Dear Editor, Stefano Tavani, and Reviewer, Marco Mercuri,

Thank you again for the constructive review. The criticism related to our lack of focus on censoring is valid. Previously the effect of censoring was mostly mentioned only on lines 494-499 in the discussion. Consequently, we have included some new minor analysis results related to the effect of censoring on our trace datasets. The new results provide alternative values for e.g. power-law exponents and allow critical assessment of the previous results. However, they do not override the previous results which are thus still kept intact.

See below for specific responses to your comments. Line number references are to the PDF document with the changes highlighted.

Other changes which might not be highlighted in the PDF:

- Reference to Skyttä et al. 2023 has been promoted from a preprint to the published article
- Appendix Table B1 has been updated with power-law exponent values.
- "Fig." is used in place of "Figure" and "Figs." in place of "Figures" as instructed in the author guidelines
- Addition of Fig. 6 with censoring investigation

On behalf of all the authors, Turku, Finland, May 2023, Nikolas Ovaskainen

```
Revision of "Detailed investigation of multi-scale fracture networks in
glacially abraded crystalline bedrock at Åland Islands, Finland" by N.
Ovaskainen and coauthors

General comments

Dear Editor and Authors,

The manuscript by Ovaskainen and cohautors has been improved from its initial
version. In particular, the introduction section is now more concise and
organised, consisting in two sub- sections, one of the which focuses on the
aims and scope of the manuscript. The cumulative distributions of all fracture
lengths for all datasets are now shown in a supplementary figure. I still
believe that this is an interesting study for the implications, impressive
dataset, and the use of a new software for fracture network analysis.
Therefore, in my opinion, the study deserves to be published on Solid Earth.
However, I think that the manuscript still requires some revisions in the
methodology.

I apologize, but I still have concerns regarding the methodology used for
fitting single-scale and multi-scale cumulative length distributions. In my
opinion, the reasoning for choosing the cut-off length and the assumption that
the cumulative length distributions (above the cut-off length) are best fitted
by a power-law is circular. The authors assume a priori that part of the
cumulative length distribution can be fitted with a power-law, and as a result,
```

the algorithm returns the power- law parameters and the cut-off length. The authors then only consider the lengths above the cut- off when comparing the power law fit with lognormal and exponential fits, concluding that the power-law fit is the best.

We now highlight the inability to fit a power-law to the full data in the appendix Fig. B1, reducing the "circular logic" mentioned. We would like to note that we do not conclude that the power-law fit is best for any of the scale length distributions. Rather, the lognormal fit is the best fit for all length data, regardless of scale or categorization into azimuth sets (see R-values in Table 5). However, the significance of the goodness-of-fit determination (p-values) vary, providing avenues of comparisons between the data. Furthermore, the multi-scale nature (i.e. power-law/fractal) of bedrock fractures is well documented and is based on physical rationale. We therefore believe providing the power-law characteristics, even if power-law is not the best statistical fit, has value.

Essentially, the cut-off length found in this way is taken as representative of the censoring bias. I agree that the fit with different equations should be tested on the same range of lengths and that lengths affected by the censoring bias should not be considered in the cumulative length distribution fit. However, I do not understand why the authors only consider that specific range of lengths. Additionally, the fit is performed on only around 2% of the dataset, which may not be statistically representative. Another concern I have is why the authors give more weight to the truncation bias and not the censoring bias.

I suggest two possible solutions to address this issue:

1) Specify and convince readers that the assumption that cumulative lengths are distributed following a power law and that only lengths above the cut-off are important while the remaining lengths are not affected by a censoring bias.

2) Revise the methodological part of the manuscript. Here, I see two possible solutions:

   A. Conduct a comparison between the cumulative length distribution on all data (substituting Fig. 5 with the new fig B1 plus the power law fit). The authors can discuss the fact that the power-law can only be applied to a small range of lengths, or alternatively, the cut-off can be set as equal to the minimum trace/branch length. Negative exponential and lognormal distributions fit the cumulative length distribution better (looking at Tab B1 and Fig. B1). Additionally, the fit for the multi-scale cumulative length distribution should be performed on all data, not just on the data above the cut-off length (i.e. the gray data points in Figure 6).

B. Apply the Maximum Likelihood Estimators - Kolmogorov-Smirnoff (MLE-KS) test considering various ranges of lengths, as described in some works (e.g., Dichiarante et al., 2020; Ceccato et al., 2022).

I acknowledge that option 2B, although very accurate, would be very time consuming and the topic is not the focus of the work. My suggestion to the authors is to evaluate option 2A, highlighting that the effect of censoring and truncation has been not considered in the fitting procedure.

We now provide analysis of the effect of a censoring cut-off on the determination of the power-law exponent, cut-off and subsequent cut-off proportion in Fig. 6. This allows critical assessment of the previously determined exponents and the truncation i.e. tail cut-offs. Along with the figure, text has been added to the methods (line 256 onwards), results (lines 351-362), discussion (line 399 onwards) and conclusion (lines 538-545) sections. This approach is a combination of the suggested options 1 and 2A as we both better clarify our reasoning on focusing on truncation cut-offs for the multi-scale length analysis and provide analysis of the effect of censoring in the single scale length distribution analysis. As previously mentioned, this analysis does not rule out the previous analysis but provides e.g. alternative exponent values and avenues of discussion. We have not replaced Fig. 5 with Fig. B1 from the appendix, as suggested in option 2A, as analysis of the full length data is not a focus and we hope the provided text, analysis and reasoning address the pointed out issues.

To summarise our reasoning on the focus on truncation cut-offs here: We firstly lack technical tools to reproducibly optimise for both truncation and censoring cut-offs simultaneously for multi-scale data (lines 277-278 and 502-505). Furthermore, the effect of the tail (low lengths) on the multi-scale fitting process is much more significant than the head due to our use of the cumulative number in the fitting for the multi-scale length data (lines 278-280). To remove the higher significance of the tail lengths in future studies we recommend the use of the probability density function in the discussion section (lines 507-509). Lastly, the defining of a censoring cut-off is less studied in literature as the effect of censoring is less obvious than the horizontal trend related to truncation, which is clearly observed in the tail lengths in log-log length distribution plots.

The number of lines refers to the clean version of the manuscript.

Marco Mercuri

Specific comments

1. Lines 89-92: Yes, but please anticipate here that the decrease in connectivity with increasing scale could be a a methodological issue.

The possible methodological cause is now mentioned on lines 91-92.

2. `Lines 363-366 These lines should be revised depending on whether and how the authors decide to revise the methodology.`

The lines 399-404 have been revised to include the discussion of censoring.

3. `Line 374. The authors might be interested in a recent paper which shows the use of Bing Maps for fracture network characterisation: Mercuri, M., Tavani, S., Aldega, L., Trippetta, F., Bigi, S., and Carminati, E. (2023). Are open-source aerial images useful for fracture network characterisation? Insights from a multi-scale approach in the Zagros Mts. Journal of Structural Geology, 104866`

Thank you! We have included it as a reference on line 418.

4. `Lines 489-493 These lines should be revised depending on whether and how the authors decide to revise the methodology.`

We have included the lacking censoring analysis for the given values on line 538-539 and added a new separate conclusion regarding the censoring investigation on lines 542-545.

`Technical corrections`

- `Lines 9-10: "The best fit to model the lineaments and fracture lengths with a common power- law resulted in an exponent of -1.13". The sentence is not easily understandable during a first read. Please rephrase it.`

The sentence has been rephrased to be clearer.

- `Lines 11-12: I suggest removing "could"`

Replaced with “can”.

- `L 24-25 There is a repetition of lines 19-20. I suggest removing these lines.`

The repeating part “, but it is crucial … characterisation.” was removed.

- `I suggest rephrase this sentence into something like "requires collecting fracture and lineament data using a combination of methods and preferably from multiple scales of observation" for improving conciseness.`

We rephrased the sentence on lines 33-34.

- `L76 Please remove "e.g."`

It has been removed.

- `L126. I suggest changing "e.g." with "due, for example, to" or something similar.`

The sentence has been rephrased to overall be more clear.

- `L144. Maybe the reference in the text should be written like "published by Ovaskainen et al. (2022)". Please check the guidelines for authors.`

Thanks, the reference was written wrong. It has been fixed.

- `L217. Please rephrase in "represent non connected nodes" for better clarity`

We rephrased as suggested.