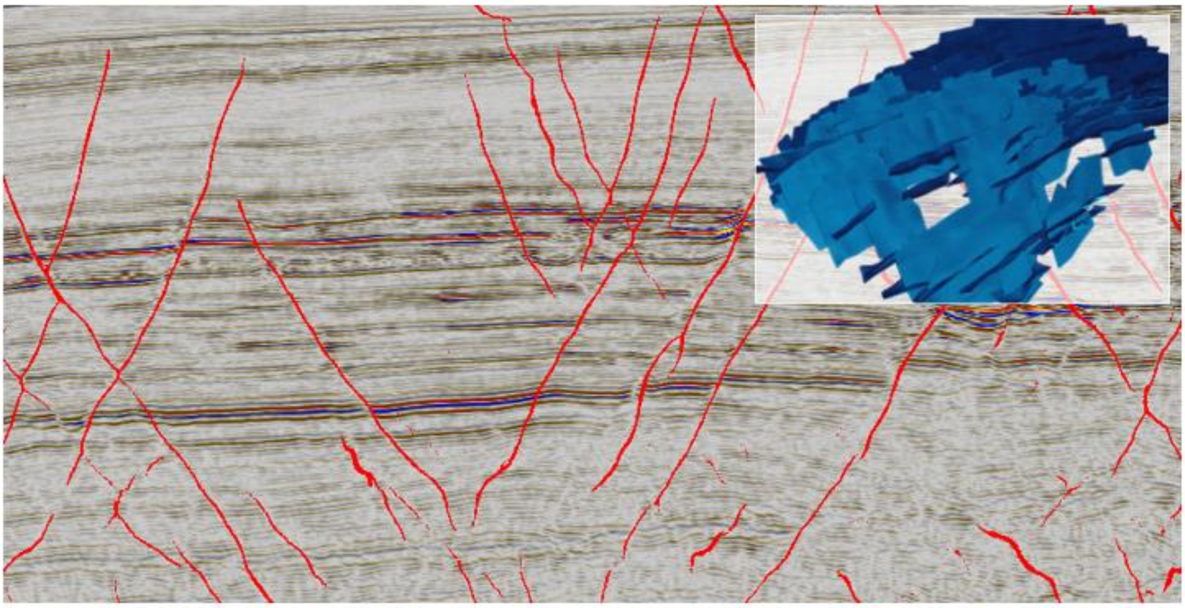
Fault segmentation based on the machine learning results

Schlumberger Norway Technology Center (SNTC), Stavanger.

One of the biggest challenges in the oil and gas industry is how to improve efficiency and reduce the time taken to get to production of oil and gas.

An example of a task that would benefit from improved efficiency is the fault interpretation process. Traditional fault interpretation method is a laborious and time-consuming task often taking weeks of effort for a skilled geoscientist. Automation of the process can help to reduce the interpretation time and allow a geoscientist to focus on more important duties instead of spending most of their time manually interpreting.



Fault interpretation is manual and laborious.

To enable a geoscientist to quickly and easily get a fault interpretation result in an automated manner, at least two steps need to be considered:

1. Prediction of faults in seismic data.
2. Automatically extracting faults from the predictions.

Prediction of faults is currently accomplished using machine learning techniques. As well, there are fault extraction techniques available and used in traditional workflows. The challenge with fault extraction methods is that there is no available, robust algorithm which is capable of proper faults segmentation during extraction from seismic volumes. Several novel approaches are being researched within the industry to address this problem.

This project focuses on attempting to automatically segment the faults from the predictions.

This project aims to solve the following tasks:

- Given a set of seismic machine learning results, examine different feature extraction techniques.
- Define possible solution to segment faults before/during fault extraction.
- Extract segmented fault planes and compare it against original seismic data.

Data

We will provide one public seismic dataset from offshore northwest Australia, together with its associated fault likelihood label file. The data will be delivered in either CSV or binary generic format, such that the files (input seismic data file and label file) can be easily read programmatically in any programming language:

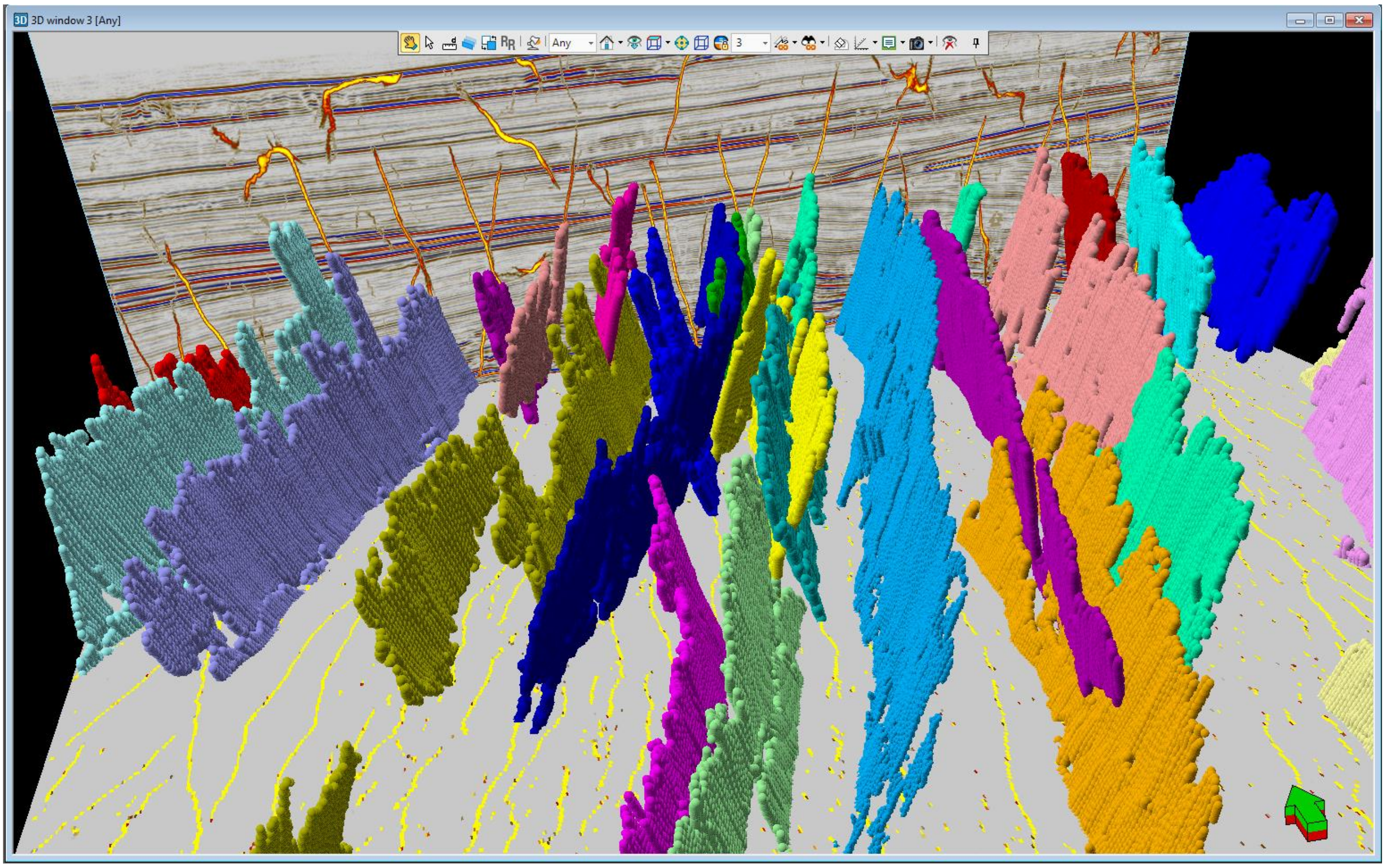


***Figure**: Input 3D image (in red-white-blue color scale), overlaid with fault labels (in yellow). Also shown are some extracted fault planes (in point-cloud format), using a deterministic statistical approach. Clearly, the extraction process has not managed to skeletonize all fault predictions in the data.*

Each 3D file will contain circa 1000 x 1000 x 1500 = 1.5 billion samples in 32-bit format (which is 6 GB each, in binary format). The data is already available on a computer at NTNU.

For this project (to develop one or more Machine-Learned fault segmentation models), we see two optional tracks:

a) **6-month duration track:** Generate a set of synthetic 3D fault planes, and their associated ground truth labels. Then use that synthetic dataset to train one or more ML models to recognize and skeletonize those faults. Then qualitatively test those models on the observed 3D fault likelihood volumes
b) **12-month duration track**: As in track a), but then compare/benchmark your new methods against at least one other method in the industry. For instance this one (with code available on GitHub): https://inside.mines.edu/~dhale/papers/Hale12FaultSurfacesAndFaultThrowsSeg.pdf

Even more datasets, with the same details, but covering images from the underground from other parts of the world, can be provided as needed.

Contact: Olesya Zimina (ozimina@slb.com), Victor Aare (vaarre@slb.com)