Solid Earth
Discussions

Open Access

EGU

Interactive comment on "How do we see fractures? Quantifying subjective bias in fracture data collection" *by* Billy J. Andrews et al.

William Dunne (Referee)

wdunne@utk.edu

Received and published: 10 February 2019

"How do we see fractures? Quantifying subjective bias in fracture data Collection" by Billy J. Andrews, Jennifer J. Roberts, Zoe K. Shipton, Sabina Bigi, Maria C. Tartarello, and Gareth O. Johnson. https://doi.org/10.5194/se-2018-135

The paper tackles an important general topic in scientific research and choses to use the characterization of natural fracture networks on two-dimensional surfaces as the framework for analysis and discussion. The topic is "the effect of the biases of the observers on the observations that will form the data population for a scientific analysis". The approach is to have a range of participants apply four different established methods for characterizing the networks. The participants and their differences in data

populations that they gathered, are considered in terms of factors such as amount of professional experience, individual or group work, or a participant's preference to gather detailed data carefully or larger data sets quickly. The paper provides extensive and appropriate documentation, and focuses its analysis particularly on the impact to the state of the gathered data as a function of whether observers tend to be "more detailed" or "less detailed" with respect to their data gathering. For the particular case, the authors provide recommendations for which sampling methods best overcome observer bias, sizing the sampling technique, best approaches to data gathering by a group, and integrating project goals into the data-gathering plan to minimize observer bias. **

Major Comments: Section 3 - A careful set of data are collected about participant performance for factors such as amount of data collection, type of data collection, patterns of data collection, time taken to collect data, and data collection performance as a function of individual or group data collection. These data are well documented. However, the analysis of this data in Section 3 is somewhat vague with statements such as "reasonably consistent", "a suggestion in the data", "differences are not enough to be confident", or "the trend is very weak". No framework for a quantitative and/or qualitative approach is established at the outset of the data presentation and analysis in this section. Presenting this framework and then utilizing it would be a critical for improving the rigor of the present paper. Presenting the framework will likely lead to similar results and will do so in a manner that creates greater confidence in the results presented in this Section.

Page 17, Line 27 to Page 18, Line 5 (End of Discussion) - This text should be replaced by more ambitious text that speaks both more generally than just the mechanics of resolving data gathering differences between observers in the context of "detail" and also connects to real-world situations that apply to the readers beyond just those for the particulars of gathering fracture-related data. So, it is certainly worthwhile constructing experiments that directly test for effects related to subjective bias or operator bias con-

cerning the collection of geological data. Yet, how do experimental results apply to real situations involving data collection? For example, how do the results provide value to an instructor working with a group of students who are performing field data collection for the first time vs. to an individual or team that are applying a rules-based data collection process with specific training prior to the first field deployment to ensure familiarity with the rules-based approach vs. to a computer-based observer utilizing virtual 3-D outcrops from photogrammetric data who has no prior field experience with the data set vs. to building a data set by crowd-sourcing. In this context, the present paper would be a stronger contribution if it explicitly considered the application of its outcomes to real-world circumstances of value and interest to readers. Replacing the existing text at the end of Discussion and embracing this opportunity for expanding the import of the narrative should bring greater recognition to the contribution of the authors and greater interest from the readership. Also, this revised text would address comments made on Page 14 – Line 29, Page 15 – Line 9, and Page 15 – Line 27, where the authors need to extend their work and provide more guidance about the meaning and application of their results.

The Discussion also has a few key locations where the work of others should be included and considered. Please see "Other Comments" for details.
****

Other Comments: Page 2, Lines 8-11 – It seems odd to list four methods and only provide citations for one of the four methods. Quality citations exist for all of the methods and the manuscript would be more useful for readers if each method was paired in the text here with at least two appropriate citations.

Page 2, Lines 11-14 – The annotated PDF for this review of the paper provides suggestions for strengthening the statement of the purpose of this contribution.

Page 2, Lines 25-29 - These two sentences consider observational resolution and limitations to the quality of observations that can be made as a function of the exposed

rock. These two points would relate to both objective and subjective uncertainty, and as such seem out of place in the narrative flow. Given the text in Lines 22 to 25 that is focusing on subjective uncertainty, any text, if any is needed, after Line 25 in this paragraph should only consider factors the relate to subjective uncertainty. It might be best to eliminate this text and just continue with the text in the new paragraph starting on Line 30 that focuses on the subjective uncertainty and further introduces the paper.

Page 3, Lines 15-18 – Suggested text revisions are offered to completely and correctly state the contribution of Zeeb et al., (2013) to defining the number of measurements needed to provide an estimated value for a characteristic that is statistically significant.

Page 4, Line 3 – It would be useful to explicitly state for the reader why plotting topology data in a triangular diagram is useful.

Page 6 Line 15 - Specify the dimensions of A3 paper as it is not a universally used paper size.

Page 7 Lines 4 to 6 – Suggestions offered in the annotated PDF for this review to improve the clarity and purpose of this text related to methodology and then the approach to statistical characterization.

Page 7, Line 19 - How is "reasonable amount of consistency" quantitatively defined or qualitatively recognized?

Page 7 Line 30 to Page 8 Line 4 - Suggestions are offered to improve the precision and the clarity of the text describing locations and causes of increased observational uncertainty as a function of the participants for Scanline 6.

Page 8, Lines 23-24 - How often were participants "internally consistent"? What is the measure/criterion for "internal consistency"? What is the measure/criterion for defining the occurrence of "often"?

Page 9, Line 6 - What defines or qualifies "varied considerably"? What is the standard or basis for comparison?

Page 9 Line 14 - Words added because an operational preference for a participant to report more smaller fracture traces does not necessitate that the reported small fracture traces are the actual small fractures.

Page 10, Line 30 - The meaning of "the joint highest" is not clear.

Page 10, Line 30 - What defines "a suggestion in the data"? This characteristic or attribute is not defined or explained. More explanation is needed here.

Page 11, Lines 3-4 - While I agree, the paper would be better if the authors explained why they believe "differences are not enough to be confident that this is due to working in groups rather than differences in the fracture network".

Page 11, Line 15 - "correlation is weak" - Is this statement backed by statistical analysis or is that a judgment call by the authors. Additional explanation is needed here.

Page 11, Line 29 - What is the statistical or qualitative basis for stating that "however the trend is very weak". Further explanation is needed.

Page 12, Lines 19-25 - Suggested revisions offered to focus the narratives on the reporting by participants. It is the values as reported by the participants rather than the values themselves that is the focus of this work and the connection of the participants to the values should be explicitly maintained throughout the narrative.

Page 12, Line 25 - The narrative should be more direct and avoid the use of the word "suggest" that is vague and lacking in framework.

Page 12, Line 26 - Revision offered to the latter part of this sentence to clearly and explicitly relate "spreads" in Table 7 to main text, and then clearly state the interpretation that the authors have derived from considering this population of spreads as a function of sampling method and subjective bias.

Page 12, Line 27 - Apologies for my confusion but how does "most robust" relate to "displaying considerable variability". Previously, "robustness" related to similar reported

values by participants with limited variability. What am I missing in the text here?

Page 13, Lines 11-12 - Why does the variation in reported values by different participants for the same sampling method directly correlate to the size of the tool for the method, and hence the sample size. Is it not the case that if a statistically valid sample size is going to be collected by each participant that both the sample size and the tool size need to be specified prior to measurement by all participants? Or is there an implication that the dimension of the scanline or window needs to be set on the basis of minimum values to be expected from the range of values due to subjective bias? I suspect that the core problem with this clause of this sentence is that it should come at the end of the paragraph after the key observations are offered, so that a summative comment can then be made and justified. So, this text should be relocated. The comment does also need some improvement in text to provide greater clarity.

Page 14, Line 2 - the participants are not less or more detailed. Their observations are. Suggested text revisions are offered to clarify this point. This approach should be adopted at other appropriate locations in the narrative.

Page 14, Line 5 - A clear conclusion and useful point is reached at the end of this paragraph. What are recommendations for operationalizing the observation with respect to a future data-sampling campaign? How does the needed level of detail for a campaign fit into this operationalization? After all, not every sampling campaign necessarily needs the same level of detail as a function of campaign goals. Or put differently, more or less detail is not always best!

Section 4 - Subsections misnumbered because two Section 4.2's are identified.

Page 14, "First" Section 4.2 - This short section is out of place and effectively encompassed in the later sections including particularly Section 4.4. It should be deleted.

Page 15 Lines 22-23 - Is this suggestion about a preference for using field-based data rather than photo-based data a first occurrence in the literature. If not, prior work should

be cited and most likely, briefly discussed.

Page 16, Lines 9-10 - The conclusion about consistency of results for single observers is dependent on the observer not changing their approach to data gathering as a function of experience gained by data gathering, subsequent training, and/or subsequent interaction with other data gatherers. The conclusion seems too simplistic vs. reality. It might be applicable to "single events" such as one day of fieldwork or a single workshop, but is likely to be less applicable with the passage of time and the occurrence of multiple events, particularly if they have differing goals.

Page 16, Lines 11-13 - Text revisions offered to be less judgmental and to more clearly state "driving philosophies" for "less detailed" vs. "more detailed" participants.

Page 16, Lines 14-15 - The work by others around this point should be cited here and included. Much of it may be in the literature for Psychology, but a useful entry point may be contributions involving Shipley & Tikoff.

Page 16, Lines 29-30 - This "part" sentence is a little odd. It is probably not needed (could place the colon after "collect"). Yet, if it is going to be retained, it should "go large" and not "small" (why focus on folks who do paleostress analysis?). The recommendations are relevant to all persons collecting structural data or utilizing the data products/analyses of others (go large!).

Abstract and Conclusions - Likely will need minor revisions if the main text is revised along these recommendations. ***

Additional Comment: Please see separately submitted PDF with annotations showing comments suggesting detailed improvements to the text of the main document.

Please also note the supplement to this comment:
https://www.solid-earth-discuss.net/se-2018-135/se-2018-135-RC2-supplement.pdf

Interactive comment on Solid Earth Discuss., https://doi.org/10.5194/se-2018-135, 2019.

C7