Deep Learning Application for Characterizing Petroleum Systems: A Case Study from the F3 Area, North Sea

Scotty Salamoff[1], Julian Chenin[1], Benjamin Lartigue[1], Nguyen Phan[1], Paul Endresen[1] ([1]Bluware Corporation)

Summary

Detailed geomodelling within high-resolution, three-dimensional (3D) seismic data is a time-consuming and arduous process. However, recent advances in deep learning practices are accelerating the speed at which geologic features can be mapped. While most geoscientific deep learning applications have focused on mapping features such as faults and salt, we propose a novel, interactive deep learning methodology that enables the interpreter to characterize a petroleum system by labeling and training networks on associated elements proven by exploration well data. This study uses available data from the complex Central Graben Basin within the North Sea, which contains many producing fields.

The F3 seismic survey contains several seismic representations of petroleum system elements such as migration chimneys and dry gas shows. Dry gas migrates vertically through overlying strata and along faults. Results from well-trained deep learning networks can accurately map various petroleum elements of the basin, which is traditionally very challenging and time-consuming. These results were obtained in a fraction of the time compared to traditional interpretation workflows and enables geoscientists to better characterize regional trends while also making observations at the petroleum system scale.

Introduction

Petroleum system mapping is a critical stage of hydrocarbon exploration. A persistent challenge facing the geoscience community is the ability to evaluate large datasets quickly and accurately for reservoir potential. Petroleum system elements are useful for describing the actions of hydrocarbon systems by focusing upon the characteristics and events of a specific geologic system (Waples, 1994). These models, if designed and integrated properly, can represent a powerful exploration tool (Mancini et al., 2003). However, these models must often go through several revisions, and creating them quickly is a significant task. In recent years, several approaches with sophisticated neural network architectures have shown potential when applied to tedious seismic interpretation tasks (LeCun et al., 2015; Bandura et al., 2018). These deep learning workflows accelerate such tasks, shifting the focus of geoscientists from exhaustively digitizing features on a workstation to critical evaluation and risk/resource analysis. By combining the accelerated workflows offered by deep learning with established play analysis procedures, geoscientists can generate and evaluate these models in a fraction of the time it takes using traditional interpretation methods, to better understand the generation, accumulation, and entrapment of hydrocarbons.

Geologic Background

The study area lies in the triple rift system of the Central Graben, Viking Graben, and Moray Firth Basins where the structural province is defined by Late Jurassic and Early Cretaceous extensional tectonics due to failed rifting (Gautier, 2005; Silva et al., 2019). This plays a fundamental role in the distribution of hydrocarbons (Schroot and Schüttenhelm, 2003; Gautier, 2005; Silva et al., 2019).

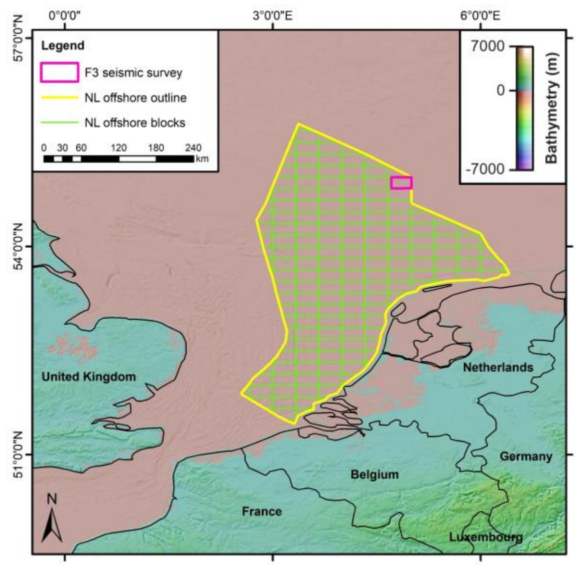

Figure 1: Location of the study area (pink boundary) within the North Sea. Seafloor bathymetry map taken from topex.ucsd.edu.

The Zechstein Group is composed of carbonate and evaporite rocks, with several salt structures within the study area providing localized structural control (Figure 2, zone 1) (Silva et al., 2019). Zone 2 within Figure 2 contains the Germanic Trias, Altena, Rijnland and Chalk Groups. The Lower Triassic Germanic Trias Group is composed of shales and siltstones interbedded with sands, whereas the Upper portion contains mostly anhydrous evaporates (Gautier, 2005, Silva et al., 2019). The Middle Jurassic Altena Group is characterized by thick marine shales and the Early Cretaceous Rijnland Group predominantly contains siliciclastics (Silva et al., 2019). Finally, the Late Cretaceous Chalk Group is composed of chalk and argillites with polygonal faulting (Figure 2) (Silva et al., 2019). The last significant regional event was during the Mid-Miocene, resulting in the Mid-Miocene unconformity (Figure 2, orange horizon) (Schroot and Schüttenhelm, 2003, Silva et al., 2019). The Cenozoic North Sea Supergroup has the most significant presence within the dataset and is characterized by strong subsidence and significant halokinesis from the Zechstein Group salt (Figure 2, zones 3-6) (Silva et al., 2019).

The hydrocarbons in the study area originate from a singular petroleum system known as the Kimmeridgian Shales Total Petroleum System (TPS) (Schroot and Schüttenhelm, 2003). TPS source rocks were deposited from the Late Jurassic to Early Cretaceous, during a period of intensive extension and rifting (Schroot and Schüttenhelm, 2003). Several large faults and vertical gas migration chimneys exhibit how the hydrocarbons migrated into the North Sea Supergroup shallow reservoirs (Figure 2) (Schroot and Schüttenhelm, 2003; Gautier, 2005).

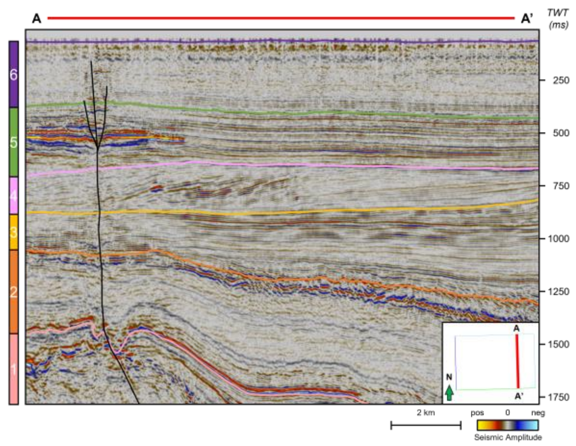

Figure 2: Regional interpretation of the F3 seismic survey, North Sea. The F3 reservoir is visible on the left side of the line, at ~500 ms. Other petroleum system elements visible on the line include migration cloud, fault (charge conduit), and leak-off.

We present a methodology that enables a high degree of interactivity between the deep learning process and the geoscientist. Therefore, the network acts as an extension of the interpreter to augment mapping capabilities of important geological features such as faults, reservoirs, migration clouds, and salt. In this case study, we will use a series of networks to map various petroleum system elements within the F3 seismic survey.

Methodology and Deep Learning Network

Machine learning techniques to help analyze and interpret geologic patterns have shown significant promise over recent years (LeCun et al., 2015; Ronneberger et al., 2015; Bandura et al., 2018; Chopra and Marfurt 2018; Silva et al., 2019; Chenin et al., 2020). However, our methodology differs from existing ones due to the high level of interactivity it provides between the deep learning process and the interpreter. Therefore, we refer to this process as interactive deep learning.

The geoscientist does not rely on “black box” utilities to generate high-quality results since all network parameters are exposed. With our workflow, the data labeling, network training, and prediction happen in real-time. Geoscientists work in tandem with the network and provide instantaneous feedback to the algorithm’s prediction (or inference). They can simultaneously train and provide active reinforcement to the deep learning network until the desired prediction is achieved. Because of this immediate feedback, quality control is embedded within the labeling and training processes. This removes the disadvantage of tedious and time-consuming quality control and reiteration of black box outputs.

Figure 3 highlights the important differences between interactive and traditional deep learning approaches. Proprietary technology such as compression and random access to seismic data makes interactive deep learning possible. Other currently available deep learning methods require significant data preparation steps before initiating training. This is due to the static nature of TensorFlow record files, which cannot be modified on the fly. Our methodology shows how this unnecessary burden is avoided (Figure 3).

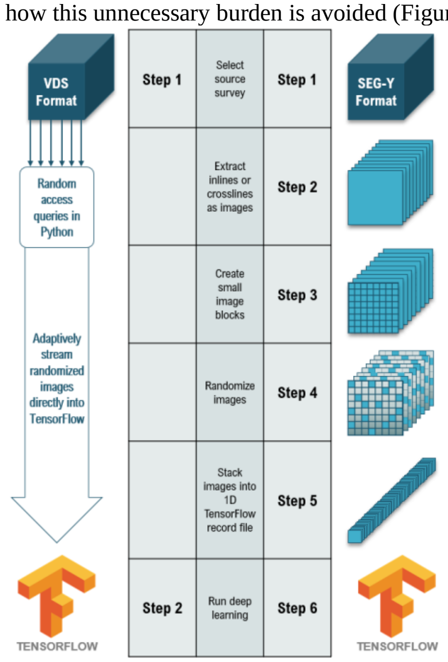

Figure 3: Comparison between traditional and interactive deep learning for seismic interpretation. With traditional deep learning, the process starts with SEGY data and image file creation and randomization. With interactive deep learning, steps 2 through 5 are eliminated.

The deep learning architecture presented is based on a Convolutional Neural Network (CNN) that reduces the dimension between the features (seismic data) and labels (interpreter input). This reduction uses “valid padding” during the convolutions to maximize the information sent to the network (Figure 4). Padding refers to the number of pixels added to an image when processed by the kernel of a neural network. Padding “same” can be detrimental to the network since it fills the convolutions with zeros and can omit some of the data near the edges of the patches. Following our methodology, deep learning models can now be operated interactively, responding in real-time to interpreter input.

Optimal CNN inputs require the following:

1. Patches from the seismic (features) and input interpretations (labels) by the geoscientist.

2. Random access to the seismic features and interpreter label pairs. The features are in 3D, whereas the labels are in 2D. Random data access is key for enabling immediate feedback on the algorithm’s inference.

3. There should be an equal and sequential number of truth and non-truth samples (e.g., the same amount of non-fault examples as faults).

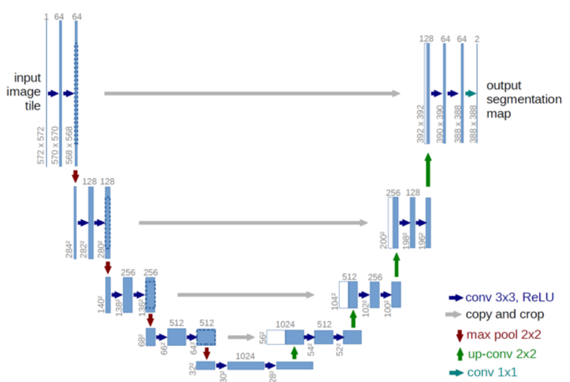

Figure 4: Example of a CNN architecture from Ronneberger et al., 2015 using a combination of “convolution” and “down-sampling/max pooling” operations. The number or combination of these operations can vary for many different reasons and depends on the type of problem being solved. It is important to note that our method uses slightly different input and output sizes compared to the model in the figure.

F3 3D Seismic Data

The F3 post-stack, time migrated survey is located roughly 180 km offshore Netherlands and covers approximately 340 km^2 in the Central Graben Basin, North Sea (Figure 1) (Silva et al., 2019). This data set is SEG positive polarity, where the sea floor is observed as a trough.

Results

A total of 7 inlines (IL) and 5 crosslines (XL) were labeled to identify reservoir and migration cloud petroleum system elements, which is 0.63% of the dataset. For salt identification, 7 ILs and 5 XLs were labeled, which amounted to labeling 0.58% of the data. Each petroleum system element network was trained separately and used different label distributions. Table 1 shows network details and individual training times. With a 16 GB NVIDIA RTX Tesla M60 graphics card, each epoch took roughly 45 seconds.

	Salt	Reservoir	Migration Clouds
GPU	1x RTX Tesla M60	1x RTX Tesla M60	1x RTX Tesla M60
# Of Patches	8192	8192	8192
# Of Epochs	50	70	70
Time Per Epoch (s)	45 s	45 s	45 s
Total Training Time (mins)	~37 mins	~53 mins	~53 mins

Table 1: Details and timing for the individual networks.

Upon completing the initial training cycle for all networks, the interpreter then reviews the inference superimposed on the amplitude data and optimizes their label sets accordingly. For example, in areas where the interpreter sees a false positive prediction, they can reinforce that the prediction is false by declining to co-locate labels with the erroneous prediction. The network continues training, and the labeling process is repeated until the desired level of accuracy is achieved. Figure 5 shows examples of inference from separate networks trained on different petroleum system elements.

Discussion

Overall, our interactive methodology substantially improves interpretation efficiency while helping to reduce human error. As a result of the interpreter working in tandem with a deep learning engine, the network will suggest similar features within the seismic section, such as potential reservoir running room that may have been originally overlooked by the geoscientist. The network can be trained on any feature visible in the seismic data and differentiates between subtle changes in seismic character that exemplifies multiple components of a complex system. The subtle geologic features identified by the algorithm enables geoscientists to gain detailed insights into the complex petroleum systems of the North Sea in a fraction of the time it takes using traditional interpretation methods.

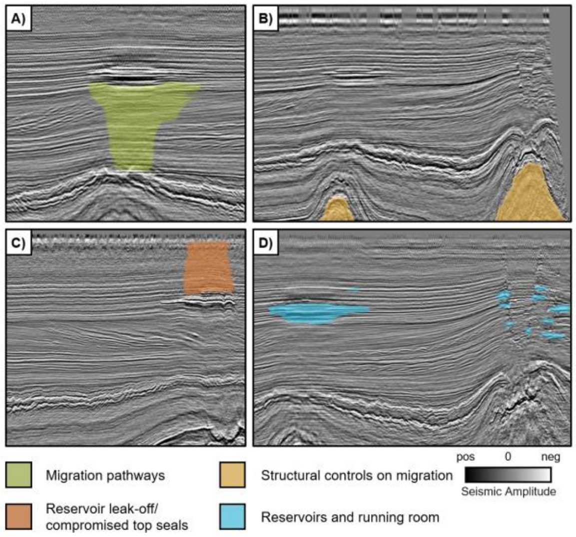

Figure 5: Network inference from A) migration cloud; B) salt (structural focus); C) leak-off; and D) charged reservoir-quality rock.

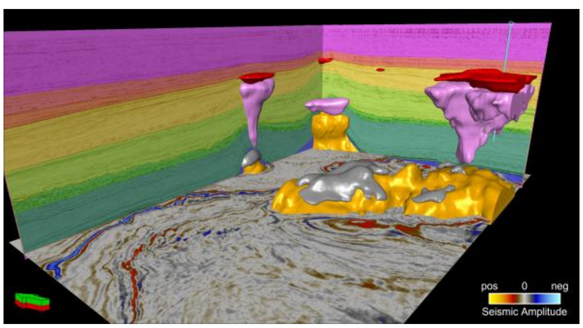

Figure 6: Raw output objects visualized with seismic amplitude data. Salt (orange) with deformed shale (grey) provides structural control for the migration clouds (pink), some of which underlie the two main reservoirs (red) identified using deep learning in the area. Smaller gas accumulations along fault planes are also shown, and the approximate location of exploration well F3-01 is shown in light blue.

Conclusions

Interactive deep learning has the potential to significantly accelerate the process of petroleum system element mapping and reservoir identification. This new deep learning methodology ultimately improves the quality of the interpretation while reducing human error. Geoscientists work in tandem with the algorithm until the results are satisfactory. Therefore, inferences are as reliable or more reliable as manual interpretation. These results have important implications for characterizing the remaining reservoir running room within the basin while capturing the basin's complex petroleum system.

Acknowledgements

The case study uses the F3 dataset – made available by NLOG of the Netherlands and distributed by terranubis.com under a Creative Commons 3.0 license.

We also thank the Bluware geoscience and data science teams for their helpful advice and feedback. This work was performed using InteractivAI, a commercial software product owned and marketed by the Bluware Corporation.