Author Response to RC1 (Billy Andrews)

Major comments responses

[1] The authors provide an extensive and interesting suite of student interpretations, however, a primary concern whilst reviewing the manuscript was the framework and analysis used to describe the results. My concerns can be split into two: (a) the statistical analyse the data and (b) the language and framework used when describing these results.

[1a] Throughout the manuscript the authors use a combination of mean, standard deviation, median and IQR, however it is not always clear how the results are distributed. Standard deviation is often used as a measure of uncertainty; however, I have two potential issues with this. The first being that in some cases the results don't appear to be normally distributed (e.g. interpretation density in Fig 9). In the case of a skewed distribution standard deviation should be avoided, and I advise the use of Inter Quartile Range (IQR) instead. It would also be useful to see what the maximum uncertainty, and not just the IQR or standard deviation, with minimum/maximum values reported. My second issues with the use of standard deviation is that a specific standard deviation value could represent either a high or low value of uncertainty depending on the mean at that point. For example, in your fault throw data for F3 (Fig 7b) the standard deviation remains ~30 in the north of the profile. The median throws in this section however ranges from ~50 to ~120, meaning the level on uncertainty is considerably higher at a smaller throw. I suggest giving the reader more confidence, and to better understand the risk in different areas that the following should be considered when analysing the results: i. The shape of the distribution? Is the distribution the same along the whole profile? ii. If the distribution is not normal, then consider using median and IQR to describe the results. iii. The normalisation of results would aid the comparison of uncertainty; this should either be through the use of coefficient of variance, or in the case of non-normally distributed data then a quartile-based coefficient of variance (IQR/median).

AUTHORS COMMENT: Thank you for the suggestions! They enabled us to make the manuscript more robust with regards to the statistical description of our findings! We agree with the reviewer in that the use of IQR instead of standard deviation is a better way to analyse the fault throw uncertainty.

CHANGE IN MANUSCRIPT: We have redrafted Figure 7 (now 6) as boxplots of the fault throw uncertainty data and have adapted our description of the findings accordingly.

[1b] Throughout the manuscript there are several instances where the authors use subjective language when describing the results (e.g. 'significantly reduced', 'observe large', 'roughly correspond'). The description of the results would be improved through the quantification, or further description, of what the authors mean. It would also give readers more confidence in the results. In addition to this there are occurrences where statements such as 'a few students' are used (e.g. Page 7 Line 7). In these cases, the authors should be explicit in what a few means '7 students'.

AUTHORS COMMENT: Thank you for pointing this out! We have adapted the manuscript to remove the use of subjective language where possible. We have improved the statistical rigor in our description of results.

CHANGE IN MANUSCRIPT: A multitude of changes throughout the manuscript. Please see the revised manuscript or the change-tracked manuscript.

[2] Discussion section 'key findings': This paper provides an impressive analysis of a 1st pass data set aiming to quantify uncertainty in 3-D seismic interpretation, however, I felt the authors needed further elaboration on the reasons behind the presented results. The manuscript would be improved through reducing the key findings to a set of points (potentially even using bullet points), and adding a section which investigates the factors behind the presented uncertainty. From your results this appears to be split into two 'themes':

AUTHORS COMMENT: We have structured our discussion into four parts to first present to the reader the key findings of our study, and then to address the implications on what we believe to be the most important applications of seismic interpretation. We have added a summary section though, to provide the reader with a clear take-home message following our discussion.

CHANGE IN MANUSCRIPT: We have done various changes throughout our discussion, please view the revised or change-tracked manuscript.

[2a] Human factors: Several sections allude to how students build a mental model during interpretation (e.g. taking information from outside areas of degraded seismic image quality to inform decisions). It is somewhat lacking that the current literature in this area is missing, and that this aspect remains unexplored. Do those who use outside trends to inform areas of poor coverage end up with a better interpretation? Another important comment you raise can be found on Page 13 line 5-7 where you allude to the order students undertook the analysis. This is another important aspect to consider and should be explored further, possibly referencing other work which shows people can vary interpretations through time (e.g. Scheiber et al., 2015 for lineament extraction). The role of human factors on the collection of 3D seismic data should be further explored.

AUTHORS COMMENT: We have no real way of telling which students used outside trends to inform their interpretation in areas of poor coverage. We also have no information about the exact order students interpreted specific faults or fault blocks, as this would have required capturing and analysing their screen throughout the entire interpretation process. We therefore try to abstain from speculation about psychological effects for which we have no evidence for the dataset. We think these factors are vitally important to understanding all sources of interpretation uncertainty, but our manuscript is concerned with quantifying the effect size of this uncertainty and connect it to seismic data quality.

CHANGE IN MANUSCRIPT: -

[2b] Technical factors: These are highlighted strongly in that uncertainty is higher in areas of lower seismic resolution. Seismic resolution will always be lower surrounding fault, due to the increased amount minor structures and local deformation, and such we can expect uncertainty to remain high in such areas. We know from other aspects of fault science that 'intersection zones' or larger offset faults tend to have a wider zone of damage, and hence zone of reduced seismic image quality. Can we use some of this information to aid the assignment of risk in these areas?

I was often left asking 'why is this the case?' and the answers weren't forthcoming in the discussion. Although I have provided examples of what I feel should be expanded upon, there where sections which could also be expanded and linked to published literature. If this section is added, some of the implication sections could be slightly scaled back in particular section 4.4 which I feel has the least direct link to your results.

AUTHORS COMMENT: The assignment of risk or uncertainty based on fault offset is not straight forward in our opinion. Our study results have shown, that the seismic data quality around e.g. the

southern part of Fault 3 is low in combination with low throw. Seismic data quality is affected by so many aspects from the sedimentology and structural geology of the subsurface imaged to the imaging itself and the processing. Our study attempts to bypass exactly this to allow for estimation of uncertainties by using seismic data quality itself as a proxy for it (Fig. 8). Thus we believe the in-depth discussion of the multitude of impacting factors on seismic data quality out of scope for this manuscript, especially because it is very challenging research to identify the amount of impact every of the aforementioned aspects have.

CHANGE IN MANUSCRIPT: We have added references regarding the impact of fault zones on seismic data quality for readers to follow.

[3] Conclusions: I felt this manuscript really lacked was a clear finale. The authors present an extensive set of results, which have clear implications to the interpretation of 3D seismic, however, in my opinion fail to leave the reader with a clear take home message. This point links to the previous regarding the discussion and believe the discussion should be slightly restructured as above and a set of conclusions included which pulls together the findings, highlights the clear importance of these results (including beyond 3D seismic interpretation e.g. modelling from other sources) and raise future research directions. This would tie an important contribution together and provide readers with a clear take home message.

AUTHORS COMMENT: We have added a summary/conclusion section to the end of the paper to provide the readers with a clear tame home message

CHANGE IN MANUSCRIPT: Please see the revised manuscript for the Summary section.

[4] Figure quality/readability: Some of this may be due to the uploaded PDF, however, I found several figures difficult to read and often containing areas where text was too small. Some examples include line-weights of sub-sections, text size of longitude/ latitude and labels within panels and occasionally the chosen color scheme used was difficult to read either on the screen or when printed off. Detailed points are raised in specific comments.

AUTHORS COMMENT: Thank you for pointing this out. We have adapted the figures to improve text sizes and legibility of colours and line weights where applicable.

CHANGE IN MANUSCRIPT: Figure changes throughout the manuscript. Please see the revised manuscript.

Detailed comments responses

1 - Introduction

[1 (17-21)]: The introductory paragraph of the MS should be expanded to further define uncertainty. Conceptual uncertainty is first stated to be important in the 2nd paragraph, however, non-specialised readers would benefit from an explicit introduction to the different types of uncertainty (e.g. Bond, 2015; Tannert et al., 2007) and potentially how this effects the mental model of the interpreter.

AUTHORS COMMENT: We agree that a more specific description of the subjective uncertainty of interpretation and how it contrasts with objective measurement uncertainty would increase clarity of the manuscript.

CHANGE IN MANUSCRIPT: Our work is thus concerned with quantifying the scope of uncertainties in seismic interpretation, which represents inevitably biased, human judgment under uncertainty

(Tversky and Kahneman, 1974). This "subjective" uncertainty is in contrast to more "objective" uncertainty (Tannert et al., 2007; Bond, 2015) related to the geophysical acquisition of the data itself.

2.1 – Gullfaks geology and seismic data

2 (31-33): I suggest you make it clear that the study focuses on an interpretation boundary of a student exercise here, it currently sounds like it focuses on an area of a larger data set.

AUTHORS COMMENT: We agree, and improved clarity of what area is part of the study both in text and in Figure 1.

CHANGE IN MANUSCRIPT: Our study focuses on the structurally simpler western part of the domino system, where we investigate the uncertainty of the three faults F1-F3 and the fault blocks A-E depicted in Figure 1b.

3 (Fig1): (a) Colour and line weight for section line and interpretation box is unclear, both in colour and in B&W. I advise a change in colour and that the line weight is increased. The text size in the insert to this panel is far too small, as is the longitude and latitude numbers. The addition of a scale bar to this panel would also aid the reader. (b) A scale should be added to this panel. (c) The formation names are poor quality in the uploaded PDF, and also slightly on the small side.

AUTHORS COMMENT: We have adapted the color and line weights of annotations in Figure 1a to improve legibility.

CHANGE IN MANUSCRIPT: Figure changes.

3 (2-3): Many questions come to mind with respect to the level of experience of the interpreters and in part the limitations of your dataset, which includes undergraduate students only. Some of these include: Did everyone have the same level of training? What was the 'specialisms' in the sample set (i.e. how much seismic interpretation, structural geology, stratigraphy etc. was covered and was this equal in the students)? Also how long was spent by each student (If you have this data it would be interesting to see if those who spent more time interpreted differently to those who did not)? How comfortable where the students with using petrel & integrating well and seismic data?

AUTHORS COMMENT: The students all went through the undergraduate studies at Aberdeen, and this was their introduction course to 3-D seismic interpretation using Petrel. We do not have access to the time spent by each student, as they did significant amounts of the interpretation outside formal class times. We agree that this would be an interesting variable to analyse though! We have clarified the student's level of knowledge in seismic interpretation and Petrel.

CHANGE IN MANUSCRIPT: While the students had prior training in structural geology and interpretation of 2-D seismic data, this was the student's first hands-on course in 3-D seismic interpretation using the Petrel software as part of their undergraduate program.

3 (5): How much assistance was given in this? What was the variability in the interpreted horizon when assistance was given and how does this compare to the Top Ness. Can the difference between the Top Cretaceous and Base Cretaceous/Top Ness horizons show the effect of training in reducing uncertainty? Also, if there is little variability in the Top Cretaceous, due to the supervision, will this not effectively 'pin' one end of the fault sticks to a lower range of displacements, effetely adding to the increase in U/C with depth attributed to a degradation of seismic image quality (I agree image quality decreasing with depth will also be a factor).

AUTHORS COMMENT: Comparing the difference in interpretations between Top Cretaceous (TC) and both the Base Cretaceous Unconformity (BCU) and Top Ness horizon (TN) to assess the effect of training or supervision on interpretation uncertainty is unlikely to give clear information: The TC horizon is very easy to pick due to its lack of significant deformation and we believe it would not provide a valid comparison with the strongly deformed TN. The faults investigated do not reach above the BCU and thus do not affect the TC.

CHANGE IN MANUSCRIPT: -

3 (7-8): Was there any difference in interpretation from students who used these different methods? How often was seeded tracking or manual interpretation used?

AUTHORS COMMENT: We have no information about the specific interpretation tools used by students at any given time. It would make an interesting study though to assess bias introduced by software tools!

CHANGE IN MANUSCRIPT: -

4 (Fig2): Increase the text size on the axis for clarity.

AUTHORS COMMENT: We have increased the text size for axis labels / ticks.

CHANGE IN MANUSCRIPT: Changes in Figure 2.

4 (2-3): Suggest the text about 90 interpretations be removed as is only mentioned here and does not seem required.

AUTHORS COMMENT: While the information is not further discussed in the manuscript, we believe it relevant to mention any (subjective) prior filtering we have applied to the dataset.

CHANGE IN MANUSCRIPT: -

2.3 – Data analysis

(**10**): This is an impressive data set; however, I would be interested to see how this is spread between the students. I suspect, and you allude to on page 4 line 15, that the number of fault sticks interpreted varies extensively between students, and that this is an important aspect of uncertainty. This could also then be further analysed to see if there is a correlation between number of fault sticks and level of uncertainty.

AUTHORS COMMENT: Thank you for pointing this out! We conducted a Bayesian estimation of the differences for the overall standard deviation of fault stick placement between the students falling into the categories of above and below 50% of fault stick interpretation frequency. We have added this to the results section.

CHANGE IN MANUSCRIPT: Analysis of the effect of fault stick interpretation frequency between the students with below-median and above-median fault stick interpretation frequency on overall fault standard deviation was analysed using Bayesian estimation (Kruschke, 2013). We observed a difference in mean standard deviation of $35.8~m$, $20.2~m$ and $81.5~m$ for Fault 1, 2 and 3 respectively, with probability of differences being larger than zero being $99.3~\%$, $87.4~\%$ and $99.9~\%$, making the differences for Fault 1 and 3 statistically credible.

4 (15): How is interpretation density defined?

AUTHORS COMMENT: We have removed the possibly confusing wording of interpretation density (number of interpretations per volume) and just refer to it as frequency.

CHANGE IN MANUSCRIPT: In the following analysis we present 2-D and slices of 3-D histograms of fault interpretations, showing interpretation frequency across the domain.

3 - Results

5 (Fig3): I wonder how Fault 1 and Fault 2 are defined in the northern interpretation bin once they are merged.

AUTHORS COMMENT: This is dependent on the student interpretation. If they interpreted Fault 1 as continuing or Fault 2, respectively.

CHANGE IN MANUSCRIPT: -

5 (13): I worry that this is affected not only on the placement, but also on how many fault-sticks each student included. In areas of relatively certain offsets, which will likely be increased by the image quality, I would imagine more sticks will be chosen, thus increasing the apparent 'certainty' of the result.

AUTHORS COMMENT: Indeed, as this is a basic histogram of the fault sticks, high frequency interpretations can affect the plot. But these differences in frequency are along strike and should have a negligible effect on the fault placement uncertainty patterns observed orthogonal to strike.

CHANGE IN MANUSCRIPT: -

6 (Fig4): I find the addition of the mean fault plane & k-values from Fossen and Hesthammer (1998) confusing as is, however, it is an important point which you make on Pg 11 ln 31-32. It would be made clearer to the reader if this mismatch was raised in the results, and later discussed in the 'Key findings'. A reminder that stereonet plots go from N to S actually on the figure and not just in the figure caption would also be helpful in this figure.

AUTHORS COMMENT: We have added an explanation of the Fossen & Hesthammer (1998) to the main text in the results section to improve clarity. We have also merged Figure 3 and 4 into 3, which will enable the reader to identify more easily which Stereonet plot belongs into which bin.

CHANGE IN MANUSCRIPT: We have added Bingham mean poles from \citet{fossen_structural_1998} for all three faults in the plot (light blue) for comparison.

6 (6-7): I am struggling to pull three clusters out of the stereonet data presented in fig 4a, and instead can only see two. I agree the data should be split into three due to the sinusoidal shape based in the geographic location, however, this information is instead better portrayed in Fig 3a. I advise you reword accordingly.

AUTHORS COMMENT: The clusters correspond to the three bins. You can see the clusters when separated into the three separate Stereonet plots. This should be more clear to the reader now that Figure 3 and 4 were merged. We have updated the figure reference accordingly.

CHANGE IN MANUSCRIPT: Fault 1 shows three distinct clusters of orientations (Fig. 3A, a-d), [...]

7 (Fig5): I would like to know the skewness of the distributions, particularly if this changes down dip, this will impact how valid the use of standard deviation is (See major comment 1). I also wonder which fault show the most variability with depth and why. Comparing using either a coefficient of

variance (if distributions are normal) or quartile-based coefficient of variance could pull out more trends between the faults.

Also, although standard deviation increases with depth, how well the data fits the regression line seems to decrease, particularly for F1. For F2 and F3, and to some extent F1 there seems specific horizons which show increased/decreased spread which is not in agreement with the linear regression. Is there an underlying control here? (e.g. stratigraphic layer with good/poor seismic response?). Visually I would consider changing the 'picks above BCU. . .' from light grey as it is difficult to see, the regression lines for F2 and F3 are also unclear when viewed on the screen (fine when printed).

AUTHORS COMMENT: Median skewness of the distributions show approximate symmetry to moderate skewness of the distributions, but with significant standard deviation (see Table 1). We concluded from that, that the use of standard deviation is adequate for the purpose of this paper. We use here a basic categorization of -0.5 to 0.5 as approximately symmetric and ±0.5 to ±1.0 as moderately skewed. Overall, the variation of skewness with depth is quite high. We are aware that collapsing all the information into a single scatterplot is a significant simplification of the data, but we think this to be adequate for the first-look scope of the paper.

The quality of the seismic response of horizons is a major controlling factor of interpretation uncertainty and will certainly have a strong impact on uncertainty. As we have shown in Figure 9, strong bounding seismic reflectors significantly reduce interpretation uncertainty of faults. In this plot we collapse the entire fault uncertainty along the y-axis of the domain, thus averaging across a large area which makes the correlation with specific physical features challenging, and we did not pursue for this manuscript.

Table 1: Median skewness for fault position with depth.

	Fault 1	**Fault 2a**	**Fault 2b**	**Fault 3**
Median Skewness	-0.06 ± 0.36	0.25 ± 0.36	-0.20 ± 0.33	0.21 ± 0.27

CHANGE IN MANUSCRIPT: Changed grey Figure 5 text to black for better legibility.

7 (7): How many students did this? This is a source of error/uncertainty and I feel it should not be dismissed. What training/geological information was provided to the students and from this should they have factored in the 'geological unreasonableness' of the interpretations?

AUTHORS COMMENT: We did not count the number of interpretations that do so, but removed any fault stick point above the mean BCU from the analysis automatically. But we believe that most students simply just did not terminate their fault interpretations correctly, as no geological model was built afterwards. We do not believe that most students who did this actually interpreted what they thought were meaningful faults above the BCU conceptually, but rather due to lack of attention to this detail, as this was not a focus of their coursework.

CHANGE IN MANUSCRIPT: -

7 (12): I question why probability is quoted here, you have 78 interpretations, so feel that the numbers represent the total number of students who interpreted that network.

AUTHORS COMMENT: Used student counts for clarity.

CHANGE IN MANUSCRIPT: Five modes of FN topology make up the bulk of fault network topologies, while others were only interpreted by 3 or less students respectively.

7(14): I feel this needs to be linked back to interpretation and not to 'probable'. Probable suggests that if 100 random people where to be selected then X% would choose option Y, which I think is misleading as there are more human factors involved here. I also feel it is prudent to describe in the MS the level of exposure students has with 'complex' fault topologies.

AUTHORS COMMENT: Changed probable to frequent for clarity.

CHANGE IN MANUSCRIPT: Note that the most frequent FN (Fig. 6b, A) is different from the reference expert FN interpretation …

8 (Fig. 6): In part (a) I would advise that the y-axis is changed to # of students and not a percentage (see comment #). In part (b) I wonder how statistically different A & C are in the students data? Is there a distinct gap? (as topologically they are the same, and geometrically similar).

AUTHORS COMMENT: Added student numbers to the probability mass plot for clarity. We also enhanced the clarity of the related schematic fault networks.

CHANGE IN MANUSCRIPT: Figure 6 changed.

8 (6): How do you define 'relatively constant' uncertainty? How is it measured? See major comment 1.

AUTHORS COMMENT: We are referring to no apparent changes in the uncertainty pattern of the fault throw along strike. We clarified that in the manuscript.

CHANGE IN MANUSCRIPT: Fault throw uncertainty shows no apparent pattern change along the strike (around $\pm 30~m$), while rising sharply to about $\pm75~m$ at the southern edge of the seismic cube.

9 (Fig 7): This figure makes a very important point, that uncertainty can vary spatially, however, a number of questions are raised in how the results are presented. My main concern is the use of median and standard deviation (Again see Major comment 1). Why is median used? If it is because the distribution is skewed, which I suspect it is, then it is not statistically robust to use standard deviation. I would also like to see the min and max values here (aka what is the maximum risk in this data set?). I suggest redrafting to either show standard deviation surrounding the mean, with min and max values displayed, or to show the IQR around the median again with min/max values. I prefer the second method and suspect similar trends would be observed.

Visually I would consider increasing the text size of the annotations. Is standard deviation in any way related to throw? A +/- 30 meters on a 120 m offset fault is much better (25%) than on a 50 m offset fault (60%) Is quoting exact values the best way to compare uncertainty?

AUTHORS COMMENT: Thank you for pointing this out! We have adapted the plots to boxplots to show median fault throw with IQR, minima and maxima, as well as outliers to improve the robustness of our analysis. Overall the same trends visualize, but the difference in fault throw uncertainty at the merge of Fault 1 and 2 are much more visible. We have modified the results description accordingly.

We have increased font sizes in the figures.

CHANGE IN MANUSCRIPT: Results of the fault throw analysis are plotted in Fig. \ref{fig:07}. The boxplots show median fault throw with the associated interquartile range (IQR), extrema and outlier values along fault strike direction. The throw profile of Fault 1 (Fig. \ref{fig:07}a) shows a distinct sinuous shape spatially associated with its interaction with Fault 2. This shows one bin with high median fault throw of approx. $180~m$ and high fault throw uncertainty before strongly decreasing in fault throw values down to a median of about $40~m$ and one of the lowest IQR along the fault. Median fault throw then rises steadily towards the South while increasing in uncertainty. Notice the increase of uncertainty at both ends of the dataset, with increasing median throw in the South and decreasing in the North. The throw profile of Fault 3 (Fig. \ref{fig:07}b) shows two distinct levels of throw: In the Northern part of the fault median throw values are high at around 90 to $105~m$, associated with comparatively lower uncertainty than in the South. Towards the South, median fault throw decreases down to about $40~m$ while the IQR values increase, being largest at the Southern edge of the dataset.

10 (9): How many students interpreted the fault further to the East?

AUTHORS COMMENT: Three students interpreted Fault 3 further to the East.

CHANGE IN MANUSCRIPT: -

10 (11) to 11 (7): This section suffers from a lack of statistical analysis, a framework to describe these results would increase the rigor of this section. The data shows some very important trends, probably the most important point of the manuscript, and with a more robust statistical analysis the reader would have more confidence in the results and following discussion.

AUTHORS COMMENT: We have increased the rigor in our statistical description of the results to increase clarity. Thank you for pointing this out!

CHANGE IN MANUSCRIPT: Several changes throughout the paragraph – please see Section 3.3 in the revised manuscript or the change-tracked manuscript.

4 – Discussion

11 (10-18): You open this paragraph with a statement that you show that u/c is correlated to seismic reflector strength, then backtrack on line 13 to discuss human factors. I would suggest that either this paragraph is split and both sections elaborated, or that the topic sentence incorporates both concepts. See Major comment 2 11 (26): How strong is the Top Ness horizon? Does this effect how well it is interpreted?

AUTHORS COMMENT: We do not see this as much as backtracking from the first statement, but rather we are addressing the complexity involved: Fault interpretation uncertainty correlates with seismic reflector strength, as our study results show, but it is dependent on many more variables (multiple correlation).

CHANGE IN MANUSCRIPT:

13(31-32): How did Fossen and Hesthammer (1998) get their pole? What was there scale of observation (i.e. did they have the data to extrapolate the sinusoidal shape)? The work on this should be included in this part of the discussion.

AUTHORS COMMENT: Fossen and Hesthammer (1998) calculated the pole using Bingham analysis of data with some significant spread. Their study was conducted on the basis of "reprocessed 3D seismic (ST8511)".

CHANGE IN MANUSCRIPT: We have added information on the origin of the data to the manuscript.

12 (Fig 9): How do you define interpretation density; units should be added if applicable? Visually this figure could do with a general text size increase, with many areas of text being too small. I would also suggest a change of colour for the boxes in part (1).

AUTHORS COMMENT: The term density refers to probability density of the interpretations. This just means that the integral of the histogram sums up to 1 and is a dimensionless number. This normalization makes the histograms comparable despite fluctuations in fault stick counts across the cube.

CHANGE IN MANUSCRIPT: We have increased text sizes and modified colours in Figure 9 to improve legibility.

15 (5-7): This is a potentially important point and raises a very important question 'what order did students interpret the cube?' If students are spending more time on a certain area, where data is of better quality then there are more factors to consider in why your results are different. Also does the style of interpretation change with time?

I advise either that the key findings section be reduced to a summary (e.g. set of bullet points) and separate section added to explore the reasons behind the uncertainty, probably split into 'technical' (e.g. image quality) and 'human' (e.g. different mental models) and that appropriate literature be added to this discussion.

AUTHORS COMMENT: Indeed! For such an analysis the entire interpretation process would need to be captured (e.g. screen capturing) and analysed, which was not an available option for this study. We have therefore no information about the temporal evolution of student's interpretation.

CHANGE IN MANUSCRIPT: -

14 (27-29): I think it would be unwise to suggest normal distributions, even in areas of good seismic data. I suspect in nearly all cases the distributions will be skewed. Most faults display an asymmetric damage zone, and such will also show an asymmetric signature in seismic, should the flat tail be towards the hanging wall?

AUTHORS COMMENT: The skewness of a Normal distribution can be an important factor for the detailed modeling of structures and should be considered wherever possible. But to our knowledge no detailed study has ever been carried out in trying to identify geological controls on skewness and kurtosis for probability distributions describing faults (or other geological structures). Stochastic geomodeling also needs to achieve a difficult trade-off between the level of detail and computational feasibility, as the correct sampling from the joint distribution constructed from complex parameter distributions (e.g. fat-tailed distributions) becomes increasingly challenging given today's computational constraints (e.g. Betancourt 2017, 2018). Coupling knowledge about asymmetric damage zones with interpretation experiments would be a very interesting future research topic though!

CHANGE IN MANUSCRIPT: -

15 (3): I found this an underwhelming end to a really neat data set. Although the implications for machine learning are indeed relevant, I feel the MS is crying out for a conclusion section which ties the findings together and includes the 'next stages' in tackling uncertainty in 3D seismic interpretation. The section itself also seems somewhat out of the remit of this work, and could

conceivably either be reduced or cut to make space for a discussion into the reasons behind the results as suggested previously.

AUTHORS COMMENT: We have added a conclusion section to the manuscript.

CHANGE IN MANUSCRIPT: Please see the revised manuscript for the added conclusions section.

Please find additional minor comments/suggested text edits on the attached MS (many of which are included in the specific comments).

AUTHORS COMMENT: We have considered and incorporated numerous minor comments and text edits from the attached commented MS.

CHANGE IN MANUSCRIPT: Please see the revised manuscript or change-tracked manuscript.

References

Betancourt, M. (2017). A Conceptual Introduction to Hamiltonian Monte Carlo. ArXiv:1701.02434 [Stat]. Retrieved from http://arxiv.org/abs/1701.02434

Betancourt, M. (2018). The Convergence of Markov Chain Monte Carlo Methods: From the Metropolis Method to Hamiltonian Monte Carlo. Annalen Der Physik, 1700214. https://doi.org/10.1002/andp.201700214

Bond, C. E., Gibbs, A. D., Shipton, Z. K., & Jones, S. (2007). What do you think this is? "Conceptual uncertainty" in geoscience interpretation. GSA Today, 17(11), 4. https://doi.org/10.1130/GSAT01711A.1

Fossen, H., & Hesthammer, J. (1998). Structural geology of the Gullfaks Field, northern North Sea. Geological Society, London, Special Publications, 127(1), 231–261. https://doi.org/10.1144/GSL.SP.1998.127.01.16

Kruschke, J. K. (2013). Bayesian estimation supersedes the t test. Journal of Experimental Psychology: General, 142(2), 573–603. https://doi.org/10.1037/a0029146

Lindsay, D. R. (2011). Scientific writing. CSIRO publishing.

Tannert, C., Elvers, H.-D., & Jandrig, B. (2007). The ethics of uncertainty: In the light of possible dangers, research becomes a moral duty. EMBO Reports, 8(10), 892–896. https://doi.org/10.1038/sj.embor.7401072

Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. Science, 185(4157), 1124–1131. https://doi.org/10.1126/science.185.4157.1124

Author Response to RC2 (Awad Bilal)

Major comments responses

[1] While reviewing the manuscript, there have been some instances where the scientific approach was either oversimplified or need further clarifications. For example,

[1a] The authors attributed uncertainty in fault and horizon interpretations to mainly image quality. However, it was not clear to me what does seismic quality means in the manuscript (i.e. low amplitude and homogenous reflectors? low amplitude, chaotic reflectors? or noisy seismic. etc.). In addition, there was no information about the seismic data in the materials and methods section. I appreciate the work done in section 3.3 (seismic data quality) but I worry about the use of RMS attribute (reflector strength) to assess seismic data quality. Strong seismic reflections do not necessary means high seismic quality. Clear definition of what the authors consider as high or poor seismic quality is therefore crucial.

I suggest the authors start by describing the seismic stratigraphy of the sequences and the seismic characteristics of the analysed Top Ness horizon. Areas of poor seismic quality can then be highlighted together with the possible reasons behind low quality data (e.g. depositional environment, inherited low quality i.e. processing. . . etc). In order to give the reader more confidence, I suggest including a new figure (as Fig. 2 or 3), comprising two, N-S and E-W seismic sections, showing the interpreted horizons and seismic quality.

AUTHORS RESPONSE: Thank you for pointing out the lack of definition of the term seismic data quality. We have added a paragraph to the manuscript to explain our definition of the term. As our manuscript is mainly concerned with a first-look statistical description of the uncertainties encountered in 3-D seismic interpretation, we think it out of scope of this work to do a detailed study on the sources of seismic data quality and how it affects interpretation uncertainty. Nevertheless, this would be a logical next step for future research.

CHANGE IN MANUSCRIPT: Throughout this work we make use of the term seismic data quality not in the strictly geophysical sense of quality factors surrounding seismic data acquisition and processing, but rather in the sense of the interpretability of the seismic data. If the seismic data lacks clear, continuous reflectors in a region, but shows a noisy image difficult to interpret – no matter what the source of this may be – we describe it as an area of low seismic data quality. Similarly, if reflector strengths are high and continuous (for horizon interpretation) or clearly offset (for fault interpretation), we speak of high seismic data quality.

[1b] The authors observed that uncertainties in the horizon (Top Ness; Fig. 8) interpretation are significantly reduced surrounding wells with a general trend of increasing uncertainties from west to east. I suggest that this is further analysed and discussed. Important aspects to be considered are; (1) what is the seismic characteristics of the interpreted horizon? (2) what is the structural configuration (faulted, non-faulted) of the horizon; (3) how confident the students/authors are with both formation tops and well locations; and (4) is there any uncertainty in formation tops that can be compared/correlated to the uncertainties analysed in this paper. Additional factors that might also have affected interpretation include the level of assistance to students, time spent in the interpretation, the methodology and order of interpretation and how familiar the students were with using Petrel.

The fact that uncertainty increases away from well location is a general statement and, in many cases, it is invalid unless other factors are considered. Horizon interpretation uncertainty away from well location largely depends on the continuity of the horizon (also dependent on the depositional environment) and structural complexities. There are many cases where uncertainty is significantly low for hundreds of kilometres away from well location. Other cases show high uncertainty in near to the well or even in the Formation tops of the well.

AUTHORS RESPONSE: We have expanded our discussion regarding the horizon uncertainty surrounding well tops and faults in the results section to address some of the seismic characteristics of the horizon we believe relevant to the patterns (1). We believe the structural configuration is sufficiently characterized in the manuscript (2).

(3) We have no information about the inherent uncertainty in the formation tops and well locations.

(4) We have no information about the inherent uncertainty in the welltops. As all students received the same welltops, thus our analysis can be seen as "anchored" on the uncertainty in the dataset they received.

You raise a very good point regarding that the horizon interpretation uncertainty away from wells depends on the continuity of the horizon. We have incorporated this aspect into our manuscript.

CHANGE IN MANUSCRIPT: Figure 8a shows the average Top Ness horizon basemap for all interpretations combined. Overall the horizon interpretations are increasing in depth from SE towards the NW of the domain. Figure 8b shows the associated standard deviation of the average Ness horizon interpretation, with an overlay of mean fault intersections and well locations. We observe large horizon uncertainties in vicinity to both Faults 1 and 2 throughout the dataset. An increase of horizon uncertainty is poignant at the southern end of the domain where Faults 1 and 2 are beginning to merge again. In the North the horizon uncertainty surrounding Fault 3 decreases rapidly with distance from the fault, with two welltops and packages of high reflector strength (see Fig. 9A) constraining the uncertainty. As the seismic data quality decreases towards the south (see Fig. 9B), the uncertainty in the horizon interpretation increases in the Eastern part of the dataset. Interpretation uncertainties are significantly reduced surrounding well locations in the western part of the study domain, where seismic reflectors of the Top Ness horizon are overall stronger and more continuous (see Fig. 9A). This pattern does not hold true in the East of the dataset, where reflector continuity is overall low and noise in the seismic dataset is high.

[1c] I worry about the uncertainty of fault throw. For example, Figure 7b. shows that uncertainty is higher at smaller median throw (southern part of the fault), while it is expected to see high uncertainty at larger throw value. The figure also shows that same amount of median throw can have both high and low uncertainties (see central part of the fault Y 6784000). This can also be observed in Fig. 7a. Generally, uncertainty in throw mainly depends on the uncertainty of the correlation of the stratigraphic markers in the hanging wall and footwall blocks of the fault. For example, uncertainty in fault throw will be low where the stratigraphic correlation between fault blocks is good independently of throw value. In more detail, the uncertainty in cut-off angle in each sides of the fault depends on many aspects such as structural complexity, fault plane definition on the seismic, seismic quality near faults. The new publication of Ze and Alves titled 'Impacts of data sampling on the interpretation of normal fault propagation and segment linkage' might be of interest.

AUTHORS RESPONSE: We do not see problem of small median fault throws having larger uncertainties in this dataset. As we have described in the manuscript, the throw of Fault 3 becomes

more uncertain towards the South. We have detailed the decrease in data quality from North to South, making both fault placement (shown in Fig 3A) and horizon placement (shown in Fig 7b) more uncertain, which explains how the uncertainty of fault throw also increases in this area. The lack of clear stratigraphic markers on both sides of Fault 3 in the South, and the existence of good stratigraphic markers in the North emphasize this (as seen in Fig 8A and B).

CHANGE IN MANUSCRIPT: -

[2] I believe that the manuscript would be greatly improved and give more confidence to the reader if the 'Key findings' section of the discussion part is discussed in more depth. The fact that uncertainty in seismic interpretation increases in areas of poor seismic quality and linkage zones is relatively well known and expected. I appreciate the work done on quantifying these uncertainties, but I advise that the results of the paper are discussed in more detail. One important aspect that could be highlighted is the reasoning behind the variability in seismic interpretations, taking into consideration other factors that might affect seismic interpretation. Human bias in seismic interpretation and the expertise of the interpreters can also be addressed. It would also be interesting to see whether the result would change if the interpretation was carried out by experienced seismic interpreters using seismic filters to improve seismic quality together with various attributes (i.e. the effect of the level of experience). The authors can also analyse the effect of training (supervised top Cretaceous horizon interpretation) in reducing horizon interpretation uncertainty.

AUTHORS RESPONSE: We have overhauled the discussion / key findings part of our manuscript based on your and the other referees' comments.

Collecting and analysing extensive 3-D interpretations from experts would indeed form a great study but is definitely out of scope for this experiment. Also, we believe that the effect of training can not really be analysed from comparing the results presented in the manuscript with interpretations of the Top Cretaceous horizon, due to its stark contrast in structural simplicity and high reflector continuity compared to the Top Ness horizon throughout the dataset.

CHANGE IN MANUSCRIPT: Please see the revised manuscript or change-tracked manuscript for all the changes in the discussion.

[3] While reviewing the manuscript, there have been several instances where the description was not clear, or the authors use a great deal of qualitative language. I would prefer some quantifications of what few, major etc mean. These together with other minor corrections are highlighted in the attached original manuscript.

AUTHORS RESPONSE: Thank you for making us aware of this! We have removed the use of subjective language throughout the manuscript and have incorporated a large amount of your minor corrections.

CHANGE IN MANUSCRIPT: A multitude of changes throughout the manuscript. Please see the revised manuscript or the change-tracked manuscript.

[4] Conclusions: This paper describes a very important aspect of 3D seismic interpretation. However, a clear and concise summary with a clear link to the results and discussion would highlight the significance of the paper and results. Hence, I suggest the authors add a short summary or conclusions as bullet points with a clear take-home message.

AUTHORS RESPONSE: We have added a clear and concise conclusion to the manuscript.

CHANGE IN MANUSCRIPT: Please see the revised manuscript for the added conclusion section.

[5] Figures: Overall, the figures of the manuscript are clear and well-presented but please make sure of the resolution in the final printed version. Text size in some figures needs to be increased. Several figures have neither a scale bar nor north arrow. I also have suggested re-drafting and/or adding three new figures. Detailed comments on these new figures and on each figure are highlighted in the attached original manuscript.

AUTHORS RESPONSE: Thank you for pointing this out! We have increased text sizes and generally improved legibility of the figures throughout the manuscript. We have added scale bars and North arrows where applicable. We improved our Figures throughout the manuscript due to the reviewers input and merged the histogram and stereonet plots to improve the clarity for the reader. We decided against adding additional figures as we believe our improved Figures will adequately provide the reader with the relevant information and it would further bloat the manuscript. As our study mainly focuses on the structural uncertainties, we believe the example interpretation in combination with the average basemap and the seismic sections in the Figure 8 provide the reader with the context they need to follow the manuscript.

CHANGE IN MANUSCRIPT: Please the see the revised manuscript for changes in Figures.

Detailed comments responses

Abstract

Overall, it provides a clear summary but I believe it will be updated according to the reviewer's comments.

AUTHORS RESPONSE: We did not make changes to the summary, as the scientific content of the paper did not change.

CHANGE IN MANUSCRIPT: -

Introduction

Overall, the text is clear but I suggest the authors further clarify uncertainty in seismic interpretation and make it accessible to non-specialists. The authors could simply address that conceptual uncertainty mainly comes from human bias in seismic interpretation and depends on the expertise of the interpreter. Also, other types of uncertainties such as geophysical uncertainty, uncertainties related to formation tops and check-shot data could also be highlighted.

AUTHORS RESPONSE: We agree that a more specific description of what kind of uncertainty we are investigating (and how it related to "objective" uncertainty) would increase the clarity of the manuscript.

CHANGE IN MANUSCRIPT: Our work is thus concerned with quantifying the scope of uncertainties in seismic interpretation, which represents inevitably biased, human judgment under uncertainty (Tversky and Kahneman, 1974). This "subjective" uncertainty is in contrast to more "objective" uncertainty (Tannert et al., 2007; Bond, 2015) related to the geophysical acquisition of the data itself.

[L12] Please add the name of the 3D seismic survey

AUTHORS RESPONSE: We added the name and details of the 3-D seismic survey in the dataset section 2.1. Thanks for pointing that out!

CHANGE IN MANUSCRIPT: -

[L14] I would be more specific here. Seismic noise at depth? Near faults?

AUTHORS RESPONSE: We have expanded the specific description of seismic noise and referenced relevant Figure with seismic sections giving an example.

CHANGE IN MANUSCRIPT: The dataset depicts a comparatively simple geometry of planar domino-style normal faults, but the seismic dataset exhibits high amounts of noise, especially in its eastern half, and generally increasing with depth and in fault proximity (see Fig. 9).

[L19] The authors have shown a simple fault interpretation of Fossen and Hesthammer (1998) as a reference example. How about horizon interpretation? Do you have an interpreted horizon? Please find further comments on this below (i.e. Fig. 8).

AUTHORS RESPONSE: We use the fault interpretations from Fossen & Hesthammer (1998) as a conceptual reference interpretation – i.e. their fault network topology and the reported fault orientations – we do not have access to their unpublished interpretation. We clarified this in the manuscript.

CHANGE IN MANUSCRIPT: We use the interpretation of Fossen & Hesthammer (1998) as a reference expert example (in the sense of Macrae et al., 2016) to compare fault network topology and fault orientation uncertainty with the student interpretations.

2.1 – Gullfaks geology and seismic data

The paragraph lacks information about the 3D seismic data. Key information includes (1) the name of the 3D seismic survey; (2) acquisition year; (3) whether the data is originally in depth or time? and (4) key information about processing history as this will further clarify the seismic quality issue. I would also add the key tectonic phases evolution of the area (i.e. how many phases and ages).

AUTHORS RESPONSE: We agree that this information is of importance so assess sources of uncertainty within the seismic data itself. However, we have limited information on the dataset and its origin. We have added all relevant information that is available to us to the section.

We disagree on adding further detail on the tectonic evolution of the area into the paper and refer to the comprehensive description in published literature, as it would further bloat the manuscript.

CHANGE IN MANUSCRIPT: The 3-D seismic survey of the Gullfaks field, ST85R9211, was recorded in 1985 and reprocessed in 1992. It was recorded in time and converted to depth using TWT depth conversion and was migrated using a Prestack Kirchhoff migration

L30-33 – Please clarify the research focus area both here and in Figure (1). I also suggest clarifying the scale of the interpretation area. Could you also highlight that the overall structure of the area is very complex particularly in the accommodation zone and collapsed footwall structures in the horst complex zone, while the western dominosystem (research focus) is comparatively simple.

AUTHORS RESPONSE: We have clarified the difference in complexity of the structural domains and improved clarity of the scale of the interpretation area in Figure 1.

CHANGE IN MANUSCRIPT: The field consists of three structurally distinct domains: a structurally simple domino system in the western part, and the structurally more complex accommodation zone and Horst complex towards the east (see Figure 1c). Our study focuses on the structurally simpler

western part of the domino system, where we investigate the uncertainty of the three faults F1-F3 and the fault blocks A-E depicted in Figure 1b.

P3_Figure 1 - Please increase text size particularly in Fig. 1a and the Formation names in Fig. 1c. The location of the 3D seismic is not clear. The schematic fault map (b) is oversimplified. Please show fault displacement (fault polygons) and dip direction. Fault polygons would greatly help the reader to understand linkage zones as well as major and minor faults. I wonder how Fault 1 was identified in the northern part where it merges with Fault 2. Please see further comments highlighted on the attached original manuscript.

AUTHORS RESPONSE: We have improved the clarity of Figure 1. We have simplified the schematic fault map (Fig 1b) on purpose, as it was adapted from a heave-map of a different horizon from Fossen & Hesthammer (1998). We have added dip direction markers to the plot to improve clarity. We have shaded the schematic faults to make fault identification easier for the reader.

CHANGE IN MANUSCRIPT: Changes to Figure 1.

2.2 – Interpretation dataset

Many questions came to my mind with respect to seismic stratigraphy. What are the seismic characteristics of the Top Ness horizon and why it was selected for analysis? Please add this key information to give greater confidence to the reader with respect to the results.

AUTHORS RESPONSE: Our study focuses on the quantification of structural seismic interpretation uncertainties and we believe that a detailed discussion of seismic stratigraphy is out of scope for this work. The Top Ness horizon was interpreted by the students as part of their seismic interpretation course, thus the selection of the horizon was out of our hands. We agree though, that this would be very interesting topic for future research.

CHANGE IN MANUSCRIPT: -

P3_L4 – Please show the location of the wells in Figure 1. How confident you are with well locations and Formation tops.

AUTHORS RESPONSE: We have no way to assess the uncertainty related to the well locations and formation tops. We agree that this information is vital to an ideal, uncertainty-integrated workflow, but our study also focuses on the "subjective" uncertainty produced by the interpretation. All students started out with the same set of well information.

CHANGE IN MANUSCRIPT: -

P3_L8 – Why did students use manual interpretation? how often seeded tracking and manual interpretation were used?

AUTHORS RESPONSE: Due to the amount of noise throughout the dataset, manual interpretation would be necessary in many parts of the cube. We have no precise information about the frequency of individual interpretation tools. Our information is from the course instructions the students were given, and by interviewing a few students about their interpretation process. We have clarified this in the manuscript.

CHANGE IN MANUSCRIPT: Afterwards the students started to independently interpret the Base Cretaceous Unconformity (BCU) and Top Ness horizon (which is part of the Brent group) around well locations, followed by connecting the horizon interpretation in-between wells. The students were

instructed to mainly use guided auto-tracking, as well as occasional seeded tracking and manual interpretation where possible or necessary, depending on seismic data quality.

P3_L9 – Do you mean interpolated auto-tracked surfaces? Did students actually use gridding? I would be more specific

AUTHORS RESPONSE: We removed the reference "gridded" as it can cause confusion in this context. It referred to the regularly gridded surfaces Petrel creates from horizon picks.

CHANGE IN MANUSCRIPT: The students then interpolated surfaces from the horizon interpretations using Petrel's Make Surface function.

P4_L3 – Why the Top Ness horizon was selected for analysis. Why not BCU

AUTHORS RESPONSE: Due to work volume constraints in the course, students were instructed to only interpret the Top Ness horizon below the BCU. We chose to focus our analysis for this manuscript on the Top Ness horizon, due to its higher structural complexity compared to the BCU.

CHANGE IN MANUSCRIPT: -

P4_Figure 2 - Please increase figure and text size. Also show both North arrow and a

scale bar.

AUTHORS RESPONSE: We have increased the text size and added North arrow and scale bar to the figure

CHANGE IN MANUSCRIPT: Figure 2 changes.

2.3 – Data analysis

It was not clear to me the key steps of the analysis. I would suggest adding a diagram showing steps of the workflow applied as this would greatly help the reader in particularly non-specialists. Please see further minor comments highlighted on the attached original manuscript.

AUTHORS RESPONSE: We decided against the addition of a workflow diagram, as we used common data analysis approaches to analyse our data and it would further bloat the manuscript. We provide our custom Python functionality in an open-source code repository for readers interested in the exact implementation. We hope our manuscript changes made particular parts of the data analysis clearer to the reader.

CHANGE IN MANUSCRIPT: Please see the revised or change-tracked manuscript for the changes made to the data analysis section.

P5_L8 – I am not aware of RMSA. I am sure RMS only (without 'A') would be fine.

AUTHORS RESPONSE: We define RMSA as a shorthand for RMS Amplitude to use throughout the manuscript for brevity.

CHANGE IN MANUSCRIPT: -

P5_6 - Figures 3 & 4 – I would suggest merging these figures or having both of them in the same page to make it easy for the reader to follow. Please make sure of the quality of the figure in the final printed version of the paper.

AUTHORS RESPONSE: Thank you for the suggestion! We have merged the two figures into one to improve clarity for the reader. Both figures are vector files, any quality issues might be due to pdf compression done by Copernicus or caused by the pdf-reader software used.

CHANGE IN MANUSCRIPT: Merged Figure 3 and 4 into Figure 3.

P6_L2-3 – What do you mean by spread? Can you make this clear?

AUTHORS RESPONSE: Spread is used in its statistical meaning identical to scatter or dispersion.

CHANGE IN MANUSCRIPT: Changed spread to dispersion for clarity.

P7_L13-14 – Why F3 was omitted? If it was interpreted in a similar fashion, why does uncertainty increases southwards as shown in Figure 3c?

AUTHORS RESPONSE: In all but a single student interpretation F3 does not interact with F1 or F2 and was thus not drawn in the fault network diagrams of the five most common fault network topologies, as it never changed. The uncertainty of F3's position and orientation increases towards the south, without impacting fault network topology.

CHANGE IN MANUSCRIPT: -

P8_Figure 6 – Can you show students number in the Y-axis of Fig. 6a. Please add fault numbers (e.g. F1, F2) and dip direction in Fig. 6b. I was wondering how the authors differentiated between A and C and how fault 1 was defined in the northern part where it merges with fault 2 (Fig. 6b).

AUTHORS RESPONSE: We have added student numbers to the probability mass plot, as well as fault numbers. We think that dip directions would unnecessarily clutter the figure without adding critical information. We have added dip directions to the overview plot in Figure 1.

We differentiate A and C by the location of the fault stick placements in the dataset. North of the merging of Fault 1 and 2, the fault has been classified as belonging to the fault that terminates the other. E.g. if Fault 2 terminated at Fault 1, and Fault 1 continued northward, it was classified as Fault 1. We have enhanced the Figure schematics to improve clarity.

CHANGE IN MANUSCRIPT: Changes to Figure 6 (now Figure 5)

2.3 – Fault throw and horizon uncertainty

It was not clear to me whether the throw analysis was carried out along fault strike or a strike-parallel window.

What are the reasons behind selecting faults 1 and 3 (which was interpreted in a similar fashion and omitted in the previous section) for throw analysis?

AUTHORS RESPONSE: We have chosen to show the fault throw analysis for Faults 1 and 3 as the results of our analysis are reliable for them. The complex nature of Fault 2 and its subparts made it quite difficult for our algorithm to confidently analyse the fault throw. We are working on improving our algorithm for future work to apply it confidently to more complex faults. Another reason is, that Fault 2 and 3 are good examples of how fault throw uncertainty is changing across datasets.

CHANGE IN MANUSCRIPT: -

P9_Figure 7a. - How about uncertainty in the linkage zone? I think that the increase in uncertainty near survey edge is more likely related to the complex linkage zone to the south where faults 1, 2a and 2b merge as shown in Figure 1. It is well known that fault interaction can influence fault

propagation and hence, displacement profiles. Do you think that the displacement profile shown in Fig. 7a would change if fault 2 was added to fault 1. It will be very interesting to see both faults.

AUTHORS RESPONSE: We have modified Figure 7 to represent the data as boxplots along fault strike, which much better visualizes the changes in fault throw and its uncertainty, especially in the linkage zone of F1 and F2. Thank you for pointing out the linkage zone to the south – we've added this to the results.

CHANGE IN MANUSCRIPT: Changes in Figure 7 and throughout the manuscript.

P9_Figure 7b. – As indicated in the general comments above, this figure shows higher uncertainty associated with small median throw. Also, same amount of median throw show both high and low uncertainty. Please make of the analysis. Please see further comments highlighted in the attached original manuscript.

AUTHORS RESPONSE: We have redrafted Figure 7 as boxplots to better visualize the fault throw data. Fault throw in the Southern part of Fault 3 does show lower median with higher uncertainty (IQR), which is not unrealistic. We speculate this to be due to the problematic seismic data quality (in terms of interpretability), thus the variation in where students put the fault and their Top Ness horizon is increased (as seen also in Figure 8 surrounding Fault 3).

CHANGE IN MANUSCRIPT: Please see the change-tracked manuscript for the full amount of changes surrounding fault throw.

P9_L4-9 - One key question came to my mind while reading the last paragraph of page 9 (description of Figure 8) is that where is the actual interpretation of the Top Ness Formation. I strongly suggest that the best interpreted horizon is added. so the reader can see the actual interpretation of both horizon and faults. It would be even better if it was interpreted by experienced seismic interpreter.

AUTHORS RESPONSE: We have to disagree with adding the best interpreted horizon, as we believe this is not very meaningful in this context. We have a large number of interpretations, and a priori no interpretation is necessarily better than any other as we lack knowledge of the real subsurface situation. This is why we present the mean Top Ness horizon in Figure 8a: to visualize a single interpretation in combination with the uncertainty in Figure 8b.

CHANGE IN MANUSCRIPT: -

P10_Figure 8 – Please redraft to include (a) the actual (best) interpreted horizon, (b) average horizon, and (c) the standard deviation horizon. Also, please add north arrow and scale bar in all parts of the figure. Please see further comments highlighted in the attached original manuscript.

AUTHORS RESPONSE: Please see the above response in regards why we believe that adding a single best horizon is not meaningful in the context of this manuscript.

CHANGE IN MANUSCRIPT: Added scale bar and North arrow to Figure 8.

3.3 - Seismic data quality
As indicated in the general comments above. I worry about the title of this section as well as the framework and results. The strength of seismic reflector does not necessarily mean good seismic quality. I also worry about the use of a time slice (Fig. 9.1 & 9.2) to assess seismic quality based on the strength of a single seismic reflector as the results can be significantly different at different levels or even the same level.

AUTHORS RESPONSE: We have chosen the time slice because we believe it serves as a valuable example of how the uncertainty in the interpretation of a single fault can vary in response to the seismic data quality (in terms of structural interpretability). As RMSA attribute is calculated across a vertical window, it includes information about reflectors above and below the depth slice shown.

CHANGE IN MANUSCRIPT: -

I believe that interpretation of fault 3 was aided in the northern part of the fault (Fig. 9.1 & inline section A) because of both the sharp reflection termination and the good correlation between seismic reflection packages on both sides of the fault. Generally, fault interpretation dose not actually depends on individual reflections, instead packages of reflections across the fault.

AUTHORS RESPONSE: This is the conclusion we are reaching in our manuscript: that sharp seismic reflector termination (as indicated by the through in RMSA response in Fig. 9a) strongly decreases fault interpretation uncertainty. We use Fault 3 to show how interpretation uncertainty increases when the noise of the dataset decreases and how we can possible use the RMSA response as a proxy for uncertainty.

CHANGE IN MANUSCRIPT: -

I would suggest the authors further test their results and see if the same approach can be applied to Fault 1 and 2 where the high amplitude to the south (Fig. 9) coincide with high uncertainty in both fault (Fig. 3) and horizon (Fig. 8) interpretation. Moreover, the authors could test whether the results are same if (1) the time slice was replaced by RMS attribute extraction on the actually interpreted horizon; or (2) a different time slice (e.g. 2100 or 2300) was used. Further minor comments on this and Figure 9 are highlighted in the attached original manuscript.

AUTHORS RESPONSE: We chose to present our analysis on Fault 3 as it serves as a meaningful example of the impact of noise / data quality on the interpretation uncertainty of a simple fault without many fault interactions. Data quality is generally poorer surrounding Faults 1 and 2 (as also seen in the increased spread in Figure 3) and the structural complexity is higher (merging and splitting of the faults). This is why we chose Fault 3 as a simple first example because Faults 1 and 2 are much more complex. We agree that there is much more opportunity for further research on how structural complexity interacts with fault interpretation uncertainty and data quality and we are aiming to do further work on this in the future.

(1) is a very good point! In more complex cases extracting the RMSA along a horizon probe would be ideal. But we lack an "actually interpreted" horizon (or rather we have too many valid ones). We could of course use the mean or median horizon, but this is not necessarily meaningful. As the pattern of increasing noise from North to South holds true for the dataset at all depths surrounding Fault 3 we believe it to be adequate to use a depth slice instead of a horizon probe at an arbitrary interpretation.

(2) We chose this specific depth slice as it encompasses easily interpretable results in the North with deceasing interpretability towards the South to show the trend visualized in Fig. 9a-d on how interpretation uncertainty correlates with seismic data quality.

CHANGE IN MANUSCRIPT: Several changes in section 3.3, please see the revised manuscript or difference-tracked manuscript.

4 – Discussion

Please see the attached original manuscript for further minor comments on the discussion part.

AUTHORS RESPONSE: We have incorporated many of your minor comments into the manuscript.

CHANGE IN MANUSCRIPT: Please see the revised and change-tracked manuscript for a detailed overview of all changes made.

Quantification of uncertainty in 3-D seismic interpretation: implications for deterministic and stochastic geomodelling and machine learning

Alexander Schaaf[1] and Clare E. Bond[1]

[1]Geology and Petroleum Geology, School of Geosciences, University of Aberdeen, AB24 3UE, UK

Correspondence: a.schaaf@abdn.ac.uk

Abstract. In recent years uncertainty has been widely recognized in geosciences, leading to an increased need for its quantification. Predicting the subsurface is an especially uncertain effort, as our information either comes from spatially highly limited direct (1-D boreholes) or indirect 2-D and 3-D sources (e.g. seismic). And while uncertainty in seismic interpretation has been explored in 2-D, we currently lack both qualitatitive and quantitative understanding of how interpretational uncertainties of 3-D datasets are distributed. In this work we analyze 78 seismic interpretations done by final year undergraduate (BSc) students of a 3-D seismic dataset from the Gullfaks field located in the northern North Sea. The students used Petrel to interpret multiple (interlinked) faults and to pick the Base Cretaceous Unconformity and Top Ness horizon (part of the Mid-Jurassic Brent Group). We have developed open-source Python tools to explore and visualize the spatial uncertainty of the students fault stick interpretations, the subsequent variation in fault plane orientation and the uncertainty in fault network topology. The Top Ness horizon picks were used to analyze fault offset variations across the dataset and interpretations, with implications for fault throw. We investigate how this interpretational uncertainty interlinks with seismic data quality and the possible use of seismic data quality attributes as a proxy for interpretational uncertainty. Our work provides a first quantification of fault and horizon uncertainties in 3-D seismic interpretation, providing valuable insights into the influence of seismic image quality on 3-D interpretation, with implications for deterministic and stochastic geomodelling and machine learning.

1 Introduction

Geosciences, and geology in particular, are concerned with integrating various sources of data, often of limited, sparse and indirect nature, into scientific models. The use of limited data combined with our limited knowledge of the highly complex earth system, invariable infuses any model with uncertainty. Especially as geology inherently relies heavily on interpreted data that often requires reasoning about processes that occur over geological time scales (Frodeman, 1995), which further increase the space of uncertainty.

The interpretation of 3-D seismic data is an integral part of constructing structural geomodels of the subsurface and plays a major role in the energy industry. Due to the indirect, noisy and non-unique nature of seismic data processing into images, interpretation is inherently uncertain. Our work is thus concerned with quantifying the scope of uncertainties in seismic interpretation, which represents inevitably biased, human judgment under uncertainty (Tversky and Kahneman, 1974). This "subjective" uncertainty is in contrast to more "objective" uncertainty related to the geophysical acquisition of the data itself (Tannert et al., 2007; Bond, 2015). Previous work has shown that significant conceptual uncertainties and biases are encountered during the interpretation process of 2-D seismic lines (Bond et al., 2007, 2011; Macrae, 2013; Bond, 2015; Alcalde et al., 2017a, c), as well as the impact of seismic image quality on the interpretation (Alcalde et al., 2017b). But subsurface structures are naturally three-dimensional and the use of 3-D seismic data is ubiquitous in industry (Biondi, 2006) ~~which~~. This raises the need to further our understanding of the distribution of interpretational uncertainties in 3-D space (Abrahamsen et al., 1992; Thore et al., 20 . Additionally, the process of interpretation between seismic lines and cubes is fundamentally different, and thus might lead to conceptually different uncertainties to be dominant (e.g. the need to connect fault evidence between seismic lines introduces significant uncertainty in widely spaced 2-D interpretation compared to fault interpretation in seismic cubes; see Freeman et al., 1990).

In this work we investigate the scope of uncertainties in 3-D seismic interpretation. We qualitatively and quantitatively analyze interpretations of 78 final year undergraduate students (BSc) conducted on a 3-D seismic cube of the Gullfaks field. The dataset depicts a ~~domino system of fault blocks with a comparatively simple structural style~~comparatively simple geometry of planar domino-style normal faults, but the seismic dataset ~~comes with~~ exhibits high amounts of noise~~throughout, which~~, especially in its eastern half, and generally increasing with depth and in fault proximity (see Fig. 8A, B). This inhibits straightforward interpretation of ~~the~~ major faults and horizons and limits use of structural seismic attributes. We analyze the spatial variation in fault stick interpretations and the subsequent uncertainty in fault orientation. Horizon interpretations are analyzed and combined with fault interpretations for a description of fault throw uncertainty. Additionally, we investigate the differences in fault network topology (see Morley and Nixon, 2016; Peacock et al., 2016, 2017) in the interpretation ensemble to better estimate the uncertainty of interpreting fault networks in 3-D seismic data. We use the interpretation of Fossen and Hesthammer (1998) as a reference expert example (in the sense of Macrae et al., 2016) ~~for comparison~~ to compare fault network topology and fault orientation uncertainty with the student interpretations. We integrate our findings of interpretation uncertainties with its relation to seismic data quality and discuss the implications for both deterministic and stochastic geomodelling, as well as machine learning applied to seismic interpretation.

2 Materials and Methods

2.1 Gullfaks geology and seismic data

We give here a brief overview over the regional and structural geology of the study area; a more in-depth description of the structural geology of the Gullfaks Field can be found in Fossen and Hesthammer (1998).
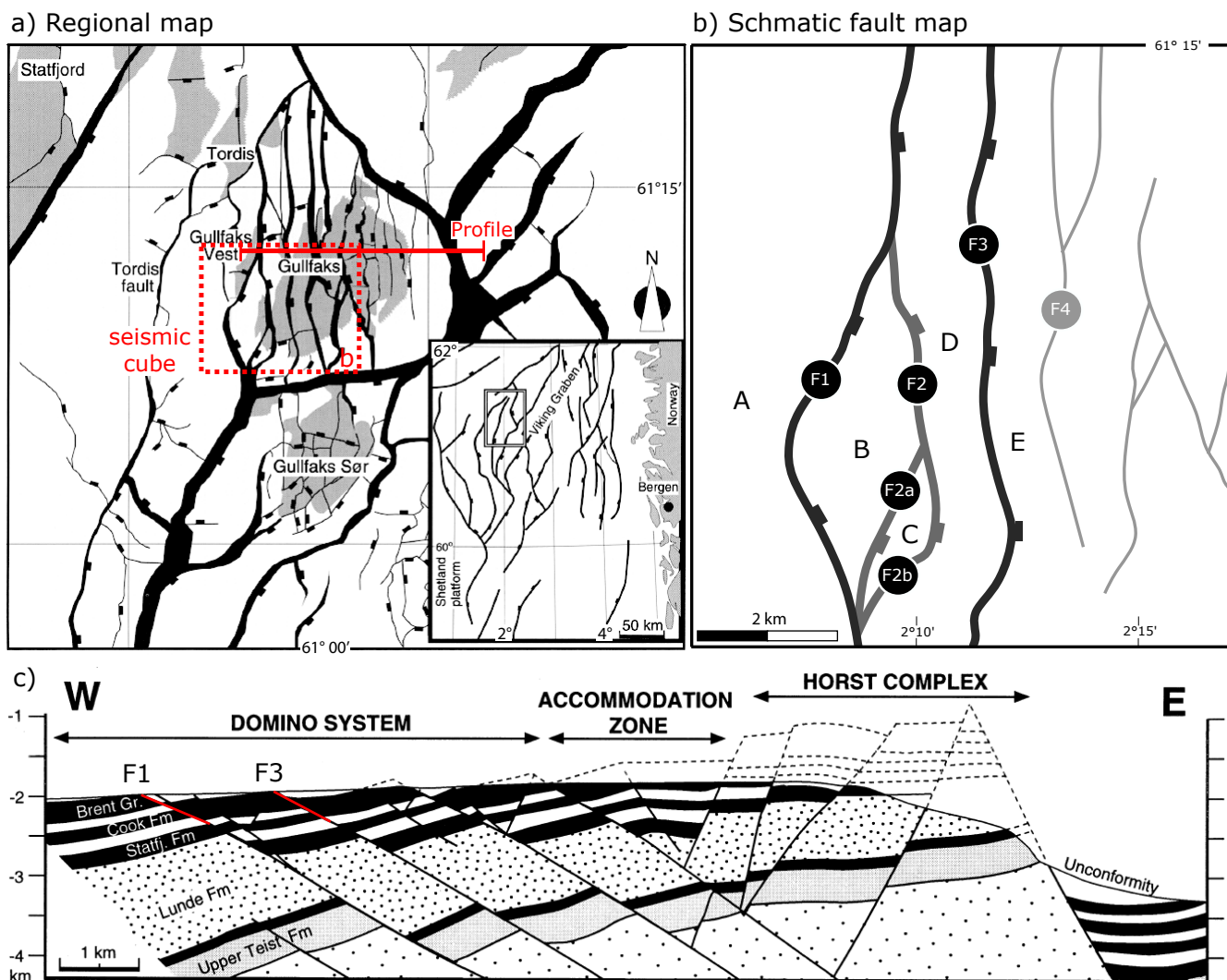
Figure 1. (a) Regional overview of the ~~Gullfaks~~ Gullfaks-Statfjord area located in the northern North Sea, showing the location of the Gullfaks oil field; (b) Fault map of the Statfjord ~~formation~~Formation, showing the main faults of the Gullfaks field and their labeling; (c) Cross-section across the Gullfaks field, depicting the three distinct structural ~~subsets~~domains, major faults and stratigraphy (modified from Fossen and Hesthammer, 1998).

The Gullfaks Field is a subset of the NNE-SSW-trending, $10 - 25$ km wide Gullfaks fault block, located in the western part of the Viking Graben within ~~the~~ ~~th~~ northern North Sea (see Figure 1a). The Gullfaks field's reservoir units reach from the late Triassic Hegre Group, over the Early Jurassic Statfjord Formation, Dunlin Group up to the Brent Group (Hesthammer and Fossen, 1997). The reservoir units are separated from the Upper Cretaceous sediments above by the Base Cretaceous ~~unconformity~~ Unconformity (Fossen and Hesthammer, 1998). The field consists of three structurally distinct ~~subsets: a~~ domains: a structurally
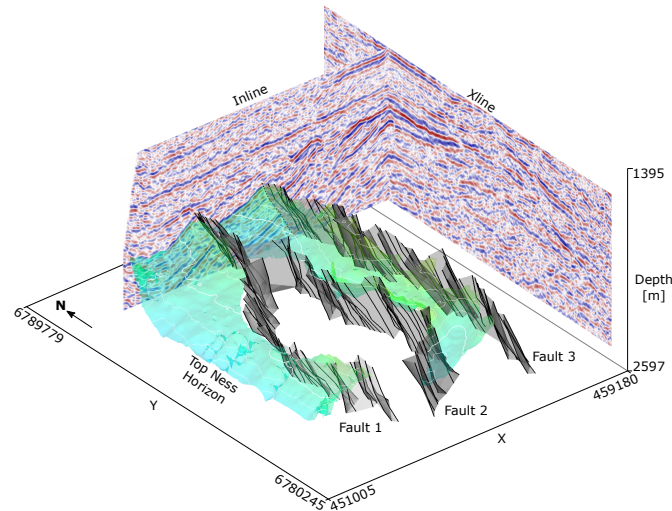
Figure 2. Example interpretation from a single student, showing the three major faults considered in our study as well as part of the Top Ness horizon interpretation. BCU and additional faults towards the East are hidden for better visualization of the key elements.

simple domino system in the western part, ~~an~~ and the structurally more complex accommodation zone and ~~a~~ Horst complex towards the east (see Figure 1c). Our study focuses on the ~~major~~ structurally simpler western part of the domino system, where we investigate the uncertainty of the three faults F1 - F3 and the fault blocks A - E depicted in Figure 1b~~, which are part of the western domino system~~. Note that Fault 1 and 2 merge in the northern half and at the bottom of the domain and that Fault 2

5   splits into two smaller faults F2a and F2b.

The 3-D seismic survey of the Gullfaks field, ST85R9211, was recorded in 1985 and reprocessed in 1992. It was recorded in time and converted to depth using TWT depth conversion and was migrated using a Prestack Kirchhoff migration.

2.2   Interpretation dataset

The analyzed interpretations were produced as part of the Surface and Subsurface Digital Imaging course within the undergrad-

10   uate program Geology and Petroleum Geology at the University of Aberdeen. While the students had prior training in structural geology and interpretation of 2-D seismic data, this was the student's first hands-on course in 3-D seismic interpretation using the Petrel software as part of their undergraduate program. The fourth year undergraduate (BSc) students loaded the seismic data into Petrel together with 14 wells, while ensuring proper georeferencing. The following interpretation process focused on first interpreting the Top Cretaceous horizon with initial support by the lecturing staff. Afterwards the students started to

15   independently interpret the Base Cretaceous Unconformity (BCU) and Top Ness horizon (which is part of the Brent group) around well locations, followed by connecting the horizon interpretation in-between wells~~using mainly~~. The students were instructed to mainly use guided auto-tracking, as well as occasional seeded tracking and manual interpretation where possible or necessary, depending on seismic data quality. Afterwards fault interpretations were conducted of major faults. The stu-

4

dents then interpolated ~~gridded~~ surfaces from the horizon interpretations using Petrel's *Make Surface* function. Polygons were created based on fault locations to create a Top Ness surface subdivided into the fault blocks.

For our study we collated the interpretation data from the students Petrel projects into a joint project, where interpretations were sorted and labeled. Of a total of 90 student interpretations, we used 78 in our study, as the other 12 either lacked relevant

5 interpretations or were corrupted Petrel project files. Interpreted Top Ness horizon surfaces and fault sticks where then exported systematically to allow for automated data processing and analysis. An example student interpretation is shown in Fig. 2, containing the three major faults considered, as well as the Top Ness horizon in two of the fault blocks A and D defined in Fig. 1b.

2.3 Data analysis

10 To process and analyze the large amount of interpretation data, the exported Petrel surfaces and fault sticks were wrangled using custom Python functionality and labeled in an open tabular data format. The result is a set of 4460878 data points, belonging to 78 student interpretations, with 228 unique faults considered, which consist of a total 10052 individual fault sticks. For data processing and analysis we made heavy use of the open-source Python packages *pandas*, *SciPy* and *NumPy* (McKinney, 2011; Jones et al., 2001; Oliphant, 2006).

15 For the purpose of visualization and statistical analysis of the fault interpretations across the collective interpretations, the domain was discretized into regular bins ($nx = 60$, $ny = 60$, $nz = 24$). In the following analysis we present 2-D and slices of 3-D histograms of fault interpretations, showing interpretation ~~density and~~ frequency across the domain. Fault orientations for individual faults were computed by fitting a plane (using singular-value decomposition) to all fault stick points of a single interpretation over all grid cells (collapsed along the x-axis). This makes the analysis less ~~dependant~~ dependent on the fault

20 stick interpretation density, which varies extensively in-between students. The resulting normal vector can be converted into strike and dip values. For visualization the fault orientation data was subdivided into three regular bins oriented E-W across the structures. The fault throw analysis is based on the Top Ness horizon and is computed individually at each interpreted fault stick for each fault and each interpretation. The fault throw for each interpretation is then averaged into 20 bins along the Y-axis to make the analysis independent of the number of fault sticks interpreted by each student. The nearest data points

25 on both the ~~hanging and foot wall~~ hangingwall and footwall of the Ness horizon fault blocks were selected as seeds. From these seeds the surface data approximately orthogonal to strike was used within a strike-parallel window of three grid cells ~~.~~ (approx. $570\ m$). A relative gradient filter was then used to exclude points with gradients to their nearest neighbors outside of the inner-quartile-range (IQR) of the selected subset. The resulting data is fitted with a linear regression for each fault block and the intersection with the fault stick used to calculate fault throw.

30 Throughout this work we make use of the term seismic data quality not in the strictly geophysical sense of quality factors surrounding seismic data acquisition and processing, but rather in the sense of the interpretability of the seismic data. If the seismic data lacks clear, continuous reflectors in a region, but shows a noisy image difficult to interpret—no matter what the source of this may be—we describe it as an area of low seismic data quality. Similarly, if reflector strengths are high and continuous (for horizon interpretation) or clearly offset (for fault interpretation), we speak of high seismic data quality.

For the assessment of seismic reflector strength an RMS Amplitude ~~(RMSA) seismic attribute~~ seismic attribute (RMSA for short) was calculated using Petrel~~'s RMS Amplitude function using a window length of 24 traces~~. RMS Amplitude represents the square root of the arithmetic mean of squared amplitude values across a specified seismic trace window.

3   Results
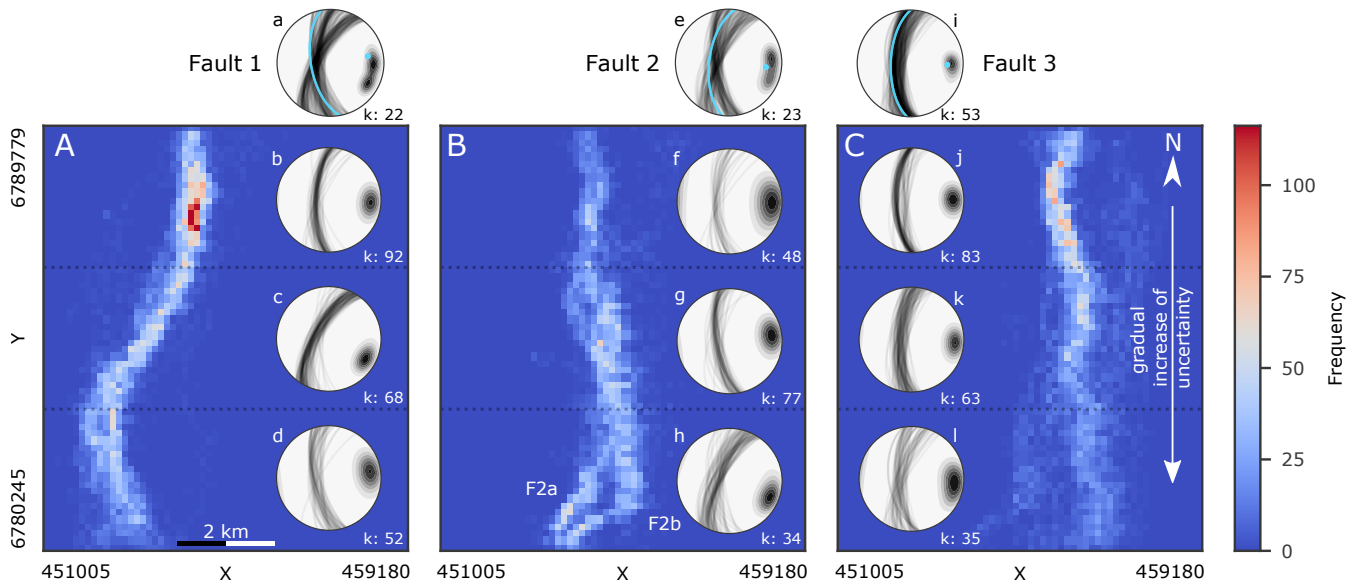
5   ### 3.1   Uncertainty in fault interpretation

Figure 3. 2D Histograms for the three Faults 1 (~~a~~A), 2 (~~b~~B) and 3 (~~e~~C) for depth slice at 2 km $\pm$ 0.1.~~Horizontal lines separate the three bins used for plotting subsets~~ Stereonet plots of Faults 1-3 (columns), with all fault strike orientations along the fault length plotted combined in the first row (a, e, i). To discriminate changing trends in actual fault orientation from interpretation uncertainty rows two to four plot data from bins separated by dashed blue lines, from the northern bin (c, g, k), middle (d, g, l) and south (e, h, m). Blue planes and poles in ~~Fig~~top row (a, e, i) show Bingham analysis mean pole from Fossen and Hesthammer (1998) for reference.~~3.~~

Figure 3 shows 2-D histograms for the three major faults taken into consideration within this study. The histograms cover the entire extent of the seismic cube, with the frequency of fault stick points counted per bin in a depth slice at 2 km $\pm$ 0.1. Fault 1 shows a sigmoidal shape in the N-S direction of the seismic depth slice (see Fig. 3~~a~~A), with high frequency densities in the northern part and lower intensities found in the southern part. Plotting all fault plane orientations within a single stereonet reveals three distinct clusters of planes (Fig. 3a), which when separated into three equal bins along the N-

10   S axis correspond to the components of the sigmoidal fault shape (Fig. 3b-d). We have added Bingham mean poles from Fossen and Hesthammer (1998) for all three faults in the plot (light blue) for comparison. Interpretations of Fault 2 show a split into two sub-faults F2a and F2b in the southern part (see Fig. 3~~b~~B), as also interpreted by Fossen and Hesthammer (1998,

see Fig. 1). The spatial uncertainty of the fault interpretations appears slightly lower in the northern part of the seismic slice, but also shows evidence of a small separated fault block in the central part of the fault. The interpretations of Fault 3 show a strong increase in ~~spread~~ dispersion towards the southern part of the seismic slice (Fig. 3~~e~~C). The same trend of increasing uncertainty can be observed in the fault plane orientation (Fig. 3j-l). Additionally, the histogram shows the occasional interpretation

5 of the fault branching towards Fault 2(b), towards the West, and towards Fault 4 to the East (see Fig. 1b). The effect of fault stick interpretation frequency between the students with below-median and above-median fault stick interpretation frequency on overall fault standard deviation was analysed using Bayesian estimation for two groups (Kruschke, 2013). We observed a difference in mean standard deviation of $35.8\ m$, $20.2\ m$ and $81.5\ m$ for Fault 1, 2 and 3 respectively, with probability of the differences being larger than zero being $99.3\,\%$, $87.4\,\%$ and $99.9\,\%$ respectively, making the differences for Fault 1 and 3

10 statistically credible.

~~Stereonet plots of Faults 1-3 (columns), with all fault strike orientations along the fault length plotted combined in the first row (a, e, i). To discriminate changing trends in actual fault orientation from interpretation uncertainty rows two to four plot data from bins shown in Fig. 3, from the northern bin (c, g, k), middle (d, g, l) and south (e, h, m). Blue planes and poles in top row (a, e, i) show Bingham analysis mean pole from Fossen and Hesthammer (1998) for reference.~~

15 Figure 3 plots fault plane orientations calculated from students fault stick interpretations. Fault 1 shows three distinct clusters of orientations (Fig. 3~~a~~A, a-d), which can be separated by subdividing the study domain into three equal bins along the N-S axis. Fault 1 shows a striking decrease in k-values (a measure of tightness of the orientation clusters) from North to South (92, 68, 52). The pattern does not hold true for Fault 2, which shows low k-values both in the North and South. Fault 3 shows similar behaviour to Fault 1, with a strong increase in dispersion from North to South (83, 63, 35).

20 Overall the observed uncertainty (standard deviation) of the fault plane along the W-E axis appears to be increasing linearly with depth for all three faults (Fig. 4). Note that the overall mean standard deviations between the faults vary greatly: 791, 384 and $575\ m$ for Faults 1, 2 and 3, respectively. The extremely high variation of standard deviations seen in the upper part of Fig. 4 (faded data points) is due to a few students extending their fault interpretations above the Base Cretaceous ~~unconformity~~ Unconformity (BCU), making the data points at that depth statistically unreliable due to low sample numbers and geologically

25 questionable. Any fault stick interpretations above the ~~unconformity~~ BCU were thus excluded from the least-squares linear regression (R-values for Fault 1, 2 and 3: 0.75, 0.85 and 0.61).

The ensemble of interpretations show 11 different fault network (FN) topologies (see Fig. 5a). Five modes of FN topology make up the bulk of fault network topologies, while ~~all others occur with a probability of less than 0.05 each~~ others were only interpreted by 3 or less students respectively. The sketches in Fig. 5b represent these five most ~~likely~~ interpreted FNs (Fault 3

30 is omitted for brevity, as it was interpreted by all students in a similar fashion, and only a single student connected Fault 2 with Fault 3). Note that the most ~~probable~~ frequent FN (Fig. 5b, A) is different from the reference expert FN interpretation of Fossen and Hesthammer (1998), which corresponds to either the second or fourth most common FN interpretations (Fig. 5b, B and D). The major source of uncertainty in this specific FN appears to be interpreting *both* F2a and F2b, and which one abuts the other. The few students who branched off the southern part of Fault 3 towards the West interpreted Fault 2b as part of Fault

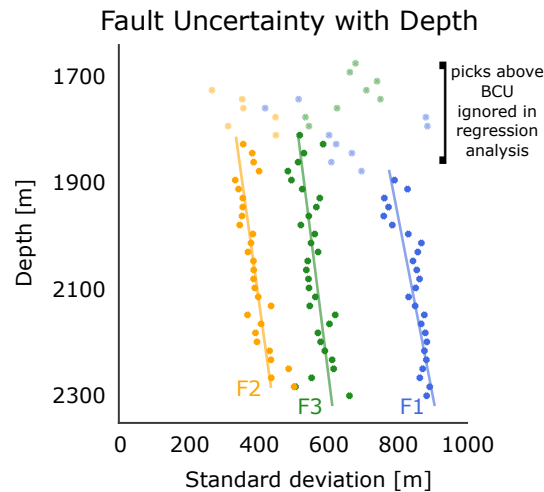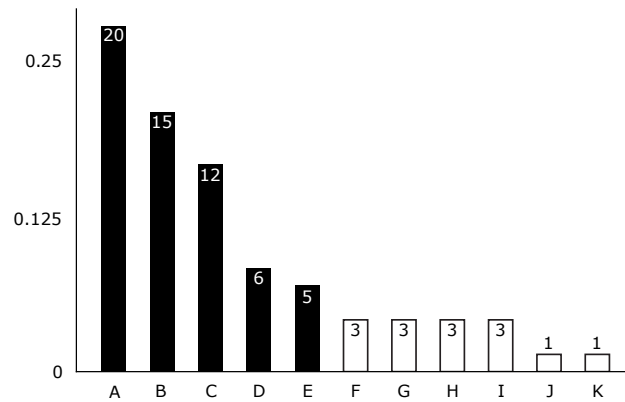35 3, but did not connect it to the FN of Fault 1 and 2.

Figure 4. Scatter plot of mean (collapsed y-axis) standard deviation along x-axis of mean fault surfaces for Fault 1, 2 and 3. Chaotic patterns of uncertainty at shallow depths (faded data points) are likely due to sporadic numbers of interpretations, as students often stopped interpreting the faults before reaching the Base Cretaceous ~~unconformity~~ Unconformity (BCU). Overall, uncertainty of fault interpretations is increasing linearly with depth (linear regression only takes into account interpretations below the BCU, R-values for Fault 1, 2 and 3 respectively: 0.75, 0.85 and 0.61).

3.2 Fault throw and horizon uncertainties

Results of the fault throw analysis are plotted in Fig. 6~~, showing~~. The boxplots show median fault throw with the associated ~~standard deviation~~ interquartile range (IQR), extrema and outlier values along fault strike direction. The throw profile of Fault 1 (Fig. 6a) shows a distinct sinuous shape spatially associated with its interaction with Fault 2. ~~Fault throw uncertainty remains~~
5 ~~relatively constant along the strike (around $\pm30\,m$), while rising sharply to about $\pm75\,m$ at the southern edge of the seismic cube~~This shows one bin with high median fault throw of approx. $180\,m$ and high fault throw uncertainty (IQR $144\,m$) before strongly decreasing in fault throw values down to a median of about $40\,m$ and one of the lowest IQR along the fault ($36\,m$). Median fault throw then rises steadily towards the South while also increasing in uncertainty (culimating in an IQR of $119\,m$ in the South). Notice the increase of uncertainty at both ends of the dataset, with increasing median throw in the South and
10 decreasing in the North. The throw profile of Fault 3 (Fig. 6b) shows two distinct levels of throw~~uncertainty. In the northern half fault throw uncertainty remains relatively constant at around $\pm36\,m$. This changes in the southern part of the fault, where fault throw uncertainty rises to up to $\pm74\,m$~~: In the Northern part of the fault median throw is high at $109\,m$, compared to $53\,m$ in Southern part. IQR increases from $40\,m$ in the North to $57\,m$ in the South, with highest IQR observed at the Southern end of the survey ($79\,m$).
15   Figure 7a shows the average Top Ness horizon basemap for all interpretations combined. Overall the horizon interpretations are increasing in depth from SE towards the NW of the domain. Figure 7b shows the associated standard deviation of the average Ness horizon interpretation, with an overlay of mean fault intersections and well locations. We observe large horizon
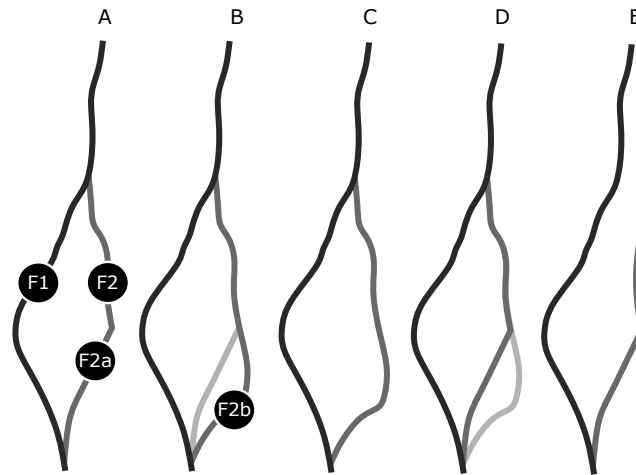
Figure 5. Probabilities of unique fault network topologies (a) with corresponding schematic fault networks (b) of the five most likely networks.

uncertainties in vicinity to both Faults 1 and ~~2.~~ 2 throughout the dataset. An increase of horizon uncertainty is occurring at the southern end of the domain where Faults 1 and 2 are beginning to merge again. In the North the horizon uncertainty surrounding Fault 3 decreases rapidly with distance from the fault, with two welltops and packages of high reflector strength (see Fig. 8A) constraining the uncertainty. As the seismic data quality decreases towards the south (see Fig. 8B), the uncertainty in the horizon interpretation increases in the Eastern part of the dataset. Interpretation uncertainties are significantly reduced surrounding well locations in the western part of the study domain, ~~while the effect is diminished towards the east, the reasoning of which is discussed later.~~ where seismic reflectors of the Top Ness horizon are overall stronger and more continuous (see Fig. 8A). This pattern does not hold true in the East of the dataset, where reflector continuity is overall low and noise in the seismic dataset is high.
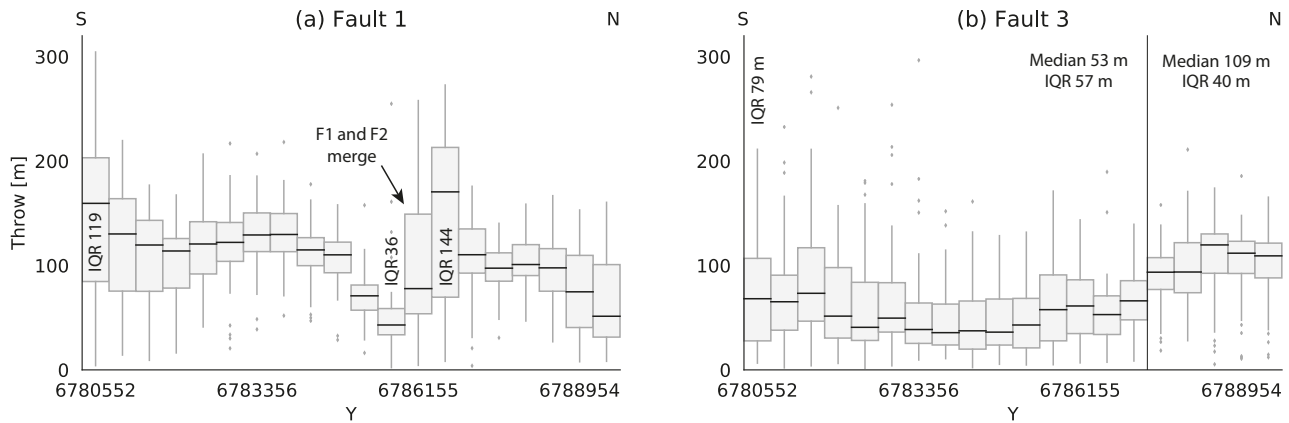
Figure 6. Results of fault throw analysis for Faults 1 (a) and 3 (b) visualized as boxplots, showing ~~the~~ median fault throw ~~in~~ (black ~~,~~ lines) with ~~respective standard deviation in~~ interquartile range (grey boxes), minima and maxima whiskers (grey lines) and outliers (grey dots).

3.3 Seismic data quality

To assess influences of seismic data quality on fault interpretation uncertainty we made use of the RMS Amplitude (RMSA) attribute as a proxy for reflector strength (strong horizon reflectors aiding the interpretation of faults result in high RMSA values). Specifically, we investigated the example of Fault 3, as it shows significant gradual changes in interpretation uncertainty across the seismic dataset (Fig. 3~~c~~ C and 6b). Figure 8 shows four averaged RMSA responses with corresponding fault stick interpretation histograms (Fig. 8a-d) from the locations shown as white boxes in Fig. 8.1. In the northern extent of the seismic slice, Fault 3 is closely bounded by strong horizon reflectors, as shown in in the RMSA slice (Fig. 8.1), seismic slice (Fig. 8.2) and inline section A, focusing the students interpretations, as seen in the corresponding histogram of fault interpretations (Fig. 8a). The histogram shows a bimodal distribution, as some students interpreted the fault further towards the East, where another fault is present (Fault 4, see Fig. 1b). The overall uncertainty related to the interpretation of the actual Fault 3 ~~appears to be Gaussian~~ is approximately normal distributed (skewness of 0.33), with the ~~standard deviation roughly corresponding to~~ width of the trough seen in the RMSA response containing more than $\pm$ 1.8 standard deviations (92.8 %) of the fault stick placement uncertainty. Further towards the south of the dataset, the RMSA response diminishes east of Fault 3, while remaining strong on the western side (Fig. 8b). The fault interpretations show a ~~stronger spread~~ 64 % increase in standard deviation, with thicker tails in the histogram (79 % increase of Pearson kurtosis), especially towards the West, where interpretations are then bounded by strong seismic reflectors. Further south the seismic response degrades and is noisy (Fig. 8.1 and B), ~~which is reflected by the homogeneous~~ leading also to a lack of signal in the RMSA values (Fig. 8c). The corresponding fault ~~interpretations show an~~ stick placements increase in uncertainty ~~,~~ (increase in standard deviation by 289 %), now appearing nearly uniformly distributed ~~but showing some consistency in the interpretation of Fault 3 (slight crest~~ with a slight crest (Fig. 8c). At the southern end of the slice, RMSA responses increase again (Fig. 8d). The distribution of fault interpretations shows a bimodal
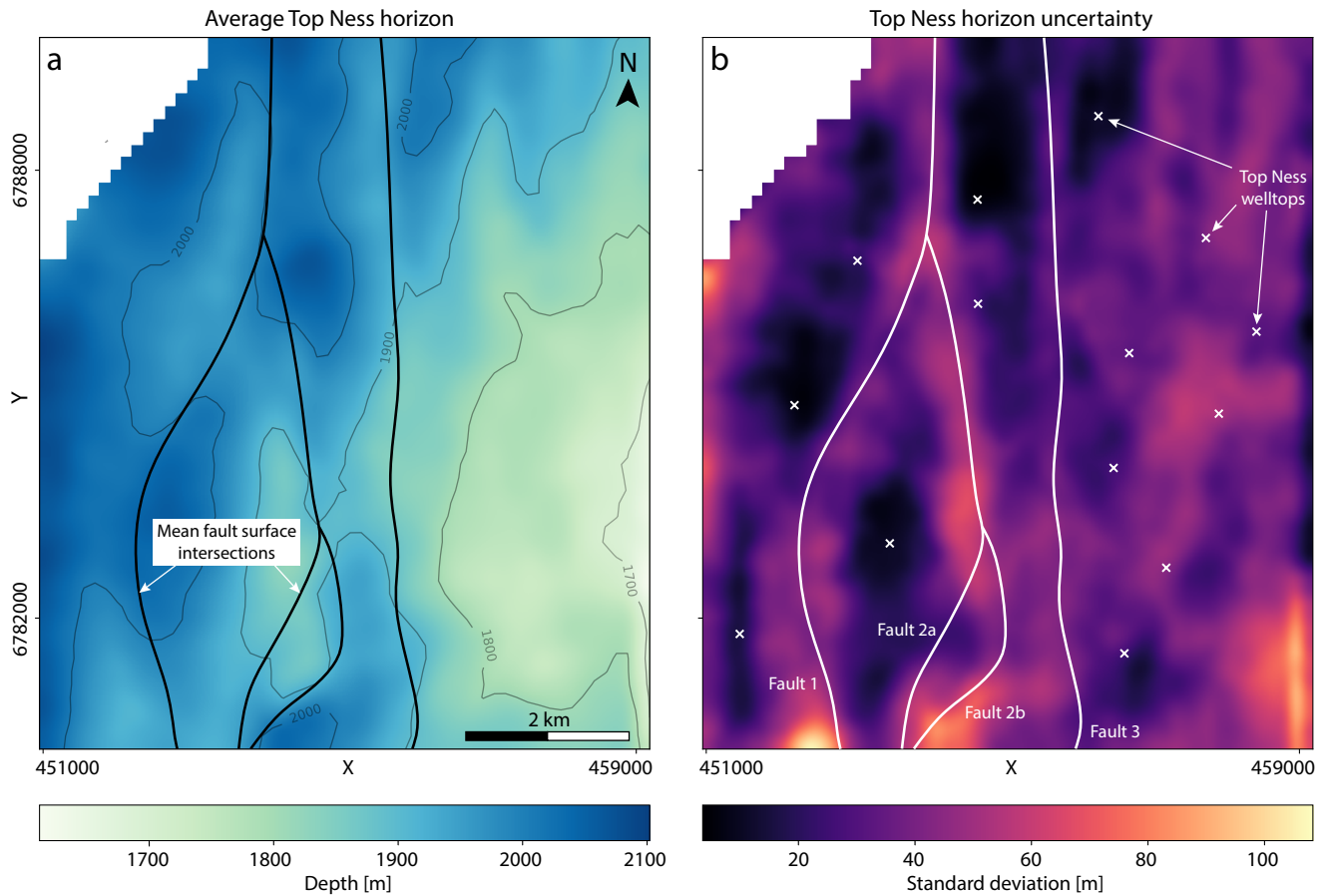
Figure 7. (a) Basemap of average Top Ness horizon from all interpretations; (b) Standard deviation of the average Top Ness horizon from all interpretations with overlaid mean fault surfaces (white lines) and welltop locations (white crosses).

distribution, as seen before in the map view 2-D histograms shown in Fig. 3. ~~The histogram roughly resemble two very broad, overlapping Normal distributions.~~

4 Discussion

4.1 Key findings

5 Our work has shown that uncertainties in fault stick placement correlate with seismic reflector strengths. In areas of high data constraint this uncertainty is strongly constrained between areas of high RMSA response (Fig. 8a). Our analysis shows how interpretation uncertainty increases with a decrease in data constraints (Fig. 8b). This trend culminates in near-maximal uncertainty in areas of ~~high seismic noise~~ low seismic image data quality (Fig. 8c). The spread in fault stick placement appears to not

11
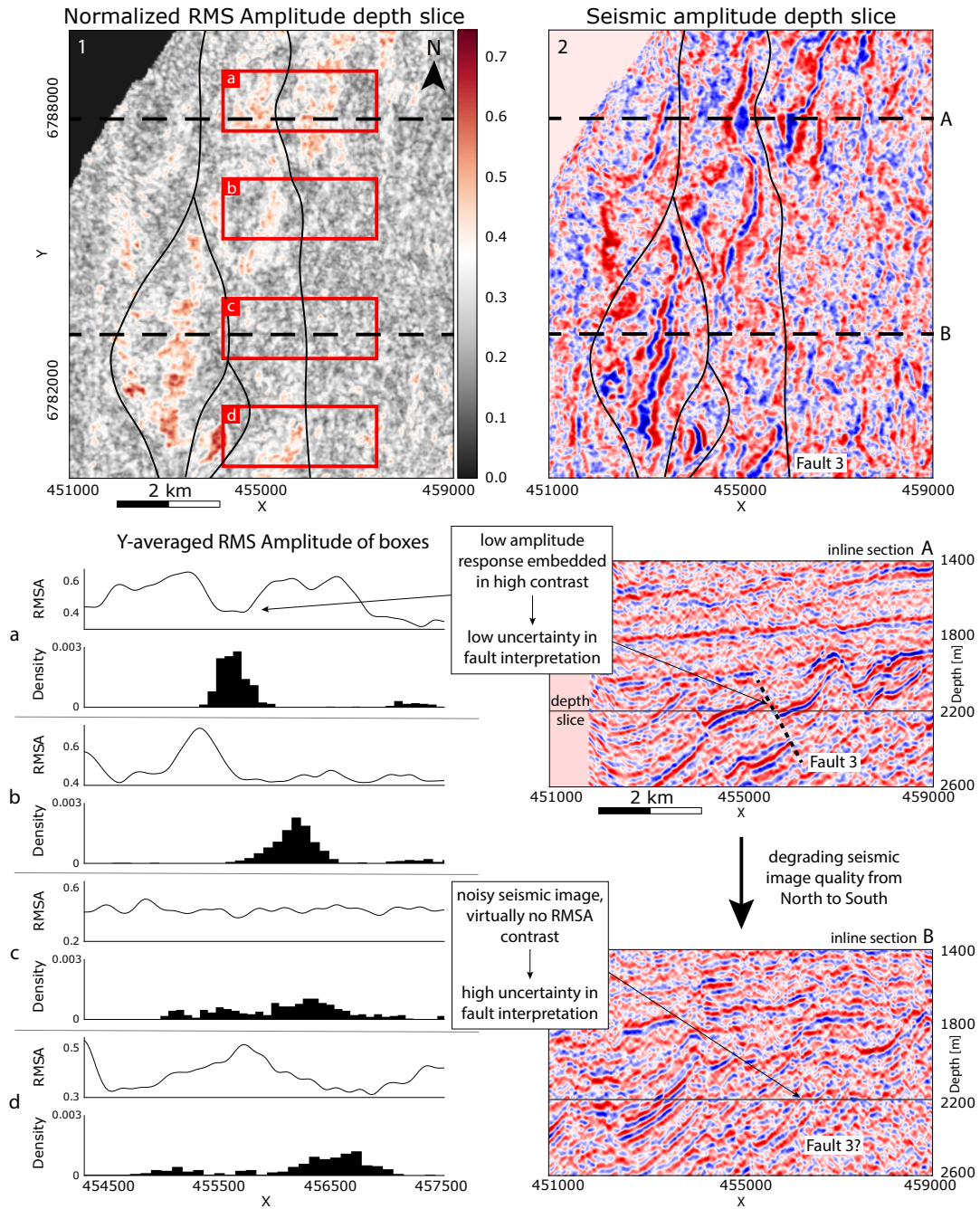
Figure 8. Depth slice of RMSA attribute (1) and seismic cube (2) with locations (black horizontal lines) of two inline sections (A, B). Four averaged RMSA responses with corresponding student Fault 3 stick interpretations are plotted in a-d, with their locations noted in the RMSA depth slice (1). The inline sections A and B correspond to the RMSA responses a and b, respectively.

be entirely driven by seismic noise, but rather appears to be, at least partly, guided by the surrounding interpretations in areas of higher data quality (Fig. 8b, d), allowing interpreted faults to conform to common fault shape models (e.g. to avoid sudden shifts in fault plane orientation). In areas of low data quality the corresponding uncertainty in fault placement ~~seems related more to~~ is also influenced by more conceptual uncertainties of fault network topology—e.g. interpretations of Fault 3 branching

5   off towards the East or West (as seen in Fig. 3~~a, d).~~ C). Recent research by Alcalde et al. (2017a) has shown the importance fault model availability plays during seismic interpretation (availability bias; see Tversky and Kahneman, 1973, 1974), which should only increase in relevance in areas of low seismic image data quality.

While the domino structure of the study area makes overall tectonic conceptual uncertainty less significant (Bond et al., 2007), the low seismic data quality makes it a challenging interpretation project, as reflector continuity is low and the dataset noisy

10   in large parts of the survey. So despite the increased information density provided by 3-D seismic surveys, significant fault network topology uncertainties remain (Fig. 5).

Our study also shows that uncertainties in the placement of faults sticks appears to increase linearly with depth (Fig. 4). This is an important finding for approximating uncertainty trends with depth, and is especially important as seismic image quality tends to decrease with depth. ~~Additional studies on different seismic datasets could yield useful information about how these~~

15   ~~fault uncertainties might depend on seismic image data quality, as seen here, and how uncertainties may be different in other tectonic and stratigraphic settings.~~

Our analysis of Top Ness horizon interpretation uncertainties shows the correlation of uncertainty with fault proximity (Fig. 7b). Horizon interpretations surrounding Fault 3 in the Northern part show only slight increases in uncertainty towards the fault, as they are strongly constrained by wells and high ~~data quality~~ reflector continuity (see Fig. 7b and 8A). With decreasing

20   seismic data quality towards the South we also see an increase in uncertainty surrounding Fault 3, which also shows in the increasing fault throw uncertainty seen if Fig. 6b. This trend is not as evident for the throw across Fault 1, with the Top Ness horizon being better constrained on both sides, with overall higher seismic image data quality. Qualitative comparison of median fault throw for both Fault 1 and 2 with depth maps in Fossen and Hesthammer (1998) displaying fault heave can be made under the simplifying assumption of constant fault dip. Fault heave qualitatively mimics the patterns we found in our

25   uncertainty study: gradual decrease from North to South for Fault 3, and the stark differences where Fault 1 and 2 merge, as well as the increase towards the South. This comparison consolidates the confidence in our automated fault throw analysis and hints at the validity of aggregating large number of interpretations of even non-experts to assess geological features.

We have also shown uncertainty in fault orientations (Fig. 3), and the inadequacy of summarizing fault orientation using a deterministic mean pole, as in Fossen and Hesthammer (1998), who calculated it using Bingham analysis of data from a

30   different 3-D seismic dataset. This inadequacy results from sinusoidal fault map pattern (Fig. 3~~a~~A) and curved 3-D geometries on-top of uncertain fault stick placements.

We ~~also see~~ observed strong decreases in uncertainty surrounding wells in the West of the study area (Fig. 7b). The trend of increasing uncertainty in horizon location from ~~west to east~~ West to East could be attributed to the decrease in seismic image data quality and the thus much lower reflector continuity. But we would also expect for the horizon uncertainties to be

35   reduced within the immediate surroundings of the wells. One possible explanation for this could be that students focused their

interpretation efforts on the higher quality western part of the seismic cube due to the time constraints on the interpretation project.

4.2 Implications for deterministic modelling

Our results and analysis suggest that the uncertainties recorded in 2-D seismic interpretation experiments (e.g. Bond et al., 2007; Bond, 2015) are similarly seen in the interpretation of 3-D seismic image data. Akin to the 2-D experiment and analysis of Alcalde et al. (2017b), we show a correlation between seismic image quality and interpretation uncertainty. We have quantified the impact of fault uncertainty on fault network topology and fault and horizon uncertainty on fault throw. The fault network topology in a 'traditional' deterministic geomodel is important as it determines the number of fault blocks and hence the degree to which stratigraphic units are separated (by faults). This information is imperative to the understanding of reservoir compartmentalization in hydrocarbon reservoirs, connectivity and flow characteristics of ground water aquifers and for geothermal projects. Simply, reservoir performance can be significantly affected by fault network topology and understanding ~~the uncertaintiesin fault network topology~~these uncertainties, and hence reservoir connectivity, can be critical to the planning of production strategies (e.g. Manzocchi et al., 2008a, b; Lescoffit and Townsend, 2005; Tveranger et al., 2008). The type of fault network topology information available in Fig. 5 could be used to inform reservoir modelling to guide multiple production strategies (e.g. multiple deterministic models are made) and for informed history matching during field operation when reservoir models have been developed from single deterministic models and need to be updated.

Although not developed in detail in this paper, our analysis of fault throw uncertainties highlight that the use of fault throw information, e.g. to predict fault sealing properties such as shale gouge ratio (Yielding, 2002; Vrolijk et al., 2016), could be significantly affected by uncertainty in the interpretation of 3-D seismic ~~image~~ cubes. Our work highlights areas where uncertainties in fault throw are likely to increase: with increasing distance from wells in ares of low reflector continuity, where seismic image quality is poor (and correspondingly with increasing depth), and where faults join or abut. Modelling of uncertainties in fault throw and using information such as that derived here to inform where uncertainties are likely to be greater could provide the basis for more informed modelling of fault seal parameters, such as through stochastic modelling discussed below, or integrated as uncertainty parameters into deterministic geo- and reservoir models. In summary, fault throw uncertainty together with fault network topology uncertainty has the potential to significantly alter predicted fluid-flow patterns in the crust with implications for water-resources, reactive element transfer (e.g. to inform nuclear waste disposal engineering), or hydrocarbon and energy production.

4.3 Implications for stochastic modelling

Although we can outline how uncertainty information could be used to better inform use of deterministic models and their inherent uncertainties, advances in both computational capabilities and implicit structural geomodelling have allowed for major improvements in the incorporation of uncertainties into structural geomodels by means of stochastic simulations. At its core stochastic structural geomodelling requires adequate disturbance distributions to obtain reasonable estimates of geomodel uncertainty (see Wellmann and Caumon, 2018). This uncertainty parametrization can be used to better establish and account

for interpretation uncertainties on-top of deterministic modelling workflows using a hybrid approach based on a single deterministic or multiple deterministic models. For example, a deterministic fault network topology model with fault throw uncertainties parameterized, or multiple deterministic models to characterize the most probable fault network topologies (e.g. Fig. 5) with fault throw uncertainties parameterized. Such hybrid approaches may provide the best solutions when the time and computational costs of full stochastic modelling are too high and/or when elements of the uncertainty in the geomodel are not best represented as simple stochastic functions, such as different conceptual models (e.g. for fault network topologies). Stamm et al. (2019) have recently explored how both fault throw and fault sealing uncertainty can be incorporated into stochastic geological modeling workflows, and studies like ours can help inform stochastic parametrization with how fault throw uncertainties can change along strike depending on changes in the seismic data quality.

The work of Pakyuz-Charrier et al. (2018) discusses the importance of proper parametrization of input data measurement uncertainty when constructing stochastic geomodels, but little is known about the uncertainties in interpreting the dense 3-D seismic datasets to obtain such input data for structural geomodels. Our work not only provides a first look at how significant these uncertainties can be, but additionally provides a first order approximation for parametrization of fault and horizon interpretation uncertainties within stochastic geomodels . Student based on seismic image quality surrounding fault interpretations. Future research into how our findings could be integrated with the seismic expression of fault zones (e.g. Botter et al., 2014; Iacopini et al., 2016) could further our ability to parametrize stochastic geomodels directly from seismic data.

While student interpretations will most certainly reside within the upper range of interpretation uncertainty, but we argue that it nevertheless provides significant value to the parametrizationof stochastic parametrization, especially to the emergent Bayesian approaches to stochastic structural geomodelling (Caers, 2011; de la Varga and Wellmann, 2016; Wellmann et al., 2017), as it could provide informed—but not overly constrained—prior parametrization that can be reduced by case-specific geological likelihood functions and auxiliary data integration. We also agree with Caers (2018) on the need for a rigorous methodology of falsification in geomodelling. The integration of adequately parameterized seismic interpretation uncertainties into a Bayesian geomodelling framework could enable a quality control of the interpretation by probabilistic assessment of the stochastic geomodel against geological likelihood functions (e.g. fault length and throw relationships, fault population distributions, analog studies).

Our findings underline the complexity involved in the adequate parametrization of interpretation uncertainty in stochastic geomodelling: while Normal distributions may capture uncertainty adequately in areas of good seismic imaging (Fig. 8a), skewed fat-tailed distributions (e.g. Cauchy; Fig. 8b) or even Uniform distributions (Fig. 8c) are reasonable choices with degrading seismic quality. When implicitly modelling 3-D geological surfaces, many approaches are based on both surface points and strike and dip information. The latter carry significantly higher amounts of information (Calcagno et al., 2008; Laurent et al., 2016; Grose et al., 2017) than the surface points, thus emphasizing the need to quantify their uncertainty if it is used to generate and constrain 3-D stochastic geomodels. The use of van Mises-Fisher distributions to model uncertainty of orientation vectors was shown as a robust way to describe surface orientation uncertainty (Pakyuz-Charrier et al., 2018),

and our analysis could provide valuable information for their parametrization in areas of high interpretation uncertainty within sedimentary basins.

4.4 Implications for machine learning

Recent efforts in automating seismic interpretation through the use of, mainly, Neural Networks (NN) has been increasingly successful in interpreting high quality (synthetic) seismic data (e.g. Huang et al., 2017; Dramsch and Lüthje, 2018; Wu et al., 2018). But NNs inherently do not take into account any geological reasoning skills that could make sense of areas of low seismic image data quality, but rather infer abstract features from their training data. And while NNs can be constructed and trained probabilistically, and thus enable uncertainty quantification of their outputs, they are likely to require additional information about the structures to be interpreted in areas where interpretation suffers from low data quality. In our analysis of the uncertainties in interpretations of the Gullfaks field 3-D seismic cube the interpretations of faults in areas of low seismic image quality conform to known fault geometries and topologies and are likely informed by interpretations of adjacent higher quality seismic image data, rather than their uncertainty distributions being simply correlated to seismic image quality (Fig. 8). Such evidence highlights the nuances in geological interpretation, and the difficulties in creating algorithms that represent the complexities of human thought processes. Irrespective of how easy it is, or otherwise, to apply our findings to inform automated interpretation efforts, there is value in studies such as the one presented here in generating understanding as to when and where such machine learning processes maybe applied effectively and how they could be improved by integrating geological knowledge, maybe in the form of *logic rules* that influence NN weights and biases (Hu et al., 2016). Lu et al. (2018) show the use of Generative Adversarial Networks (GANs) in improving fault interpretation of low-resolution seismic image data by generating supersampled high-quality seismic images from lower quality data. One possibility could be to train similar NNs on the structural geology represented by ideal synthetic seismic image data and artificially noisy perturbations of the same to let NNs learn how ideal structures that underlie noisy seismic image data might look like. They could then possibly generate possible higher-quality realizations of noisy, uncertain areas of seismic images to support interpretation efforts. Overall, the complex interplay of the underlying geology, computational and conceptual challenges of the machine learning approaches and the human-induced uncertainties will require strongly interdisciplinary approaches to combine state-of-the-art algorithms and geological domain knowledge to further automate the laborious and uncertain task of 3-D seismic interpretation.

5 Summary

Our study provides a first look and quantification of the scale of uncertainties involved in the structural interpretation of data-dense 3-D seismic data. We have found that:

– Fault placement uncertainty shows strong dependency on seismic data quality. The use of a seismic attribute (RMS Amplitude) showed promising first results to be used as a proxy for estimating fault interpretation uncertainties. This can be especially valuable for the parametrization of stochastic geomodels based on single seismic reference interpretations.

- The common use of Normal distributions as perturbance distributions in stochastic geomodelling seems inadequate in areas where interpretation uncertainty is high (low seismic data quality). Instead, uncertainty parametrization should always be directly linked to the surrounding seismic data quality and the use of near-uniform distributions can be recommended in areas of extremely poor interpretability.

5
- We found that relative trends of fault placement uncertainty can be approximated linearly with depth, although we recommend more research to further investigate the influence of seismic data quality in combination with changes in reflector continuity with depth.

- The student interpretations displayed significant uncertainties in fault network topologies—despite the information-dense 3-D seismic data—which can have critical impacts on decision making based on geomodels constructed from seismic
10
interpretations.

Additional interpretation uncertainty studies on different seismic data are recommended to further our knowledge about how interpretation uncertainties depend on seismic image data quality, especially in different tectonic and stratigraphic settings, and to further the integration of uncertainties into future geomodelling efforts and decision making.

References

Abrahamsen, P., Lia, O., and Omre, H.: An Integrated Approach to Prediction of Hydrocarbon in Place and Recoverable Reserve With Uncertainty Measures, in: European Petroleum Computer Conference, Society of Petroleum Engineers, https://doi.org/10.2118/24276-MS, 1992.

Alcalde, J., Bond, C. E., Johnson, G., Butler, R. W., Cooper, M. A., and Ellis, J. F.: The Importance of Structural Model Availability on Seismic Interpretation, Journal of Structural Geology, 97, 161–171, https://doi.org/10.1016/j.jsg.2017.03.003, 2017a.

Alcalde, J., Bond, C. E., Johnson, G., Ellis, J. F., and Butler, R. W.: Impact of Seismic Image Quality on Fault Interpretation Uncertainty, GSA Today, https://doi.org/10.1130/GSATG282A.1, 2017b.

Alcalde, J., Bond, C. E., and Randle, C. H.: Framing Bias: The Effect of Figure Presentation on Seismic Interpretation, Interpretation, 5, T591–T605, 2017c.

Biondi, B.: 3D Seismic Imaging, Investigations in Geophysics, Society of Exploration Geophysicists, https://doi.org/10.1190/1.9781560801689, 2006.

Bond, C., Gibbs, A., Shipton, Z., and Jones, S.: What Do You Think This Is? "Conceptual Uncertainty" in Geoscience Interpretation, GSA Today, 17, 4, https://doi.org/10.1130/GSAT01711A.1, 2007.

Bond, C. E.: Uncertainty in Structural Interpretation: Lessons to Be Learnt, Journal of Structural Geology, 74, 185–200, https://doi.org/10.1016/j.jsg.2015.03.003, 2015.

Bond, C. E., Philo, C., and Shipton, Z. K.: When There Isn't a Right Answer: Interpretation and Reasoning, Key Skills for Twenty-First Century Geoscience, International Journal of Science Education, 33, 629–652, https://doi.org/10.1080/09500691003660364, 2011.

Botter, C., Cardozo, N., Hardy, S., Lecomte, I., and Escalona, A.: From Mechanical Modeling to Seismic Imaging of Faults: A Synthetic Workflow to Study the Impact of Faults on Seismic, Marine and Petroleum Geology, 57, 187–207, https://doi.org/10.1016/j.marpetgeo.2014.05.013, 2014.

Caers, J.: Modeling Uncertainty in the Earth Sciences: Caers/Modeling Uncertainty in the Earth Sciences, John Wiley & Sons, Ltd, Chichester, UK, https://doi.org/10.1002/9781119995920, 2011.

Caers, J.: Bayesianism in the Geosciences, in: Handbook of Mathematical Geosciences: Fifty Years of IAMG, edited by Daya Sagar, B., Cheng, Q., and Agterberg, F., pp. 527–566, Springer International Publishing, Cham, https://doi.org/10.1007/978-3-319-78999-6_27, 2018.

Calcagno, P., Chilès, J. P., Courrioux, G., and Guillen, A.: Geological Modelling from Field Data and Geological Knowledge: Part I. Modelling Method Coupling 3D Potential-Field Interpolation and Geological Rules, Physics of the Earth and Planetary Interiors, 171, 147–157, https://doi.org/10.1016/j.pepi.2008.06.013, 2008.

de la Varga, M. and Wellmann, J. F.: Structural Geologic Modeling as an Inference Problem: A Bayesian Perspective, Interpretation, 4, SM1–SM16, https://doi.org/10.1190/INT-2015-0188.1, 2016.

Dramsch, J. and Lüthje, M.: Deep-Learning Seismic Facies on State-of-the-Art CNN Architectures, in: SEG Technical Program Expanded Abstracts 2018, SEG Technical Program Expanded Abstracts, pp. 2036–2040, Society of Exploration Geophysicists, https://doi.org/10.1190/segam2018-2996783.1, 2018.

Fossen, H. and Hesthammer, J.: Structural Geology of the Gullfaks Field, Northern North Sea, Geological Society, London, Special Publications, 127, 231–261, https://doi.org/10.1144/GSL.SP.1998.127.01.16, 1998.

Freeman, B., Yielding, G., and Badley, M.: Fault Correlation during Seismic Interpretation, First Break, 8, 1990.

Frodeman, R.: Geological Reasoning: Geology as an Interpretive and Historical Science, Geological Society of America Bulletin, 107, 960–0968, https://doi.org/10.1130/0016-7606(1995)107<0960:GRGAAI>2.3.CO;2, 1995.

Godefroy, G., Laurent, G., and Bonneau, F.: Structural Interpretation of Sparse Fault Data Using Graph Theory and Geological Rules, p. 15, 2018.

5 Grose, L., Laurent, G., Aillères, L., Armit, R., Jessell, M., and Caumon, G.: Structural Data Constraints for Implicit Modeling of Folds, Journal of Structural Geology, 104, 80–92, https://doi.org/10.1016/j.jsg.2017.09.013, 2017.

Hesthammer, J. and Fossen, H.: Seismic Attribute Analysis in Structural Interpretation of the Gullfaks Field, Northern North Sea | Petroleum Geoscience, Petroleum Geoscience, 3, 13–26, 1997.

Hu, Z., Ma, X., Liu, Z., Hovy, E., and Xing, E.: Harnessing Deep Neural Networks with Logic Rules, arXiv:1603.06318 [cs, stat], 2016.

10 Huang, L., Dong, X., and Clee, T. E.: A Scalable Deep Learning Platform for Identifying Geologic Features from Seismic Attributes, The Leading Edge, 36, 249–256, https://doi.org/10.1190/tle36030249.1, 2017.

Iacopini, D., Butler, R., Purves, S., McArdle, N., and De Freslon, N.: Exploring the Seismic Expression of Fault Zones in 3D Seismic Volumes, Journal of Structural Geology, 89, 54–73, https://doi.org/10.1016/j.jsg.2016.05.005, 2016.

Jones, E., Oliphant, T., Peterson, P., et al.: SciPy: Open Source Scientific Tools for Python, 2001.

15 Kruschke, J. K.: Bayesian Estimation Supersedes the t Test, Journal of Experimental Psychology: General, 142, 573–603, https://doi.org/10.1037/a0029146, 2013.

Laurent, G., Ailleres, L., Grose, L., Caumon, G., Jessell, M., and Armit, R.: Implicit Modeling of Folds and Overprinting Deformation, Earth and Planetary Science Letters, 456, 26–38, https://doi.org/10.1016/j.epsl.2016.09.040, 2016.

Lescoffit, G. and Townsend, C.: Quantifying the Impact of Fault Modeling Parameters on Production Forecasting for Clastic Reservoirs, pp.
20 137–149, https://doi.org/10.1306/1060761H23166, 2005.

Lu, P., Morris, M., Brazell, S., Comiskey, C., and Xiao, Y.: Using Generative Adversarial Networks to Improve Deep-Learning Fault Inter-pretation Networks, The Leading Edge, 37, 578–583, https://doi.org/10.1190/tle37080578.1, 2018.

Macrae, E. J.: Uncertainty in Geoscience Interpretation: Statistical Quantification of the Factors That Affect Interpretational Ability and Their Application to the Oil and Gas Industry, p. 264, 2013.

25 Macrae, E. J., Bond, C. E., Shipton, Z. K., and Lunn, R. J.: Increasing the Quality of Seismic Interpretation, Interpretation, 4, T395–T402, https://doi.org/10.1190/INT-2015-0218.1, 2016.

Manzocchi, T., Carter, J. N., Skorstad, A., Fjellvoll, B., Stephen, K. D., Howell, J. A., Matthews, J. D., Walsh, J. J., Nepveu, M., Bos, C., Cole, J., Egberts, P., Flint, S., Hern, C., Holden, L., Hovland, H., Jackson, H., Kolbjørnsen, O., MacDonald, A., Nell, P. a. R., Onyeagoro, K., Strand, J., Syversveen, A. R., Tchistiakov, A., Yang, C., Yielding, G., and Zimmerman, R. W.: Sensitivity of the Impact of Geological
30 Uncertainty on Production from Faulted and Unfaulted Shallow-Marine Oil Reservoirs: Objectives and Methods, Petroleum Geoscience, 14, 3–15, https://doi.org/10.1144/1354-079307-790, 2008a.

Manzocchi, T., Matthews, J. D., Strand, J. A., Carter, J. N., Skorstad, A., Howell, J. A., Stephen, K. D., and Walsh, J. J.: A Study of the Structural Controls on Oil Recovery from Shallow-Marine Reservoirs, Petroleum Geoscience, 14, 55–70, https://doi.org/10.1144/1354-079307-786, 2008b.

35 McKinney, W.: Pandas: A Foundational Python Library for Data Analysis and Statistics, Python for High Performance and Scientific Com-puting, pp. 1–9, 2011.

Morley, C. K. and Nixon, C. W.: Topological Characteristics of Simple and Complex Normal Fault Networks, Journal of Structural Geology, 84, 68–84, https://doi.org/10.1016/j.jsg.2016.01.005, 2016.

Oliphant, T. E.: A Guide to NumPy, vol. 1, Trelgol Publishing USA, 2006.

Pakyuz-Charrier, E., Lindsay, M., Ogarko, V., Giraud, J., and Jessell, M.: Monte Carlo Simulation for Uncertainty Estimation on Structural Data in Implicit 3-D Geological Modeling, a Guide for Disturbance Distribution Selection and Parameterization, Solid Earth, 9, 385–402, https://doi.org/https://doi.org/10.5194/se-9-385-2018, 2018.

5 Peacock, D., Nixon, C., Rotevatn, A., Sanderson, D., and Zuluaga, L.: Interacting Faults, Journal of Structural Geology, 97, 1–22, https://doi.org/10.1016/j.jsg.2017.02.008, 2017.

Peacock, D. C. P., Nixon, C. W., Rotevatn, A., Sanderson, D. J., and Zuluaga, L. F.: Glossary of Fault and Other Fracture Networks, Journal of Structural Geology, 92, 12–29, https://doi.org/10.1016/j.jsg.2016.09.008, 2016.

Stamm, F. A., de la Varga, M., and Wellmann, F.: Actors, Actions and Uncertainties: Optimizing Decision Making Based on 3-D Structural
10   Geological Models, Solid Earth Discussions, pp. 1–35, https://doi.org/https://doi.org/10.5194/se-2019-57, 2019.

Tannert, C., Elvers, H.-D., and Jandrig, B.: The Ethics of Uncertainty: In the Light of Possible Dangers, Research Becomes a Moral Duty, EMBO reports, 8, 892–896, https://doi.org/10.1038/sj.embor.7401072, 2007.

Thiele, S. T., Jessell, M. W., Lindsay, M., Wellmann, J. F., and Pakyuz-Charrier, E.: The Topology of Geology 2: Topological Uncertainty, Journal of Structural Geology, 91, 74–87, https://doi.org/10.1016/j.jsg.2016.08.010, 2016.

15 Thore, P., Shtuka, A., Lecour, M., Ait-Ettajer, T., and Cognot, R.: Structural Uncertainties: Determination, Management, and Applications, GEOPHYSICS, 67, 840–852, https://doi.org/10.1190/1.1484528, 2002.

Tveranger, J., Howell, J., Aanonsen, S. I., Kolbjørnsen, O., Semshaug, S. L., Skorstad, A., and Ottesen, S.: Assessing Structural Controls on Reservoir Performance in Different Stratigraphic Settings, Geological Society, London, Special Publications, 309, 51–66, https://doi.org/10.1144/SP309.4, 2008.

20 Tversky, A. and Kahneman, D.: Availability: A Heuristic for Judging Frequency and Probability, Cognitive Psychology, 5, 207–232, https://doi.org/10.1016/0010-0285(73)90033-9, 1973.

Tversky, A. and Kahneman, D.: Judgment under Uncertainty: Heuristics and Biases, Science, 185, 1124–1131, https://doi.org/10.1126/science.185.4157.1124, 1974.

Vrolijk, P. J., Urai, J. L., and Kettermann, M.: Clay Smear: Review of Mechanisms and Applications, Journal of Structural Geology, 86,
25   95–152, https://doi.org/10.1016/j.jsg.2015.09.006, 2016.

Wellmann, J. F. and Caumon, G.: 3-D Structural Geological Models: Concepts, Methods, and Uncertainties, in: Advances in Geophysics, vol. 59, 1st edn., 2018.

Wellmann, J. F., de la Varga, M., Murdie, R. E., Gessner, K., and Jessell, M.: Uncertainty Estimation for a Geological Model of the Sandstone Greenstone Belt, Western Australia – Insights from Integrated Geological and Geophysical Inversion in a Bayesian Inference Framework,
30   Geological Society, London, Special Publications, p. SP453.12, https://doi.org/10.1144/SP453.12, 2017.

Wu, X., Shi, Y., Fomel, S., and Liang, L.: Convolutional Neural Networks for Fault Interpretation in Seismic Images, in: SEG Technical Program Expanded Abstracts 2018, SEG Technical Program Expanded Abstracts, pp. 1946–1950, Society of Exploration Geophysicists, https://doi.org/10.1190/segam2018-2995341.1, 2018.

Yielding, G.: Shale Gouge Ratio — Calibration by Geohistory, in: Norwegian Petroleum Society Special Publications, edited by Koestler,
35   A. G. and Hunsdale, R., vol. 11 of *Hydrocarbon Seal Quantification*, pp. 1–15, Elsevier, https://doi.org/10.1016/S0928-8937(02)80003-0, 2002.