Response to Referee 1

In manuscript 'Coevolving edge rounding and shape of glacial erratics; the case of Shap granite, UK', the author reports on a field study that attempts to understand the effects of plucking, abrasion/attrition and fracturing during (sub)glacial transport. The study considers the complexities of the dynamic equilibrium between clast shape and edge roundness as a function of distance from the Shap granite rock outcrop. The author discusses many interrelated aspects of the processes that determine size as well as shape and edge roundness on the erratics and offers theoretical insights into why and how these develop with increasing transport distance.

The approach taken by the author is comprehensive and models of fracture mechanics are employed to predict what could happen to blocks of granite after being plucked and transported by the ice away from the outcrop. The theoretical considerations appear sound and plausible, from the perspective of shape retention as well as mean edge roundness versus distance.

Thank you for noting the approach is comprehensive, sound and plausible.

My main issue with the work, which seems otherwise carefully designed (and the manuscript well written), is that the data from the field do – in my opinion - not always convincingly match those that would be expected from the theoretical considerations. The main issue here – I think – is with the author's stated assumption that subaerial weathering after erratic deposition is negligible. Looking at the inset diagram in Figure 4, I do not think that assumption is justified. The edge rounding at the outcrop location is in the order of 50% of the values from blocks that were transported between 5-10 km. At least part of this must be attributed to subaerial, i.e. post depositional, weathering, I'd say. This view would be supported by looking at the field photograph of Fig. 6. It shows an erratic essentially split in half. The fragments remain in situ, so there is no apparent post-fracture transport in a subglacial traction zone. Still, looking at the 'fresh' edges, they seem remarkably rounded.

Thank you for noting the project was carefully designed. The issue of post-glacial weathering is an important one and I have made changes to the manuscript to acknowledge that fact. Edge rounding close to the outcrop was rapid as the newly fractures edges would have been sharp (typically 90^{o}), as is evident if a block of Shap granite is fractured by hammer. However, boulders rarely display prominent phenocrysts which fact indicates a very slow rate of loss of mass due surface weathering. Many boulders exhibit very smooth or polished surfaces, apparently due to ice action. Occasionally shallow striations are preserved. Weathering must contribute to edge rounding but as rounding does increase away from the outcrop over a relatively short distance, it is reasonable to ascribe the rounding to glacial transport. Note that at line 410 it was stated 'many blocks close to source initially exhibited near right-angle edges (Fig. 4)' which is incompatible with a considerable degree of weathering.

The author refers to a study by Parsons and Lee (2005) to justify the assumption, but I am not sure if that work does actually allows this. The paper says something about the texture as well as the composition (feldspar) as important factors in weathering potential, but – the way I understand it – it does not say that weathering of this particular granite on this timescale may be considered negligible.

Parsons and Lee (2005) were focussed primarily on the chemical weathering processes of Shap granite and not on the geomorphological implications. The reviewer is correct that Parson and Lee (2005) do not specifically state that the Shap granite surfaces overall are slow to weather. However, they obtained surface grus samples close to the Shap granite outcrop and they also examined specimens taken from an ice-sculpted exposure. On the specimens, they noted that the large phenocrysts generally stand out, at best, a few mm above the rest of the rock. They also calculated that the albite surfaces (which are prone to weathering) had retreated at around 1micron yr^{-1} since deglaciation, leading to the protrusion of the more resistant phenocrysts to potentially produce a weakened surface structure (susceptible to weathering) consisting of feldspars. However, they expressed surprise at the slow rate of surface retreat. This information together with the authors own observations of the height of protrusion of feldspar crystals indicate that post-glacial surface weathering, although present, is not a substantive factor in the loss of rock mass from the Shap granite boulder surfaces. Very sharp edges must have lost a little material due to surface weathering (see reply to previous comment) but, as stated above, the trend in edge rounding over a relatively short distance (and hence no climatic gradient) can best be related to glacial transport.

In fact, in the present submission the author himself later uses biotite mineral weathering to explain some of the edge rounding characteristics of the Shap erratics. Furthermore, some granites - albeit not necessarily Shap – are known to weather to grus rapidly (in fact, this is a particular problem for archaeological monuments, and several papers are dealing with this phenomenon).

The reviewer is mistaken here. I did not use the weathering characteristics of biotite to explain edge rounding. I indicated that biotite has a relatively low Moh hardness in contrast to the hardness of the feldspar crystals and this might contribute to the enhanced edge-rounding and a low tensile strength of the granite. Some other granites might readily weather to a grus. However, grus and flakes are remarkably absence (in any quantity) in the study area. You certainly cannot find small ridges of grus around boulders, which would be the case if surface weathering was prominent.

Figures 4 and 7 are the main results figures and show clouds of measurements of edge roundness and shape that are not always aligned with the theoretical characteristics. Using envelopes and arrows in the figures, the reader is led to certain inferences and conclusions, but I have to say that I do not always find them convincing. Some of the clouds are – in my view (unless I miss the point) – not sufficiently clustered to draw any firm conclusions. I can see weak relationships based on the distribution of data points, but they are perhaps not as obvious as the author claims. I am not sure if this is related to the choice of presentation. For a start, I found it quite difficult to match an increasing distance from the source with a particular cloud of data points. Perhaps the author can give a bit more guidance here by better colour-coding (or redesigning) and cross-referencing with Fig. 3. I am not sure if the chosen diagrams are the most meaningful when trying to analyse the relationships between shape and edge roundness as a function of distance.

Due to fracture, edge-rounding data are inevitably subject to scatter, and this is a basic tenant of the manuscript. Nonetheless, the trends in Fig. 4 are significant and are interpretable, as is summarized within the inset. I have rewritten the text and redrafted the edge rounding Fig. 4 to make the inferences clearer by changing the data point colour scheme and by adding the distances for each data set. The reviewer questions if the figures are the best way to present the data, but s/he does not suggest any alternative. I chose a way of plotting the data in Fig. 4 that spreads it out, such that the inevitable scatter does not clutter the image. The Zingg diagram in Fig. 7 is often regarded as the best way to display data of this nature (Howard 1992). The Hofmann (1994; 1995) or the Sneed and Folk (1958) triangular diagram is an alternative method and the Hofmann diagram is presented in the Supplementary Information.

Hofmann, H.J., 1994. Grain-shape indices and isometric graphs. *Journal of Sedimentary Research*, A64, 916-920, 1994.

Hofmann, H.J., 1995. Grain-shape indices and isometric graphs – Reply. *Journal of Sedimentary Research*, A65, 721-723, 1995.

Howard, J.L., 1992, An evaluation of shape indices as palaeoenvironmental indicators using quartzite and metavolcanic clasts in Upper Cretaceous to Palaeocene beach, river and submarine fan conglomerates: Sedimentology, v. 39, p. 471–486.

Sneed, E.D., Folk, R.L., 1958, Pebbles in the lower Colorado River, Texas—A study of particle morphogenesis: Journal of Geology, v. 66, p. 114–150.

I am also not fully convinced of the argument with regard to tensile versus compressive stresses. The calculations about shear stresses at the base of the ice is arguably a bit simplistic, and the author’s use of terminology (e.g. subglacial traction versus lodgement, see Evans et al, 2006) is not entirely up to date. With regard to forces on erratics in a subglacial environment, for example, it would make a difference if the size of the block was such that it could be contained entirely in the traction zone. In other words, if the deforming layer is thick relative to erratic size, I imagine that tensile forces may be dominant. However, I can also imagine that compressive forces are more dominant if the thickness of the traction zone is smaller than the size of the erratic. If the traction zone is relatively constant in thickness over significant distances under the glacier, it would seem that the shift from compressive to tensile dominance could be a function of distance, assuming a progressive reduction of erratic size. This could have implications for the model of fracture dynamics (stochastic versus silver ratio). This could also have implications for the way edge rounding coevolves. I would find it harder to envisage how edge rounding could then be increasing exponentially with transport distance, for example.

In the original submission I referred to 'subglacial traction' as an active process and to 'lodgement till' as a product. I accept the referees point that the terminology is not up to date and have relabelled 'lodgement till as 'subglacial traction till' after Evans *et al*., 2008 and added the reference.

The reviewer indicates the calculation of shear stress is simplistic. This section was included because I felt it would be inevitable that a reviewer would want some assurance that the boulders had been transported by ice. Without reverting to a complex ice flow numerical model for the region, it seems appropriate to make use of a simple example to illustrate key principles. Anything come complex verges on fantasy, as there are no data to justify such detailed analysis. I have added a line to explain at 564-566: "Although a more complex and complete appreciation of the stress environment of a boulder would be preferred, a simple force balance is utilized instead. Simplicity is dictated by the absence of data to inform a more complex model."

The reviewer is thanked for bringing to my attention a degree of potential confusion re the compressive force which might dominate if the thickness of the deforming layer is small compared to the size of the boulder. Tensile forces will likely dominate if the thickness of the deforming layer is large compared to the size of the boulder. In the original submission I ignored this distinction for two reasons. Firstly, primarily I wished to note that the tensile strength is much less than the compressive strength and, as many boulders are fractured at a right angle to the inferred direction of ice movement, reference to tensile stress was appropriate. Secondly, given the nature of the data and the absence of any information on the thickness of the deforming layer, venturing into such subtle considerations seemed a step to far. That said, I have revised manuscript to note the issue with respect to the thickness of the deforming layer (lines 551-556): "If the thickness of the deforming ice layer at the basal boundary is small relative to the size of the boulder, then the compressive force is likely to dominate. However, if the converse applies then the tensile force likely will dominate. Herein, given that there is no information as to the thickness of the deforming layer, the distinction is not considered because, in most cases, blocks will fracture at a lower stress due to tension in contrast to compression."

So, all in all, I found the careful discussion of the theoretical outcomes of this field experiment convincing, but I think there is a flaw in one of the main assumptions, i.e. that subaerial modification of the erratics is negligible, which essentially meant that the data are not always aligned with the theory. Then there is a consideration to be made about the consequences of tensile versus compressive forces in how the fracture process develops. Also, the way the results are presented in diagram is perhaps not the most efficient and user-friendly.

The points made here have been addressed at other places within this reply to reviewer. However, it should be noted that although data do not always 'align with theory', this fact is due to the inevitable difficulty of sampling and measurement rather than the vagaries of subaerial modification. Perhaps theory is in need of revision.

I think the study has merit, but there are some issues that would need addressing before it can be published.

Thank you. I trust the revised manuscript answers your queries.

Response to Referee 2

The author has conducted a thorough field-based analysis of erratic morphology (edge rounding-shape) to investigate how blocks evolve as a function of glacial transport distance. The paper is based on sufficiently large dataset of samples, and the analysis is generally well-written and structured. These data are linked to theoretical models of block evolution (stochastic, silver ratio), and are interpretated considering the relevant literature.

Thank you for your overall positive comments.

The main conclusions appear sound, while the evidence for a clear but irregular relationship between transport distance and edge roundness is particularly strong. However, the analysis and presentation of the data could be improved in a few key places.

Thank you for noting that the conclusions are sound and indicating that some improvements in analysis and presentation would enhance the publication.

The exclusion of **subaerial weathering** (Line 132) as an influencing factor could be justified more fully. There are a range of studies quantifying surface weathering of granite boulders (e.g., using the Schmidt hammer). Given the timescales of exposure (c. 20 ka), I would be surprised if losses were negligible, particularly given the presence of blocks with exposed phenocrysts (Line 488). Kirkbride (2008) "*Boulder edge-roundness as an indicator of relative age: A lochnagar case study*" is also relevant here, as this work correlates the degree of edge rounding with the duration of subaerial exposure, which suggests that transport distance/duration might not be the only control on your data.

The relevant text is at line 132. I am aware of the range of studies using the Schmidt hammer to quantify surface weathering, but no substantive study has been conducted on Shap granite. I also was surprised by the apparent absence of surface weathering on most of the Shap granite boulders I examined by eye (*c.*, 10,000). Spalling of surfaces due to weathering and local deposits of grus around boulders are almost non-existent. Weathering of Shap granite does not seem to be controlled by such factors as altitude, exposure or burial. Rather, it seems to be controlled by subtle differences in composition of the rock. A detailed study of that control is outwith the focus of the submission. However, the issue of post-glacial weathering is an important one and it is worth elaborating here. Edge rounding close to the outcrop was rapid as the newly fractures edges are sharp (typically 90^{o}), as was reported in the original submission and as is evident if a block of Shap granite is fractured by hammer. However, boulders rarely display prominent phenocrysts, which fact indicates a very slow rate of loss of mass due surface weathering. Many boulders exhibit very smooth indeed polished surfaces, apparently due to ice action. Occasionally shallow striations are preserved. Weathering must contribute to edge rounding but as rounding does increase away from the outcrop over relatively short distances, it is reasonable to ascribe the rounding to glacial transport. Note that at line 410 it was stated 'many blocks close to source initially exhibited near right-angle edges (Fig. 4)' which is incompatible with a considerable degree of weathering.

Parsons and Lee (2005: cited in the original submission), were focussed primarily on the chemical weathering processes of Shap granite and not on the geomorphological implications. Parson and Lee (2005) do not specifically state that the Shap granite surfaces overall are slow to weather. However, they obtained surface grus samples close to the Shap granite outcrop and they also examined specimens taken from an ice-sculpted exposure. On the specimens, they noted that the large phenocrysts generally stand out, at best, a few mm above the rest of the rock. They also calculated that the albite surfaces (which are prone to weathering) had retreated at around 1micron yr^{-1} since deglaciation, leading to the protrusion of the more resistant phenocrysts to potentially produce a weakened surface structure (susceptible to weathering) consisting of feldspars. However, they expressed surprise at the slow rate of surface retreat. This information together with the authors own observations of the height of protrusion of feldspar crystals indicate that post-glacial surface weathering, although present, is not a substantive factor in the loss of rock mass from the Shap granite boulder surfaces. Very sharp edges must have lost a little material due to surface weathering but, as stated above, the trend in edge rounding over a relatively short distance (and hence no climatic gradient) can best be related to glacial transport.

Although I am familiar with Kirkbride (2008), I did not cite it as it is not a study of Shap granite and surface weathering does not seem to be that important on Shap granite as it appears to be on some other granites. However, I have now cited Kirkbride (2008) in the Method to bring it to the reader's attention.

In turn, could the scatter in Figure 4A, and the statistical overlap of the site values (inset) represent the combined effect of re-setting (parent to child block) *and* subaerial weathering, rather than the former alone?

The scatter is predominantly due to 1) the difficulty of sampling a potentially large population of boulders within a reasonably prescribed area and 2) resetting by fracture as the referee notes. Point 2 was already emphasized in the manuscript at line 413-416. The original manuscript contains the statement: "However, fracture away from the outcrop introduces new sharp edges (Figs. 4 & 7), such that larger radii characterizing an individual edge-rounded block just before fracture are augmented by smaller radii. This change is reflected in the scatter of radii values found with increased distance from the outcrop (Fig. 4)." This existing statement seems to answer the referees' question. Subaerial weathering over such relatively short distances (where climate and other environmental controls do not change much, nor change systematically) cannot explain the spatial trend in rounding, although it may contribute in some small way to some of the scatter. This latter point has been made in the revised script of the Discussion where it is acknowledged that weathering may contribute in a minor fashion to the edge-rounding; inserted at line 387: "which include a small degree of weathering." and at line 397 " from glacial wear, as well as a little post-glacial weathering."

While I found the main conclusions plausible, the **presentation of the results**, in particular Figures 4 and 7, could be improved as follows:

- Both Figures 4 and 7 use a discrete colour for each site, which makes it difficult to interpret distance-related trends. An alternative approach would be to use a sequential colour scale (based on distance to source) which might be more intuitive and would give more information to the reader. This would also address issues for colour blind readers (green − red).

  Because there is a lot of overlap between groups, a sequential colour scale does not discriminate the different groups very well whereas discrete colours do. The red symbols have been changed and the figure coloured checked for colour-blindness.

  The curves shown in Figure 4 are described as "linear", but given the dependent variable is a function of the square of the independent (h, l), should these not be described as non-linear?

Only a sub-group of curves in Fig. 4B are empirical linear fits between h and r_c and are not intended to show the functional relationship between the two parameters, which (of course) are not independent. The functional relationships are shown as non-linear functions. Rather, given the data scatter and the close plotting of different groups of data, the linear functions are included to help the reader see the increasing value of r_c for a given value of h as distance increases.

- The subplot in Figure 4A is a key result, addressing one of the main hypotheses, so it might be presented more clearly and powerfully as its own figure, rather than as a subplot of another. The error bars should also be described here – SD, SE, MAD?

  Although a key hypothesis was "*Sg ice-transported blocks would display systematic changes in edge-rounding and shape"* the trend in mean edge rounding is not a result I wish the reader to focus upon. The inherent variability is largely lost within such plots as the inset of Fig. 4. Previously too much attention has been afforded such plots, so I purposefully included it as an inset to focus attention on the variability that often underlies such functions, as shown in Fig. 4A and 4B. The error bars are s.d. and this information was in the caption of the original submission.

- A proportion of the discussion in Section 3.2 is based on the central tendency values shown in Figure 7. While there are some issues with presentation here (e.g., Blasterfield is listed twice in the legend; use of colour, as noted above), I am not overly convinced by the strength of the arguments, given the scatter of the data, and the use of central tendency values without consideration of uncertainty. Are these mean or median values? How sensitive are they to the outliers? A better approach would be to use S/M and M/L standard deviation to produce hotspots or envelopes for each location, rather than a single value.

The legend has been revised and the colours changed to align with colour blindness. The central tendency values are mean values. Given the broad spread of data for each location, plotting sub-samples showed that the central tendency values are not overly affected by outliers. I was unsure what the reviewer wished to see in terms of the standard deviations for different data groups. Adding further detail to the plot would make it confused and the standard deviations cannot be plotted on a Zingg diagram, as detailed below.

The purpose of the figure is to demonstrate, as is explained in the main text, that despite transport, there is a distinct central tendency to the shape of Shap boulders. The principle of the present approach is to objectively identify extremes of shape that nevertheless remain statistically representative of both the sampled population and the shapes. To keep the paper succinct, I have done this only for the regional data using the method of Oakey et al., 2005, and the 95% contour is plotted. Shapes outside of this circle can be regarded as statistically not representative of the overall spread in shapes. Once an extreme value is determined, all three dimensions of the representative boulder at that point can be determined (not considered in the current manuscript). On the Zingg diagram, plotting of individual points, including the mean are meaningful values whilst the standard deviation values are not (Oakey *et al*., 2005 cited in main text).

Minor comments

Line 65-66: Consistent formatting 1) or 2: ?

Done

Figure 1: The base map is from OS Map, but what is the source for the granite pluton? Is this derived from original mapping, from the BGS, or another source?

Thank you noting this omission. The source (BGS) has been added to the caption.

Line 101-102: Phrasing could be a little clearer. "This analysis is focused on erratics with easterly transport vectors, defining …"

The sentence in question is "The focus solely is on those erratics the final transport vectors (direction and distance) of which are roughly due east, defining a simple linear direction over which changes in the nature of the erratic populations might be measured." I have inserted (red text) an indication of the two values determining the vectors and inserted 'roughly' as there is naturally some spread of the vectors around 90° E.

Figure 3: Map design-information could be improved here. Source for erratic locations (subplot A) should be listed in the caption. The location of the source pluton is appropriately generalised as a point in subplot A, but is less so for the larger scale subplot B. A better approach would be to use the geometry shown in Figure 1. The extent of the sublots (B, C, D, E) should also be shown in the small-scale plot (A), while alternatives base maps (e.g., topography?) might be more useful than OS data.

Figure 3 has been redrawn and the caption modified to address the comments. Placing the sampling locations on panel A would result in crowding. Rather I have added latitudes and longitudes so the reader can find the approximate locations. I have not changed the base maps. Although I considered a topographical map the contours obscure, and NEXTMap images do not allow the reader to quickly locate 'themselves' with reference to towns and roads.

Line 109: second "complexity" can be removed.

Deleted

Line 133: The hypothesis and aim could be combined here "… systematic changes in edge-rounding and shape as a function of distance from the pluton".

I have kept the hypothesis and the aim separate, as I wish to explore coevolution, which is not within the hypothesis the referee proposes.

Line 164: Should *Method* be capitalised here?

This statement refers to the Method section below, so capitalization seems appropriate.

Line 180 – 181: Although the study is based on a good selection of sites and individual samples, the paper would be improved with clearer justification for the selection of sites, and particularly their distribution. Given the presence of Sg erratics much further easter and south (Fig. 1A), how were the current sites chosen?

Thanks for seeking clarification here. The nature of the sampled easterly-directed plume is explained in the Introduction, but a link back was missing, c. line180, so I have added a reference to Fig. 3A at line 183. I have also added text at line 186 "The sites selected were known to have sufficient erratics within defined areas for sampling. However, to obtain similar sample sizes, the areas searched for the final two locations necessarily increased as the surface density of blocks decreased eastwards."

Line 181: "To obtain similar sample sizes…". Apologies if I have missed this somewhere in the manuscript or supplementary, but how many blocks were sampled at each location? Can the underlying data be made available in the supplementary files, or in an open-access location?

The sample sizes are noted in the next paragraph, beginning line 191. The data availability statement indicates that the data are available upon request. I am not a fan of putting small datasets in open-access repositories. A large data set which forms part of this current study is available on-line (Carling *et al*., 2013, PGA, 134).

Line 186-187: "The sample size was found to be sufficient for the aims of the project". How was this determined? Is this a qualitative appraisal, or have you assessed statistical power?

I do not go into detail within the manuscript as the explanation is rather long. I sampled a minimum of 30 erratics at each location, which meant >90 edge measurements at each location. Choosing a sample size in advance using statistical theory is not possible as the population of erratics is not known. I chose open heathland sites which are undisturbed, and generally searched 1 to 2km^2 in order to find 30 erratics. In that sense the sampling is 100%; i.e. I recorded all the erratics present.

If I consider the total erratic population (unknown) as a rule of thumb, a sample of 100 in 20,000 is often considered satisfactory (0.5%). I had previously mapped the location of 10,000 erratics, so if this is the target population then I measured around 0.45%, in accord with the rule of thumb, but the true population must be larger.

Subsequent upon sampling, I decided that the population sampled would only be acceptable if interpretable patterns arose, as reported in the manuscript. Otherwise, the sampled population would be increased. To enlarge the sample size, brings problems, as it would require sampling ever-increasing large geographical areas for each nominal point location (i.e. Wasdale Old Bridge, Haybanks, Blasterfield, sites near Barnard Castle in Teesdale and Levy Pool) which means that any function defined by distance from the granite outcrop would be obscured as the search area increased around each nominal point location. I considered plotting every point against distance but, in that manner, there are insufficient points at all the new point locations. Having a relatively small sample size within a small geographical area means that each sample population is comparable within the group and the sample sizes are statistically large enough to compare between groups. I have revised the sentence at line 193 and added two references which deal with estimating sample size from large 'unknown' populations. "The sample size was found to be sufficient (Daniel, 1999; Conroy, 2018) for the aims of the project and, moreover, interpretation of data trends became possible once the sample size n > 30 at each location."

I have added a new section to the Supplementary Information explaining the sampling strategy.

"Line 188: "Changes in block size with distance from the pluton are not considered herein using field data". Given this argument, there is a case for excluding Section 3.3.

Although the sampling strategy noted above precluded a sufficiently large enough sample to empirically explore size reduction with distance, the theoretical framework for fracture does allow an analysis and useful conclusions can be drawn. Section 3.3 is based on theory not on field data.

Line 189: Similar to above, what is meant by "statistically significant sample size"? How was this determined?

This statement no longer exists in the revised text and the issue of sampling has ben answered above.

Line 413: The description of parent to child evolution, and changing radii is convincing, but could have been augmented with a schematic timeline e.g., illustrating the process of resetting for a representative block. This would be similar to Figure 8, but for radii, rather than number of fractures.

I considered this option for radii when preparing the first submission, but it is difficult to realize given the rapid increase in new edges with variable degrees of edge-rounding with distance. I consulted a statistical modeller but, after some time, he gave up on the challenge! This aspect of the study will have to await possible inclusion in a future paper on coevolving shape and edge rounding.

Line 557: What is mean by “modest yet dynamic ice cover”. Does modest relate to ice thickness? By dynamic are you referring to ice velocity, thermal regime, a fluctuating ice margin?

At line 569, I have added a value of (*c*., 100m+ thick) ice, based on Hallet’s (1996) analysis. The thickness is returned to in the next paragraph of the manuscript.

Overall, this work is suitable for publication, but would be improved with refinements to presentation, and wider discussion of alternative drivers of edge rounding, which I think could play a role here.

Thank you finding the submission suitable for publication subject to modifications. I trust the changes I have made and my responses above are satisfactory.