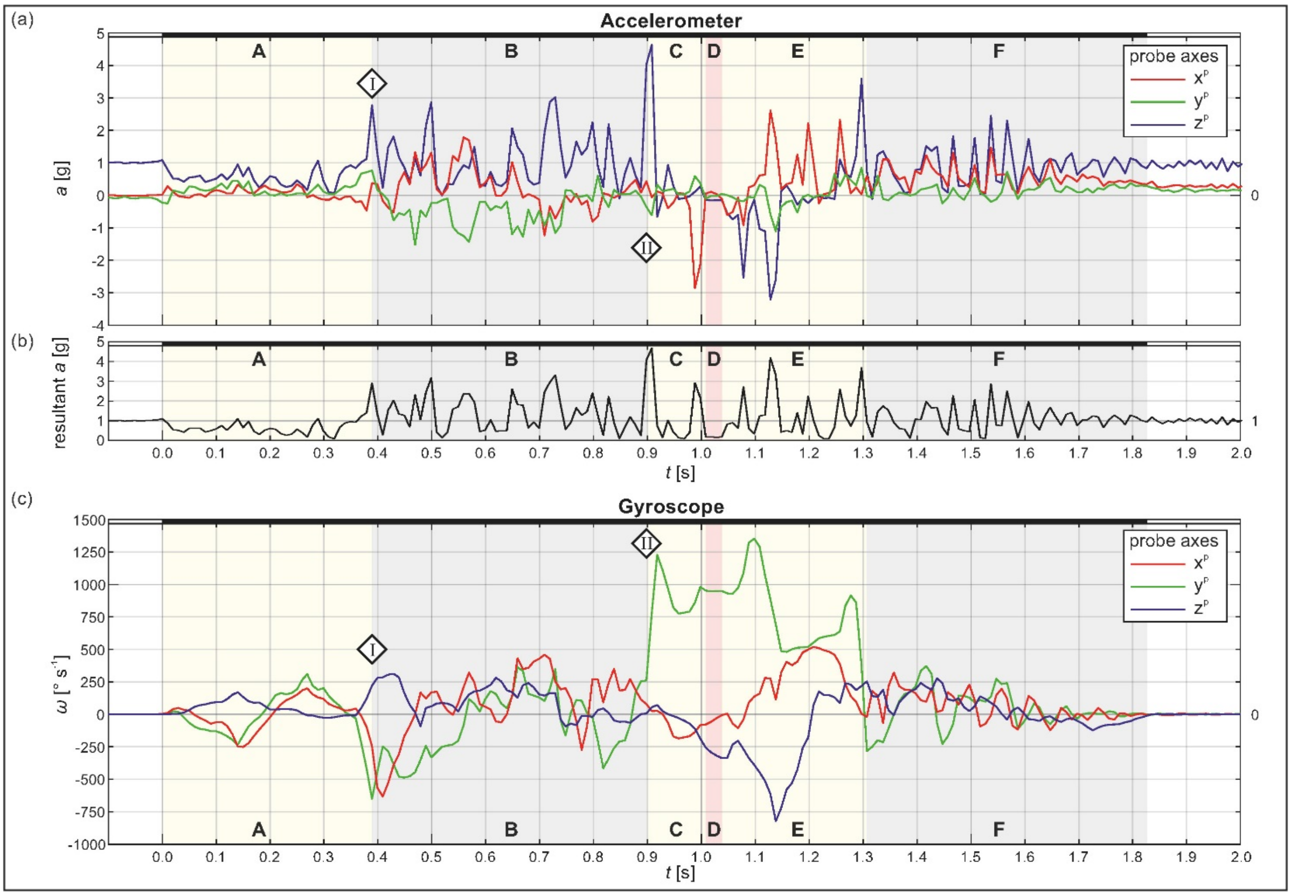


Figure 5a,c - add the flume reference as well.

To improve the orientation within the figure, the flume bottom (grey dashed line) has been added to Fig. 5 (a). Adding the flume bottom in Fig. 5 (c) as well would reduce the clearness of the trajectory visualisation. Fig. 5 (a) was modified as follows:

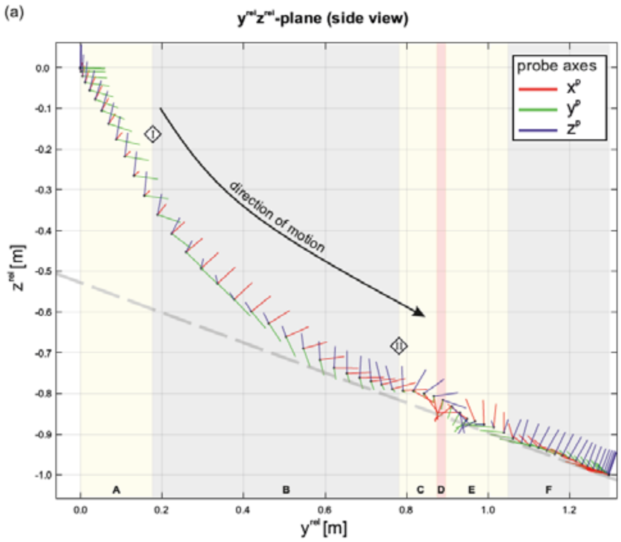
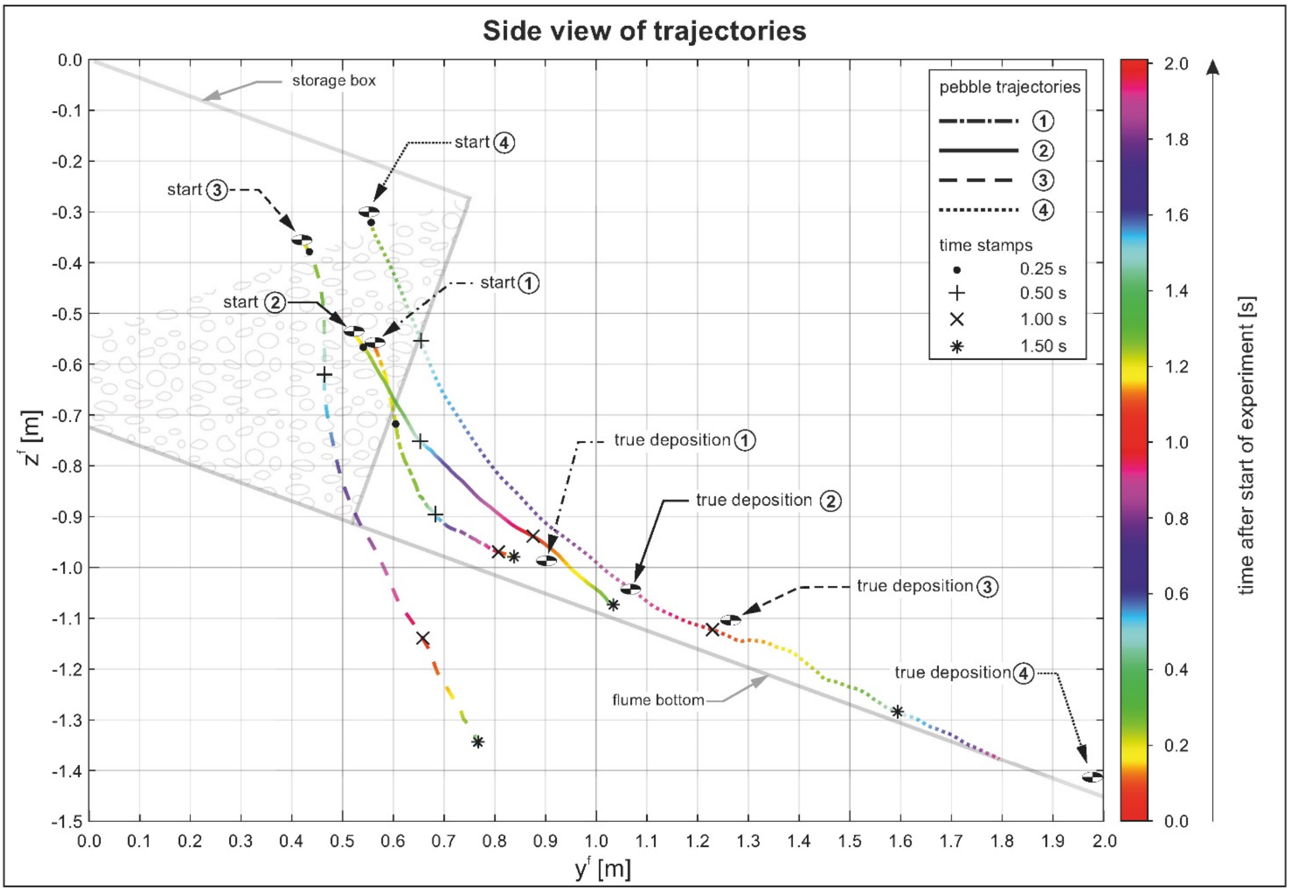


Fig 6 - color coding has a few cycles so it is not injective and a bit hard to follow, maybe add time stamps at the end of each cycle?

Repetition of cycles allows for a more precise identification of time compared to only one cycle. As the trajectory is continuous in time, the colour coding might not be injective, but it is distinct: The first time, red occurs again along the trajectory, must equal 1 s. However, we can add timestamps although we think that this is content overload.



Additional References:

Heller, V. (2011). Scale effects in physical hydraulic engineering models. In: Journal of Hydraulic Research 49 (3), pp. 293–306. DOI: 10.1080/00221686.2011.578914.

Iverson, R. M. (2015). Scaling and design of landslide and debris-flow experiments. In: Geomorphology 244, pp. 9–20. DOI: 10.1016/j.geomorph.2015.02.033.

Kesseler, M.; Heller, V.; Turnbull, B. (2020). Grain Reynolds Number Scale Effects in Dry Granular Slides. In: Journal of Geophysical Research: Earth Surface 125 (1). DOI: 10.1029/2019JF005347.

Yang, W. Y.; Chang, T. G.; Somg, I. H.; Cho, Y. S.; Heo, J.; Jeon, W. G.; Lee, J. W.; Kim, J. K. (2009). Signals and Systems with MATLAB. Springer-Verlag Berlin Heidelberg: Berlin, Heidelberg.

Turnbull, B.; Bowman, E. T.; McElwaine, J. N. (2015). Debris flows: Experiments and modelling. In: Comptes Rendus Physique 16 (1), pp. 86–96. DOI: 10.1016/j.crhy.2014.11.006.