The past present and future of AI in geosciences

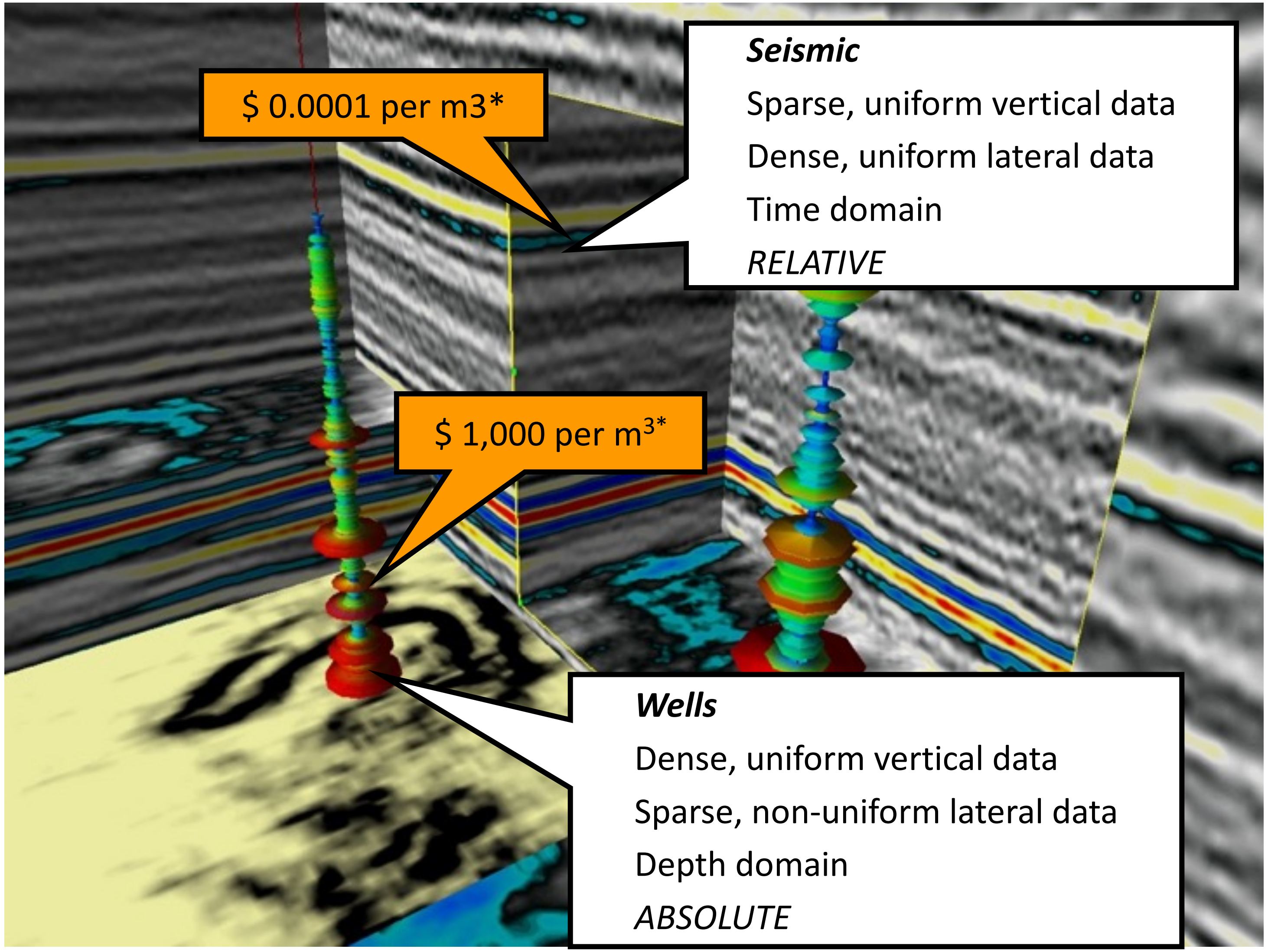
What have we learned and where are we going ?

Dr. Lucy MacGregor

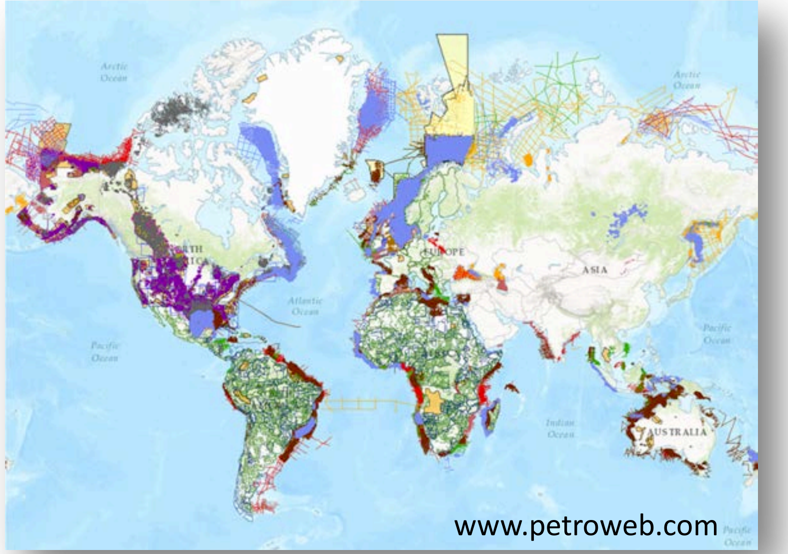
Introduction

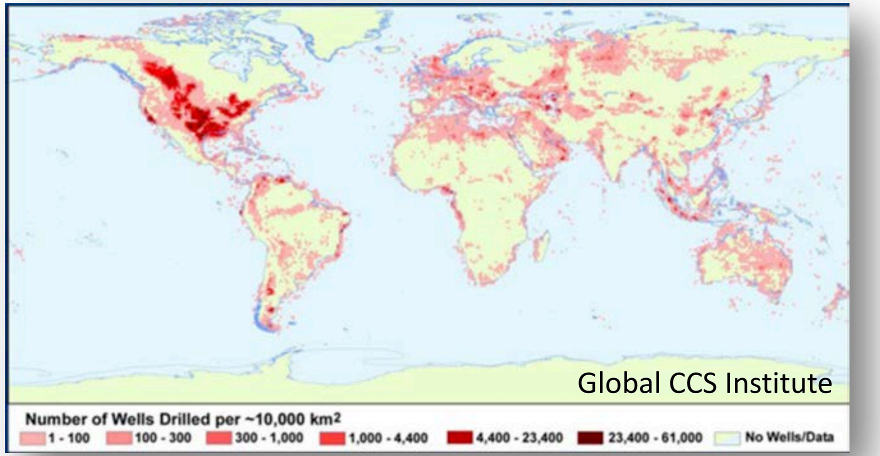
What data are we dealing with ?

- Seismic data
- Well log data
- Electromagnetic data
- Gravity data
- Magnetics data
- Geological information
- Cores
- Sediment samples
- .....



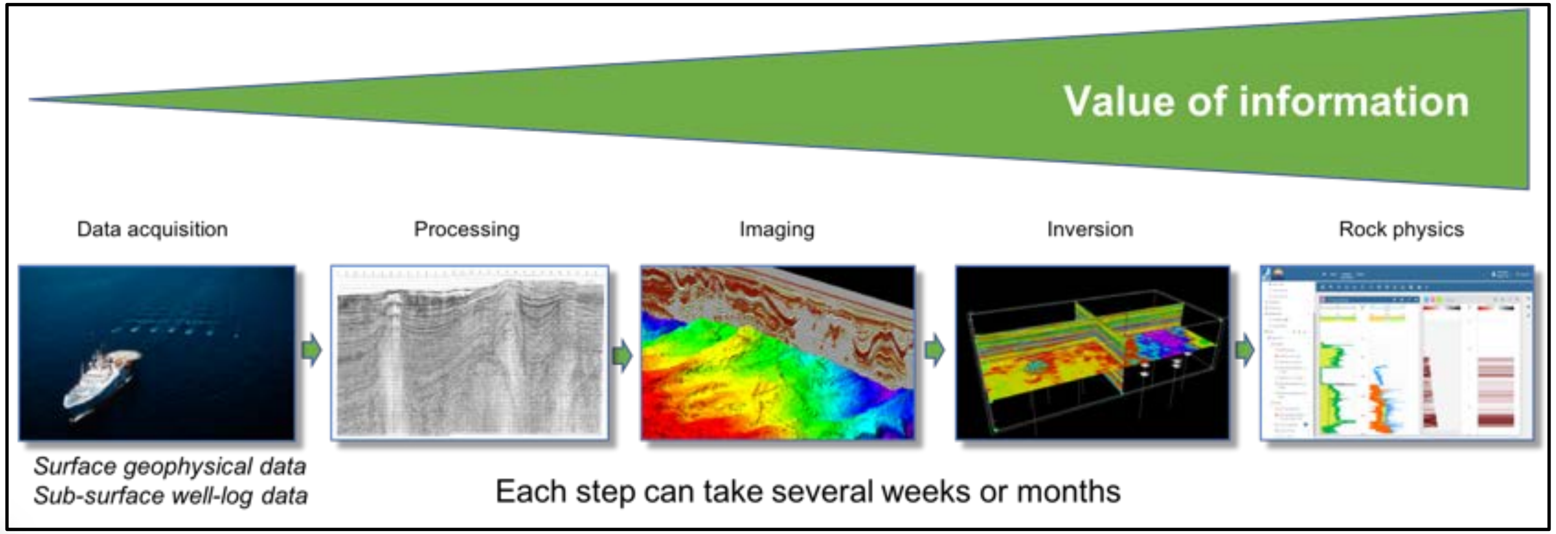
The problem with data:

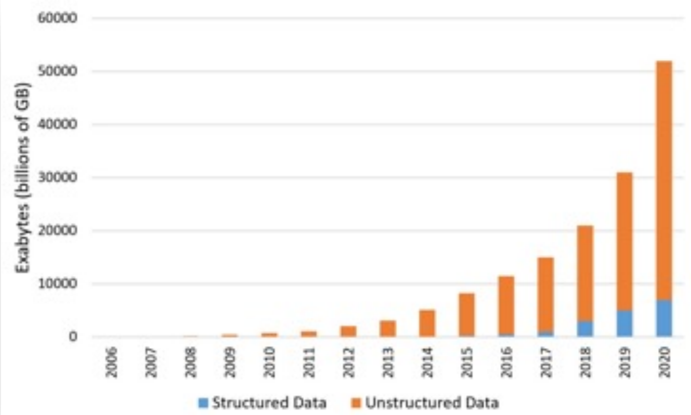




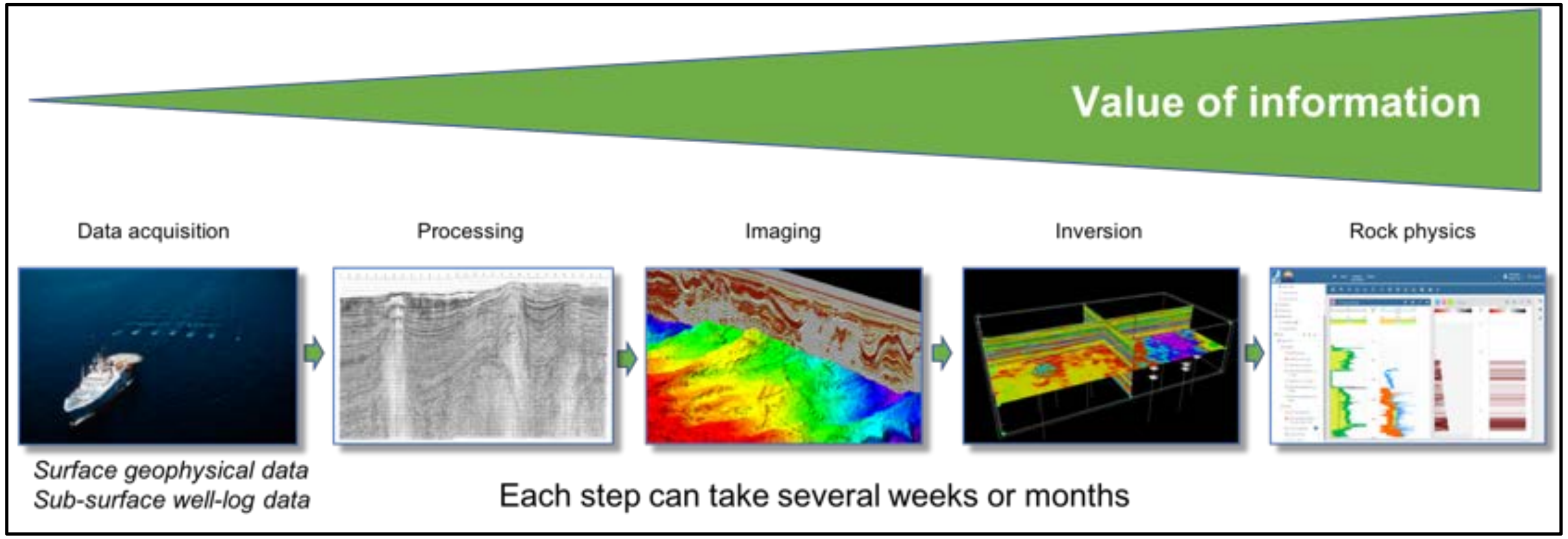
We have a lot of data !

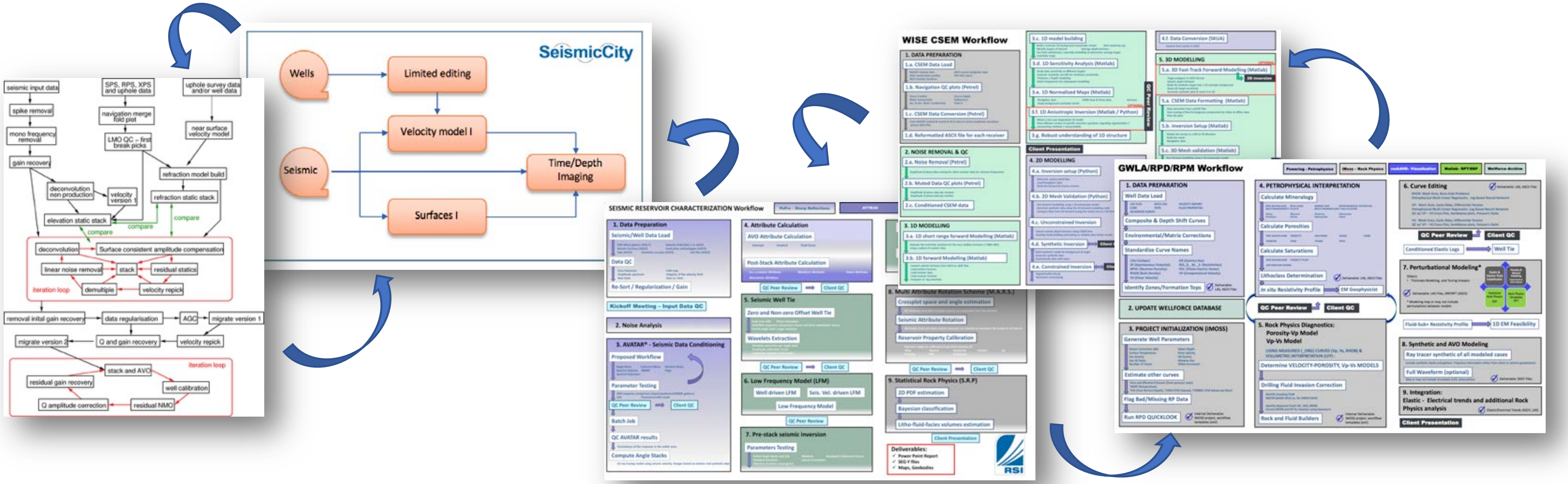
....but it's a bit of a mess...





The problem with data:





Overview

Past
- Machine learning in geoscience
- Neural networks

Present
- Prediction of properties from logs

Future
- Uncertainty management in geomodelling

Overview

Past

- Machine learning in geoscience
- Neural networks

Present

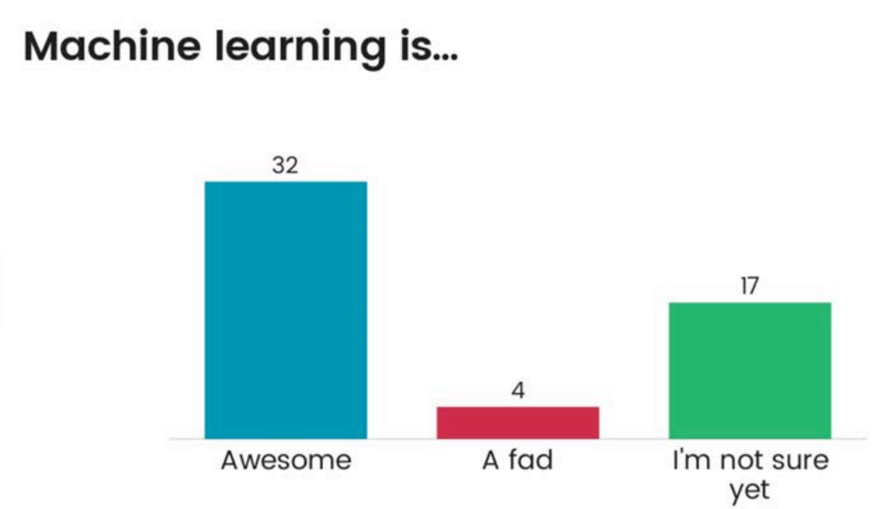
- Prediction of properties from logs

Future

- Uncertainty management in geomodelling

Machine learning in geoscience

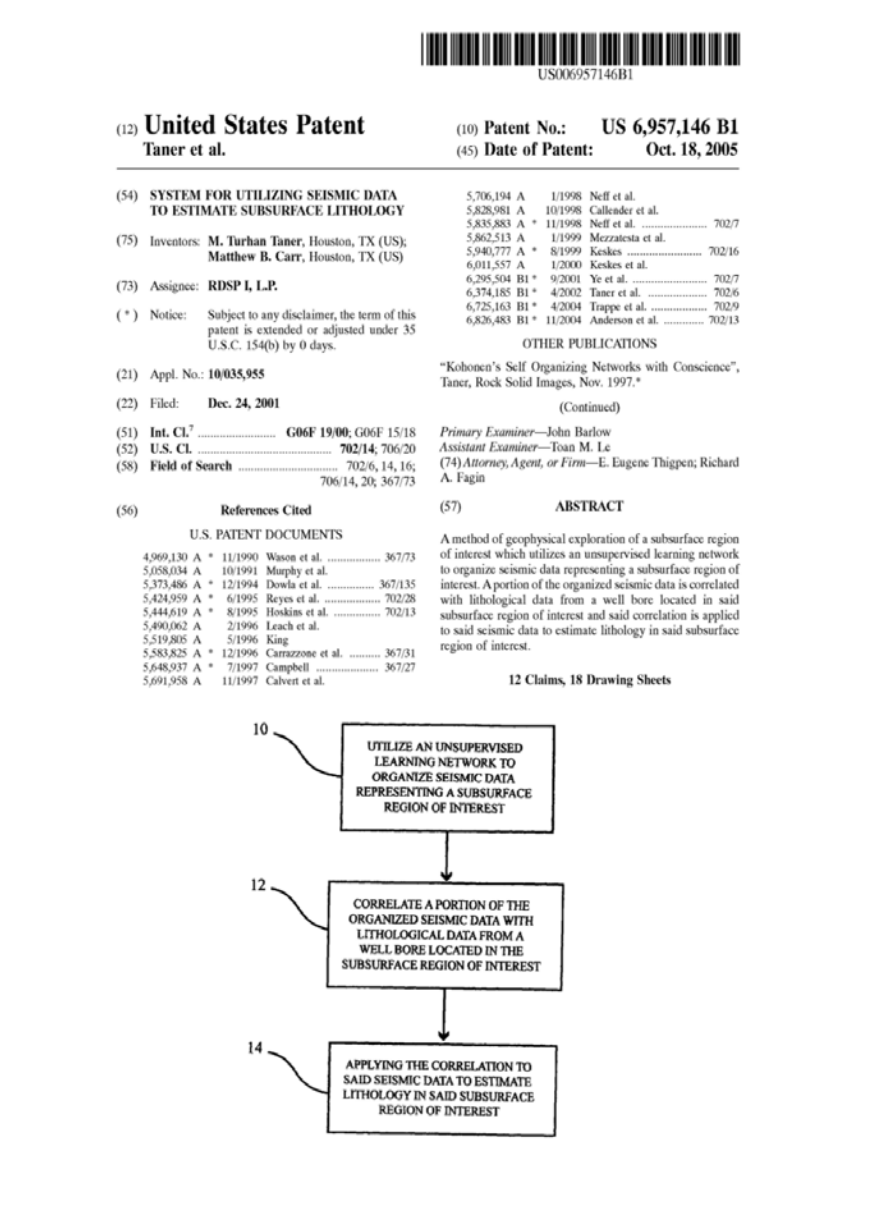
- Most of the steps required in geoscience interpretation involve pattern recognition or classification.
- Many machine learning algorithms boil down to a regularized inversion for a function linking predictors and outputs.
- As geoscientists we've inverting things for years !
- However as an oil industry we've been overtaken by e.g. google, amazon….



www.agilescientific.com, Survey at Paris EAGE workshop on AI

An example: LITHANN®

- Neural network classification of seismic attributes.
- Developed by Tury Taner in the late 1990s as part of RSI’s LFP consortium and patented.
- Uses a Kohonen self organizing map
- Still in use today !



US006957146B1

(12) **United States Patent**
Taner et al.

(10) **Patent No.:** **US 6,957,146 B1**
(45) **Date of Patent:** **Oct. 18, 2005**

(54) **SYSTEM FOR UTILIZING SEISMIC DATA TO ESTIMATE SUBSURFACE LITHOLOGY**

(75) Inventors: **M. Turhan Taner**, Houston, TX (US); **Matthew B. Carr**, Houston, TX (US)

(73) Assignee: **RDSP I, L.P.**

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 0 days.

(21) Appl. No.: **10/035,955**

(22) Filed: **Dec. 24, 2001**

(51) **Int. Cl.**[7] **G06F 19/00**; G06F 15/18
(52) **U.S. Cl.** **702/14**; 706/20
(58) **Field of Search** 702/6, 14, 16; 706/14, 20; 367/73

(56) **References Cited**

U.S. PATENT DOCUMENTS

4,969,130	A *	11/1990	Wason et al.	367/73
5,058,034	A	10/1991	Murphy et al.	
5,373,486	A *	12/1994	Dowla et al.	367/135
5,424,959	A *	6/1995	Reyes et al.	702/28
5,444,619	A *	8/1995	Hoskins et al.	702/13
5,490,062	A	2/1996	Leach et al.	
5,519,805	A	5/1996	King	
5,583,825	A *	12/1996	Carrazzone et al.	367/31
5,648,937	A *	7/1997	Campbell	367/27
5,691,958	A	11/1997	Calvert et al.	
5,706,194	A	1/1998	Neff et al.	
5,828,981	A	10/1998	Callender et al.	
5,835,883	A *	11/1998	Neff et al.	702/7
5,862,513	A	1/1999	Mezzatesta et al.	
5,940,777	A *	8/1999	Keskes	702/16
6,011,557	A	1/2000	Keskes et al.	
6,295,504	B1 *	9/2001	Ye et al.	702/7
6,374,185	B1 *	4/2002	Taner et al.	702/6
6,725,163	B1 *	4/2004	Trappe et al.	702/9
6,826,483	B1 *	11/2004	Anderson et al.	702/13

OTHER PUBLICATIONS

“Kohonen’s Self Organizing Networks with Conscience”, Taner, Rock Solid Images, Nov. 1997.*

(Continued)

Primary Examiner—John Barlow
Assistant Examiner—Toan M. Le
(74) *Attorney, Agent, or Firm*—E. Eugene Thigpen; Richard A. Fagin

(57) **ABSTRACT**

A method of geophysical exploration of a subsurface region of interest which utilizes an unsupervised learning network to organize seismic data representing a subsurface region of interest. A portion of the organized seismic data is correlated with lithological data from a well bore located in said subsurface region of interest and said correlation is applied to said seismic data to estimate lithology in said subsurface region of interest.

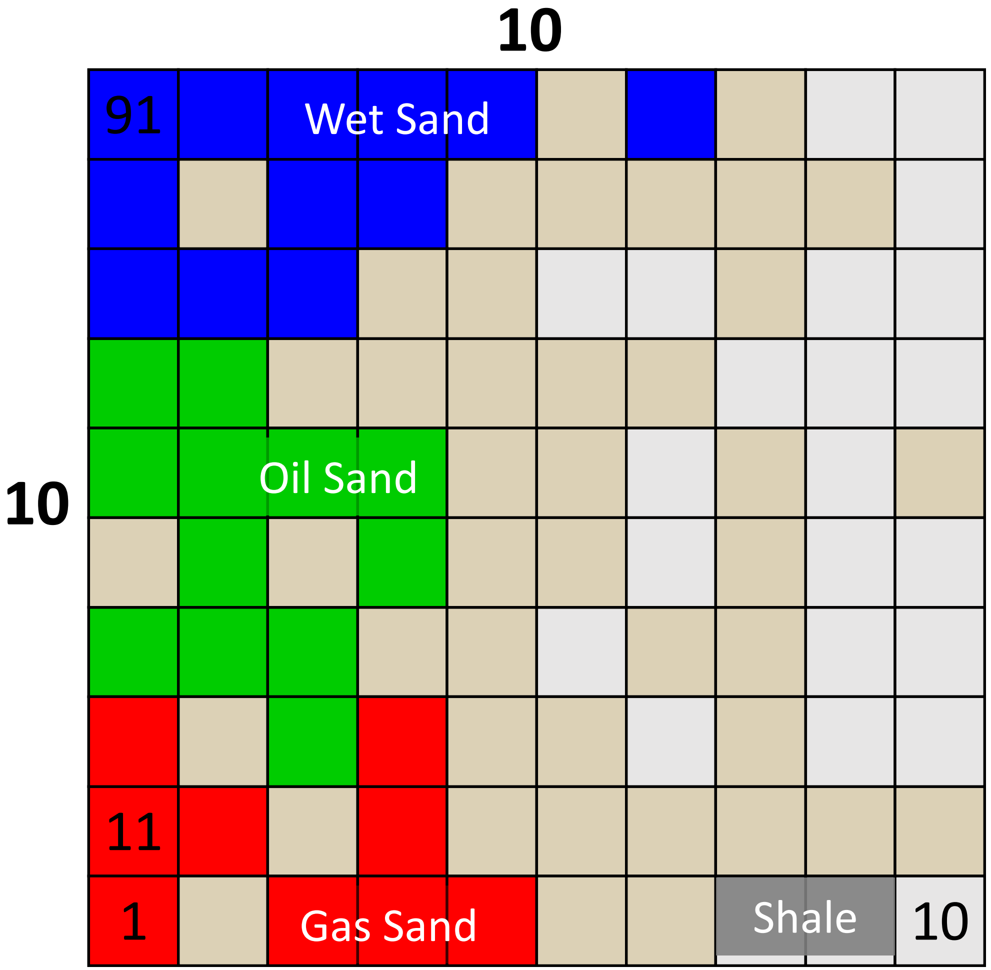
12 Claims, 18 Drawing Sheets

10 UTILIZE AN UNSUPERVISED LEARNING NETWORK TO ORGANIZE SEISMIC DATA REPRESENTING A SUBSURFACE REGION OF INTEREST

12 CORRELATE A PORTION OF THE ORGANIZED SEISMIC DATA WITH LITHOLOGICAL DATA FROM A WELL BORE LOCATED IN THE SUBSURFACE REGION OF INTEREST

14 APPLYING THE CORRELATION TO SAID SEISMIC DATA TO ESTIMATE LITHOLOGY IN SAID SUBSURFACE REGION OF INTEREST

How Do KSOM's Work?



Kohonen self-organizing map algorithm posts ANN classifications of multiple seismic attributes within the confines of a 2D topology

Samples (seismic facies) exhibiting a similar multi-attribute signature are mapped into contiguous regions of Kohonen space.

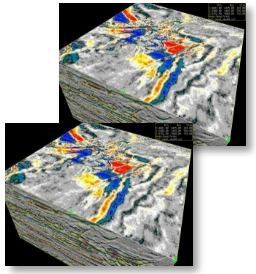
10 X 10

(100 Classes)

2D Network Topology

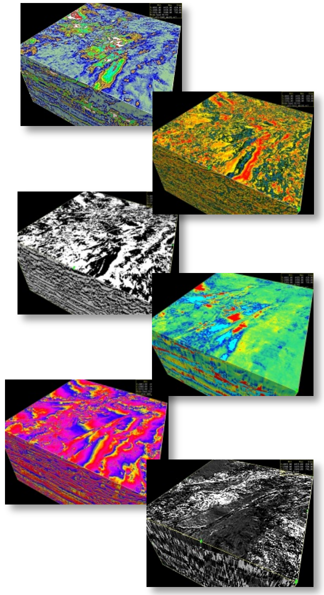
LITHANN® workflow

Conditioning



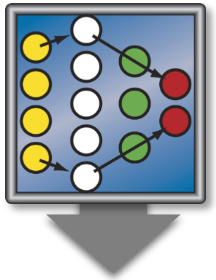
Multiple pre-stack
&/or poststack
seismic volume(s)

Calculation

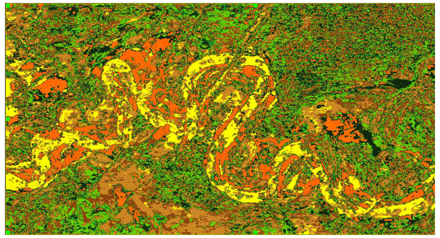


instantaneous, elastic,
geometric, waveshape,
attribute volumes

Classification

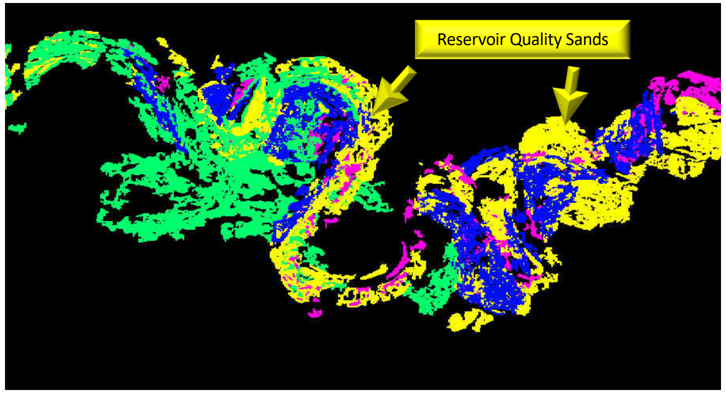


1 classified volume:
2-D topology, 64, 100, 144, or 225
classes
unsupervised or
supervised



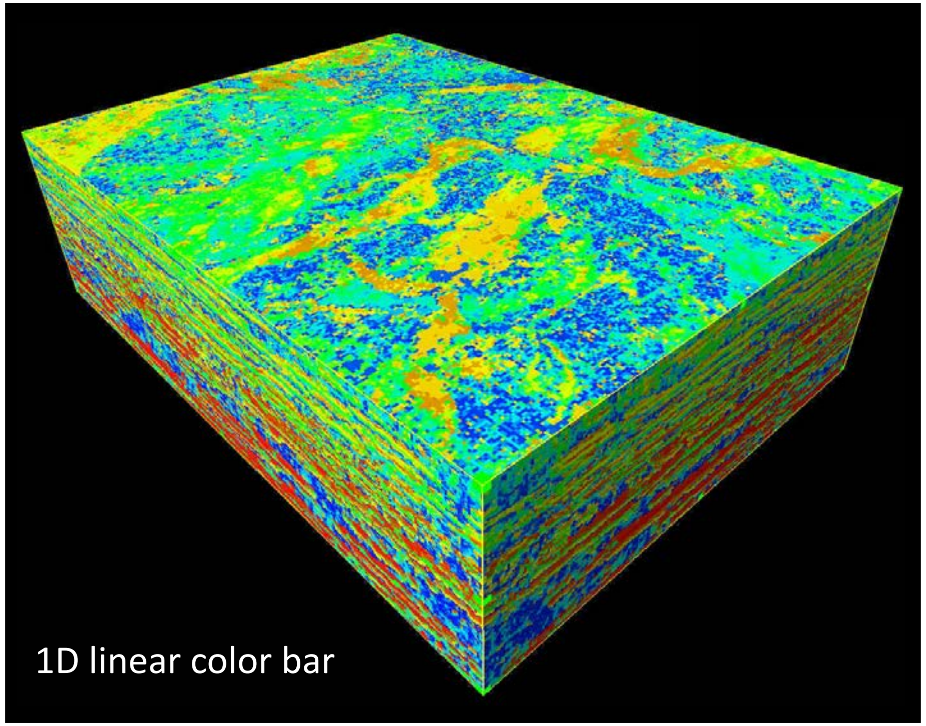
Neural net

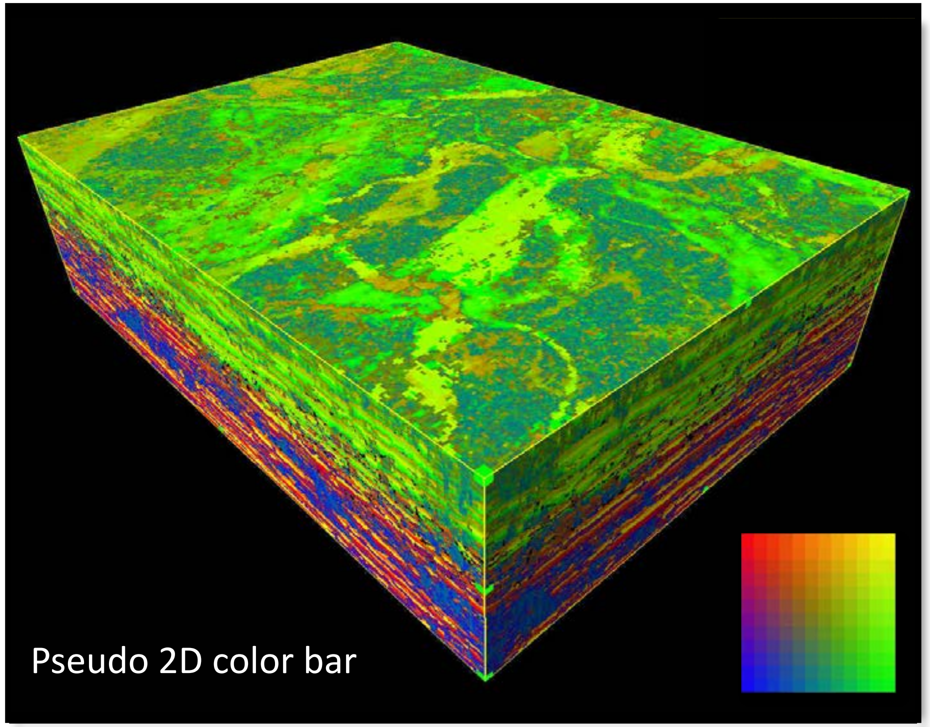
Calibration

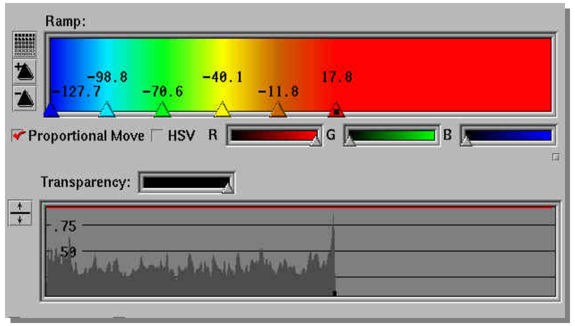


1 seismic facies volume:
3 calibrated classes

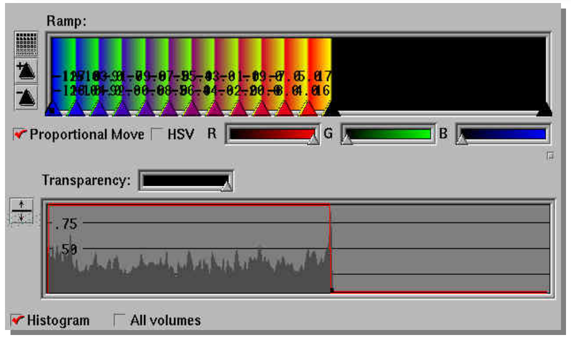
Seismic Facies Visualization







Apparent decrease in spatial resolution



Apparent increase in spatial resolution

Overview

Past
- Machine learning in geoscience
- Neural networks

Present
- Prediction of properties from logs

Future
- Uncertainty management in geomodelling

There has been an explosion of ML libraries:

- Most are open source: access to numerous advanced algorithms and a community of users
- (Relatively) easy to use
- ….but the key is training data

scikit learn

PyTorch

XGBoost

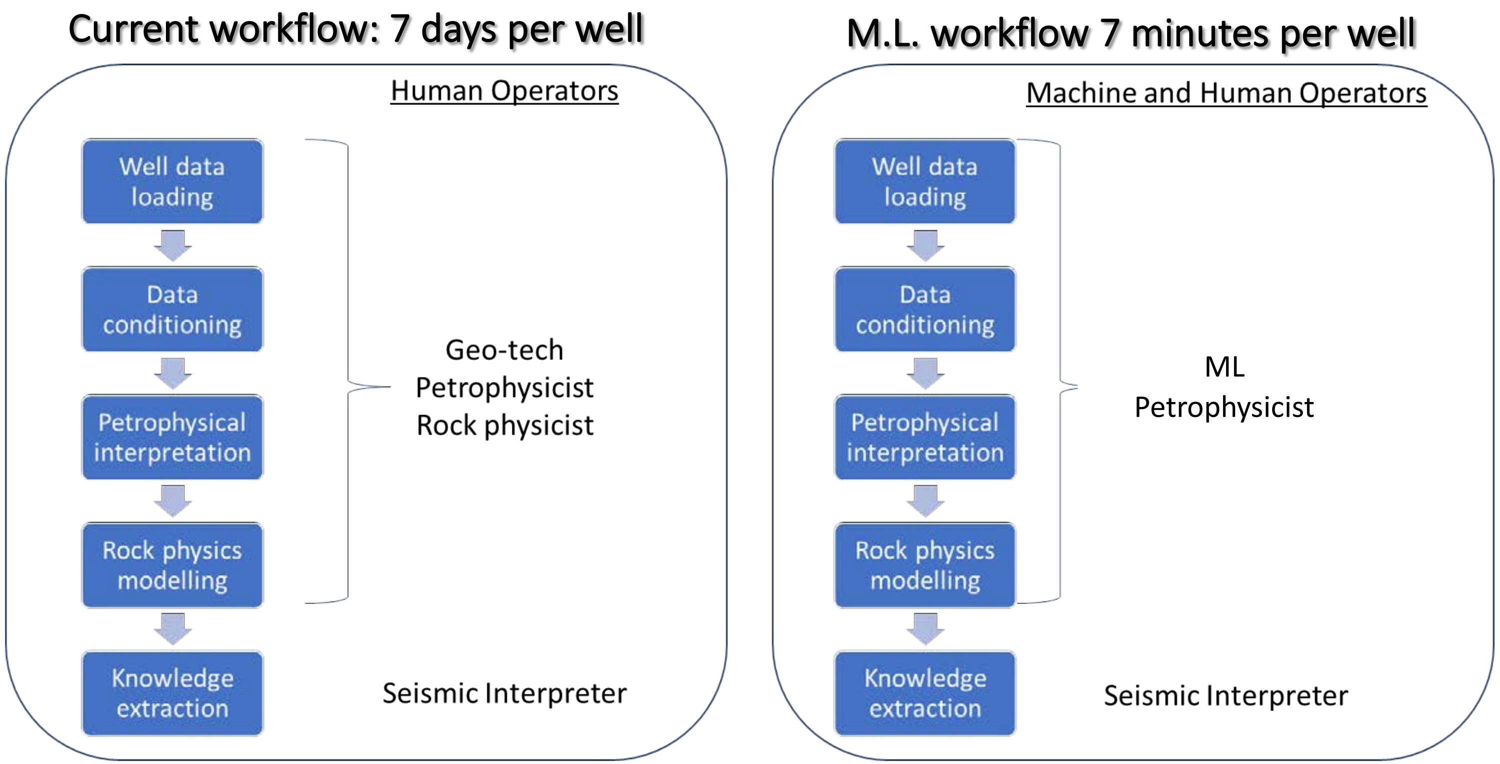
theano

K Keras

Caffe

Project SWOOP: Streamlined WOrkflows for Optimsed Petrophysics

The project was conducted in partnership with EPCC – Edinburgh University's AI team in the high performance computing centre.



OGIC Oil & Gas Innovation Centre

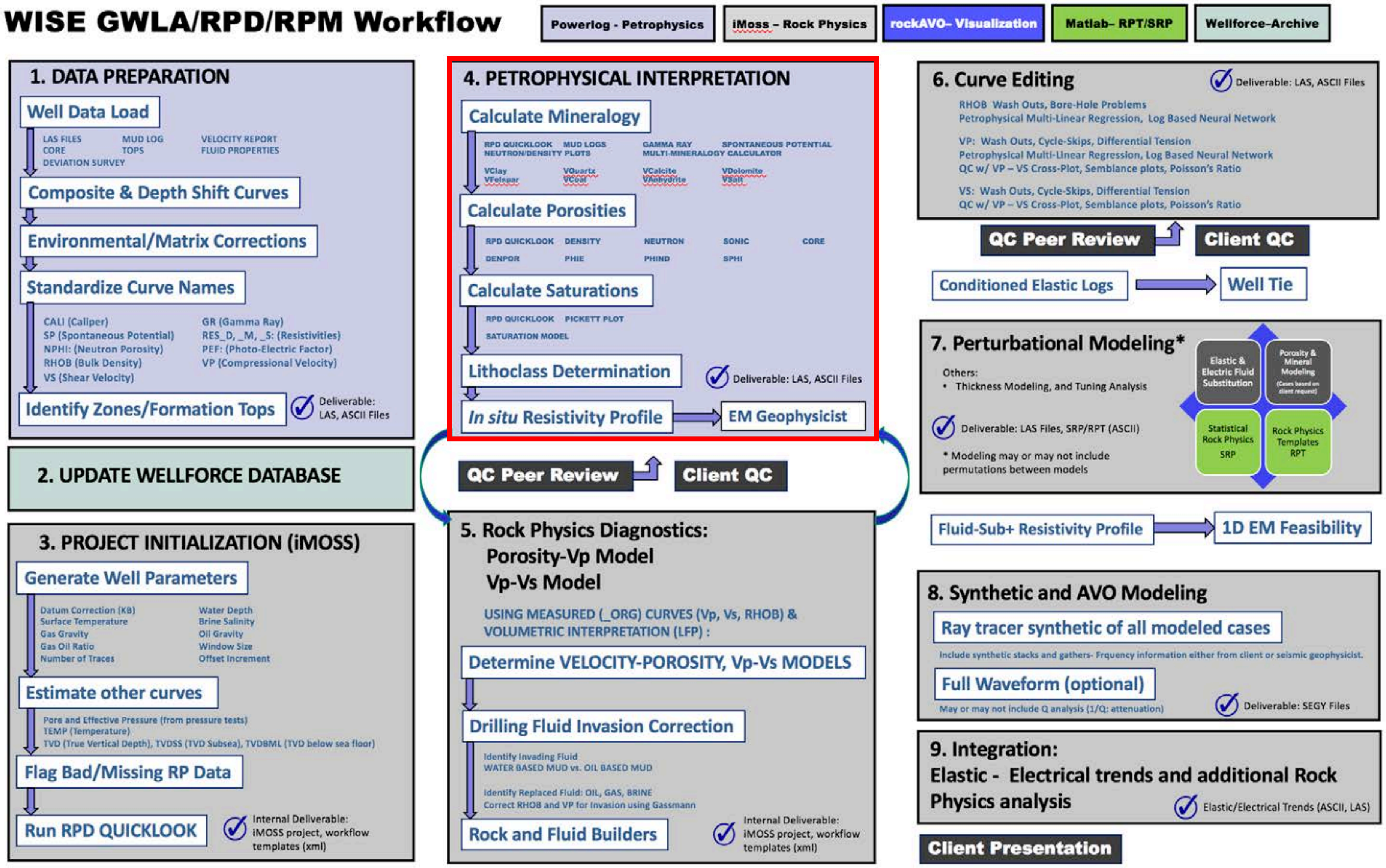
fostering and funding innovation

epcc

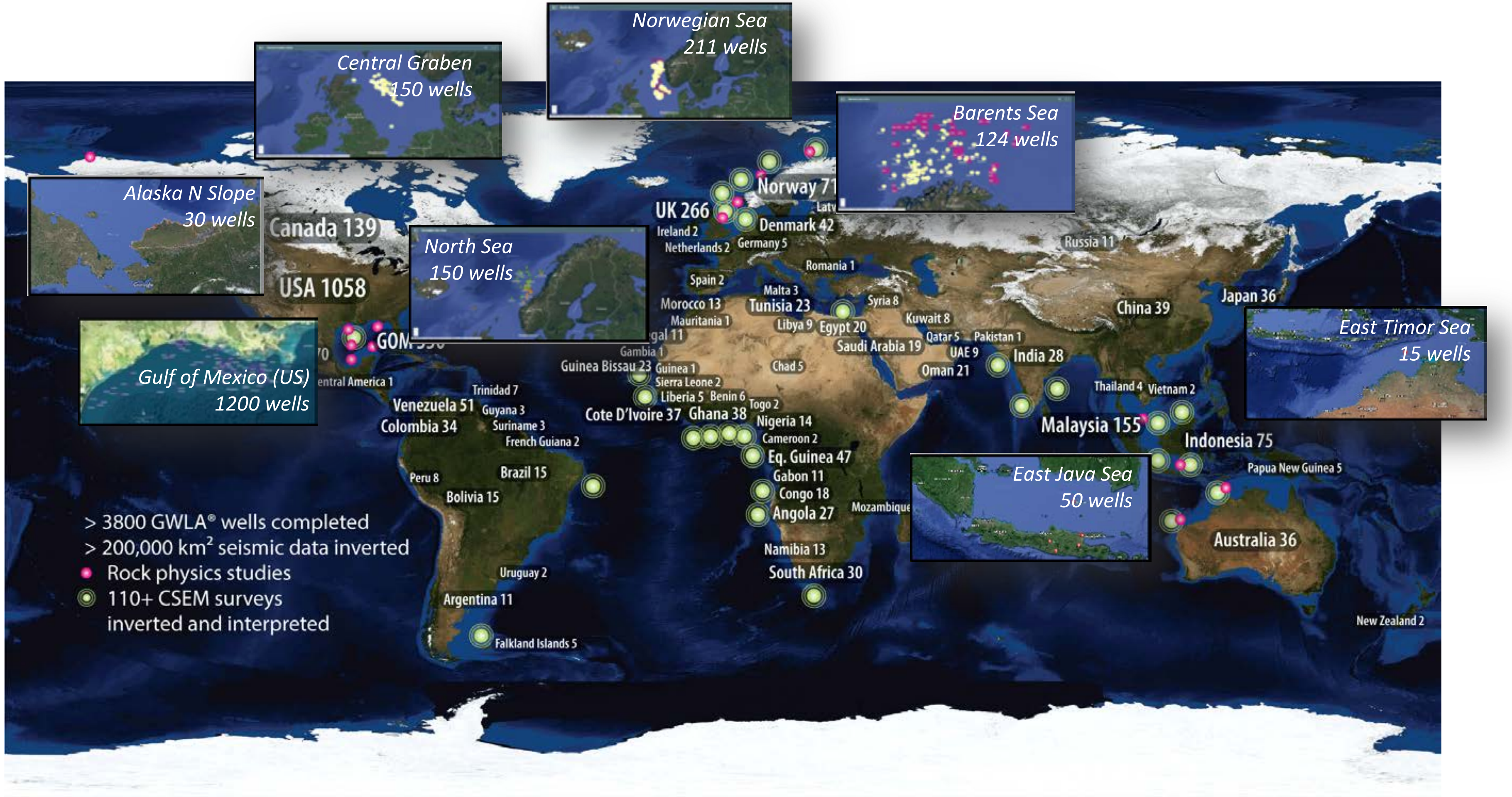
RSI

- Streamlining petrophysical workflows with machine learning, MacGregor, L., Brown, N., Roubickova, A., Lampaki, I., Berrizbeitia, J., Ellis, M. & Nichols, K., 2019, EAGE Conference on Reservoir Geoscience, December 2018
- Machine learning on Crays to optimise petrophysical workflows in oil and gas exploration, Brown, N., MacGregor, L., Roubickova, A., Lampaki, I., Ellis, M., Vera de Newton, P., 2019, Journal of Distributed and Parallel Computing, subm.

Petrophysics/rock physics workflow



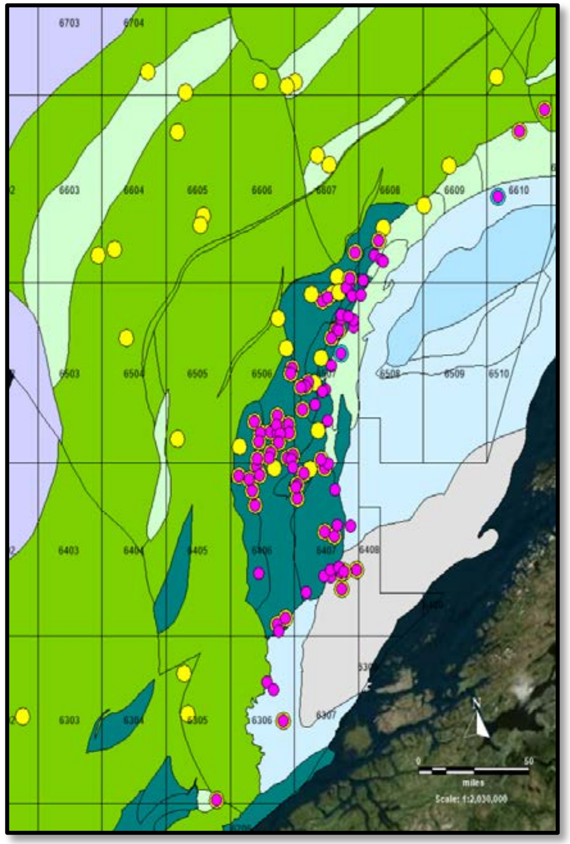
Training data: must be CONSISTENT and COMPLETE



2000 wells in the existing multi-client database

Starting point – training data and ML algorithm

Training data:



64 wells from mid-Norway

- Which algorithm ? We have tried:
  - Dirichlet regression (didn't work very well)
  - Deep Neural Network (current second favourite)
  - Multi-layer perceptron
  - Boosted trees (current favourite)
- How to choose ?
  - Currently looking into a parameter optimisation approach.

Optimising hyperparameters

Number of boosting rounds to perform

- num_rounds: between 10 and 150

Controls model complexity

- max_depth: between 1 and 30
- min_child_weight: between 1 and 20
- gamma: between 0.1 and 1

Maximum depth of the tree (increasing makes the model more likely to overfit.)

Minimum sum of the weight needed for children, the larger this is the more conservative the model

Minimum loss reduction needed to make a new leaf node, the larger this is the more conservative the model

Adds randomness for robustness

- eta: between 0.025 and 0.8
- subsample: between 0.2 and 1
- colsample_bytree: between 0.1 and 1

Step size shrinkage used to prevent overfitting

Ratio of amount of data to sample when growing trees, used to prevent overfitting

Subsample ratio of columns when constructing each tree

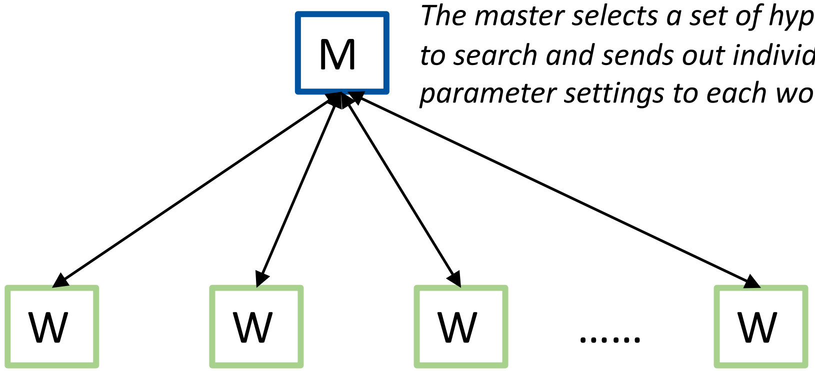
Clay: num_rounds=142, max_depth=3, min_child_weight=3.5, gamma=0.6, eta=0.05, subsample=0.7, colsample_bytree=0.9
Quartz: num_rounds=120, max_depth=5, min_child_weight=17.5, gamma=0.2, eta=0.15, subsample=1.0, colsample_bytree=0.1
Calcite: num_rounds=135, max_depth=3, min_child_weight=14.5, gamma=1.0, eta=0.075, subsample=0.4, colsample_bytree=0.75

Use hyperopt to search through the parameter space

- Select five different splits between training and test data and for each parameter setting train and test the model with each of these
- Looking to optimise the mean absolute error between prediction and truth here and hyperopt will direct its searching based on these values
- Whilst each of these hyperparameters has a specific meaning, they are interlinked and the best settings are not always obvious (so best to find empirically.)

Speeding up hyperparameter searching

- Searching for hyperparameters can take a long time – with the largest number of input curves around 20 minutes per parameter
  - Which means only around 72 parameters can be searched in a day
  - Annoyingly the XGBoost library is not parallelised over multiple CPU cores and the GPU support is quite immature (it tends to run out of memory on our GPU machine.)
- Using Cirrus, an HPC machine and even in a single node there are 36 physical cores. But hyperopt is not able to take advantage of these cores.
- Therefore we added in support for MPI parallelisation into hyperopt to enable it to search the space in parallel across all cores

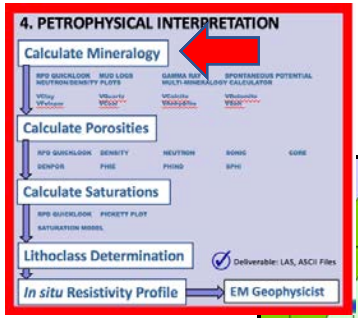


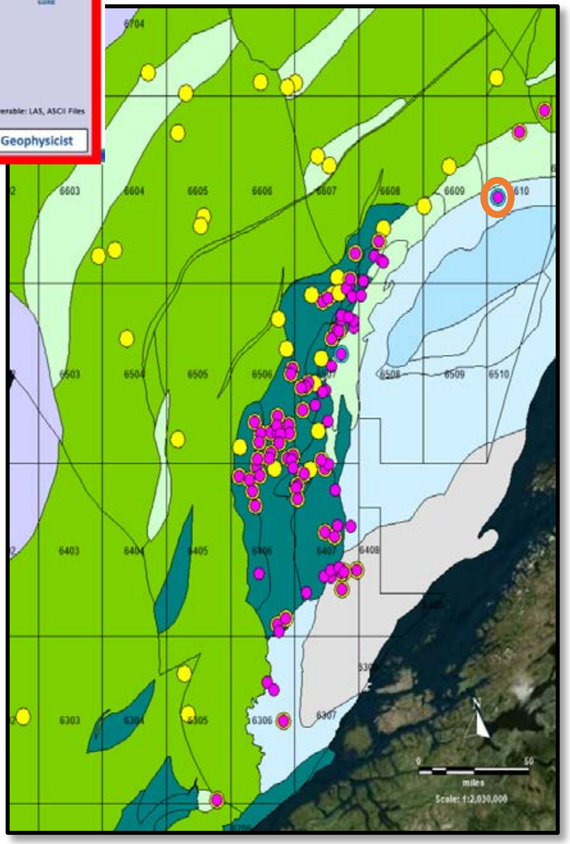
The master selects a set of hyperparameters to search and sends out individual parameter settings to each worker.

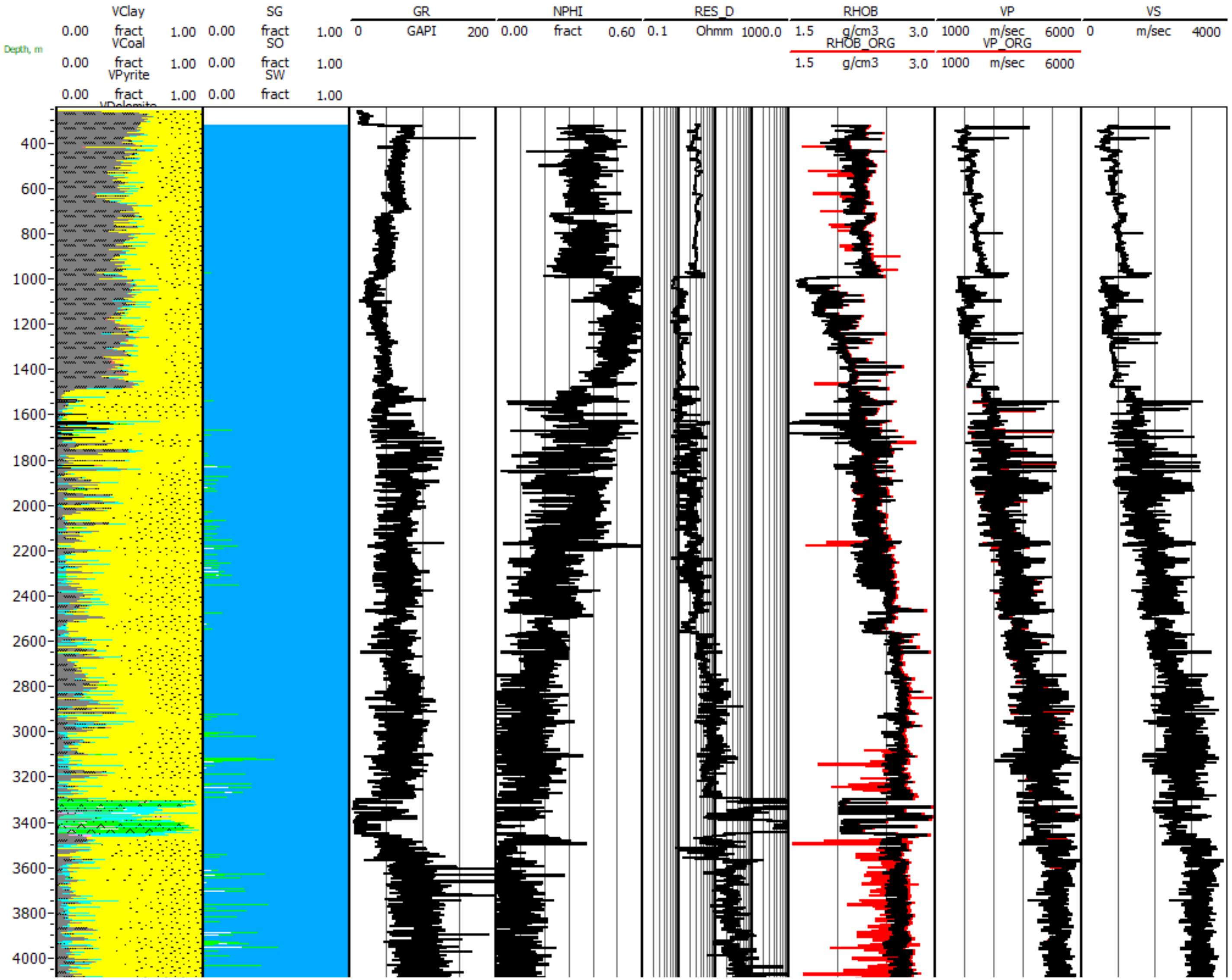
Workers execute with the parameters supplied by the master and send back the corresponding absolute error. The master then looks at all of these, decides a new set of hyperparameters and sends these out to again be processed

- Based on this parallelisation we can run around 2,500 parameter searches in a day now on a node of Cirrus
  - We can scale up to more nodes (Cirrus has 280 nodes, so based on this it would be possible to do around 20,000 searches in day) but this isn't needed at the moment
  - We tend to find the optimum parameters in around 500 searches, so in reality it means we only need to run the script for about 4 or 5 hours

Mineralogy: Test on a single well

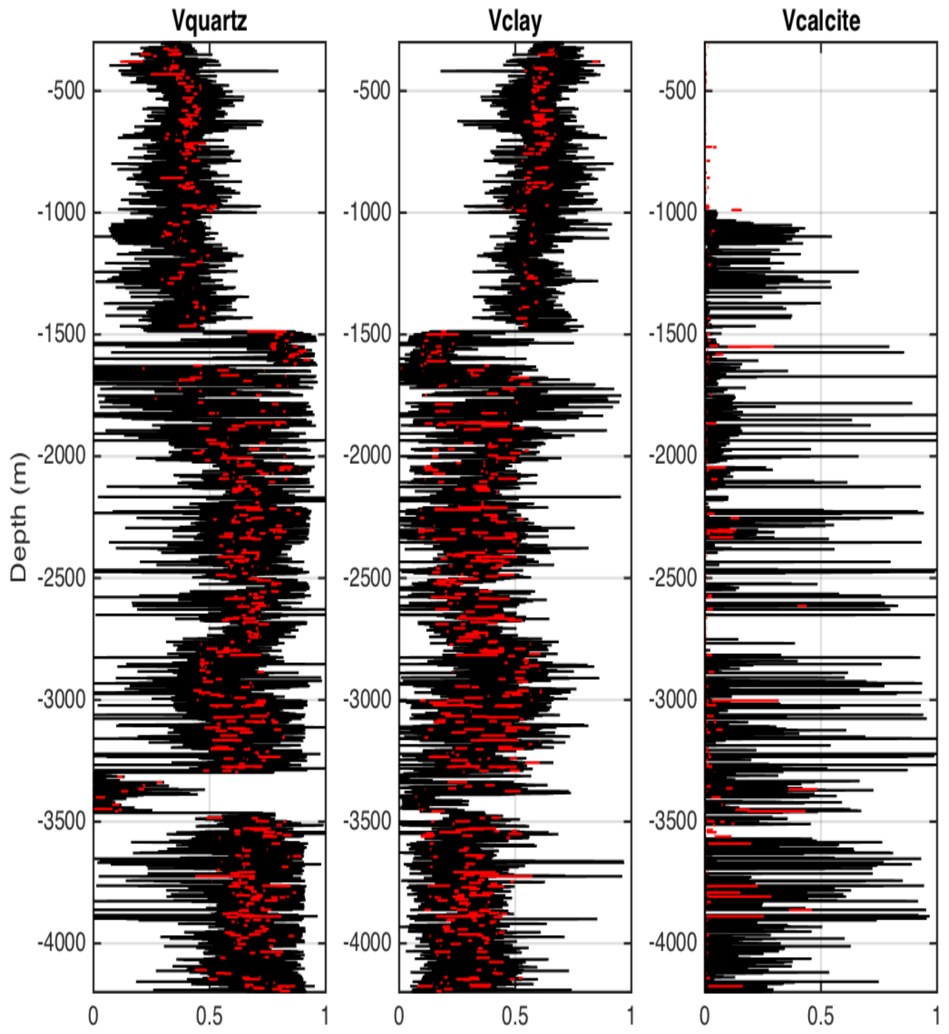


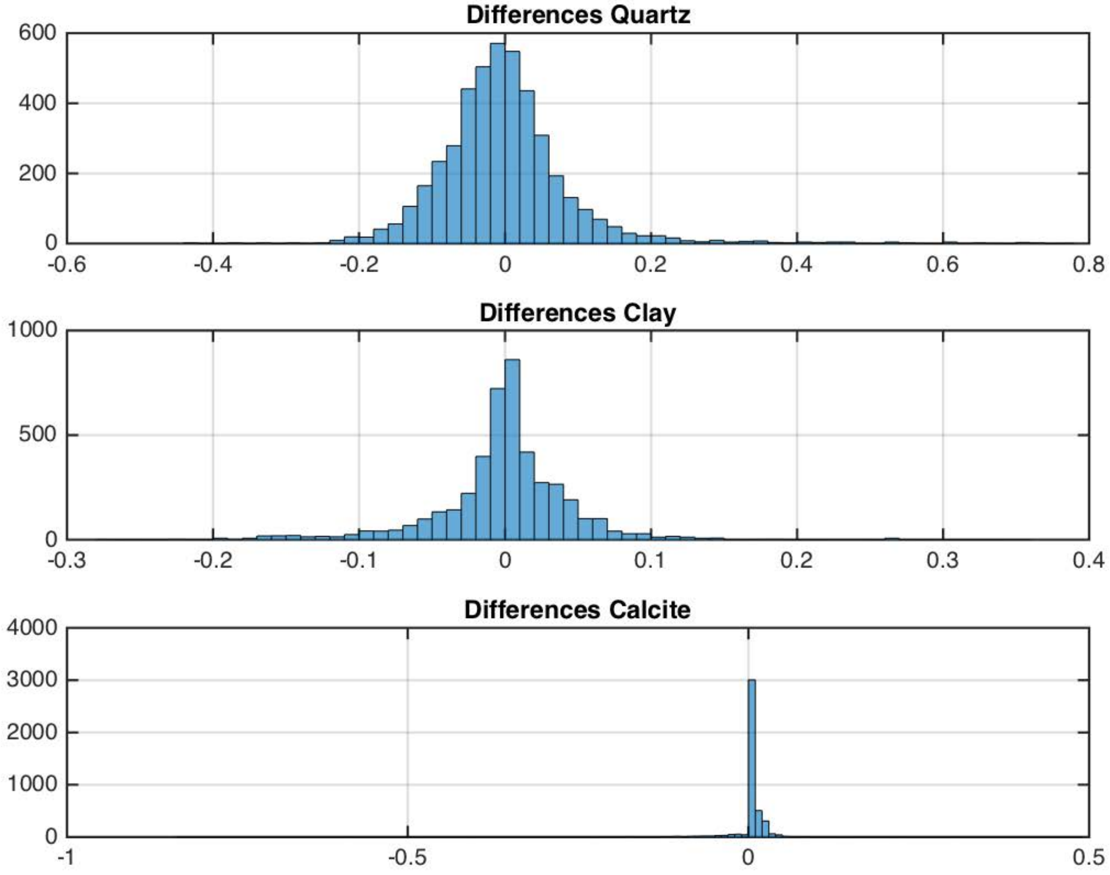




- Look at a single well: for this case the training data is as good as it can be !
- 80% of the points for training, 20% for testing.

Results

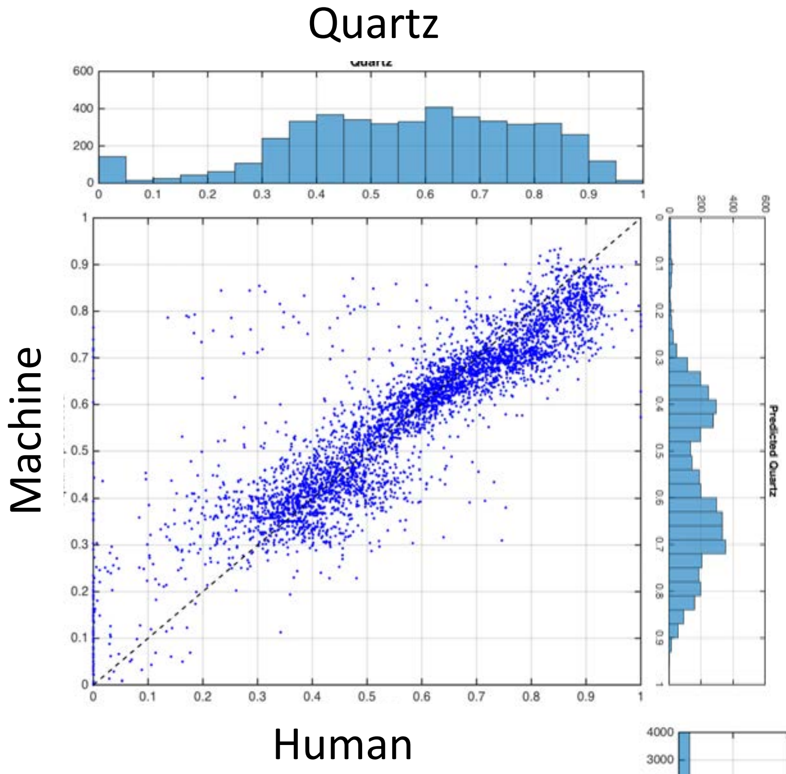




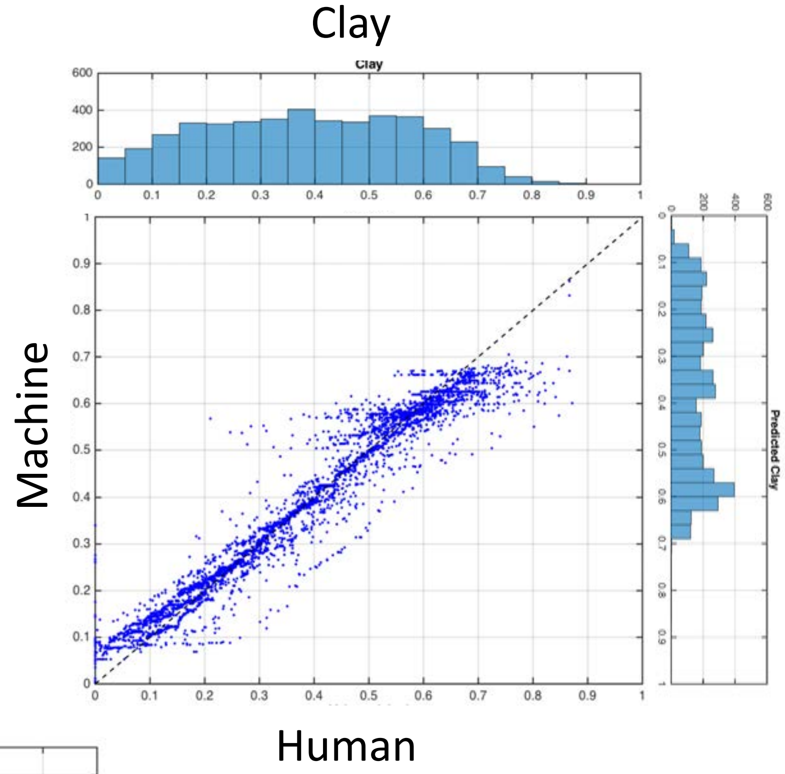
Original petrophysical workflow

ML workflow

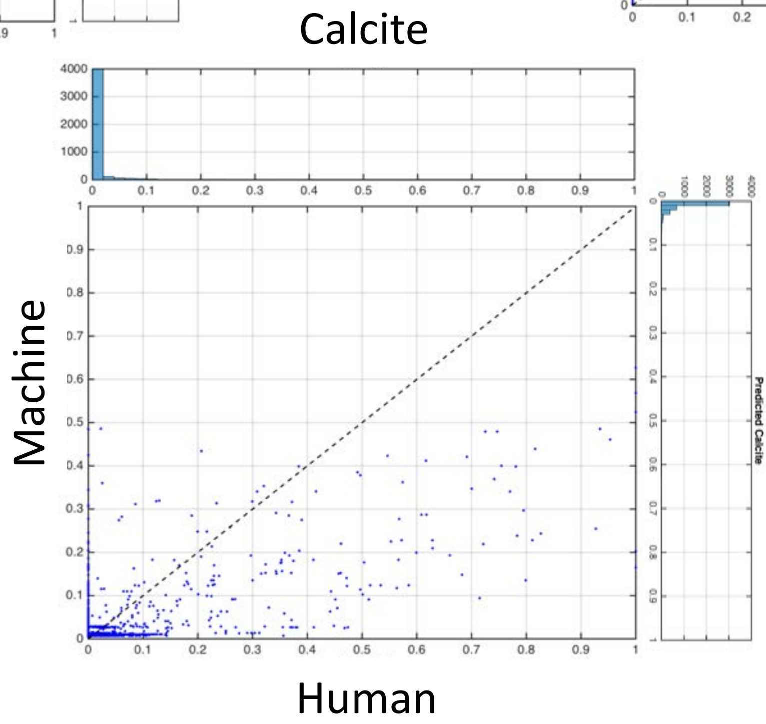
Results



Quartz
Machine
Human
Predicted Quartz



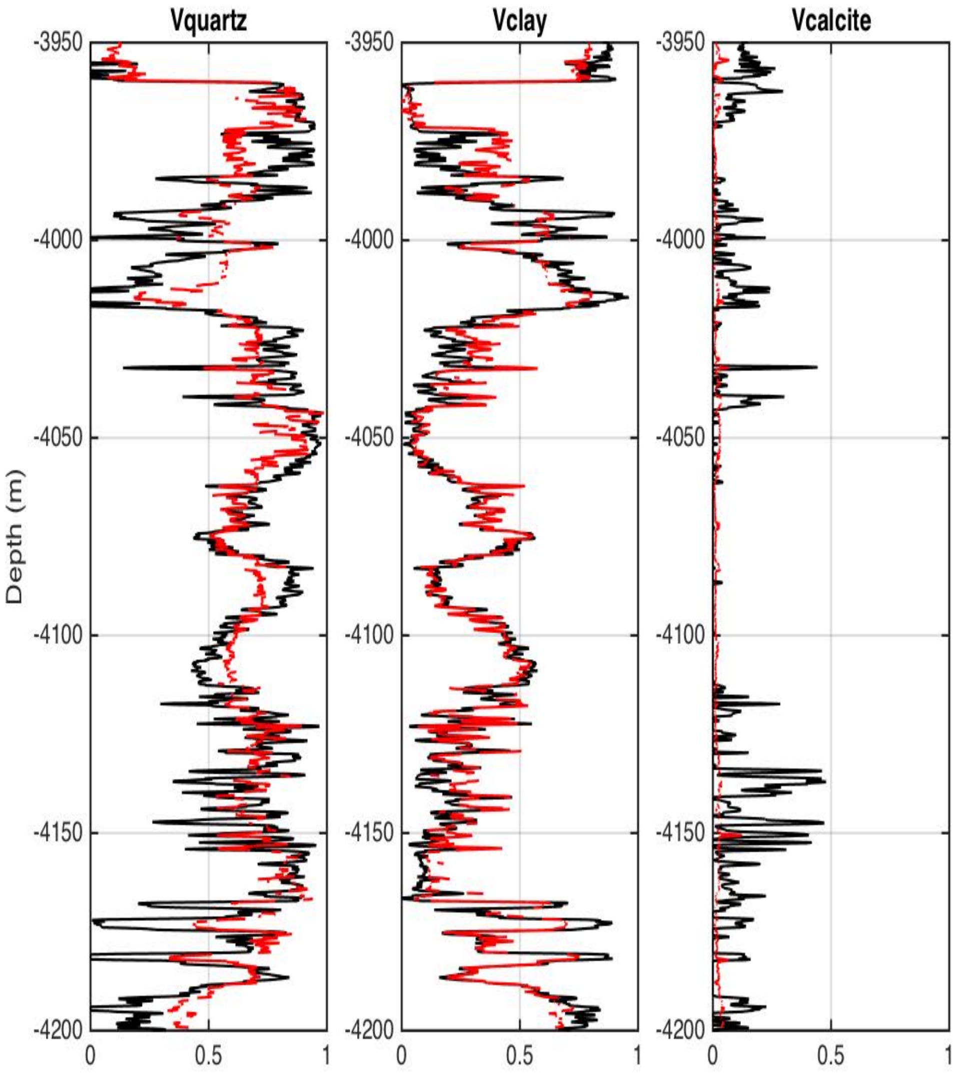
Clay
Machine
Human
Predicted Clay



Calcite
Machine
Human
Predicted Calcite

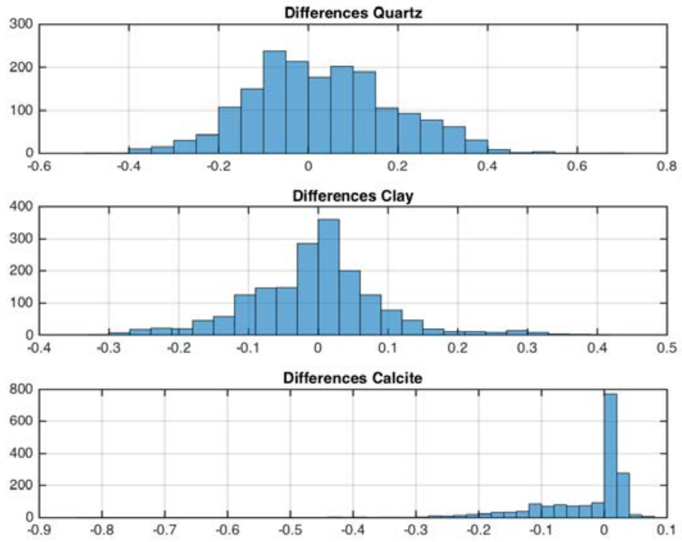
So far so good: what about predictions across wells ?

Example: 6507/3-5S

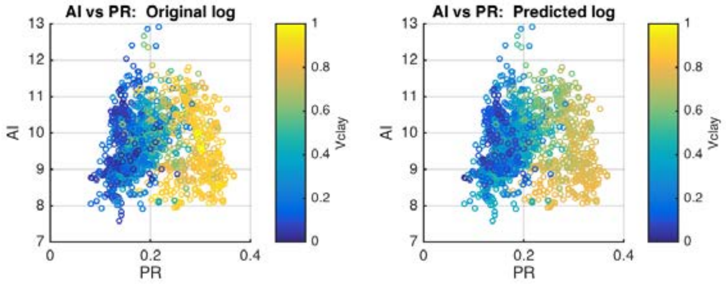


Original petrophysical workflow

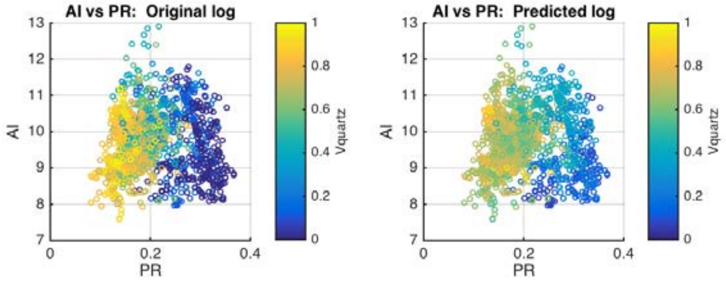
ML workflow



Clay

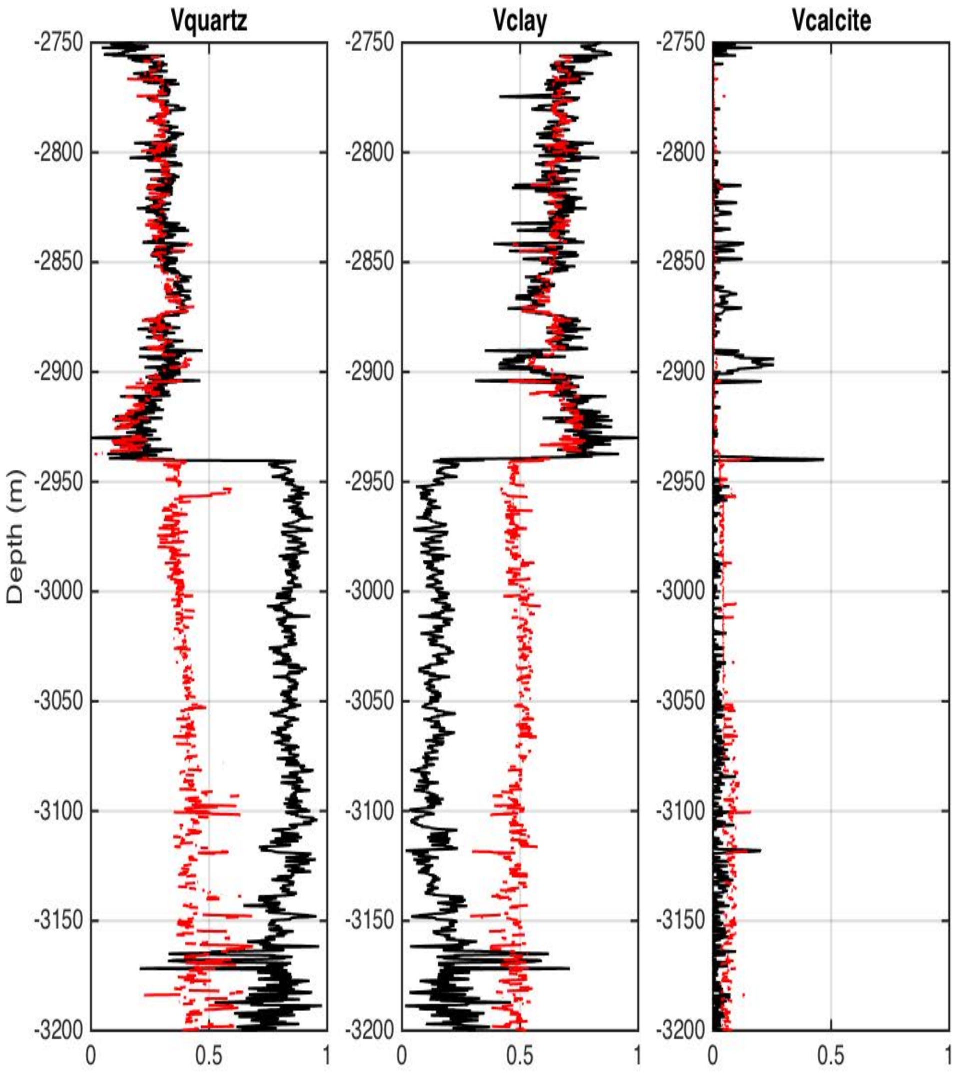


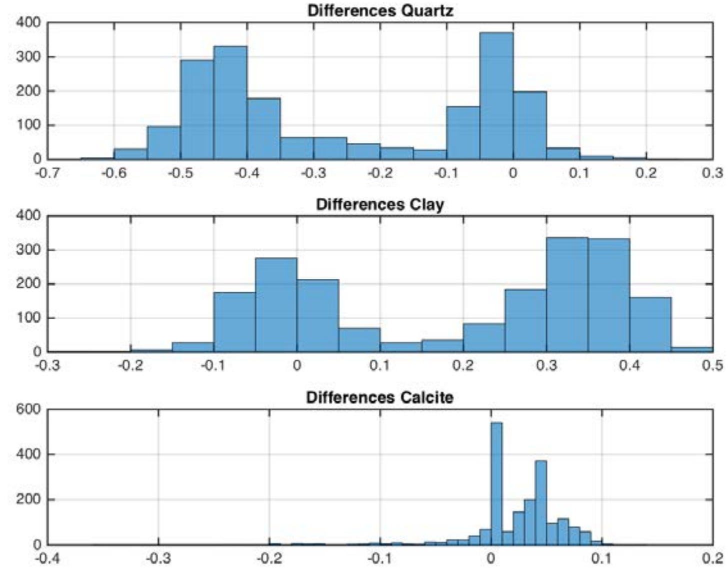
Quartz



So far so good: what about predictions across wells ?

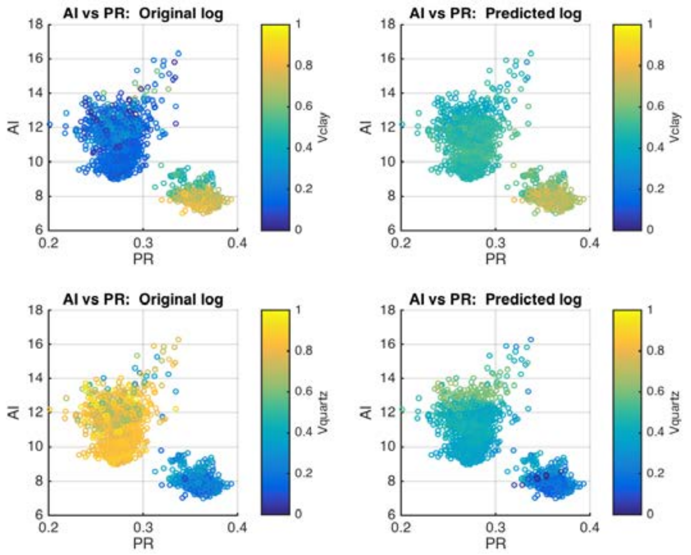
Example: 6306/5-2



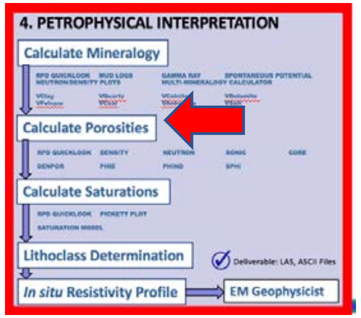


Clay

Quartz

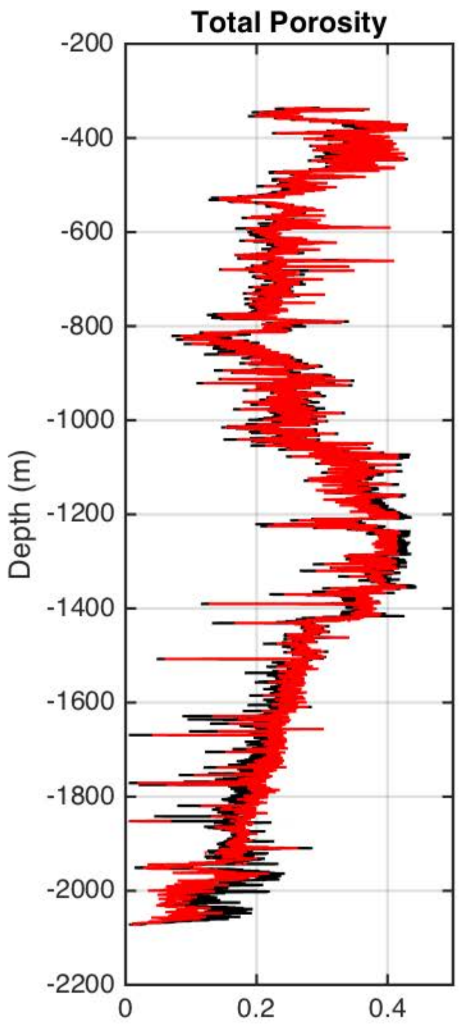


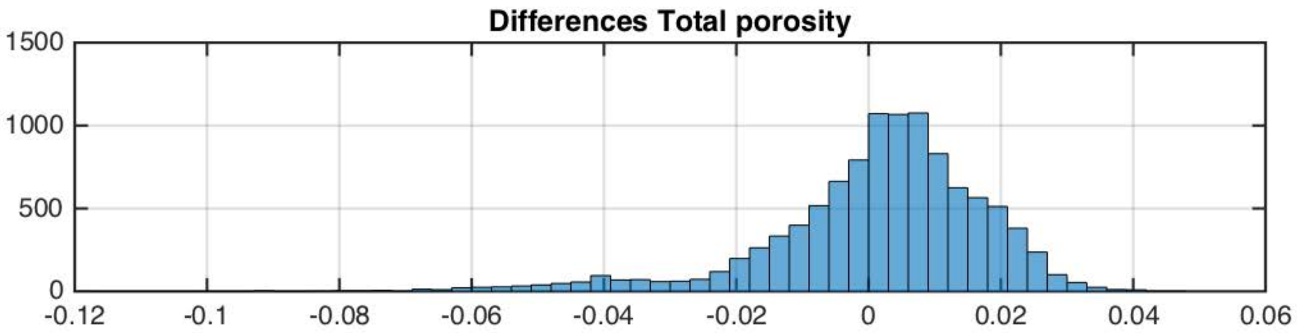
Predicting total porosity

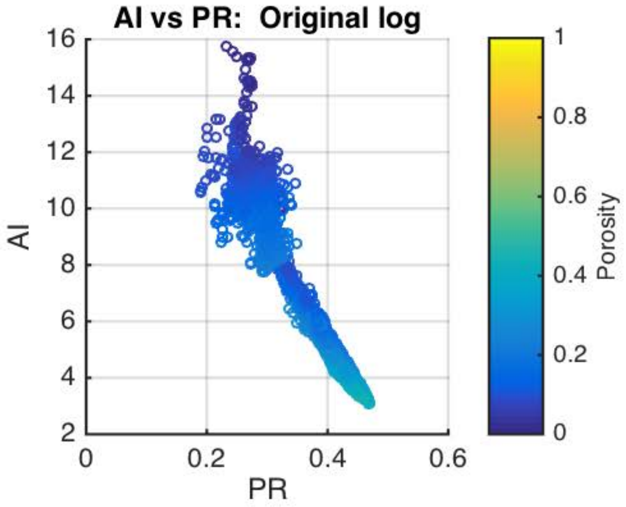


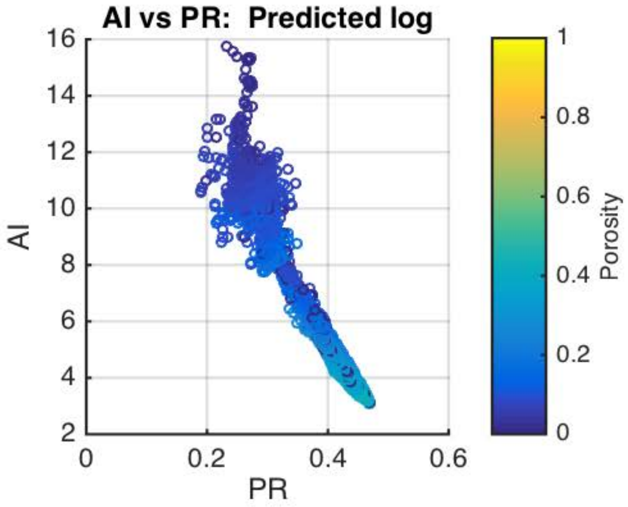
Example: 6306/6_2

Predictors: GR, Nphi, Resistivity, Rhob, Vp



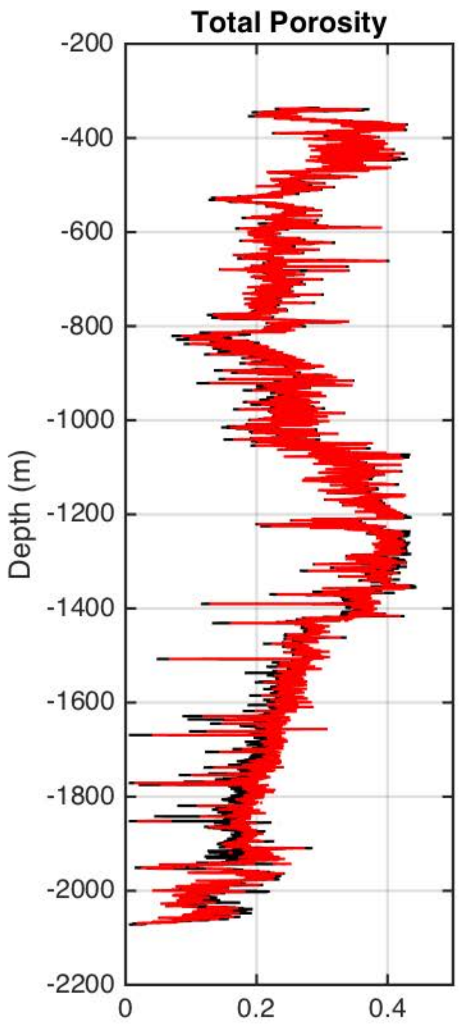






Original petrophysical workflow
ML workflow

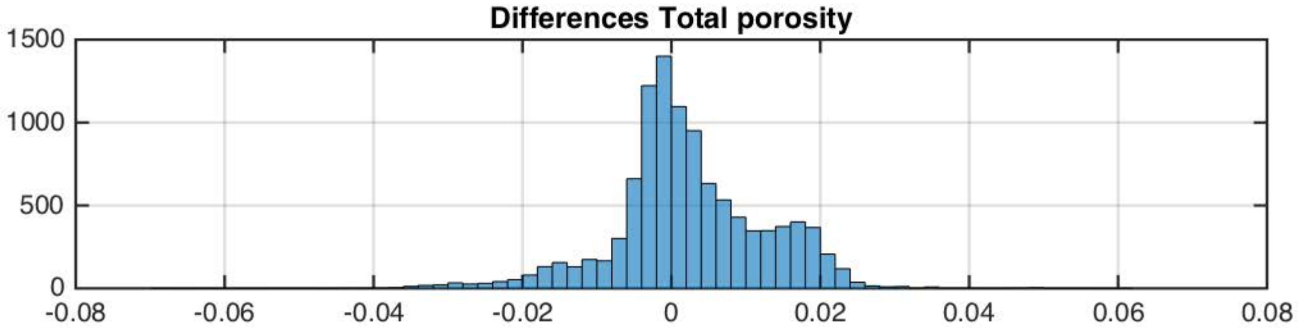
Predicting total porosity: add mineralogy

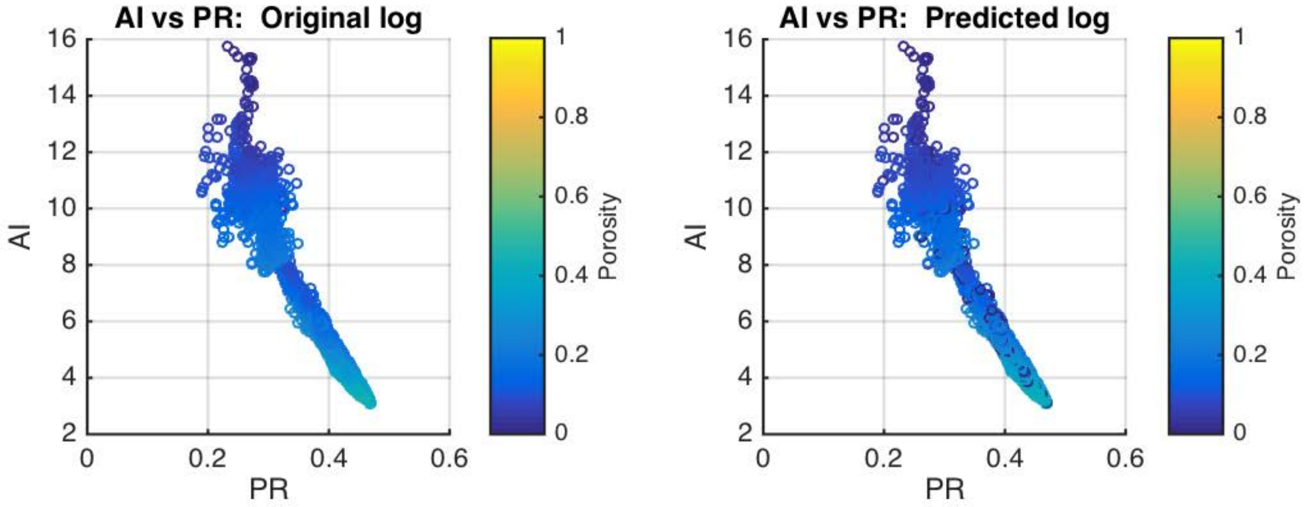


Original petrophysical workflow
ML workflow

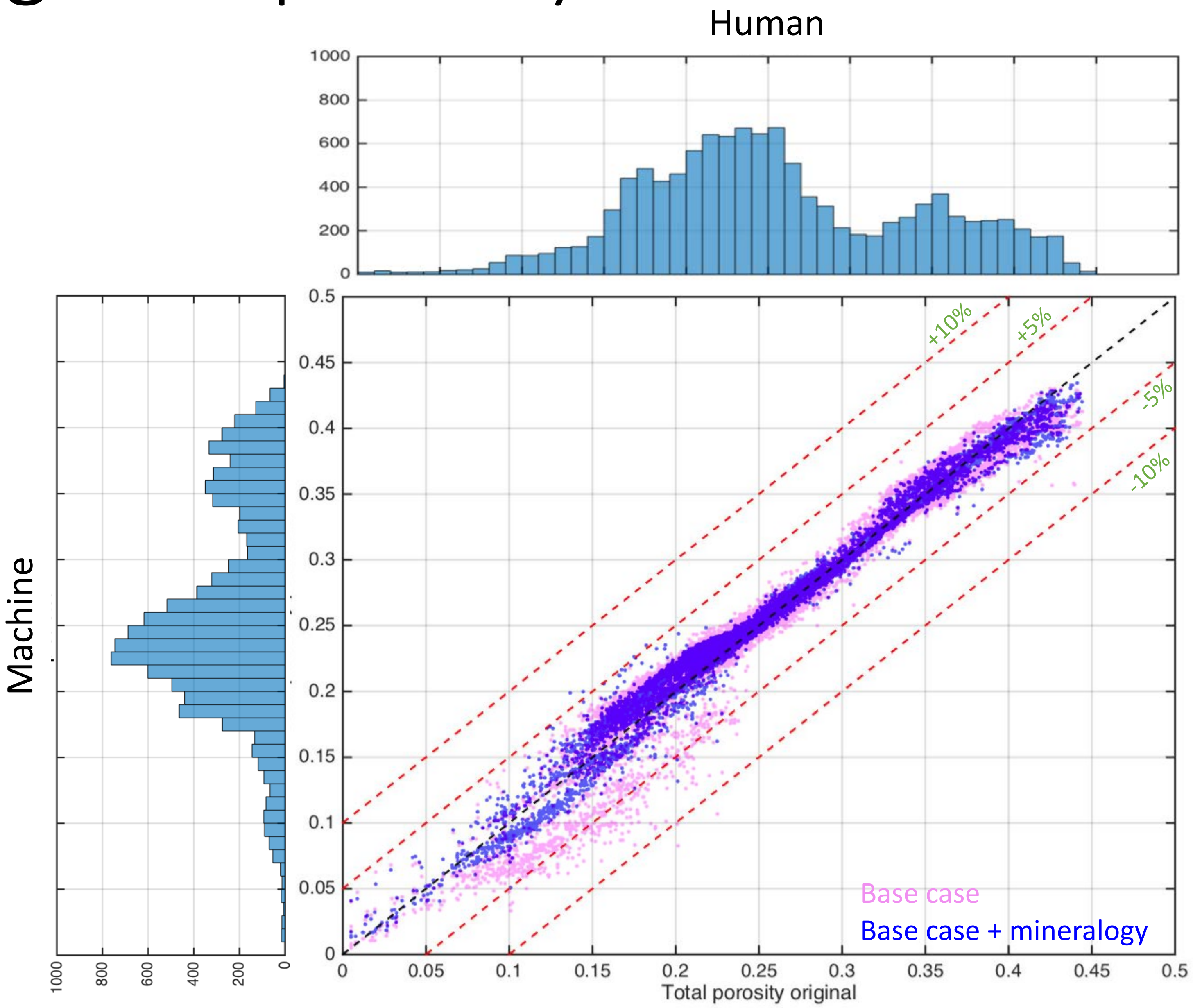
Example: 6306/6_2

Predictors: GR, Nphi, Resistivity, Rhob, Vp, full mineral suite

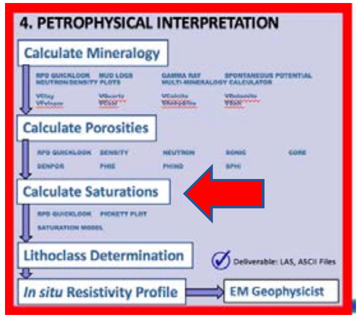




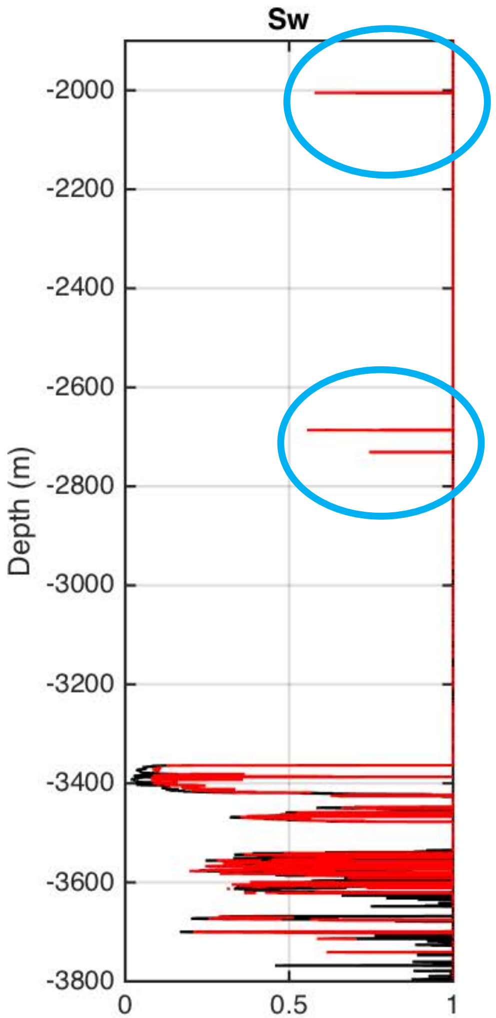
Predicting total porosity



Predicting saturation: Sw



Well 6507/3_3



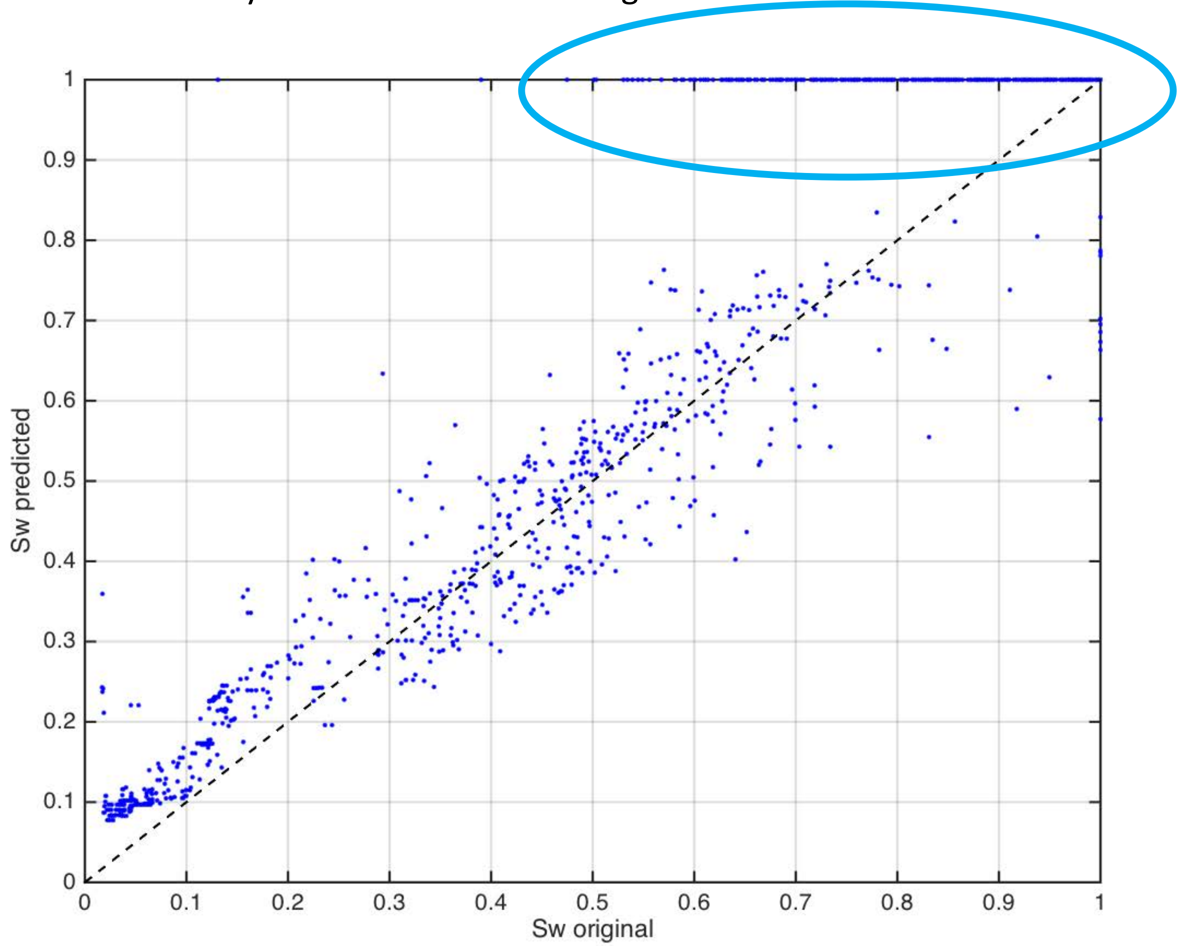
False positive (lithology)

Gas shows

Original petrophysical workflow

ML workflow

Training data bias: far more water than hydrocarbon in the training dataset

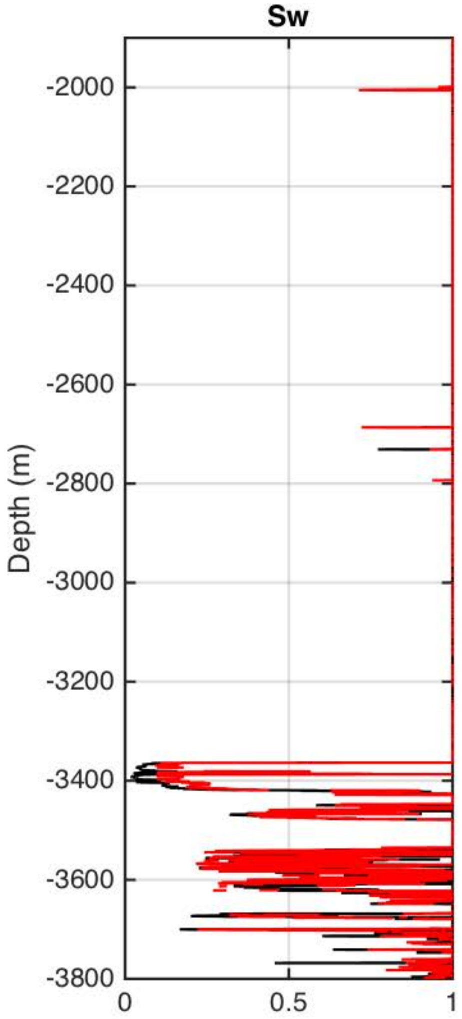


Predictors: GR, Nphi, Resistivity, Rhob, Vp, total porosity

Balance the training data

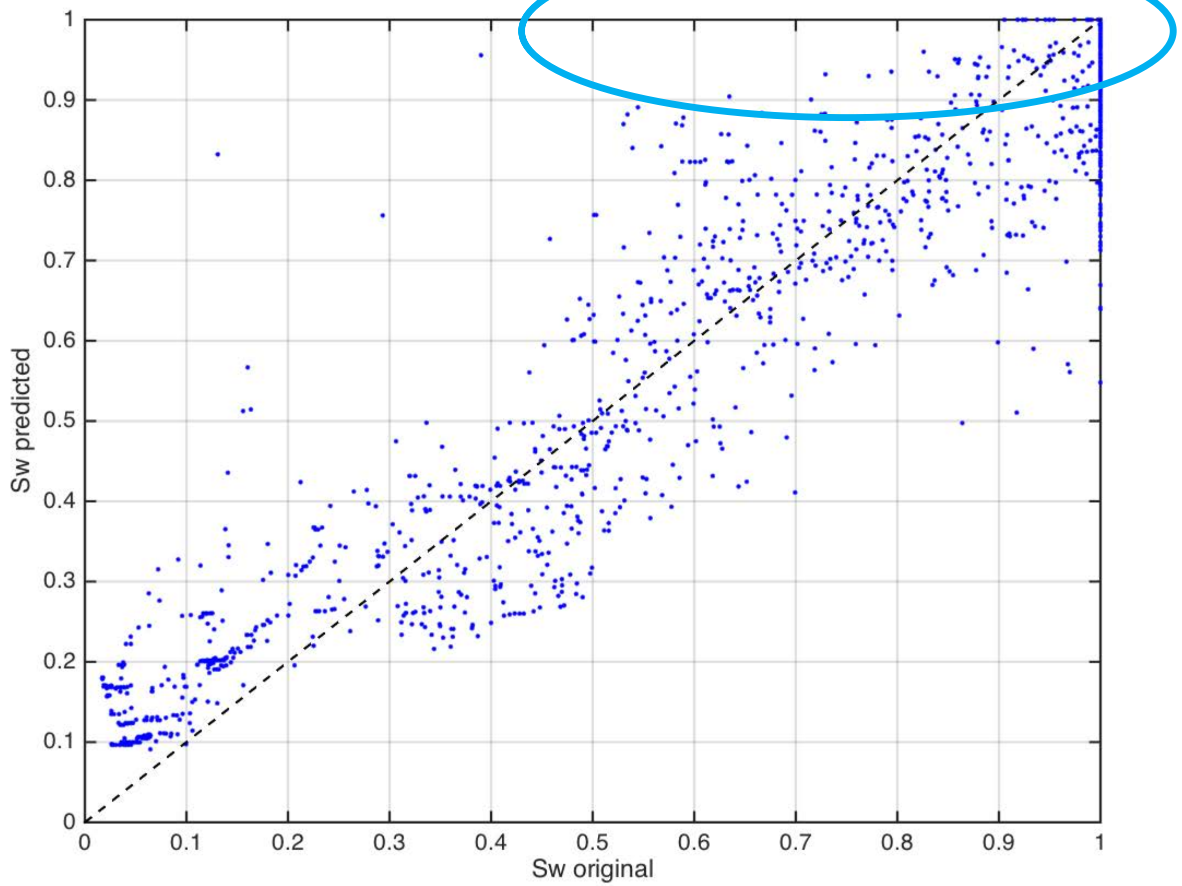
Well 6506/3_3

Predictors: GR, Nphi, Resistivity, Rhob, Vp, total porosity



Original petrophysical workflow
ML workflow

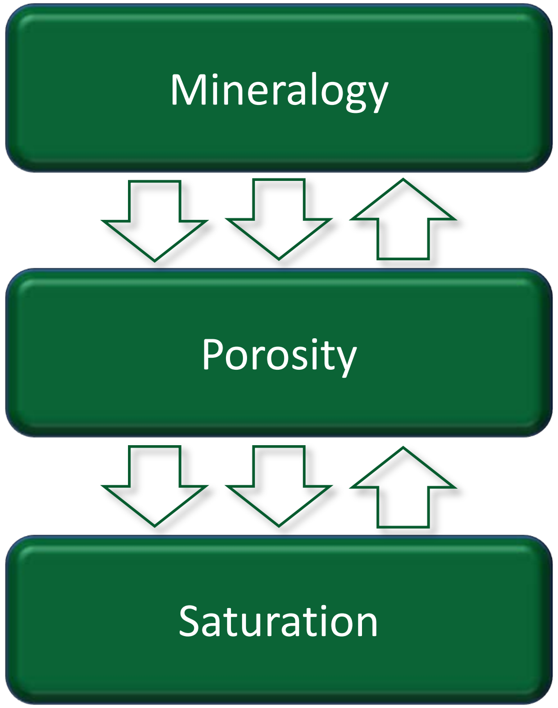
Improvement with a more balanced training dataset



Observations

- Can we use machine learning to streamline workflows ?
  - Yes – results are promising
  - …but we need to understand the effect of biased training data and assess how general a model can be built.
  - Needs to be iterative

- How good a model do we need ?
  - It doesn't need to be perfect to be useful. None of our models are perfect !
  - We're comparing against a human interpretation – who is right ?
  - How good is good enough and how do we characterize that ?



Overview

Past

- Machine learning in geoscience
- Neural networks

Present

- Prediction of properties from logs

Future

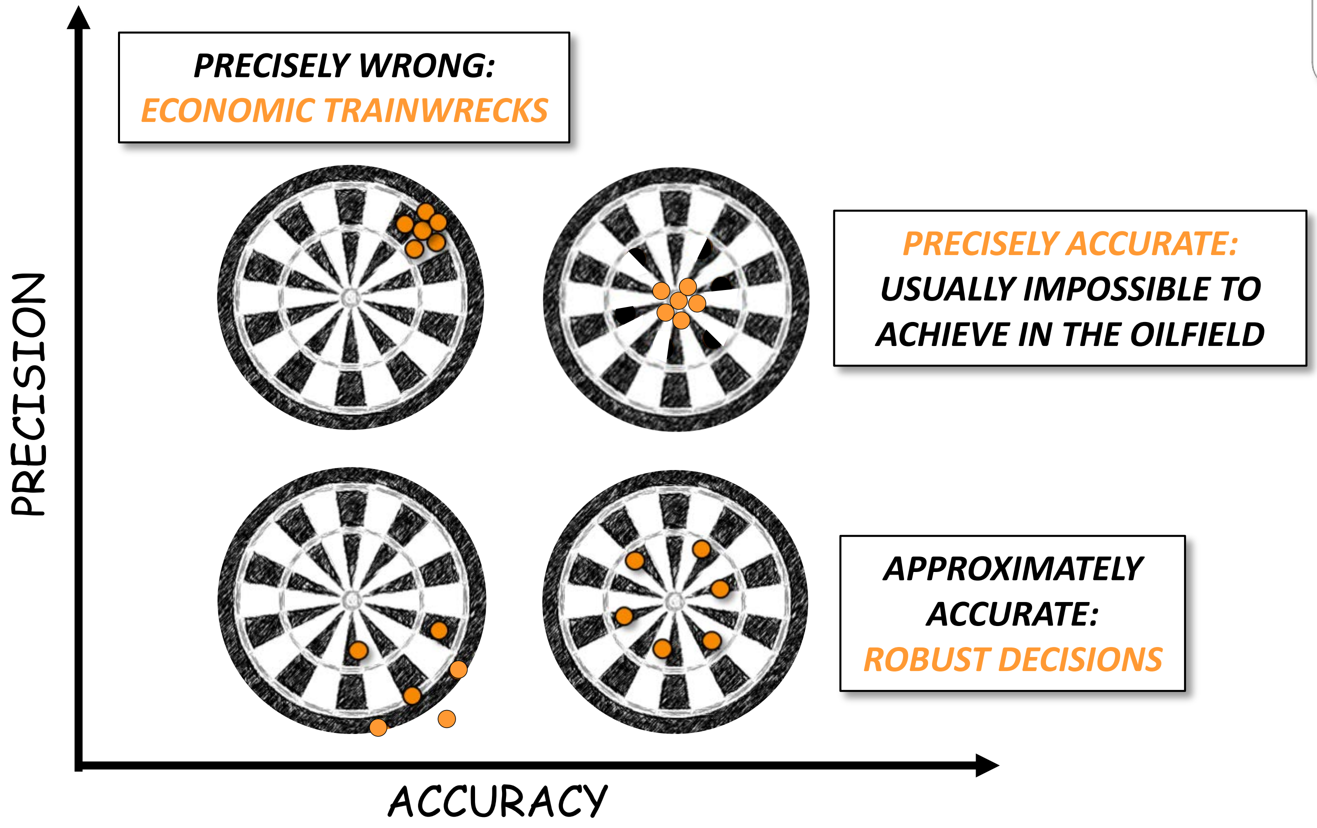
- Uncertainty management in geomodelling

Geo-modelling in the oil and gas industry

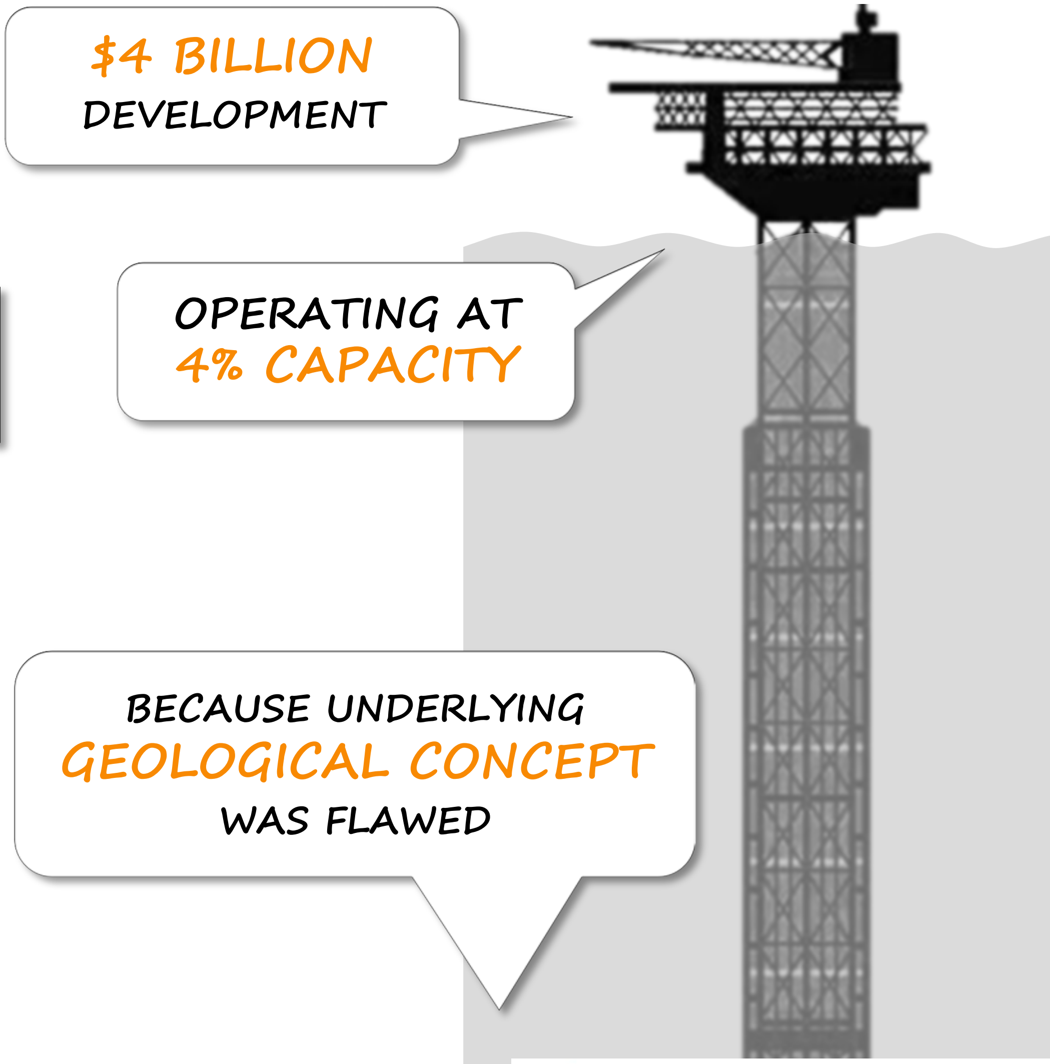
"all things equal, the best geologist is the one who's seen the most rocks"

Geology is a science of case based reasoning, but ***not*** one where we can determine a ***precise*** answer

Accuracy versus precision

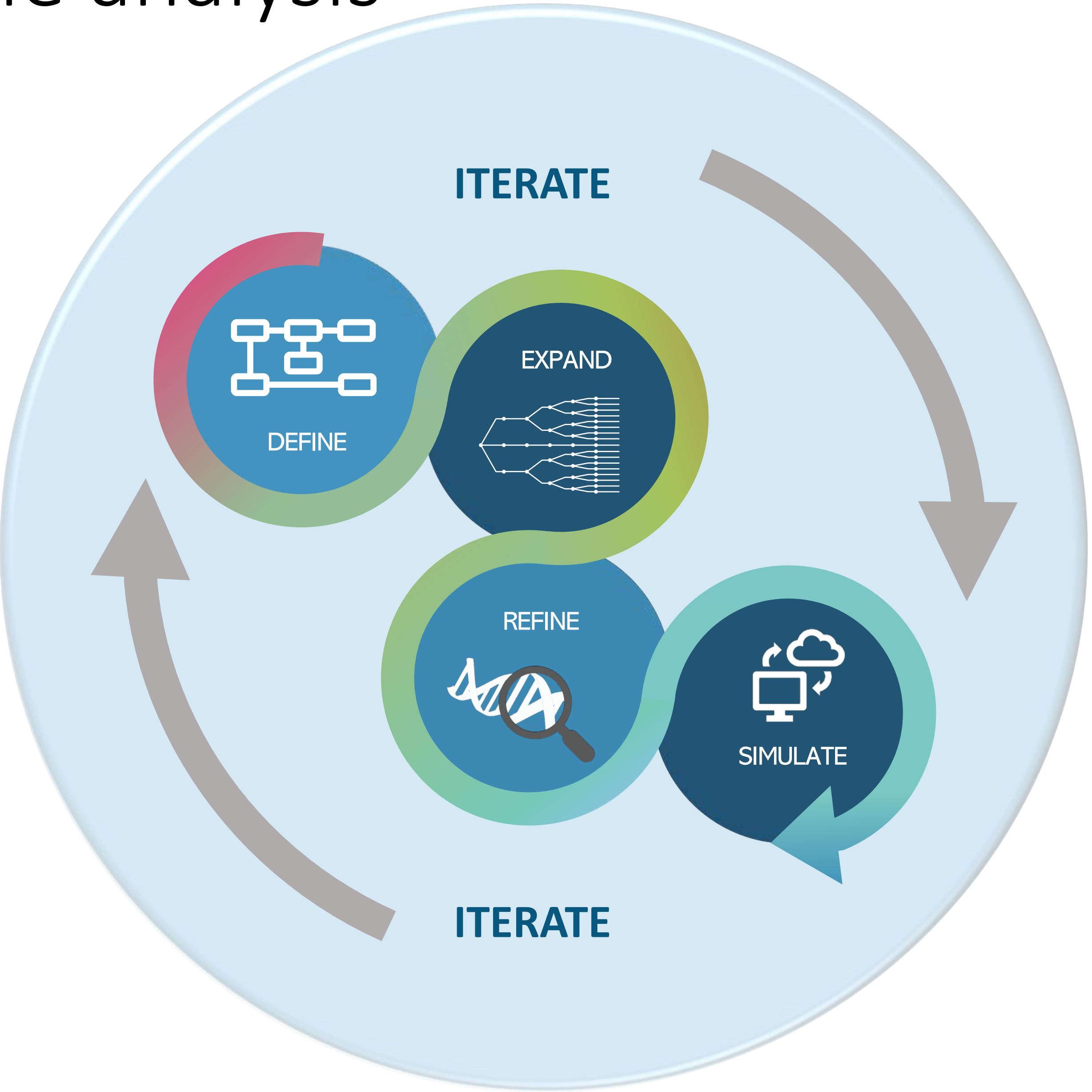


FAILURES ARE EXPENSIVE



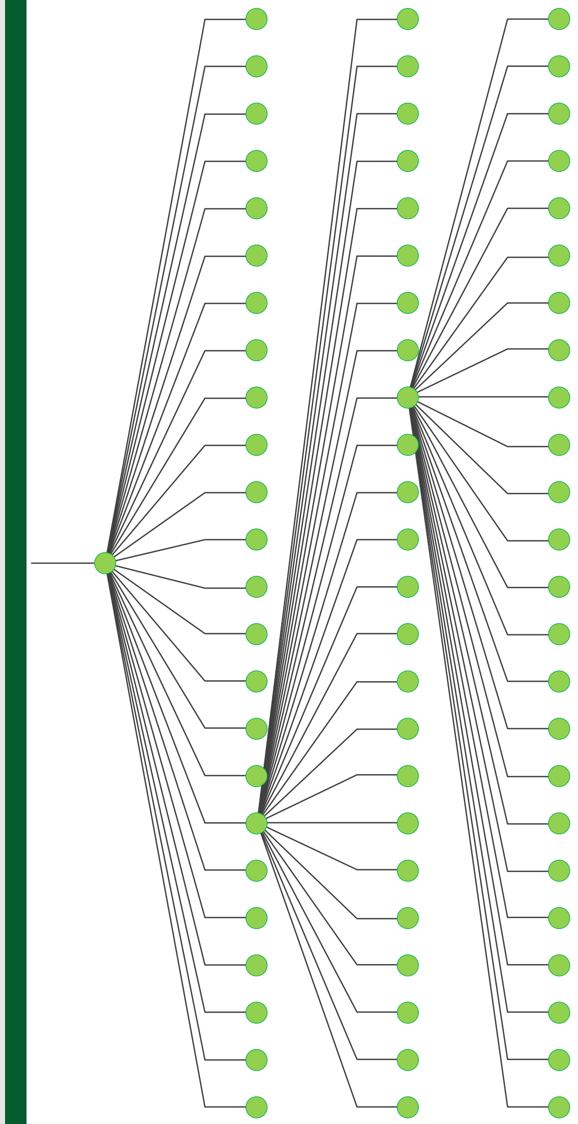
Cognitive Geology

Model ensemble analysis



Cognitive Geology

Creating the ensemble

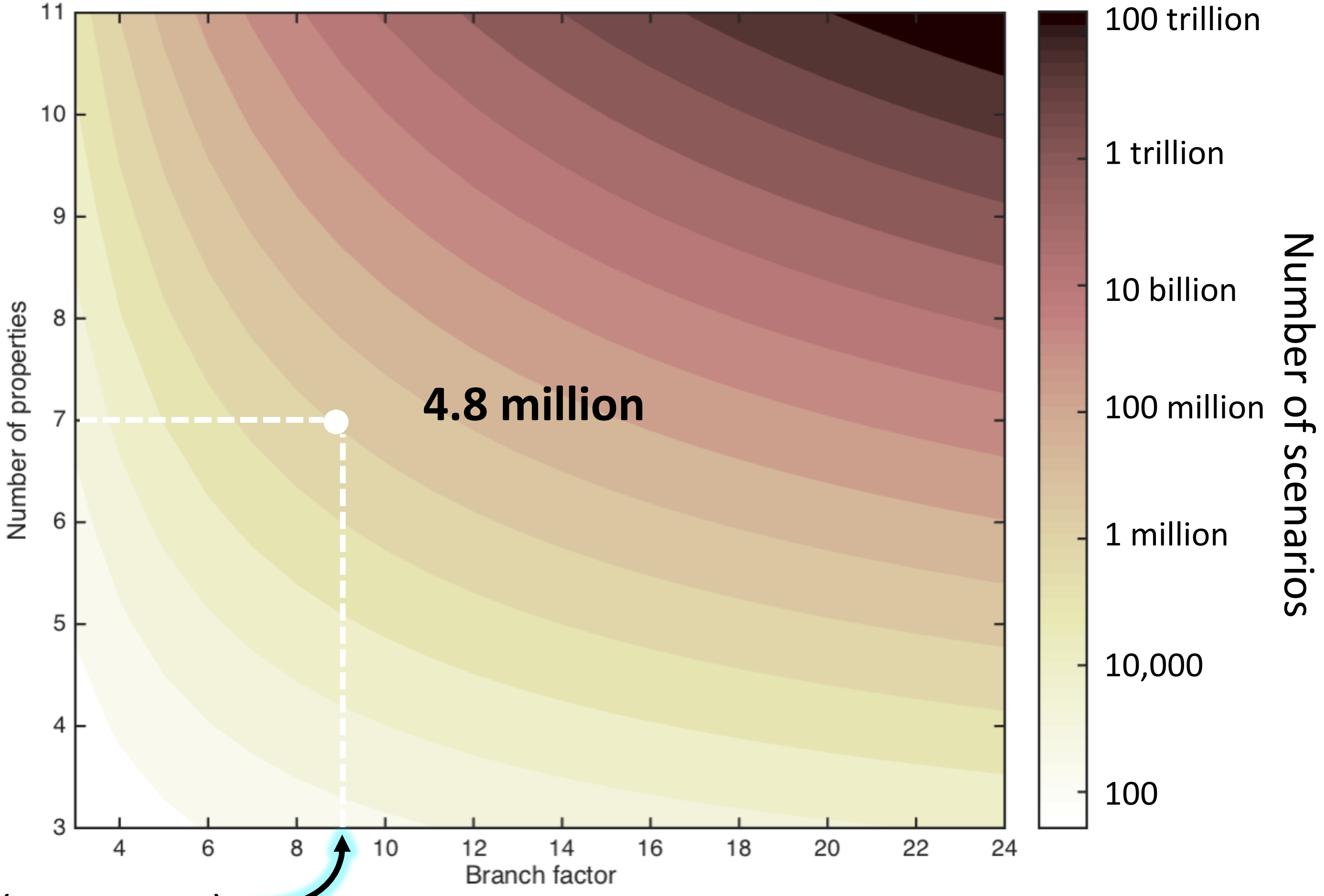


Number of scenarios = i^n

i=branch factor

n=number of properties

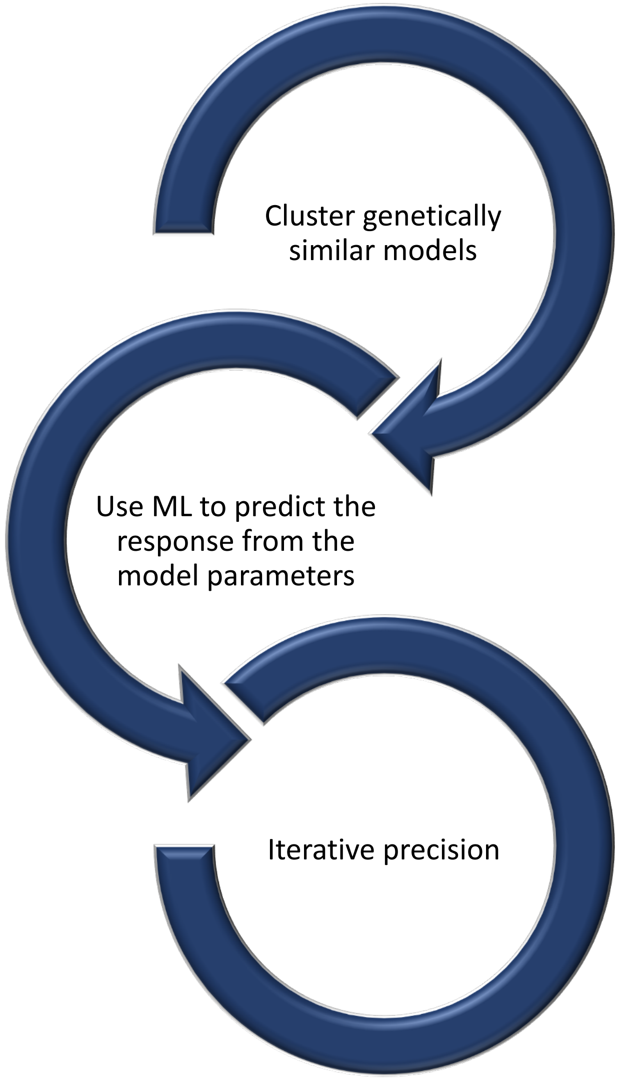
Grid
Facies
NTG
Porosity
XY Perm
Kv/Kh
Swirr



Percentiles (P10 to P90)

Cognitive Geology

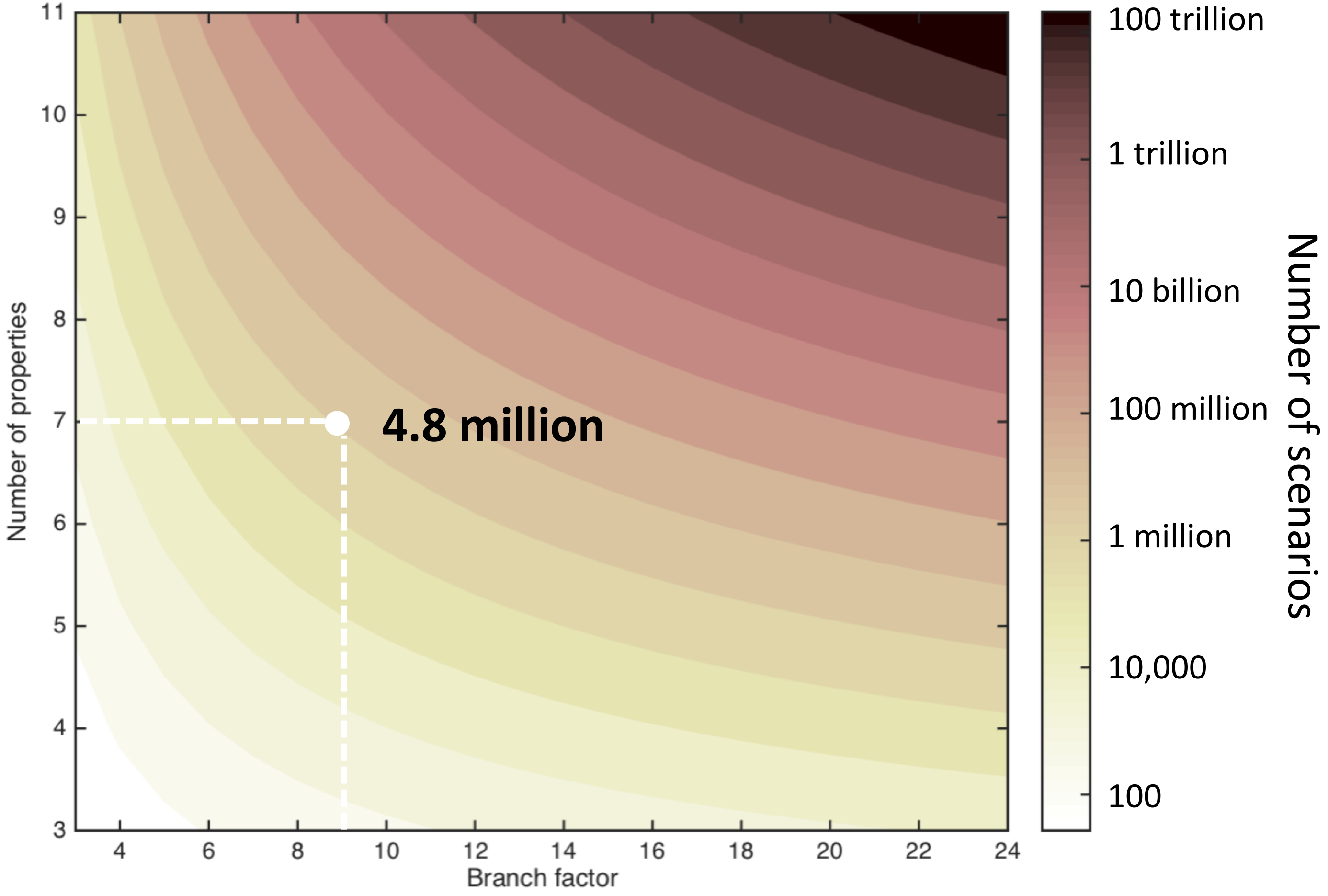
Managing the ensemble !



Number of scenarios = i^n

i=branch factor

n=number of properties



Cognitive Geology

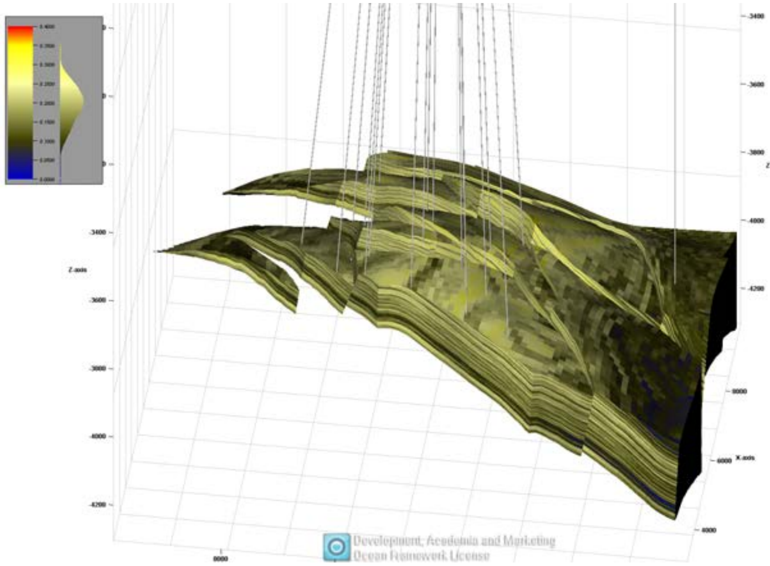
A really (really) simple example

Suppose we want to calculate the amount of oil in a reservoir:

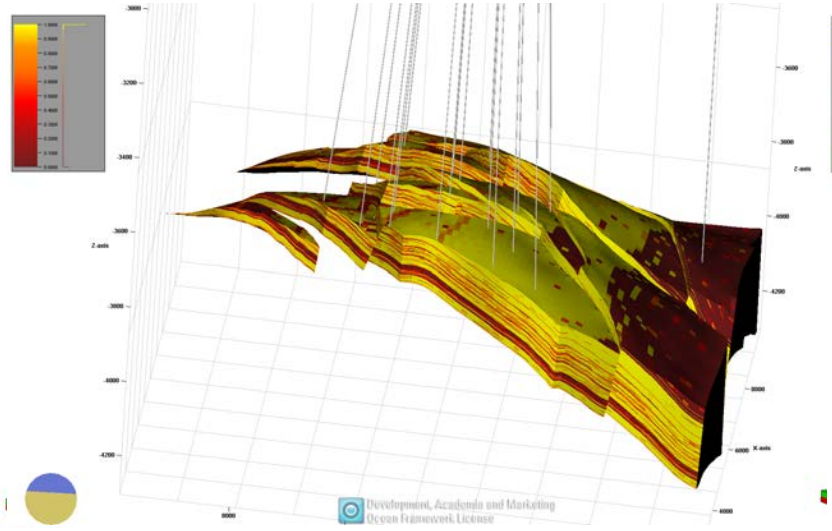
Oil in place = Net to Gross x Porosity x Oil saturation

Number of scenarios for each property: 24

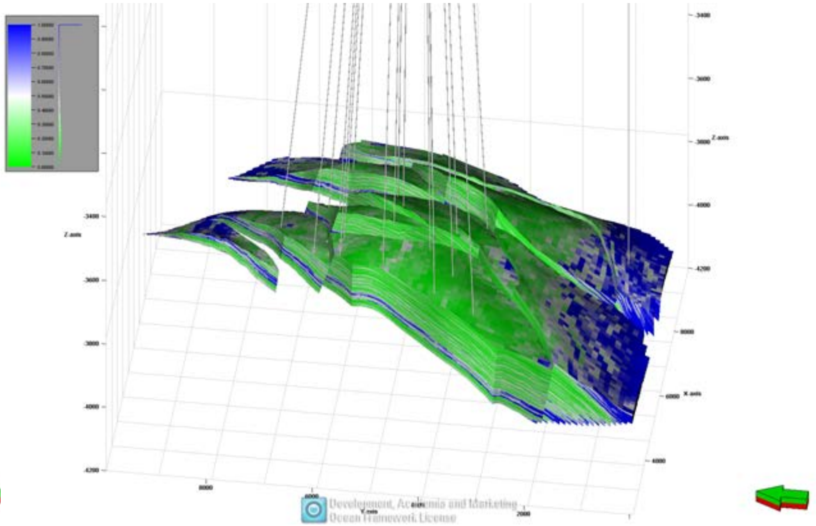
Total number of oil in place values = 24^3 = 13824



Porosity



NTG



Saturation

Cognitive Geology

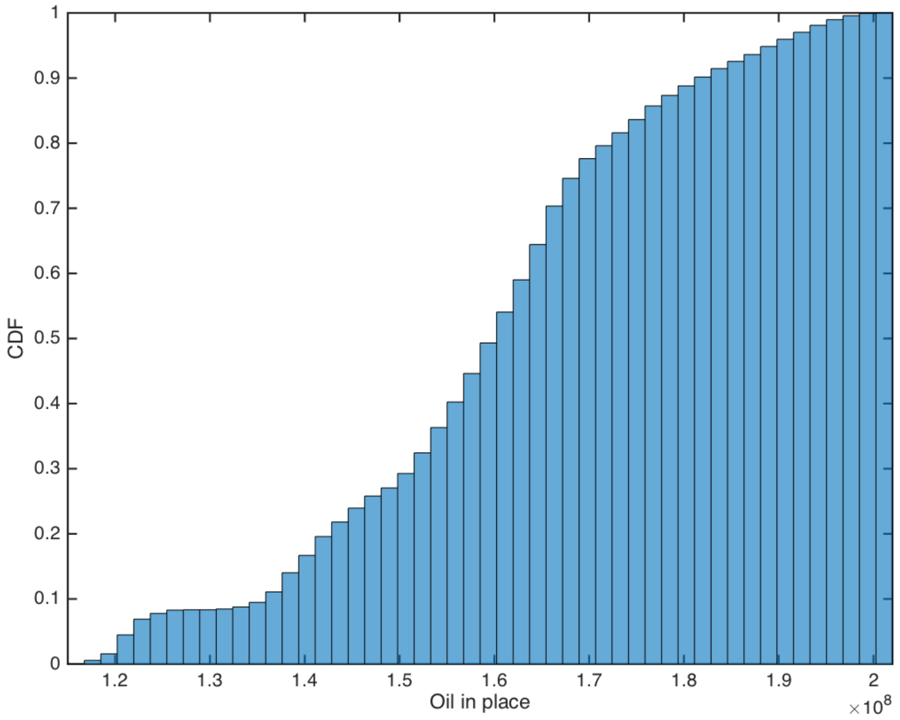
A really (really) simple example

Suppose we want to calculate the amount of oil in a reservoir:

Oil in place = Net to Gross x Porosity x Oil saturation

Number of scenarios for each property: 24

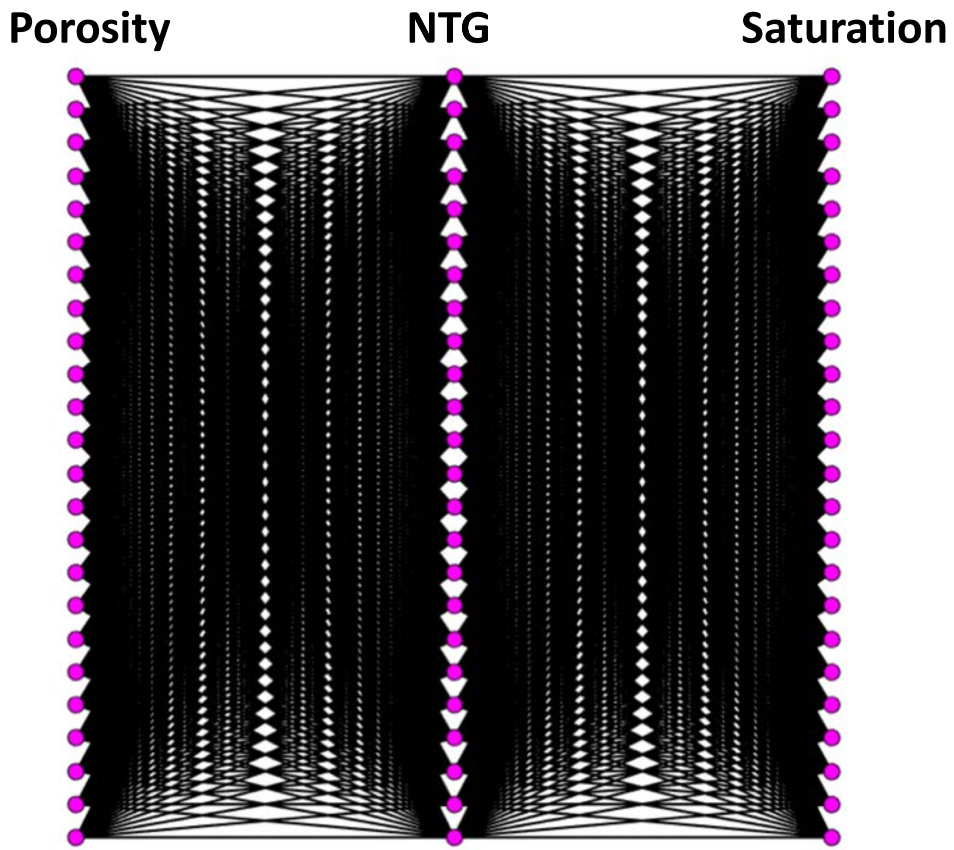
Total number of oil in place values = 24^3 = 13824

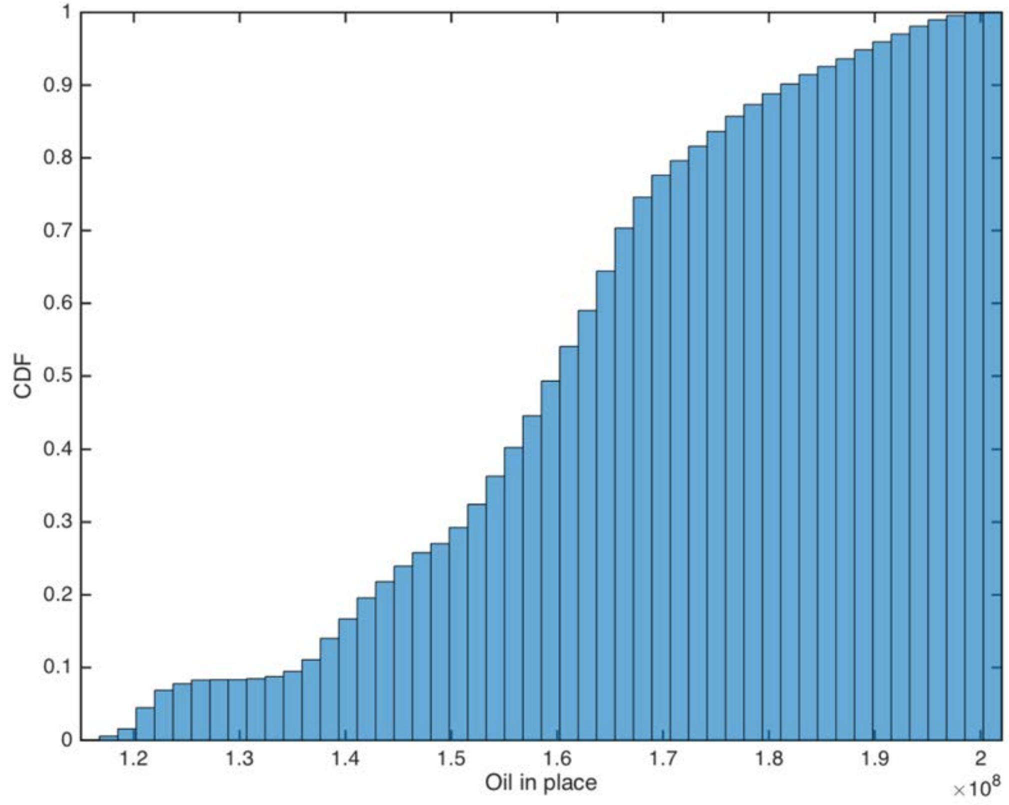


Cognitive Geology

How many models do we need ?

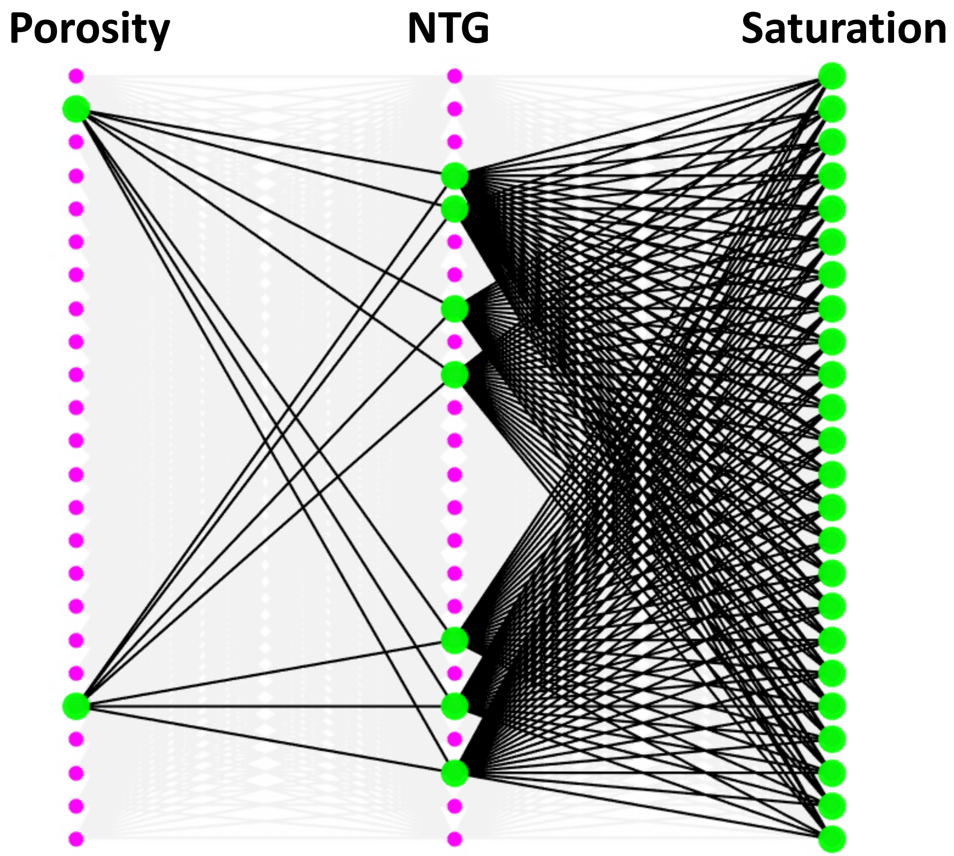
13824 models

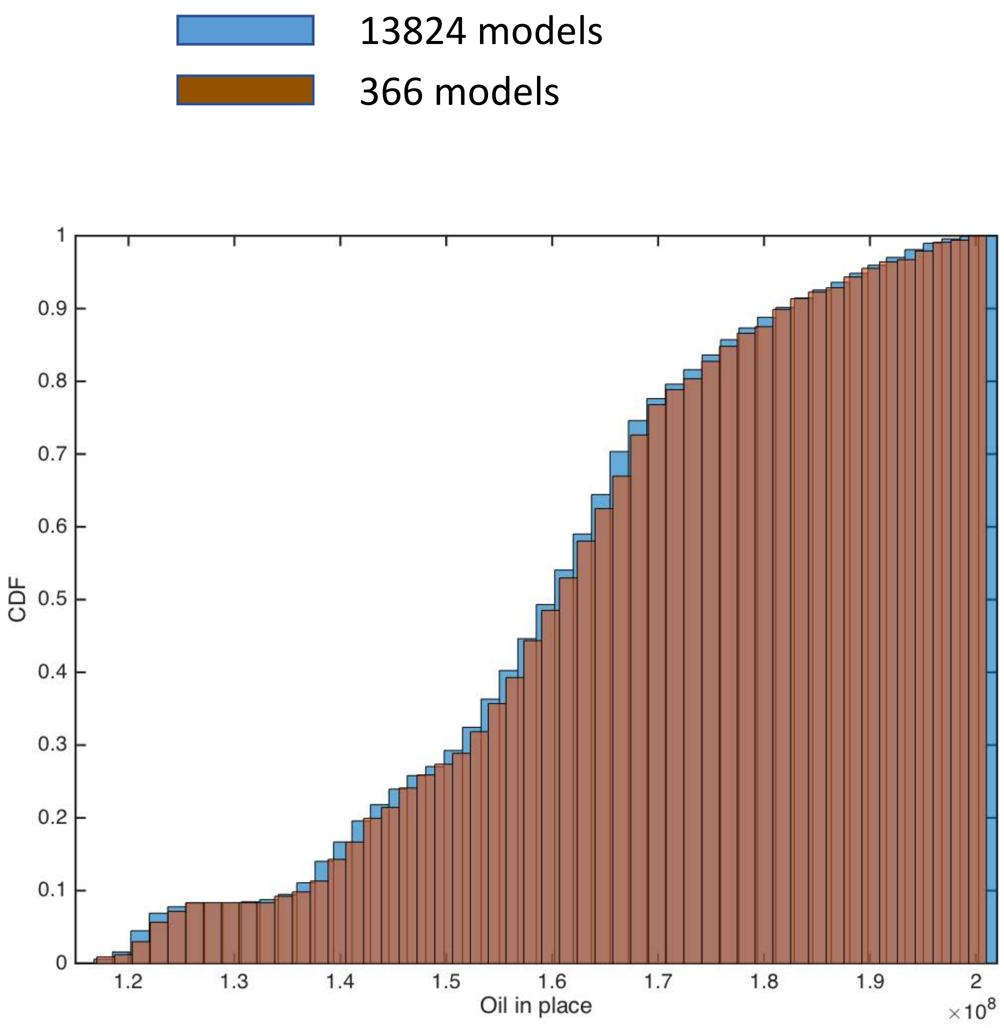




Cognitive Geology

How many models do we need ?





Really simple example…..but shows that the approach is promising.

Cognitive Geology

Questions and challenges

- How much training data is required?   WHAT training data is needed ?
- How general will a trained machine learning approach be? For example, whereas you can imagine that results may be good within a single geological basin, will they be equally good on the other side of the planet?
- How accurate will a machine generated interpretation be? It does not need to be perfect to be useful: something that is right 70-80% of the time could still represent a huge time saving in the workflow.
- Can we characterize/understand the uncertainty in the results?