NERC Centre for Doctoral Training in Oil & Gas (2018 start)

Project Title: Quantitative assessment of interpretational uncertainty in geological mapping with machine learning.

Host institution: Heriot-Watt University
Supervisor 1: V. Demyanov (HWU)
Supervisor 2: D McCarthy (BGS)
Additional Supervisor (s): D Arnold (HWU)

Project description:

This project aims to tackle the fundamental problem of how to adequately capture and preserve geological uncertainty in reservoirs modelling workflows. Traditional reservoir characterisation workflows integrate interpretations from data of different nature – seismic, wireline, core. These interpretations are done by relevant domain experts, who are often separated into siloes, focused only on their aspect of the data, and furthermore can be subject to the experiential bias. When data is finally combined together to create a reservoir model, it is typically done by a single person who must create a coherent representation of the reservoir and its associated uncertainties.

Typically the resulting models are often "best" fits to each bit of data and do not adequately quantify the uncertainties subject to multiple possible interpretations given the sparse nature of the data. Instead we should seek to generate models that combine data in different possible ways to capture the fullest representation of the uncertainty and preserve the knowledge of uncertainties associated with each input data type and its interpretation.

The project will develop a way to elicit a wide range of geological concepts as models directly from the data by discovering a variety of data combinations (data types and how they are used/fused together) through using machine learning (ML), while preserving geological knowledge. This will enhance the estimation of uncertainty in our reservoir modelling workflows based on finding a range of unique data combinations/fusions that are coherent (geologically realistic) and unbiased. Achieving this will make a step change enhancement in subsurface uncertainty modelling practice, by identifying a wider possible spread of geological scenarios and providing a quantitative way in assessing their probability. This process will also be significantly faster than a manual approach to developing different ways to use the data to build models.

A key element of the project will be to embed geoscience understanding into data driven workflows with machine learning (ML). The challenge is that recent advances in computer science need to be adapted for best implementation in geoscience to capture the context understanding and thinking as it is performed by domain experts. Rigorous ML approach will ensure there is no preferential bias in the way multiple interpretations are elicited from the data. Proper application of ML is able to quantify the impact of the tendencies hidden in the data and how their combinations define possible sedimentological settings. Feature selection approach will be implemented to enhance interpretability of the ML model outcomes. Feature selection is a powerful tool, which is often misused in ML geoscience applications due to the lack of understanding the nature of the domain data. Geologically consistent features need to be derived from the relevant data with account for the associated uncertainty to enhance performance of ML prediction and retain geological realism in ML predictions. This will be a step-change to many recent studies where ML application is done mechanistically without proper respect to what the actual drivers for geological interpretation are in the presence of depositional scenario uncertainty. The latter requires introduction of solid geoscience understanding into ML prediction model design. Embedded geoscience knowledge will ensure geological consistency of the multiple elicited interpretation leads.

The outcome of the PhD project will be a more robust handling of geological uncertainty through novel workflows that use modern machine learning techniques with embedded geoscience understanding. In particular, this will involve adapting Artificial Intelligence (AI) methods, such as deep learning and semi-supervised learning, feature selection, to collate interpretable patterns with

the geological context to ensure depositional consistency and geological sense of the predictive models.

The feasibility of the novel scientific approach will be justified with a real field study using a modern dataset from the Sea Lion Field, which features modern 3D seismic, as well as log and core data of a comprehensive nature. Once established, this method could be applied to any project where modelling of the subsurface is required, e.g. CCS and the UK-GEOS projects.

CDT Research theme(s): **Exploitation in Challenging Environments** - development of AI technology to enhance managing of sedimentological risks by data and knowledge integration.

Research context:

Reservoir exploration and development decisions in hydrocarbon exploration are potentially very costly, and are typically based on sparse data sets (boreholes, seismic, outcrop analogues, etc.). Therefore, these decisions are put at a significant risk by a large degree of uncertainty, which is partially due to interpretational bias. Development of a reliable predictive models that support decision-making with accurate uncertainty management is made difficult due to differences in uncertainty handling across data types and scales of measurements.

At present there is no coherent method in which information from core, wireline and seismic data is integrated in a holistic objective way with respect to the associated uncertainties. Traditional statistical methods used in geoscience applications are constrained with many assumptions (linearity, stationarity, Gaussianity), which are not likely to be met in many geological settings. Many recent applications of ML to different aspects of geophysical or sedimentological data are usually done from the computer science perspective without enough understanding of the geological context behind the data. Therefore, many attempts to use ML: for classification or regression problems demonstrate under performance and lack of interpretability.

The premise of the proposed PhD is based on the top MSc project in Applied Petroleum Geoscience at HWU in 2017, which was selected for oral presentation at the Conference for Mathematical Geosciences. This was followed by a recent successful pilot study, performed in collaboration with the BGS, on a test application of Random Forest to assess interpretational uncertainty in facies classification with wireline data from the Sea Lion Field, North Falkland Basin. The outcomes of this project will be submitted to the 57th British Sedimentological Research Group Annual Meeting.

This proposed PhD project is based on the solid fundamental science foundation that poses an advance multidisciplinary challenge across: geology, geophysics, petrophysics, computer science (AI) and statistics (UQ). This truly multi-disciplinary project will explore uncharted territory, will fundamentally integrate geoscience and data science and will lead to a number of student-led publications.

Research costs:

PhD tuition fees and a stipend, PC hardware and conference costs. The project will have an opportunity to collect additional field analogue data from relevant outcrops (Spain, HWU REM field trip) with access to drone technology. The total additional field, travel and hardware costs will be 15k over 4 years.

Career routes:

The PhD graduate with a geoscience background will master data analytics, AI and statistics skills, to become a fine example of new generation of geoscientists developed by the CDT to progress into a successful career path in oil & gas industry or subsurface energy regulators. The academic progress of the PhD will be ensured by a strong supervision team between HWU and BGS that combines geological, reservoir geomodelling, statistical and AI modelling expertise. The student will be exposed to peer-learning within a multi-disciplinary Uncertainty team at HWU who are working on a range of industry and RC funded projects.