HOW FAULT INTERPRETATION METHOD MAY INFLUENCE THE ASSESSMENT OF A FAULT-BOUND CO2 STORAGE SITE

E.A.H.Michie*[1], A. Braathen[1], B. Alaei[2]

[1] University of Oslo, Oslo, Norway

[2] Earth Science Analytics, Oslo, Norway

* Corresponding author e-mail: e.m.haines@geo.uio.no

Abstract

Interpretation of faults in the subsurface hinges on utilizing an optimum picking strategy, i.e. the seismic line spacing. Differences in line spacing lead to significant changes in subsequent fault analyses such as fault growth, fault seal and fault stability, all of which are crucial when analyzing a fault-bound CO_2 storage site. With the ever-advancing technologies, machine learning techniques, such as Deep Neural Networks (DNN), used for fault extraction are becoming increasingly common, however their limitations and corresponding uncertainty is still largely unknown. Here, we show how fault extraction using DNN compares with faults that have been picked manually, and with using different line spacing. Uncertainty related to both manual and automated fault extraction methods are heavily reliant on seismic quality. As such, faults that are well-imaged show a closer similarity between those that have been manually picked and automatically extracted. In cases of poorly imaged faults, DNN picking on narrower line spacing creates a fault that is more irregular and with a lower predicted stability than the smoother and simpler fault model created by manual picking. Thereby, DNN creates fault surfaces that are less stable than those that have been picked manually, which is assumed to be associated with the increased irregularity of the fault segments. We conclude that fault picking by DNN without in-depth expertise works for well-imaged faults; poorly imaged faults require additional considerations and quality control for both manually and DNN picked faults.

Keywords: *Fault interpretation, fault stability, machine learning, uncertainty, CO_2 storage*

1. Introduction

In order to achieve targets to reduce emissions of greenhouse gases as outlined by the European Commission [1], methods of carbon capture and storage can be utilized to reach the 2°C goal of the Paris Agreement [2]. One candidate for a CO_2 storage site has been identified in the Norwegian North Sea, which is the focus of this study: the saline aquifer in the Sognefjord Formation at the Smeaheia site [3] (Figure 1). The Alpha prospect identified for this site is located within a tilted fault block bound by a deep-seated basement fault: the Vette Fault Zone (VFZ) [3], and hence a high fault sealing capacity is required to retain the injected CO_2. Further, it is necessary for the fault to have no reactivation potential. Both of these parameters hinge on generating an accurate geological model, performed using suitable picking strategies.

The process by which seismic is interpreted has developed significantly over the years. Initially, seismic interpretation involved the manual picking using printed seismic sections [4], which has since developed to provide users with the ability to interpret using a suite of digital environments [5]. The ease and accuracy of seismic interpretation is continually increasing, associated with advanced geophysical and rock physics tools, as well as the increased use of automated technologies. While technology has progressed to allow user to quickly interpret horizons using facilities such as auto-tracking, the ability for machine learned algorithms for automated fault extraction has, until recently, been lacking. New technology has emerged that uses Deep Learning (i.e. Deep Neural Networks inspired machine learning) to automatically extract faults from seismic, with minimal manual seismic fault interpretation [6]. However, it is crucial to understand any uncertainties when using these automated methods, and how they may impact any further fault analyses.

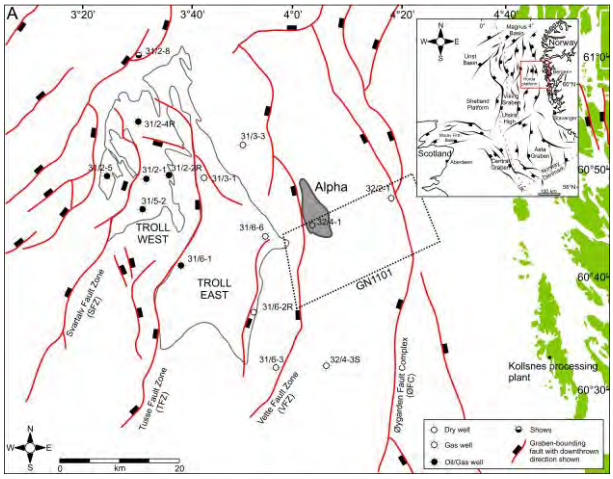

Figure 1. Location of the potential CO2 storage site, known as Smeaheia.

2. Methodology

In this contribution, we compare manual fault interpretation with that from supervised DNN fault extraction for the prospect-bounding Vette Fault within the Smeaheia potential CO_2 storage site. Further, we also compare differences between interpretation using different line spacing, as this has proven to be crucial for in-depth fault analyses such as fault stability analysis. Specifically, we have examined how the predicted dilation tendency varies when faults are picked on every line (25 m), 2nd line (50 m), 4th line (100 m), 8th line (200 m), 16th line (400 m) and 32nd line (800 m). Dilation tendency is the relative probability of a plane to dilate within the current stress field (Table 1), taking into consideration the cohesion and frictional coefficient of the fault rock, which are set as 0.5 MPa and 0.45, respectively. Dilation tendency is a ratio between 0 and 1, where the higher the value, the more likely a fault will go into tensile failure.

	Gradient (MPa/m)	Stress (MPa)	Depth (m)	Direction (degrees)
SHmin	0.0146	23.07	1699.5	090
SHmax	0.0146	23.07	1699.5	180
Sv	0.0215	32.37	1699.5	
PP	0.01	16.94	1699.5	

Table 1: *In situ* stress data used for geomechanical analysis, from [6].

GN1101 3D seismic survey was used in this study. The survey has a 25x12.5 m inline and cross line spacing and 4 ms vertical sampling interval. Data is prestack time migrated. Data quality is overall good. Seismic imaging challenges caused by large faults (i.e. the Vette Fault Zone) caused some poor imaging around fault zones. Wells around the survey were used to build a simple velocity model for depth conversion purposes.

3. Results

3.1 Manual Fault Interpretation

Although fault stability is influenced by external factors, specifically the *in situ* stress conditions, it is also heavily influenced by intrinsic fault attributes, namely strike and dip. Since the stress conditions used in this study are isotropic, fault dip has a primary control on fault stability over fault strike. Here, we show how fault dip, and hence geomechanical analysis, varies with picking strategy.

3.1.1. Dip

Fault dip varies down the VFZ. There is low fault dip within the top 1000 m, particularly in the Northern section, where the fault penetrates younger stratigraphy, specifically the Cromer Knoll and the Shetland Groups. Here, the dip decreases to approximately 35 degrees, but can be as low as 15 degrees at the very top of the fault (Figure 2). The fault then steepens in dip to approximately 70 degrees at 1500 – 4000 m depth, beyond which the dip decreases again to approximately 40 degrees at the base of the fault.

Fault dip is also shown to vary according to picking strategy. The shallowly dipping portion at the top of the fault is smoothed with increasing picking distance, such that the lowest dip for faults picked on every 400 m and 800 m line spacing is 35 degrees. However, the shallowest dip for faults picked on every 25 m and 50 m line spacing is 15 degrees. Further, small, bulls-eye areas of steeper dip are also removed and smoothed when line spacing is increased (Figure 2, red circles). Similarly, the steeper portion of the fault is smoothed as the line spacing used for picking is increased. This decreases the range of dips, and smooths any bulls-eye patches of steeper or shallower dip (Figure 2, black circles).

Although rigorous quality control has been performed to improve continuity between each inline, there remains several places where slight differences in picking has occurred between lines. This human error leads to an increased irregularity of the fault surface, often creating these bulls-eye areas of inconsistent dip, associated with the triangulation algorithm trying to honour each point along the fault segments. Since fault stability is influenced by fault dip, these areas will be brought through to geomechanical modelling. The uneven nature of the fault surface is most severe when every inline line has been picked on. The irregularity decreases with increased picking spacing.

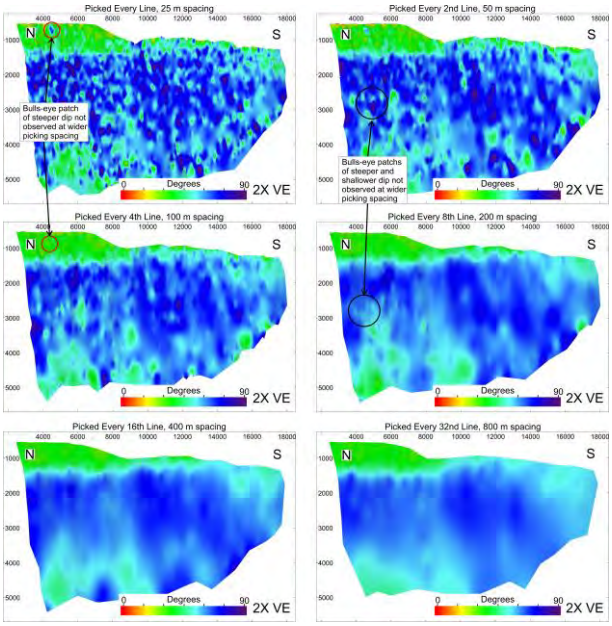

Figure 2. Fault plane diagrams showing fault dip attribute displayed on the fault surfaces for each manual picking strategy: 1, 2, 4, 8, 16 and 32 lines. Fault dip is observed to vary with line spacing used for fault picking. A highly irregular fault surface is observed when every line is used for picking, when compared to the overly smooth surface when every 32nd line is used for picking. X- and y-axes represent the lateral and vertical scales in metres, respectively. Note that unconstrained triangulation is used for fault surface generation.

3.1.2. Dilation Tendency

Since dip varies with picking strategy, as does the predicted fault stability. Along fault-strike there are minor patches where the fault is more stable than the surrounding values and patches where the fault is less stable. These patches are most apparent when every line is picked on, with irregularity decreasing in severity until every 100 m to 200 m line spacing is used for picking, where the frequency of these irregular patches is reduced. Since the fault surface is smoothed with greater picking spacing (i.e. >200 m line spacing), the results for fault stability are also smoothed, reducing the range of values of the predicted dilation tendency (Figure 3). Hence, interpretation of fault stability will vary with picking strategy, and may in fact lead to incorrect fault stability assumptions. For example, areas where the fault is close to failure are only observed when a narrower line spacing picking strategy is used (Figure 3). These areas are smoothed out and not visible when a coarser line spacing picking strategy is used. However, if these irregular areas are not a product of human error or triangulation method, the overall stability would be overestimated within this location if a coarser line spacing was used. Patches where the fault is more or less stable than the average surrounding values that occur when a narrower line spacing picking strategy is used, could be a product of human error and/or triangulation method, but may also in fact be geologically plausible due to the inherent irregularity of faults in nature. Therefore, a question is presented regarding optimum picking strategy that retains sufficient detail but remove any data that is caused by human error and/or triangulation method. Human error combined with triangulation method complications are most apparent at narrower line spacing. Conversely, over-smoothing occurs at coarser line spacing. For the studied fault system, we propose an optimum picking strategy of every 100 m line spacing. Best-practice is however likely to be case-dependent.

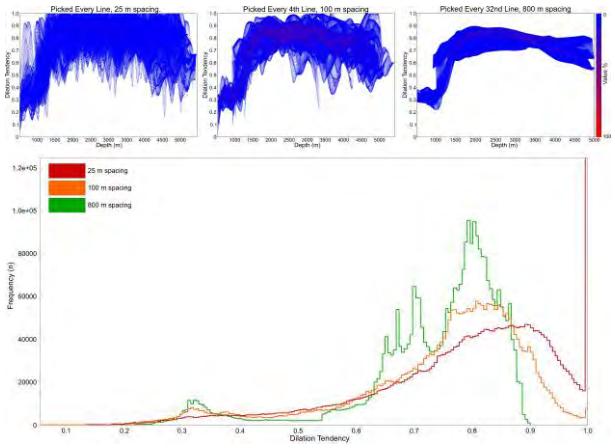

Figure 3. Top: plots showing dilation tendency with depth, for scenarios manually picked on every line (left), every 4th line (middle) and every 32nd line (right). Colour intensity reflects the frequency of those values, where blue is 1% and red is 100% frequency. Bottom: Histogram showing frequency of dilation tendency for scenarios picked on every line (red), every 4th lines (orange) and every 32nd line (green). Note that when every line is picked, a large portion of the values are above 1 (i.e. in failure). This decreases as the spacing decreases.

3.2. Deep Learning Fault Interpretation

For the DNN approach, we used supervised learning which means providing some fault picks to be used for training with seismic full stack data. The trained DNN models are checked against input picked faults through confusion matrix and visual review. We have also predicted faults using a pretrained fault model which trained on 20 surveys (excluding the GN1101 survey). We have then generated ensemble fault results of four DNN models. This enables us to see the confidence of predicted faults. Such information is useful when used to verify fault geometric attributes such as dip in this study. Comparisons have been made between fault surfaces that have been picked using traditional manual picking methods and machine learning techniques, at different picking intervals. Here we show how fault surfaces and the subsequent attributes and fault analyses vary when picked on every line, every 4th line and every 32nd line. We then describe how faults identified by different seismic quality influence the results from machine learned automated fault extraction. The Vette Fault Zone has relatively poor seismic resolution, with a wide fault zone shown by decreased seismic quality. Conversely, minor faults surrounding the Vette Fault Zone show significantly improved seismic resolution. Differences in seismic quality show variations in the results of DNN models.

3.2.1. Vette Fault Zone: Poorly imaged fault

Starting with the coarse picking strategy of every 32nd line, we can see significant disparities in the modelled fault surfaces between manual verses machine learned fault picking (Figure 4). Despite the overall smoothing that tends to exist when a coarse line spacing is chosen as the picking strategy, there remains a high propensity of the fault to appear highly irregular when machine learning techniques are used, compared to manual picking. The irregularity of the fault surface increases when the spacing for the picking strategy is decreased, as seen when examining the results for picking on every 4th line and on every line. As shown previously, when fault segments are manually picked using every crossing line, this can lead to a rugose fault surface, despite rigorous fault QC'ing. However, a fault surface with increased irregularity is formed when machine learned techniques are used. This is observed by the increased triangles, wider spread in triangle size as well as clustering of different sized triangles that is formed through machine learned techniques, related to irregular fault segments. The increased irregularity formed from both used machine learned techniques and also a narrower line spacing may lead to potential inaccuracies during any further fault analyses performed if not real.

Not only does the irregularity of the fault surface vary with picking strategy and between machine learning and manual methods, the extents of the fault surface produced is also observed to differ. Specifically, there are places where the height of the fault surface is observed to be

increased or decreased when machine learning methods are used. For example, the top of the fault is observed to extend to shallower levels in only a portion of the Southern section of the fault when machine learning methods are used, which is not observed through manual interpretation. Since identifying the location of the base of the fault is highly ambiguous due to poor seismic resolution, machine learned techniques creates a fault surface with an increased irregular fault base, where the depth of the fault base varies more across the entire fault (as seen in figures 4 and 5).

3.2.1. Dip

Fault surface irregularity is significantly higher when machine learning techniques are employed over manual interpretation, as described above. This is reflected in the increased number of irregular 'bulls-eye' patches of varying dip values displayed on the fault, observed on all three line spacing scenarios: every line, every 4th line and every 32nd line (Figure 4). Moreover, these patches of irregular dips tend to be steeper than the surrounding. Conversely, the patches of irregular dip when manual interpretation is performed is a combination of both steeper and shallower dips than the surrounding. The fault is smoothed such that no patches of irregular dip is observed when the fault is picked manually using a picking strategy of 32nd line spacing. The irregularities remain at the 32nd line spacing when machine learned techniques are used. However, they show both lower and higher dip compared to the background and are larger in size compared to irregularities observed in dip images of every 4 lines (Figure 4).

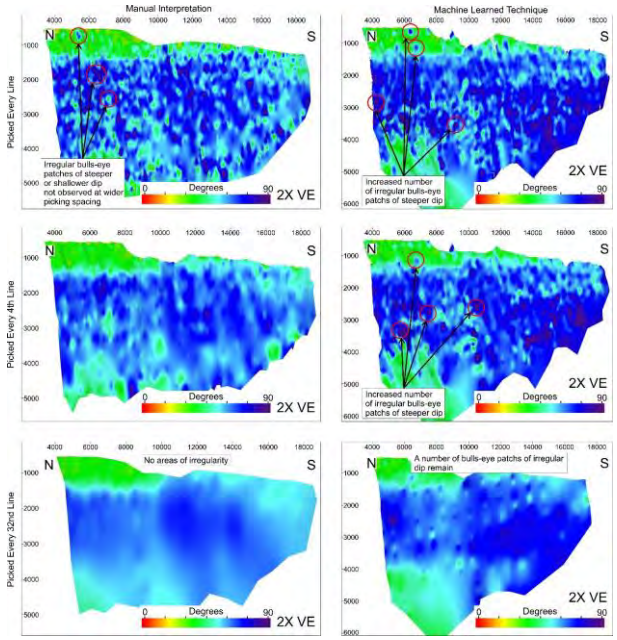

Figure 4. Fault dip attribute displayed on the fault surfaces for manual interpretation verses machine learned techniques picking on every line, every 4th line and every 32nd line. Fault dip is observed to vary with line spacing used for fault picking. A highly irregular fault surface is observed when every line is used for picking, when compared to the overly smooth surface when every 32nd line is used for picking. Moreover, fault extraction using DNN creates an increased irregularity to the fault surface when compared to manual interpretation. X- and y-axes represent the lateral and vertical scales in metres, respectively. Note that unconstrained triangulation is used for fault surface generation.

3.2.2. Dilation Tendency

Upon examining how the predicted dilation tendency of the Vette Fault Zone changes with manual verses machine learned picking techniques, we can see distinct differences. Regardless of picking strategy spacing, we can observe an increased predicted dilation tendency when machine learned techniques are employed over manual interpretation (Figure 5). Manual interpretation using different line spacing shows a gradual increase in predicted dilation tendency with a decrease in line spacing, where a fault is predicted to be at the failure envelope at 25 m line spacing, and further away from the failure envelope at 800 m line spacing. This trend of a decrease in fault stability with a decrease in line spacing is also observed when machine learned techniques are used. However, in all scenarios the fault is predicted to be at the failure envelope, regardless of line spacing used for fault surface generation. Specifically, all three scenarios (25 m, 100 m and 800 m line spacing) show areas on the fault where the dilation tendency is 1 or over. This means that any increase in pore fluid pressure, e.g. through CO_2 injection, is likely to cause the fault to fail (under these specific input parameters).

We can observe that the bulls-eye patches of higher dips correspond to those areas of high dilation tendency (Figure 4 verses Figure 5), and hence it is these areas that have an increased likelihood of failure upon injection of CO_2. Since these irregular high dip patches occur to a lesser degree when manual interpretation occurs, the likelihood of the fault to fail is interpreted to be lower through manual interpretation, particularly for those scenarios where the fault has been picked using a coarser line spacing.

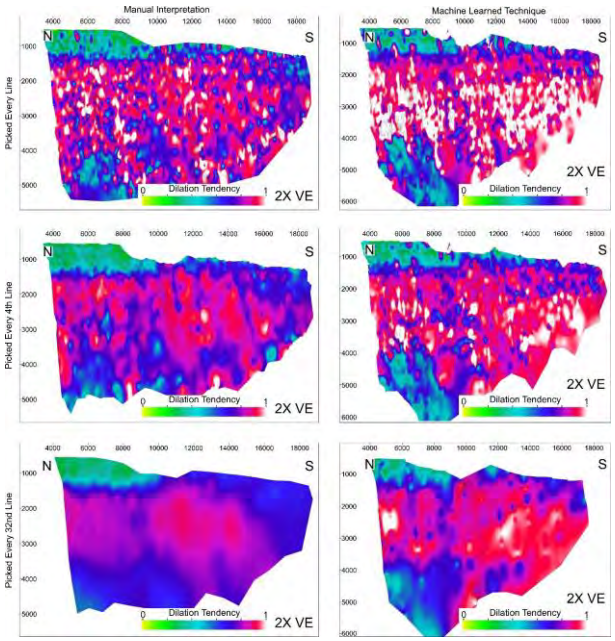

Figure 5. Dilation tendency attribute displayed on the fault surfaces for manual interpretation verses machine learned techniques picking on every line, every 4th line and every 32nd line. Since fault dip varies with line spacing and with picking technique (manual verses machine learned), as does dilation tendency. The increased irregularity to the fault surface when DNN methods are used leads to a fault that would be interpreted to be more unstable in all cases. X- and y-axes represent the lateral and vertical scales in metres, respectively. Note that unconstrained triangulation is used for fault surface generation.

3.2.2. Well imaged minor faulting

While machine learning techniques have shown to be challenging for areas of poor seismic quality, other smaller faults that are better imaged show improved identification. Specifically, minor faults (up to 100 m displacement) within the footwall of the Vette Fault Zone show accurate identification and have a significantly reduced segmentation (Figure 6), despite in several places not showing any sharp cutoffs, but rather identified by subtle folding. To qualitatively and quantitatively assess this improved fault extraction of the minor faults, we compare calculated dilation tendency using machine learning techniques with manual interpretation for one fault within the footwall of the Vette Fault Zone: fault 'FW 01' (see [3] for location details of this fault). We can observed that FW 01 has significantly less segmentation than those picked for the Vette Fault Zone, and in fact, the majority of lines pick only one segment for this fault. Moreover, the predicted dilation tendency is very similar between the machine learned and manually interpreted faults, which would lead to the same overall interpretation of fault stability (Figure 6). The only slight difference between machine learned and manual interpretation is the size of the fault: machine learned techniques do not extrapolate deeper than manual interpretation, which is simply a product of the poor seismic resolution at greater depths.

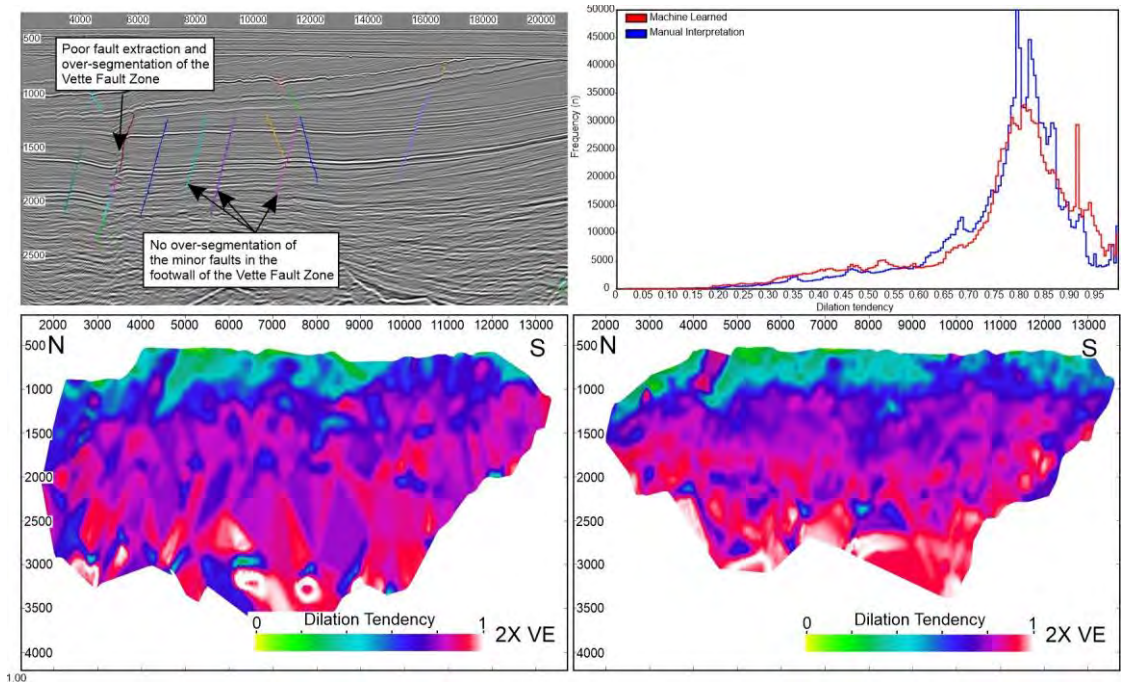

Figure 6. Comparison of manual interpretation verses machine learned interpretation of a fault in the footwall of the Vette Fault Zone. Bottom left: Fault surface produced through manual interpretation. Bottom right: Fault surface produced through machine learning. Highly similar surfaces are produced from both techniques, with the interpretation of the fault stability being almost identical.

4. Discussion

Ensuring the correct picking strategy has been chosen when manually interpreting faults, and understanding the uncertainties involved in both manual and automated fault extraction methods, is crucial to provide the most likely estimate of any subsequent fault stability analysis. Manual picking on every line creates significant

irregularities to the fault surface, which is brought through to such analysis. Choosing this method will lead to the assumption that the fault is at/near failure, which may be inaccurate, despite any assumptions that picking using every available data may provide the best-case example. Conversely, picking using a wide line spacing, such as every 32nd line, creates an overly smoothed fault, which is interpreted to be more stable, which is likely to be incorrect. Since differences in picking strategies can over- or underestimate the fault stability, it is crucial to get the picking strategy correct for accurate predictions of fault stability analysis when assessing a storage site for CO_2 storage. We have shown that picking using every 4th line spacing is likely to create the most geologically accurate representation of the faults in the subsurface; incorporating inherent irregularities of the fault surface, while adding some smoothing to reduce the impact of human error (Figure 7). Further, an important factor to be considered is the line spacing of surveys that are used for interpretation. In our case with 4 line spacing, we are ignoring any irregularity less than 75 m in inline direction.

It is important to note that although injecting CO_2 into the Sognefjord Formation will increase the pore pressure, which in turn increases the likelihood of the fault to fail; such analysis will not have the ability to indicate precisely where on the fault failure may occur. This analysis simply provides an indication of the likelihood of failure.

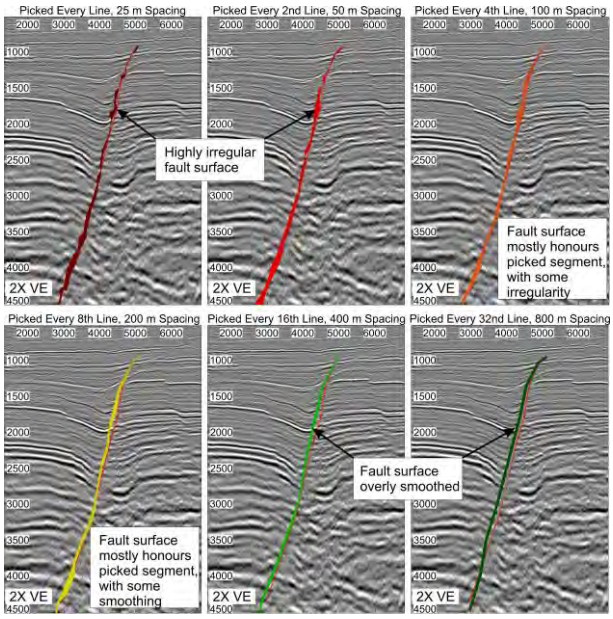

Figure 7. Differences in fault surface generation depending on picking strategy: 25 m, 50 m, 100 m, 200 m, 400 m or 800 m line spacing. Picked fault segment shown as red line. Note the smoothing that occurs at greater line spacing, and the irregularity at narrower line spacing.

Seismic data quality is an issue, which influences both manual and DNN results. The survey used in this study is prestack time migrated which is not the best imaging approach in cases with structural complexity as well as large lateral variations in velocity. Fault shadow effect is observed in the data which deteriorate the quality of seismic data underneath fault planes and causes challenges for both manual and DNN based fault picking.

Utilising modern advances in fault picking, i.e. automated fault extraction using supervised DNN, is a fast approach to the normally time-consuming manually interpretation strategy. However, as with any new technology, it is crucial to understand the influence of how, and to what extent, machine learning may influence subsequent fault analyses such as fault stability. In our case study, DNN based results showed more irregularity for the fault surfaces, especially for the poorly imaged Vette Fault. It could be argued that fault surfaces are highly irregular in nature and hence the overly irregular faults produced could be due to automated extraction picking every kink or bend in a fault, that may be missed by manual interpretation. Large amounts of data exist in geophysical studies, hence a neural network might find hidden irregularities in the data that manual interpreters may have overlooked [8]. However, it may also be due to areas of false positive interpretation by the DNN models that we applied, particularly in areas of poor seismic resolution. As with manual fault picks, errors associated with machine learning techniques will be integrated in the fault model. Identifying short-comings through quality control may indicate the need for continued machine learning retraining to improve the model, by either further hyperparameters tuning and/or improving the input labels.

Our applied DNN models provided vertically segmented images for the Vette Fault. The networks use features on seismic data that can represent discontinuity. This segmentation can happen, for example, in areas with weak discontinuities on seismic data (e.g. packages with very low reflectivity). Hence, this could be an explanation for the increased irregularity. In manual interpretation, we usually ignore such weak reflectivity areas and extend picks over them. One approach to overcome this issue was to take ensemble of multiple DNN results. Figure 8B and D show the improvement of vertical and lateral continuity by using ensemble results.

Areas with a severe decrease in seismic quality, creating a high degree of ambiguity often lead to poor picking through machine learned techniques. For example, areas surrounding the relay zone to the Southern end of the Vette Fault (known as Vette_2, see [3]) have been poorly identified (Figure 8B). The high complexity assumed for the basin-scale Vette Fault results in a poorer seismic resolution at the fault. This leads to poor predictions of the location of the fault, and some areas of the fault being missed. In areas of poor quality seismic, manual interpretation is also model-based and can be non-unique. On the contrary, well-imaged minor faulting show excellent correlation between manual and machine learned fault picking, with the resulting predicted dilation tendency being very similar between these two methods. Hence, improved seismic quality will reduce the uncertainty of DNN based fault picking.

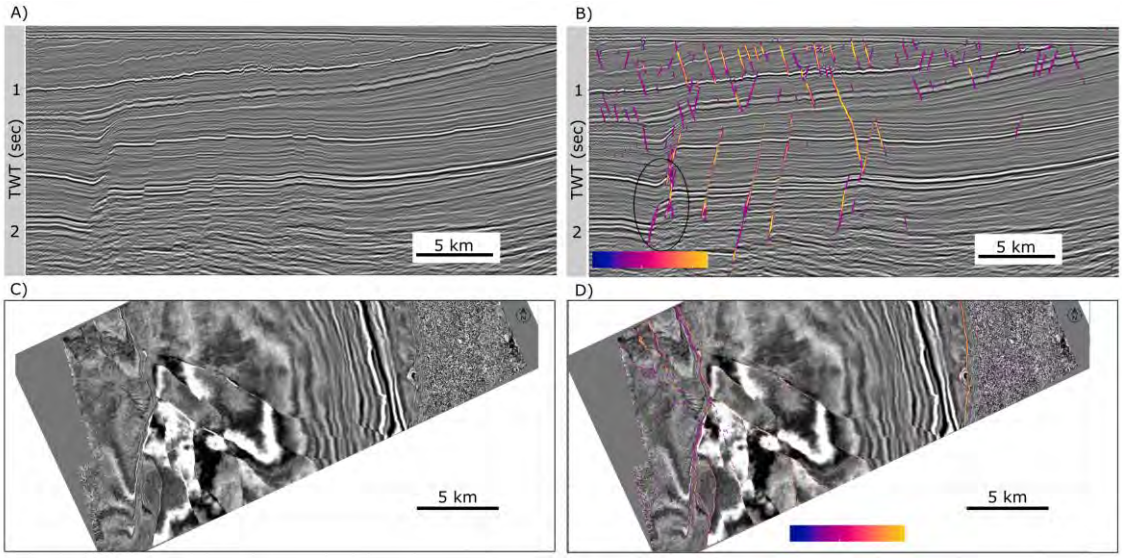

Figure 8. Examples of fault predictions using the supervised DNN models. The predicted faults are ensemble of four different models and the colour represents the confidence of predicted faults. The yellow colour means that all four models predicted faults. A) Inline without predicted results B) Inline with predicted faults. The area inside the ellipse shows poor quality seismic zone. C) Time slice without predicted faults D) Time slice with predicted faults. Note the segments of the Vette fault with less confidence in the prediction.

5. Summary

Line spacing chosen to pick fault segments will influence any subsequent analysis, e.g. fault stability, with the results varying with picking strategy. Manually interpreting using a wider line spacing creates a fault that is predicted to have an increased stability. Conversely, picking using every line spacing creates a highly irregular fault, such that the stability is predicted to be significantly reduced, and in fact will lead to the prediction of an unstable fault.

Automated methods of fault extraction is sensitive to the quality of seismic data. Poorer imaging of faults creates fault surfaces with increased irregularity when compared to manual interpretation, leading to higher predicted dilation tendency values in all line spacing scenarios. Using ensemble models, a larger coverage of faults was imaged using DNN with additional information of the confidence of predicted fault. Further fine-tuning of hyperparameters and fault label picks can potentially improve the results. On the contrary, picking of well-imaged, smaller faults show noticeable similarity in results between manual and automated methods.

Acknowledgements

This is a contribution of the FRISK project, supported by the Research Council of Norway (RCN# 295061). Support from the NCCS Centre is acknowledged, performed under the Norwegian research program Centres for Environment-friendly Energy Research (FME). The authors acknowledge the following partners for their contributions: Aker Solutions, Ansaldo Energia, CoorsTek Membrane Sciences, EMGS, Equinor, Gassco, Krohne, Larvik Shipping, Lundin, Norcem, Norwegian Oil and Gas, Quad Geometrics, Total, Vår Energi, and the Research Council of Norway (RCN# 257579/E20). Badley Geoscience Ltd. and Earth Science Analytics are thanked for their academic licenses of T7 and EarthNet, respectively.

References

[1] EU Commission, 2018. Communication from the commission to the European parliament, the European Council, the Council, the European economic and social committee, the committee of the regions and the European investment bank, in: A Clean Planet for All. A European Strategic Long-Term Vision for a Prosperous, Modern, Competitive and Climate Neutral Economy. Brussels.

[2] Rogelj, J., Den Elzen, M., Höhne, N., Fransen, T., Fekete, H., Winkler, H., Schaeffer, R., Sha, F., Riahi, K. and Meinshausen, M., 2016. Paris Agreement climate proposals need a boost to keep warming well below 2 C. *Nature*, *534*(7609), pp.631-639.

[3] Mulrooney, M.J., Osmond, J.L., Skurtveit, E., Faleide, J.I. and Braathen, A., 2020. Structural analysis of the Smeaheia fault block, a potential CO2 storage site, northern Horda Platform, North Sea. *Marine and Petroleum Geology*, p.104598.

[4] Sheriff, R.E., 1982. *Structural interpretation of seismic data* (No. 23). American Association of Petroleum Geologists.

[5] Al-Shuhail, A.A., Al-Dossary, S.A. and Mousa, W.A., 2017. *Seismic Data Interpretation Using Digital Image Processing*. John Wiley & Sons.

[6] Zhao, T. and Mukhopadhyay, P., 2018. A fault detection workflow using deep learning and image processing. In *SEG Technical Program Expanded Abstracts 2018* (pp. 1966-1970). Society of Exploration Geophysicists.

[7] Statoil, 2016. Subsurface Evaluation of Smeaheia as part of 2016 Feasibility study on CO_2 storage in the Norwegian Continental Shelf. OED 15/1785. Document A – Underground report Smeaheia (Internal Report – Available on Request Only).

[8] Russell, B., 2019. Machine learning and geophysical inversion—A numerical study. *The Leading Edge*, *38*(7), pp.512-519.