Review Report

Schaefer et al., Laboratory Simulation of Earthquake-Induced Damage in Lava Dome Rocks, TEKTONIKA, 2023

Table of Contents

1st Round of Revisions

Decision Letter

Dear Dr. Schaefer,

Thank you for submitting "Laboratory simulation of earthquake-induced damage in lava dome rocks" to Tektonika. We have received 2 comprehensive reviews of your manuscript which you should now be able to access. We have selected "request revisions" for your manuscript.

The first reviewer raised concerns regarding the amount of damage recorded in most samples through such low amplitude stress cycling (with implications for the conclusions). In particular, the reviewer highlighted that most of your data could be interpreted within the hypothesis that stress cycling does not strongly affect damage state and suggested that the manuscript thus required restructuring and more information on measurement sensitivity; they also suggest extra experiments may be beneficial, but we understand that this may not be feasible given time/funding constraints. The second reviewer indicated they would like to see more quantitative microstructural evidence of the damage (which could help address some of the concerns of the first reviewer).

Overall, from reading both the reviewer comments and manuscript, your data does contribute new and important insights to the literature regarding the build-up of damage in volcanic rocks, but there does seem to be a bit of a jump between incremental damage accumulation and how it may result in failure. This will likely be rectified by addressing the reviewers comments and adding more nuance to the discussion.

We hope that you will be able to address the reviewers' concerns and submit a revised manuscript soon (please note that Tektonika does not impose a time limit on submitting revised manuscripts). Feel free to reach out if anything is unclear.

All the best,

Noah Phillips and Craig Magee

Comments by Reviewer 1

Section A: Overview of manuscript

A1) Overall evaluation, general comments & summary

A1.1) Reviewer’s comments

A1.1.1) General evaluation and publication suggestion – Required:

Please use this space to describe, in your own words, the core subject of the submission and your overall assessment of its suitability for publication.

The manuscript describes experiments on a suite of samples from Unzen volcano to explore damage caused by earthquakes, weakening the dome structure and potentially leading to collapse events. The authors deform their suite of samples with well-considered load hold and oscillation experiments, designed to investigate how different stress fields and perturbations and sample characteristics might influence the degree of damage imparted on a sample.

The manuscript has many strengths. The experiments are well thought through and nicely investigate scenarios often not considered. The manuscript is well written and gives nice explanations of the method and results. I therefore recommend acceptance with moderate revisions.

A1.1.2) What does the submission need to be publishable? (select as needed; comment for all cases)

☐ No changes required

☒ Rewriting

☐ Reorganising

☒ More data/figures

☐ Condensing

☐ Reinterpretation

☐ Other

Comments:

Overall, the main ways the manuscript could be improved are by making providing more evidence for links made between the experimental data and sample characteristics, and by exploring the implications of these in more detail.

The sample characteristics are only described quite vaguely and more detail (ideally with quantification) would make the discussion of the data more convincing. I agree that the sample characteristics will have an effect on the damage imparted, but without more description of vesicle sizes/distribution, microfracture densities/lengths in the samples it is difficult to follow which characteristics are controlling the damage degree of damage imparted on each sample. The addition of a table may help here too for the reader to refer back to if the authors would rather not include this information in the text.

Re-writing to add detail to the text would improve the manuscript and ensure that each point is backed up by evidence. Much of the writing is clear but some sentences could be improved by rephrasing to simplify the sentence structure and keep terms consistent throughout.

A1.1.3) Can the submission be improved by reducing/adding any of the following? (select as needed; comment for all cases)

- ☒ Text
- ☒ Table
- ☐ Figures
- ☐ Supplementary material

Comments:

Adding additional text that explores the implications of the results would improve the manuscript. What does it mean for hazards if the earthquake simulation experiments impart more damage on samples? This has been slightly explored but expanding these ideas would help. What else needs to be done in the future to expand on your experiments? The introduction would benefit from more detail about how domes might be expected to fail (e.g. why is tension important?).

Including a table with the information on sample characteristics would provide the reader with an easy way to see the connections between sample characteristics and the experimental data. These characteristics are difficult to see in Figure 2 but annotation might avoid the need for more images.

A1.1.4) Please complete the following section if you recommend that the submission is NOT appropriate for publication (select as needed; comment if a box is selected)

☐ Quality is poor

☐ Research is not reproducible

☐ Other

A2) Summary of main merits and main points of improvement

A2.1) Reviewer’s comments

Please describe below in a few sentences (100 to 300 words) the main merits of the submission and suggestions for improvements.

The main merits I have found are...

This is a well-written manuscript that clearly explains the background and rationale behind the experiments completed. The methods are thoroughly written in a step by step manner and the data is quite well presented. The experiments are interesting and cleverly explore the potential of earthquakes to impart damage on samples, impacting dome stability.

The main points of improvement I have found are...

The manuscript would benefit from more detail about the sample characteristics as these are later used to explain the experimental results and would make these points more compelling. Inclusion of a table would also help here.

A greater exploration of the implications of these results would improve the manuscript as many interesting points have been made but only briefly. For example: Why does strain accumulation decrease after 5-7 events and what does that mean for dome stability? How and why does pore geometry affect the damage accumulation?

A2.2) Author’s responses:

Section B: Detailed evaluation of manuscript

B1) Title and abstract

B1.1) Reviewer's comments

*These statements are a **guide** to what good* Titles *and* Abstracts include*. Please select YES or NO to the statements below if you wish and detail in the free form box below your reasons for any box checked with NO, or to comment on any other matter.*

The *Title* describes the main topic of the manuscript **accurately** — [YES]

The *Title* describes the main topic of the manuscript **succinctly** — [YES]

The *Title* includes **appropriate key terms** — [YES]

The *Abstract* includes a **clear aim and rationale** — [YES]

The *Abstract* supports the rationale with **sufficient background information** — [YES]

The *Abstract* includes a **well-balanced description of the methods** — [YES]

The *Abstract* describes the **main results sufficiently and adequately** — [NO]

The *Abstract* clearly describes the **importance/impact of the study** — [YES]

The *Abstract* clearly states the **conclusions of the study** — [YES]

The *Abstract* is **clear** and **well structured** — [YES]

Comments:

The title nicely fits the manuscript and the abstract gives a clear outline of the study.

In the abstract there are places where rephrasing could add more detail without increasing the abstract length. This is true for much of the manuscript where sentences can include quite vague language – see line by line comments.

The abstract gives a good explanation of the results but the key points you pull out could better match the contents of the manuscript. The abstract mentions vesicular materials buffering earthquake damage for example, which is not mentioned elsewhere. If you would like this to be a key point it needs to be mentioned more explicitly in the text, and it would be interesting to explore further in the discussion too. The orientation of fabric is also mentioned only very briefly in the main text – again to be a leading point in your abstract this needs to be mentioned in more detail.

B2) Introduction

B2.1) Reviewer's comments

*These statements are a **guide** to what good* Introductions *include. Please select YES or NO to the statements below if you wish and detail in the free form box below your reasons for any box checked with NO, or to comment on any other matter.*

The *Introduction* provides **sufficient background and context** for the study — [YES]

The *Introduction* describes the **aim/hypothesis/rationale** clearly, providing **sufficient context** — [YES]

The *objective/hypothesis/rationale* **flows logically from the background** information — [YES]

The *Introduction* describes the study's **objective and approach** (last paragraph) — [YES]

The *Introduction* contains **relevant, suitable citations** — [YES]

The *Introduction* is **organized effectively** — [YES]

Comments:

This is a good clear introduction that really nicely sets the scene for the rest of the manuscript. Lots of background information is neatly tied together for the reader and the introduction flows logically.

The introduction contains quite a lot of information about Unzen and some of this might be better placed in the next section on Unzen itself, instead just mentioning the dome and its instability – see line by line comments.

B3) Data and methods

B3.1) Reviewer's comments

*These statements are a **guide** to what good* Method *sections include and good practices for* Dataset *accessibility. Please select YES or NO to the statements below if you wish and detail in the free form box below your reasons for any box checked with NO, or to comment on any other matter.*

The *Methods* are described **concisely and with enough detail** for reproducibility — [YES]

Necessary information about **data sources/acquisition/processing** is included — [YES]

Data used are accessible via either supplementary files or links in the data availability statement — [NO]

The *Dataset and/or Methods* are **organized effectively** — [YES]

Comments:

The methods are well explained in lots of detail.

There is no data availability statement in the manuscript.

B4) Results

B4.1) Reviewer's comments

*These statements are a **guide** to what good* Result sections *include. Please select YES or NO to the statements below if you wish and detail in the free form box below your reasons for any box checked with NO, or to comment on any other matter.*

The *Results* findings are **supported by data** — [NO]

The *Results* findings are presented **clearly and succinctly** — [YES]

The text in the *Result* section **cites tables and figures appropriately** — [YES]

The *Results* directly **relate to the study objectives** — [YES]

The *Results* present **data for all the approaches** described in the *Methods* section — [YES]

The *Results* **text belongs to the Results section**, not to *Introduction*, *Methods*, or *Discussion.* — [YES]

The *Results* section is **organised effectively** — [YES]

Comments:

The results section gives clear descriptions of the findings based on evidence from the data, citing figures appropriately. The results are organised in a logical manner.

There are some results e.g. UNZ9a and UNZ9b that are not specifically mentioned in the results section but are then discussed in the following discussion section. Can you include the results in the appropriate section and then explore them further in the discussion.

To make some of the links to sample characteristics made in the discussion and mentioned in the abstract and conclusion, the results section needs to contain more detailed/clearer descriptions of the sample characteristics such as vesicles and phenocryst sizes, density of microfractures etc to allow the reader to reach the same conclusions.

B4.2) Author's responses

B5) Discussion and conclusions

B5.1) Reviewer's comments

These statements are a ***guide*** *to what good* Discussions *and* Conclusions *include. Please select YES or NO to the statements below if you wish and detail in the free form box below your reasons for any box checked with NO, or to comment on any other matter.*

The *Discussion* is **focused on the objectives** of the study — [YES]

The *Discussion* **addresses all major results** of this study, which are shown in *Results* — [NO]

The *Discussion* section makes **comparisons with other studies** that are relevant and informative — [YES]

The *Discussion* section properly identifies all **speculative statements** — [YES]

The *Discussion* section presents the **implications of the study** persuasively — [NO]

The *Discussion* section **highlights novel contributions** appropriately — [YES]

The *Discussion* section **addresses the limitations** of the study appropriately — [YES]

The *Discussion* section is **organised effectively** — [YES]

The *Conclusions* are **consistent** with and **summarise** the rest of the manuscript — [YES]

The *Conclusions* are **supported by the data** in *Results* and **follow logically** from the *Discussion* — [YES]

The *Conclusions* are **clear and concise** — [YES]

Comments:

The discussion includes a great section on the possible sources of repeated stressing of rock and the impacts that this might have on overall strength. The conclusions section gives a great outline of the study with writing that is more detailed than elsewhere.

More explanation and exploration is needed in the discussion section to be able to link the sample characteristics to the experimental data in a compelling way for the reader. Why do these characteristics control the damage evolution seen? What mechanisms are acting here?

It could be clearer what the implications of the study are. In the conclusion you state that your results have important implications for determining the damage accumulation of rock masses during stressing events and for monitoring volcanic stability – what are these implications? Can you give more detail?

B6) Figures, tables and citations

B6.1) Reviewer’s comments

These statements are a ***guide*** *to what good* Figures *and* Tables *include and how they are presented. Please select YES or NO to the statements below if you wish and detail in the free form box below your reasons for any box checked with NO, or to comment on any other matter.*

Tables and *Figures* are **ordered logically** and **numbered sequentially** — [YES]

Tables and *Figures* have **captions that explain** all their major features — [YES]

Tables and *Figures* have **captions that complement** the information in the main text — [YES]

Tables and *Figures* present data that **relate** to the study objective — [YES]

Tables and *Figures* present data that are **consistent** with and support the description of results — [YES]

Tables and *Figures* have **succinct and informative titles** — [YES]

Figures are **accessible** (elements are clearly labelled, accessible colour palettes, colour contrasts, font size legible, etc.…) — [NO]
Please, check our [Figure guidelines]

Figures with **maps or cross-sections** contain all **elements to be understood** (north arrow orientation, scale, visible coordinates, sufficient coordinate grid intercepts) — [YES]

Figures with **maps** have **sufficient location information** (in the map or caption) — [YES]

Cross-sections have clear labels for **scale and coordinates** at ends and within-section kinks — [N/A]

All georeferenced elements are provided in common format (.shp, .geotiff, .kml) [in an open-access repository] — [N/A]

Citations throughout are relevant, suitable, and comprehensive — [YES]

Comments:

The manuscript has nicely presented figures that display the data quite clearly. Figures are neat and well put together, and contain much useful information that adds to the manuscript.

Not sure if I'm missing something, but where is the data point for UNZ1 at 70% in Figure 4d? Is it perfectly hidden behind another data point, or had the sample failed at this point? Something needs adding to the caption here.

Altering some of the symbols used in figures and simplifying axes would help the data to be displayed more clearly – see line by line comments.

Section C: Additional comments

C1) Minor/line-numbered comments

C1.1) Reviewer’s comments

Abstract

Line 24: Why is this rarely systematically measured? Can you add a few words to explain here?

Line 30: Are these samples that were already fracture-dominated or are these fractures gained during your experiments? Are you referring to microfractures or larger scale fractures?

Line 31: How do these buffer earthquake damage? It is a lovely idea that needs to be mentioned elsewhere if you want to include it in the abstract.

Line 33: I agree that the direction of fabric will be important – can you be more specific about your findings here?

Line 35: Can you rephrase to shorten this sentence or split into two? Such as “During each “earthquake” experiment of multiple dynamic stress-oscillations samples accumulate inelastic strain. The strain imparted during each successive event is initially high and then reduces after 5-7 events, except when stressing results in failure, where strain increments increase and strain rate accelerates.”

Line 39: This is a great rationale statement – can you include another similar statement further up to explain why you have carried out your study or what the implications are?

Introduction

Line 47: How rapid is this growth? Can you quantify this?

Line 48: Is this magmatic or hydrothermal activity increasing or decreasing stability? Or both? Give more detail here

Line 57: “Manifested”? Or “embodied”? Can you simplify the grammar here?

Line 57: What sort of stress field changes? Not clear if these are caused by the earthquakes or by the deformation that has occurred. How do these dynamically triggered earthquakes occur? Give more detail here.

Line 60: What sort of volcanic environments? Of domes? I initially read this to just mean rocks in the areas around volcanoes.

Line 63: Can we understand the degree of damage and the likelihood and timing of failure? Maybe rephrase to “and” rather than “or”? Would also help to make it clear that these are all linked and help each other, as understanding the degree of damage will help understand the likelihood of failure etc.

Line 66: These authors are in alphabetical order rather than date order.

Line 74: Coalesce rather than coalescence.

Line 78: How slow is relatively slow? Can you give a range that is used and explain why we need faster rates too?

Line 80: The start of this sentence is a little clunky. “Investigations of cyclic/fatigue loading of rock” is slightly better?

Line 81: I don’t follow this sentence, can you rephrase?

Line 83: The frequency of events? Or of waves?

Line 86: Why is tensile failure of vital importance? Can you explain?

Lines 92-102: Can you just give the necessary detail about Unzen here and then go on to explain more in the Unzen section? The details on earthquake activity or the tsunami either need moving up, as examples in the introduction of failures, or moving down in the Unzen section. It’s a little confusing to go into so much detail here without some sentences properly introducing Unzen and where it is (like there are at the start of Section 2!).

Line 115: Steam in the foreground

Section 2

Line 135: Rephrase to “A GPS survey from 1996 to 1999 found the displacement parallel to the direction of the steepest slope was largest where the slope is the steepest, suggesting that the lobes were still hot and viscously deforming.”

Line 145: How are these causing instability? Can you explain this link here or in the introduction?

Line 149: It would be nice if this blocky nature could be better highlighted before now – is there an image in Figure 1 that could help? This is phrased as though the reader already knows it is blocky and this is a key thing to place in their minds. Mention further up e.g. around line 126?

Methods and Materials

Line 174: The rock samples?

Lines 176-196: It is difficult to take the different characteristics away from this paragraph, perhaps because they are also not easy to tell apart in the images in Figure 2. More detail is needed here to link your points to the experimental data later on in a compelling way. I would consider the density and size of microfractures, size and distribution of vesicles and phenocrysts and label these features in Figure 2. Including this information in a table would prevent you needing to add lots of extra words to the text while adding the evidence you need.

Line 179: What do you mean by thin? How big is that?

Line 189: Are these vesicles bigger than in the other samples? By how much? Are the distributed differently?

Line 191: How do these compare to the small subrounded vesicles you have mentioned in other samples?

Line 203: Can you annotate these features in the figure?

Line 207: How much smaller are they? Significantly so or just a little?

209: Are these short microfractures shorter than those mentioned elsewhere?

Line 222: Brazilian tests

Line 240: Edit so words all fit on one line without losing some of their letters to the line below. Separating the two experiment types with a thicker line might make the table easier to interpret.

Line 250: What are inter-earthquake periods? Is this needed here or can you just take this phrase out and simply mention characterising the samples?

Lines 256-260: Rephrase to “Specimens were kept beneath a sub-critical stress (as determined from creep experiments) to ensure that changes in strain were due primarily to earthquake loading and not brittle creep.” Or similar

Line 289: Format so words are not broken across multiple lines (as in Table 1)

Line 292: Steps rather than step?

Line 307: Why did you choose two minutes apart? Was there a rationale behind this?

Discussion

Line 436: Seems like this is a key point as it was mentioned in the Abstract, if so it needs mentioning in your results. Would be nice to give more of an explanation about it here too.

Line 441: What are these highly heterogeneous textures? And what do they mean for your results? That different parts will deform quite differently? What are the implications of this? How representative are your samples of this heterogeneity?

Line 444: And are these fabrics oriented unfavourably? What do you expect?

Line 469: Did you take a look at the samples afterwards to see how this sample did fail? It would be neat if it visibly followed some of these pre-existing microfractures that you have identified. How do you know it is the microfractures that are the key characteristic here?

Line 476: What is the importance of the pore geometry? The discussion would benefit from exploring this a little and that would nicely help highlight the implications of your study too.

Line 477: To rely on the sample characteristics here they need to be better described (and ideally with better images or quantification) earlier on to be compelling. Can you add the needed detail above to draw on here?

Line 478: Why is this? Can you give more explanation?

Line 480: Amounts rather than mounts.

Conclusion

Line 512: More detail is needed on this point and more explanation for you to turn it into a key point here. How does it mean the damage evolution varies?

Line 525: It would be nice to explore these insights more in your discussion.

Figure 1

Can you annotate some of the features mentioned e.g. the steam? This is difficult to spot!

Can you annotate the views seen in the photos in parts c and d? Maybe arrows or boxes to show the view direction or similar?

Figure 2

Can you add annotations of the features mentioned here?

Figure 3

The inset with the earthquake needs to be larger for text to be read easily. Text also looks like it might be blurry when enlarged.

Figure 6

So is the x-axis here effectively time, with 2 min holds between each event? Can you label it as that or as the event number?

Figure 7

I find these symbols difficult to follow as the red can be either a fill or a line/stroke and it takes a lot of thought to match them back to the creep data and Figure 6. Can you stick to shapes that you can just change to all red for the during oscillations data?

Comments by Reviewer 2

General comments

This manuscript presents experiments in which dacite dome samples from Mt Unzen (Japan) were held under a constant uniaxial stress and subjected to 21 small-amplitude stress cycles. These authors hope to better understand the impact of small stress perturbations, such as those during earthquakes, on volcanic rock, with implications for volcanic hazards. I do, however, have several issues with the experimental design and the interpretation of the data, outlined below.

Ignoring, for now, the experiment performed on UNZ1, one could quite easily conclude from these data that the stress cycling imposed on the samples does not particularly influence their damage state. Indeed, the strain accumulated following the stress cycling is extremely small. This is to be expected, however, since very-short-duration excursions at a stress only a couple of MPa above the load-hold stress (which is on-purpose at a low stress with respect to the UCS) would not be expected to impart much in the way of additional damage, if any. However, and throughout the manuscript, the authors use these experiments to reinforce their hypothesis that small stress perturbations, like those during earthquakes, can result in strain accumulation that can lead to failure. Although very small strains were likely accumulated during the stress-cycling in these experiments, I'm unconvinced that the magnitude of these strains, which must be close to the precision of the displacement transducers in some cases, warrants discussion of failure leading to collapse. Indeed, one could instead conclude from these data (all data excluding UNZ1) that small stress perturbations (when the hold stress is low) do not particularly damage rock and, therefore, that earthquakes may not disturb intact rock (held at a low differential stress) sufficiently to promote failure and collapse. For the experiment on sample UNZ1, the authors observe much more in the way of the accumulation of strain and, in the tensile experiment on sample UNZ1, the sample failed during the stress cycling. I suspect that much of the discussion centred on failure and collapse is because of these two experiments performed on UNZ1. However, and as discussed in my specific comments below, I think that the strain accumulation seen in these samples is very likely due to the fact that these samples were weaker than the "average" strength values reported from the earlier Kendrick et al. (2021) paper. In other words, and for example, I think that 60% of the average UCS (determined using three experiments) is much closer to the failure stress of sample UNZ1-3. In this case, the small-amplitude stress cycles likely bring the sample close to the failure stress, and thus impart much more in the way of damage. Therefore, I think the following. (1) That the stress cycles can likely impart strains of an "interesting" magnitude if the sample is held closer to the failure stress. Cycling the stress (only by a couple of MPa) at a low stress (compared to the failure stress) would be not expected to particularly damage the sample, which I think is what most of the data show (all the data apart from UNZ1). In my opinion, the larger strains seen during the cycling in UNZ1 is the consequence of sample-to-sample variability (something brittle creep is very sensitive to) which, in this case, likely meant that the sample was closer to failure than expected (i.e. at a stress higher than 60% of its failure stress). (2) Based on the above reasoning (also discussed below), I think the authors should remove all of the discussion about how microcracked samples are likely to be more damaged than pore-dominated samples following stress perturbations, which are said to "buffer" earthquake damage. I don't think that the authors can conclude this from these data, due to the low number of experiments on samples characterised by a high sample-to-sample variability. As stated above, I think that the behaviour of UNZ1 is the result of the fact that sample UNZ1-3 was just weaker than the average UCS, rather than because it contains more microcracks than the other samples. Many more experiments would be required to make such statements, in my opinion.

Implementing the changes described above would require a restructuring of the manuscript and, potentially, a couple of extra experiments that try to explore the idea that small-amplitude stress-cycling close to the failure stress can push the sample to failure, with interesting implications for the influence of earthquakes on the stability of volcanic structures (assuming that the duration and stresses of the "laboratory earthquakes" scales with natural earthquakes, see my comments below).

Specific comments

Brittle creep experiments have shown that water, and even water vapour in the air, can influence creep strain rates. It's very important therefore to, at least, describe how the samples were dried following their preparation (the ambient humidity of the room would also be welcome, if available). Ideally, the samples would have been vacuum-dried prior to experimentation.

Some of the strains reported in this study are very, very small. I think the authors should provide the precision of their measurement of displacement in the methods section of their manuscript.

It is likely that brittle creep will occur at any stress above the threshold stress required for dilatant microcracking, sometimes referred to as C'. This can be determined from load-to-failure uniaxial compressive strength tests using, for example, AE data. Because the authors are keen to keep their load-stress below that required for brittle creep, I think it would be very informative to provide C' under uniaxial conditions and discuss how far their load-stress is above/below this stress. In my opinion, this would represent an easier, and more robust, method to determine the stress required to "to generate irreversible damage in the form of creep" (Line 263) than performing a stress-stepping brittle creep experiment. For example, I suspect that 60% of the UCS is already above C' (although the strain rate at this stress is likely too low to observe in a 6-hour stress step). While this is not really a problem for the experiments presented, I think the authors should make it clear that it's very likely that brittle creep can occur at 60% of the UCS, because this stress is above C', but that it's not likely to progress at a rate that is particularly measurable in the laboratory (as shown in Figure 3a). However, and as outlined in my general comments, I think that it's much more interesting to perform experiments closer to the failure stress (and definitely above C'). In theory, creep and stress-cycling below C' would be elastic and recoverable.

During the load oscillation experiments, the stress was held for 30 minutes prior to the stress cycling. This duration was always sufficient to ensure that the strain rate had decreased sufficiently (the duration of the pronounced deceleration at the start of a creep test is a function of the strain/stress rate to that point, and the creep stress)? Looking at Figure 5, there appears to be an appreciable creep strain rate in the UNZ1 experiment (at least when compared to the other experiments).

The samples were subjected to 21 "earthquakes" during the step 3 of the load oscillation experiments. Was there a reason for choosing exactly 21 "earthquakes"?

What is the duration of an "earthquake" event? 14 seconds? How does this compare with the durations of natural events?

The maximum stress fluctuation of the "earthquake" events is 10% of the load-hold stress. For most of the samples, this is a couple of MPa. And how does this compare with stress perturbations caused by natural events? Is a couple of MPa not high compared to natural stress perturbations caused by earthquakes?

The authors should provide the average porosity and standard deviation for the samples presented in Table 2.

The authors present the "average strain rate" for their creep experiments in Figure 4. This average strain rate was calculated from the start of the load-hold step (as mentioned in the figure caption). The problem here is that the strain rate during a brittle creep experiment is, in fact, never constant, which questions the worth of an "average creep strain rate". The strain rate during a creep test will first decelerate to a minimum, before accelerating to failure. Therefore, the best (only?) way to compare creep experiments is to compare their "minimum creep strain rates" (that way, you're comparing like-with-like). Although stress-stepping creep experiments have their advantages, one disadvantage is that the strain rate at the end of a given step is only close to this minimum creep strain rate, and the proximity to this minimum depends on, amongst other things, the duration of the step. Comparing the "average strain rates" of these experiments is a little dangerous, because rocks with different porosities and strengths may well be characterised by different strain rate decelerations during the initial stages of a creep experiment (indeed, they are loaded to different creep stresses). As mentioned above, the duration of this deceleration may well depend on several factors. Therefore, there is a risk that the authors are not comparing like-with-like in Figure 4a and 4b. The strain rate at the end of the step, rather than the average strain rate for the entire 6 hours, is probably more informative. One way previous authors have circumvented this problem is to plot the evolution of the strain rate during the test, rather than a single “average” point, or to run a series to single-step creep experiments to obtain minimum creep strain rates.

What do the authors mean by “average strain” (Line 332)? Is this not the total strain accumulated during the 6-hour step? The strain should be increasing during the step: why have the authors calculated an average of this? Would it not be better to simply plot the total strain accumulated during the 6-hour step? This measure does, however, also suffer from the same issue as the strain rate (see my above point): the stronger rocks loaded to higher stresses will likely have longer-duration decelerations.

The authors discuss how their sample that contains more microcracks accumulated the most strain during the oscillations, while their sample that contains pores, and not many microcracks, accumulated the least. The implication, according to these authors, is that microcracked rocks will suffer more than pore-dominated rocks during earthquakes. I don’t think the authors can conclude this with these data, as mentioned in my general comments above. Brittle creep is highly sensitive to damage state. In previous studies of brittle creep, authors have prepared many samples and then removed those that do not lie within a very narrow porosity window. Even then, minimum strain rates can vary from sample-to-sample (see Figure 14b in the cited Brantut et al. (2013) paper). I notice, which is not unusual for volcanic rocks, that the porosities of the samples cored from the same block varies by quite a lot. This is true for the samples prepared for this study, and those described in Kendrick et al. (2021), experiments used to determine the average UCS of each block. When studying a process, I think its best to minimise factors that can influence the interpretation of the data, or perform a large number of experiments to compensate. Although the UCS of block UNZ1 is characterised by a small standard deviation, the porosities of the three UNZ1 samples used to determine the average UCS of the block were 20.4, 17.0, and 18.2% (from Kendrick et al., 2021). Although a similar UCS is indeed promising for the interpretation of the creep experiments, these data do underline a certain heterogeneity in samples cored from the same block of material. The other blocks, however, are characterised by much larger standard deviations in their UCS (Table 2). For example, the average UCS for UNZ9b is about 23 MPa, but the variation between three experiments was +/- about 12 MPa. My point is that, 60% of the “average” UCS is almost certainly different to the UCS of a given sample. Therefore, and although the authors waited 30 minutes to check that the creep strain rate was not appreciably different to that expected (although the experiment on sample UNZ1 appears to have a faster creep strain rate prior to the oscillation step than the other experiments), whether or not a certain sample fails, or not, during an oscillating creep experiment is much more likely, in my opinion, to be related to sample-to-sample variability. In other words, how much the UCS of that particular sample varies from the average UCS determined for the block. Therefore, I think it is very dangerous to draw conclusions such as “this suggests that vesicular materials may buffer earthquake damage more effectively than fractured materials” with these few experiments, given the problems discussed above.

The authors discuss strain rates during the oscillation step (step 3) and these data are presented in Figure 7. These strain rates were calculated using the data from the entire oscillation period? If so, are they not overestimated (since this average also includes the stress cycles)? Why was the “starting” strain rate calculated from the first 5-10 minutes rather than the full 30-minute period before the oscillation step? Why is the creep strain rate not compared before and after the oscillation step? In fact, it is remarked that “following the end of the oscillation period, the creep rate closely approximates the pre-oscillation rates” on Lines 387-388. Does this not suggest that little to no damage was accumulated during the oscillation step?

Line 356: “total strain increases in general with successive earthquake events”. The strain increases by about 0.00005 (Figure 6b). How many microns is that? Not only are these strains very small, but they must also be very close to the precision of the displacement transducer. The same comment applies to the compression experiments in Figure 6a. Apart from experiment UNZ1, which appeared to have a faster creep strain rate prior to the oscillation step (perhaps this sample is weaker than 17.7 MPa?), the strain accumulated during the other experiments is very small, about 0.0002. As outlined above, I think that low-amplitude stress perturbations from a low starting stress would not be expected to appreciably damage a sample.

Lines 411-413: But doesn’t Figure 4a show that the strain rate increases from 50 to 60%? In any case, I suspect that the strain rate does increase, but by an amount that is not easily measurable in the laboratory, especially at the displacements accumulated in only 6 hours.

The authors do not offer a reason why their samples are damaged by the stress oscillations. Microscopically, what is happening?

Authors' Reply to Reviewer 1

Section A: Overview of manuscript

A1) Overall evaluation, general comments & summary

A1.1) Reviewer's comments

A1.1.1) General evaluation and publication suggestion – Required:

Please use this space to describe, in your own words, the core subject of the submission and your overall assessment of its suitability for publication.

The manuscript describes experiments on a suite of samples from Unzen volcano to explore damage caused by earthquakes, weakening the dome structure and potentially leading to collapse events. The authors deform their suite of samples with well-considered load hold and oscillation experiments, designed to investigate how different stress fields and perturbations and sample characteristics might influence the degree of damage imparted on a sample.

The manuscript has many strengths. The experiments are well thought through and nicely investigate scenarios often not considered. The manuscript is well written and gives nice explanations of the method and results. I therefore recommend acceptance with moderate revisions.

A1.1.2) What does the submission need to be publishable? (select as needed; comment for all cases)

- ☐ No changes required
- ☒ Rewriting
- ☐ Reorganising
- ☒ More data/figures
- ☐ Condensing
- ☐ Reinterpretation
- ☐ Other

Comments:

Overall, the main ways the manuscript could be improved are by making providing more evidence for links made between the experimental data and sample characteristics, and by exploring the implications of these in more detail.

The sample characteristics are only described quite vaguely and more detail (ideally with quantification) would make the discussion of the data more convincing. I agree that the sample characteristics will have an effect on the damage imparted, but without more description of vesicle sizes/distribution, microfracture densities/lengths in the samples it is difficult to follow which characteristics are controlling the damage degree of damage imparted on each sample. The addition of a table may help here too for the reader to refer back to if the authors would rather not include this information in the text.

Re-writing to add detail to the text would improve the manuscript and ensure that each point is backed up by evidence. Much of the writing is clear but some sentences could be improved by rephrasing to simplify the sentence structure and keep terms consistent throughout.

Due to limitations of the study we have not included a thorough characterization of the microstructural components (see comment above in section A1). Both Coats et al. (2018) and Kendrick et al. (2021) have further details related to Unzen sample microstructures, which come from the same blocks as the ones used in our experiments, and are referenced in the text. We have added some details and annotations to Figure 2 to hopefully clarify the details being requested, in particular the distribution of the componentry of the rocks.

A1.1.3) Can the submission be improved by reducing/adding any of the following? (select as needed; comment for all cases)

- ☒ Text
- ☒ Table
- ☐ Figures
- ☐ Supplementary material

Comments:

Adding additional text that explores the implications of the results would improve the manuscript. What does it mean for hazards if the earthquake simulation experiments impart more damage on samples? This has been slightly explored but expanding these ideas would help. What else needs to be done in the future to expand on your experiments?

Please see response in section A2 below.

The introduction would benefit from more detail about how domes might be expected to fail (e.g. why is tension important?).

We have added the following lines to the introduction: "Volcanic domes can fail in both compression and tension. Compressional failures are related to gravitational forces on the dome interior due to dome growth, in addition to gas and magma pressure (Voight 2000). Tensile failure is most relevant to exterior parts of the dome, e.g., emplacement on slopes or uneven ground or bulging of a dome resulting in extensional stresses on the outer portions (Voight 2000), or pressurized gases within the pores that cause rocks to fail in tension (Sato et al. 1992; Hornby et al., 2019)"

Including a table with the information on sample characteristics would provide the reader with an easy way to see the connections between sample characteristics and the experimental data. These characteristics are difficult to see in Figure 2 but annotation might avoid the need for more images.

We have added annotations detailing the sample microstructural characteristics to Figure 2.

A1.1.4) Please complete the following section if you recommend that the submission is NOT appropriate for publication (select as needed; comment if a box is selected)

- ☐ Quality is poor
- ☐ Research is not reproducible
- ☐ Other

Comments:

[Free form box]

A2) Summary of main merits and main points of improvement

A2.1) Reviewer's comments

Please describe below in a few sentences (100 to 300 words) the main merits of the submission and suggestions for improvements.

The main merits I have found are...

This is a well-written manuscript that clearly explains the background and rationale behind the experiments completed. The methods are thoroughly written in a step by step manner and the data is quite well presented. The experiments are interesting and cleverly explore the potential of earthquakes to impart damage on samples, impacting dome stability.

The main points of improvement I have found are...

The manuscript would benefit from more detail about the sample characteristics as these are later used to explain the experimental results and would make these points more compelling. Inclusion of a table would also help here.

A greater exploration of the implications of these results would improve the manuscript as many interesting points have been made but only briefly. For example: Why does strain accumulation decrease after 5-7 events and what does that mean for dome stability? How and why does pore geometry affect the damage accumulation?

Our discussion has been modified to emphasize that although many samples didn't fail catastrophically, permanent damage in the form of increased inelastic strain shows that earthquake events can damage rock and promote instability over time in critical zones. Specially, we emphasize that fracture dominated samples, which cannot dissipate earthquake energy as efficiently as porous samples, are more prone to eventual catastrophic failure. Highly fractured rock masses can be identified at field-scale, and thus can define potential zones that would be more prone to catastrophic failure during an earthquake event. We also mention the use of deformation monitoring and nonlinear fatigue damage cumulative models, applicable for both constant and variable loading conditions to define the transition from a uniform velocity phase (creep) to an accelerated phase. Additionally, we have removed the discussions of pore geometry since these were not explored in detail.

Section B: Detailed evaluation of manuscript

B1) Title and abstract

B1.1) Reviewer's comments

These statements are a ***guide*** *to what good* Titles *and* Abstracts include*. Please select YES or*

NO to the statements below if you wish and detail in the free form box below your reasons for any box checked with NO, or to comment on any other matter.

The *Title* describes the main topic of the manuscript **accurately** — [YES]

The *Title* describes the main topic of the manuscript **succinctly** — [YES]

The *Title* includes **appropriate key terms** — [YES]

The *Abstract* includes a **clear aim and rationale** — [YES]

The *Abstract* supports the rationale with **sufficient background information** — [YES]

The *Abstract* includes a **well-balanced description of the methods** — [YES]

The *Abstract* describes the **main results sufficiently and adequately** — [NO]

The *Abstract* clearly describes the **importance/impact of the study** — [YES]

The *Abstract* clearly states the **conclusions of the study** — [YES]

The *Abstract* is **clear** and **well structured** — [YES]

Comments:

The title nicely fits the manuscript and the abstract gives a clear outline of the study.

In the abstract there are places where rephrasing could add more detail without increasing the abstract length. This is true for much of the manuscript where sentences can include quite vague language – see line by line comments.

The abstract gives a good explanation of the results but the key points you pull out could better match the contents of the manuscript. The abstract mentions vesicular materials buffering earthquake damage for example, which is not mentioned elsewhere. If you would like this to be a key point it needs to be mentioned more explicitly in the text, and it would be interesting to explore further in the discussion too. The orientation of fabric is also mentioned only very briefly in the main text – again to be a leading point in your abstract this needs to be mentioned in more detail.

Due to concerns raised by the other reviewer, the earthquake buffering statement has been removed from the abstract. We have modified the abstract to include more details about fabric orientation.

B2) Introduction

B2.1) Reviewer's comments

*These statements are a **guide** to what good* Introductions *include. Please select YES or NO to the statements below if you wish and detail in the free form box below your reasons for any box checked with NO, or to comment on any other matter.*

The *Introduction* provides **sufficient background and context** for the study — [YES]

The *Introduction* describes the **aim/hypothesis/rationale** clearly, providing **sufficient context** — [YES]

The *objective/hypothesis/rationale* **flows logically from the background** information — [YES]

The *Introduction* describes the study's **objective and approach** (last paragraph) — [YES]

The *Introduction* contains **relevant, suitable citations** — [YES]

The *Introduction* is **organized effectively** — [YES]

Comments:

This is a good clear introduction that really nicely sets the scene for the rest of the manuscript. Lots of background information is neatly tied together for the reader and the introduction flows logically.

The introduction contains quite a lot of information about Unzen and some of this might be better placed in the next section on Unzen itself, instead just mentioning the dome and its instability – see line by line comments.

Some details of Unzen have been removed, but we think some sentences in the introduction to set the stage of Unzen and emphasize the importance of the study should remain.

B3) Data and methods

B3.1) Reviewer's comments

*These statements are a **guide** to what good* Method *sections include and good practices for*

Dataset *accessibility. Please select YES or NO to the statements below if you wish and detail in the free form box below your reasons for any box checked with NO, or to comment on any other matter.*

The *Methods* are described **concisely and with enough detail** for reproducibility — [YES]

Necessary information about **data sources/acquisition/processing** is included — [YES]

Data used are accessible via either supplementary files or links in the data availability statement — [NO]

The *Dataset and/or Methods* are **organized effectively** — [YES]

Comments:

The methods are well explained in lots of detail.

There is no data availability statement in the manuscript.

This has been added.

B4) Results

B4.1) Reviewer's comments

*These statements are a **guide** to what good* Result sections *include. Please select YES or NO to the statements below if you wish and detail in the free form box below your reasons for any box checked with NO, or to comment on any other matter.*

The *Results* findings are **supported by data** — [NO]

The *Results* findings are presented **clearly and succinctly** — [YES]

The text in the *Result* section **cites tables and figures appropriately** — [YES]

The *Results* directly **relate to the study objectives** — [YES]

The *Results* present **data for all the approaches** described in the *Methods* section — [YES]

The *Results* **text belongs to the Results section**, not to *Introduction*, *Methods*, or *Discussion.* — [YES]

The *Results* section is **organised effectively** — [YES]

Comments:

The results section gives clear descriptions of the findings based on evidence from the data, citing figures appropriately. The results are organised in a logical manner.

There are some results e.g. UNZ9a and UNZ9b that are not specifically mentioned in the results section but are then discussed in the following discussion section. Can you include the results in the appropriate section and then explore them further in the discussion.

UNZ9a and UNZ9b were already mentioned in the load hold results section, which is what is being referred to in the discussion section that the reviewer has indicated. However, we have now added additional information about UNZ9a and UNZ9b in the oscillation results section too.

To make some of the links to sample characteristics made in the discussion and mentioned in the abstract and conclusion, the results section needs to contain more detailed/clearer descriptions of the sample characteristics such as vesicles and phenocryst sizes, density of microfractures etc to allow the reader to reach the same conclusions.

We have now added details of the microstructures as specified below in the line by line comments.

B5) Discussion and conclusions

B5.1) Reviewer's comments

*These statements are a **guide** to what good* Discussions *and* Conclusions *include. Please select YES or NO to the statements below if you wish and detail in the free form box below your reasons for any box checked with NO, or to comment on any other matter.*

The *Discussion* is **focused on the objectives** of the study — [YES]

The *Discussion* **addresses all major results** of this study, which are shown in *Results* — [NO]

The *Discussion* section makes **comparisons with other studies** that are relevant and informative — [YES]

The *Discussion* section properly identifies all **speculative statements** — [YES]

The *Discussion* section presents the **implications of the study** persuasively — [NO]

The *Discussion* section **highlights novel contributions** appropriately — [YES]

The *Discussion* section **addresses the limitations** of the study appropriately — [YES]

The *Discussion* section is **organised effectively** — [YES]

The *Conclusions* are **consistent** with and **summarise** the rest of the manuscript — [YES]

The *Conclusions* are **supported by the data** in *Results* and **follow logically** from the *Discussion* — [YES]

The *Conclusions* are **clear and concise** — [YES]

Comments:

The discussion includes a great section on the possible sources of repeated stressing of rock and the impacts that this might have on overall strength. The conclusions section gives a great outline of the study with writing that is more detailed than elsewhere.

More explanation and exploration is needed in the discussion section to be able to link the sample characteristics to the experimental data in a compelling way for the reader. Why do these characteristics control the damage evolution seen? What mechanisms are acting here?

It could be clearer what the implications of the study are. In the conclusion you state that your results have important implications for determining the damage accumulation of rock masses during stressing events and for monitoring volcanic stability – what are these implications? Can you give more detail?

See response in section A2.

B6) Figures, tables and citations

B6.1) Reviewer's comments

*These statements are a **guide** to what good* Figures *and* Tables *include and how they are presented. Please select YES or NO to the statements below if you wish and detail in the free form box below your reasons for any box checked with NO, or to comment on any other matter.*

Tables and *Figures* are **ordered logically** and **numbered sequentially** — [YES]

Tables and *Figures* have **captions that explain** all their major features — [YES]

Tables and *Figures* have **captions that complement** the information in the main text —

[YES]

Tables and *Figures* present data that **relate** to the study objective — [YES]

Tables and *Figures* present data that are **consistent** with and support the description of results — [YES]

Tables and *Figures* have **succinct and informative titles** — [YES]

Figures are **accessible** (elements are clearly labelled, accessible colour palettes, colour contrasts, font size legible, etc.…) — [NO]
Please, check our [Figure guidelines]

Figures with **maps or cross-sections** contain all **elements to be understood** (north arrow orientation, scale, visible coordinates, sufficient coordinate grid intercepts) — [YES]

Figures with **maps** have **sufficient location information** (in the map or caption) — [YES]

Cross-sections have clear labels for **scale and coordinates** at ends and within-section kinks — [N/A]

All georeferenced elements are provided in common format (.shp, .geotiff, .kml) [in an open-access repository] — [N/A]

Citations throughout are relevant, suitable, and comprehensive — [YES]

Comments:

The manuscript has nicely presented figures that display the data quite clearly. Figures are neat and well put together, and contain much useful information that adds to the manuscript.

Not sure if I'm missing something, but where is the data point for UNZ1 at 70% in Figure 4d? Is it perfectly hidden behind another data point, or had the sample failed at this point? Something needs adding to the caption here.

Yes, this is the case (the datapoint is partially obscured by another) and it has been added to the caption.

Altering some of the symbols used in figures and simplifying axes would help the data to be displayed more clearly – see line by line comments.

Figure modifications have been made where noted in line by line comments.

B6.2) Author's responses

[Free form box]

Section C: Additional comments

C1) Minor/line-numbered comments

C1.1) Reviewer's comments

Abstract

Line 24: Why is this rarely systematically measured? Can you add a few words to explain here?

Changed to: "However, the accumulation of damage in rocks under the frequency and amplitude of shaking that rocks experience during earthquake events in such scenarios is rarely systematically measured due to technical limitations"

Line 30: Are these samples that were already fracture-dominated or are these fractures gained during your experiments? Are you referring to microfractures or larger scale fractures?

Changed to "incipiently microfracture-dominated samples tended to be more susceptible to damage than vesicle-dominated samples"

Line 31: How do these buffer earthquake damage? It is a lovely idea that needs to be mentioned elsewhere if you want to include it in the abstract.

This has been removed from the abstract due to concerns from the other reviewer, though it is still discussed in the discussion section later.

Line 33: I agree that the direction of fabric will be important – can you be more specific about your findings here?

Changed to say "The orientation with respect to loading direction dictates the magnitude of strain accumulation under load oscillations."

Line 35: Can you rephrase to shorten this sentence or split into two? Such as "During each "earthquake" experiment of multiple dynamic stress-oscillations samples accumulate inelastic strain. The strain imparted during each successive event is initially high and then reduces after

5-7 events, except when stressing results in failure, where strain increments increase and strain rate accelerates."

Changed as suggested.

Line 39: This is a great rationale statement – can you include another similar statement further up to explain why you have carried out your study or what the implications are?

We have added the following in the third sentence: "Here, we characterize damage evolution during laboratory experiments on a suite of dacitic rocks from Unzen volcano, Japan to help inform accumulated damage and landslide susceptibility of lava domes during regional earthquake events."

Introduction

Line 47: How rapid is this growth? Can you quantify this?

Growth rate can vary hugely from volcano to volcano, and depending whether all the erupted lava accumulates or is shed during an eruption. For example at Unzen the dome accumulated volume at a rate of approx. 200m^3 per month, but also suffered hundreds of gravitationally driven PDCs (per month), such that the erupted volume is estimated at closer to 1000m^3 a month (see Yamamoto et al., 1993). We have added the Yamamoto reference in this line for reference.

Line 48: Is this magmatic or hydrothermal activity increasing or decreasing stability? Or both? Give more detail here

Modified as follows: "are subject to magmatic (Kerr, 1984) and hydrothermal activity (including mineral alteration; Ball et al., 2015) as well as earthquake loading (Belousov, 1995; Walter et al., 2007) that may decrease volcanic slope stability"

The references describe instability processes in detail.

Line 57: "Manifested"? Or "embodied"? Can you simplify the grammar here?

Modified.

Line 57: What sort of stress field changes? Not clear if these are caused by the earthquakes or by the deformation that has occurred. How do these dynamically triggered earthquakes occur? Give more detail here.

We have clarified the wording as follows: "This can lead to stress field changes around volcanoes (Farías & Basualto, 2020; Fujita et al., 2013) and can trigger additional earthquakes at volcanoes (Bell et al., 2021; Enescu et al., 2016)."

Line 60: What sort of volcanic environments? Of domes? I initially read this to just mean rocks in the areas around volcanoes.

Modified to say “at several volcanoes”

Line 63: Can we understand the degree of damage and the likelihood and timing of failure? Maybe rephrase to “and” rather than “or”? Would also help to make it clear that these are all linked and help each other, as understanding the degree of damage will help understand the likelihood of failure etc.

Modified as follows: “Understanding the degree of damage volcanic rocks may experience during earthquakes, which contributes to understanding the likelihood and timing of material failure, can aid in hazard mitigation at volcanoes”

Line 66: These authors are in alphabetical order rather than date order.

Corrected.

Line 74: Coalesce rather than coalescence.

Corrected.

Line 78: How slow is relatively slow? Can you give a range that is used and explain why we need faster rates too?

Modified as follows: “Typically, rock mechanics experimentation is conducted under relatively slow (e.g., typically 10^{-5} s^{-1} for compressive strength tests following guidance from ASTM, 2014) and/or constant loading rates, which do not reflect the rates of earthquake loading.”

Line 80: The start of this sentence is a little clunky. “Investigations of cyclic/fatigue loading of rock” is slightly better?

Modified as suggested

Line 81: I don’t follow this sentence, can you rephrase?

Modified as follows: “Investigations of cyclic/fatigue loading of rock, which apply constant or variable stress amplitudes, show that the accumulation of damage with each stressing cycle can eventually lead to failure (Cerfontaine & Collin, 2018).”

Line 83: The frequency of events? Or of waves?

“frequency range” has been added to clarify

Line 86: Why is tensile failure of vital importance? Can you explain?

We have now added details in the introduction that explain this as follows:
"Volcanic domes can fail in both compression and tension. Compressional failures are related to gravitational forces on the dome interior due to dome growth, in addition to gas and magma pressure (Voight 2000). Tensile failure is most relevant to exterior parts of the dome, e.g., emplacement on slopes or uneven ground or bulging of a dome resulting in extensional stresses on the outer portions (Voight 2000), or pressurized gases within the pores that cause rocks to fail in tension (Sato et al. 1992; Hornby et al., 2019)."

Lines 92-102: Can you just give the necessary detail about Unzen here and then go on to explain more in the Unzen section? The details on earthquake activity or the tsunami either need moving up, as examples in the introduction of failures, or moving down in the Unzen section. It's a little confusing to go into so much detail here without some sentences properly introducing Unzen and where it is (like there are at the start of Section 2!).

We have removed a sentence here to avoid too much background information.

Line 115: Steam in the foreground

Modified.

Section 2

Line 135: Rephrase to "A GPS survey from 1996 to 1999 found the displacement parallel to the direction of the steepest slope was largest where the slope is the steepest, suggesting that the lobes were still hot and viscously deforming."

Revised as suggested.

Line 145: How are these causing instability? Can you explain this link here or in the introduction?

The following has been added: "that could contribute to instability via mechanical weakening of material as primary minerals transform into weaker secondary products such as clay minerals (e.g., Ball et al., 2015)."

Line 149: It would be nice if this blocky nature could be better highlighted before now – is there an image in Figure 1 that could help? This is phrased as though the reader already knows it is blocky and this is a key thing to place in their minds. Mention further up e.g. around line 126?

Done.

Methods and Materials

Line 174: The rock samples?

Yes, clarified.

Lines 176-196: It is difficult to take the different characteristics away from this paragraph, perhaps because they are also not easy to tell apart in the images in Figure 2. More detail is needed here to link your points to the experimental data later on in a compelling way. I would consider the density and size of microfractures, size and distribution of vesicles and phenocrysts and label these features in Figure 2. Including this information in a table would prevent you needing to add lots of extra words to the text while adding the evidence you need.

Due to limitations of the study we have not included a thorough characterization of the microstructural components (see comment above in section A1). Both Coats et al. (2018) and Kendrick et al. (2021) have further details related to Unzen sample microstructures, which come from the same blocks as the ones used in our experiments, and are referenced in the text. We have added some details and annotations to Figure 2 to hopefully clarify the details being requested.

Line 179: What do you mean by thin? How big is that?

This descriptor has been removed.

Line 189: Are these vesicles bigger than in the other samples? By how much? Are the distributed differently?

We have added details about the pore size.

Line 191: How do these compare to the small subrounded vesicles you have mentioned in other samples?

We have added a note that the UNZ-13 vesicles are preferentially oriented local to large phenocrysts.

Line 203: Can you annotate these features in the figure?

Annotations have been added to Figure 2.

Line 207: How much smaller are they? Significantly so or just a little?

Size comparison details have been added.

209: Are these short microfractures shorter than those mentioned elsewhere?

Length details have been added.

Line 222: Brazilian tests

Modified.

Line 240: Edit so words all fit on one line without losing some of their letters to the line below. Separating the two experiment types with a thicker line might make the table easier to interpret.

Yes we agree that the words should all be one line. This is a “modifiable” table as requested by the journal, and which can be modified as described during typesetting.

Line 250: What are inter-earthquake periods? Is this needed here or can you just take this phrase out and simply mention characterising the samples?

We have clarified this.

Lines 256-260: Rephrase to “Specimens were kept beneath a sub-critical stress (as determined from creep experiments) to ensure that changes in strain were due primarily to earthquake loading and not brittle creep.” Or similar

Modified.

Line 289: Format so words are not broken across multiple lines (as in Table 1)

See response above.

Line 292: Steps rather than step?

Modified.

Line 307: Why did you choose two minutes apart? Was there a rationale behind this?

The two minutes was both similar to the rhythmic seismicity measured at Unzen by Lamb et al. (2015)- we have added this to the text. Originally, we had also hoped to compare the 20 inter-seismic periods, but these were too appreciably small to compare in the end.

Discussion

Line 436: Seems like this is a key point as it was mentioned in the Abstract, if so it needs mentioning in your results. Would be nice to give more of an explanation about it here too.

Because we did not conduct any post-experiment microstructural

We mention failure of this sample in the first paragraph of the results section. We do not have any additional post-failure data (e.g., thin sections) that can be added, which is why we support our postulation with data from Kendrick et al. (2021), which use specimens from the same blocks.

Line 441: What are these highly heterogeneous textures? And what do they mean for your results? That different parts will deform quite differently? What are the implications of this? How representative are your samples of this heterogeneity?

Comparing our sample porosity and banded nature to other Unzen lavas (e.g., Coats et al., 2018; Kendrick et al., 2021), we consider our samples to be representative of the heterogeneities at lab-scale. We mention the heterogenous nature of domes on a macroscopic scale to acknowledge laboratory experimentation limitations.

Line 444: And are these fabrics oriented unfavourably? What do you expect?

We suspect that given the lava flow direction and the dip of the slope on which the dome has been emplaced that spatial fabrics (shapes and arrangements of the components of a rock and the orientation of textures and structures) may be along-slope and thus oriented unfavourably from a landslide hazard perspective. However, this assumption is speculative so we would prefer it to not be included in the manuscript without further field observations which are not possible at this time due to the volcano's exclusion zone. We have now included a note about some in-situ shear banding that was documented by Wallace et al. (2019) (see their Figure 1g).

Line 469: Did you take a look at the samples afterwards to see how this sample did fail? It would be neat if it visibly followed some of these pre-existing microfractures that you have identified. How do you know it is the microfractures that are the key characteristic here?

Originally, we were hoping to use acoustic data as a proxy for damage pre and post experiments, however it was too noisy. By the time we established this it was difficult to produce thin sections due to the majority of authors changing institution and the laboratory being packed and relocated to another country. We do admit this is a fault in this study but we are fairly confident that the sample variability of the 5 suites of samples used would have made numerical quantification of textural differences impossible, however, we have refenced a recent paper by Lamur et al. (2022) that uses similar oscillation experiments but on a more homogeneous and repeatable rock and finds longer fractures and higher fracture density in the post-experimental andesite from Volcán de Colima.

Line 476: What is the importance of the pore geometry? The discussion would benefit from exploring this a little and that would nicely help highlight the implications of your study too.

This refers to the ability of pore space to accommodate the accumulating strain (by closing), and thus to buffer seismic energy. We have clarified this.

Line 477: To rely on the sample characteristics here they need to be better described (and ideally with better images or quantification) earlier on to be compelling. Can you add the needed detail above to draw on here?

We believe this is now addressed with the additions described above.

Line 478: Why is this? Can you give more explanation?

We have added the followed statement: “While a larger number of experiments are required to reduce uncertainty related to inherent sample-to-sample variability, buffering of seismic energy in unsaturated porous material due to pore compaction has been observed in other volcanic materials (Andrews et al., 2007; Lockner and Marrow, 2008)”

Line 480: Amounts rather than mounts.

Modified.

Conclusion

Line 512: More detail is needed on this point and more explanation for you to turn it into a key point here. How does it mean the damage evolution varies?

We have removed “pore geometry” from this list since we do not explore it thoroughly in the text, and have instead added “microfractures.” In the discussion we describe how microfractured samples experience more permanent strain, how porous samples experience less permanent strain, and how pore anisotropy (e.g., dense bands vs. not dense bands and their orientation) influence strength.

Line 525: It would be nice to explore these insights more in your discussion.

See response in A2 section above.

Figure 1

Can you annotate some of the features mentioned e.g. the steam? This is difficult to spot!

Done.

Can you annotate the views seen in the photos in parts c and d? Maybe arrows or boxes to show the view direction or similar?

Done.

Figure 2

Can you add annotations of the features mentioned here?

Done.

Figure 3

The inset with the earthquake needs to be larger for text to be read easily. Text also looks like it might be blurry when enlarged.

Done.

Figure 6

So is the x-axis here effectively time, with 2 min holds between each event? Can you label it as that or as the event number?

The caption has been modified to provide clarity.

Figure 7

I find these symbols difficult to follow as the red can be either a fill or a line/stroke and it takes a lot of thought to match them back to the creep data and Figure 6. Can you stick to shapes that you can just change to all red for the during oscillations data?

The symbols have been changed and the figure font size has been increased.

Authors’ Reply to Reviewer 2

Michael Heap

General comments

This manuscript presents experiments in which dacite dome samples from Mt Unzen (Japan) were held under a constant uniaxial stress and subjected to 21 small-amplitude stress cycles. These authors hope to better understand the impact of small stress perturbations, such as those during earthquakes, on volcanic rock, with implications for volcanic hazards. I do, however, have several issues with the experimental design and the interpretation of the data, outlined below.

Ignoring, for now, the experiment performed on UNZ1, one could quite easily conclude from these data that the stress cycling imposed on the samples does not particularly influence their damage state. Indeed, the strain accumulated following the stress cycling is extremely small. This is to be expected, however, since very-short-duration excursions at a stress only a couple of MPa above the load-hold stress (which is on-purpose at a low stress with respect to the UCS) would not be expected to impart much in the way of additional damage, if any. However, and throughout the manuscript, the authors use these experiments to reinforce their hypothesis that small stress perturbations, like those during earthquakes, can result in strain accumulation that can lead to failure. Although very small strains were likely accumulated during the stress-cycling in these experiments, I’m unconvinced that the magnitude of these strains, which must be close to the precision of the displacement transducers in some cases, warrants discussion of failure leading to collapse. Indeed, one could instead conclude from these data (all data excluding UNZ1) that small stress perturbations (when the hold stress is low) do not particularly damage rock and, therefore, that earthquakes may not disturb intact rock (held at a low differential stress) sufficiently to promote failure and collapse. For the experiment on sample UNZ1, the authors observe much more in the way of the accumulation of strain and, in the tensile experiment on sample UNZ1, the sample failed during the stress cycling. I suspect that much of the discussion centred on failure and collapse is because of these two experiments performed on UNZ1. However, and as discussed in my specific comments below, I think that the strain accumulation seen in these samples is very likely due to the fact that these samples were weaker than the “average” strength values reported from the earlier Kendrick et al. (2021) paper. In other words, and for example, I think that 60% of the average UCS (determined using three experiments) is much closer to the failure stress of sample UNZ1-3. In this case, the small-amplitude stress cycles likely bring the sample close to the failure stress, and thus impart much more in the way of damage. Therefore, I think the following. (1) That the stress cycles can likely impart strains of an “interesting” magnitude if the sample is held closer to the failure stress. Cycling the stress (only by a couple of MPa) at a low stress (compared to the failure stress) would be not expected to particularly damage the sample, which I think is what most of the data show (all the data apart from UNZ1). In my opinion, the larger strains seen during the cycling in UNZ1 is the consequence of sample-to-sample variability (something brittle creep is very sensitive to) which, in this case, likely meant that the sample was closer to failure than expected (i.e. at a stress higher than 60% of its failure stress). (2) Based on the above reasoning (also discussed below), I think the authors should remove all of the discussion about how microcracked samples are likely to be more damaged than pore-dominated samples following stress perturbations, which are said to “buffer” earthquake damage. I don’t think that the authors can conclude this from these data, due to the low number of experiments on samples characterised by a high sample-to-sample variability. As stated above, I think that the behaviour of UNZ1 is the result of the fact that sample UNZ1-3 was just weaker than the average UCS, rather than because it contains more microcracks than the other samples. Many more experiments would be required to make such statements, in my opinion.

Implementing the changes described above would require a restructuring of the manuscript and, potentially, a couple of extra experiments that try to explore the idea that small-amplitude stress-cycling close to the failure stress can push the sample to failure, with interesting implications for the influence of earthquakes on the stability of volcanic structures (assuming that the duration and stresses of the “laboratory earthquakes” scales with natural earthquakes, see my comments below).

Our experiments show that the fracture dominated samples have higher strain than those that are vesicle dominated, which is consistent with others who have emphasized the impacts of these physical properties on their mechanical behavior (e.g., Heap et al., 2014). However, we acknowledge that more testing would make this conclusion more robust and have limited discussions related to vesicle-dominated earthquake buffering to the discussion. Based on hesitations of the reviewer, we have also mentioned the possibility of unaccounted sample-to-sample variability and have emphasized that more work is required to confirm these speculations. We have also added two references which demonstrate seismic energy dissipation in porous volcanic material, which support our assertions (Andrews et al., 2007; Lockner and Morrow, 2008). The wording in the text has been modified as follows:

“Although a larger number of experiments would be useful to reduce uncertainty related to inherent sample-to-sample variability, we suggest that unsaturated porous material can more efficiently buffer seismic energy due to pore compaction, as has been observed in other volcanic materials (Andrews et al., 2007; Lockner and Morrow, 2008).”

We are not able to conduct further experiments due to funding limitations and changes in both laboratory and co-authorship locality. However, a recent publication by Lamur et al. (2023) subjected andesitic rock to stress oscillations under various stresses (60, 65, and 70%) and with variable amplitude stress fluctuations (2.5, 5, 7.5%). Overall, they found that the amplitude of the oscillations seemed to slightly correlate with the total amount of inelastic strain experienced by the rock, with samples subjected to larger amplitudes tending to record higher total accumulated strains after the final straining event. We have added a discussion of this point as follows:

“Our results reinforce the importance of microstructural variations (e.g., pores versus microfractures) on material strength and in defining a rocks’ susceptibility to damage accumulation. Highly fractured rock masses identified at field-scale could thus be used to indicate zones more prone to landslide susceptibility during earthquake shaking. Oscillation experiments by Lamur et al. (2023) on andesitic volcanic rock also show that the amplitude of oscillations slightly correlates with the total amount of inelastic strain experienced by the rock, with samples subjected to larger amplitudes tending to record higher total accumulated strains after the final straining event. Defining the oscillation stress amplitude and resulting damage to material at various depths (increasing confinement), and how microstructures control this damage, would be informative.”

Andrews, D. J., Hanks, T. C., & Whitney, J. W. (2007). Physical Limits on Ground Motion at Yucca Mountain Physical Limits on Ground Motion at Yucca Mountain. Bulletin of the Seismological Society of America, 97(6), 1771-1792. https://doi.org/10.1785/0120070014

Heap, M. J., et al. "Microstructural controls on the physical and mechanical properties of edifice-forming andesites at Volcán de Colima, Mexico." Journal of Geophysical Research: Solid Earth 119.4 (2014): 2925-2963.

Lamur, A., Kendrick, J.E., Schaefer, L.N., Lavallee, Y., Kennedy, B.M. (2023). Damage amplification during repetitive seismic waves in mechanically loaded rocks. Scientific Reports, 13(1), 1271. https://doi.org/10.1038/s41598-022-26721-x

Lockner, D. A., & Morrow, C. A. (2008). Energy dissipation in Calico Hills tuff due to pore collapse. In: AGU Fall Meeting Abstracts (Vol. 2008, pp. T51A-1856).

Specific comments

Brittle creep experiments have shown that water, and even water vapour in the air, can influence creep strain rates. It’s very important therefore to, at least, describe how the samples were dried following their preparation (the ambient humidity of the room would also be welcome, if available). Ideally, the samples would have been vacuum-dried prior to experimentation.

The following sentence has been added: “Following preparation, specimens were oven dried at 60°C for 24 hours and then placed into a vacuum chamber for one week prior to porosity and permeability measurements”

Some of the strains reported in this study are very, very small. I think the authors should provide the precision of their measurement of displacement in the methods section of their manuscript.

Load was recorded by an Instron Dynacell 2527 load cell at 100 Hz, which has an accuracy of ±0.1% of the full load capacity (100 kN). Strain is recorded using an Instron LVDT (Linear Variable Differential Transformer) Deflection Sensor, ±0.00001 mm or

±0.05% of the measured displacement, whichever is the largest. This information has been added to the manuscript.

It is likely that brittle creep will occur at any stress above the threshold stress required for dilatant microcracking, sometimes referred to as C'. This can be determined from load-to-failure uniaxial compressive strength tests using, for example, AE data. Because the authors are keen to keep their load-stress below that required for brittle creep, I think it would be very informative to provide C' under uniaxial conditions and discuss how far their load-stress is above/below this stress. In my opinion, this would represent an easier, and more robust, method to determine the stress required to "to generate irreversible damage in the form of creep" (Line 263) than performing a stress-stepping brittle creep experiment. For example, I suspect that 60% of the UCS is already above C' (although the strain rate at this stress is likely too low to observe in a 6-hour stress step). While this is not really a problem for the experiments presented, I think the authors should make it clear that it's very likely that brittle creep can occur at 60% of the UCS, because this stress is above C', but that it's not likely to progress at a rate that is particularly measurable in the laboratory (as shown in Figure 3a). However, and as outlined in my general comments, I think that it's much more interesting to perform experiments closer to the failure stress (and definitely above C'). In theory, creep and stress-cycling below C' would be elastic and recoverable.

Using the uniaxial data reported in Kendrick et al. (2021), we calculated C' to be between 43-57% of UCS using the stress at which the volumetric strain deviates from that observed from hydrostatic loading, as outlined in Heap et al. (2009). Thus, 60% of the UCS as used in our oscillation experiments is above C'. We have added a table to the supplementary information to report these values and have added the following text to the manuscript:

"Performing experiments at 60% of a sample's expected failure load is above the threshold stress required for dilatant microcracking, or where axial stress begins non-linearly decreasing as a function of axial strain (referred to as C'; Heap et al., 2009), as calculated from uniaxial compressive strength tests from Kendrick et al. (2021) (see Supplementary Table 1). This assures that the rock is above the point of elastic and recoverable damage, but below a brittle creep rate that progresses at a rate measurable in the laboratory test times"

Kendrick, J.E., Schaefer, L.N., Schauroth, J., Bell, A.F., Lamb, O.D., Lamur, A., Miwa, T., Coats, R., Lavallée, Y. and Kennedy, B.M., 2021. Physical and mechanical rock properties of a heterogeneous volcano: the case of Mount Unzen, Japan. Solid Earth, 12(3), pp.633-664.

Heap, M.J., Baud, P., Meredith, P.G., Bell, A.F. and Main, I.G., 2009. Time-dependent brittle creep in Darley Dale sandstone. Journal of Geophysical Research: Solid Earth, 114(B7).

During the load oscillation experiments, the stress was held for 30 minutes prior to the stress cycling. This duration was always sufficient to ensure that the strain rate had decreased sufficiently (the duration of the pronounced deceleration at the start of a creep test is a function of the strain/stress rate to that point, and the creep stress)? Looking at Figure 5, there appears to be an appreciable creep strain rate in the UNZ1 experiment (at least when compared to the other experiments).

The 30-minute hold was sufficient for the accelerated creep phase at the onset of the load-hold to slow to a negligible rate. This equilibration time was indeed longer for UNZ1, approximately 20 min, as seen in Fig. 5, which still falls within our hold threshold. This can be attributed to closure of fractures perpendicular to applied stress in this fracture-dominated sample.

The samples were subjected to 21 "earthquakes" during the step 3 of the load oscillation experiments. Was there a reason for choosing exactly 21 "earthquakes"?

Originally, we had hoped to compare strain rates between earthquake events, and 21 earthquakes would allow us to compare 20 inter-seismic periods. However, these were too appreciably small to compare in the end.

What is the duration of an "earthquake" event? 14 seconds? How does this compare with the durations of natural events?

This is a stacked waveform from natural earthquake events recorded at Unzen. Duration is not scaled, we simply scaled the amplitude to our desired stress fluctuation, as detailed in section 3.2.2.

The maximum stress fluctuation of the "earthquake" events is 10% of the load-hold stress. For most of the samples, this is a couple of MPa. And how does this compare with stress perturbations caused by natural events? Is a couple of MPa not high compared to natural stress perturbations caused by earthquakes?

Here we apply simulated "earthquakes" to lava dome rocks, these events were recorded at approx. 300m depth below the lava dome and therefore these stress differential fluctuations are appropriate for what these rocks may have experienced in a shallow (low confinement) setting.

The authors should provide the average porosity and standard deviation for the samples presented in Table 2.

These values have been added.

The authors present the "average strain rate" for their creep experiments in Figure 4. This average strain rate was calculated from the start of the load-hold step (as mentioned in the figure caption). The problem here is that the strain rate during a brittle creep experiment is, in fact, never constant, which questions the worth of an "average creep strain rate". The strain rate during a creep test will first decelerate to a minimum, before accelerating to failure. Therefore, the best (only?) way to compare creep experiments is to compare their "minimum creep strain rates" (that way, you're comparing like-with-like). Although stress-stepping creep experiments have their advantages, one disadvantage is that the strain rate at the end of a given step is only close to this minimum creep strain rate, and the proximity to this minimum depends on, amongst other things, the duration of the step. Comparing the "average strain rates" of these experiments is a little dangerous, because rocks with different porosities and strengths may well be characterised by different strain rate decelerations during the initial stages of a creep experiment (indeed, they are loaded to different creep stresses). As mentioned above, the duration of this deceleration may well depend on several factors. Therefore, there is a risk that the authors are not comparing like-with-like in Figure 4a and 4b. The strain rate at the end of the step, rather than the average strain rate for the entire 6 hours, is probably more informative. One way previous authors have circumvented this problem is to plot the evolution of the strain rate during the test, rather than a single "average" point, or to run a series to single-step creep experiments to obtain minimum creep strain rates.

We have changed the values provided in Figure 4 from the average strain and strain rate over the full hold period (6 hours) to just the final hour of the hold period (one hour, being the last hour), as suggested by the reviewer. The values have only changed slightly.

What do the authors mean by "average strain" (Line 332)? Is this not the total strain accumulated during the 6-hour step? The strain should be increasing during the step: why have the authors calculated an average of this? Would it not be better to simply plot the total strain accumulated during the 6-hour step? This measure does, however, also suffer from the same issue as the strain rate (see my above point): the stronger rocks loaded to higher stresses will likely have longer-duration decelerations.

This is the mean of the absolute strain during any given step, used to show how strain increased from step to step in the different samples rather than stacking all the stress-strain curves in one plot which was messy. We have added the line: "Average strain (or the mean absolute strain during a given step) and strain rate (or equivalent diametric strain for indirect tensile experiments) in each load-hold step for both compression and tension experiments are shown in Figure 4."

The authors discuss how their sample that contains more microcracks accumulated the most strain during the oscillations, while their sample that contains pores, and not many microcracks, accumulated the least. The implication, according to these authors, is that microcracked rocks will suffer more than pore-dominated rocks during earthquakes. I don't think the authors can conclude this with these data, as mentioned in my general comments above. Brittle creep is highly sensitive to damage state. In previous studies of brittle creep, authors have prepared many samples and then removed those that do not lie within a very narrow porosity window. Even then, minimum strain rates can vary from sample-to-sample (see Figure 14b in the cited Brantut et al. (2013) paper). I notice, which is not unusual for volcanic rocks, that the porosities of the samples cored from the same block varies by quite a lot. This is true for the samples prepared for this study, and those described in Kendrick et al. (2021), experiments used to determine the average UCS of each block. When studying a process, I think its best to minimise factors that can influence the interpretation of the data, or perform a large number of experiments to compensate. Although the UCS of block UNZ1 is characterised by a small standard deviation, the porosities of the three UNZ1 samples used to determine the average UCS of the block were 20.4, 17.0, and

18.2% (from Kendrick et al., 2021). Although a similar UCS is indeed promising for the interpretation of the creep experiments, these data do underline a certain heterogeneity in samples cored from the same block of material. The other blocks, however, are characterised by much larger standard deviations in their UCS (Table 2). For example, the average UCS for UNZ9b is about 23 MPa, but the variation between three experiments was +/- about 12 MPa. My point is that, 60% of the “average” UCS is almost certainly different to the UCS of a given sample. Therefore, and although the authors waited 30 minutes to check that the creep strain rate was not appreciably different to that expected (although the experiment on sample UNZ1 appears to have a faster creep strain rate prior to the oscillation step than the other experiments), whether or not a certain sample fails, or not, during an oscillating creep experiment is much more likely, in my opinion, to be related to sample-to-sample variability. In other words, how much the UCS of that particular sample varies from the average UCS determined for the block. Therefore, I think it is very dangerous to draw conclusions such as “this suggests that vesicular materials may buffer earthquake damage more effectively than fractured materials” with these few experiments, given the problems discussed above.

See reply above under “reviewer general comments” for our detailed response to this comment. With regard to the comments of checking for differences between samples during the 30min initial creep (“… although the authors waited 30 minutes to check that the creep strain rate was not appreciably different to that expected (although the experiment on sample UNZ1 appears to have a faster creep strain rate prior to the oscillation step than the other experiments)…”) – we only used this to check for intra-sample consistency, not compare one sample suite to another, thus UNZ1 consistently took longer to equilibrate than other samples, therefore UNZ1 samples were considered to behave consistently with one another.

The authors discuss strain rates during the oscillation step (step 3) and these data are presented in Figure 7. These strain rates were calculated using the data from the entire oscillation period? If so, are they not overestimated (since this average also includes the stress cycles)? Why was the “starting” strain rate calculated from the first 5-10 minutes rather than the full 30-minute period before the oscillation step? Because of equilibration time as he describes in the earlier comment. Why is the creep strain rate not compared before and after the oscillation step? Does this not suggest that little to no damage was accumulated during the oscillation step?

The strain rates are calculated during a stable 5- to 10-minute period of the first load hold step (just prior to the first oscillation) in the oscillation tests, after the initial equilibration of the strain rate (as discussed elsewhere).

Calculating the changes in strain rate between each oscillation was our original intention but the load hold was too short to observe any strain change in the inter-earthquake period (2 min), given the precision of our apparatus. We believe this may be an equilibrium issue and would potentially require longer holds between oscillations. As such, we opted to demonstrate the strain rate increase during the oscillation period by calculating the average strain rate during this time, which does indeed include the strain accumulated during oscillations. It demonstrates that strain rate is higher during intermittent stressing (with an average stress at the same value) than during constant stress. Figure 7 caption explains this calculation.

It is remarked that “following the end of the oscillation period, the creep rate closely approximates the pre-oscillation rates” on Lines 387-388. We expect that the strain rate had not fully decelerated (as pointed out by the reviewer) in the initial load hold period, and therefore that the first and last load hold periods therefore represent materials not in a comparable creep state. However, the increased accumulated strain after the tests seen in Fig 7e-f demonstrate that imparting of damage can be inferred in other ways than the pre- and post- oscillation strain rate.

Line 356: “total strain increases in general with successive earthquake events”. The strain increases by about 0.00005 (Figure 6b). How many microns is that? Not only are these strains very small, but they must also be very close to the precision of the displacement transducer. The same comment applies to the compression experiments in Figure 6a. Apart from experiment UNZ1, which appeared to have a faster creep strain rate prior to the oscillation step (perhaps this sample is weaker than 17.7 MPa?), the strain accumulated during the other experiments is very small, about 0.0002. As outlined above, I think that low-amplitude stress perturbations from a low starting stress would not be expected to appreciably damage a sample.

Strain is recorded using an Instron LVDT (Linear Variable Differential Transformer) Deflection Sensor, which has an accuracy of ±0.00001 mm or ±0.05% of the measured displacement, whichever is the largest. Therefore, these strain values fall within the instrument precision. We have calculated the standard error associated with both the load hold and oscillation measurements for the reported strain and strain rate, which are now available as supplementary material Tables S2 and S3. The error is appreciably small, therefore we are confident reporting these strain and strain rate results.

Microstructural investigations in Lamur et al. (2023) found, under similar laboratory conditions, and which induce similar total accumulated oscillation strain (e.g., see Lamur et al., 2023 Figure 5 which shows total accumulated oscillation strain on the order of 10^{-5} to 10^{-3}), that post-experimental andesitic rock had longer fractures and a higher fracture density. While post-experimental microstructural investigations were not conducted in this study due to the natural variability of our starting materials, we expect that our dacitic samples experienced similar damage in these low-amplitude stress perturbations.

Lines 411-413: But doesn’t Figure 4a show that the strain rate increases from 50 to 60%? In any case, I suspect that the strain rate does increase, but by an amount that is not easily measurable in the laboratory, especially at the displacements accumulated in only 6 hours.

The strain rate does not systematically increase from 50 to 60%, for example the strain rate for sample UNZ1 decreases substantially between these steps. Samples also do not systematically increase with each % stress increase- this is true for UNZ1, UNZ9a, UNZ13, and UNZ14. These tests demonstrate that over the span of our earthquake experiments, which were ~2 hours, the strain rate does not increase appreciably, thus allowing us to state that the strain rate changes in the oscillation experiments were indeed from oscillation events.

The authors do not offer a reason why their samples are damaged by the stress oscillations. Microscopically, what is happening?

We suspect, similar to Lamur et al. (2023), that the stress oscillations cause the development and propagation of cracks and/or pore compaction, but because we do not have microscopic data post-testing, we do not state explicitly that this is the effect. We have referenced this study in the discussion.

Recommendation: Revisions Required

2nd Round of Revisions

Decision Letter

Lauren Schaefer, Jackie E. Kendrick, Yan Lavallée, Jenny Schauroth, Oliver D. Lamb, Lamur Anthony, Takahiro Miwa, Ben M. Kennedy:

Thank you for taking the time to address the reviewer comments and for re-submitting your work to Tektonika. Associate Editor Noah Phillips and I have reached a decision regarding your submission to Tektonika, "Laboratory simulation of earthquake-induced damage in lava dome rocks". We will be delighted to publish your manuscript following some minor corrections:

Main comment 1: We request you place your experimental data (processed time series data is OK) into a data repository and add a link to a bespoke Data Availability section. One of Tektonika's publication requirements is 'reproducibility', meaning all elements necessary to understand, evaluate, replicate, and build upon the submitted work must be made available whenever possible; this is in line with FAIR data principles and the ethos of Tektonika.

Main comment 2: Please add an author contribution section.

Main comment 3: Regarding Reviewer 2's (Mike Heap's) comments, we understand your maintanence that stress cycling influences damage state based on your results. We also appreciate the reviewers comments that the number of analysed samples is limited and their possible pre-existing heterogeneity compared to previous benchmarked tests restricts interpretation. We can see that your additions highlight this uncertainty, but throughout the manuscript we would request that this is also reflected in the language of the entire text; specifically, we ask you to 'tone down' the study findings. For example, Line 800 could be rephrased to,

"*Our results suggest fractured-dominated samples may be more susceptible to oscillation-induced strain than vesicle-dominated samples in both compression and tension.*"

Very minor line-by-line comments:

Line 21: experience should be experienced

Line 58: needs an "are" - and ***are*** subjected to…

Line 99: spelling – prompting

Line 143: missing period and space after 1995

Line 269: missing “to” – rock subjected ***to*** prolonged loading

Line 486: replace “such as” with “which”

Following these minor corrections, we will be happy to publish your article.

Kind regards,

Craig and Noah

Authors’ Reply

31 March 2023

Dear Dr. Magee and Dr. Phillips,

Thank you for the opportunity to publish in Tektonika pending minor comments. Please see responses below.

____

Main comment 1: We request you place your experimental data (processed time series data is OK) into a data repository and add a link to a bespoke Data Availability section. One of Tektonika's publication requirements is 'reproducibility', meaning all elements necessary to understand, evaluate, replicate, and build upon the submitted work must be made available whenever possible; this is in line with FAIR data principles and the ethos of Tektonika.

I have created a data release for the manuscript in the USGS Science Base system. There is an internal review process for these data releases, and I am expecting a two-week turnaround time. The links are not active until the data release has been officially released, but a doi has been created and this has been added to the acknowledgement section (doi:10.5066/P94CMZVO). The data release will have .txt documents of time, stress, and strain for each of the 10 load hold and 10 oscillation experiment data for each of the specimens described in Table 1. A screen shot of the pending data release is available at the end of this letter, and I have additionally provided all the .txt files that will be provided in the data release.

Main comment 2: Please add an author contribution section.

An author contribution section is now available in the manuscript. We have also added an additional acknowledgement that was missing from the first revision.

Main comment 3: Regarding Reviewer 2's (Mike Heap's) comments, we understand your maintanence that stress cycling influences damage state based on your results. We also appreciate the reviewers comments that the number of analysed samples is limited and their possible pre-existing heterogeneity compared to previous benchmarked tests restricts interpretation. We can see that your additions highlight this uncertainty, but throughout the manuscript we would request that this is also reflected in the language of the entire text; specifically, we ask you to 'tone down' the study findings. For example, Line 800 could be rephrased to,

"Our results suggest fractured-dominated samples may be more susceptible to oscillation-induced strain than vesicle-dominated samples in both compression and tension."

The wording has been changed as suggested.

Very minor line-by-line comments:

Line 21: experience should be experienced

Line 58: needs an “are” - and are subjected to…

Line 99: spelling – prompting
Line 143: missing period and space after 1995
Line 269: missing "to" – rock subjected to prolonged loading
Line 486: replace "such as" with "which"
These comments have all been revised as suggested.

____

I will remain in touch as the data release internal review process proceeds.

Best,

Lauren Schaefer