Machine learning to optimise the petrophysical workflows in oil and gas exploration

Nick Brown, Data Architect at EPCC

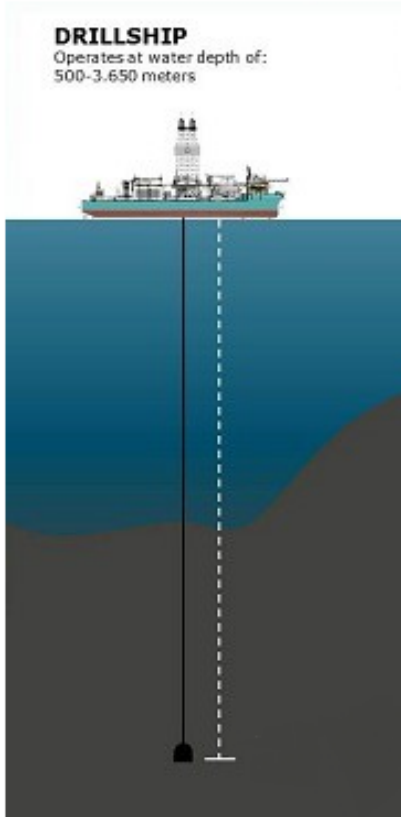
n.brown@epcc.ed.ac.uk

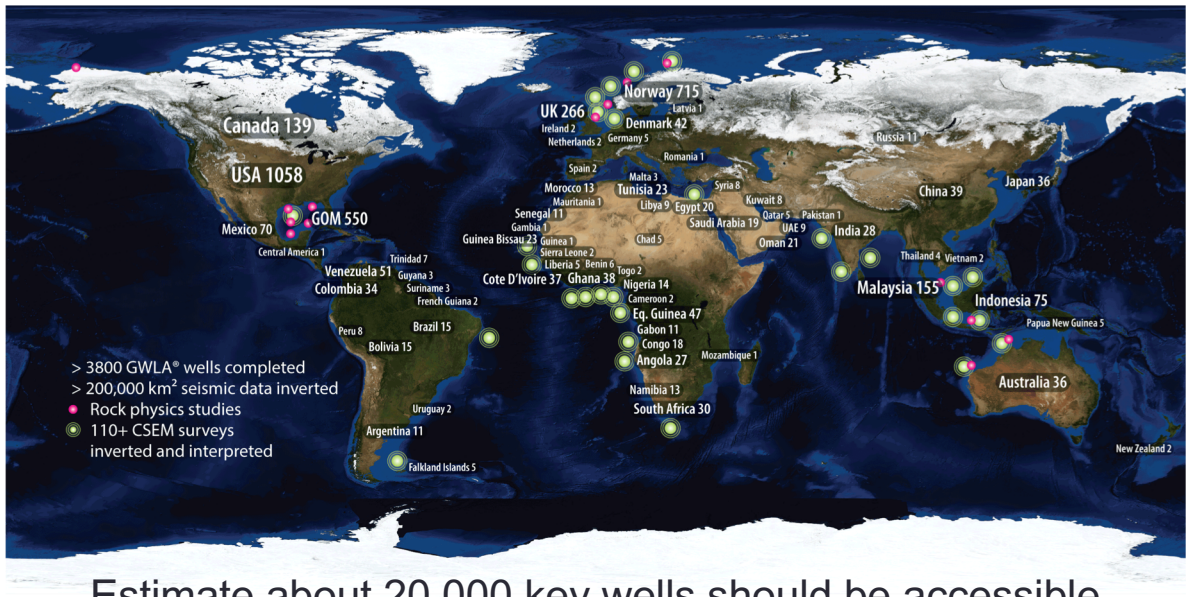
epcc

THE UNIVERSITY OF EDINBURGH

Driven by a collaboration with Rock Solid Images (RSI)

- RSI develop a successful software package known as RockAVO™ which enables oil and gas companies to virtually explore geological regions to inform decision making around exploration and exploitation
  - This is fed with well atlases, which are interpreted from the raw well log data

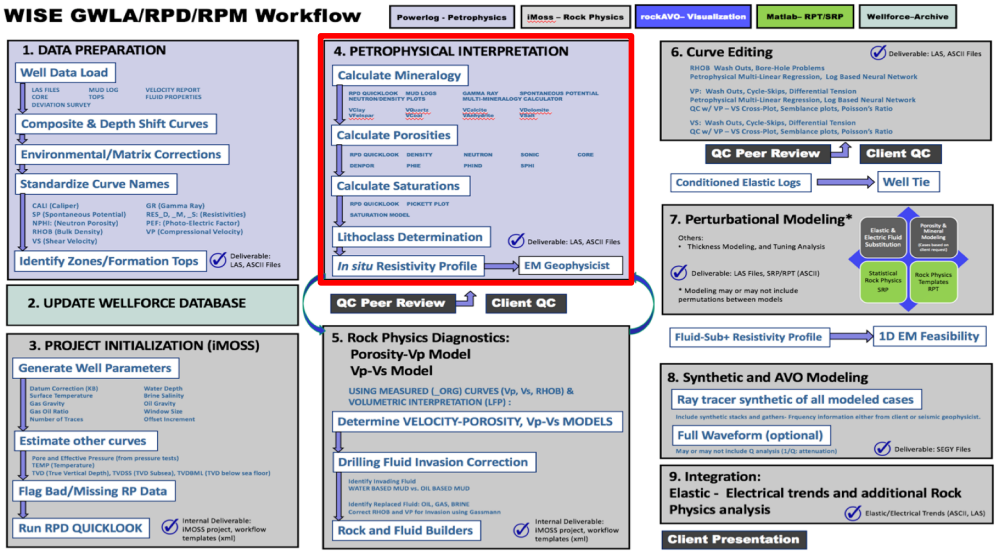




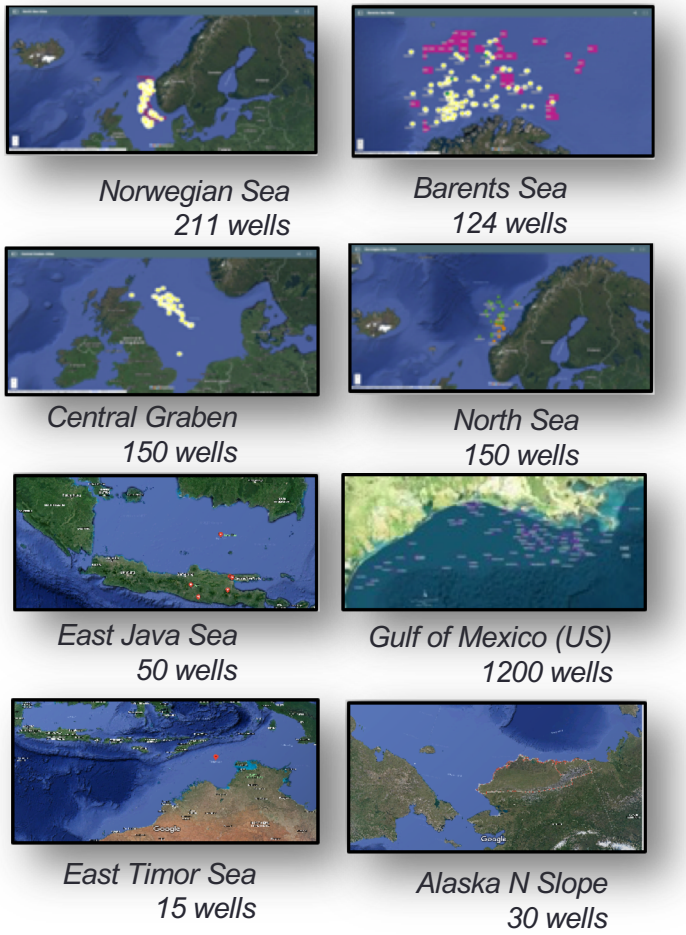
Estimate about 20,000 key wells should be accessible

Where does machine learning come in?

- A time consuming manual interpretation process is needed to convert raw well log data into atlases
  - Over 7 days per well
- This is fundamentally a pattern recognition problem, so can we use machine learning?
  - For instance could we go from over 7 days to 7 minutes per well?



Existing data base

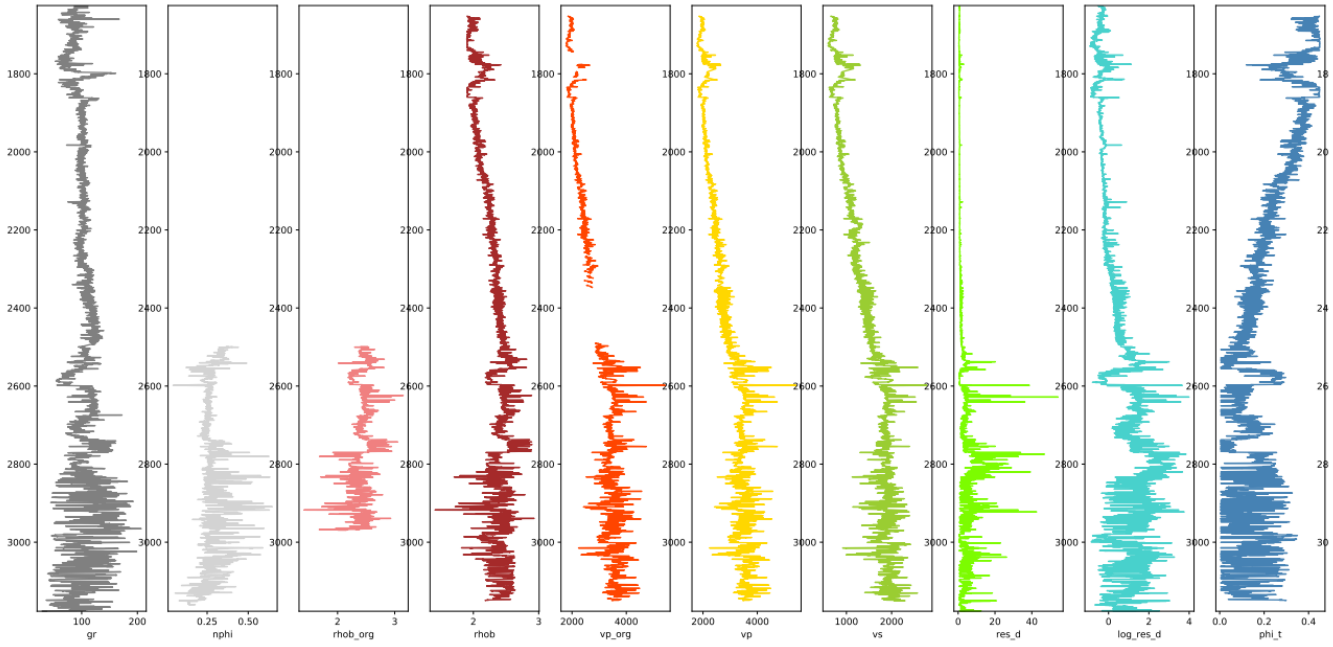


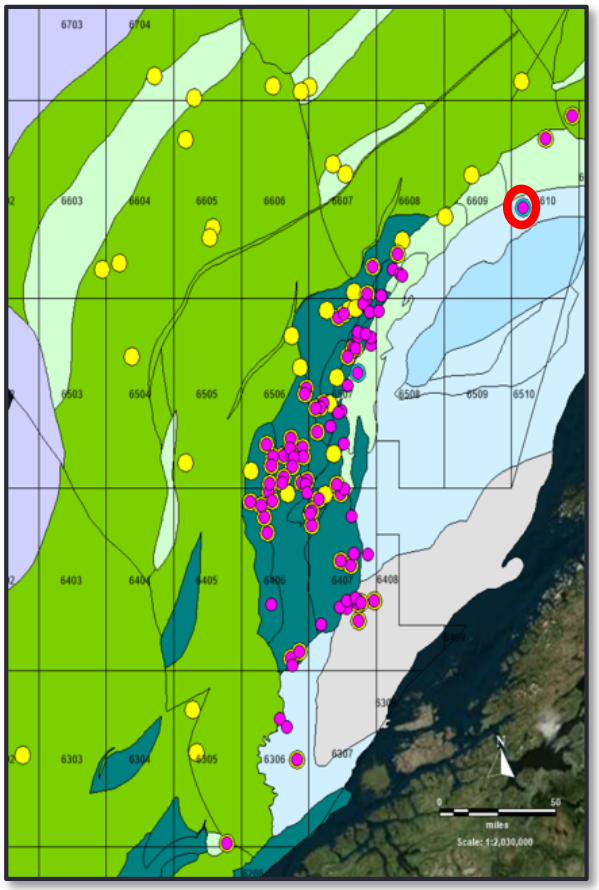
2000 wells in the existing database

epcc | THE UNIVERSITY OF EDINBURGH

The real world data can be messy!

- The data itself it incomplete and unpredictable
- Our truths are themselves interpreted, and the human doesn't always get it right!



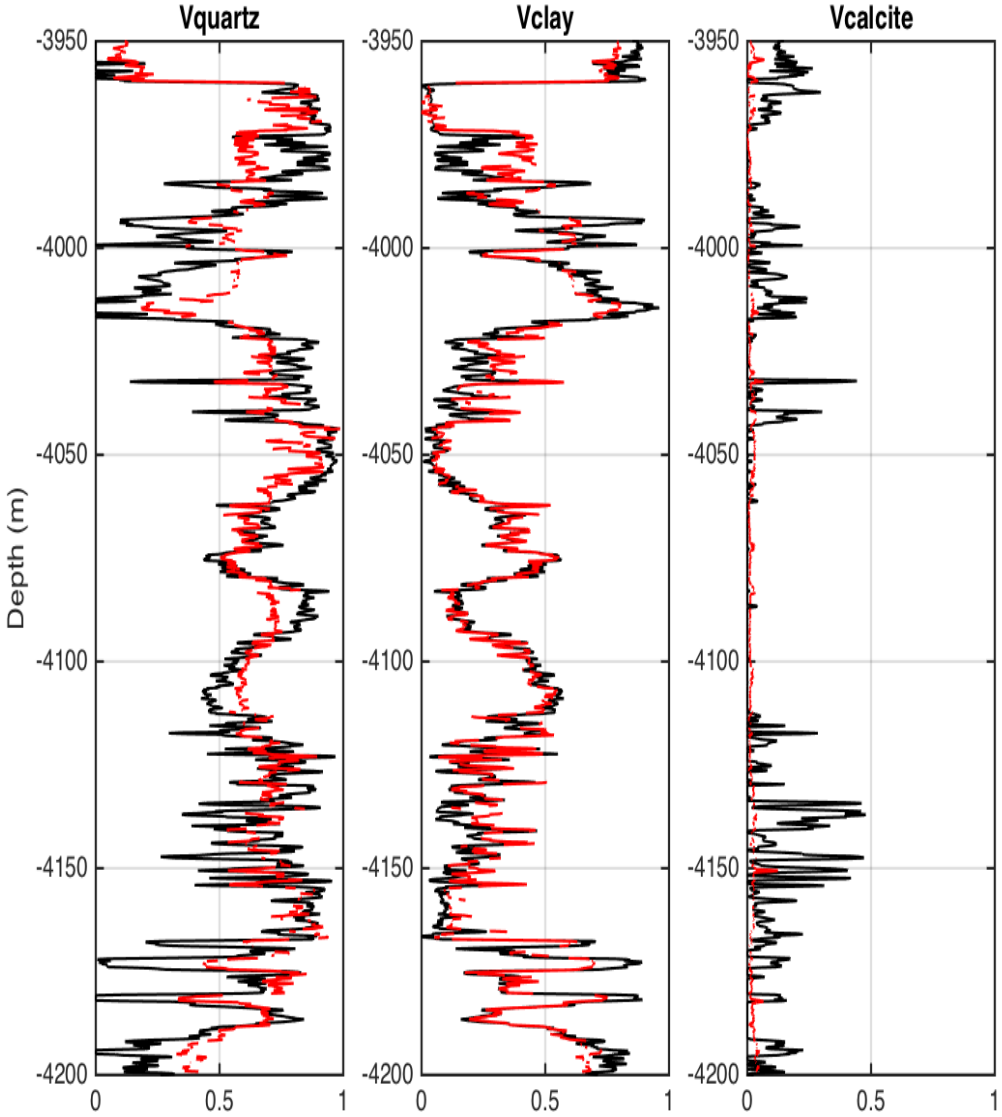


42 wells from mid-Norway

epcc

THE UNIVERSITY OF EDINBURGH

Predicting well mineral content by depth

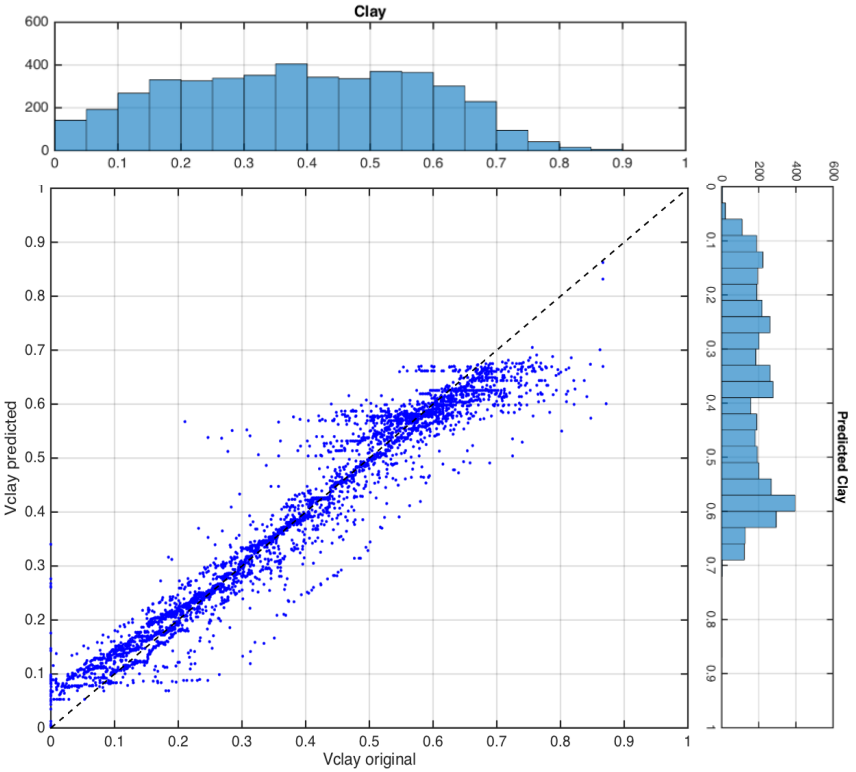


- Supervised learning using labelled data, split into 80% training and 20% test
- Fraction between 0 and 1 of minerals
- Regression using boosted trees (xgboost library)
  - Missing data limits the number of levels that neural networks can predict
- Geology is inherently biased
  - Lots of clay and quartz, other minerals far less abundant

epcc

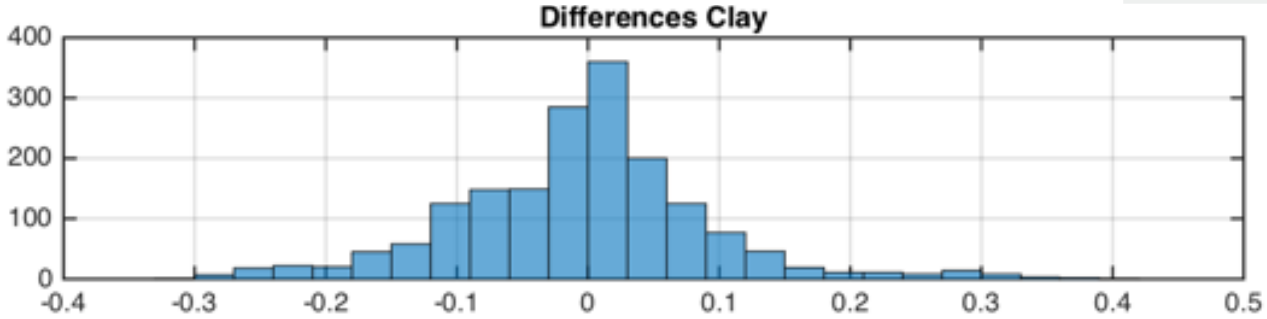
THE UNIVERSITY OF EDINBURGH

Minerology content - Clay



- From experimentation, our feeling is that DNNs give best accuracy, but they can only operate on a subset of points due to the missing data

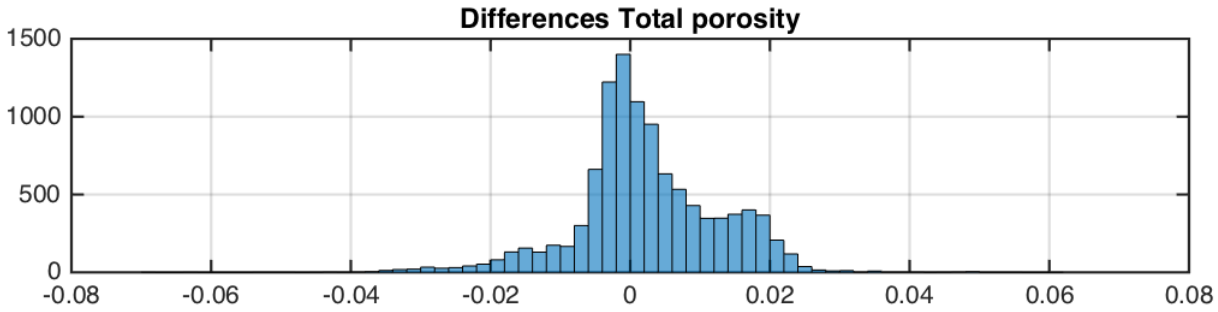
	MSE	MAE	Std dev
MLP (Sklearn)	0.0337	0.159	0.093
DNN (PyTorch)	0.0065	0.031	0.074
Boosted trees	0.0140	0.094	0.071
Boosted trees with missing data	0.0093	0.067	0.062

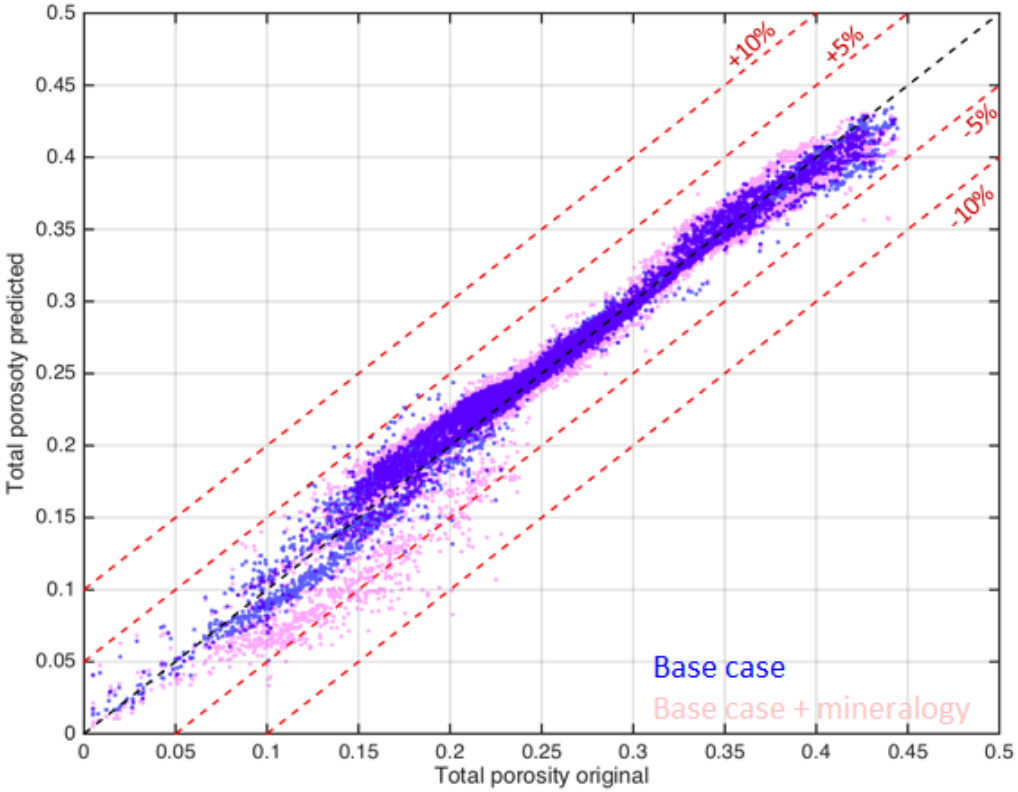


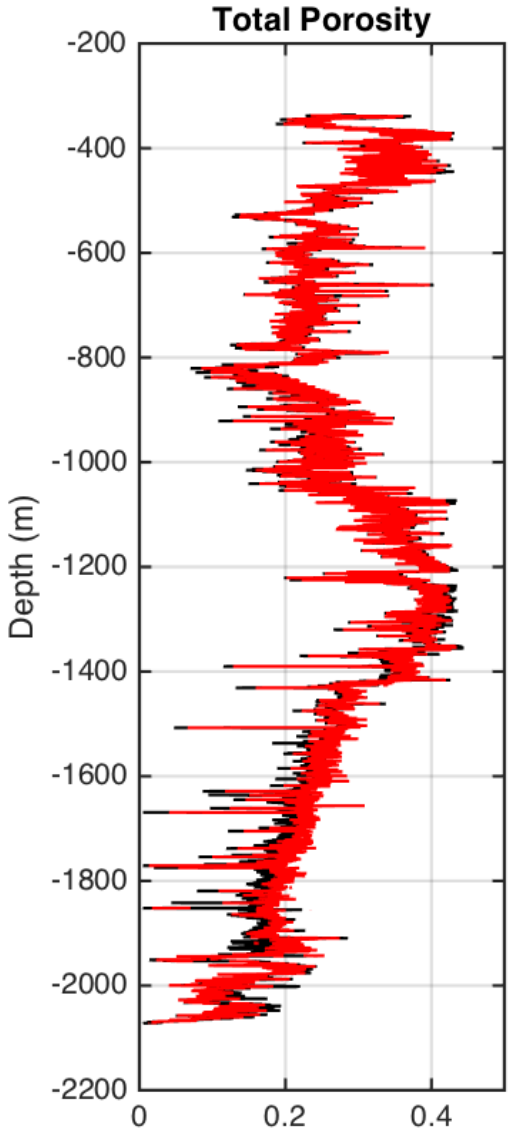
epcc

THE UNIVERSITY OF EDINBURGH

Predicting rock porosity



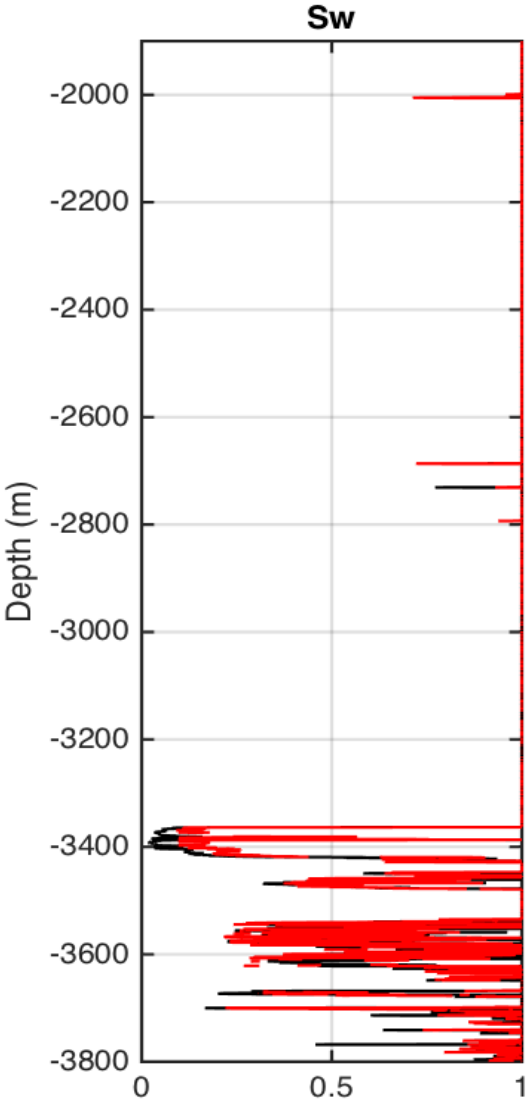




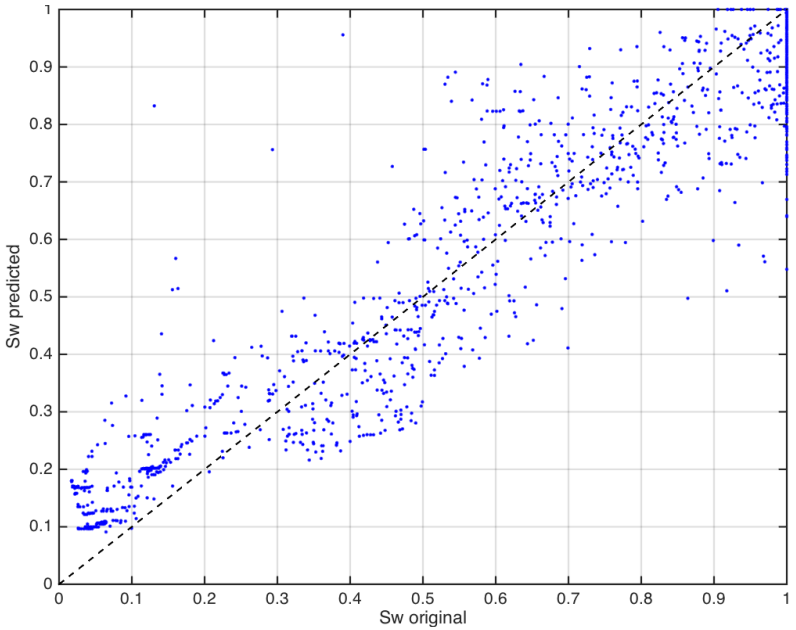
epcc

THE UNIVERSITY OF EDINBURGH

Predicting hydrocarbon saturation



- These are small reservoirs in the overall well, and the data is more accurate
  - Use DNN for classification (i.e. hydrocarbon or water)
  - Based on this use boosted trees for magnitude prediction of the water
  - 1-water is the hydrocarbon concentration



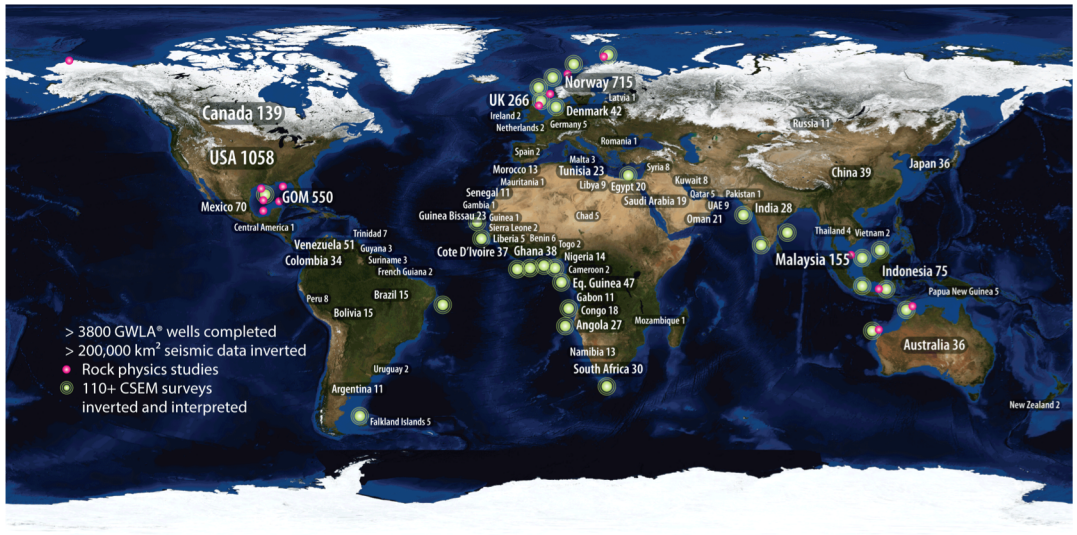
- Data is biased towards water
- Could use NN error function to address this?

epcc

THE UNIVERSITY OF EDINBURGH

Conclusions

- The good news for the petrophysicists is that they won't be replaced any time soon!
  - This is a complex and involved process which requires their knowledge and expertise, but we believe ML can assist here



  - Data set size is crucial, we think that 42 wells isn't likely enough to capture all patterns – especially with the rarer minerals

- This works as a general tool
  - A first pass to refine the data fairly significantly before interpretation
    - Optimise the use of the human
  - A quick pass for a client to check whether further analysis is like to result in success

epcc

THE UNIVERSITY OF EDINBURGH