

81 I suggest making this comment more nuanced and adding a reference [...]

Thank you, we changed the line as suggested.

106 Here the flow in the fracture network is assumed to be a function of connectivity, but in the preceding list of fracture types many of the elements may not be conducive of fluid flow, for example some faults and opening-mode fractures that are sealed. Likewise, if you have a situation where sets are of different ages, early sets may be sealed (or partly sealed) and later ones more open. An example is an outcrop of veins abutted or crossed by later joints. These abutting and crossing relations will have a different impact on flow than a bunch of intersecting joints. Maybe in say: “If all the fractures are open, a network with prevalence of I nodes...” This may not be central to the point that you are making in this paper, but it’s such a common and misleading logical jump in fracture network studies that the clarification is useful. See the discussion in Forstner and Laubach, 2022, J. Struct. Geol. Also, if the rock itself is porous, even a network that has only I nodes can markedly augment fluid flow (Philip et al., 2005, SPE Res. Eval. Eng.), here length distribution is the key parameter (not connectivity) which just makes your focus on length even more important.

We agree on the importance of clarifying, we expand the line as following:

106 In a non-porous rock with all open fractures, a network with a prevalence of I nodes will be less connected, and fluid flow will be more restrained. However, for many of the fracture types that were previously discussed, this is indeed not the case (e.g. sealed faults and opening-mode fractures) (Forstner and Laubach, 2022). Furthermore, if the rock is porous then length distribution becomes the key parameter for controlling fluid flow (Philip et al., 2005).

122 It seems like these values might also be meaningless for ‘stochastic modeling’? Do you clarify this in the Discussion?

The discussed values indeed are useful for stochastic modeling both from a statistical and numerical point of view (i.e. DFNs). In the text this was not explained clearly. We provide the following edit to clarify.

119 Circular scan lines methods on the other hand do offer an unbiased estimate of the mean length, however, being non-parametric, they do not yield neither the distribution type (e.g. normal, exponential, etc.) nor distribution shape parameters (e.g. standard deviation, etc.). This in turn, makes the estimate completely useless to quantitatively compare different results, and carry out any downstream statistical and/or numerical modelling, such as DFN stochastic fracture modelling.

129-132 On first read, I found the transitions here confusing. For clarity I think you ought to warn the reader here that you are going to demonstrate the time-length dimensional shift in 3.3. Something like ‘Survival analysis is usually used in the time domain. In section 3.3. we show how a time-length dimensional shift is valid. Here we briefly introduce the terms as they are used in the time domain.’ These are key lines defining terms. I think they could use some clarification. What do you mean by ‘the event of interest is commonly defined as *death*’? Is a clarifying word missing? The ‘event of “x” is’? Or do you need a some more information at the start of the paragraph: “Survival analysis is used to analyze data in which the time until the event is of interest (for example, the time until death in some medical or biological contexts).” This would perhaps be a good point to introduce the idea that you are substituting distance for time?

Thank you for pointing out that you found the transition confusing. We changed the text as follows hoping to make it clearer.

122 To solve these problems, we propose to use survival analysis, a specialized field of statistics, specifically developed to deal with censored data. Survival analysis focuses on the analysis of time of occurrence until an event of interest (Kalbfleisch and Prentice, 2002). The advantage of survival analysis over the methods discussed above is that it considers censored data as the carrier of the crucial information that the event did not occur up to the censoring time, thus allowing for an unbiased estimation of all statistical parameters and models. However, although in literature the terms survival times, time-to-event, or more generally lifetimes (Lawless, 2003) seem to imply that time is the only valid variable, any non-negative continuous variable, such as length, is valid (Kalbfleisch and Prentice, 2002;

Lawless, 2003). In the following sections of this chapter, we will start describing the canonical theory behind survival analysis in function of time, and then we will show how the same theory can be applied in space, to sets of length or distance measurements.

129 Since this technique is rooted in medical and biological applications, the nomenclature from this type of literature is carried along. The event of interest (for which we measure the time-to-event) is often defined as death, while a loss indicates that the observation has been lost because it was hindered by a secondary event, called a censoring event. Censoring can be ...

133 Which 'length' do you mean here?

Changed with:

133 the event happens after the end of the study period and thus we observe the partial lifetime of the event.

173 1D or 2D? How does this conversion work?

We did not understand if the comment is referred to the type of intersection between the fractures 173 4. the censored event as the intersection between the fracture trace and the boundary (marked by a B node), or if it is referring to the figure below indicating that it is not clear what the figure entails. For the former it is a 2D intersection. For the latter, then we can expand the figure caption and text as follows:

Figure 6. Censoring effect on an example of a simple fracture network and corresponding survival diagram. The survival diagram is a 1D representation of the fracture length. On the Y axis the fracture number is indicated and on the X axis the length is measured. Solid lines indicate the actual measured length while dashed lines indicate the possible continuation of the fracture. Yellow pentagons represent the censoring of the boundary.

174 Figure 6 represents an abstraction of the fracture network by just representing fractures by their length. Each fracture in the network is numbered (Y axis) and the

corresponding fracture length is represented by a bar. Bars with a yellow pentagon indicate that the fracture n is censored and thus the measured length is shorter than the true length. By applying ...

190 'simplest'

204 'it has its limitations'

223 '...that can enable the researcher to obtain an informed...'

235 'both figures'

395 'it'? Maybe 'they are'?

409 'in key of time' is an odd phrase. Check.

411 'useless' seems harsh. I'm not convinced this extra remark is needed. Anyway, there may be other parameters (like segmentation) that have similar effects to outcrop size that would benefit from the approach you propose, even if outcrops were arbitrarily large.

Changed in the main text. Thank you for the corrections.

350 Given the limitations of any spacing statistic, I think it would be worthwhile mentioning here that good field practice with scanlines should be to keep track of the sequence of fracture occurrences, in other words, the spatial arrangement, as you've pointed out in other work (and also Marrett et al. 2018, J Struct Geol). Your analysis here seems like it would be equally apt for spatial arrangement data collection and analysis.

We added a brief mention of this in the text as such

371 Finally we would like to point out that the censoring analysis is a secondary part in the analysis for spacing. It is worth noting that analysing the spatial arrangement of the fractures in the network (such as Marrett et al. 2018 and Bistacchi et.al 2020) is of fundamental importance. The presented datasets are equally apt to this type of analysis; however, we decided not to include this analysis and focus mainly on censoring to avoid increasing the length of an already dense text.

376 I agree with this way of proceeding re: defining length. Does your method work as well with lengths defined via branches; is there a reason to choose one or the other? Maybe this gets out of scope, but the way you mention it here might make a reader wonder.

Yes, we chose to measure the lengths of the entire segments instead of branches because they entail two different things. Branches offer a useful topological abstraction of the network (making it possible to classify node intersections), but they do not have a real geological or physical meaning. As we defined in section 2, 2D fractures traces are the intersection of discontinuity surfaces with a secondary surface. Branches on the other hand are defined as a segment of a fracture trace between any two nodes (either I-I, I-Y etc..). Considering the geological origin of a trace, by using branches we would be essentially segmenting fracture planes in smaller sub-planes. This, however, is only an artifact given by the topological definition of a branch and thus the obtained branch length distribution does not carry any real physical meaning.

This discussion, as interesting as it is, may be a bit out of scope and we tried our best to summarize it in the discussion as follows:

376 **Branches offer a useful topological abstraction of the network (making it possible to classify node intersections), but they do not carry a real geological or physical meaning and as such a distribution obtained by fitting branch-length will have a different meaning compared to a length distribution.**

388 This is a big claim that length is always underestimated. What if you have a process that produces only short fractures (or even fractures that are shorter than your outcrop size). Hooker et al. 2013, J. Struct. Geol. describes one set (of several) that only contains very short lengths. Maybe some caveats are in order here.

Yes, however we firmly support it and we expanded the discussion to motivate it further:

423 Measured lengths of censored fractures will always be shorter than their true lengths and, by using the first simple approach, the dataset is essentially “polluted” by shorter fractures thus always decreasing the measured mean. The second simple method will also lead to an underestimation of the mean because of the size bias. However, this second method can be less impacted by censoring. For example, if a fracture population has a very small standard deviation (i.e. almost all fractures have the same length) and/or fractures are occurring in an outcrop that is much bigger than the characteristic fracture length, then removing censored values would not have a great impact on the estimation. But, even if small, the underestimation will always be present. Overestimation of the mean length would be possible in these scenarios when we do not consider censoring as independent from the length distribution (for example if only fractures shorter than a certain value are censored). However, this would violate both the core underlying hypothesis of random censoring, and standard geological experience, and thus we do not deem it possible under these imposed limits.

405 Although testing this hypothesis is something that people studying fracture lengths in the context of geomorphology ought to consider. Particularly large or open fractures can affect the size, shape, and occurrence of outcrop. See Eppes et al. 2024, Earth Surface Dynamics, doi.org/10.5194/esurf-12-35-2024.

We added this remark in

406 ... independent processes. Nonetheless in some applications (Eppes et al. 2024) this assumption may not hold, and a more in-depth study may be required to prove the independence hypothesis before proceeding.

445 This assumes that measurements are only carried out at one scale of resolution. But this need not be the case. See Ortega et al. 2006, AAPG Bulletin (for aperture sizes) and Forstner et al. for lengths.

Changed to

445 **Because of this limitation, for a constant resolution scale, the modelled length distribution**

451 And for some fracture systems, the smaller fractures are more prone to be mineral filled and potentially less obvious features on images. This size/visibility effect can also manifest in the picking of long fractures if the long traces are segmented.

Added as another factor contribution to censoring, thank you for the suggestion

In the example case studies, with such big clear outcrops, can you analyze a small area within the larger area and verify that you are accounting for the censoring correctly?

This is a tricky question that we thought about while writing the paper. It is not easy to see if censoring is correctly accounted for by just subsampling the outcrop (as large as it is). The problem is that essentially, we do not have a controlled environment. First, we are estimating only a limited suite of statistical models and thus we cannot say if the best estimated model is the true underlying model (and in fact we will never be able to tell). So even if for the first case study the lognormal may seem perfectly fitting, we cannot be certain that it is the true underlying statistical model. Moreover, the spatial distribution of the fractures in the outcrop space is also not uniform thus with the same sub area dimension you will have different model estimations depending on the position of the sub area. Thus, we found it difficult to obtain a satisfactory estimate of how well censoring is accounted and how well survival analysis works depending on the censoring percentage. We are relying on the fact that survival analysis has been used and is still being used in countless applications and show that it is working also for lengths. However, we believe that synthetic experiments can and should be carried out to explore further the effects of censoring, violations of the underlying hypothesis on the final estimation and the overall precision and reliability of survival analysis (we talk about this in 407-435). We decided to not include or explore synthetic results because it would have drastically increased the length of the MS and blurred its focus.