General comments

The manuscript by Schaaf and Bond 2019 describes an interesting and important concept of 3D seismic interpretation uncertainty. The authors analysed an extensive data set, comprising 78 seismic interpretations carried out by final year undergraduate students of a 3D seismic dataset from the Gullfaks Field, northern North Sea. Uncertainties related to fault and horizon interpretations are explored and quantified using Python tools. The results of the research are then discussed with implications for deterministic and stochastic geomodelling as well as machine learning.

Overall, I believe the paper would be of great interest for the readers, and it could be accepted but with strong modifications. There are some instances where the scientific approach was oversimplified and didn't actually consider other controlling factors. I appreciate the general discussion done but I believe that the 'Key finding' section of the discussing part could be significantly modified and discussed in more depth. The manuscript also lacks a clear and concise summary or conclusions. The above general comments are discussed in further detail below and/or in the attached PDF file of the manuscript with my comments highlighted. For revision, please use both files. I look forward to reading the revised final version of manuscript.

Major comments

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

a) 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.

b) 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.

c) 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 hangingwall 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. also, 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.

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.

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.

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.

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.

Detailed comments

P1 – Abstract

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

1. 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.

L12 – Please add the name of the 3D seismic survey

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

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).

P2 – Section 2.1 Gullfakes 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).

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 domino-system (research focus) is comparatively simple.

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.

P3 – Section 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.

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

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

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

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

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

P4 – Section 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.

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

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.

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

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?

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).

P8 – Section 3.2 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?

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.

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.

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.

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.

P10 – Section 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 dose 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. 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. 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.

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