Response to the reviewers: Computational Sedimentation Modelling Calibration: a tool to measure the settling velocity at different gravity conditions, By N. J. Kuhn and F. Trudu

Response to Referee #1

Dear authors,
Thank you for the clarifications made to the manuscript.

Before publication, I urge the authors to consider presenting the Earth data in the main part of the paper and make a full comparison between the Earth data, low gravity data and the Ferguson and Church model/equation, preferentially in one figure. Without this comparison, many of the statements seem unfounded. I am sure they are not, but as it is presented now, the data does not prove the statements made, e.g. "...values predicted using models calibrated with data collected at terrestrial gravity underestimate settling velocity on Mars."

It is necessary to show that the Earth data matches the results on the Ferguson and Church model/equation before you can state that it does not work for low gravity environments. If you do not, the reader cannot be certain your calibration parameters C_1 and C_2 of the equation are correct for the tested particles, they should be validated. A wrongly chosen C_1 and C_2 could be an explanation for the difference in results of the model/equation and the low gravity results.

Based on the response to the review, I understand the authors want to present this manuscript as a method paper. In my opinion the method cannot be proven without Earth data. Furthermore, the method is in my opinion not novel enough to present as only a method paper. An earlier version of this setup has already been presented in Kuhn (2014) and since then only experienced minor improvements. This setup is also not the final version, as improvements are noted for the future.

I am concerned that the resistance to do the comparison and showing the Earth data in the main paper is a result of salami-slicing of the data from the parabolic flight. Please prove me wrong.

Response: We had performed a Ferguson and Church parameter optimization prior to testing our instrument and included the results now in the Methods section (section 2, lines 156 to 171). C_1 has to stay at 18 because it captures viscosity-induced drag, which is not affected by gravity. For terrestrial gravity, Ferguson and Church calculated a value for C_2 of 0.4. Fitting C_2 to settling velocities observed at terrestrial gravity generated a value of 0.36. The small difference illustrates the suitability of the Ferguson and Church model to simulate the settling velocities of the particles we selected for this test. We attribute the small difference of C_2 to small inaccuracies of particle shapes and sizes. For the further calculations in the study, we kept using the value suggested by Ferguson and Church because it reflects the nature of the error made in studies that apply non-calibrated models to sediment textures observed in high-resolution imagery from Mars. For our study, the effect of this choice is limited because of the small difference between non-calibrated and calibrated value.

We agree with the referee that the error between observation and calculation is small, but this was expected because the experiment was aimed at testing the instrument aimed at particle tracking in a parabolic flight environment. We clarified this aim in the text by describing the limitations of previous instruments in more detail (Introduction, lines 76 to 90)). We also added a figure (Figure 1) that illustrates why the effect of gravity is larger for fine than coarse sand: in the fine sand range drag values drop steeply so that the gravity (and associated settling velocity) -induced error would be much greater for a 200 μm particle than a 2 mm particle. The reason we did not use smaller particles is their limited visibility in GoPro videos. A video system that would provide sufficient resolution to capture the movement of fine sand would have to be custom-designed at costs of several ten-thousand Swiss Francs. We therefore decided to test the general suitability of parabolic flights to capture highly accurate tracks of individual settling particles first before developing an improved imaging system. Where appropriate, we clarified this aim throughout the text. We also reject the notion of salami-slicing our results, because we do not have any data on fine sands while a special issue of Earth Surface Dynamics focusing on analogue planetary environments appeared to be a good match for the scope of this study.

Response to Referee #3

The role of gravity in geological processes is an important topic for a better understanding of these processes on other planets. This manuscript describes results of experiences led in parabolic flights. The results suggest a difference with models, namely an underestimation under Mars gravity. Overall, the protocol of experiences and related parameters are well described, and the results well explained, making this paper useful for the community. Nevertheless, I have a series of comments that I would like to be answered before any eventual publication.

One of the key results is the difference between the experiences and the model from Ferguson and Church 2004. Yet, the differences are not dramatic. For instance, Table 4 for Mars simulated gravity, the first line indicates 17.2 for the experience, and 16.1 for the model, making of this a <10% difference. This is the case for most Mars and Moon results. Yet, the error bars are not enough well explained nor plot in the key diagram of figure 7. One one hand, line 223 it is indicated that the errors on velocities are limited to less than 3%. Later, on line 253 it is written that for the moon "The maximum error of the observed velocities ranges from 3.8% to 10.2%". So why would the error be up to 10% for the experiences of the moon, but stated in general as 3% earlier?

Response: The referee may have misunderstood the meaning of the data presented in the different parts of the text: the *"the errors on velocities are limited to less than 3%"* refers to the accuracy of the measurement listed in Table 3, which is the relative percentage between the velocity value calculated by frame count and distance covered by the particle and the value obtained by the least squares method. The 3% do therefore not refer to an error when using drag coefficient values from Earth on Mars, but the inaccuracy associated with our measurement method. The other percentage values refer to the different gravity scenarios. The way they are cited in the referee's statement are incomplete, e.g. for Mars the complete sentence reads "*The maximum error of the observed velocities ranges from 3.8% to 10.2%, which is lower than the deviations obtained for hyper and Martian gravities.*"

On the other hand, the approximation on the gravity is large. First, it is mentioned line 134 that the "angle offers approximately 33s of Martian gravity" and the table S3 indicates a variability of the g between 3.4 and 4, with a variability from flight to flight. This makes more than a difference of 10%. How was this "approximation" taken into account? Is there a diagram of g with time that could help the reader to evaluate if this approximation is of second order or not? Otherwise, it could also explain some of the difference with the model. while the results for the latter were produced for a precise Mars gravity, we can imagine that the some of the experiences may have been dominated by a gravity that was not exactly that of Mars.

Response: During the parabolic flights variations in gravity occur. This was expected and our instrument was designed to cope with them. Gravity was logged at 0.1 second intervals and gravity loggers were being synchronized with the videos at one-second intervals. Consequently, for our calculations we used the actual gravity values measured during the short periods (max. 3 seconds) the particles settled through the settling chamber. The procedure has been described in the Methods section 2.3 *"The gravity logger data, which has a time frequency of 10 Hz, are then matched to the tracking records by joining them to the image with the nearest recorded time."*. The gravity data for the individual calculations are already reported in the results (Table 2).

Again, we play here with <10% differences, and given the difficulty of measurements during these flights, this point should be well discussed. I would recommend adding error bars on the diagram of figure 7 which would make it more scientific and more convincing.

Response: This issue is addressed in the reply to referee 1. To reiterate, the aim of the experiment was to test inasmuch individual particles can be tracked in a parabolic flight environment. This is a prerequisite to studies on smaller particles and studies aimed at generating data for fluid dynamics modelling. The small differences were expected because of the size of the particles.

Also, I would be better convinced by the results if Mars and Moon gravity results would go on the same direction, and hypergravity in the other. This questions where are Earth data in this trend? Should be precisely similar to the model, right? Regard to this point, the authors mention after a previous review that they have included the data within Earth gravity (outside flights) in the Supp Table 1, but I am not sure to really understand that table, or perhaps the caption should be clarified if those are indeed under Earth gravity results. Yet, the results made with terrestrial data could be plot on the figure 7 to make clear that the experience is well set up and provide no difference with the model in that case.

Response: We are not sure what the referee is addressing in this comment. The differences between observations and calculations are reported in the manuscript (Tables 4 and 5, and the new figure 8, which was formerly figure 7). The calculations are based on a model suggested and calibrated for gravity on Earth. For terrestrial gravity the model works well (see new text in section 2.2), so the only comparison that generates information with regards to the effect of gravity on drag is the difference between model and observation for the gravities that differ from Earth. However, the aim of the experiment and the manuscript is not to study this error in detail but to present parabolic flights as an environment to study processes such as sediment settling on Mars. Finally, Table 1 in the Supplementary material does not refer to an Earth-Mars comparison but reports the data on the accuracy test we performed on the method we used to calculate settling velocities from the videos. The procedure is summarized in the caption and section 3.1.

Actually on Figure 7, the purple bars are useless and misleading. They must be removed because they are redundant from the orange bars, and they correspond to a %, not a velocity as in the Y-axis label.

We would like to keep figure 7 (now figure 8) because it shows in a not uncommon way the absolute differences as well their percentages. The legend clarifies that the purple bar is a percentage.

In the abstract and several other locations in the discussion, the authors mention that the model underestimate the value from the experiences, but, for the Moon, it is an overestimation. So I would suggest either use underestimate only when mentioning Mars gravity, or write it differently, for instance that the experiences provide substantial differences with the model.
The results at hyper gravity should be better introduced and commented. Why are they done, for which body? Any future terrestrial exoplanet? Or is this just to test the model?

Under- and overestimation: we could not find a statement where underestimation did not just refer to Mars and/or hypergravity. The data from all the gravities were reported to give a full account of the capabilities and limitations of the instrument. Again, this manuscript has been aimed at a special issue on analogue environments with a focus on Mars. The data we have at this stage are limited and do therefore not allow much comment on the reasons for the observed and other potential differences.