We thank Roberto Emanuele Rizzo and William Dunne for their detailed, very constructive, and helpful comments on this manuscript. The following section lists our responses to their comments and the corresponding changes made to the manuscript.

5

## Reply to 'Reviews of Mapping the fracture network in the Lilstock pavement, Bristol Channel, UK: manual versus automatic by Weismüller et al.' by Roberto Emanuele Rizzo.

10 We thank Roberto Emanuele Rizzo for the detailed and helpful comments on the manuscript.

A main criticism was that the manuscript was quite hard to read, particularly in the results section where it goes back and forth discussing the data obtained from manual and automatic methods. In the new version of the manuscript we have taken special care to clarify which
15 methods we are referring to in the narrative. We revised the results section accordingly and further divided the section into sub-sections for a clearer structuring therein. Throughout the manuscript, we further improved the narrative for a better readability.
Another major point of criticism was that some of the claims presented in the manuscript were not always well justified. We agree that some of our claims required further explanations and
20 citations and provided those for the addressed cases whenever possible. For the cases where we were not able to provide a sufficient amount of evidence or citations, or when the latter would raise questions out of the scope of this study, we removed the claims from the manuscript accordingly.

25 Detailed comments on the individual points raised by the reviewer are provided below.

Lines 147-148: Can you please provide an example for this, e.g. image/figure

Reply: The section was moved a few lines up according to the suggestion of the other reviewer.
30 We added a figure showing examples of the ground truthing to the supplement (new S1).

Lines 150 – 151: I don't disagree, but I would not say that is "obvious"; for example someone else can claim that they choose to manually pick all fractures because this yields the best results, due to the topography of an outcrop or light exposure during
35 image acquisition.

Reply: We agree that "obvious" was not the best choice of words. The "obvious" was replaced according to the suggestion of the other reviewer to clarify, that we refer to the need of a more rapid technique.
40
 How long is a "reasonable time"? How do you account for it?

Reply: The initial choosing of words did not provide clear information, we have revised the sentence for clarity and provide quantitative estimations. More information about the time
45 required and extrapolated to the complete outcrop was added to the section below.

In this regards, can you please give an overall estimate on the time needed by the automatic process to extract all the fractures from one of the tiles (including all the steps for preprocessing an image and the number of trials needed before finding the right set of
50 parameters to extract the fractures) and compare this on how long it takes manually?
In my opinion this would be a very interesting information. In addition, for someone that has never done fracture tracing would learning and using the software make the job faster/easier or it would take longer than do it directly manually? Can you please comment?

55
Reply: We agree that these points are very interesting for the reader. This was also one of the main comments of the other reviewer, therefore, we added a narrative in the methods parts explaining the required time under different aspects for each of the two methods in detail, incorporating the tracing itself and the time required for the removal of artifacts. Further points

60 were added in the discussion giving an estimate what time would be required to trace the whole dataset using either technique. Also, we have added a short explanation to highlight the pros and cons for someone who has never done fracture tracing for both techniques. As explained in the manuscript, the dataset requires the interpreter to make judgment calls on several occasions, thus this task might be challenging for someone who has never done fracture

65 tracing before. With the help of the automatic trace extraction, the results would be unbiased and easier to reproduce by another interpreter. Timewise, an unexperienced interpreter must either face the learning curve of the GIS software (or any other software used to manually trace the fractures) or MATLAB for which the automatic trace extraction code is written. Depending on personal preferences, either learning curve might be steeper. For small datasets and an

70 unexperienced interpreter, it might be faster to do everything manually, but if larger datasets are to be processed, the time invested for learning the automatic trace extraction is compensated by the time saved during the application of the method.

75 Line 156: Can you please report which software have you used during the manual digitalization of fractures? Was it one of the vector graphics editors (e.g., Adobe Illustrator, Corel Draw, Inkscape), or a geographic information system-based software?

Reply: This information was accidentally removed during an earlier revision of the manuscript.
80 We now clearly mention the use of ArcGIS at this point.

Lines 170 – 173: You never mentioned before any "intensity" parameter; can you please clarify? In addition, when discussing the needs for applying an 'intensity threshold' do you mean that this process is needed in order to make the automatically detected maps

85 look similar to the original photograph from which you extracted the fractures? If this is the case, this sentence needs to be clarified.

Reply: The "intensity" in this case refers to the chosen parameter combination during the automatic extraction. We have slightly revised the sentences to clarify and cited Prabhakaran

90 et al. (2019) for further information on the topic for the interested reader.

Line 175: For completeness, can you please briefly explain the 'polygon sampling' strategy?
Reply: We added a narrative to briefly explain the polygon sampling strategy as discussed in
95 Nyberg et al. 2018.

Lines 179 – 180: The sentence "2D fracture networks ..." needs a citation. Can you please clarify the meaning of "spatial graph"? In addition, can you please clarify the
100 terms 'node' and 'edge'? In the context of fracture networks, these can mean different things, depending on whether the fractures themselves are viewed as graph vertexes (and therefore they are the 'nodes'), or if they are considered as links (the 'edges') between fracture intersections and terminations.

105 Reply: We added a citation as suggested (Sanderson et al. 2019) and also added some sentences to better explain the meaning of "spatial graph" and our use of the terms 'node' and 'edge' with respect to their use in literature.

110 Lines 187 – 188: On lines 156 – 157 you mentioned that when manually tracing the

fracture these are "traced as polylines". From the use of the prefix 'poly-' I have understood that one fracture trace is already made of many segments. So, can you please explain the need to further subdivide the fracture traces?

115   Reply: The term "polyline" is used according to its definition within the ArcGIS environment: "A Polyline object is a shape defined by one or more paths, in which a path is a series of connected segments." Therefore, it is correct that one manually mapped fracture trace consists of a polyline made of many segments that share the ID of the polyline. However, NetworkGT only allows single segments (stored with an individual ID) as input, therefore we had to dissolve

120   the polylines first. We have added a narrative to point this out for the reader along with a more detailed explanation.

Lines 191 – 192: Can you please clarify if the correction for node degree > 3 was your improvement on the software, or it is already and option within of the original

125   NetworkGT?

Reply: In the manuscript we state that nodes with a degree > 4 (X-node) are not supported/implemented in the code of NetworkGT and, thus, returned as error. To fix these resulting errors, we developed our own method to correct these errors resulting from the

130   application of NetworkGT. We revised the sentence to make this point clear.

Lines 192 – 194: Can you please clarify and give more details on the use of the "spatial join function"? How does it work?

135   Reply: The spatial join function is provided in the ArcGIS Analysis Toolbox and can be used to join features based on different parameters of their spatial location. We added a few sentences to give more details on the topic and provide a better explanation on how we applied this function to our data.

140   Lines 214 – 215: What is Passchier et al., interpretation? The paper is not published yet, so unless you clarify it is not possible to know how these features have been interpreted. Even if the paper was published, it would be more helpful if you could briefly explain this interpretation here, otherwise the reader would be forced to read another paper to understand what you mean.

145
Reply: In Passchier et al. (in prep.) we introduce the initial interpretations of the generations on a larger scale of the outcrop, which helped to identify the criteria on which we base our interpretation of the generations in this manuscript. However, the criteria we used and present in this work are self-sustained, explained in detail in section 3.4 and do not require the citation

150   of Passchier et al. (in prep). We now refer to this citation as a companion paper only once and earlier in the manuscript and rewrote the narrative to clarify, that the interpretation of the generations is based on the criteria explained in this manuscript.


155   Line 218: You can avoid the brackets here because you are directly mentioning the authors.

Reply: We revised the sentence according to the suggestions of the other reviewer. In the revised sentence, the brackets are now necessary.

160   Lines 223 – 224: Can you please explain why have you chosen to show exclusively fracture intensity maps? To my knowledge NetworkGT allows also for fracture density maps, why have you opted for not showing these?

Reply: In the initial version of the manuscript we only presented fracture intensity because
165   fracture intensity shows the total persistence of fracture segments within the domains, while fracture density only gives the amount of fractures segments per unit area. Fracture segment

density was assumed to be of lesser interest because we already stated that automatic trace extraction produces more trace segments than manual tracing (of whole fractures and splitting them at the nodes), therefore, higher densities for the automatic extraction are just the logical consequence.

To provide a better comparison of the methods presented in this manuscript we revised the section following the suggestion of the other reviewer and now also incorporate fracture segment densities throughout the domains. We have added two figures comparing the results for both methods, further material to the supplement, a new subsection introducing the results of $P_{20}$ for both methods and a narrative to explain why we now investigate fracture segment density and intensity with implications for the two methods.

Lines 228 – 229: This sentence is not clear. Do you mean that number of trace segments is higher than the actual count of fractures? Or do you mean that the automatic extraction of fractures overcount the total number of fractures compared to the manual interpretation?

Reply: The automatic code generates segments, while a manual interpreter maps the complete fracture from tip to tip which represents the path along several connected segments. We revised the sentence for clarity and now better explain the difference of the results and how to process the data to make a comparison possible by splitting the manually traced fracture traces into segments analogue to the ones generated automatically.

Line 239: If not resolvable from the drone images, was the 'minimum length' measured directly on outcrop?

Reply: The section was rewritten according to the suggestion of the other reviewer. Now it is clarified that we are referring to a cutoff length within the sampling windows. These short segments are the result of the tracing and cannot be verified because they are below the resolution of the ortho-mosaics.

Lines 239 – 247: Are all these results relative to the manually traces fractures, the automatic, or both. Not clear. Moreover, can you please add few lines to describe more in details how all the statistical parameter that you show are useful to describe the fracture network? What having a positive kurtosis means? Similarly, can you please better explain what do you mean by symmetric and asymmetric branch distribution?

Reply: We revised the whole section and divided it into sub-sections to make it more clear to which results (manual or automatic) we are referring to. The section describing the statistical parameters has been extended to provide more detailed explanations of the parameters and their meaning or implication for the network.

Line 246: Data of fracture trace length distribution are only shown for the automatic trace detection method, but not for the manually derived network. Is there a reason for this choice? Since you are discussing both methods, and relative results, I would advise to add a similar figure to Fig.7 where showing data from the manually traced network.
Reply: To follow the initial narrative of the manuscript, we stated that the results of both methods are comparable anyway and, thus, we only used the results of one method to describe the network. Now, that we have revised the narrative to an even larger focus on the comparison of both methods, we have added a similar figure presenting the manually traced segments as advised and also added a section that compares both methods using the results presented in the figures.

Lines 248 – 250: This sentence is not clear. Can you please review it? Are you referring to a limitation of the automatic extraction method?

Reply: Yes, this is what we are referring to. The sentence was revised according to the suggestion of the other reviewer for clarification.

Lines 250 – 253: For helping the reader to compare/contrast the results of automatic vs. manual method, I would suggest adding to Fig.5 and Fig.6 (which currently only show results for the automatic extraction) the fracture maps obtained by manually picking fractures. I acknowledge that these figures are shown later in the manuscript (figures13, 14 and 15), but having the fracture maps produced with the two methods one next to each other would be ideal.

Reply: We agree that a direct comparison of the networks resulting from the techniques is interesting for the reader. However, this would require another very large figure for the networks and the branches and nodes to be visible, and break the narrative at this point of the manuscript. We present a figure that directly compares all networks using $P_{21}$ in which the networks are shown in the background of the plots and can be compared there with the advantage of different $P_{21}$ values pointing the reader directly to locations where the methods have different results. Furthermore, we have added examples of two areas which are compared in detail using $P_{20}$ along with supplementary material that shows $P_{20}$ with the networks plotted in the background in the supplement. For the interested reader, the shapefiles are also provided as supplementary material and can compared using a GIS.

Lines 223 – 261 (all section 4.1): The whole section is a bit confusing and needs to be reviewed. It is never clear if you are referring to the results obtained by the manual of the automatic method. For clearness, I would suggest to first describe the results obtained by one method, followed by those obtained by the other, but I do not want to impose any style to your paper. As long as you make it clear which method you are referring to, find a way that you like better.

Reply: We agree that the initial version of the section was a bit confusing. We revised the whole section for clarity in our narrative and divided it into several subsections to better compare both methods, following the suggestion of the other reviewer. Within the subsections we also revised the narrative to avoid confusion between the two methods and now state clearly which method we are referring to.

Lines 275 – 277: I understand that you are referring to the plot of fracture length vs. strike, however as written here it is not very clear. Please review.

Reply: We revised the sentence to clarify that we refer to the relationship of strike and length of single fractures. Also, according to the comment of reviewer 2, we rewrote the sentence to avoid confusion with statistically defined clusters in the plots.

Line 287: Is the fault to the Southeast corner of the tile NE2, or to the southeast to the whole outcrop? Please clarify.

Reply: The sentence was revised to clarify that we are referring to the relative position of the fault to the sample window.

Lines 288 – 289: This sentence is rather confusing. Can you please review it? Particularly, can you please clarify the meaning of 'perceived appearance'?

Reply: We revised the sentence and deleted the term "perceived appearance" to avoid confusion. Now we state clearly that we refer to the generation 3 which is present in the network of the domains in the SW but not in the NE.

Lines 290 – 291: There is a 'IN' missing between 'domains' and 'the SW'.

280 Reply: We added the 'in'.

Line 314: Should be 'strike at angle'.

Reply: We replaced "in" with "at".

285

Line 390: I understand that you might don't want to engulf the paper with too many figures, however, because you discuss largely around Figure S1, in my opinion this should be incorporated in the main paper, rather than be relegated to supplementary.

290 Reply: We agree with this point and added the supplement S1 as figure to the manuscript within the revised section 4.1.

Line 396: Would it be more appropriate saying 'automatically extracted fracture network' rather than 'generated'? Usually the term generated is associated to Discrete
295 Fracture Network (DFN) models.

Reply: We agree and changed the wording accordingly to avoid confusion

Lines 401 – 402: Point (i), as written, it can be interpreted that the fractures are in the
300 'code', while the algorithm extracts these features from an image. Please amend. Point (ii), can you please expand on this point?

Reply: We revised (i) to avoid confusion. Now it is clear that the code extracts the fractures we refer to. Furthermore, we expanded (ii) by providing more explanations and details in the following sentences
305 Lines 402 – 404: As written these sentences are very difficult to understand. Please review them.

Reply: We revised the sentences for clarity. As suggested by the other reviewer, we also added a new figure with more detailed explanations to the manuscript to provide examples of the cases discussed in these sentences.
310

Line 404 – 405: Can you please give more details about this procedure? It is not clear to me if you have completely discarded erosional features from the network created when manually tracing the fractures. In addition, the sentence needs to be reviewed: missing 'AS' between 'not' and 'fractures'.
315

Reply: In this sentence we refer to wrong positives in the automatic extraction caused by a rough surface and a too high sensitivity (in terms of the chosen parameter combination) of the code. We added a more detailed explanation of the issue and rewrote the sentence to clarify.

320 Line 405 – 407: Can you please provide a full description of what do you mean by 'sensitivity'? This term has not been used before in the manuscript, therefore needs to be fully explained. In addition, I would avoid the use of vague adjectives like 'too high' or 'too low'. How much? Can you please quantify?

325 Reply: We revised the sentence and replaced "sensitivity" to clarify that we refer to the parameter combination chosen during the automatic tracing, what is explained with more details earlier in the manuscript. The sentence and complete manuscript were revised to avoid vague adjectives as suggested, using qualifications instead.

330 Line 409: As per my last comment before, can you please quantify 'Slightly smaller'.

Reply: We replaced "slightly smaller" by stating a quantity.

Lines 411 – 413: Rather confusing sentence. Do you mean that the dissimilarities between automatic and manual extraction of fractures are comparable to differences between two manual interpretations? Can you provide any evidence for this? Otherwise can you cite appropriate works?

Reply: Yes, this is what we wanted to express. We revised the sentence for clarity and added a citation of a work that compared manual interpretations of fracture networks to back our statement.

Lines 418 – 421: Please review these three sentences as they are rather unclear. It might be necessary to add a 'the' between 'requires' and 'expertise'. What do you mean by 'several generations are possible for a single fracture'? That a single fracture can be overprinted by a series of tectonic events?

Reply: The sentences were revised according to the suggestions of the other reviewer for clarity. Now it is clear that we wanted to say that several generations can be assigned to a single fracture trace during the interpretation of the generations. Furthermore, we added a more detailed section below that includes a new figure to better explain the narrative besides providing more detailed explanations of the points discussed in the section.

Lines 422 – 423: Can you please clarify? Do you mean that pre-existing fractures can cause distortions in the orientation of later-formed fractures? How widespread such distortions need to be for not be considered just noise in the data (particularly if you have hundreds or thousands of fracture data)?

Reply: Yes, this is what we wanted to express. We rephrased the sentence to make this clear for the reader. In our cases, we have mainly observed this occurrence for old fractures that terminated at their tips within the sampling window. While this might be of minor relevance for the interpretation of generations, it is necessary to be mentioned for the analysis of the fracture network evolution with respect to its connectivity. There, we see that the tips of the old fractures (I nodes) are successively connected to other fractures until barely any isolated tips remain in the network. Therefore, we interpret this connecting of the tips as a common occurrence in this fracture network.

Line 426: Do you mean that fracture length is not a useful parameter to assign one fracture to a specific set?

Reply: Yes, this circumstance is caused by the censoring of the long fractures by our map domains. We revised the sentence for clarity and added further explanations to point out that length as a parameter is biased in the presented case.

Lines 428 – 430: Are you still referring to the fracture set denominated Gen. 1? Not clear.

Reply: Yes, we are. We revised the sentence for clarity.

Lines 430 – 431: This sentence is not very clear and needs to be reviewed. What do mean by 'fractures as different appearances'?

Reply: We revised the sentence according to the suggestion of the other reviewer. Now we clearly state that we refer to the trend of the fractures in this case.

Lines 432 – 433: Can you please explain which one is this 'larger structure' that you are referring to? Not clear to me.

Reply: We refer to the fault as shown in fig, 1b. We have added the reference to the figure and "fault" for clarification. With the revised figure caption of figure 1, this is now clearer.

Lines 436 – 438: In the sentence before this you have argued that gen.1 and gen. 2 can be seen as belonging to one fracture set, however here you assume, without proving it, that these are instead two different fracture sets? Please either refer to a work that shows that gen1 and gen2 were indeed formed at different times, or please provide a full explanation for your assumption.

Reply: We revised the previous sentence according to the other reviewer. The explanation was expanded to further clarify, that we stuck to our initial interpretation because we cannot make a reliable decision in this case simply based on the abutting criteria, which allow several interpretations: either an interpretation as two consecutive generations, or as one generation in which the geometry of the abutting fractures simply represents the order of which fractures belonging to one generation developed. However, either interpretation leads the same result in the following analysis of the network development because the relative order in which the fractures developed remains the same.

Lines 442 – 444: A citation is needed when you mention a mechanical cause for the lack of gen3 fracturing in the NE area.

Reply: The sentence was revised according to the suggestion of the other reviewer. The mentioning of a "mechanical cause" was been removed and we simply refer to the possibility of other reasons that are not within the scope of this study.

Line 449: what do you mean by 'complete distance'? Whole area?

Reply: Yes, we revised the sentence for clarity.

Lines 462 – 466: All these claims are not supplemented by sufficient evidences or by citing relevant works. How was the paleo-stress oriented at time when gen.4 was formed? Are there veins filling gen1 and gen2 fractures that provide cementation for the fractures? If this is the case, why are you mentioning it just now? Otherwise, do you mean that fluids circulating through gen1 and gen2 fractures caused further cementation in the host rocks near the fractures? If this is the case, you should provide evidences or cite relevant works.

Reply: We revised the section accordingly to the suggestions of the other reviewer. Now, we explain the situation more clearly and more simply state examples of possible causes without the need of additional citations that lead to details which are out of the scope of this study.

Lines 467 – 468: This sentence needs to be reviewed as it is not clear.

Reply: The sentence was removed due to the elimination of sinuosity in this manuscript as suggested by the other reviewer.

Lines 476 – 477: Can you please further explain the meaning of 'decreasing skewness'? How does this statistical parameter relate to the geometry of the fracture network?

Reply: We added a sentence to briefly explain the reference, now also provide a clear link to the figures and tables showing the data. More details are now presented earlier in the revised section 4.1.

Lines 4879 – 481: I feel that it needs to be made clearer how the variations in the count of Y nodes and I nodes are related.

8

445 Reply: This part of section is supposed to discuss trends in the data visible throughout our domains and not intended to focus on the relationships of nodes. The appearance and relationships of different node types is discussed in detail a few lines below, still within the same section.

Line 483 – 484: How can you establish that a representative domain has been sampled?
450 Can you show examples using your study case? It can be easy to argue that on outcrop the possibility of sampling a 'complete fracture network' are relatively scarce and relegated to few ideal outcrops.

Reply: This is the point we wanted to make at this position to highlight the necessity of a
455 complete interpretation of the network. We present the criteria based on which we have selected our domains in section 2.2 and added further text at this part of the discussion to debate what aspects need to be considered when the domains are supposed to be representative for the complete network at this point.

460
Line 502: What do you mean by 'undirected fractures'? Can you please clarify?

Reply: We referred to generation 5 fractures which do not follow a clear orientation mode in all areas. We removed "undirected" to avoid confusion and revised the statement to one of a more
465 general nature.

Lines 504 –505: Can you please provide examples that can prove this claim? And, can you please produce a conceptual model that exemplifies the described process?

470 Reply: This claim was non-unique and unsupported and not necessary to aid the narrative of the manuscript. Therefore, we have removed it from the manuscript according to the suggestion of the other reviewer.

Line 511: Do you mean that all subareas show comparable branch lengths? As now
475 written is a bit unclear.

Reply: Yes, we revised the sentence for clarity.

Conclusions: While I like bullet points in the conclusion to show the main findings, I
480 also feel that you should have few sentences that wrap up and recap your work. In addition, some of the points listed are a bit vague, specifically point 3 and point 5, as written, do not add any value to the work. So please either reformulate or delete them.

Reply: We added a introduction to the conclusion to recap our work presented in the
485 manuscript. To account for the comment of the other reviewer on this chapter, we have also revised the bullet points. Former point 3 was reformulated and point 5 deleted.

Figure 1. Last sentence is not clear. Do you mean areas the labels show the areas where you have acquired the data?
490
Reply: The figure was revised earlier, and the changes of the captions were lost. We updated the caption, it now states that the map domains in which we traced the fractures are marked as the yellow (not red) squares in the figure.

495 Figure 4. Please add that P21 indicates 'Fracture Intensity'.

Reply: This clarification was added to the caption.

Figures 5 and 6: This is just a personal taste, so you can ignore it, but I would find
500 useful if you can show in these figures the location of the analysed tiles.

505  Reply: The location of the analyzed tiles is shown in figure 1. This is now clarified by the revision of the caption of figure 1. Therefore, we do not include their locations additionally in the figures as suggested, because this would a repetition and furthermore reduce the size of the figures with the important content at this place, leading to a worse readability.

Figure 7. Since this plot refers to the automatic extraction method, can you please include in the caption that these are branch lengths and not trace (whole) fracture lengths? AS mentioned in the comments, I would suggest integrating this figure with
510  the length distributions in the manually extrapolated network.

Reply: The caption was revised accordingly, and an additional figure showing the length distributions of the manually traced segments was added to the manuscript as suggested.

515  Figure 8. Similar to previous comment. Please indicate that these orientations refer to branches.

Reply: We now state explicitly that we are referring to branches in the caption.

520

525

530

535

540

545

550

555

# Replies to 'Review of SE-2020-67 by William Dunne'

We thank William Dunne for the detailed and very constructive comments on the manuscript.

A major point of his comments is that the document needs to achieve a stronger alignment of its actual text with its purpose of comparing the manual and automatic methods, because the comparison is not sufficiently developed. We agree and carefully revised the manuscript to better develop this point throughout the whole narrative, particularly in the sections comparing the two methods.

Further key matters that needed to be addressed are the following ones:

(1) actual presentation of data related to time required to perform each method.

Reply: In the revised version of the manuscript, we extended the narrative at several points to compare the different times required for both methods, discuss different parameters that influence the required time and provide estimates to map the complete outcrop using both techniques.

(2) a more effective utilization of fracture intensity (P21 in this case) as a discriminator of quality and/or match between the two data sets, including a careful explanation and use of the attributes that compose P21

Reply: We revised and extended the sections presenting and discussing P21. We now provide more details on the method itself, explain the attributes composing P21, how it can be applied to our dataset and what the results indicate with respect to the different methods. We have added an additional figure to show the differences of P21 for both methods and discuss them in detail. To further back the P21 analysis, we also included a preceding section where we introduce and discuss P20 (fracture trace densities) with additional figures as preceding section to the P21 section. Overall, we provide a more detailed explanation on the Pij system and compare and analyze the results of P20 and P21 with respect to the different methods and the structure of the data of fracture trace segments.

(3) avoid two unsupported interpretations in the Discussion;

Reply: We removed any unsupported interpretations throughout the manuscript and provided further explanations and citations for interpretations that are supported but were lacking the necessary arguments in the previous version of the manuscript.

(4) rebuild the Conclusions so that they are actually about the purpose of the manuscript and highlight key outcomes related to that purpose.

Reply: We revised the conclusions and added a short introduction to recap our work. The bullet points were extended and rearranged, to first evaluate the raw data, then compare the aspect of time required for both methods followed by the differences and similarities (P20, P21) and pros and cons of both methods. Once the main key findings with respect to the different methods are addressed, more general points are listed that highlight the particular findings for the particular fracture network throughout the domains.

Detailed comments on the individual points raised by the reviewer are provided below.

Major Comments:

Section 3.2, Line 150 and elsewhere in the text - Repeatedly and with good reason, the point is made that the manual tracing of the fractures in the

11

digital maps is very time consuming and limits a researcher's ability to effectively utilize the large data sets that can now be created through tools such as UAVs and digital imaging equipment, so the development of effective automatic characterizing protocols will greatly enhance the amount and quality of information for analysis by researchers.

615 Yet, the manuscript does not offer any data to justify this statement. For example, stating the amount of time needed to create each of the five manual samples vs. the five digital samples should be simple and effective. Further, then a comment/short narrative could be added into the discussion about the amount of time that would be needed to characterize the entire "bench" automatically and how, for such a modest

620 amount of time, one would have a much richer data set to tackle........ Consequently, the manuscript should be revised to explicitly document the difference in time usage between the two methods, and then should consider the implications of having the quicker, more powerful automatic approach in the discussion.

625 Reply: We agree that these points are very interesting for the reader. Therefore, we added a narrative in the methods parts explaining the required time under different aspects for each of the two methods in detail, incorporating the tracing itself and the time required for the removal of artifacts. Further points were added in the discussion giving an estimate what time would be required to trace the whole dataset using either technique and discuss

630 advantages and disadvantages of both methods with respect to the required time in different scenarios. Also, we have added a short explanation to highlight the pros and cons for someone who has never done fracture tracing following the suggestion of the other reviewer. Following the revisions throughout the manuscript, we revised the conclusions accordingly to better present the most important findings with respect to the time required using both

635 methods.

Line 156 - It is very important to explicitly state that the manual data set is derived the digital imagery and was not collected in the field (Correct?). This point in the text is a good place to present this point clearly.

640

Reply: We agree and now state explicitly, that the data was derived from the digital imagery and that we used ArcGIS as requested by the other reviewer.

Lines 224 to 225, Fig. 4, and S1-P21 difference illustration - This comparison of fracture

645 intensity between the manual and automatic data gathering approaches is presented as being a primary tool for relating the information and quality from the two data sets. Yet, this comparison is significantly underexplained. For example: (1) Why is a P21 comparison such an effective choice for comparing the two data sets? (2) Just how good is the match, particularly as nothing is said to explain and/or characterize the

650 difference illustration in the S1 illustration? (3) Why is no basic explanation provided of what a reader is seeing in Fig. 4, such as the use of meter-square sample areas, or consideration of the sensitivity from P21 varying spatially from 0 to 18 m-1? (4) Why the lack of an actual narrative comparing the two sets of imagery qualitatively, particularly if locations exist where the match is poor and that needs explanation to, for example,

655 show the overall strength of the approach?

Reply: We agree that the narrative was underdeveloped int the section and revised it carefully. Besides $P_{21}$ we now also provide examples of $P_{20}$. Both are now properly introduced along with explanations why we use them and what their results tell us about the network. Explanations

660 and examples are added for the new and existing figures to back the narrative and make it more comprehensive for the readers.

Lines 226 to 227 - An expectation is presented that a larger number of traces would generate a greater fracture intensity (P21) than a smaller number of traces. This expectation is not well rooted. Intensity P21 is a function of total trace length in a sample area, and that is a function of both the number of fracture traces AND their individual lengths (the authors show an understanding of this point in Lines 234/235, but have not utilized that understanding here). So, a fracture population with more traces may, in fact, be less intense because its traces are individually shorter than the other population with fewer but longer traces. Therefore, any expectation of differences in P21 between the two fracture trace populations would need to consider both the number of traces and some aggregate representation of the population of lengths, such as the mean length. Further, if the number of traces is thought to be the key parameter, a much more detailed presentation about the total number of traces in each sample window for each of the two sampling procedures is needed. Overall, the underlying logic of this comparison needs to be better developed and then more completely explained, if utilized.

Reply: We agree that our reasoning in the old version of the manuscript was not well presented one extended and clarified the statements, along with a more detailed narrative, additional figures and an extended Table 1. Besides $P_{21}$ we now also provide examples of $P_{20}$ (see the reply to the prior comment) to compare the different numbers of traces generated by both methods in the sample windows for a better comparison of the trace segments as they result from both methods. We now explain that we use $P_{21}$ to compare the trace lengths present in the sample areas, which allows us to better compare the different traces generated by both methods with respect to their lengths representing the fractures, giving us better control over the interpretation of the results of both methods. This topic is now addressed within a whole sub-section to provide better and more extensive explanations on how we utilize Pij and P21 in that particular section, and how we use it to compare the different methods.

Line 367 - Given that the discussion covers "classification into fracture generations" and "network analysis", it is not at all clear why "Passchier et al. (XXXX)" needs to be cited, particularly as it is a very shaky citation with no status or occurrence in the references.

Reply: In Passchier et al. (in prep.) we introduce the initial interpretations of the generations on a larger scale of the outcrop, which helped to identify the criteria on which we base our interpretation of the generations in this manuscript. However, the criteria we used and present in this work are self-sustained, explained in detail in section 3.4 and do not require the citation of Passchier et al. (in prep). We now refer to this citation as a companion paper only once and earlier in the manuscript and rewrote the narrative to clarify, that the interpretation of the generations is based on the criteria explained in this manuscript.

Sinuosity values vary so little from one sample window to another and with respect to the sequence of fracture development in the sample window patterns, would it not be better to eliminate all description/discussion of sinuosity from the manuscript, so as to simplify and focus it?

Reply: We agree that the differences of sinuosity, spatially or between the methods, are minor and not significant, thus, we eliminated sinuosity from the manuscript entirely, because it does not add valuable information.

Lines 411 to 413 - An important comparison of difference is made in this sentence, Yet, no evidence or citation from other work is provided to support that this comparison is
720  correct. So, the statement is an unsupported speculation and really needs to be better than that for the purposes of this manuscript.

Reply: We infer that the automatic code at this stage represents a good option for creating an initial fracture trace map that only differs from a manual interpretation to a degree that is
725  comparable to the deviation of two manual interpretations of the same fracture network. This assumption is backed by e.g. Long et al. (2018), who compared different manual interpretations of fracture networks and is now cited in the revised version of the manuscript. The sentence was revised for clarity and a citation added to support our point.

730
Lines 419 to 421 - Here is another key point that is incompletely developed and explained. Particular examples of "human bias" should be identified with explanation.
Then the highlighted text can be eliminated and replaced with text that has greater meaning and clarity. Further, the replacement text will need to be a few sentences
735  rather than just one sentence, given the importance of this point
.
Reply: We agree that is an important point. We extended the section and added an additional figure to provide several particular examples and more detailed explanations discussing human bias, e.g. several examples of non-unique interpretations to aid the narrative in the new
740  section. The highlighted text was replaced by a more elaborate section.


Line 516 and Conclusions - Given the title of the manuscript and the setup of the abstract and introduction, these conclusions show a surprising lack of content related to
745  comparing manual and automatic methodologies. The conclusion should be reorganized to begin with comparison outcomes (e.g., time usage, P21 comparison, managing manual input into automatic interpretation by parameter selection, etc.). It should delete any text related to superiority of manual to automatic unless a substantial addition is made to the manuscript about that matter. Then it can list particular outcomes
750  for the samples of this particular fracture pattern in this particular location and geological setting, which are not the central focus of the contribution of the manuscript.


Reply: We have added a short introduction to the conclusion to recap our work presented in the manuscript. To account for this comment, we have also revised the bullet points.  Former
755  point 3 was reformulated and point 5 deleted. The bullet points were extended and rearranged, to first evaluate the raw data, then compare the aspect of time required for both methods followed by the differences and similarities (P20, P21) and pros and cons of both methods. Once the main key findings with respect to the different methods are addressed, more general points are listed that highlight the particular findings for the particular fracture network
760  throughout the domains. Throughout the conclusions we do avoid any text related to superiority of manual to automatic trace extraction that is not discussed in detail during the manuscript as advised.

Lines 423 to 425 - Again, examples with identification in figures are needed to support and
765  document this point.

Reply: See comment to lines 419-421: We revised and supplemented this section by the addition of another figure to give more detailed examples and for a better documentation of the narrative.

14

770

Comments:
Line 21 - As this paper has a focus on methodology, the abstract should
briefly present an explanation here as to why automatic assignment of fractures into
generations cannot yet be done.

775

Reply: We agree and added an explanation to the abstract to briefly explain the differences of
the methods causing this circumstance.

Lines 31-32 - The fracture geometries are set and not "evolving" so the explanation for
780    the connectivity increase needs to be replaced/improved.

Reply: We revised the sentence to clarify, that each domain has slightly different fracture
network characteristics and greater connectivity occurs where the development of later, shorter
fractures has been distorted less by the abundance of pre-existing, longer fractures as
785    observed in our data.

Lines 51 to 53 - This sentence should be moved to the end of this section (Line 59), so
that this paper's purpose and approach is stated completely first.

790    Reply: We moved the sentence accordingly.

Line 52 - What is the status of this companion paper (it is not in the reference list)?
Can it be cited or does reference need to be made an unpublished source? Or?

795    Reply: In Passchier et al. (in prep.) we introduce the initial interpretations of the generations
on a larger scale of the outcrop, which helped to identify the criteria on which we base our
interpretation of the generations in this manuscript. However, the criteria we used and present
in this work are self-sustained, explained in detail in section 3.4 and do not require the citation
of Passchier et al. (in prep). We now refer to this citation as a companion paper only once and
800    earlier in the manuscript and rewrote the narrative to clarify, that the interpretation of the
generations is based on the criteria explained in this manuscript.

Line 84 - Joints and not jointing are unfilled.
Reply: We replaced "jointing" with "joints".
805

Line 104 - "v" - The significance of the radial pattern to the NE is not provided, so this
text is superfluous to the later part of this point. If the radial pattern has significance, it
should probably be a separate "vi".

810    Reply: We assigned the radial pattern to an own bullet point, because the results highlight its
significant impact on the fracture network connectivity. The position of the pattern we refer to
in the domains and the bench is now clarified by referring to our domains NE1 and NE2.


815    Lines 132 to 140 - Text revised to create a narrative that more contrast imagery results
for flight altitudes of 100m or 10m vs. 25m, including the removal of the green
highlighted text.

Reply: All suggested revisions were implemented.
820

Lines 133 to 137 - This highlighted text as written breaks up the narrative flow. It should
likely be added to the end of the text on Line 127.

Reply: We moved the text accordingly.

825    Lines 145 to 148 - This text should be moved up to the end of Line 127 to complete the general description of methodology. Thus, the subsection will finish with the determination of the optimal flight height and the creation of the overview.

Reply: We moved the text accordingly.
830

Line 154 and elsewhere in the text - Please remember that "this" and "these" are neither pronouns or nouns.

Reply: We reviewed the manuscript and corrected this mistake throughout the whole text.
Line 172 - The intensity belongs to the product, the fractures, and not to the process in
835    this case (clearly, fracturing can be intense, but that is not the intended meaning here, as the description is of a "final fracture population".).

Reply: We agree and replaced "fracturing" with "fracture" for clarity.

840    Line 201 - Again, can this contribution actually be cited? What is its status? If it does not have an accepted or published status, what can be used as a citation in its place?

Reply: In Passchier et al. (in prep.) we introduce the initial interpretations of the generations on a larger scale of the outcrop, which helped to identify the criteria on which we base our
845    interpretation of the generations in this manuscript. However, the criteria we used and present in this work are self-sustained, explained in detail in section 3.4 and do not require the citation of Passchier et al. (in prep). We now refer to this citation as a companion paper only once and earlier in the manuscript and rewrote the narrative to clarify, that the interpretation of the generations is based on the criteria explained in this manuscript.
850

Line 214 - Again, can this contribution actually be cited? What is its status? If it does not have an accepted or published status, what can be used as a citation in its place?

855    Reply: In Passchier et al. (in prep.) we introduce the initial interpretations of the generations on a larger scale of the outcrop, which helped to identify the criteria on which we base our interpretation of the generations in this manuscript. However, the criteria we used and present in this work are self-sustained, explained in detail in section 3.4 and do not require the citation of Passchier et al. (in prep). We now refer to this citation as a companion paper only once and
860    earlier in the manuscript and rewrote the narrative to clarify, that the interpretation of the generations is based on the criteria explained in this manuscript.


Line 222, Section 4.1 - This section would benefit from a revised title and the addition
865    of two subsection headings to facilitate reader understanding: (1) The entire section is about P21 so change the title to be explicit about trace intensities rather than just traces. (2) 4.1.1 Intensity comparison between two methods (3) 4.1.2 Characterization of intensities for automated data - these two subsection titles clearly separate the purposes of these two portions of the text.
870

Reply: We agree that the section required a revision and rewrote the text to more clearly compare both methods and took special care to avoid possible confusion of which method we are referring to respectively. Since we also added an analysis using $P_{20}$ the manuscript, the new sections were titled as follows: 4.1 "Fracture trace segments", in which we discuss general
875    statistics of the segments created by both methods along with a new figure as suggested by the other reviewer. The new subsections are titled 4.1.1 "Fracture trace segment densities", which compares $P_{20}$ for both methods, 4.1.2 "Fracture trace segment intensities", which compares $P_{21}$ for both methods was revised to better explain $P_{21}$ as suggested in an earlier comment.

880 Previously also included in 4.1 were the results of the network topology analysis of the automatic traces. They are now separated more clearly from the previous comparisons in a new section 4.2, that precedes the manual interpretation and network analyses.

Line 224 - Not everyone knows P21, so here is a good location to simply and explicitly
885 define the term. Note: P21 tends to be our best available measure of fracture abundance for natural rock outcrops and does have wide usage, but it is not universally known.
Reply: We revised the complete section (see replies to earlier comments) and now properly introduce the $P_{ij}$ system, in particular $P_{21}$ and $P_{20}$ to the readers along with more detailed
890 explanations of why and how we use both to compare the segments of both tracing/extraction methods.

Lines 227 to 231 - If this text is meant to explain why the P21 intensity for the automatically
895 collected data as compared to the manually collected data, it does not achieve that outcome. What is this text attempting to say? It is not clear?

Reply: This text was supposed to explain that the automatic trace extraction is expected to result in a larger number of segments, because fracture traces are segmented at intersections,
900 while a manual interpreter traces complete fractures from tip to tip. We revised the sentences for clarity and moved them out of the P21 section into the section providing more general differences of the networks resulting from the different methods to avoid confusion (also due to the restructuring of section 4.1).
More clear comparisons of P21 are now provided in the new section, along with a new figure
905 (previous supplement) to further help the narrative and better compare both methods of fracture trace mapping, also with an extended table 1.

Lines 231 to 233 - This sentence is not about the P21 comparison between manual and automate, but rather a description of the differences in P21 for the automated data set.
910 It should be the first sentence of the next paragraph, which is about spatial variations in P21 for the automated data sets.

Reply: We have revised and restructured section 4.1 according to an earlier comment. This sentence belongs to the paragraph dealing with the P21 analyses, which are primarily used to
915 compare both methods. The sentence commented on is separated out at the end of the paragraph to avoid confusion with the comparisons between the methods, because it describes differences in domains and not between methods. It now serves as connection to the following section which discusses the spatial variation within the domains.

920
Lines 239 to 241 - The description of maximum length needs to allow for censoring by the sample window perimeters. The identified maximum lengths cannot be stated to be the actual maximum length of any trace that is sampled in a window because of censoring. Now, it is possible to consider the maximum length of traces that are fully
925 contained in a sample window, but that needs to be stated explicitly.

Reply: We agree, this is correct. We added sentence to clarify that maximum lengths may be censored by the sampling windows. This point is also examined later in the manuscript in the discussion chapter.
930
Line 254 and Figure "18" - To preserve order of figure citation, Figure 18 should become Figure 8, and Figures 8 to 17 should be renumbered.

17

Reply: We have added several new figures to the manuscript and updated the numbering accordingly. The figure referred to in this comment is presented earlier as suggested in the revised version of the manuscript.

Line 259 and new Figure 9 (old Figure 8) - This figure illustrates length-weight fracture trace abundance as a function of orientation, so it does not show "fracture strike" in any simple manner. The text is revised in this line to better describe what is being illustrated.

Reply: We agree and implemented the suggested revisions.

Lines 264 to 265 - This sentence is an interpretation of observations before the observations are fully presented, so it is out of place and should be deleted.

Reply: We agree and deleted the sentence.

Lines 265 to 266 - Redundant and unneeded sentence

Reply: We agree and deleted the sentence.

Line 281 - How are these "denser clusters" recognized or statistically defined? A reader must be able to use the criterion/criteria to identify the clusters themselves. If reproducible methodology does not exist, the statement about orientations modes for clusters should be deleted.

Reply: Our statement was based on a qualitative visual interpretation of the plot of fracture length vs orientation. We agree that this is not a good foundation to describe the distribution of gen. 5 and therefore deleted the sentence.

Line 288 and Line 289 - text revised in these two locations to more completely and correctly describe the pattern characteristic of the fractures in the NE1 sample window, and in the NE windows vs. the SW windows, respectively.

Reply: We agree and implemented the suggested revisions.

Line 289 - text revised to more simply and clearly describe the lack of a relative age relationship.

Reply:  We agree and implemented the suggested revisions.

Lines 291 to 293 - Sentence revised to more clearly identify and state differences with citation of Table 4 being placed in the parentheses at the end of the sentence to not distract from sentence meaning.

Reply: We agree and implemented the suggested revisions.

Line 301 - Adding the word "evolution" to this sentence is important for clarity and meaning. Throughout the text – "mapping boundary" should be "map boundary"

Reply: We agree and added "Evolution" to the sentence and. We replaced "mapping boundary" with "map boundary" throughout the manuscript.

Lines 369 to 370 - Are these two sentences needed here? Their contents are not

18

990 used immediately. Their content should be added where it is needed further into the document.

Reply: We agree and deleted the sentences at this part of the manuscript. References to Table 10 and the figure are placed at another position where they better complement the narrative
995 further into the document.

Lined 373 - Cite published work that supports the lack of sampling bias and similarity of results for exposure of this quality for different operators. Solid Earth has a publication about this matter, for example.
1000

Reply: Citations on this topic were already provided earlier in the manuscript. To better back our narrative in the lines commented on, we have also added the citations at this place.

Line 382 - "excessive" requires some more explanation and/or examples to assist
1005 reader comprehension

Reply: We added an explanation with examples to assist reader comprehension on the topic in the following line. Even more details on the topic are now discussed in the manuscript supported by an additional figure in chapter 5.2.
1010

Lines 388 to 395 - Not convinced that this discussion of the comparison of P21 between the manually and automatically acquired datasets is particularly useful because it considers number of traces separately from the distribution of tracelengths, which is
1015 somewhat arbitrary.

Reply: In the revised version of the manuscript we now present $P_{20}$ earlier in the manuscript to also consider segment density (number of traces/unit area) and provide a direct comparison of the numbers of traces in the revised Table 1. However, in the box-counting method for $P_{21}$,
1020 the box considers the length of cut segments per unit area independently from the number of segments (graph edges) and the manually traced fractures have been pre-processed earlier to resemble the same data structure as the automatically traces ones. Therefore, we deem this to be a valid comparison. To avoid confusion, this is now clearly stated earlier in the manuscript.
1025

Lines 411 to 413 - An important comparison of difference is made in this sentence, Yet, no evidence or citation from other work is provided to support that this comparison is correct. So, the statement is an unsupported speculation and really needs to be better than that for the purposes of this manuscript.
1030

Reply: The sentence was revised according to the other reviewer and we added citation of a work that compared manual interpretations of fracture networks to back our claim. We infer that the automatic code at this stage represents a good option for creating an initial fracture trace map that only differs from a manual interpretation to a degree that is comparable to the
1035 deviation of two manual interpretations of the same fracture network (e.g. Long et al., 2018).


Lines 426 to 438 - This paragraph needs an introductory sentence to establish its purpose and why three different characteristics are being "juxtaposed" in one paragraph.
1040 An example sentence is offered.

Reply: We implemented the example sentence as suggested.

Lines 452 to 453 - This one sentence paragraph is not really needed as written. A
1045 clause is added to the opening sentence about generation 5 to preserve overall narrative

flow.

Reply: We agree and implemented the revisions as suggested.

1050 Line 460 to 461 - Text revised and added, so as to achieve greater clarity with a more complete and needed explanation for readers.

Reply: We agree and implemented the revisions as suggested.

1055 Line 464 to 466 - Text revised to more clearly explain situation and to more simply state examples of possible causes.

Reply: We agree and implemented the revisions as suggested.

1060

Lines 467 to 473 - Recommend the elimination of this text because sinuosity is not a distinguishing characteristic for this study.

1065 Reply: We agree, the text commented on and the whole aspect of sinuosity was eliminated from the manuscript as suggested in an earlier comment.

Lines 504 to 505 - Interpretation is non-unique and unsupported, so it should not be included in the manuscript.

1070 Reply: We agree that this is indeed an unsupported statement and deleted the interpretation from the manuscript.

Please see the annotated PDFs for the main text and figures for additional detailed
1075 comments about the syntax.

Reply: We would like to thank the reviewer William Dunne for the detailed remarks on the syntax provided in the annotated PDF, we have corrected the syntax accordingly to the comments and suggested revisions provided therein.

1080

Captions: Figure 1 – Lines 724-725 states "The study areas on "the bench" that are mapped in detail are marked in 725 red." Yet, the only red in the illustration are the red lines for faults. Is the color incorrectly identified or is something missing from this illustration?

1085

Reply: The figure was revised prior to the version presented in the manuscript and the change of the captions was lost. We corrected the caption to state that the map domains are marked as the yellow (not red) squares in the figure.

1090 Figure 2 - Line 731 - the object in 2a to 2e should be identified in the caption. Also, 2a to 2f should have scale bars even if they are approximate.
Reply: We added a reference to the persons in the images and scale bars for the images on the left.

1095 Figure 4 - Line 737 - Please note suggested addition to the text at the end of the caption.

Reply: We implemented the suggested addition in the caption.

1100 Figure 7 - the labels for the X and Y axes of the three sets of graphs should be larger, so that they are easier to read when viewing the entire figure.

Reply: We enlarged all labels on the axes for a better readability.

1105  Figure 9 (10) caption improved by explicitly identifying that data belong to manually collected set and providing a better description of the rose diagrams.

Reply: We implemented the suggested revisions of the caption.

1110  Figure 10 (11) Caption - Revise to match revised Figure 9 (10) caption. Eliminate the addition sentence because the network cannot be described as being "oriented NW-SE". The pattern for NE2 can be described as having a strong length-weighted orientation mode that trends NW-SE, but even that information is likely not needed in the caption here.
1115
Reply: We revised the caption match the one of the associated figure, the numbering of the figures was updated to match the new order (now figures 14 and 16).
In the second part of the caption, we were not referring to the orientation of the actual network, but the map domains themselves. We have revised the caption for clarity.
1120
Figure 11 (12) - the key at the bottom of the figure should use larger dots so that the colors are easy for readers to distinguish. While these solid circles will be much larger than the actual ones in the plot, that is not an issue because assigning the colors easily to the generations for the readers is the goal.
1125
Reply: We agree and enlarged the dots in the key of the figure.

Figure 13 (14) caption - make the change in the first sentence of the caption for Figures 14(15), 15(16), 16(17), and 17(18).
1130
Reply: We made the suggested revisions for all mentioned figures and a remark was added for the NE domains that gen 3 fractures were not identified within these two domains.

Tables - Orientation data in all tables - These data should be rounded off to the nearest
1135  integer. The use of two significant figures to the right of the decimal point is false precision, particularly for the manually traced lines.

Reply: We agree and rounded the orientation data in all tables off to the nearest integer.

1140  Table 1 – Title expanded to explicitly identify that the characteristics stated in the table related to the automatically generated fracture network

Reply: We have added the characteristics of the manual interpretation to the table to provide a more direct comparison. The title of the table was revised according to the comment and the new content.
1145




1150




1155

# List of relevant changes in the manuscript

**Abstract**

- We added an explanation why automatic assignment of fractures into generations cannot yet be done, because this is an important aspect of the methods in the focus of the manuscript.

**2.2 Data set**

- To better explain the reasoning for the selection of our domains, we split one of the points listed in two and added two more points.

**3.1 UAV**

- Sentences were rearranged within the section for a better narrative.

**3.2 Fracture trace mapping**

- We have added several sentences for a better and more detailed narrative around the time aspect of both methods and the steps involved in each.

**3.3 Network analysis**

- We have added more detailed explanations for
  - the software and dataset used for the manual tracing,
  - the polygon sampling approach,
  - our use of the term "spatial graph" along with a citation,
  - the processing steps to make the datasets of manually and automatically-traced fractures comparable
  - the post processing steps to correct error nodes and branches from the NetworkGT toolbox and the spatial join function used during the corrections.

**3.4 Manual classification into fracture generations**

- We removed the citation of the companion paper Passchier et al. (in prep), because the interpretation of fracture generations in this manuscript is self-sustained and self-explanatory, thus, does not require a citation of this so far unpublished work.

**4.1 Fracture traces**

- We revised the complete section to provide a clearer structure of the narrative and more detailed explanations to compare the results of the methods. We took special care to avoid confusion between the two methods by revising the narrative and clearly stating which one of the methods we are referring to.

- We added several figures as suggested by the reviewers and an analysis of the fracture segment density P20 to support or analysis of the fracture segment intensity

22

P21 for both methods. We expanded the narrative around both methods and use the new figures to more clearly explain and highlight differences and similarities of the results.

1210
- The section was split into subsections, of which 4.1 compares general statistics of the fracture segments as they result from manual and automatic tracing. We added a new figure with the same plots for the manually traced segments as previously only provided for the automatically extracted segments. We evolved the narrative around the figures and now better explain the meaning of the statistical parameters with
1215     respect to the networks.

- We added a more detailed introduction and explanation of the Pij system.

- The new sub-section 4.1.1 is used to present, explain in detail and compare the
1220     results of both methods for the P20 segment density in two selected domains using two new figures and references to new figures in the supplement showing P20 for all domains.

- The new sub-section 4.1.2 presents the P21 analyses of both methods and all
1225     domains as previously present in the manuscript. We added several sentences to better explain and justify our use of P21, guide the reader through the figures and highlight differences and similarities of the methods.

- The previously presented network characteristics of the automatically traced
1230     segments with respect to the map domains are now presented in a separate section 4.2


**4.3 (former 4.2) with a revised title: "Manually interpreted fracture generations and their**
1235 **spatial variation" to better match the content of the section**

- The numbering of the sections was updated due to the restructuring of the former section 4.1

1240
- Sinuosity as parameter for the trace segments was eliminated in this section and the entire manuscript as suggested


**5.1 Manual vs automatic tracing**

1245
- We added a detailed discussion of the time required for tracing the segments using the two methods along arguments for or against either method under different conditions or with different goals in mind. Furthermore, we estimate and discuss how long it would take for either method to completely map the whole outcrop.

1250
- We added further points for a more detailed discussion of P20 and P21


**5.2 Classification into fracture generations**
1255

- To better develop and discuss the point of human bias when tracing and interpreting fracture generations, we added a new figure (24) and further details, explanations,

23

and discussion points. We now show examples of human bias while tracing and interpreting fracture generations in the dataset and discuss them in detail.

### 5.3 Network analysis

- We added a short section to discuss the selection of representative domains.

### 6. Conclusions

- The conclusions were revised according to the now better developed narrative around the time aspect of the methods and the comparison of the methods

- We added an introduction to review the presented work.


- Former point 3 was reformulated and point 5 deleted. The bullet points were extended and rearranged, to first evaluate the raw data, then compare the aspect of time required for both methods followed by the differences and similarities (P20, P21) and pros and cons of both methods. Once the main key findings with respect to the different methods are addressed, more general points are listed that highlight the particular findings for the particular fracture network throughout the domains.

### Figures:

- Former figure 18 is now presented earlier

- Former supplement S1 is now presented in the manuscript (new fig. 9)

- Additional figures were added: 5, 6, 7, 24

### Tables:

- Orientation data in the tables was rounded off to the nearest integer

- Data of the manually traced segments was added to Table 1

# Mapping the fracture network in the Lilstock pavement, Bristol Channel, UK: manual versus automatic

Christopher Weismüller[1], Rahul Prabhakaran[2,3], Martijn Passchier[4], Janos L. Urai[4], Giovanni Bertotti[2], Klaus Reicherter[1]

[1] Neotectonics and Natural Hazards, RWTH Aachen University, Aachen, Germany
[2] Department of Geoscience and Engineering, Delft University of Technology, Delft, the Netherlands
[3] Department of Mechanical Engineering, Eindhoven University of Technology, the Netherlands
[4] Structural Geology, Tectonics and Geomechanics, RWTH Aachen University, Aachen, Germany

*Correspondence to* Christopher Weismüller (c.weismueller@nug.rwth-aachen.de)

**Abstract.** The 100,000 m² wave-cut pavement in the Bristol Channel near Lilstock, UK, is a world-class outcrop, perfectly exposing a very large fracture network in several thin limestone layers. We present an analysis based on manual interpretation of fracture generations in selected domains and compare ~~this~~ it with automated fracture tracing. Our dataset of high-resolution aerial photographs of the complete outcrop was acquired by unmanned aerial vehicle, using a survey altitude optimized to resolve all fractures. We map fractures and identify fracture generations based on abutting and overprinting criteria~~,~~ and present the fracture networks of five selected representative domains. Each domain is also mapped automatically using ridge detection based on the complex shearlet transform method. The automatic fracture detection technique provides results close to the manually ~~mapped~~ traced fracture networks in shorter time, however, with a bias towards closely spaced Y over X nodes. The assignment of fractures into generations cannot yet be done automatically ~~yet~~ because the fracture traces extracted by the automatic method are segmented at the nodes, unlike the manual interpretation in which fractures are traced as a path from fracture tip to tip consisting of several connected segments. This segmentation makes an interpretation of relative age impossible because the identification of correct abutting relationships requires the investigation of the complete fracture trace following a clearly defined set of rules~~.~~ Generations one and two are long fractures that traverse ~~our~~ all domains. Generation three is only present in the southwestern domains. Generation four follows an ENE-WSW striking trend, is sub-orthogonal to generations one and two and abuts on them and generation 3, if present. Generations five is the youngest fracture set with ~~diffuse~~ a range of orientations, creating polygonal patterns by abutting ~~on all older~~ at all other fracture generations. Our mapping results show that the northeastern domains only contain four fracture generations, thus the five generations of the outcrop identified in the southwestern domains are either not all present in each of the five domains or vary locally in their geometry, preventing the interpreter to link the fractures to their respective generation over several spatially separate mapping domains. ~~The domains have different P₂₁ f~~Fracture intensities ~~which is lowest in~~ differ between domains where the smallest is in the NE with 7.3~~1~~ m/m² and ~~highest~~ greatest in the SW with 10.~~01~~ m/m², coinciding with different fracture orientations, and distributions of abutting relationships. Each domain has slightly different fracture network characteristics~~. The network analysis shows that~~ and greater connectivity ~~connectivity increases, depending on the evolving fracture generations.~~ occurs where the development of later, shorter fractures is not affected by the stress shadowing of pre-existing, longer fractures.

## 1. Introduction

Recent technological advances allow us to collect large amounts of remote sensing and outcrop data, e.g. using LiDAR, unmanned aerial vehicles (UAV), and structure from motion (SfM) tools (Bemis et al., 2014; Bisdom et

al., 2017; Hansman and Ring, 2019; Müller et al., 2017; Niethammer et al., 2012; Vasuki et al., 2014; Weismüller et al., 2019; Westoby et al., 2012). These high-resolution datasets can be acquired in short periods of time, optimizing resolution and enabling the detailed digitalization of km-scale outcrops in sub-cm resolution. The mapping and interpretation of features in the datasets is time-consuming because ~~of~~ these very large areas ~~available which~~ must be mostly interpreted manually. Automated tools can aid during the ~~process~~interpretation, but do not yet match the quality and reliability of manual interpretations, and thus must be used with care (Duelis Viana et al., 2016; Vasuki et al., 2014).

Several sampling techniques have been developed to extract data ~~on~~ about fracture sets that do not necessarily require a complete mapping of the whole area, e.g. line sampling (Priest and Hudson, 1981), polygon or areal sampling (Wu and D. Pollard, 1995), circular scanline sampling (Mauldon et al., 2001; Rohrbaugh et al., 2002; Watkins et al., 2015) or rectangular window sampling (Pahl, 1981).

In this work, we use a UAV to collect aerial images of the fractured limestone at the Lilstock coast in the Bristol Channel, UK. Sample pictures and tests from varying height in combination with ground-truthing show that we can resolve all visible mode I barren fractures in the pavement, which are the target of this study. With the right compromise between flight altitude and resolution, we photographed the whole outcrop of 400x150 m within two days with similar lighting conditions, creating a set of raw images depicting a fracture network which is too large to be manually interpreted. ~~In a companion paper (Passchier et al., XXXX) we present a fracture mapping study, and propose criteria for identifying the different generations of fractures.~~ Our aim in this study is to provide a complete interpretation and fracture network analysis of the selected sub-areas of the outcrop, providing a benchmark of our analyses. We hand-interpret several domains and use them as basis for supervised automatic tracing to create a large fracture dataset. We present our interpretation of several fracture generations in the selected areas to review the evolution of the fracture network in time steps, using topological branch and node analysis (e.g. Dimmen et al., 2017; Morley and Nixon, 2016; Nyberg et al., 2018; Procter and Sanderson, 2018; Sanderson and Nixon, 2015). By comparing these areas, we give further estimates ~~on~~ about the spatial variation within the larger fracture network. In a companion paper (Passchier et al., ~~in prep~~in prep) we present a manual fracture mapping study of the whole area, ~~and propose~~discuss the criteria for identifying different fracture generations, and study the spatial heterogeneity of different fracture generations ~~criteria for identifying the different generations of fractures.~~

## 2. Study area

### 2.1 Geology

The study site is a wave-cut platform on the southern coast of the Bristol Channel in Somerset (Fig. 1a), close to the hamlet Lilstock. The outcrop exposes joints, faults, and fractures in a lower Jurassic sedimentary sequence of bituminous shale, marl, and limestone (Fig. 1b). The site is renowned among structural geologists and the subject of various publications studying faults, fractures, fracture relationships, or basin inversion (Crider and Peacock, 2004; Dart et al., 1995; Gillespie et al., 2011; Glen et al., 2005; Peacock et al., 2018; Peacock and Sanderson, 1991; Procter and Sanderson, 2018; Rawnsley et al., 1998). We focus on a single fractured limestone layer that has been previously referred to as "the bench" (Loosveld and Franssen, 1992) or "Block 5" (Engelder and Peacock, 2001).

Several tectonic phases are inferred from the structures found in the Bristol Channel basin, starting with two events of N-S extension, in the Jurassic to Early Cretaceous and Late Cretaceous to Oligocene, followed by ~~the~~ Alpine N-S contraction from the late Oligocene to Miocene and the progressive relaxation during the Late or post-Miocene (Rawnsley et al., 1998). These events are recorded as conjugate E-W striking normal faults (Brooks et al., 1988) caused by the extension and conjugate strike slip faults along with inverted normal faults, which resulted from the subsequent compression (Dart et al., 1995; Glen et al., 2005; Kelly et al., 1999; Nemčok et al., 1995). Th~~e scope of this~~ is study focuses on ~~are~~ mode I ~~fractures~~ joints in the limestone layers, which have been studied by Agheshlui et al. (2018), Azizmohammadi and Matthäi (2017), Belayneh et al. (2006a, 2006b), Belayneh and Cosgrove (2004), Engelder and Peacock (2001), Gillespie et al. (2011), Loosveld and Franssen (1992), Matthäi et al. (2007), Matthäi and Belayneh (2004), and Peacock and Sanderson (1991, 1994, 1995). According to Engelder and Peacock (2001), minor tectonic events post-date the inversion and resulted in the jointing of the limestone: the first joint set of the Bristol Channel was possibly caused by a stress field consistent with second Alpine event during the late Oligocene to Miocene (Rawnsley et al., 1998). A second phase of jointing that follows a propagation sequence consistent with an anticlockwise shift (NW-SE to NE-SW) of the regional maximum horizontal stress (Loosveld and Franssen, 1992) and~~,~~ a period of E-W compression (Hancock, 1969). ~~and u~~ Unfilled ~~jointing~~ joints are caused by basin-wide relaxation of the Alpine compression (Rawnsley et al., 1998). Engelder and Peacock (2001) further point out that the youngest joint set ~~shows a correlation~~ correlates with the contemporary tectonic stress field (Bevan and Hancock, 1986; Hancock and Engelder, 1989), and that the youngest NW-striking joint sets in NW Europe could be caused by exhumation in a late stage Alpine stress field (Hancock and Engelder, 1989).

## 2.2 Data set

We selected five 140 m² large rectangular domains (Fig. 1b) to be mapped and interpreted, both manually and automatically, from the generated ortho-mosaics that have a spatial resolution of ~ 1 cm/pixel. Three domains are located on the south-western part of "the bench", referred to as SW1, SW2, and SW3 (from W to E), and domains on the north-eastern part, referred to as NE1 and NE 2 (from NW to SE) (Fig.1b). The ~~reasoning behind the selection of the~~ reasons for selecting these specific regions are:

    i.     all domains are within a single fractured limestone layer "the bench", allowing us to compare the spatial variation of the fracture network within ~~the~~ this single layer,

    ii.    fractures in "the bench" have erosion-enhanced aperture and hence are easily detected in UAV-photographs because of the high contrast between fracture and rock,

    iii.    "the bench" offers the largest exposure of a sub-horizontal single layer,

    iv.    a qualitative review indicated that the fracture network complexity changes over short distances in the southwestern part of "the bench",

    v.    ~~fractures show a distinct radial pattern in the NE and our~~ our domains NE1 and NE2 were chosen to partly overlap with the area studied in Gillespie et al. (2011), allowing a comparison of the results~~.~~,

    vi.    in the area overlapping with the one of Gillespie et al. (2011) a qualitative review reveals a radial pattern of fractures that converge towards the fault (see Fig. 1b) S of the domain, which may impact the local fracture network geometry and topology,

vii.     areas with preferably little erosion of the surface of the pavement help to reduce the number of false positives during the automatic extraction,

~~v.~~viii.     areas with no obvious voids in the network which may be caused accumulations of remaining seaweed covering the fractures or missing parts of the pavement due to erosion.

1425

The size of the domains (140 m² each) ~~has been~~was chosen ~~in~~ to be as large as possible, to prevent over-censoring of longer fractures, while still allow manual mapping in an appropriate amount of time.


## 3.   Methods

### 3.1  UAV

1430   We used a DJI Phantom 4 UAV with a 12 MP camera to capture several sets of overlapping photographs of the pavement from an orthogonal perspective to create ortho-rectified mosaics. These mosaics were used as a raster layer for detailed mapping of the fracture network. To optimize the quality of the raw image data, the aerial photographs were acquired on days with suitable weather conditions: low wind velocities, no rain, sunny conditions with almost clear skies and warm temperature, that lead to the complete evaporation of the seawater on

1435   the surface, but not in the cavities of the fractures, further increasing the contrast between fracture and matrix. Because the near-flat topography does not cast shadows, sunny conditions in combination with the contrast due to the wet fractures are superior to the typically recommended cloudy skies that usually result in a better image quality as diffuse light casts fewer shadows with lower dynamic range. The setting ~~at~~along the coast of the Bristol Channel necessitates data acquisition at low tide, since most of the pavement is covered by the sea during high tide and the

1440   outcrop is washed clean but dries quickly on a clear, warm day. To optimize the volume of data acquisition during a single day, we undertook as many flights as possible during low tide in the morning and again in the afternoon. Apart from the cliff, the coast is flat without major obstacles, enabling us to automate the UAV mapping procedure by using Pix4Dcapture to pre-generate flight routes for the UAV. The automation provides better control on the degree of overlap between each photograph. We used settings of at least 80% frontlap and 60% sidelap with

1445   vertically downwards facing camera, resulting in 8-9 perspectives per point in the center of our models that are adequate to create a 2D map of a low topography surface. We used the SfM photogrammetry software Agisoft PhotoScan to match the photographs, calculate camera perspectives, and create dense point clouds. Using the point clouds, digital elevation models (DEMs) were calculated achieving resolutions < 2.3 cm/pixel for flight altitudes below 25 m above ground. These DEMs were used as reference surfaces for the calculation of the ortho-mosaics

1450   with resolutions ~ 1 cm/pixel respectively. Since we focus on the 2D geometry of the fracture networks, we were able to omit the placement of ground control points to increase the geo-referencing accuracy. The onboard GPS receiver of our UAV allows absolute horizontal positioning in the range of few meters and we only refer to the relative not absolute positioning. Quality control and ground truthing were done by field observation and measurements on objects of known size in our models. Examples showing that we were able reproduce

1455   measurements in the field with measurements taken in the resulting DEMs and ortho-mosaics are provided in S1.


To identify the optimum between flight altitude and resolution, we took sample photographs starting at 7 m up to 100 m (Fig. 2). We evaluated the spatial resolution in the photographs with respect to fractures that we can identify

1460 in the outcrop. An altitude ranging between 20 - 30 m offers the best balance between flight altitude ~~and~~ vs image accuracy, while~~thus~~ being able to cover larger areas in ~~shorter~~ less time ~~and~~ at a sufficient spatial resolution ~~depicting~~ to capture all fractures in the photographs (Fig.2). Photographs taken ~~from a~~ at the greate~~n~~ altitude of 100 m (Fig. 2a) resolve only the larger fractures, which is possible because of the high contrast between fracture and rock ~~(~~ (cf. Gillespie et al.~~, ,~~ 2011), but failing to resolve the smaller fractures sufficiently (Fig.2). ~~We used the

1465 SfM photogrammetry software Agisoft PhotoScan to match the photographs, calculate camera perspectives, and create dense point clouds. Using the point clouds, digital elevation models (DEMs) were calculated achieving resolutions < 2.3 cm/pixel for flight altitudes below 25 m above ground. These DEMs were used as reference surfaces for the calculation of the ortho-mosaics with resolutions ~ 1 cm/pixel respectively. To further test our survey altitude (Fig. 2) we additionally covered selected areas from~~By contrast, flying at the lower elevation of 10

1470 m produced ~~flight altitude, resulting in~~ DEMs with greater resolution of 8 mm/pix ~~resolutions~~ and ortho-mosaics with 4 mm/pix. Yet, n~~No~~ other fractures can be identified ~~in these models which have not been resolved from~~ in the data from these flights as compared to the data from flight altitudes of 25 m (e.g., Fig. 2b vs. 2e)~~,~~. ~~thus~~ Thus, we opted for a maximum flight altitude of 25 m ~~higher altitudes~~ to reduce survey time/battery and data volume. An overview of the entire set (Fig. 1b) ~~has been~~was created from photographs taken from a flight altitude of 100

1475 m~~, showing structures of larger scale, but is not detailed enough to accurately map fractures~~. The spatial resolution in the ortho-mosaic from 100 m is 4.21 cm/pix, which is slightly better than the 5 cm/pixel achieved by Gillespie et al. (2011), (their figures 8 and 9), who used a large-scale camera for an aerial photographic survey by airplane. ~~Since we focus on the 2D geometry of the fracture networks, we were able to omit the placement of ground control points to increase the geo-referencing accuracy. The onboard GPS receiver of our UAV allows absolute horizontal

1480 positioning in the range of few meters and we only refer to the relative not absolute positioning. Quality control and ground truthing were done by field observation and measurements on objects of known size in our models.~~

### 3.2 Fracture trace mapping

Manually tracing all fractures of the outcrop is a very time intensive task because of the large amount of fractures present in the dataset. ~~It is not possible to manually map all fractures on the whole outcrop within reasonable time;

1485 . thus~~Thus, an automatic fracture mapping technique ~~seems obvious~~for more rapid data characterization is needed. However, the reliability along with the strengths and weaknesses of automatic fracture detection must be addressed first. To do so, we use both manual and automatic techniques for the same domains, allowing a direct comparison of ~~the~~ results. This comparison helps us to identify potential human bias (e.g. Andrews et al., 2019; Peacock et al., 2019) and gives us a measure of quality of the automatic fracture detection technique, which then can be optimized

1490 to map the whole outcrop in future work.

In the manual approach, fractures are traced as polylines from tip to tip, utilizing the digital imagery provided by the UAV survey as base map in ArcGIS. We traced the widely eroded fractures along their medial axis creating as few vertices as possible, while still maintaining the original fracture geometry. Abutting fractures were mapped as polylines and snapped at the abutting intersections. Snapping is not necessary for cross-cutting fractures, but it is

1495 for intersections where a younger fracture converges to an existing junction. Tracing one short segment between two junctions of fractures involves identifying its' start and end points to providing the input in the software to create the trace. Based on our experience, this process takes at least five seconds when the case is clear to the interpreter. Extrapolated to an amount of ~ 7000 segment traces, which roughly represents the amount of segment traces required to map one domain~~s~~ in the SW, the mapping time sums up to ca. 10 hours. Yet, the estimated time

1500 does not include quality and integrity checks or corrections and reinterpretations, which may easily add another five to six hours to the required minimum time of 10 hours to trace the fractures of one domain manually.

An approach to semi-automatically map this outcrop has been published by Gillespie et al. (2011), who used ant tracking on a lower resolution dataset, resulting in a good overview of the fracture geometry, but incomplete

1505 fracture detection (their Fig. 9), which is insufficient for a detailed analysis of crosscutting relationships. To overcome these limitations, we use the technique presented inof Prabhakaran et al. (2019), that allows detailed detection of fractures in our data. The automated trace extraction is based on ridge detection using the complex shearlet transform method as described by Reisenhofer et al. (2016) and extended to fracture trace extraction from drone photogrammetry by Prabhakaran et al. (2019). The automatic processing consists of serial image processing

1510 steps applied to the UAV-images, including ridge ensemble computation, segmentation, skeletonization, and polyline fitting (Fig. 3). The automatic trace extraction requires a selection of parameters pertaining to the shearlet systems. For the Lilstock dataset, we chose a set of shearlet systems that capture fractures of multiple scales and applied it to all image tiles. Subsequently, for each image tile, a single set of shearlets was used and the intensity threshold (Fig. 3) was adjusted to convert ridge ensemble maps into binarized images by visual comparison with

1515 the source image as explained in detail in Prabhakaran et al. (2019) . Such a step was necessary owing to the spatially varying fracturing fracture intensity.

To preprocess the ortho-mosaics for the automatic trace detection we sliced them into tiles of 1000 x 1000 pixels resulting in six to seven tiles per map domain. Depending on the available hardware the processing time of the feature detection may vary and took us ca. 20 minutes per tile. However, this time does not account for the

1520 preparation of shearlet systems, trial and error taken in deciding the intensity cutoff for each image, removal of potentially produced false positives, and manually joining the detected features at the edges of the tiles. Taking these necessary steps into account, the automatic trace extraction of one tile may require one up to two hours in total and between ca. 6 and 14 hours for a map domain.

1525 **3.3 Network analysis**

To analyze the fracture networks, we chose a sampling strategy comparable to the polygon sampling explained inof Nyberg et al. (2018) by completely analyzing the mapped domainswhich allows two-dimensional areal sampling of fracture networks by in a manually defined area or subareas therein. This approach allows us to analyze the domains entirely as areas and to identify possible spatial heterogeneities between and within the map

1530 domains.

We used ArcGIS as main platform in combination with the NetworkGT plugin (Nyberg et al., 2018). The ArcGIS environment enables us to measure fracture length , calculate fracture and segment sinuosity, and measure strike directions based on the traced segments of the fractures that are depicted in the georeferenced and orthorectified mosaics.

1535 These 2D fracture networks may be described in terms of nodes vertices and edges that constitute a spatial graph (e.g. Sanderson et al., 2019.), in which the set of vertices comprises the intersections of e.g. abutting or crosscutting fractures as well as the terminating points of isolated fracture tips. and tThe sinuosity of the fracture is approximated using piecewise linear edges. Because the intersection points between these edges are also

considered as vertices, we will follow the usage of e.g. Nyberg et al. (2018) and refer to intersections of fractures as nodes.

The nodes are classified in terms of node degree, i.e., the number of edges that pass through it, and the edges are classified based on the type of branches. Sanderson and Nixon (2015) specify three node types, isolated (I), abutting (Y), and crosscutting (X) whose relative proportions define the topology of the graph. In our interpretations, the nodes of fractures clipped censored by the boundary of the selected sample area are classified as end nodes (E). The fracture segments connecting the nodes and referred to as branches, that may be classified into three types. The branch types are I – I when both nodes at the branch tips are I-nodes, C-I when connected at one tip (via an abutting or crosscutting node) but isolated at the other, and C-C when both tips end at a Y or X node.

The manually-traced fractures were split into segments between intersections To to apply these topological analysis tools provided by thefrom the NetworkGT plugin (Nyberg et al., 2018), we had to split the manually traced fractures into segments between intersections. This step is necessary for the manually-traced fractures, because they have been traced along a path from fracture tip to fracture tip and do not yet include information of intersections with other fractures along that path., and Furthermore, the traces cannot be processed using NetworkGT in that form in their initial form, because the toolbox requires singlepart features instead of multipart features (i.e. every segment with an own ID instead of several connected segments with a shared ID to represent the complete fracture trace). not necessary fFor the automatically generated traces that are already segmented at intersections as output from the technique of Prabhakaran et al. (2019), this processing step is not necessary. NetworkGT only supports nodes of the types I, Y, and X ., and nodes nNodes of higher order, i.e. more than four branches intersecting at one point are returned as error nodes due to the missing implementation of such cases. To assign the correct order of the node to the returned error nodes, the spatial join function from the ArcGIS Analysis Toolbox was used to count the number of intersecting branches at the specific nodes, allowing us to reassign their type accordingly These error nodes consisting of node degree greater than four, are corrected and distinguished as special cases of X nodes, where five (Penta), six (Hexa), seven (Hepta) or eight (Octa) branches intersect. The spatial join function was used to count the number of intersecting branches at the specific nodes and their classes were reassigned accordingly. This is possible, because the spatial join function allows matching the branches to the target error node based on their relative spatial locations. The tips of the branches are at the same spatial position as the node, which resembles an intersection, and the spatial join function returns a new field in the error nodes attribute table including the number of intersecting branches at that exact position. Furthermore, linkage Linkage to an unknown type of node yields errors in the recognition of branch connectivity. These errors were corrected using SQL queries to select undefined branches to change their connectivity accordingly to their number of connected ends.

### 3.4 Manual classification of fractures into generations

To distinguish identify different generations of the manually manually-traced fractures, we used abutting and cross cross-cutting relationships as discussed in Peacock et al. (2018), aided by general fracture attributes such as length and strike. The interpretation of fracture generations on this dataset along with detailed rules for the classification of generations is presented on a larger scale by Passchier et al. (XXXX) and therefore kept briefly within this section, focusing on the application in our domains. Fracture generations were assigned following the criteria below:

31

Gen. 1:   Longest fractures that traverse through the study areas. Absolute orientations are not used as criterion because they ~~are not reliable~~ may change in converging patterns as present in the NE of "the bench".

Gen. 2:   Subparallel orientation to gen. 1, shorter in length and abuts on to gen. 1.

Gen. 3:   Shorter than gen. 1 and 2, oriented at ~~a sharp~~an acute angle to gen. 1 and 2 and abuts on to them.

Gen. 4:   ENE-WSW striking, sub-orthogonal to gen. 1 and 2 and abutting gen. 1 and 2, ~~also~~ and both abuts and crosscuts gen. 3

Gen. 5:   Shortest ~~with diffuse~~fracture traces with a large range of orientation~~s~~ and abuts all other generations sub-orthogonally, creating polygonal shapes

Distinguishing generations 1 and 2 by length is not possible for all fractures because large ~~parts extend~~portions of individual traces continue outside of our ~~selected~~ domains. Thus, we qualitatively ~~decided to adopt the interpretation of Passchier et al. (XXXX) for fractures of generations 1 and 2 in the southwestern domains~~investigated the continuation of the fracture trace towards the outside of the map domain to aid the interpretation. A fracture set analogous to generation (gen.) 3 in the SW domains was absent~~could not be identified~~ in the NE domains~~in the NE~~. Our interpretation of fracture generations is ~~in line~~consistent with the four main joint sets ~~introduced~~ identified in Rawnsley et al. (1998) for the northeastern areas. ~~Further interpretations of the different sets of joints include~~Other interpretations of joint populations in other locations at Lilstock (Belayneh and Cosgrove, 2004; Engelder and Peacock, 2001; Loosveld and Franssen, 1992)~~, who interpreted~~identified up to six generations of ~~joints in the Lilstock pavement~~ joints.

## 4.   Results

### 4.1 Fracture ~~traces~~trace segments

We traced fractures in five different domains of 140 m² each (Fig. 1b) manually and automatically. To make both datasets comparable, the manual traces were split into segments between nodes ~~as already the case for~~to resemble the automatically extracted trace segments. This step is necessary because the automated extraction results in traces segmented at the fracture intersection nodes. The number of overall trace segments in automatically traced networks is, therefore, greater than the number of fracture traces as would be described by an interpreter mapping complete fractures that consist of several segments along a path from fracture tip to tip, which makes a direct comparison of both methods difficult.

The number of automatically-extracted segments is greater than the number of the manually-traced ones in all domains (Table 1), whereas the difference of traced segments within the domains is smaller in the NE domains and larger within the SW domains. The cumulative length distribution of segments and associated plots of log-normal standard deviations for automatically-extracted traces (Fig. 4) and manually-traced segments (Fig. 5) for each domain are similar to each other when the methods are compared qualitatively, despite the aforementioned difference in the number of segments. For both methods, the cumulative length distribution and log-normal standard deviation plots in fig 4 and fig. 5 show that fracture traces of all domains resemble the characteristic negative power law associated with fracture traces. A qualitative comparison of the histogram distributions within both, the automatically-extracted segments (Fig. 4) and the manually-traced segments (Fig. 5), shows a greater

similarity within the groups of spatially close domains in the SW or the NE than a comparison between SW vs. NE. While the overall histogram distributions of automatically-extracted segments (fig. 4) and manually-traced segments (fig. 5) appear similar and the plots have peaks at comparable segment lengths and mean segment lengths (Table 1) when compared to each other, the distributions show that the automatic extraction results in a relatively higher amount of smaller segments.

The calculated mean lengths of the automatically-traced segments are similar to each other in the SW and range from 16 to 17 cm in the three domains, while the difference from 15 to 21 cm is greater in the NE domains. The variation of the mean lengths between the SW and NE domains for the manually-traced segments is similar to the automatically-extracted ones and the mean lengths of manually-traced segments are consistently greater by a difference of 1 to 4 cm (Table 1). Both methods show that the longest segments are in NE2 as compared to the other domains, which is consistent over the greater values for the NE2 25%, 50% and 75% quartiles (Table 1).

The cutoff length for image sampling for all windows is 1 cm, therefore values smaller than 1 cm are not significant. The maximum lengths of the segments may be censored by the sampling windows. Apparent maximum lengths of the segments are largest in SW2 and shortest in NE1 which is consistent in both tracing methods, even though manual maximum segment lengths may be up to 16 cm larger than their automatically-traced counterparts (Table 1). The covariance is positive and similar in all domains and both methods, ranging from the lowest in NE1 to the highest in SW1. The kurtosis in all domains and for both methods is positive, largest in SW1 and overall larger in the SW domains than in the NE domains, which suggests that length-frequency-distributions are closest to that of a half-normal distribution (kurtosis = 3) in SW1 with a decreasing trend towards the NE, what suggests that segment lengths towards the NE have fewer outliers in the form of longer segments. This characteristic of the length-frequency distribution is shown further by the positive skewness in all domains and both methods with a decreasing trend towards the E as well, which further indicates that the branch distribution in the NE domains is more symmetric, while the distributions in the W are more asymmetric with a tail towards the right. This characteristic of the distribution suggests that the SW domains tend to have more segments that are longer than most of the segments traced or extracted in the respective domain as compared to the domains in the NE.

To compare the resulting networks of both methods and all domains spatially we calculate fracture trace segment density and fracture trace segment intensity. Therefore, we use the $P_{ij}$ system (Dershowitz and Herda, 1992) as denotation. In the $P_{ij}$ system, "i" gives the dimension of the sample which is "2" for two-dimensional maps of fracture traces. The second index "j" gives the dimension of the measurement, which is "0" for fracture density that quantifies the number of fractures per unit dimension, or "1" for fracture intensity that measures the total fracture persistence per unit dimension (Dershowitz and Herda, 1992). In the case of our results for both, $P_{20}$ and $P_{21}$, it must be noted that these do not represent actual fracture density and fracture intensity in the domains, but density and intensity of the fracture segments.

### 4.1.1    Fracture trace segment densities

To provide examples and a direct comparison of the manually- and automatically-extracted traces we selected the $P_{20}$ plots of the domains SW3 (Fig. 6 and NE2 (Fig. 7) that show the density of fracture segments windows of 0.5 x 0.5 m within the domains and the network of traces in the background. Therefore, these figures allow us to make

a direct comparison of the number of segments traced in a certain part of the domain between the two methods. Plots of $P_{20}$ and $P_{20}$ absolute and relative differences for all domains are provided in the supplement S2 and S3.

1660 In domain SW3 (Fig. 6) qualitative comparisons of the manual and automatic traces show an overall similar spatial distribution of areas with relatively higher and lower $P_{20}$ values, except for an area between 4 and 6 m in the x-direction and 3 – 4 m in the y-direction . Plots of the $P_{20}$ absolute and relative difference at the bottom (Fig. 6) show that the automatic trace extraction resulted in a higher $P_{20}$ value in this area which can be verified visually by the greater amount of automatically-extracted short segments of the underlying fracture trace network which

1665 are not represented in the manually-traced segments. Overall, $P_{20}$ is higher (58.9 m$^{-2}$) for the automatically-extracted segments than for the manually-traced segments (49.8 m$^{-2}$) (Table 1).

A similar comparison made for domain NE2 (Fig. 7) shows fewer differences in the segment density for both methods when compared to SW3 (Fig. 6). Again, the plots of the $P_{20}$ absolute and relative difference at the bottom of fig.7 show that the automatic method extracted small segments that are not represented in the manually-traced

1670 network, mostly at areas at 4 and 8 m along the x-axis and 0.5 m along the y-axis and 16 m along the x-axis and 5 m along the y-axis. The greater similarity of $P_{20}$ for both methods in this domain is also evidenced by the overall $P_{20}$ values of 30.9 m$^{-2}$ for the manually-traced network and 34.5 m$^{-2}$ for the automatically-extracted network (Table 1).

### 4.1.2 Fracture trace segment intensities

1675 The $P_{20}$ plots presented in the preceding section allow us to compare the spatial distribution of the number of segments per unit area and further enable us to isolate areas where the resulting number of trace segments differs between methods. However, the resulting networks derived from the two methods require another comparison than just $P_{20}$, because arguably a single fracture trace can be mapped as a path consisting of a different numbers of traces by different methods, while the overall path geometry of the fracture from tip to tip along the segments may

1680 be the same. Therefore, $P_{21}$ allows us to compare the lengths of segments per unit area for both methods, where the similarity between methods suggests that the same fracture traces were recognized and extracted or traced with similar lengths regardless of the number of segments that represent the fracture trace.

The trace maps are depicted as an overlay on a $P_{21}$ fracture intensity plot

1685 ~~and present the resulting trace maps as $P_{21}$ fracture intensity plots~~ (Dershowitz and Herda, 1992) for both methods and all domains in Fig. 4~~8 where $P_{21}$ is a measure of fracture abundance defined as the fracture trace length in meters per square meters~~. To provide a good comparability to the $P_{20}$ plots in figures 6 and 7, the same cell size of 0.5 m * 0.5 m was chosen. A qualitative comparison between the resulting $P_{21}$ plots of automatic segment trace extraction and manually-traced segments shows a high similarity in all domains because the distribution of cells

1690 with relatively higher or lower segment intensity within the respective domains is the same for both methods. However, smaller areas within the domains can be identified in which $P_{21}$ is higher for the automatically-extracted traces (Fig. 8). An example of higher $P_{21}$ can be observed in SW3 between 4 and 6 m in x-direction and 3 – 4 m in y-direction (Fig. 8), where $P_{20}$ was also higher and overall more smaller fracture traces were extracted automatically than traced manually (Fig. 6.). To better isolate and quantify areas in which the two methods give

1695 different $P_{21}$ results, t~~T~~he difference between the manual and automatic interpretations is depicted as a spatial map of as $P_{21}$ fracture intensity difference in ~~the supplement S1~~fig. 9. A qualitative comparison of the two methods between the domains again shows, that the differences are minor and not larger than ca. 4 m$^{-1}$ per cell. Examples

of areas with highest differences are in SW3 between 6 and 8 m along the x-axis and 7 m in y-direction (fig. 9), which are again caused by an overall larger amount of traces extracted automatically than manually (see also fig. 8). In NE2 several neighboring cells at 14 m in x-direction and between 0 and 2 m along the y-axis show the opposite case (fig. 9), where the manual interpreter traced two parallel fractures which were extracted as one automatically.

Overall, resulting $P_{21}$ values for all domains are very similar for both methods with a minimum difference of 0.01 $m^{-1}$ between the methods in SW2 and NE2 and a maximum difference of 0.25 $m^{-1}$ m in NE1 (Table 1). Unlike the results of $P_{20}$, which are greater for the automatically-extracted trace segments than the for the manually-traced segments for all domains, the resulting $P_{21}$ values might be greater or smaller for either method when compared to the other, varying in between the domains (Table 1). ~~Despite a slightly larger number of fracture segments in the automatically generated code compared to the manual interpretation, the resulting plots are very similar.~~ ~~The automated extraction results in traces segmented at the fracture intersection nodes. The number of~~ overall ~~traces~~ trace segments ~~is therefore higher than the number of fracture traces as would be described by an interpreter mapping complete fractures~~ that consist of several segments. ~~However, these segmented traces already resemble the branches as used for the network analysis and have been processed to represent branches and nodes (Figs. 5 and 6). Further a~~Analyses ~~of the automatically generated networks show that th~~e $P_{21}$ fracturing intensity for both methods shows, that ~~$P_{21}$~~ ~~in higher~~ is overall greater in the SW ~~study area~~domains compared to the NE domains, and ~~within the SW study areas is highest~~ greatest in SW2 and smallest in NE2 (Table 1).

~~A quantitative comparison of the interpreted fractures from the five domains shows that the number and mean lengths within the south-western domains and NE1 is similar, but clearly differ from NE2 (Table 1, Fig. 7): in the south-western domains and NE1 between 8184 and 8548 branches were counted with mean lengths between 0.15 m and 0.17 m, while NE2 only has 4831 branches which are therefore longer with a mean length of 0.21 m. The trend of NE2 having longer branches than the other domains is consistent over the 25%, 50% and 75% quartiles. Minimum lengths in all domains are < 1 cm, thus cannot be accurately resolved in our orthorectified mosaics.~~ and ~~maximum lengths of the branches may be censored by the sampling windows~~This trend is not present in the maximum lengths of the branches, which are largest in SW2 (0.97) and shortest in NE1 (0.57).~~ The covariance is relatively similar in all domains, ranging from the lowest (0.56) in NE1 to the highest (0.7) in SW1. The kurtosis in all domains is positive and are closest to that of a half normal distribution (3) in SW1 with a decreasing trend towards the E. All domains show a positive skewness with a decreasing trend towards the E as well, indicating that the branch distribution in the eastern domains is closer to a symmetric one, while the distributions in the W are more asymmetric with a tail towards longer branches. The cumulative length distribution plots and the plots of log-normal standard deviation (Fig. 7) show that our dataset resembles the characteristic negative power law associated with fracture traces. The resulting traces generated by the automatic technique are segmented, thus only the final cumulative network development can be analyzed. This is the case because the correct identification of abutting and crosscutting relationships is a prerequisite to identify age relationships and requires one to review the complete fractures, not just segments. The resulting networks depicted as branches and nodes are presented in Fig. 5 for the domains in the SW and Fig. 6 for the domains in the NE, also showing that all visible fractures have been identified successfully. The node topology statistics indicate that the network is well connected with a relatively small proportion of isolated nodes with the system dominated by abutting (Y) nodes (Table 2 and compare with Fig.18). Isolated nodes are few in general with decreasing numbers from W to E. The number of Y nodes lies~~

between 3919 in SW3 and 4390 in SW2 in the SW, and has both, maximum and minimum values in the northeastern domains, where NE1 has the largest number (5100) and NE2 the lowest (2517). The number of X nodes along with Penta nodes increase from SW1 towards SW3 and NE1 to NE2, with NE1 having the lowest number of X nodes of all five domains. Length weighted rose plots are depicted in Fig. 8. The rose plots depict a significant variation in the fracture strike between NE and SW regions. In the SW study areas, the fracture pattern is similar in all three areas. In the NE2 study area proximal to the fault, there are two predominant fracture sets. In NE1 the same two sets persist, but the fractures acquire a polygonal pattern.

## 4.2  Network characteristics and spatial variation of the automatically-traced fracture segments

The traces resulting from the automatic technique are segmented, thus only the final cumulative network can be analyzed, because the correct identification of abutting and crosscutting relationships is a prerequisite for identifying age relationships, requiring a review of the complete fractures, not just segments. The networks resulting from the automated data collection depicted as branches and nodes did identify all visible fracture traces successfully (Fig. 10 and 11). The node topology statistics indicate that the network is well connected because it consists of only few isolated nodes and an abundance of abutting (Y) nodes (Table 2 and compare with Fig. 12 upper). Isolated nodes decrease in number from W to E. The number of Y nodes lies between 3919 in SW3 and 4390 in SW2 in the SW domains and has maximum and minimum values in the northeastern domains, where NE1 has the largest number (5100) and NE2 the smallest (2517). The number of X nodes along with Penta nodes increase from SW1 towards SW3 and NE1 to NE2, with NE1 having the smallest number of X nodes of all five domains. Length weighted rose plots (Fig. 13) show a significant variation in correlation of fracture length to orientation between NE and SW regions. In the SW study areas, the fracture pattern is similar in all three areas (Fig. 10). In the NE2 study area proximal to the single NE-SW-trending fault that transects the entire exposure, there are two predominant fracture sets (Fig.13). In NE1, the same two sets persist (Fig. 13), but the overall fracture pattern is polygonal in geometry (Fig. 11).

## 4.2 4.3    Manually interpreted Fracture fracture generations and their spatial variation

Within our manual interpretation we classify the fractures into generations Fracture generations for the manually collected data were defined by using the rules presented in Section section 3.4. The interpreted fracture generations in the SW (Fig. 9) slightly differ from the fracture generations in the NE (Fig. 10). . Figures S4 to S8 in the supplemental data enable the reader to compare these manually derived trace maps to the digital base maps. presenting the raw data next to the interpretations are provided in S2-S6. The generations in the southwestern and northeastern domains are described below.

### 4.2.1 4.3.1   Southwestern domains

The SW domains (Fig. 9 14) include five fracture generations that differ in their distribution through out the domains. In SW1 and SW2, gen. 2 fractures are more numerous than gen. 1 fractures. In SW3, this relation is reversed. Overall, the generations 1 and 2 have similar geometry. Generation 3 is represented in all three domains with increasing presence abundance from W (SW1) to E (SW3), coinciding with an increasing fracture intensity. Generation 4 only covers represents a small number of fractures in each domain, while gen. 5 is present in large numbers, abutting on all older generations and converging into existing junctions. Average values of fracture

36

length~~, and~~ strike~~, and sinuosity~~ for the domains in the SW are presented in Table 3. To ~~identify~~ visualize the relationship between ~~potential clustering of the~~ fracture strike and length, their values ~~have been~~were plotted for each fracture and color-coded according to the associated fracture generation (Fig. 15). ~~Fractures in SW1 show three clusters (Fig. 11):~~ Fractures of gen. 1 and 2 ~~accumulate~~ trend between 90° and 120°, reaching a maximum (censored) length of 18 m (Fig. 15). Generation 3 resembles a Gauss distribution around 80° with most fractures reaching lengths up to 2 m, but also including lengths up to a maximum of 10 m. Generation 4 plots around a mode of ~~0~~60° with lengths up to 3 m. Fractures associated with gen. 5 ~~are more widely distributed, covering all strike directions~~cover a wide range of strike directions. ~~with denser clusters around 0°-20°, 160°-180°, and 70°-90°, representing dominant N-S and E-W orientations.~~ Fractures in SW2 resemble the distribution of SW1 with no obvious differences. Fracture length-related distributions in SW3 (Fig. ~~11~~15) show deviations from SW1 and SW2: the cluster with the longest fractures between 100° and 120° is dominated by gen. 1, while the respective clusters in SW1 and SW2 are interpreted as mostly gen. 2, coinciding with ~~the~~ prior qualitative observations ~~before~~.

### ~~4.2.2~~4.3.2    Northeastern domains

The network in NE2 (Fig. ~~10~~16) mainly consists of many gen.1 fractures that converge beyond the southeastern edge of the sample window towards a fault in the SE, while NE1 ~~is more balanced in respect to the presence of~~has a less unimodal distribution of length-weighted fracture orientation for the fracture generations. Notably, the northeastern fracture networks ~~in the northeast~~ lack ~~the~~gen. 3 fractures ~~associated with gen. 3 in the SW and differ strongly in their perceived appearance.~~as compared to the SW sample windows. The ~~relationship of~~ gen. 2 and 4 fractures ~~is complex, as they~~crosscut ~~but also~~and abut on each other, so they do not have a clear relative age relationship. Like the domains in the SW, gen. 5 is present in both NE domains and abuts all older generations without a preferred ~~strike~~orientation. Averages of fracture length~~, and strike~~strike ~~and sinuosity~~ for NE1 and NE2 ~~are presented in Table 4. With~~show some differences to the domains in the SW~~, fracture generations the NE are clustered as well (Fig. 12~~Fig. 17 and Table 4). Generation 1 is mostly ~~present~~occurs between 120° and 160° in both, NE1 and NE2, reaching almost 9 m maximum length. Generation 2 shows a wider cluster in NE1 with most fractures striking between 80° and 140° and a similar distribution as gen. 1 in NE2. The cluster associated with gen.3 in the SW (Fig. ~~11~~15) is not present in NE1 or NE2 (Fig. 17). Another cluster at 60° is associated with gen. 4 in NE1 and NE2, which is the same in all three areas in the SW. Generation 5 is widely spread in all directions, especially in NE1, while a maximum of gen. 5 fractures between 20° and 60° is present in NE2.

### ~~4.3~~4.4    Fracture network evolution

To analyze the impact of the fracture generations on the evolution of the overall network, ~~their~~fracture traces were split at intersections into branches and the intersections represented as nodes, to resemble the fracture networks ~~presented~~generated ~~for~~from the automatic trace detection ~~in figures 5 and 6~~(Fig. 10 and 11). The results are presented for SW1 - SW3 in figures ~~13 - 15~~18 - 20 and NE1 and NE2 in figures ~~16 and 17~~21 and 22, showing the network evolution in time steps by adding the ~~subsequent~~next generation ~~in each~~to the subsequent subfigure. The following sub-sections summarize the major changes for every domain during the network evolution to guide the reader through figures ~~13 - 17~~18 - 22 and tables 5 - 9.

### ~~4.3.1~~4.4.1    SW1

The initial time step comprises three sub-parallel branches of gen. 1 fractures (Fig. ~~13a~~18a and Table 5). Their full length~~s~~ ~~falls~~fall outside the mapped area and they are censored by the map~~abut at the~~ boundary resulting in three isolated (I – I) branches of 3 m length ~~in~~on average. With gen. 2 (Fig. ~~13b~~18b) more sub-parallel branches are added, of which a large number abut at the ~~mapping boundary~~map boundary, resulting in further isolated branches and an increased average length of 4.69 m. Some gen. 2 fractures abut at gen. 1 in acute angles, resulting in Y nodes at connected (C – C) branches. Generation 3 strikes ~~in~~at an angle towards gen. 1 and 2, either crosscutting the older fractures (X nodes) or abutting them (Y nodes) ~~as presented in~~ (Fig. ~~13c~~18c), leading to a shorter average branch length of 0.59 m. Few fractures interpreted as gen. 3 end in I nodes. Generation 4 is sparsely represented in SW1, ~~thus~~so the impact on the fracture network as compared to the previous ~~time step~~generations is minor (Fig. ~~.~~ ~~13d~~18d). With gen. 5 (Fig. ~~13e~~18e) the network becomes more spatially dense, further reducing the average branch length to 0.2 m. The fractures may abut on older generations, but also tend to join in existing junctions, leading to the development of Penta and Hexa nodes. ~~The sinuosity slightly increases from gen.1 and 2 towards gen.3. From gen.4 to gen.5, it slightly reduces again.~~

### ~~4.3.2~~4.4.2    SW2

In SW2, generation 1 is represented ~~as~~by four fractures, two of which ~~two~~ abut on the mapping boundaries at each side and two end in I nodes (Fig. ~~14a~~19a and Table 6). Most gen. 2 fractures end in either I nodes or at the ~~mapping boundary~~map boundary but can also connect to the tips of preexisting gen.1 fractures as observed on the largest gen.1 fracture in the center (Fig. ~~14b~~19b). Generation 2 leads to a decrease in average ~~(avg.)~~ branch length from 7.98 m to 4.51 m. Generation 3 has a large impact on the network, further reducing the ~~avg.~~average branch length to 0.54 m due to many fractures crosscutting older ~~generations~~fractures (Fig. ~~14c~~19c), resulting in a sharp increase of Y and X nodes. Generation 4 has a minor impact (Fig. ~~14d~~19d), not significantly altering the network parameters. Generation 5 (Fig. ~~14e~~19e) again reduces the ~~avg.~~average branch length to 0.2 m, while reducing the number of I nodes and strongly increasing the number of Y and X nodes. ~~The sinuosity has a consistent value of 1 during all stages of the network within this domain.~~

### ~~4.3.3~~4.4.3    SW3

Compared to SW1 and SW2, gen. 1 fractures are more numerous in SW3, however, they are still not well connected with most fractures ending in I or E nodes at the ~~mapping boundary~~map boundary (Fig. ~~15a~~20a and Table 7). Generation 2 is less pronounced in this area with only a small impact on the network (Fig. ~~15b~~20b). The average branch length slightly increases by 0.11 m, as few long fractures are added. With the addition of gen. 3 to the network (Fig. ~~15c~~20c), the average branch length is reduced from 3.17 m to 0.59 m, again with a strong increase of Y and X nodes. With the addition of gen. 4 (Fig. ~~15d~~20d), I nodes are reduced in favor of Y and X nodes. Generation 5 (Fig. ~~15e~~20e) further reduces the ~~avg.~~average length to 0.2 m with a sharp increase of Y and X nodes. ~~The sinuosity increases slightly from gen. 1 and 2 towards gen. 3 and slightly decreases again towards gen. 4.~~

#### 4.3.4 4.4.4    NE1

Generation 1 consists of several sub-parallel branches that mostly end in I nodes or at the ~~mapping boundary~~map boundary (Fig. ~~16a~~21a and Table 8). Abutting fractures are sparse, as only four Y and zero X nodes are observed, leading to an average branch length of 4.33 m. Generation 2 (Fig. ~~16b~~21b) strikes at an angle to gen. 1 leading to an increase of cross cutting and abutting nodes, that are present in almost equal numbers. Also, the number of I nodes is quadrupled. The ~~avg.~~average branch length is reduced to 0.79 m. With the addition of gen. 4 (Fig. ~~16c~~21c), the ~~avg.~~average branch length is further reduced to 0.68 m, accompanied by an increase of abutting (Y) and cross cutting (X) nodes. Generation 5 (Fig. ~~16d~~21d) further reduces the ~~avg.~~average branch length to 0.24 m with a strong increase of Y nodes from 547 to 4135 and X nodes from 229 to 705. ~~The sinuosity increases from 1 to 1.01 during the last step of the network development.~~

#### 4.3.5 4.4.5    NE2

In this subarea, gen. 1 is present as long subparallel branches that spread out radially from a fault outside of the southeastern map~~ping~~ boundary (Fig. ~~17a~~22a and Table 9). The branches barely abut or intersect each other or create I nodes. Most of the branches run into the mapping boundaries at opposite sides, leading to an ~~avg.~~average branch length of 4.14 m. Generation 2 crosscuts gen. 1 almost orthogonally (Fig. ~~17b~~22b) leading to an increase of Y and X nodes along with a reduced ~~avg.~~average branch length of 0.8 m. Generation 4 (Fig. ~~17c~~22c) further reduces the ~~avg.~~average branch length to 0.45 m, with a slight increase of X and Y nodes. ~~At this stage, the sinuosity of the network is maximal.~~ With the addition of gen. 5 (Fig. ~~17d~~22d), the ~~avg.~~average branch length is reduced to 0.16 m along with the addition of a large ~~additional~~ number of Y and X nodes.

A direct comparison of the node distribution for all domains is ~~visualized~~ shown in ternary plots highlighting the changing node distribution during the network development (Fig. ~~18~~12 upper). Results of the latest network state (analog to gen. 1 – 5 for ~~manually~~ manually-traced fractures) from the automatic trace extraction are depicted in Fig. ~~18~~12 lower for a direct comparison.

## 5.    Discussion

### 5.1 Manual vs automatic tracing

~~The resulting branch lengths and node counts of the automatic and manual procedures are listed in Table 10. Three detailed examples comparing the manual and automatic interpretations processed as branches and nodes are given in Figure 19.~~

Manual tracing of the fracture network is comprehensive but time-consuming. At this quality of outcrop and resolution of imaging, human bias is minor (e.g. Andrews et al., 2019; Peacock et al., 2019), and ~~very~~ similar results ~~are~~ can be produced by different interpreters (e.g. Long et al., 2018). A manual interpreter's work can ~~get~~ gain quality control from considering the work of another interpreter, while the reliability of the automatically mapped network entirely depends on ridge detection and image post-processing parameters.

Time is an important variable for both methods and was estimated in section 3.2 to compare manual and automatic trace extraction. For a network of traces in one domain a minimum time of ca. 15 hours is required for the manual tracing when quality checks are incorporated, which is comparable to the maximum time required for the complete

automatic trace extraction. During these estimated periods of time, the interpreter's attention is required all the time in the manual method, whereas in the automatic method the interpreter only has to check the input parameters and the results. -Thus,

Aautomatic mapping may save personal time of the interpreter, it is overall fast compared to manual mapping;, and additional time can be saved byit is possible to parallelizeation of the process and save additional time. hHowever, the network needs to be checked for artifacts along with a general estimation of the reliability or capability of the method before it can be applied widely. Depending on theThe chosen image processing parameters chosen, this can be a huge advantage when greatly influence the automatic extraction has minimal artifacts. If the automatic extraction returns many false positives and incorrect node connections, correcting them can be more time-consuming than an initially correct manual interpretation.

Based on an extrapolation of the time required to manually map the five domains, a complete interpretation of the whole dataset incorporating the so far unmapped areas of the bench and the other layers would take a manual interpreter several hundreds of hours of pure mapping time, and weeks to months when work hours and weekends are accounted for. An estimation of the required time for the automatic tracing to complete the whole outcrop based on the time required for the presented domains is not trivial because and several aspects need to be considered. We selected areas of good outcrop quality and high fracture visibility for our map domains. Other regions covered by the ortho-mosaics include areas where the water in the fracture cavities has already dried out, what locally reduces the contrast. To achieve results of good quality in those areas, more shearlet combinations and different thresholds and parameter settings are required. In highly eroded areas, more time is required to remove false positives. Therefore, the chosen image processing parameters greatly influence the automatic extraction and speed up the overall process when the extracted network has minimal artifacts. If the automatic extraction returns many false positives and incorrect node connections, correcting them can be more time-consuming than an initially correct manual interpretation.

In either case, an advantage of the automatic tracing is the ability to reproduce results solely by choosing the same parameters, while manual reproduction requires the interpreter to follow a clearly defined set of rules which that can become excessive depending on the complexity of the dataset. Examples are widely eroded fractures that can be traced along their edges or the median of the cavity, which may also lead to different interpretations, especially at widely eroded junctions. These junctions can either be interpreted as two younger fractures closely abutting on an older one, or only one younger fracture that crosscuts the older one. This is of particular importance for unexperienced interpreters who have to decide whether to learn how to use the software and set up the rules as explained above, or how to apply the automatic trace extraction code written for MATLAB. While the ability to learn either method strongly depends on the individual, the advantage for such users is that the automatic traces provide unbiased segments which can be used to guide the interpretation and maintain consistency and quality.

However, using the fracture mapping code described inof -Prabhakaran et al. (2019), it is only possible to generate quality segmented networks of branches and nodes for well exposed patterns. The interpretation of fractures longer than a segment between two nodes and the association to a certain generation depending on crosscutting relationships still must be done manually.

For the data presented in this work, the results of both techniques are very similar., and to better highlight the dissimilarities of the traced fractures, the differences of the $P_{20}$ and $P_{21}$ analyses are presented in as shown earlier

1930 ~~in~~ Fig.~~s~~ 46, 7, 9 and supplements S2 and S3). ~~To better highlight the dissimilarities of the traced fractures, the differences of the P_{21} analyses are presented in supplement S1.~~ $P_{21}$ plots are expected to ~~be~~ have ~~higher~~ greater values for ~~automatically~~ automatically-traced fractures, because the traces are expected to be more sinuous than the ~~manually~~ manually-traced ones, given that~~;~~ the automatic code generates traces based on the detected ridges, ~~while~~ whereas a manual interpreter tends to trace using as few vertices as possible. Regions where $P_{21}$ is ~~higher~~

1935 greater for manual traces such as in NE2 at 14 m in x-direction and between 0 - 2 m in y-direction (Fig. 9) show cases where the interpreter can draw a trace based on geological knowledge and ~~not just the presence of a ridge~~no trace was identified automatically. More general examples for cases like this are~~, e.g.~~ thinning out or merging fracture openings where the automatic detection stops but the human interpreter continues the trace, resulting in longer or more numerous traces (see also e.g. Fig. ~~19~~23 mark 7). However, when using the $P_{ij}$ system that further

1940 subdivides the domains in smaller boxes to analyze the differences between two methods, it should be noted that apparent differences might also be cause by the position of the segments and nodes relative to the gird cells. The whole network is clipped by the grid cells and intersection points of some fractures may fall into a different cell due to a different interpretation, e.g. at the widely eroded intersections of fractures. When one intersection point falls into another cell due to a different interpretation, it not only increases the count in one cell, but also reduces

1945 the count in the other which results in neighboring cells with one having a positive and the other a negative difference (e.g. in NE1 at 5 m in x and 5 m in y direction, fig. 9).

Fig. ~~19~~ 23 ~~gives~~ illustrates a direct comparison of the manual interpretation and automatically ~~generated fracture~~extracted fracture networks. Increasing complexity of the fracture network causes more differences in the interpretations, mainly in the interpretation and number of nodes, while the average lengths of the branches only

1950 differ by a few centimeters (see also Table 10). The manual interpretation favors nodes of higher degree in direct comparison to the automatic interpretation in most cases (Fig. ~~19~~23, see also Fig. ~~18~~12). The biggest difference is the ~~higher~~ greater number of Y nodes counted in the automatic network~~.~~: This difference ~~is caused by~~ results from i) the overall larger number of fractures ~~present in the automatic code~~identified by the automatic process along with ii) the bias of the code towards nodes of lower degrees. This bias is caused by ~~B~~broadly eroded fractures that

1955 lead to inaccuracies when the code is tracing one of the edges instead the medial axis of the fracture, as a manual interpreter would do. This artifact leads to an overall higher number of nodes of smaller degree in the automatic traces for the same sample window as for manual traces (see Fig. 23 mark 4 and Table 10). More ~~small~~ short ~~fractures~~ traces are detected leading to a ~~higher~~ greater number of isolated nodes (Fig. ~~19~~23 mark 5). Due to the erosion the limestone becomes rough at its surface and can develop a structure which may be interpreted as a

1960 fracture trace in the automatic extraction. These false positives are caused when the parameters of the code are chosen in a way, that they are too sensitive. In the manual interpretation, these structures~~have been~~ were interpreted as erosional surface features and not as fractures. Thus, the difference here highlights the importance to find the right ~~sensitivity~~ parameter combination of the image processing parameters, to find the middle ground between the detection of ~~as a too high sensitivity will lead to~~ wrong positives and ~~a too low sensitivity to~~ false negatives.

1965 Compared to the ant-tracking method ~~in~~ (Gillespie et al.~~,~~ ~~(~~2011) applied to a much lower resolution (5cm/pixel) dataset, we deem our results more reliable because our spatial resolution of ~1 cm/pixel allows us to resolve details smaller than ~~is slightly smaller than~~ the observed width of the eroded fractures, that may add up to several cm.~~,~~ This resolution ~~allowing~~ allows ~~is~~ us to make interpretations based on features in the same scale, as it would be done directly on the outcrop.

1970 We infer that the automatic code at this stage represents a good option ~~to create~~for creating an initial fracture trace map ~~which~~ that only differs from a manual interpretation to a degree ~~which~~ that is comparable to the ~~degree of~~ deviation ~~by another manual interpreter~~of two manual interpretations of the same fracture network (e.g. Long et al., 2018). More complex tasks, like an interpretation of age relationships based on abutting and crosscutting criteria still require manual input. Based on this, future work can include the extension of the automatic mapping

1975 routine to the whole outcrop and use the manual interpretations to define criteria to combine automatically mapped branches into fractures and to assign fractures to predefined generations.

## 5.2 Classification into fracture generations

The classification of fractures into generations ~~requires~~ depends on the expertise of the interpreter. Locally in the sample windows, assignment to one of several generations is~~, several generations are~~ possible for a single fracture

1980 trace. In those cases, the interpreter has ~~to make a decision~~to decide with possible human bias. Examples for cases where an interpretation of the generations is non-unique or different interpretations of the underlying geometry are possible because of a locally lower quality of the data (e.g. eroded areas) are provided in figure ~~19.2~~24. Figure ~~19.2~~24a shows an example of a fracture interpreted as gen. 2 by an elimination process based on the predefined rules and interpretations. During this process, the other generations were ruled out and gen. 2 remained

1985 as the most likely one in the eyes of the interpreter. We interpret these cases as the reason for outliers in figures ~~11~~15 and ~~12~~17, in which the plotted point representing a fracture outlies the rough distribution of the other fractures associated to the same generation. Even though rules for the interpretation of generations based on abutting and crosscutting criteria may be clearly defined, their implementation is not trivial as shown in Fig. ~~19.2~~24a mark 1. At the shown location, a clear identification of the abutting relationship of the 3 fractures is not possible because

1990 the junction is eroded, and several fractures appear to intersect at the same location. In other cases, the interpreter must ~~make a decision~~decide whether a junction is the result of a splaying fracture or a younger fracture abutting at an older one (Fig. ~~19.2~~24b mark 2). While the dataset of ortho-mosaics provides a good contrast between eroded fracture and rock in most areas, few cases where the location of the fracture tip is not clearly distinguishable are possible (Fig. ~~19.2~~24b mark 3). The tracing of fractures as one-dimensional lines leaves further room for the

1995 interpretation of the position of the fracture when the original fracture has been eroded widely. In those cases, the interpretation ~~has to~~must be based on an area that envelops the actual fracture instead of clear trace of the fracture at the surface (e.g. ~~f~~Fig. ~~19.2~~24c mark 4 and panel d mark 7). Intersections of fractures that are widely eroded can be interpreted as an intersection of all fractures, or several spatially close intersections of several fractures (~~f~~Fig. ~~19.2~~24c mark 5), which influences the results of analyses of the network connectivity ~~and will be discussed later~~

2000 (~~e.g.~~c.f. ~~figure~~Fig. ~~19~~23). Other complex interpretations are required, when e.g. a long fracture matches the criteria for an old generation, but its trace appears to abut on a younger fracture, or the trace bends and continues with a geometry that qualifies the fracture for another generation (Fig. ~~19.2~~24d mark 6). One possible explanation for this~~, which is the reason for outliers in figures 11 and 12: some fractures have been assigned to a certain generation, but do not match the cluster of the associate fractures. This~~ outcome is ~~possible~~ when old fractures are

2005 reactivated and abut younger fractures or fractures interact with local features, e.g. preexisting fractures, that may ~~lead to a fracture geometry somewhat atypical for a certain set, even though the development is simultaneous~~cause distortions in the geometry of later-formed fractures. This effect is visible e.g. in figures ~~5~~15 and ~~6~~17 where~~our data as~~ fractures that overlap with the point clouds of the youngest generation in the plots were assigned to~~match~~

~~the distribution of the youngest generation but have been initially assigned to~~ an older generation during the interpretation.

Examples of the types of judgment calls that an interpreter may need to make for gen. 1 traces include accounting for the effect of censoring by sample windows on considerations of fracture lengths, assigning generation to a fracture trace with generations share orientations, and deciding if nearly parallel fractures are one or two generations.  For our early generations, fracture length is ~~not a reliable criterion~~a biased parameter, because the fractures can be censored, such that they are longer than the respective dimension of our map~~ping~~ domain. Th~~e is circumstance~~circumstance of possible censoring of long gen.1 fractures highlights the impact of fracture geometry on possible interpretations and results when selecting location and dimension of ~~mapping domains~~a sampling window. The strike directions of gen.1 fractures in the SW (100° - 120°) and NE (120° - 160°) differ, raising the question whether these are two independent sets. Considering the radial/converging pattern of gen.1, that can be observed best on a larger scale on the eastern part of "the bench". Here, we interpret the fractures ~~as~~with different ~~appearances~~trends as belonging to a larger structure, and hence a single generation. The underlying reason for this interpretation is that fractures ~~associated to a structure~~that are subordinate to a larger structure like this fault (Fig. 1b) can form simultaneously but in different directions. Thus, one criterion cannot rule out the other in this special case.

In the southwestern domains generations 1 and 2 are very similar in length and strike, so they ~~and~~ can only be distinguished when gen.2 fractures bend at the tips and abut on gen.1. Considering this observed geometry, another possible interpretation is to merge the first two generations in the SW into one, where the geometry is simply recording the order in which fractures of the same generation formed. We opted for two separate generations ~~in this work~~, because ~~they are~~ interpreted as two consecutive generations and therefore this decision does not have an impact on the analysis of the succeeding network development ~~analyses in time steps~~.

Gen. 2 is more distinct in NE1 than NE2 where gen. ~~1 is spaced more narrowly~~1 has greater abundance with narrower spacing, restraining the development of younger fractures on one hand and causing them to appear like gen. 5 on the other hand, possibly leading to a mix-up in the interpretation of the generations.

Generation 3 is present in all domains in the SW but absent in the NE. The ~~fading~~disappearance of such a distinct fracture generation over a relatively short distance of 200 meters can ~~have a mechanical cause but could also~~ be explained by a wrong association of fractures with other generations or subject to reasons that are not within the scope of this study. The geometry and distribution of preexisting fractures strongly influences the development of younger generations, possibly ~~leading~~creating local variations. Considering that generation 1 and 2 in the SW could belong to the same set, gen. 2 as mapped in the NE could be ~~associated with~~equivalent to generation 3 in the SW. This assignment would mean that all domains incorporate the same generations but with local variations in their geometry. Based on our analysis with five spatially isolated domains, this ~~is a question that~~determination cannot be ~~answered easily~~easily made, but requires a continuous tracing of the fracture generations over the complete ~~distance~~outcrop area. This interpretation does not have an impact on the network evolution analysis, because the generations are merged in the same order, but the situation highlights the necessity of a complete automatic tracing and interpretation of the whole outcrop.

Given the consistency of gen. 4 traces across the five sample windows, we focus next on ~~Generation 4 also referred to as left-stepping arrays as in (Rawnsley et al., 1998) are consistent in their geometry in all 5 subareas, indicating a larger stress field as reason of the fracturing.~~

2050 ~~g~~Generation. 5 fractures, that ~~result in~~ created polygonal patterns ~~are~~ with generally shorter traces than the other fracture generations (< 0.5 m) and are present in all areas. They ~~show~~ barely show ~~maxima in strike~~ orientation modes, but a qualitative inspection of fig. 15 and 17 suggests that a large fraction of the gen. 5 fractures is oriented ~~in~~in N-S and E-W directions ~~(Figures 11 and 12)~~, which appears to be caused by the influence of pre-existing fractures~~,~~ (e.g., in the NE2, where gen. 5 strikes between 20° and 60°, which is orthogonal to gen.1 fractures,

2055 representing the shortest connection between ~~those~~them).

For most cases, we expect and observe that younger fractures are shorter than older ones ~~when~~because pre-existing fractures acted as propagation barriers, constraining the maximum lengths of the younger fractures ~~restrain their development and thus maximum lengths~~. Counterintuitive to that, our gen.1 fractures can be shorter than gen. 2, and gen. 3 fractures shorter than gen. 4 (Tables 3 and 4). We ~~accredit~~interpret this anomaly in the case of gen. 1

2060 and 2 to ~~relate~~ the selection and orientation of our domains relative to the gen. 1 fractures, so that they are more strongly censored than gen. 2 fractures ~~as previously explained~~. The orientation of gen. 3 is subparallel to preexisting generations, thus the fractures are more likely to abut. This is not the case for gen. 4, which is sub-orthogonal to generations 1 and 2 und thus more likely to ~~protrude~~ cut through existing fractures, because they were able to propagate through the older fractures for reasons beyond the scope of this project (e.g., propagation

2065 stress conditions, existence of mineral fill in the older fractures, etc.). ~~the underlying stress cannot be accommodated sufficiently by the existing fractures. Another possible explanation is a stronger cementation of gen. 1 and 2 fractures between the events causing gen. 3 and 4, that allows younger fractures to crosscut older, recemented ones.~~

~~The sinuosity of all fractures is low on average indicating that they are rather straight than curved over lengths~~

2070 ~~covered in our detailed areas. Higher sinuosity can be expected e.g., for the radially converging (gen.1 NE) fractures when measured over the complete fracture length and not over a length defined by the dimensions of our mapping domains. We infer that larger mapping areas will lead to better results but also more intensive mapping times. Fracture geometry in the pavement can change over a distance of tens of meters (e.g. between NE1 and NE2 or the domains in the SW and NE), while the geometry and distribution of older fracture sets can strongly influence~~

2075 ~~the geometry and appearance of younger ones.~~

### 5.3 Network analysis

Analyzing the distribution and geometry of the branches shows that both can change over distances of a few meters within the same limestone layer. These changes may follow a local trend, e.g. decreasing skewness of the segment distribution plots (fig.4 and 5, Table 1) ~~from W to E~~ that indicates that the branch length distribution in the NE

2080 domains is closer to a symmetric one, while the distributions in the W are more asymmetric with a tail towards longer branches, suggesting that fractures in the W may consist of longer segments as in the NE. However, ~~but~~ ~~can~~ also strong fluctuations were observed ~~fluctuate strongly~~, e.g.~~such as~~ with ~~in~~ the number of branches, influencing the magnitude of ~~and thus the resulting~~ $P_{21}$ fracture intensity, which has a difference of almost a third from the ~~highest~~greatest value in SW2 to the ~~lowest~~smallest in NE2. Node distributions are linked to the number

2085 of branches and the way they interact. The decrease of I nodes from SW to NE suggests a ~~local~~consistent spatial trend~~;~~. ~~however~~However, this ~~assumption~~observation of a trend ~~is not backed by~~does not apply to the numbers of Y nodes, which fluctuate over short distances, e.g. from 2517 to 5100 from NE2 to NE1 (Table 2), a percentage difference of 68%. This effect underlines the heterogeneity of the fracture network, even though it might appear relatively homogeneous when observed qualitatively, and the necessity ~~in~~for sampling representative domains,

44

2090   when it is not possible to map the complete fracture network. In this study, we selected the domains primarily based on the quality of the data as explained in section 2.2.

When the sampling areas are supposed to be selected in a way to represent the complete fracture network in its variety, preliminary investigations are required. First steps to identify those representative domains can consist of a qualitative analysis of the network followed by a sparse interpretation of the most prominent fracture sets to

2095   reduce the risk of their lack in the chosen domains. However, a reliable statement whether the network in a small domain is representative for the entire network can only be made when the network has been analyzed entirely.


### 5.3.1 Network evolution

Topological analyses of the fracture network evolution show that average branch lengths decrease with additional

2100   fracture generations. This outcome is expected for non-parallel fracture sets which will eventually abut or crosscut each other, contemporaneously increasing the count of Y or X nodes. We identified nodes with more than four branches intersecting at one point in both, manually and automatically extracted traces. Depending on the number of intersecting branches, these nodes are treated as special cases of X nodes: Penta- (5), Hexa- (6), Hepta- (7) or Octa- (8) nodes. Due to the widespread erosion of the fractures at the surface, it is not possible to tell

2105   macroscopically, whether these they are narrowly spaced X and/or Y nodes or true nodes of a higher degree. In a spatially dense and strongly connected fracture network, we consider it as possible.

Isolated (I) nodes are more numerous in the initial network stages, where fractures have more space to develop and propagate through the limestone without encountering stress shadows of pre-existing fractures. At later stages, most of the I nodes have beenbecome connected to other fractures reducing their overall number. Compared to Y

2110   or X nodes, the number of I nodes is much lower in general, except for the initial fracture generations. In some cases, initial I nodes of old fracture generations appear to abut to younger fractures. Reasons for this geometry might be the reactivation of the fracture, or younger fractures connecting with the tip of the pre-existing ones (c.f. fig. 24). In these cases like this, a unique interpretation is not always possible because abutting criteria become unreliable and other criteria such as length and strike must be considered to aid the interpretation.

2115   The number of Y nodes increases when generations of similar orientation interact with each other or short undirected fractures connect larger ones. X nodes are often the result of intersecting fractures with orthogonal or sub-orthogonal orientation. Nodes of higher degree (5+ branches) are the result of X nodes, to which a younger fracture (mostly gen. 5) abuts. This is possibly caused by the local stress field, in which the existing node represents a zone of weakness to which later developing fractures converge.

2120   The average branch lengths (Tables 5 – 9) show a trend of decreasing branch lengthsAs expected for a network with with an increasing number of nodes. These trends are caused by due to an increasing number of of non-parallel fractures, branch lengths decrease with more crosscutting/abutting fracturesyounger fractures that crosscut or abut on the older ones. The longer the crosscutting fracture and the larger the deviation from the other fracture sets strike, the larger greater the decrease of the average branch length of the network (longer fractures can

2125   potentially cut more fractures).

Tables 5-9 show the development of branch lengths within thehave about the same order of magnitude over all subareas studiedsample windows. Especially in the final or recent stage of the network, the average branch lengths are very similar in all sample windows. This outcome indicates that the last fracture generation has a strong impact on the overall network topology. Older fracture generations have a larger influence on network geometry, because

pre-existing fractures influence the geometry of the ~~upcoming~~ later fracture generations in terms of possible fracture lengths and distribution. However, the topology can be very similar when younger fracture generations ~~overprint~~ develop and infill the network.

## 6 Conclusions

We used an UAV to take several sets of overlapping images from different altitudes to create orthorectified mosaics of the fractured limestone pavement on the coast near Lilstock in the Bristol Channel, UK. Based on these orthorectified mosaics, we selected 5 domains based on their outcrop quality and traced the fractures therein using two methods, an automatic trace segment extraction code from Prabhakaran et al., (2019), and manual tracing of the fractures. This allows us to compare both methods in terms of time usage, similarity of the resulting segment traces and network topology. Using the manual interpretation of fracture generations, we further analyze the evolution of the network connectivity and discuss spatial variations within the larger network based on the results of both techniques to highlight differences between the five domains. The main findings of this study are listed below.

- A comprehensive dataset of the fractured pavement with a so far unpreceded resolution was created using UAV photogrammetry.
- Automatic trace extraction of the fractures in the dataset is faster than manual tracing, when the parameter combinations used for the automatic extraction are chosen correctly. Furthermore, the manual tracing requires the interpreter's attention throughout the complete process, while the automatic trace extraction only needs supervision during some of the steps and further reduces the time that the interpreter actively spends on the task.
- When the parameter combinations are chosen improperly, the automatic method may produce a great number of artifacts that require manual corrections, that may take more time than an initially correct manual interpretation.
- Automatic trace extraction results in a greater number of overall segments as seen in the fracture trace density (Fig. 6 and 7 and supplement S3, Table 1). However, the overall identified fracture traces are similar in both methods, as suggested by similar fracture trace intensities in all domains (Fig, 8 and 9, Table 1).
- Resulting network topologies are similar for both methods, however, the automatic technique is biased towards a greater number of nodes of smaller degree, while a manual interpreter tends to create less segments and connects more branches at a single node.
- Using the automatic method, an interpretation of relative age relationships between fractures is not yet possible, this requires a manual interpretation.
- The five inferred fracture generations are not equally distributed throughout our five selected areas, ~~;~~ the spatial variation ~~can be~~ is significant in the same layer.
- The selected size of the mapping area can impact the measurements ~~e.g.~~ when the largest fractures are longer than the outlines of the map boundary.
- The connectivity of the fracture network increased ~~over~~ through time. The contribution of different generations of fractures to the network connectivity depends on their number and orientation relative to

46

2170 ~~pre-existing fractures. while different generations have a different impact on the overall connectivity based on their fracture numbers and orientation to pre-existing fractures.~~

- ~~Later fracture generations are influenced by preceding fracture generations.~~
- The network topology and connectivity in this area is strongly influenced by the last generation and varies between domains, in which the greatest connectivity was observed in the SW and the smallest in the NE.
- ~~Nodes with more than four intersecting branches are possible in this fracture network.~~
2175
- ~~When mapping within a 2D boundary, the selected size of the mapping area can impact the measurements, e.g. when the largest fractures are longer than the outlines of the mapping boundary.~~
- ~~While manual mapping is superior to the automatic mapping procedure, the required time can be reduced, and results of comparable quality produced.~~
- ~~The automatic interpretation of fracture generations is not yet possible and requires manual input.~~

2180

**References**

Andrews, B. J., Roberts, J. J., Shipton, Z. K., Bigi, S., Tartarello, M. C. and Johnson, G.: How do we see fractures? Quantifying subjective bias in fracture data collection, Solid Earth, 10(2), 487–516, doi:10.5194/se-10-487-2019,
2205 2019.

Agheshlui, H., Sedaghat, M. H. and Matthai, S.: Stress Influence on Fracture Aperture and Permeability of Fragmented Rocks, J. Geophys. Res. Solid Earth, 123(5), 3578–3592, doi:10.1029/2017JB015365, 2018.

2210 Agisoft, (2017). PhotoScan Professional (Version 1.3.2) [Software]. Retrieved from https://www.agisoft.com/downloads/installer/

Azizmohammadi, S. and Matthäi, S. K.: Is the permeability of naturally fractured rocks scale dependent?, Water Resour. Res., 53(9), 8041–8063, doi:10.1002/2016WR019764, 2017.

2215

Belayneh, M. and Cosgrove, J. W.: Fracture-pattern variations around a major fold and their implications regarding fracture prediction using limited data: an example from the Bristol Channel Basin, Geol. Soc. Lond. Spec. Publ., 231(1), 89–102, doi:10.1144/GSL.SP.2004.231.01.06, 2004.

2220 Belayneh, M., Geiger, S. and Matthäi, S. K.: Numerical simulation of water injection into layered fractured carbonate reservoir analogs, AAPG Bull., 90(10), 1473–1493, doi:10.1306/05090605153, 2006a.

Belayneh, M., Masihi, M., Matthäi, S. K. and King, P. R.: Prediction of vein connectivity using the percolation approach: model test with field data, J. Geophys. Eng., 3(3), 219–229, doi:10.1088/1742-2132/3/3/003, 2006b.

2225

Bemis, S. P., Micklethwaite, S., Turner, D., James, M. R., Akciz, S., Thiele, S. T. and Bangash, H. A.: Ground-based and UAV-Based photogrammetry: A multi-scale, high-resolution mapping tool for structural geology and paleoseismology, J. Struct. Geol., 69, 163–178, doi:10.1016/j.jsg.2014.10.007, 2014.

2230 Bevan, T. G. and Hancock, P. L.: A late Cenozoic regional mesofracture system in southern England and northern France, J. Geol. Soc., 143(2), 355–362, doi:10.1144/gsjgs.143.2.0355, 1986.

Bisdom, K., Nick, H. M. and Bertotti, G.: An integrated workflow for stress and flow modelling using outcrop-derived discrete fracture networks, Comput. Geosci., 103, 21–35, doi:10.1016/j.cageo.2017.02.019, 2017.

2235

Brooks, M., Trayner, P. M. and Trimble, T. J.: Mesozoic reactivation of Variscan thrusting in the Bristol Channel area, UK, J. Geol. Soc., 145(3), 439–444, doi:10.1144/gsjgs.145.3.0439, 1988.

Crider, J. G. and Peacock, D. C. P.: Initiation of brittle faults in the upper crust: a review of field observations, J.
2240 Struct. Geol., 26(4), 691–707, doi: 10.1016/j.jsg.2003.07.007, 2004.

Dart, C. J., McClay, K. and Hollings, P. N.: 3D analysis of inverted extensional fault systems, southern Bristol Channel basin, UK, Geol. Soc. Lond. Spec. Publ., 88(1), 393–413, doi:10.1144/GSL.SP.1995.088.01.21, 1995.

2245 Dershowitz, W. S. and Herda, H. H.: Interpretation of fracture spacing and intensity, in ARMA-92-0757, p. 10, American Rock Mechanics Association, Santa Fe, New Mexico. [online] Available from: https://doi.org/, 1992.

Dimmen, V., Rotevatn, A., Peacock, D. C. P., Nixon, C. W. and Nærland, K.: Quantifying structural controls on fluid flow: Insights from carbonate-hosted fault damage zones on the Maltese Islands, J. Struct. Geol., 101, 43–57, doi:10.1016/j.jsg.2017.05.012, 2017.

Duelis Viana, C., Endlein, A., Ademar da Cruz Campanha, G. and Henrique Grohmann, C.: Algorithms for extraction of structural attitudes from 3D outcrop models, Comput. Geosci., 90, 112–122, doi:10.1016/j.cageo.2016.02.017, 2016.

Engelder, T. and Peacock, D. C. P.: Joint development normal to regional compression during flexural-flow folding: the Lilstock buttress anticline, Somerset, England, J. Struct. Geol., 23(2), 259–277, doi: 10.1016/S0191-8141(00)00095-X, 2001.

Esri (2019). ArcGIS (Version 10.7.1) [Software].

Gillespie, P., Monsen, E., Maerten, L., Hunt, D., Thurmond, J. and Tuck, D., Eds.: Fractures in Carbonates: From Digital Outcrops to Mechanical Models, in Outcrops Revitalized: Tools, Techniques and Applications, SEPM (Society for Sedimentary Geology)., 2011.

Glen, R. A., Hancock, P. L. and Whittaker, A.: Basin inversion by distributed deformation: the southern margin of the Bristol Channel Basin, England, J. Struct. Geol., 27(12), 2113–2134, doi:10.1016/j.jsg.2005.08.006, 2005.

Hancock, P. L.: Jointing in the Jurassic limestones of the Cotswold Hills, Proc. Geol. Assoc., 80(2), 219–241, doi:10.1016/S0016-7878(69)80043-X, 1969.

Hancock, P. L. and Engelder, T.: Neotectonic joints, GSA Bull., 101(10), 1197–1208, doi:10.1130/0016-7606(1989)101<1197:NJ>2.3.CO;2, 1989.

Hansman, R. J. and Ring, U.: Workflow: From photo-based 3-D reconstruction of remotely piloted aircraft images to a 3-D geological model, Geosphere, 15(4), 1393–1408, doi:10.1130/GES02031.1, 2019.

Kelly, P. G., Peacock, D. C. P., Sanderson, D. J. and McGurk, A. C.: Selective reverse-reactivation of normal faults, and deformation around reverse-reactivated faults in the Mesozoic of the Somerset coast, J. Struct. Geol., 21(5), 493–509, doi:10.1016/S0191-8141(99)00041-3, 1999.

Long, J. J., Jones, R. R., and Daniels., S. E.: Reducing uncertainty in fracture modelling: assessing the sensitivity of inputs from outcrop analogues, in: The Geology of Fractured Reservoirs, 24-25 October 2018, The Geological Society of London, The Geological Society of London, oral Presentation, 2018.

Loosveld, R. J. H. and Franssen, R. C. M. W.: Extensional vs. Shear Fractures: Implications for Reservoir Characterisation, in SPE-25017-MS, p. 8, Society of Petroleum Engineers, Cannes, France., 1992.

Matthäi, S. K. and Belayneh, M.: Fluid flow partitioning between fractures and a permeable rock matrix, Geophys.
2290    Res. Lett., 31(7), L07602, doi:10.1029/2003GL019027, 2004.

Matthäi, S. K., Mezentsev, A. A. and Belayneh, M.: Finite Element - Node-Centered Finite-Volume Two-Phase-
Flow Experiments With Fractured Rock Represented by Unstructured Hybrid-Element Meshes, SPE Reserv. Eval.
Eng., 10(06), 740–756, doi:10.2118/93341-PA, 2007.

2295

Mauldon, M., Dunne, W. M. and Rohrbaugh, M. B.: Circular scanlines and circular windows: new tools for
characterizing the geometry of fracture traces, J. Struct. Geol., 23(2–3), 247–258, doi:10.1016/S0191-
8141(00)00094-8, 2001.

2300    Morley, C. K. and Nixon, C. W.: Topological characteristics of simple and complex normal fault networks, J.
Struct. Geol., 84, 68–84, doi:10.1016/j.jsg.2016.01.005, 2016.

Müller, D., Walter, T. R., Schöpa, A., Witt, T., Steinke, B., Gudmundsson, M.T. and Dürig, T.: High-Resolution
Digital Elevation modeling from TLS and UAV Campaign Reveals Structural Complexity at the 2014/2015
2305    Holuhraun Eruption Site, Iceland, Front Earth Sci, 5, doi:10.3389/feart.2017.00059, 2017.

Nemčok, M., Gayer, R. and Miliorizos, M.: Structural analysis of the inverted Bristol Channel Basin: implications
for the geometry and timing of fracture porosity, Geol. Soc. Lond. Spec. Publ., 88(1), 355–392,
doi:10.1144/GSL.SP.1995.088.01.20, 1995.

2310

Niethammer, U., James, M. R., Rothmund, S., Travelletti, J. and Joswig, M.: UAV-based remote sensing of the
Super-Sauze landslide: Evaluation and results, Eng. Geol., 128, 2–11, doi: 10.1016/j.enggeo.2011.03.012, 2012.

Nyberg, B., Nixon, C. W. and Sanderson, D. J.: NetworkGT: A GIS tool for geometric and topological analysis of
2315    two-dimensional fracture networks, Geosphere, 14(4), 1618–1634, doi:10.1130/GES01595.1, 2018.

Pahl, P. J.: Estimating the Mean Length of Discontinuity Traces, Int. J. Rock Mech. Min. Sci., 18, 221–228,
doi:0148-9062/81/030221-08502.00/0, 1981.

2320    Peacock, D. C. P. and Sanderson, D. J.: Displacements, segment linkage and relay ramps in normal fault zones, J.
Struct. Geol., 13(6), 721–733, doi:10.1016/0191-8141(91)90033-F, 1991.

Peacock, D. C. P. and Sanderson, D. J.: Geometry and Development of Relay Ramps in Normal Fault Systems,
AAPG Bull., 78(2), 147–165, 1994.

2325

Peacock, D. C. P. and Sanderson, D. J.: Strike-slip relay ramps, J. Struct. Geol., 17(10), 1351–1360,
doi:10.1016/0191-8141(95)97303-W, 1995.

Peacock, D. C. P., Sanderson, D. J. and Rotevatn, A.: Relationships between fractures, J. Struct. Geol., 106, 41–
53, doi:10.1016/j.jsg.2017.11.010, 2018.

Peacock, D. C. P., Sanderson, D. J., Bastesen, E., Rotevatn, A. and Storstein, T. H.: Causes of bias and uncertainty
in fracture network analysis, NJG, 99(1), 113–128, doi:10.17850/njg99-1-06, 2019.

Pix4D SA (Pix4D AG) (Pix4D Ltd), (2017). Pix4Dcapture (Version 2.1.0) [iOS application]. Retrieved from
https://apps.apple.com/de/app/pix4dcapture/id953486050

Prabhakaran, R., Bruna, P.-O., Bertotti, G. and Smeulders, D.: An automated fracture trace detection technique
using the complex shearlet transform, Solid Earth, 10(6), 2137–2166, doi:10.5194/se-10-2137-2019, 2019.

Priest, S. D. and Hudson, J. A.: Estimation of Discontinuity Spacing and Trace Length Using Scanline Surveys,
Int. J. Rock Mech. Min. Sci., 18, 183–197, doi:0148-9062/81/030183-15502.00/0, 1981.

Procter, A. and Sanderson, D. J.: Spatial and layer-controlled variability in fracture networks, J. Struct. Geol., 108,
52–65, doi:10.1016/j.jsg.2017.07.008, 2018.

Rawnsley, K. D., Peacock, D. C. P., Rives, T. and Petit, J.-P.: Joints in the Mesozoic sediments around the Bristol
Channel Basin, J. Struct. Geol., 20(12), 1641–1661, doi:10.1016/S0191-8141(98)00070-4, 1998.

Reisenhofer, R., Kiefer, J. and King, E. J.: Shearlet-based detection of flame fronts, Exp. Fluids, 57(3),
doi:10.1007/s00348-016-2128-6, 2016.

Rohrbaugh, M. B., Dunne, W. M. and Mauldon, M.: Estimating fracture trace intensity, density, and mean length
using circular scan lines and windows, AAPG Bull., 86, doi:10.1306/61EEDE0E-173E-11D7-
8645000102C1865D, 2002.

Sanderson, D. J. and Nixon, C. W.: The use of topology in fracture network characterization, J. Struct. Geol., 72,
55–66, doi:10.1016/j.jsg.2015.01.005, 2015.

Sanderson, D. J., Peacock, D. C. P., Nixon, C. W. and Rotevatn, A.: Graph theory and the analysis of fracture
networks, Journal of Structural Geology, 125, 155–165, doi:10.1016/j.jsg.2018.04.011, 2019.

Vasuki, Y., Holden, E.-J., Kovesi, P. and Micklethwaite, S.: Semi-automatic mapping of geological Structures
using UAV-based photogrammetric data: An image analysis approach, 25.04.2014, 69, 22–32, 2014.

Watkins, H., Bond, C. E., Healy, D. and Butler, R. W. H.: Appraisal of fracture sampling methods and a new
workflow to characterise heterogeneous fracture networks at outcrop, J. Struct. Geol., 72, 67–82,
doi:10.1016/j.jsg.2015.02.001, 2015.

2370 Weismüller, C., Urai, J. L., Kettermann, M., von Hagke, C. and Reicherter, K.: Structure of massively dilatant faults in Iceland: lessons learned from high-resolution unmanned aerial vehicle data, Solid Earth, 10(5), 1757–1784, doi:10.5194/se-10-1757-2019, 2019.

Westoby, M. J., Brasington, J., Glasser, N. F., Hambrey, M. J. and Reynolds, J. M.: 'Structure-from-Motion' 2375 photogrammetry: A low-cost, effective tool for geoscience applications, Geomorphology, 179, 300–314, doi: 10.1016/j.geomorph.2012.08.021, 2012.

Wu, H. and D. Pollard, D.: An experimental study of the relationship between joint spacing and layer thickness, J. Struct. Geol., 17(6), 887–905, doi:10.1016/0191-8141(94)00099-L, 1995.

2380
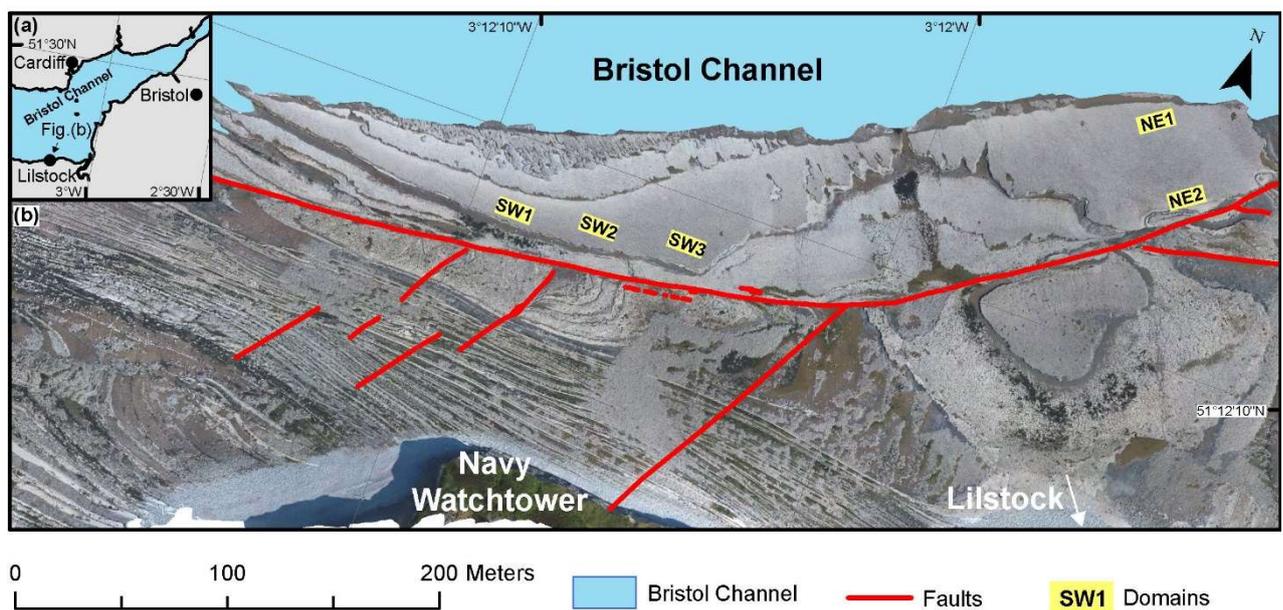


**Fig. 1: (a) Location of the study area in the Bristol Channel, Great Britain. (b Ortho-rectified photo-mosaic of UAV-photographs taken from 100 m. The ortho-mosaic shows the coast at Lilstock during low tide,**
2385 **exposing the fractured limestone. The ~~study areas~~map domains on "the bench" ~~that are mapped in detail~~ are marked as yellow rectangles ~~in red~~.**
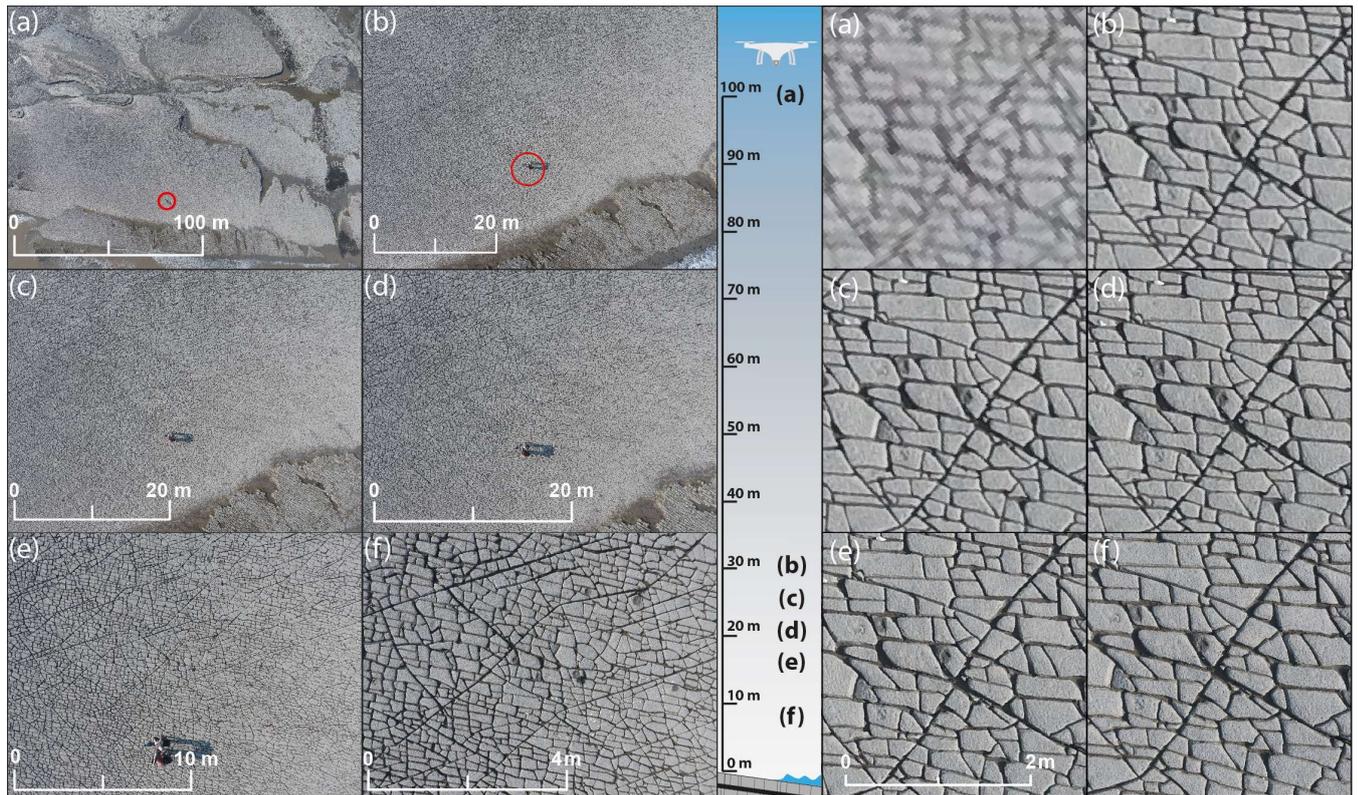
**Fig. 2: The fractured limestone captured from above by UAV. Left: View from different altitudes with altitudes decreasing from (a) to (f). The exact altitudes from which the photographs have been taken can be viewed in the sketch in the center. Right: Same photographs as on the left, zoomed in to the same degree and location in every image. With increasing altitude, the spatial resolution of the image decreases.Note: persons for scale and encircled for better visibility in (a) and (b) on the left.**

2390
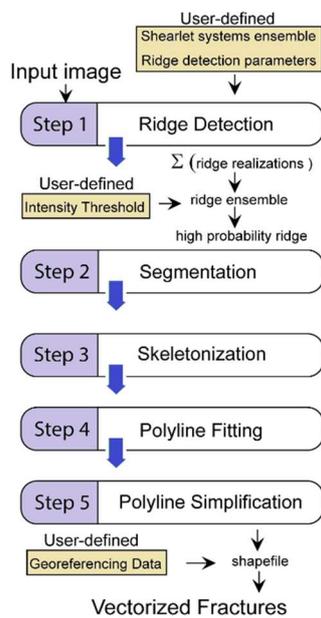


2395

**Fig. 3: Overview of the automated fracture detection process (reproduced with permission from Prabhakaran et al, 2019)**
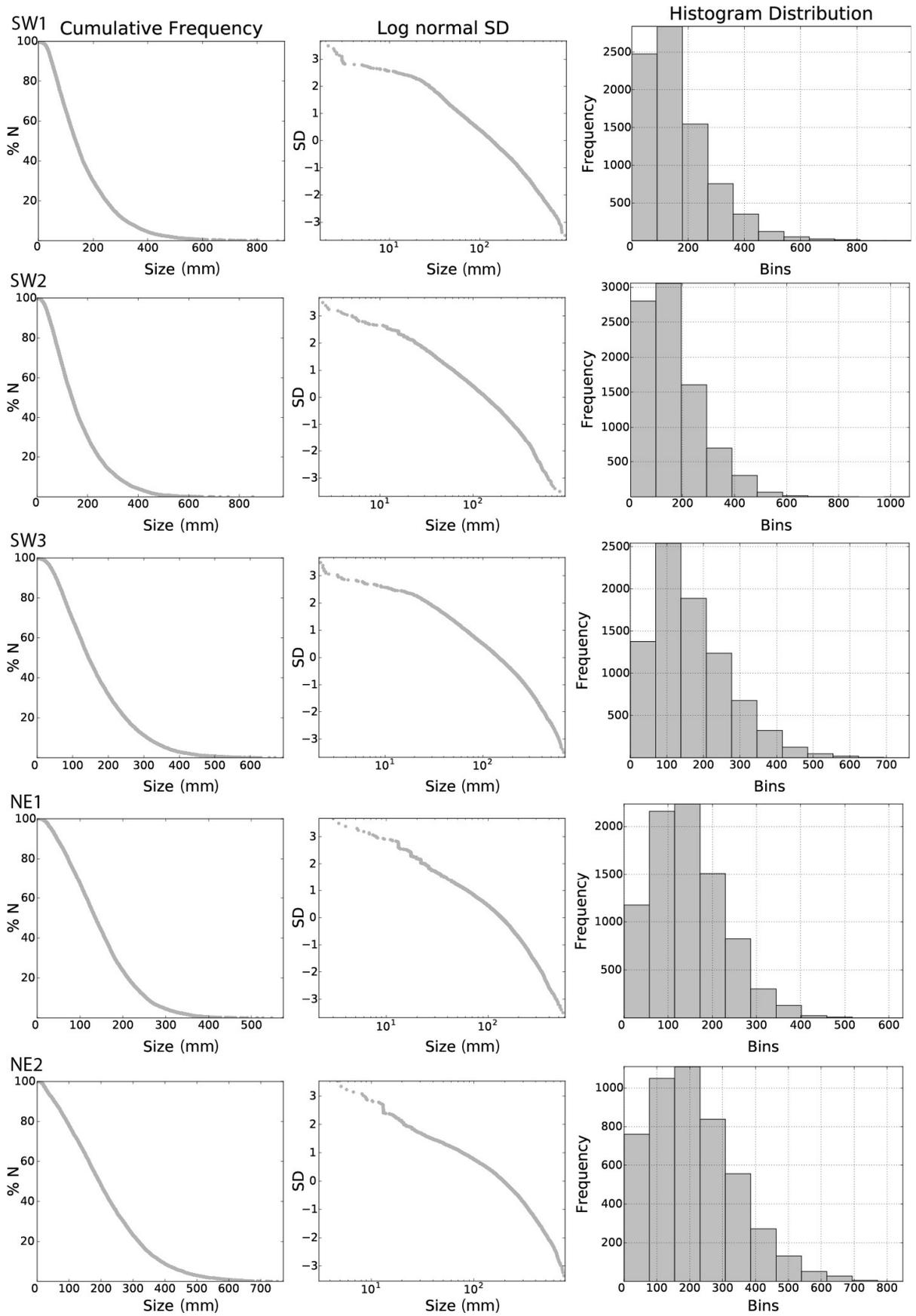
2400

54
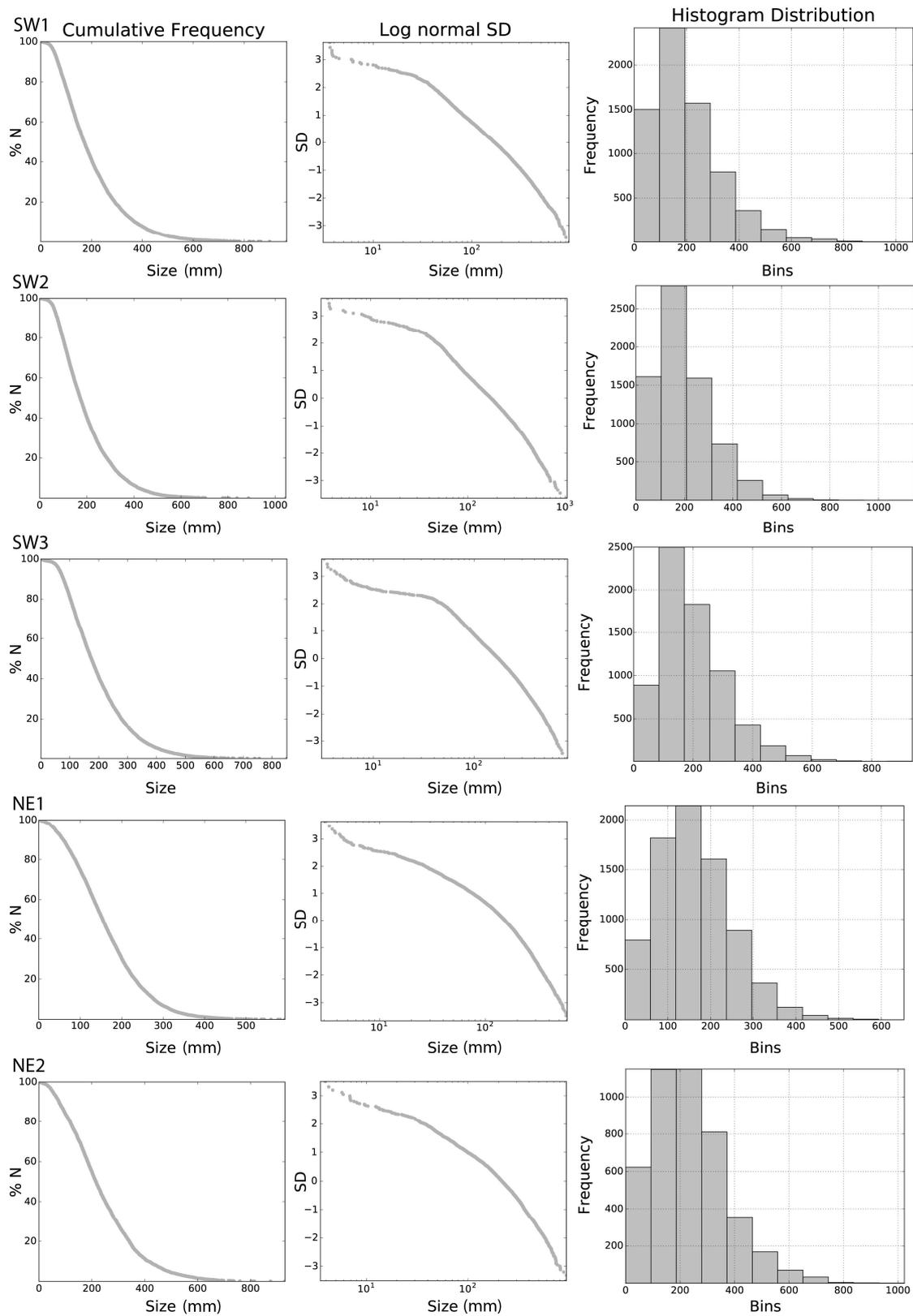
**Fig. 5: Cumulative length distribution, log normal standard deviation and histogram distributions of the manually mapped and segmented fracture traces (branches) in the five domains.**
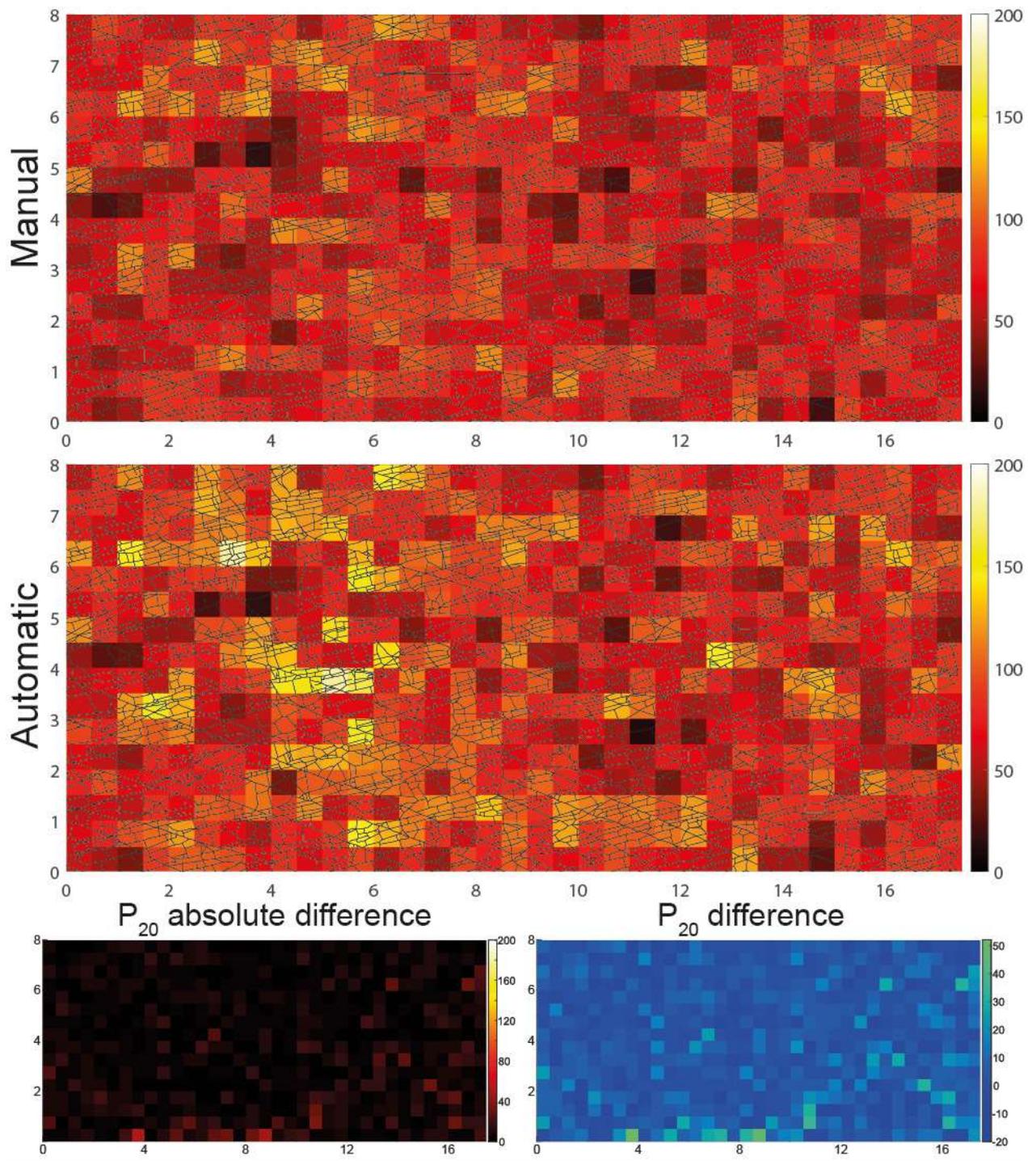
**Fig. 6: Comparison of the $P_{20}$ fracture segment density between manually an automatically-traced fracture segments for SW3. Absolute and relative differences are shown in the bottom line. Unit of the axes in (m), unit of the color bar in (1/m²).**

2405
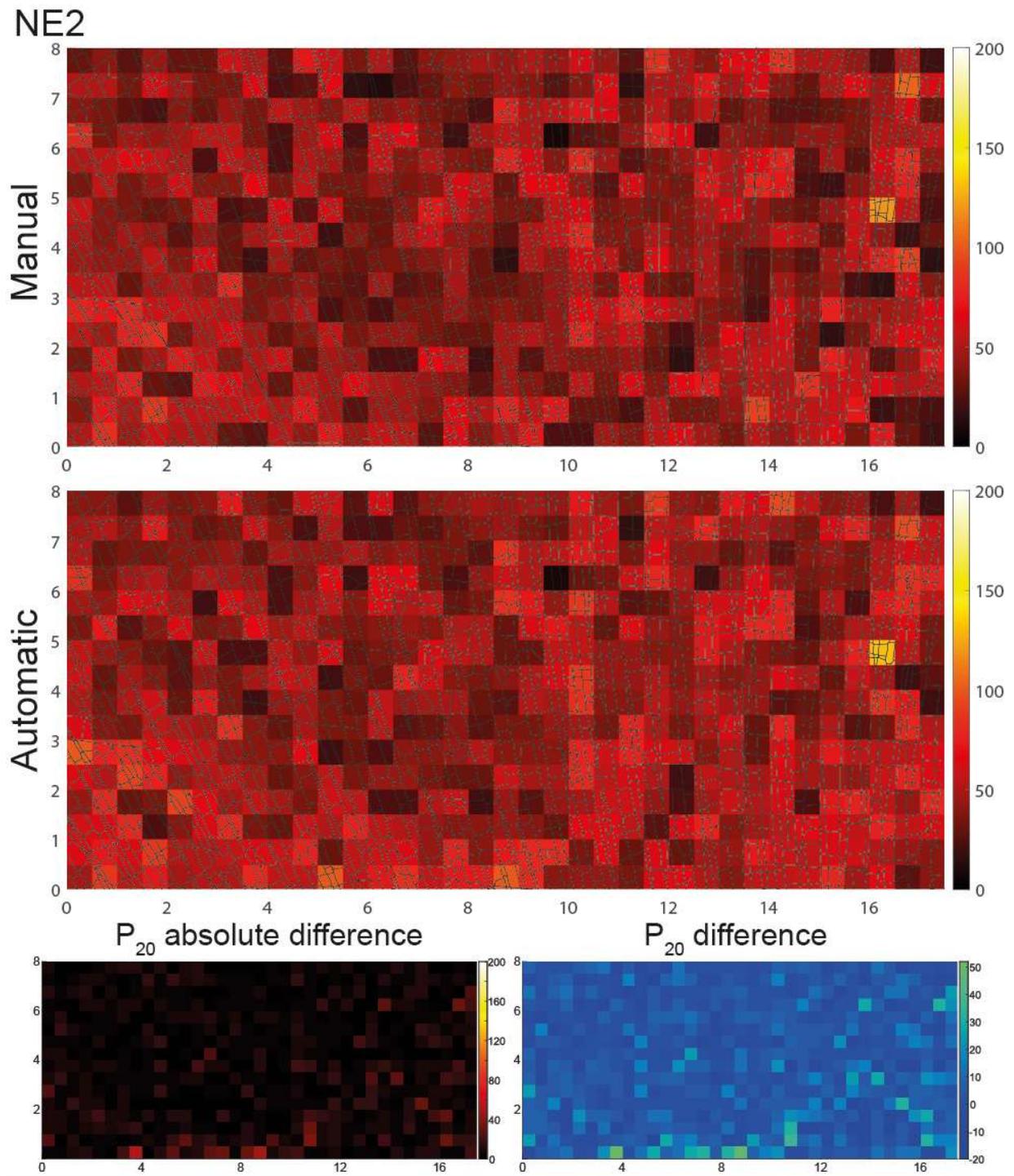
2410

**Fig. 7: Comparison of the $P_{20}$ fracture segment density between manually an automatically-traced fracture segments for NE2. Absolute and relative differences are shown in the bottom line. Unit of the axes in (m), unit of the color bar in (1/m²).**
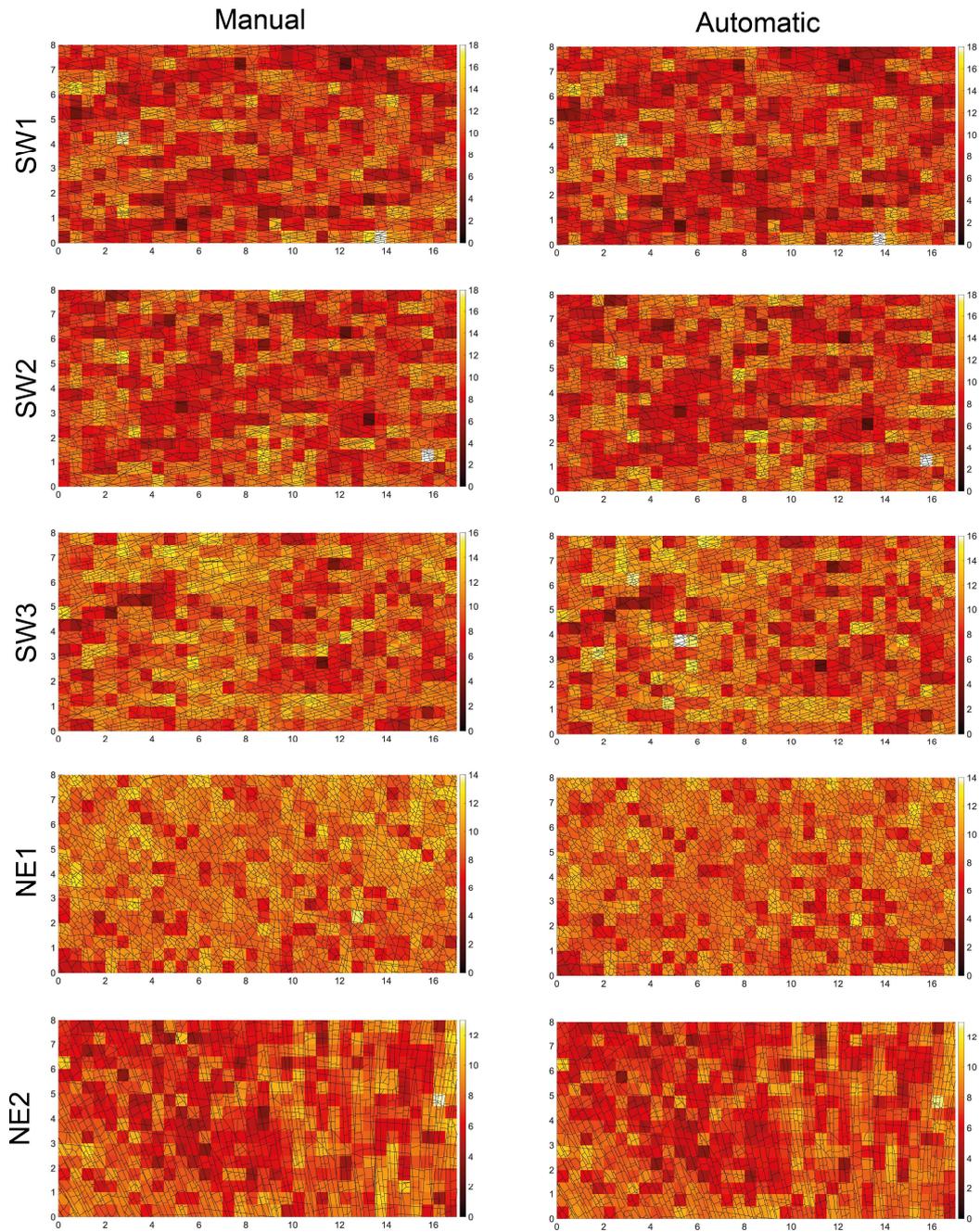
2415

**Fig. ~~4~~ 8: P$_{21}$ fracture intensity plots of the mapped fractures within the 5 subareas for the manual and automatic interpretations. Unit of the axes in (m), unit of the color bar in (m/m²), where lighter is more dense.**
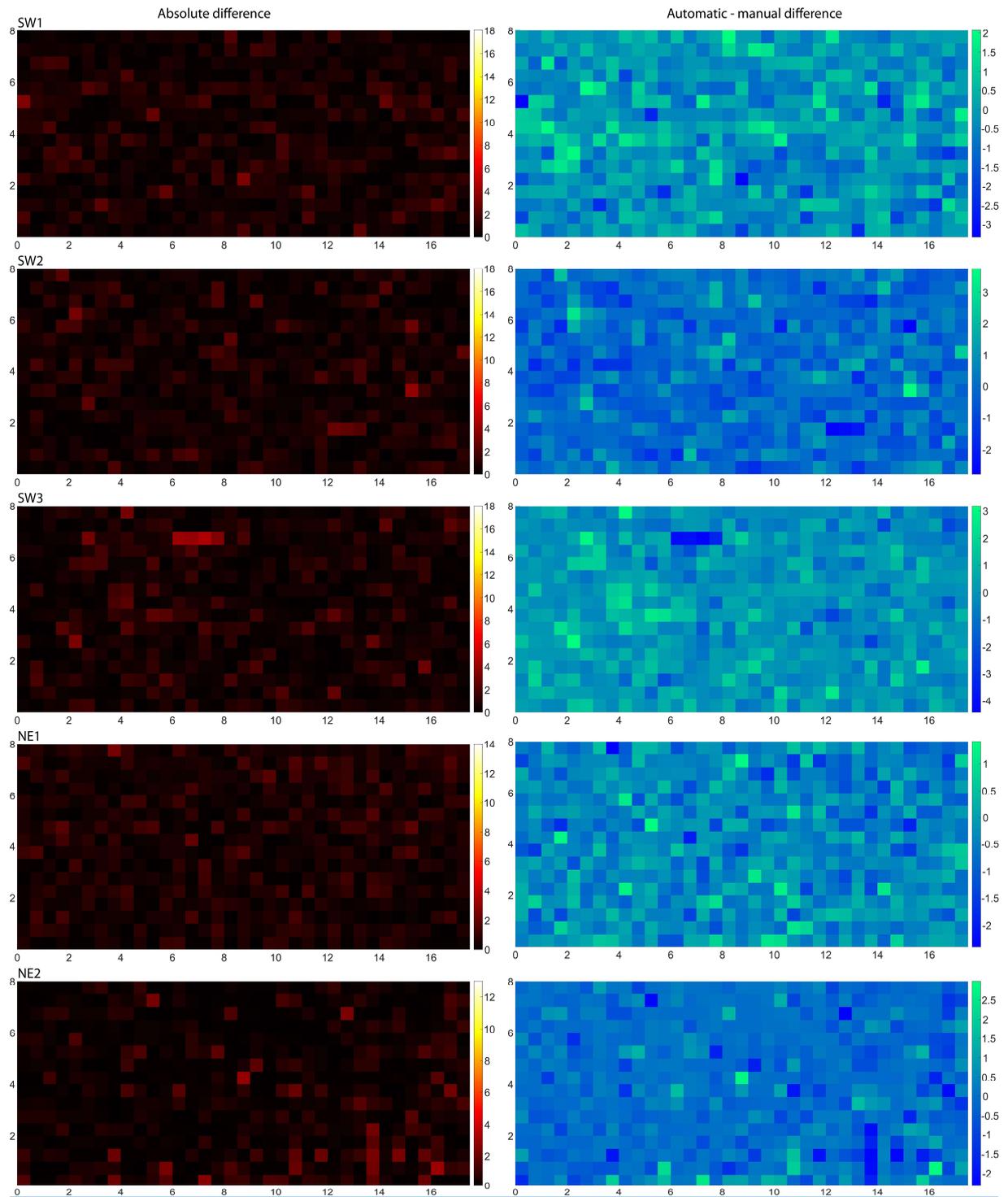
**Fig. 9: Differences of the $P_{21}$ fracture segment intensity between manually- and automatically-traced fracture segments for all five domains. Absolute differences are shown in the left column, relative differences in the right column. Unit of the axes in (m), unit of the color bar in (m/m²).**

2420

59

SW1

SW2

SW3

0   5   10 Meters   ☐ Boundaries   N

2425   ● E   ● Hepta   ● Hexa   ● I   ○ Penta   ■ X   ▲ Y   ── C - C   ── C - I   ── I - I
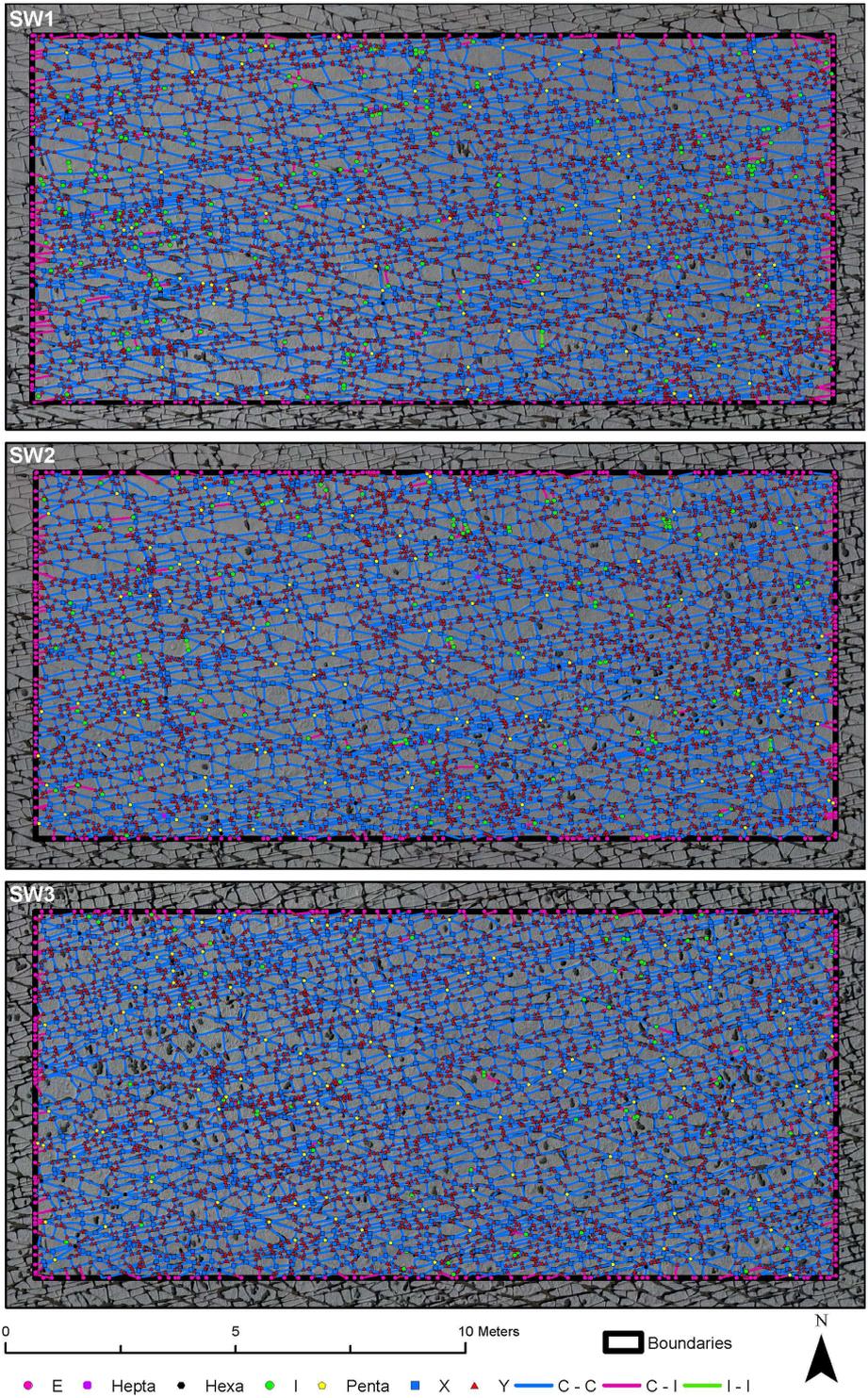
**Fig. 5 10: SW fracture networks, automatically automatically traced and plotted as branches and nodes.**

60
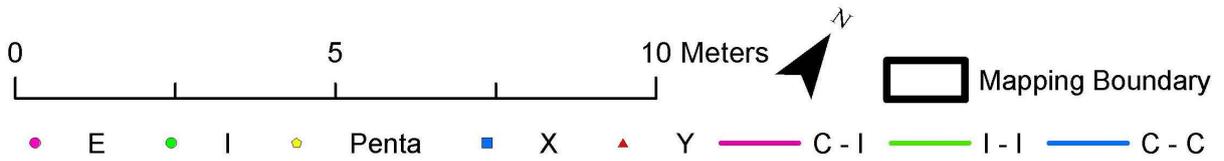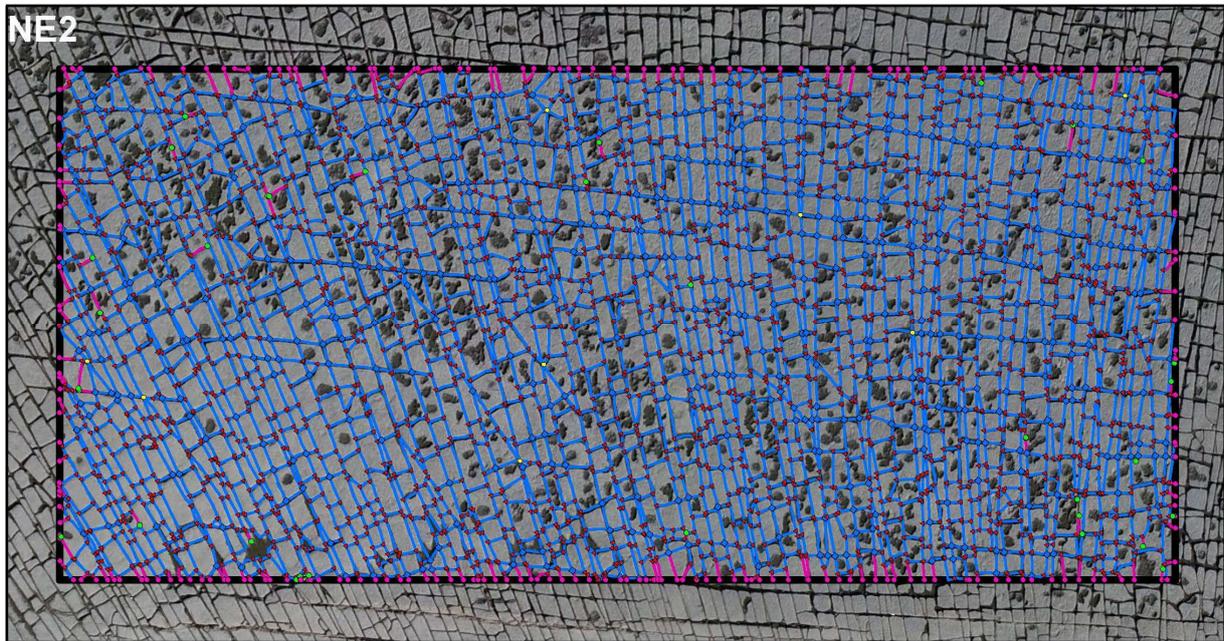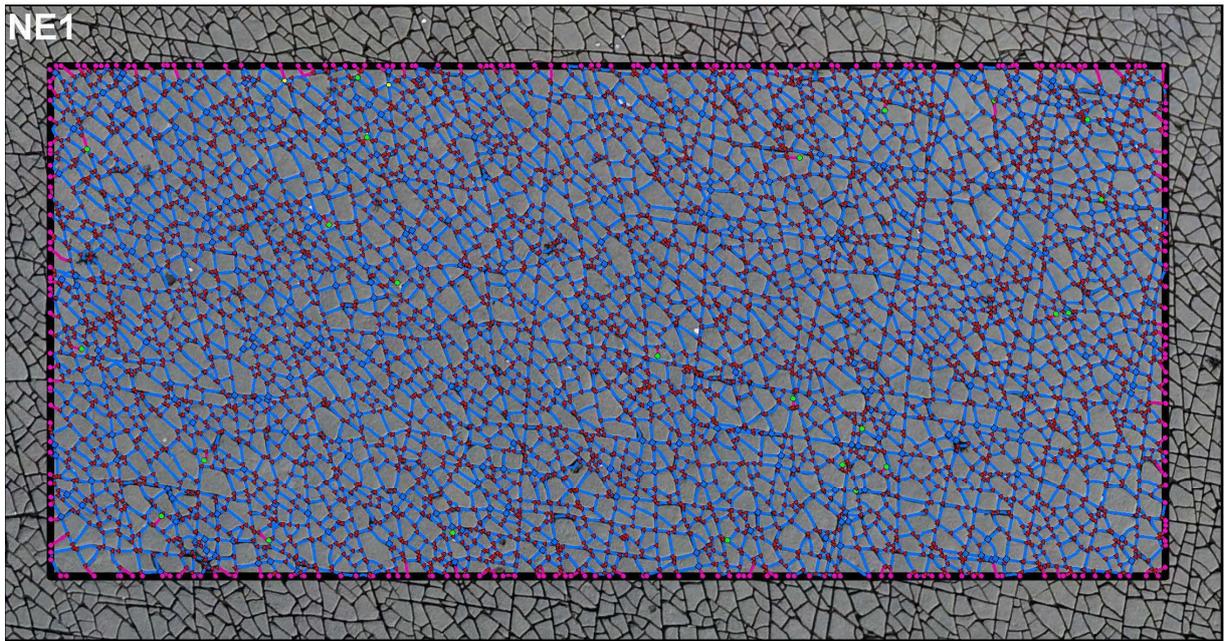
**Fig. ~~6~~ 11: NE fracture networks, ~~automatically~~ automatically-traced and plotted as branches and nodes.**

Manual interpretation
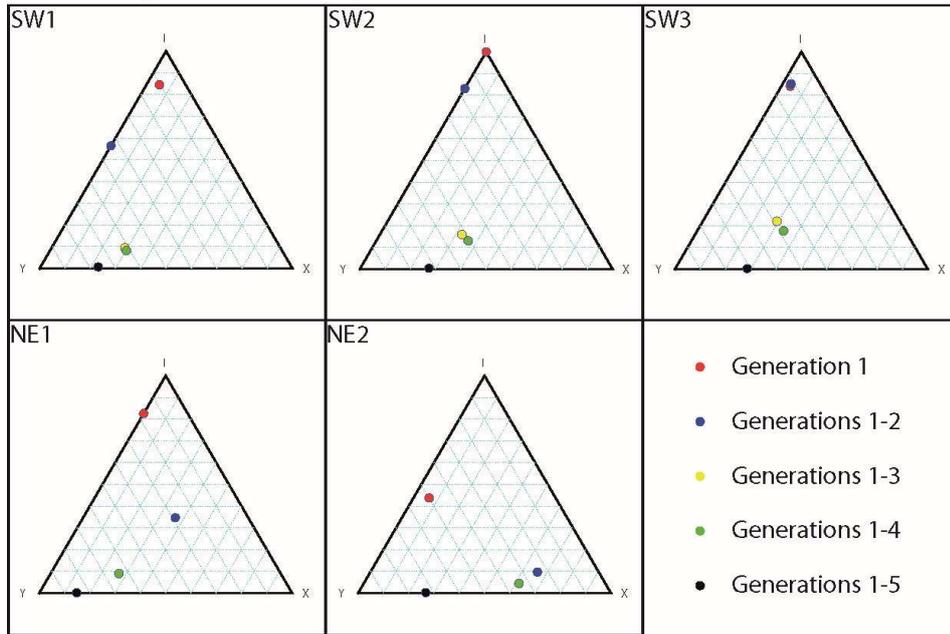


Automatically generated networks
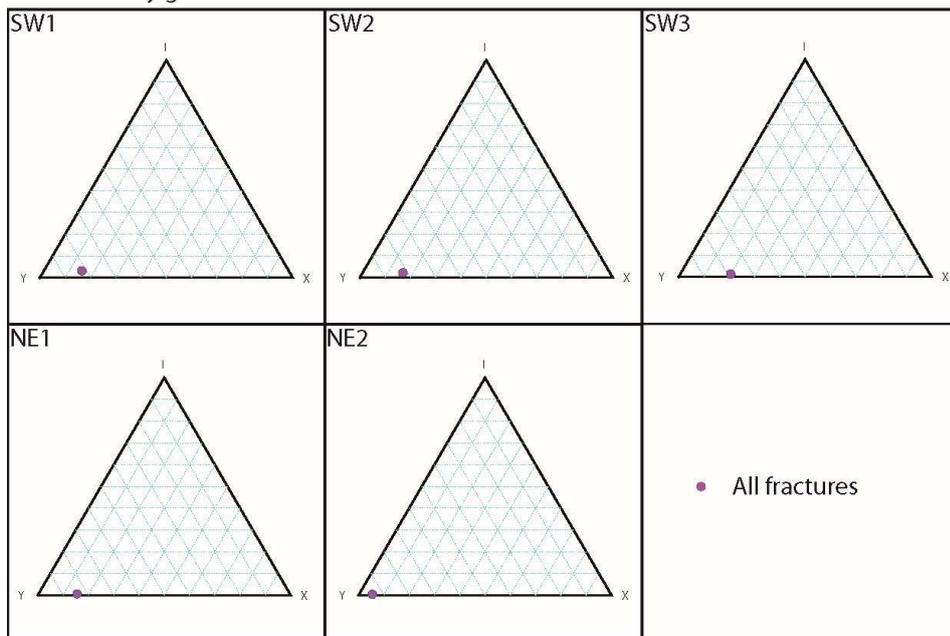


2430  **Fig. 12: Ternary plots visualizing the node distributions. For the manual interpretations, the different network stages in terms of fracture generations are plotted. The automatic networks only show the final stage of the network, which corresponds to generations 1 – 5 in the manual interpretation. Nodes with four and more than four intersecting branches have been binned as type X nodes.**

62
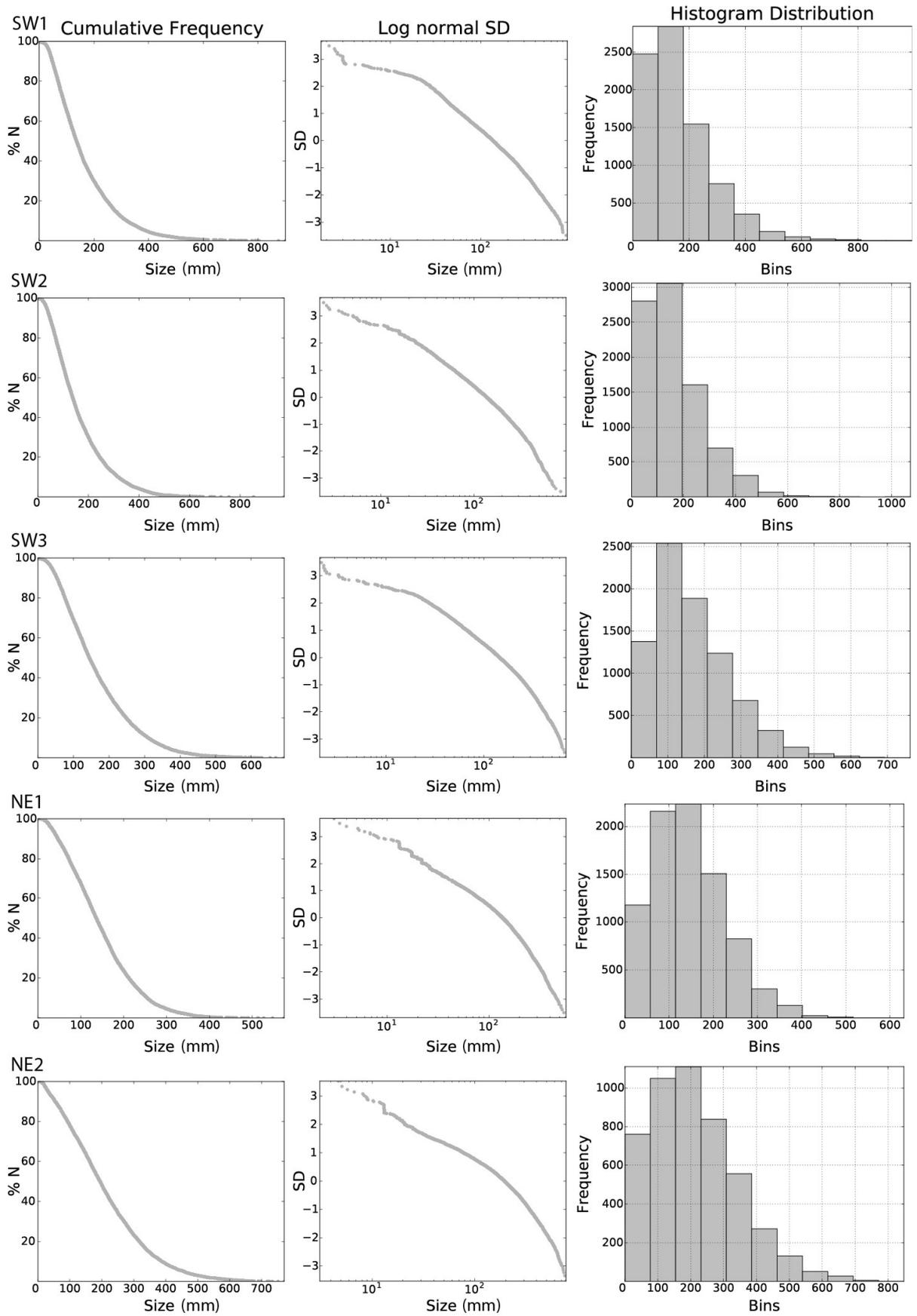
Fig. 7: Cumulative length distribution, log normal standard deviation and histogram distributions of the automatically traced fractures (branches) in the five domains.

**SW1**
0° 330° 30° 300° 60° 270° 90°
10 8 6 4 2
Mean: 16.96° n=8184

**SW2**
0° 330° 30° 300° 60° 270° 90°
10 8 6 4 2
Mean: 27.1° n=8548

**SW3**
0° 330° 30° 300° 60° 270° 90°
15 12 9 6 3
Mean: 33.44° n=8239

**NE1**
0° 330° 30° 300° 60° 270° 90°
4 3 2 1
Mean: 169.06° n=8366

**NE2**
0° 330° 30° 300° 60° 270° 90°
8 6 4 2
Mean: 161.48° n=4805
Bin Size: 10.0°
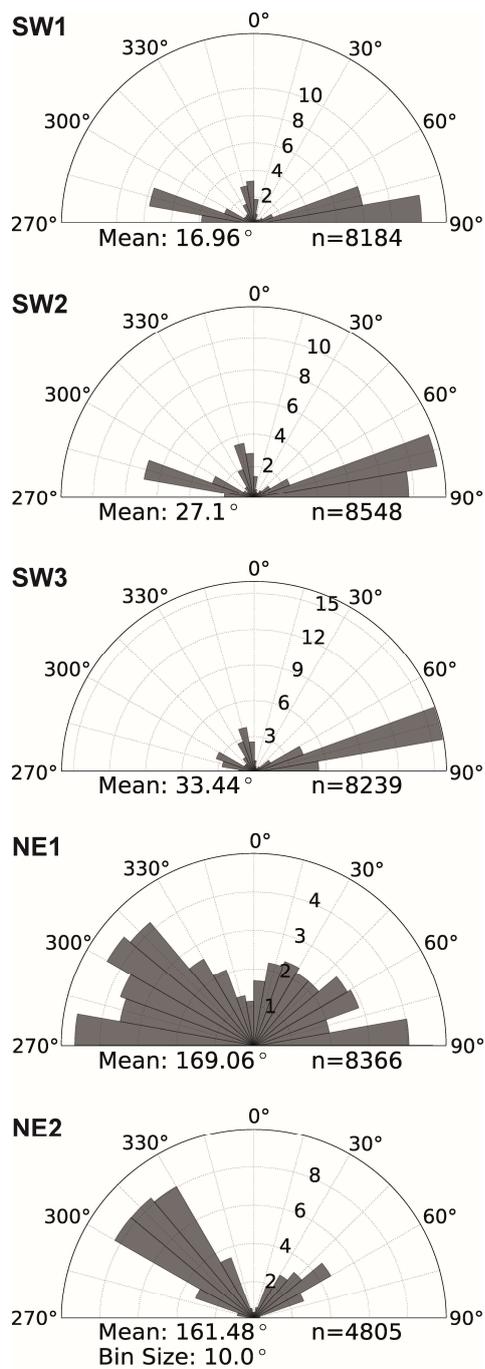
**Fig. 13: Length weighted rose plots of automatically extracted fracture trace segments (branches) in the five domains.**

2440

**SW1**

Mean: 16.96°   n=8184

**SW2**

Mean: 27.1°   n=8548

**SW3**

Mean: 33.44°   n=8239

**NE1**

Mean: 169.06°   n=8366

**NE2**

Mean: 161.48°   n=4805
Bin Size: 10.0°
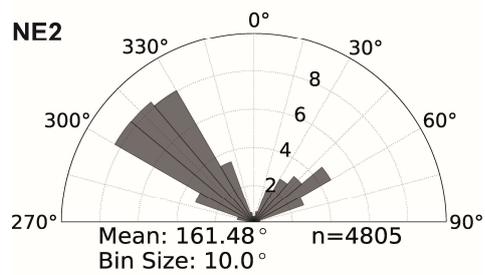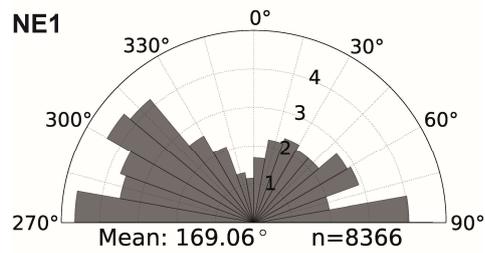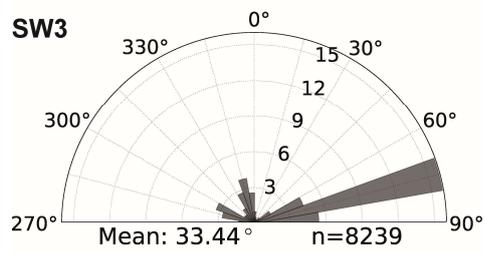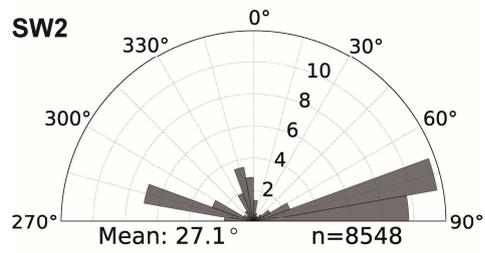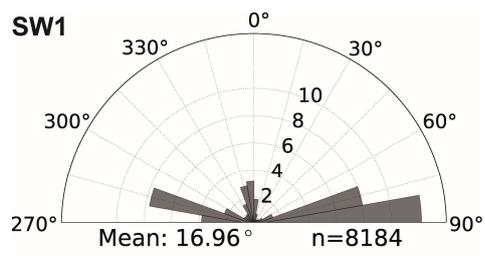
Fig. 8: Length weighted rose plots of automatically traced fractured in the five domains.

**SW 1**

Mean: 23.92°   n=2719

**SW 2**

Mean: 32.52°   n=2404

**SW 3**

Mean: 36.39°   n=2720

0   5   10 Meters

N

— Gen. 1   — Gen. 3   — Gen. 5
— Gen. 2   — Gen. 4   Bin Size: 10°
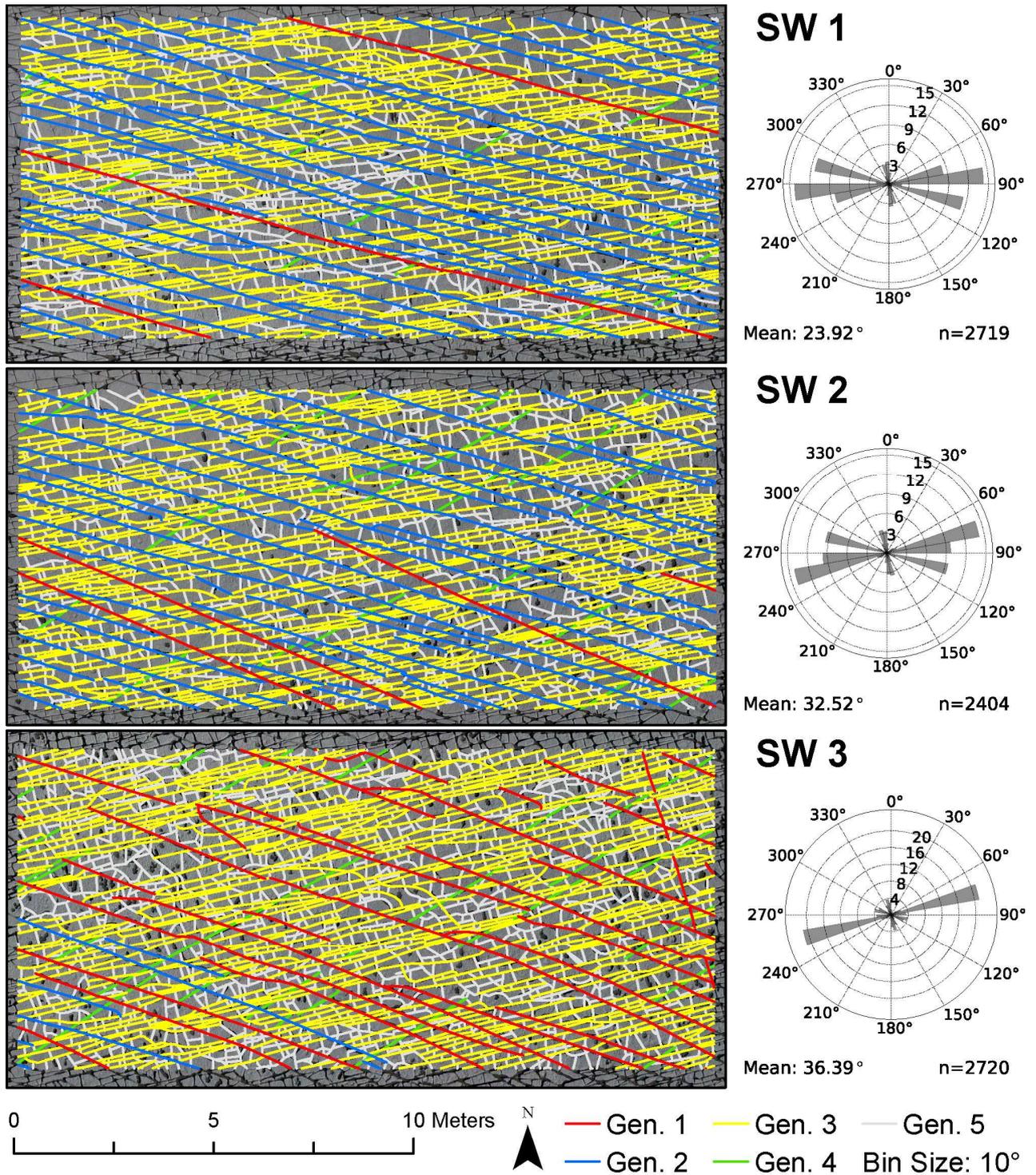
2445

**Fig.** ~~9~~ **14: Interpretation of fracture generations** _in manually interpreted data_ **in the domains in the SW.** ~~Fracture orientations are visualized as length weighted rose plots~~_Length-weighted rose diagrams showing fracture abundance as a function of the trends of fracture traces_.
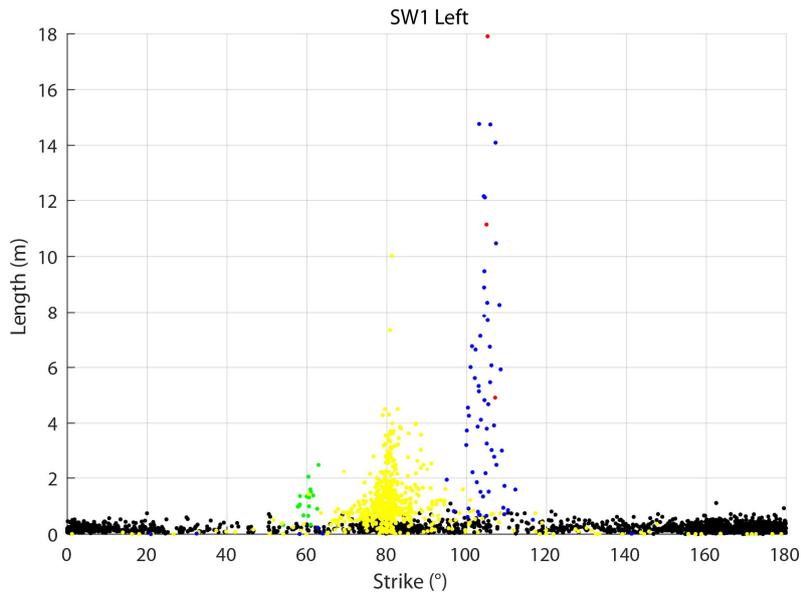
**Fig. 15: Plots of strike direction (°) on the x-axis and length (m) on the y-axis for every fracture mapped in the domains in the SW, color coded by generation.**

**NE 1 (upper)**

Mean: 156.95°     n=3972

**NE 2 (lower)**

Mean: 161.63°     n=1328

0       5       10 Meters

Gen. 1     Gen. 2     Bin Size: 10.0°
Gen. 4     Gen. 5
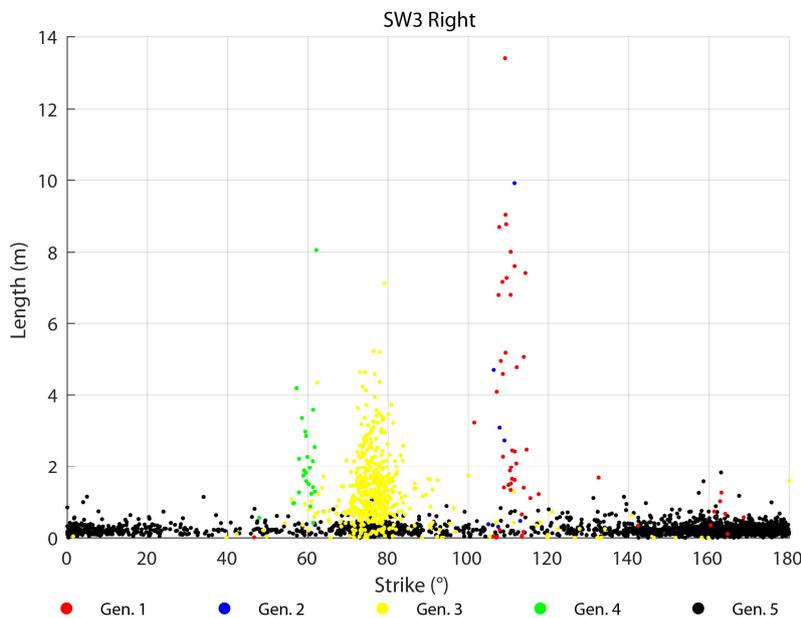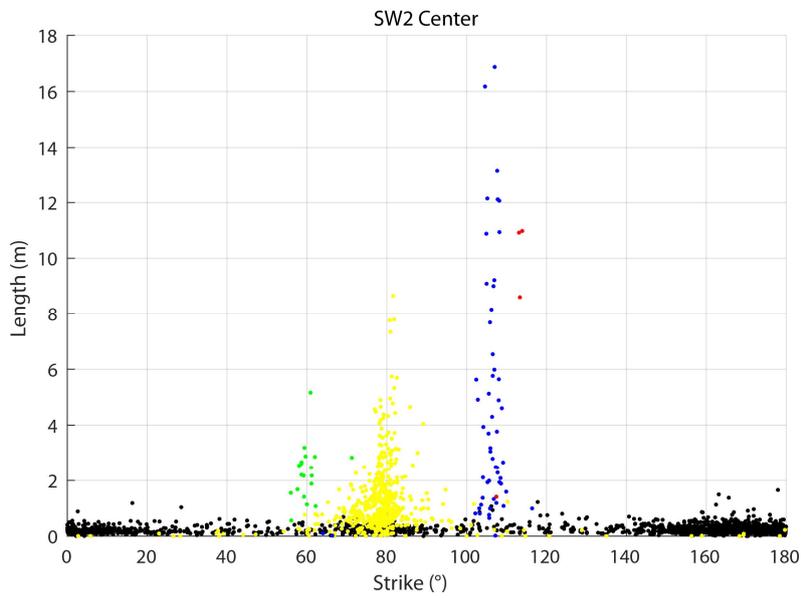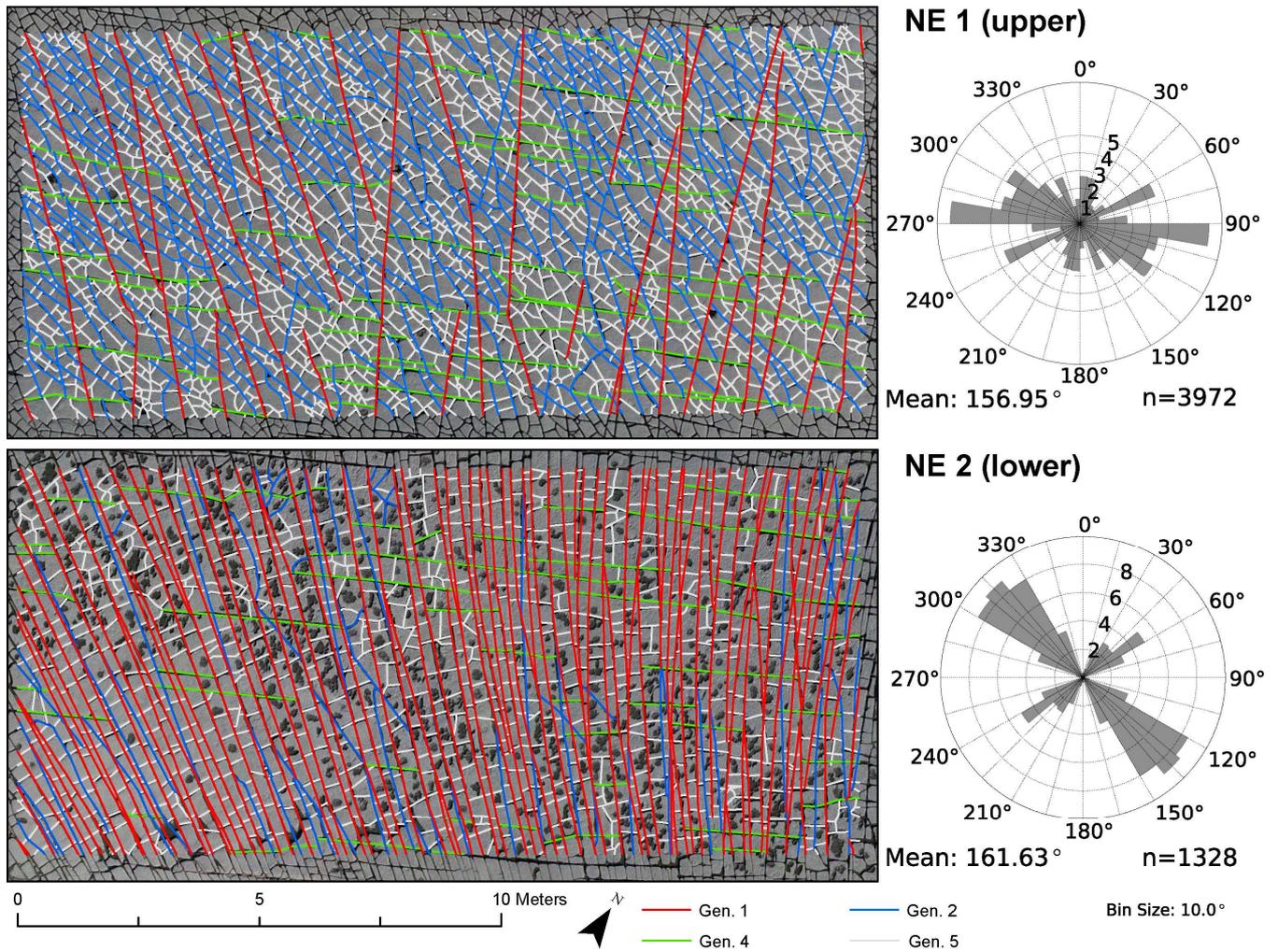
**Fig. ~~10~~16 : Interpretation of fracture generations in manually interpreted data in the domains in the NE. ~~Fracture orientations are visualized as length weighted rose plots. Note that the fracture networks on the left are oriented NW-SE.~~ Length-weighted rose diagrams showing fracture abundance as a function of the trends of fracture traces. Note: the map domains have a NW (top) – SW (bottom) orientation as indicated by the north arrow.**

2460

69

**NE1 Upper area**

**NE2 Lower area**

● Gen. 1     ● Gen. 2     ● Gen. 4     ● Gen. 5

2495 **Fig. ~~12~~ 17: Plots of strike direction (°) on the x-axis and length (m) on the y-axis for every fracture mapped in the domains in the NE, color coded by generation.**
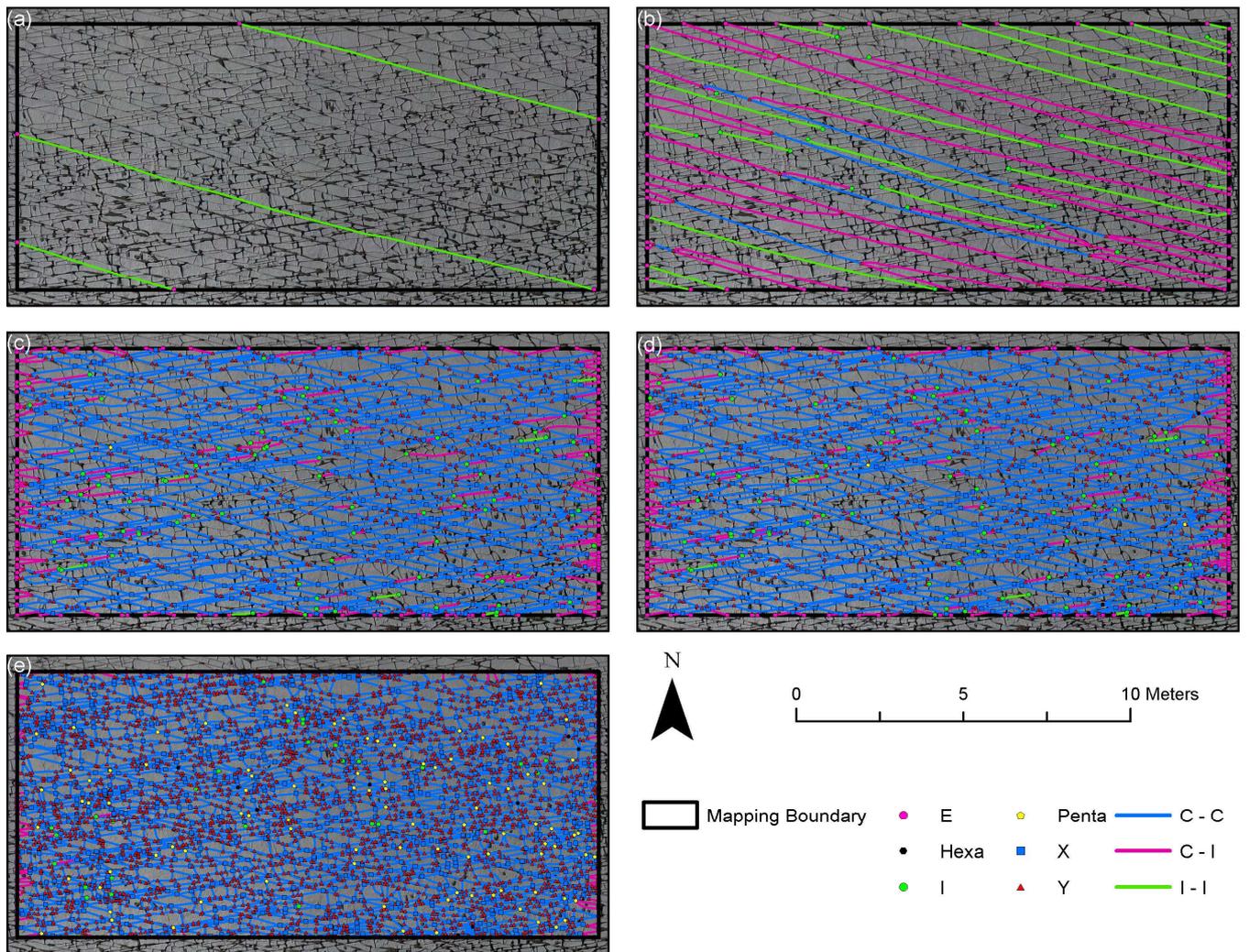
**Fig. ~~13~~ 18: Development of the fracture network in time steps for SW1 from generation 1 to 5, visualized as branches and nodes. The consecutive fracture generation is added for each panel from (a) to (e).**

2500
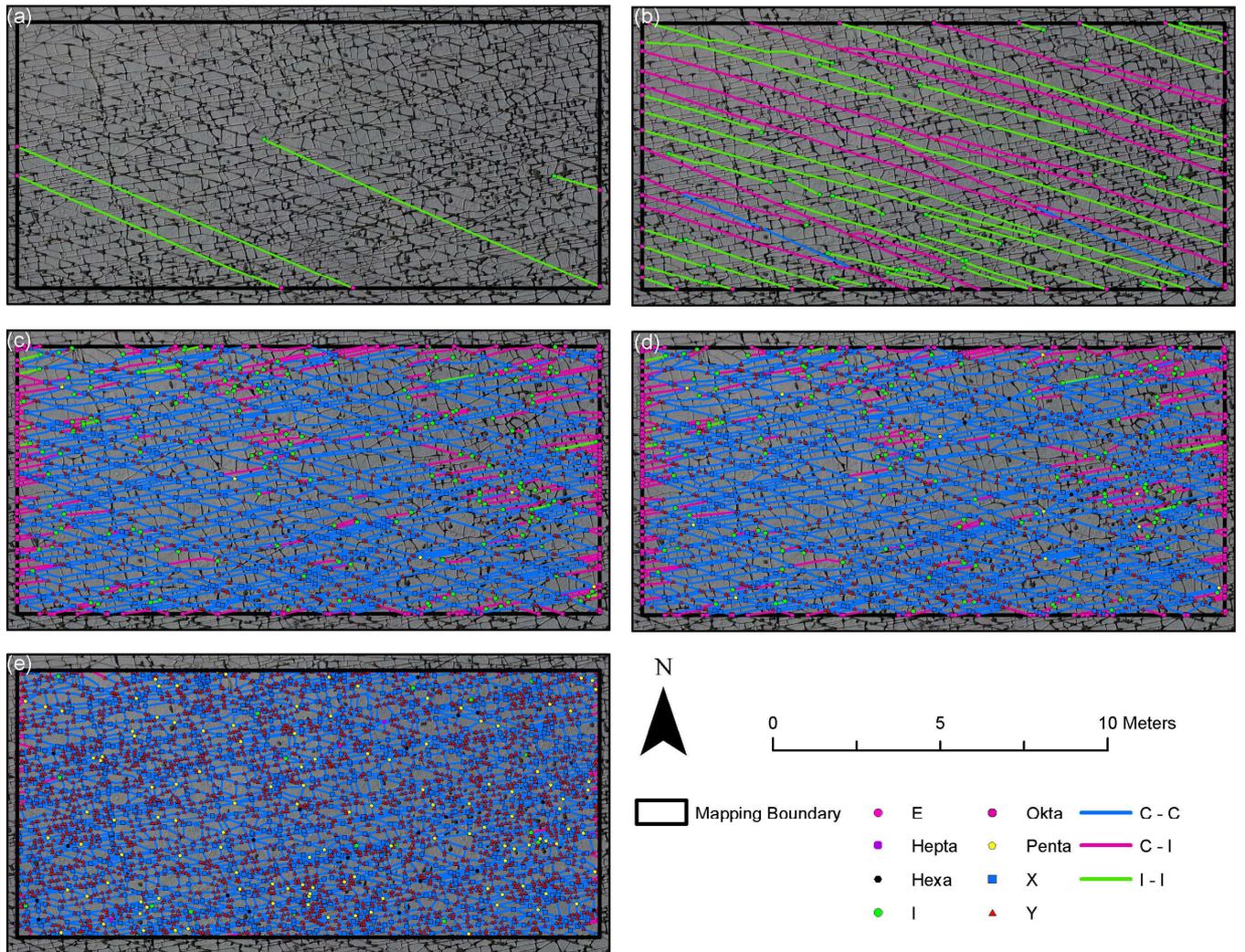
**Fig. ~~14~~ 19**: Development of the fracture network in time steps for SW2 from generation 1 to 5, visualized as branches and nodes. The consecutive fracture generation is added for each panel from (a) to ~~(e)~~E.
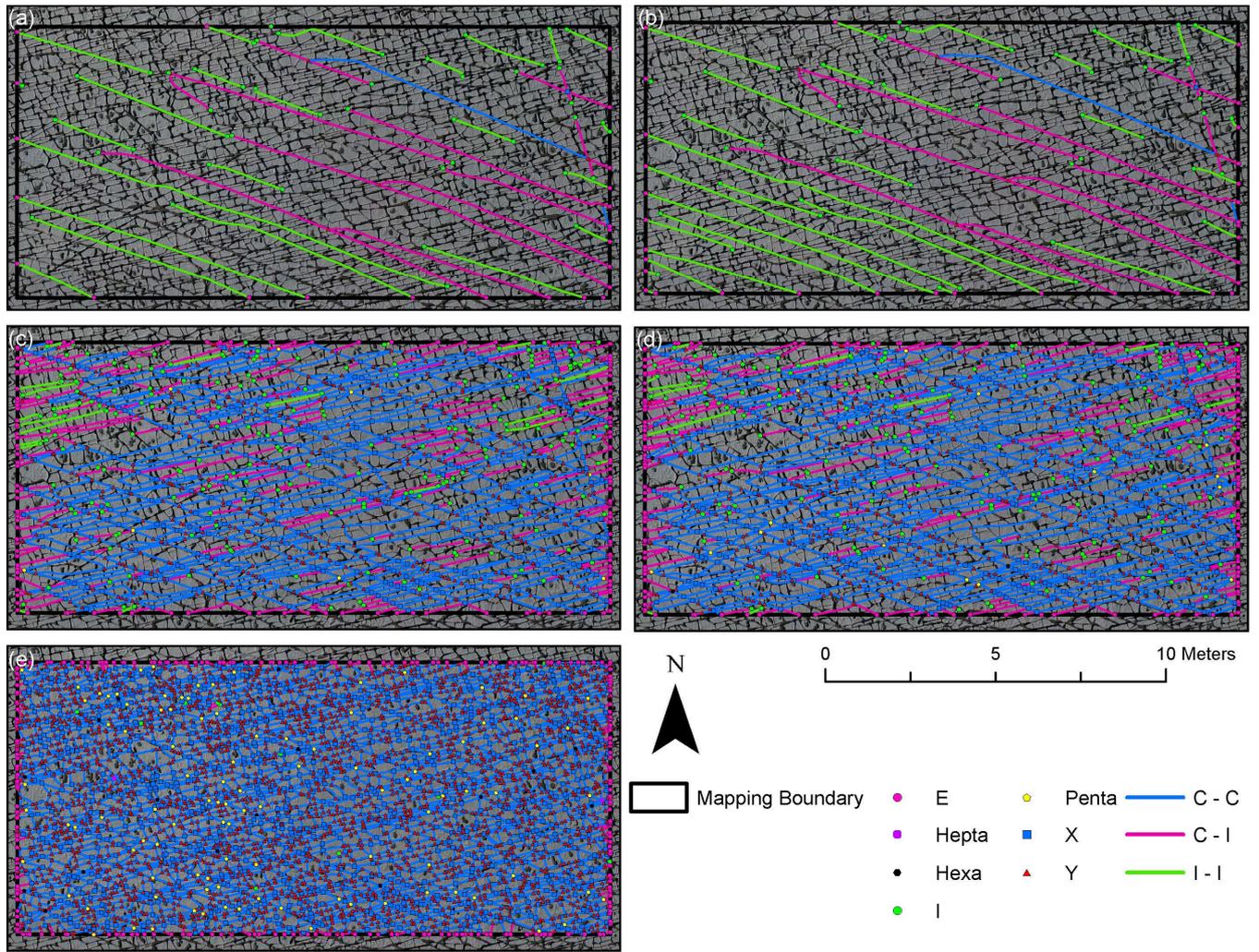
Fig. ~~15~~ 20: Development of the fracture network in time steps for SW3 from generation 1 to 5, visualized as branches and nodes. The consecutive fracture generation is added for each panel from (a) to (e).

2505

73
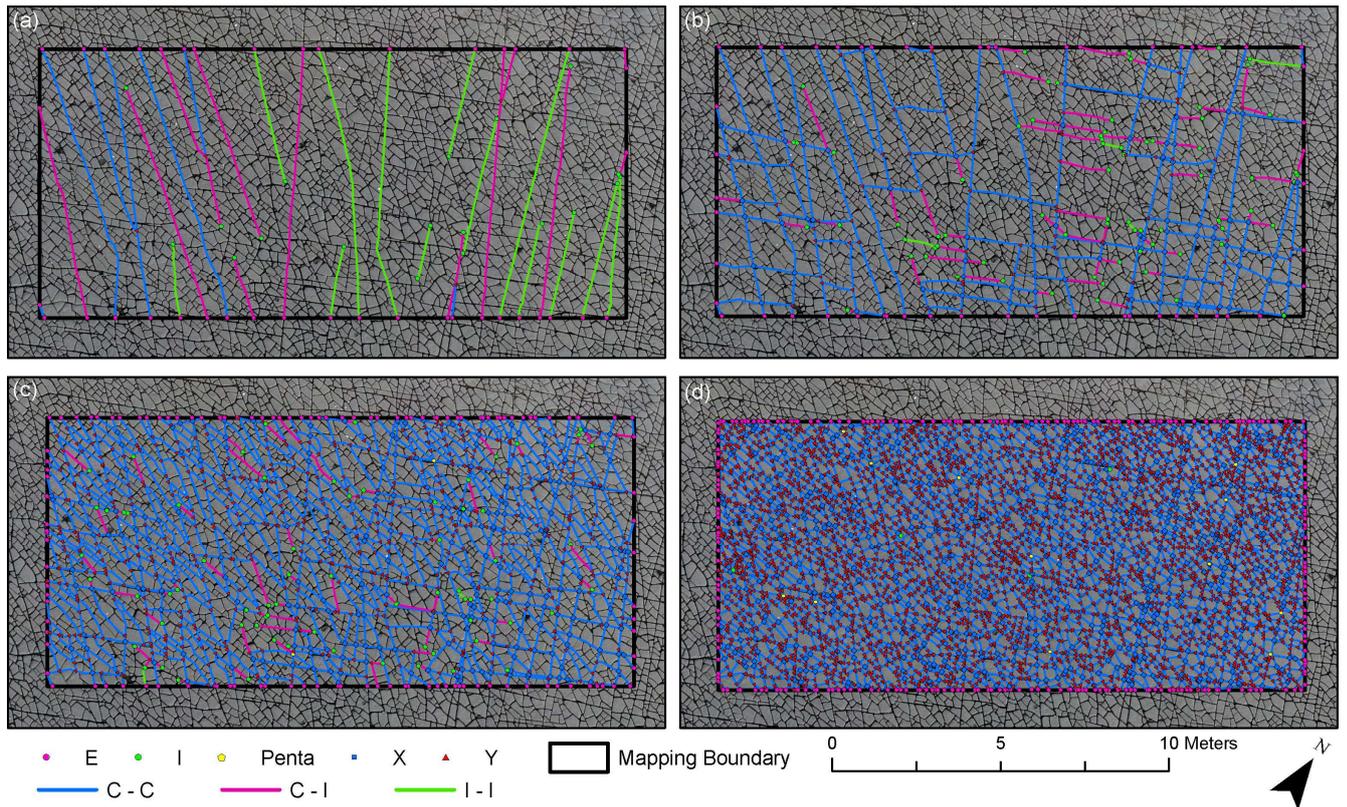
2510 **Fig. ~~16~~ 21: Development of the fracture network in time steps for NE1 from generation 1 to 5, visualized as branches and nodes. The consecutive fracture generation is added for each panel from (a) to (d). Note that no fractures belonging to gen. 3 were identified in this domain.**
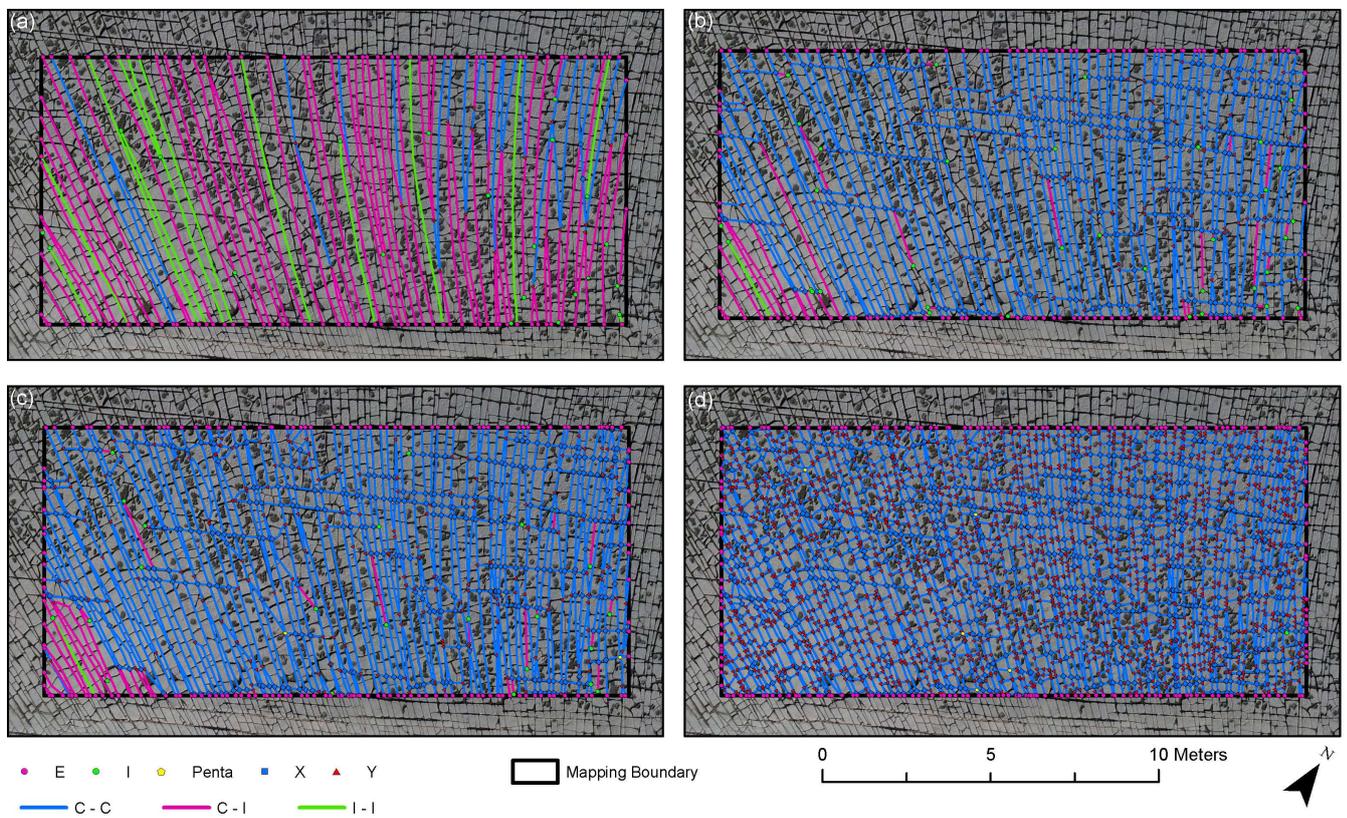
2515    **Fig. ~~17~~ 22: Development of the fracture network in time steps for NE2 from generation 1 to 5, visualized as branches and nodes. The consecutive fracture generation is added for each panel from (a) to (d). Note that no fractures belonging to gen. 3 were identified in this domain.**
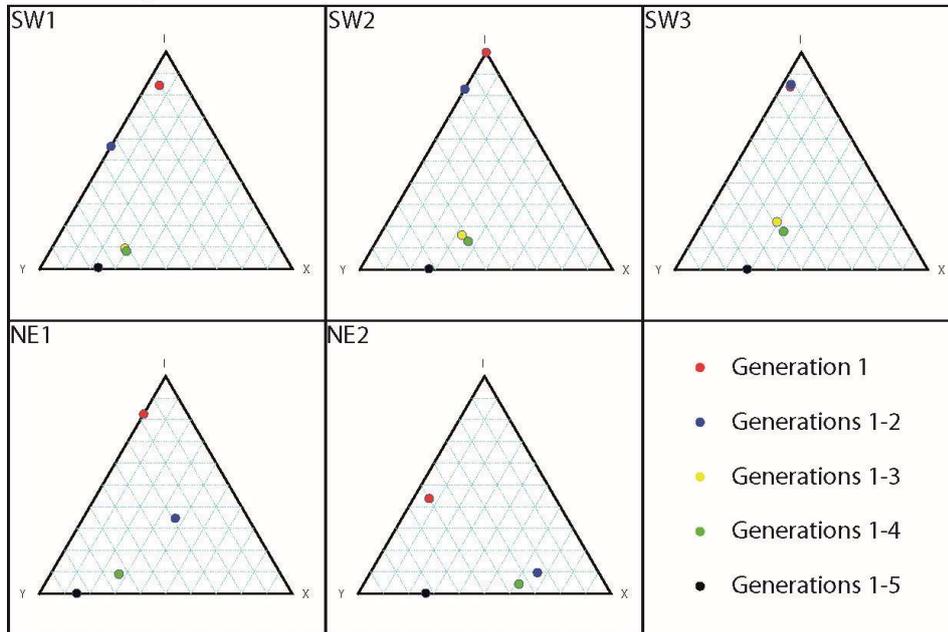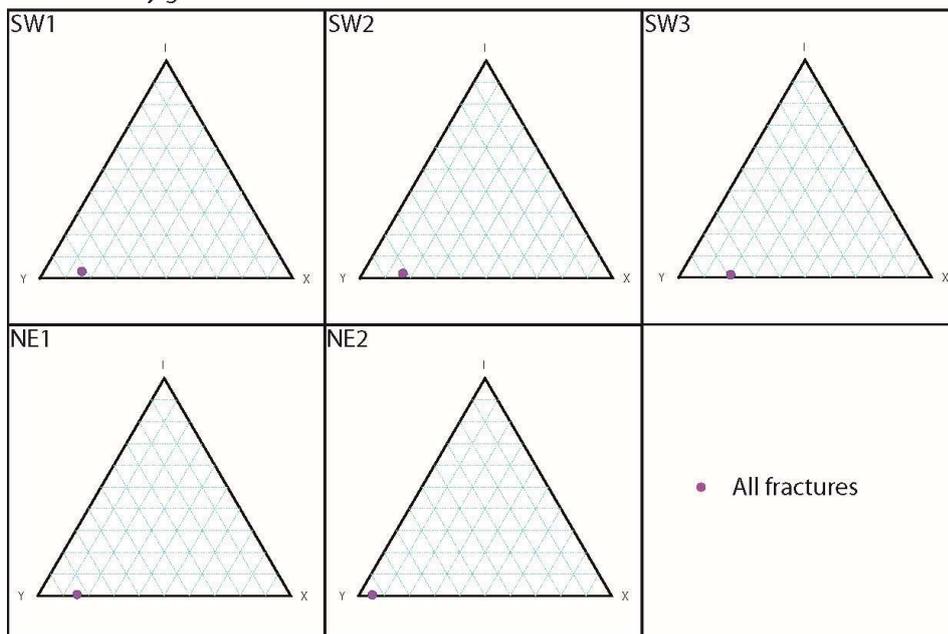


2520    ~~Fig. 18: Ternary plots visualizing the node distributions. For the manual interpretations, the different network stages in terms of fracture generations are plotted. The automatic networks only show the final stage of the network, which corresponds to generations 1 – 5 in the manual interpretation. Nodes with four and more than four intersecting branches have been binned as type X nodes.~~
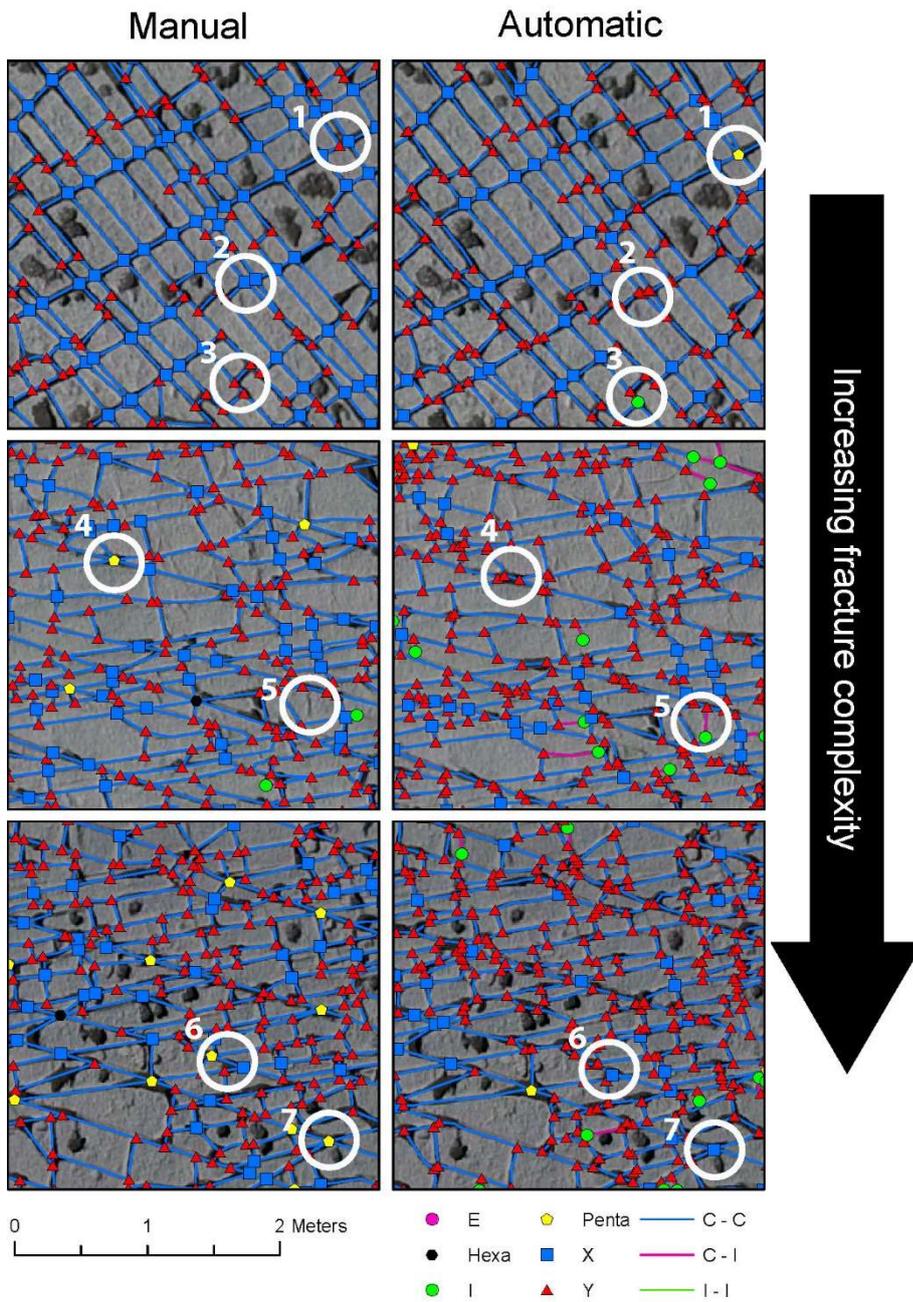
**Fig. ~~19~~ 23: Comparison of the resulting network analysis from manual (left) and automatic (right) mapping. The complexity of the fracture patterns increases top to bottom; therefore, the quality of the results is assumed to decrease. Marks 1 - 6 show differences of the manual and automatic interpretation. 1) Two closely spaced Y and X nodes (manual) or one Penta node (automatic); 2) X nodes (manual) or closely spaced Y nodes (automatic); 3) Abutting fracture (manual) or I node (automatic); 4) One Penta node (manual) or three Y nodes (automatic); 5) Interpretation as surface erosion or a small fracture; 6) Penta and X nodes (manual) or X and Y nodes (automatic); 7) Interpretation as two close fractures resulting in a Penta node (manual) or as one fracture resulting in a X node.**
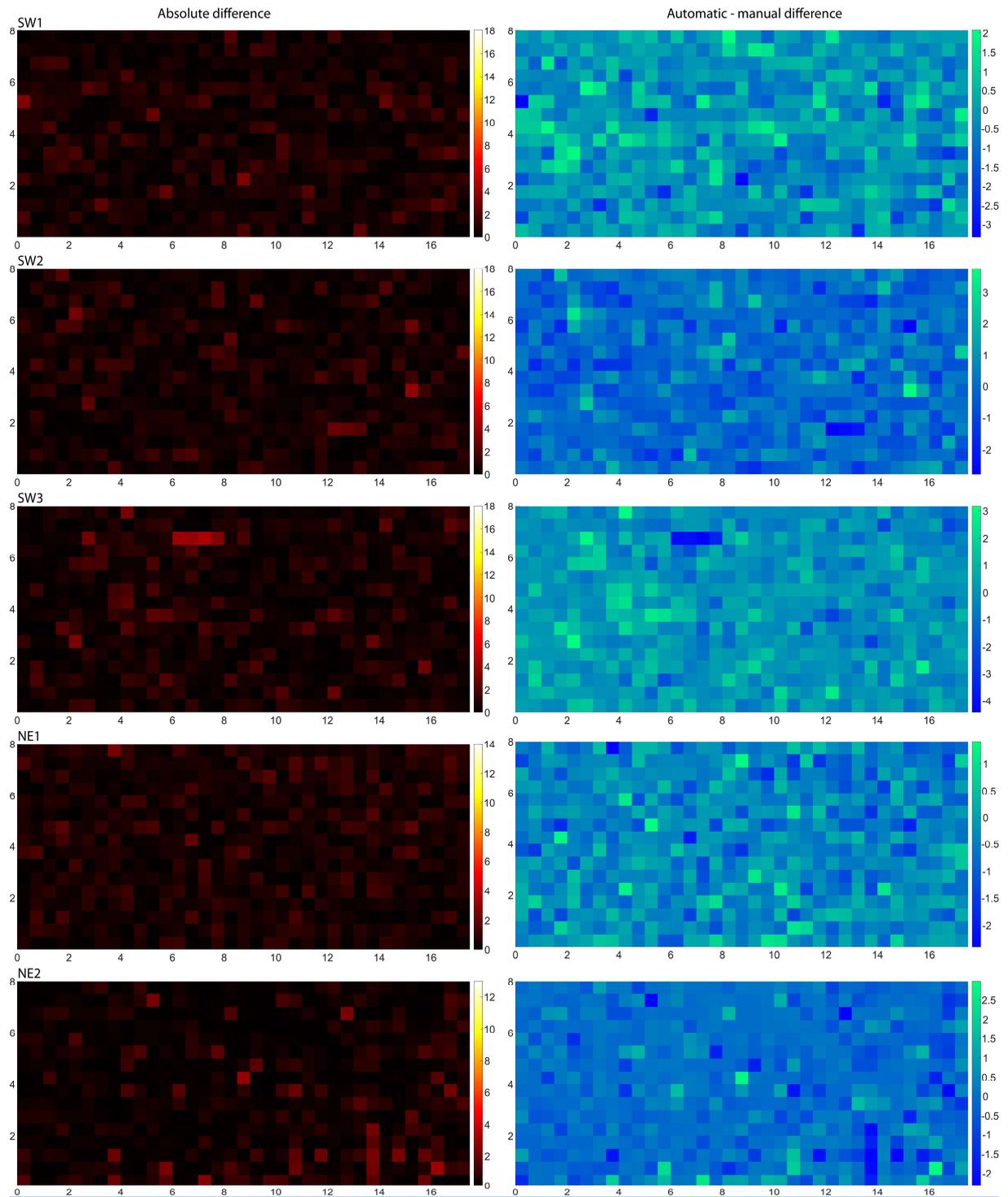
**Fig.: Differences of the P$_{21}$ fracture intensity between manually an automatically traced fractures for all five domains. Absolute differences are shown in the left column, relative differences in the right column.**

2545
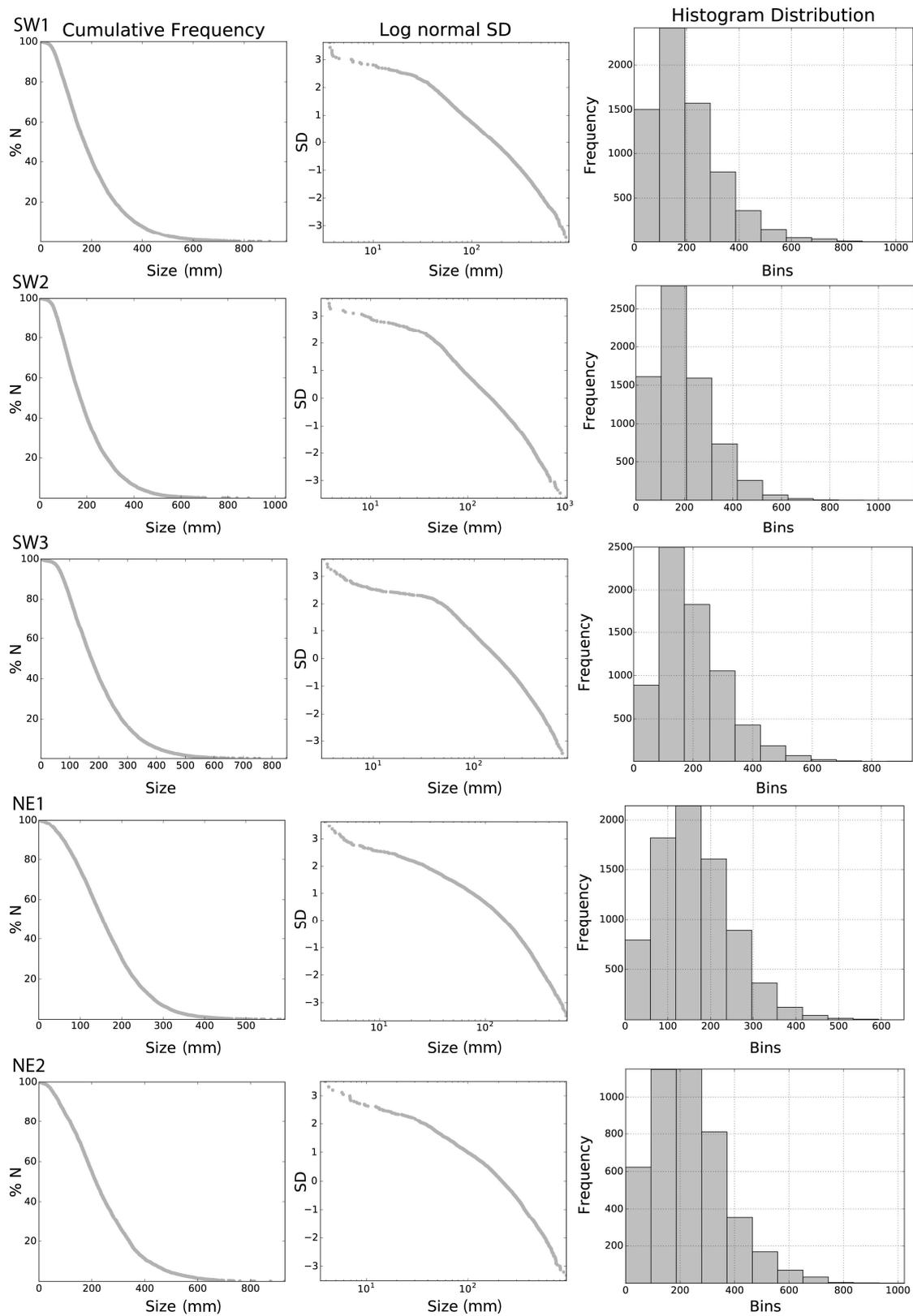
2550

78
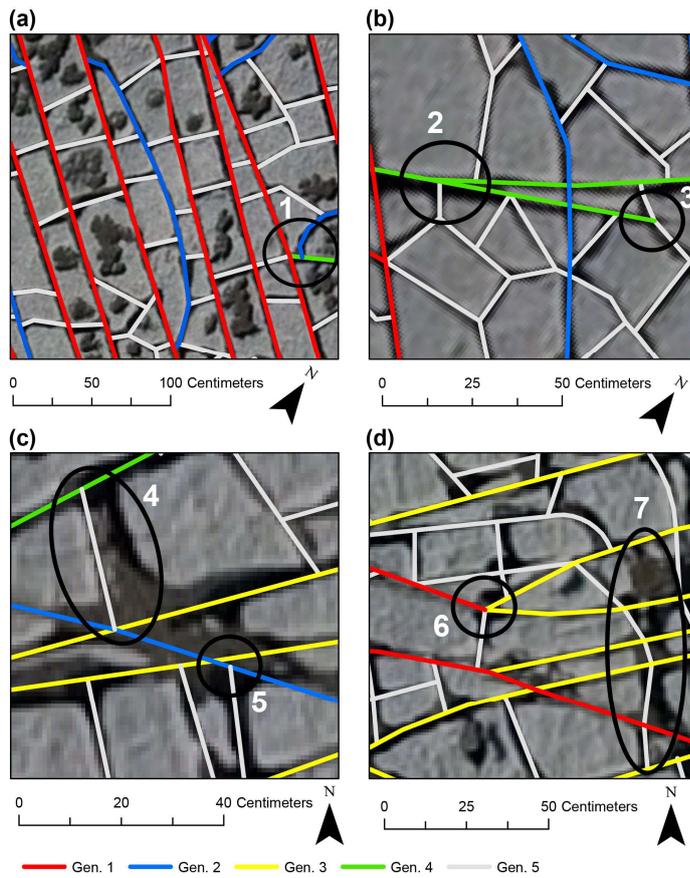
**Fig. ~~19.2~~24: Details of the manual interpretation of fracture generations showing examples that allow for different interpretations of the generation and network geometry. (a) The fracture in the center of the image was assigned to gen. 2 (blue) because it abuts twice at the same gen.1, fracture, which is atypical for fractures of gen. 2 and was decided by elimination of the other interpreted generations. Mark 1 shows an interpretation where gen.2 appears to abut on gen. 4, which then abuts on gen. 1 towards the left. This is an example in which the initial tracing of the network was unclear because of the widely eroded intersection of the fractures and has been revised during the assignment of generations. (b) Mark 2 shows a splaying gen. 4 fracture, which is a rare case and cold also be interpreted as a younger fracture abutting on gen. 4. Mark 3 shows the tip of gen.4 which apparently abuts on gen. 5. The small darker in front of the tip suggests, that the fracture may continue, or the surface of the rock has been eroded in that place or is wet, in which case the interpreter decided to end the trace. (c) Mark 4 shows a widely eroded fracture which has been traced along one of the edges instead along its median axis. The reasoning behind this decision is the assumption, that the left edge sharper and ~~more straight~~straighter than the one on the right, which suggests a stronger erosion in that part. Therefore, the left edge was interpreted as the best representation of the fracture geometry in this case. Mark 5 shows a gen. 5 fracture abutting on gen.2. However, the wide erosion of the fracture itself and the junction of gen. 2 and gen. 3 towards the left, also allow the interpretation as a triple junction of generations 2, 3 and 5. (d) Mark 6 shows a junction where the gen.1 fracture was interpreted to stop, while the trace continues as gen.3 which initially follows the rough strike direction of gen. 1 and then bends towards the strike direction of gen. 3. Mark 7 shows another strongly eroded area that hinders the interpretation of fractures.**

79

2585 **Table 1:** ~~P$_{21}$ analysis and statistics~~ Statistics and results of the P$_{ij}$ analyses of the automatically and manually-traced ~~of the~~ **fracture networks.**

| | Automatic | | | | | Manual | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | SW 1 | SW 2 | SW 3 | NE 1 | NE 2 | SW 1 | SW 2 | SW 3 | NE 1 | NE 2 |
| Branches | 8184 | 8548 | 8239 | 8366 | 4831 | 6888 | 7094 | 6975 | 7789 | 4366 |
| P$_{21}$ (m$^{-1}$) | 9.7 | 10.01 | 9.84 | 8.75 | 7.31 | 9.8 | 10 | 9.7 | 9 | 7.3 |
| P$_{20}$ (m$^{-2}$) | 58.5 | 61.1 | 58.9 | 59.8 | 34.1 | 49.1 | 50.6 | 49.8 | 55.6 | 30.9 |
| Mean (m) | 0.16 | 0.16 | 0.17 | 0.15 | 0.21 | 0.2 | 0.2 | 0.2 | 0.16 | 0.24 |
| Geom. mean (m) | 0.13 | 0.13 | 0.13 | 0.12 | 0.17 | 0.16 | 0.17 | 0.17 | 0.14 | 0.19 |
| Standard deviation | 0.12 | 0.11 | 0.1 | 0.08 | 0.13 | 0.13 | 0.12 | 0.11 | 0.86 | 0.14 |
| 25% | 0.08 | 0.08 | 0.09 | 0.08 | 0.11 | 0.1 | 0.11 | 0.11 | 0.1 | 0.14 |
| 50% | 0.14 | 0.14 | 0.14 | 0.14 | 0.19 | 0.17 | 0.17 | 0.17 | 0.15 | 0.22 |
| 75% | 0.22 | 0.22 | 0.23 | 0.2 | 0.29 | 0.26 | 0.26 | 0.26 | 0.22 | 0.32 |
| Min. | < 1 cm | < 1 cm | < 1 cm | < 1 cm | < 1 cm | < 1 cm | < 1 cm | < 1 cm | < 1 cm | < 1 cm |
| Max. (m) | 0.9 | 0.97 | 0.69 | 0.57 | 0.77 | 0.97 | 1.05 | 0.85 | 0.59 | 0.93 |
| Covariance | 0.7 | 0.67 | 0.62 | 0.56 | 0.62 | 0.64 | 0.59 | 0.56 | 0.53 | 0.57 |
| Skewness | 1.46 | 1.21 | 1.01 | 0.74 | 0.77 | 1.31 | 1.21 | 1.06 | 0.72 | 0.88 |
| Kurtosis | 2.99 | 1.82 | 1.07 | 0.55 | 0.48 | 2.33 | 2.03 | 1.42 | 0.72 | 1.01 |

| ~~-~~ | ~~SW 1~~ | ~~SW 2~~ | ~~SW 3~~ | ~~NE 1~~ | ~~NE 2~~ |
|---|---|---|---|---|---|
| ~~Branches~~ | ~~8184~~ | ~~8548~~ | ~~8239~~ | ~~8366~~ | ~~4831~~ |
| ~~P$_{21}$ (m$^{-1}$)~~ | ~~9.7~~ | ~~10.01~~ | ~~9.84~~ | ~~8.75~~ | ~~7.31~~ |
| ~~Mean (m)~~ | ~~0.16~~ | ~~0.16~~ | ~~0.17~~ | ~~0.15~~ | ~~0.21~~ |
| ~~Geom. mean (m)~~ | ~~0.13~~ | ~~0.13~~ | ~~0.13~~ | ~~0.12~~ | ~~0.17~~ |
| ~~Standard deviation~~ | ~~0.12~~ | ~~0.11~~ | ~~0.1~~ | ~~0.08~~ | ~~0.13~~ |
| ~~25%~~ | ~~0.08~~ | ~~0.08~~ | ~~0.09~~ | ~~0.08~~ | ~~0.11~~ |
| ~~50%~~ | ~~0.14~~ | ~~0.14~~ | ~~0.14~~ | ~~0.14~~ | ~~0.19~~ |
| ~~75%~~ | ~~0.22~~ | ~~0.22~~ | ~~0.23~~ | ~~0.2~~ | ~~0.29~~ |
| ~~Min.~~ | ~~< 1 cm~~ | ~~< 1 cm~~ | ~~< 1 cm~~ | ~~< 1 cm~~ | ~~< 1 cm~~ |
| ~~Max.~~ | ~~0.9~~ | ~~0.97~~ | ~~0.69~~ | ~~0.57~~ | ~~0.77~~ |
| ~~Covariance~~ | ~~0.7~~ | ~~0.67~~ | ~~0.62~~ | ~~0.56~~ | ~~0.62~~ |
| ~~Skewness~~ | ~~1.46~~ | ~~1.21~~ | ~~1.01~~ | ~~0.74~~ | ~~0.77~~ |
| ~~Kurtosis~~ | ~~2.99~~ | ~~1.82~~ | ~~1.07~~ | ~~0.55~~ | ~~0.48~~ |

2590

**Table 2: Nodes generated from the ~~automatically~~ automatically-traced fracture networks.**

|        | SW 1 | SW 2 | SW 3 | NE 1 | NE 2 |
|--------|------|------|------|------|------|
| E      | 248  | 262  | 268  | 236  | 279  |
| I      | 166  | 117  | 62   | 19   | 17   |
| Y      | 4250 | 4390 | 3979 | 5100 | 2517 |
| X      | 713  | 757  | 885  | 282  | 448  |
| Penta  | 62   | 87   | 103  | 2    | 8    |
| Hexa   | 3    | 9    | 17   |      |      |
| Hepta  |      | 2    |      |      |      |

2595

2600

**Table 3: Average (avg.) values of fracture length (m~~).~~) and strike (°)~~ and sinuosity~~ for the domains in the SW.**

| Gen. | Avg. length (m) | | | Avg. strike (°) | | | ~~Avg. sinuosity~~ | | |
|------|------|------|------|------|------|------|------|------|------|
|      | SW 1 | SW 2 | SW 3 | SW 1 | SW 2 | SW 3 | ~~SW 1~~ | ~~SW 2~~ | ~~SW 3~~ |
| 1 | 3.06 | 7.98 | 3.06 | 10~~5.78~~6 | 11~~1.94~~2 | 117.~~41~~ | ~~1.00~~ | ~~1.00~~ | ~~1.00~~ |
| 2 | 4.31 | 4.67 | 3.09 | 10~~0.71~~1 | 10~~4.99~~5 | 10~~8.58~~9 | ~~1.00~~ | ~~1.00~~ | ~~1.00~~ |
| 3 | 1.12 | 1.25 | 1.30 | 8~~2.83~~3 | 79.~~34~~ | 78.~~21~~ | ~~1.02~~ | ~~1.01~~ | ~~1.08~~ |
| 4 | 1.17 | 2.27 | 2.09 | ~~59.83~~60 | 60.~~05~~ | 59.~~41~~ | ~~1.00~~ | ~~1.00~~ | ~~1.00~~ |
| 5 | 0.22 | 0.24 | 0.24 | 11~~3.73~~4 | 118.~~11~~ | 12~~0.56~~1 | ~~1.02~~ | ~~1.02~~ | ~~1.03~~ |

2605

81

**Table 4: Average values of fracture length (m) ~~and.~~ strike (°) ~~and sinuosity~~ for the domains in the NE.**

| Gen. | Avg. length (mm) NE 1 | NE 2 | Avg. strike (°) NE 1 | NE 2 | ~~Avg. sinuosity NE 1~~ | ~~NE 2~~ |
|---|---|---|---|---|---|---|
| 1 | 4.33 | 4.13 | 14~~3.76~~4 | 13~~6.89~~7 | ~~1.00~~ | ~~1.00~~ |
| 2 | 1.64 | 1.62 | 111~~.33~~ | 12~~2.74~~3 | ~~1.00~~ | ~~1.00~~ |
| 4 | 1.77 | 0.77 | 64~~.17~~ | 58~~.07~~ | ~~1.01~~ | ~~1.01~~ |
| 5 | 0.28 | 0.19 | 7~~5.87~~6 | 55~~.33~~ | ~~1.01~~ | ~~1.01~~ |

2610

**Table 5: Evolution of the fracture network in SW1.**

| Gen. | Avg. branch length (m) | ~~Sinuosity~~ | I | Y | X | E | Penta | Hexa |
|---|---|---|---|---|---|---|---|---|
| 1 | 3.06 | ~~1.00~~ | 49 | 6 | 3 | 25 | 0 | 0 |
| 1 - 2 | 4.69 | ~~1.00~~ | 22 | 17 | 0 | 59 | 0 | 0 |
| 1 - 3 | 0.59 | ~~1.01~~ | 92 | 595 | 278 | 145 | 1 | 0 |
| 1 - 4 | 0.56 | ~~1.01~~ | 84 | 628 | 303 | 151 | 2 | 3 |
| 1 - 5 | 0.20 | ~~1.00~~ | 30 | 3194 | 849 | 250 | 96 | 9 |

2615

**Table 6: Evolution of the fracture network in SW2.**

| Gen~~.~~ | Avg. branch length (m) | ~~Sinuosity~~ | I | Y | X | E | Penta | Hexa | Hepta | Okta |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 7.98 | ~~1.00~~ | 2 | 0 | 0 | 6 | | | | |
| 1 - 2 | 4.51 | ~~1.00~~ | 49 | 10 | 0 | 49 | | | | |
| 1 - 3 | 0.54 | ~~1.00~~ | 178 | 577 | 357 | 163 | 6 | | | |
| 1 - 4 | 0.50 | ~~1.00~~ | 157 | 612 | 425 | 173 | 12 | 4 | | |
| 1 - 5 | 0.20 | ~~1.00~~ | 187 | 3057 | 1000 | 258 | 123 | 20 | 2 | 1 |

2620

**Table 7: Evolution of the fracture network in SW3.**

| Gen. | Avg. branch length (m) | Sinuosity | I | Y | X | E | Penta | Hexa | Hepta |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 3.06 | ~~1.00~~ | 48 | 7 | 2 | 25 | 0 | 0 | 0 |
| 1 - 2 | 3.17 | ~~1.00~~ | 52 | 7 | 2 | 34 | 0 | 0 | 0 |
| 1 - 3 | 0.59 | ~~1.01~~ | 203 | 448 | 262 | 145 | 8 | 0 | 0 |
| 1 - 4 | 0.54 | ~~1.00~~ | 179 | 492 | 329 | 158 | 15 | 4 | 0 |
| 1 - 5 | 0.20 | ~~1.00~~ | 12 | 2953 | 1057 | 275 | 102 | 16 | 2 |

2625

**Table 8: Evolution of the fracture network in NE1.**

| Gen. | Avg. branch length (m) | Sinuosity | I | Y | X | E | Penta |
|---|---|---|---|---|---|---|---|
| 1 | 4.33 | ~~1.00~~ | 19 | 4 | | 41 | |
| 1 - 2 | 0.79 | ~~1.00~~ | 78 | 65 | 82 | 54 | |
| 1 - 4 | 0.68 | ~~1.00~~ | 77 | 547 | 229 | 124 | 2 |
| 1 - 5 | 0.24 | ~~1.01~~ | 8 | 4135 | 705 | 283 | 12 |

2630

**Table 9: Evolution of the fracture network in NE2.**

| Gen. | Avg. branch length [m] | Sinuosity | I | Y | X | E | Penta |
|---|---|---|---|---|---|---|---|
| 1 | 4.13 | ~~1.00~~ | 21 | 24 | 3 | 131 | |
| 1 - 2 | 0.80 | ~~1.00~~ | 38 | 95 | 259 | 145 | |
| 1 - 4 | 0.45 | ~~1.01~~ | 25 | 193 | 347 | 184 | 2 |
| 1 - 5 | 0.16 | ~~1.00~~ | 7 | 1888 | 681 | 230 | 5 |

2635

2640

**Table 10: Differences of the fracture networks from manual and automatic tracing.**

|  | SW 1 | SW 2 | SW 3 | NE 1 | NE 2 |
|---|---|---|---|---|---|
| Avg. branch length (m) | 0.03 | 0.03 | 0.03 | 0.09 | -0.05 |
| I | -136 | -99 | -50 | -11 | -10 |
| Y | -1056 | -1333 | -1026 | -965 | -629 |
| X | 136 | 243 | 172 | 423 | 233 |
| Penta | 34 | 36 | -1 | 10 | -3 |
| Hexa | 6 | 11 | -1 | 0 | 0 |
| Hepta | 0 | 0 | 2 | 0 | 0 |
| Okta | 0 | 1 | 0 | 0 | 0 |

2645