EP/Y006143/1, EP/Y005732/1

Upscaling the porosity–permeability relationship of a microporous carbonate for Darcy-scale flow with machine learning

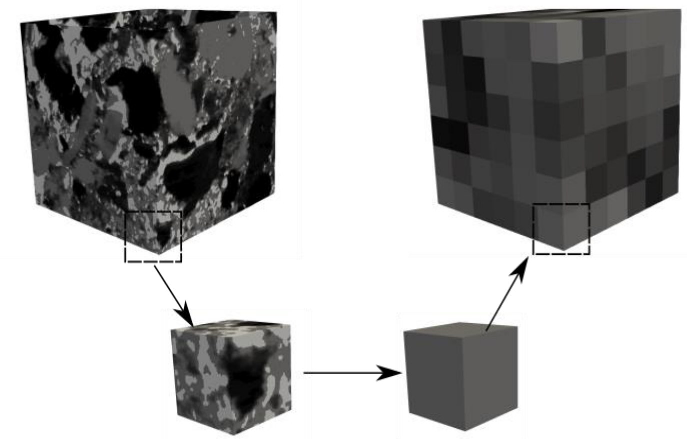
Dr Hannah Menke

Institute of GeoEnergy Engineering

Heriot-Watt University

Edinburgh, Scotland, UK

Co-authors: Dr Julien Maes, Dr Kamaljit Singh, Prof Ahmed ElSheikh



The Problem Statement: Why bother upscaling?

Experimental Imaging
- Porosity-permeability relationship

Numerical Modelling
- DNS modelling
  - Single phase Darcy-Stokes-Brinkman

Machine Learning Tools
- Multivariate Structural Regression

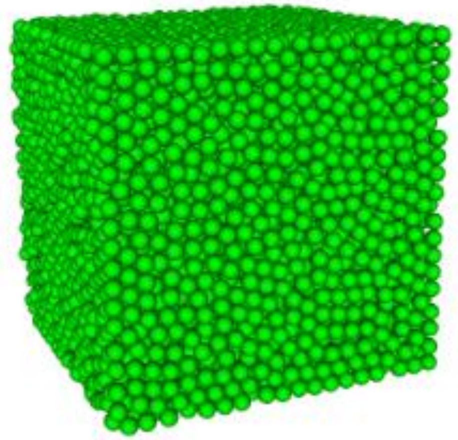
Upscaled descriptions of structure and flow

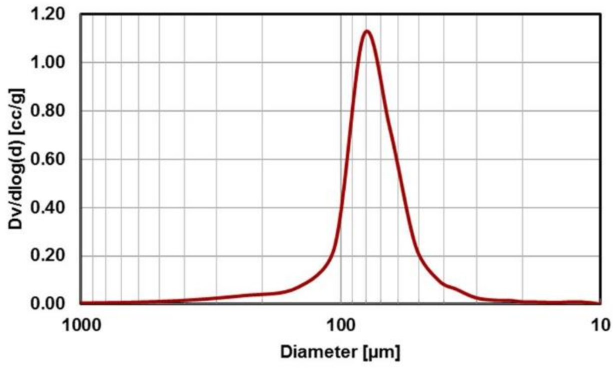
1. Full flow model (computationally expensive)
2. Simplifying either structure (network extraction) or flow properties (Kozeny-Carman)
3. Machine Learning upscaling using image features

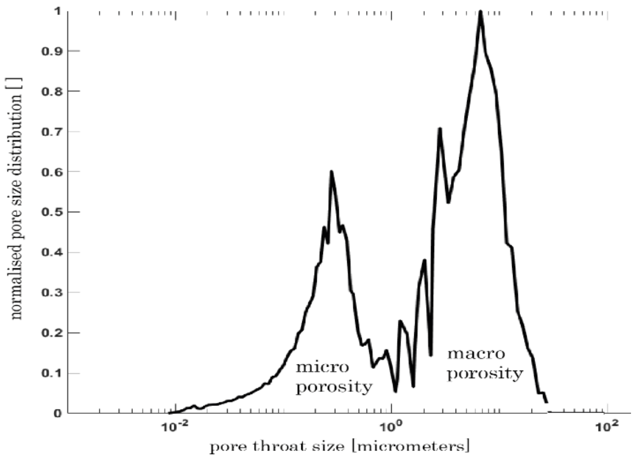
How do we upscale structure in carbonates?

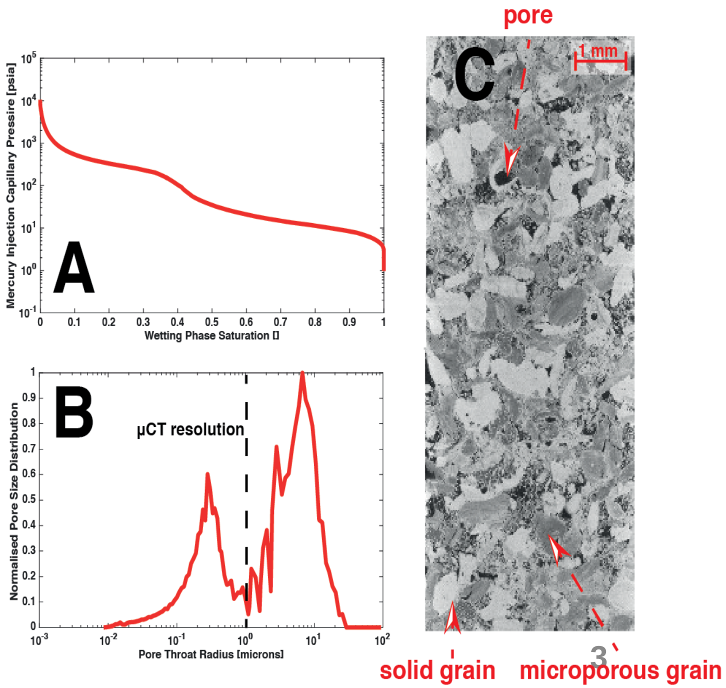
- **The porosity-permeability relationship is then defined typically using the Kozeny-Carman equation.**
  - **Based on the assumption that the pore space is effectively represented by an even packing of equally-size elliptical beads.**
  - **Does not incorporate any geological processes that would change the shape and connectivity of the pore structure**
  - **A very poor estimate for microporous rocks which have widely varying grain sizes and multimodal porosity structures**

$$K = \frac{\varphi^3}{c(1-\varphi)^2 S^2}$$









Why Decision Trees?

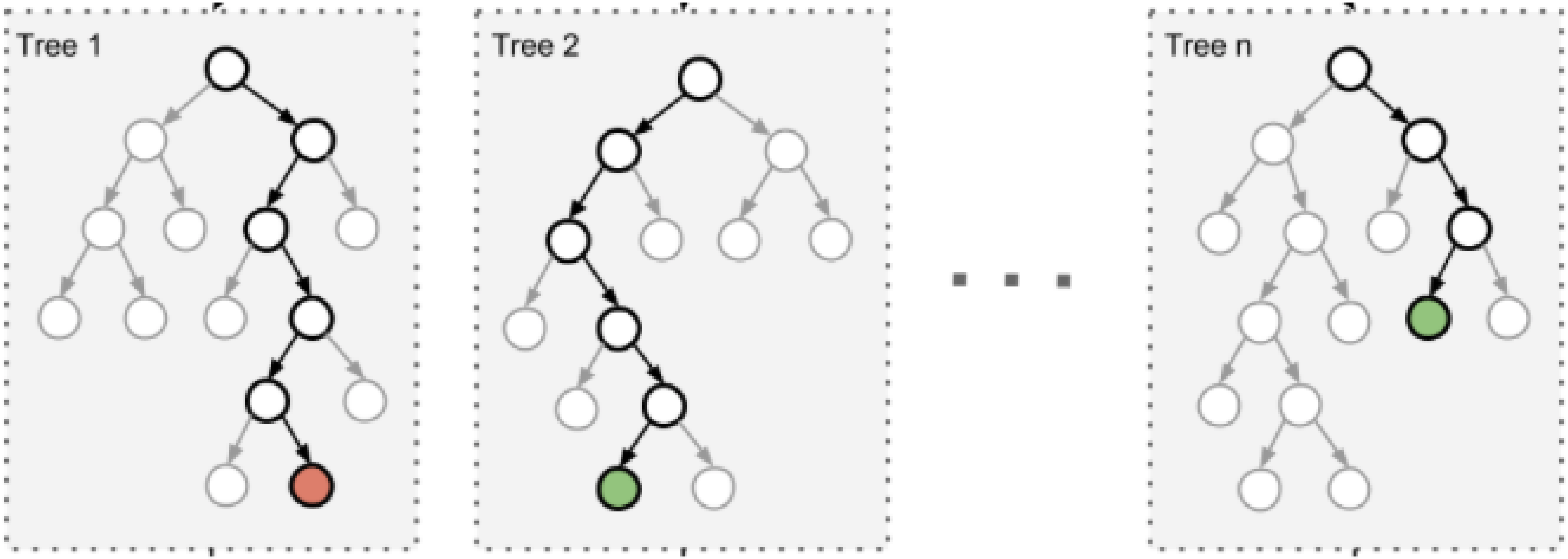
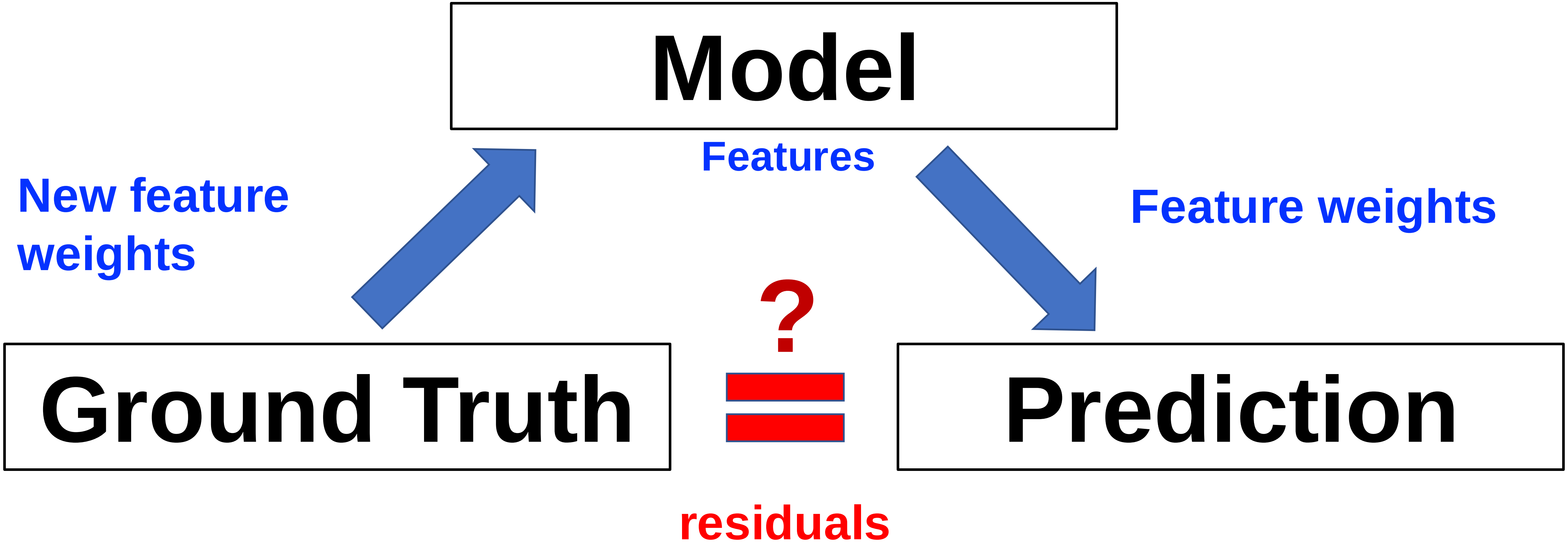


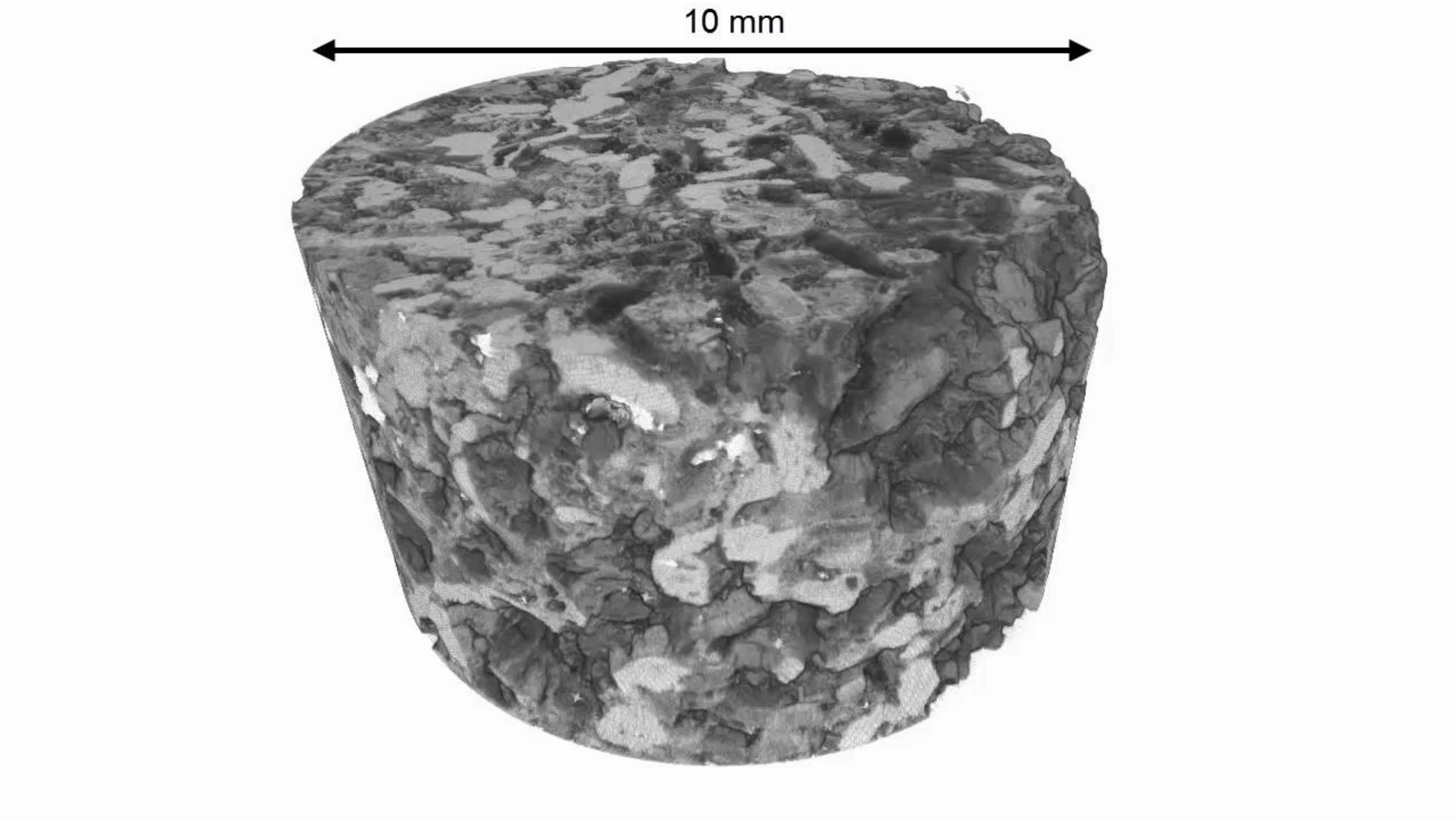
Figure 1 Decision Tree Schematic of an Extremely Randomized Forest.
Figure Credit: https://blog.statsbot.co/ensemble-learning-d1dcd548e936

- **Simple to use**
- **Low computational expense**
- **Not a black box – outputs tractable feature weightings**
- **Works well for single scale structures (e.g. Andrew 2019, SCA)**

The Rules of Machine Learning:



The Ground Truth: Multi-scale structural analysis



The Ground Truth: Multi-scale structural analysis

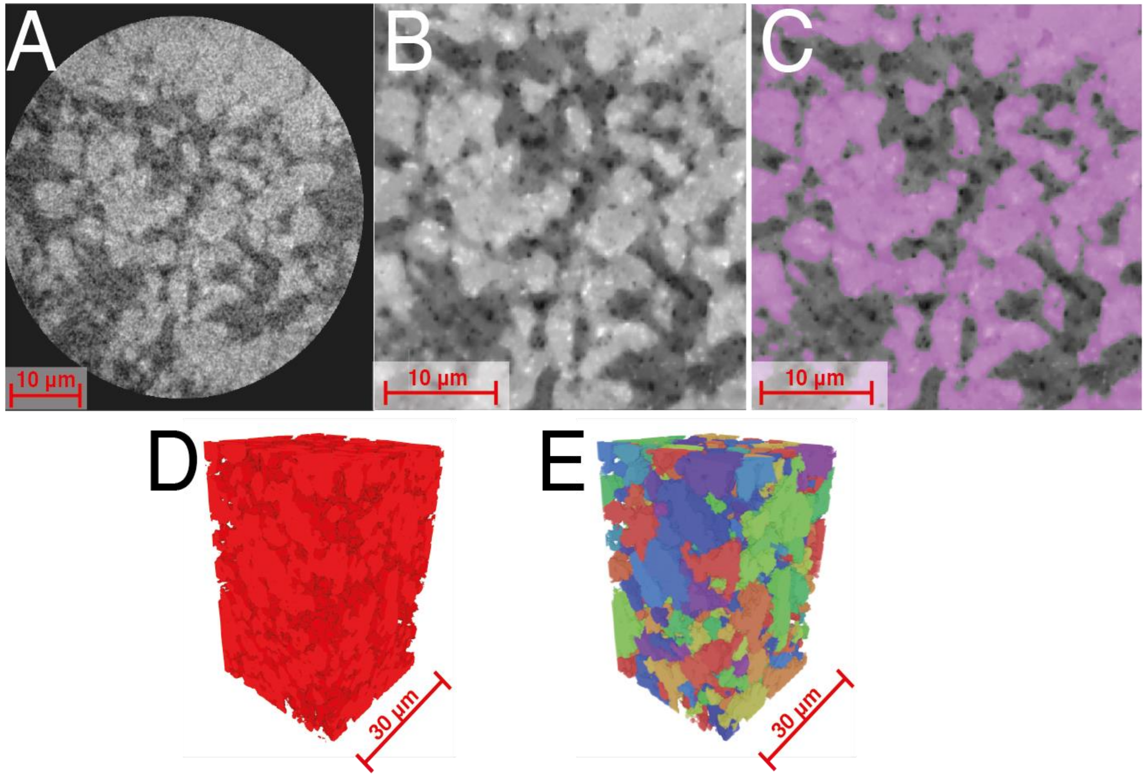


Figure 1: The raw NanoCT image (A), is filtered (B), segmented (C), rendered (D), and the grains separated (E) for grain size distribution analysis

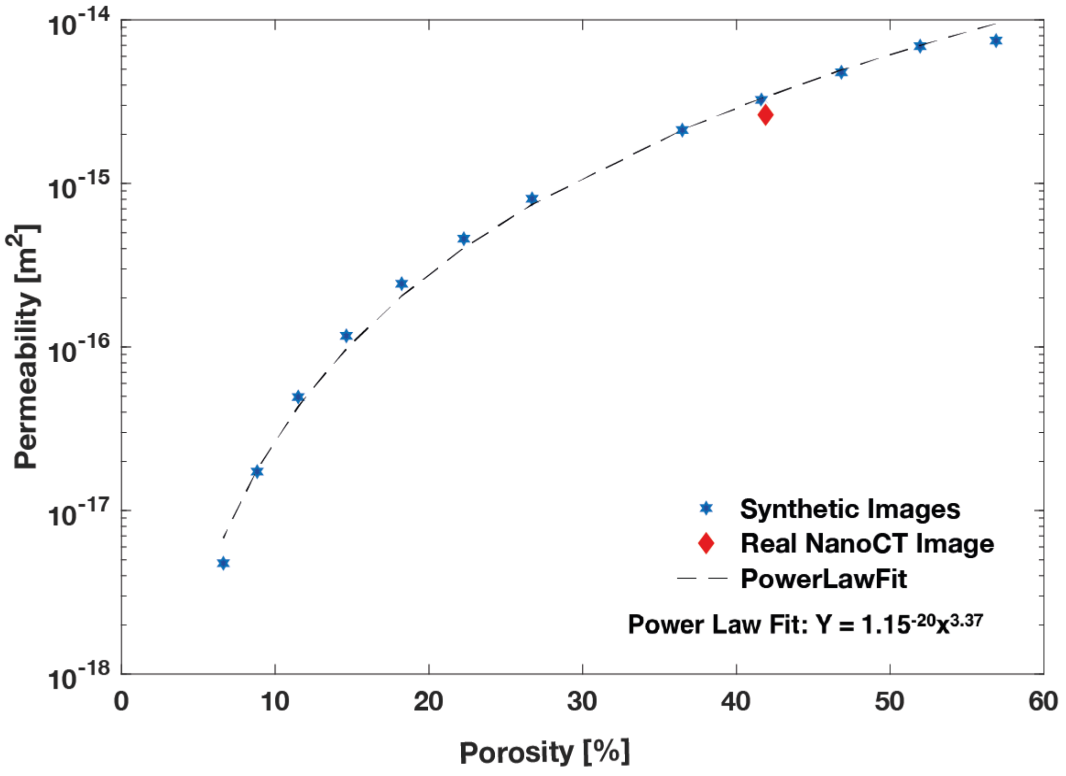
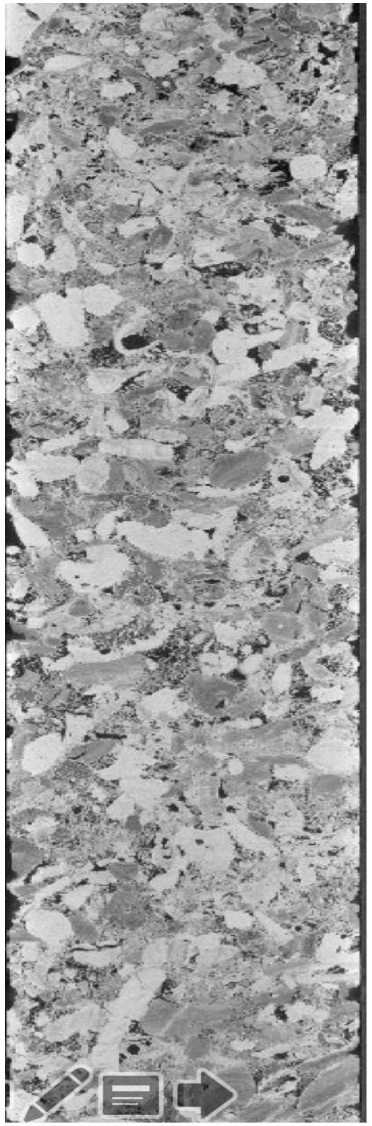
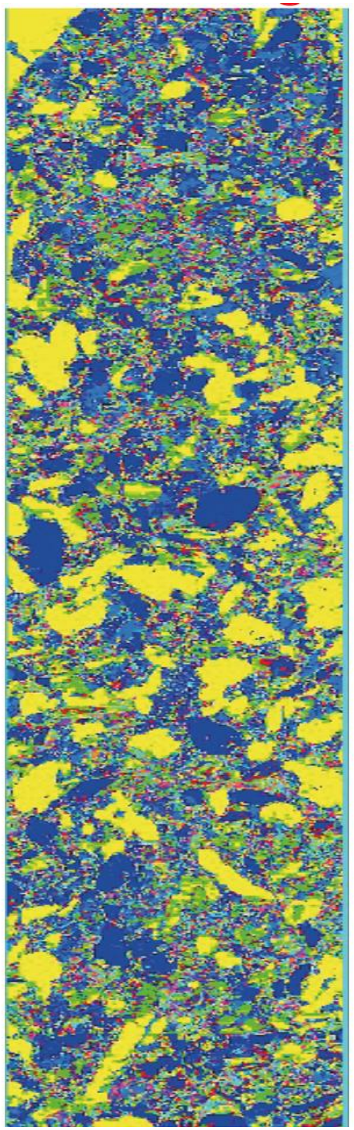


Figure 2: Custom porosity-permeability curve for Estaillades Limestone microporosity

This dataset can be downloaded at the British Geological Survey image archive: https://www.bgs.ac.uk/services/ngdc/citedData/catalogue/0543fe60-8e38-49ba-a8ec-a727e8babd25.html





Constructing the Ground Truth: Multi-scale imaging of structure

Table 1 Porosity and permeability values for micro-porosity calculated from synthetic images.

Porosity Range in Difference Image [%]	Porosity of Simulation [%]	Simulated Permeability [m²]	Fraction of Total Core Volume [%]	Segmentation Phase #
100	N/A	N/A (Pore)	9.95	1
54.45 – 99.9	56.89	7.47×10^{-15}	17.16	2
49.4 – 54.4	51.95	6.91×10^{-15}	4.63	3
44.3 - 49.3	46.83	4.79×10^{-15}	4.62	4
39.1 - 44.2	41.63	3.24×10^{-15}	4.86	5
31.6 - 39.0	36.48	2.12×10^{-15}	7.22	6
24.5 – 31.5	26.71	8.06×10^{-16}	6.93	7
20.3 - 24.4	22.27	4.59×10^{-16}	4.09	8
16.5 - 20.2	18.23	2.44×10^{-16}	3.65	9
13.1 - 16.4	14.63	1.17×10^{-16}	3.21	10
10.2-13.0	11.5	4.95×10^{-17}	2.78	11
7.8-10.1	8.83	1.73×10^{-17}	2.36	12
0.1-7.7	6.63	4.76×10^{-18}	7.73	13
0	N/A	N/A (Grain)	20.83	14
N/A	N/A	0 (Viton Sleeve)	N/A	15

Further information on these multiscale imaging and modelling methods can be found in: Menke et al., 2019 "Using nano-XRM and high-contrast imaging to inform micro-porosity permeability during Stokes-Brinkman single and two-phase flow simulations on micro-CT images." EarthArXiv

Dividing the image into sub volumes

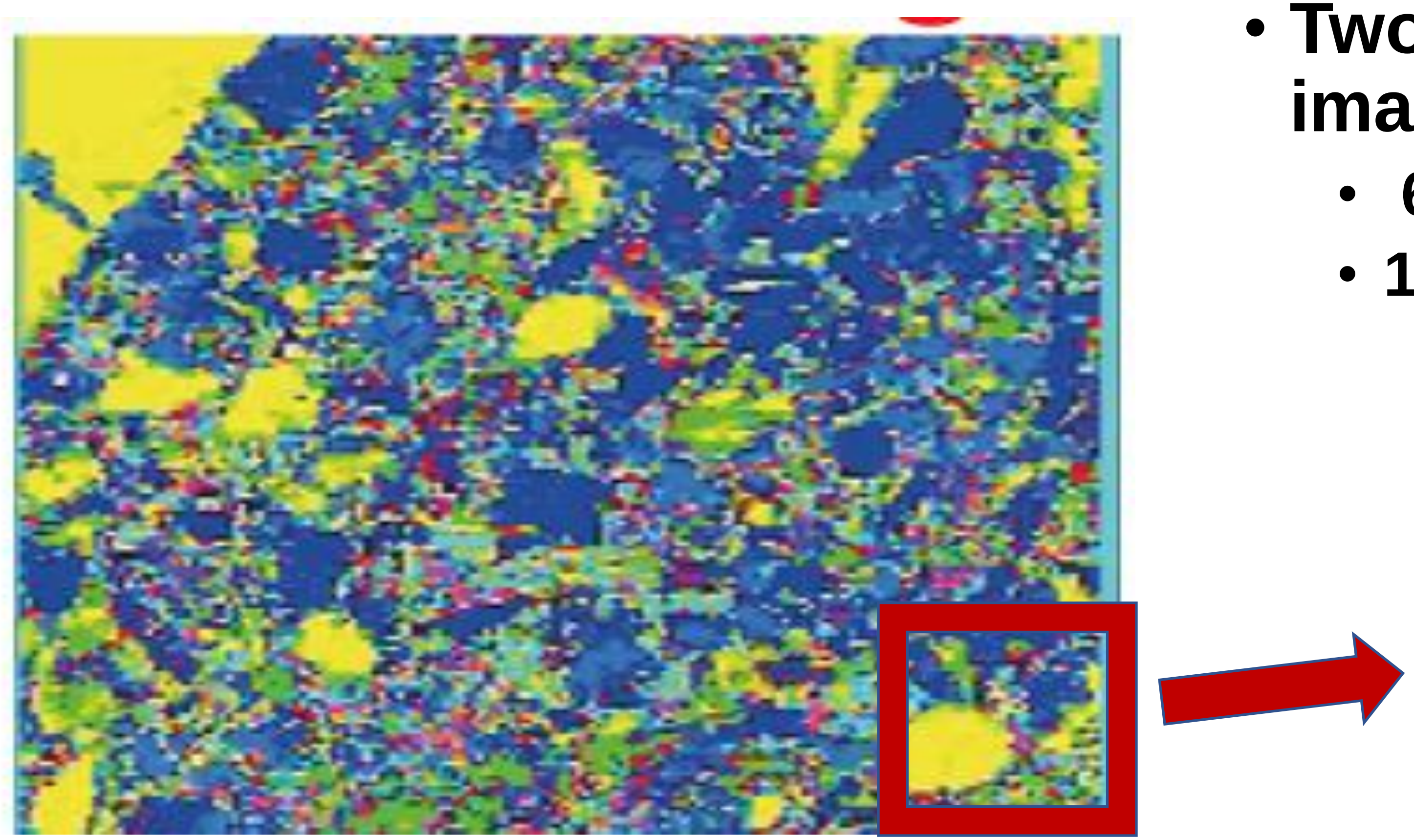


Figure 1 A 15 phase segmentation of a microporous carbonate. Figure Credit: Menke et al. 2019 EarthArXiv.

- Two separate training image sets:
  - 60^3 voxel sub volumes
  - 120^3 voxel sub volumes

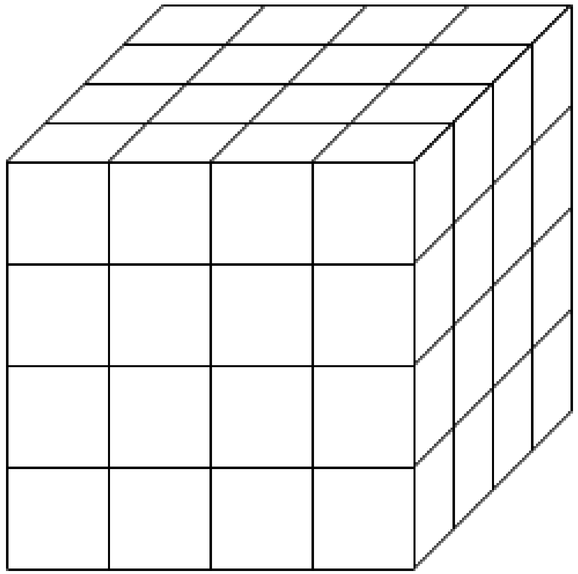


Figure 2 A matrix of sub volumes.

Feature extraction

- Feature set 1: volume fraction of each phase. (15 features)
- Feature set 2: connectivity of the phases in each orthogonal direction (first phase to connect inlet to outlet). (3 features)
- Total computational time was ~2-20s per sub volume on 24 processors for all feature extraction.

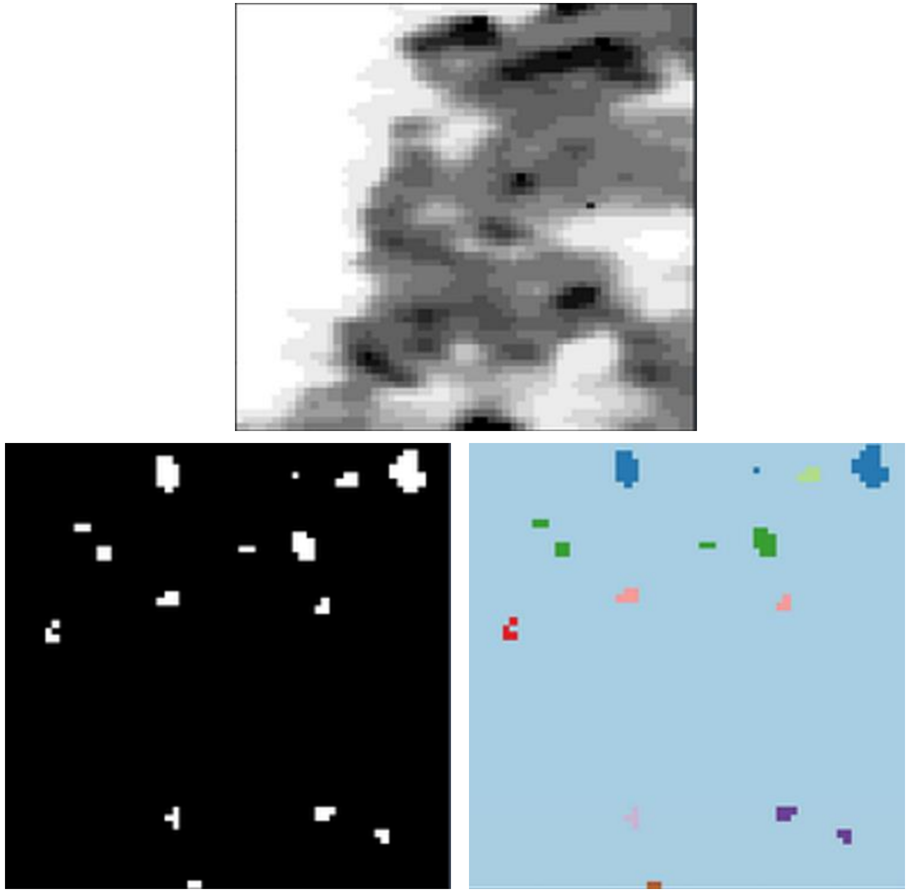


Figure 1 Labelling the connected components of the primary porosity in a sub volume. (top) 15 phase segmentation (left) primary porosity (right) labelled connected primary porosity.

GeoChemFoam - Darcy-Stokes-Brinkman

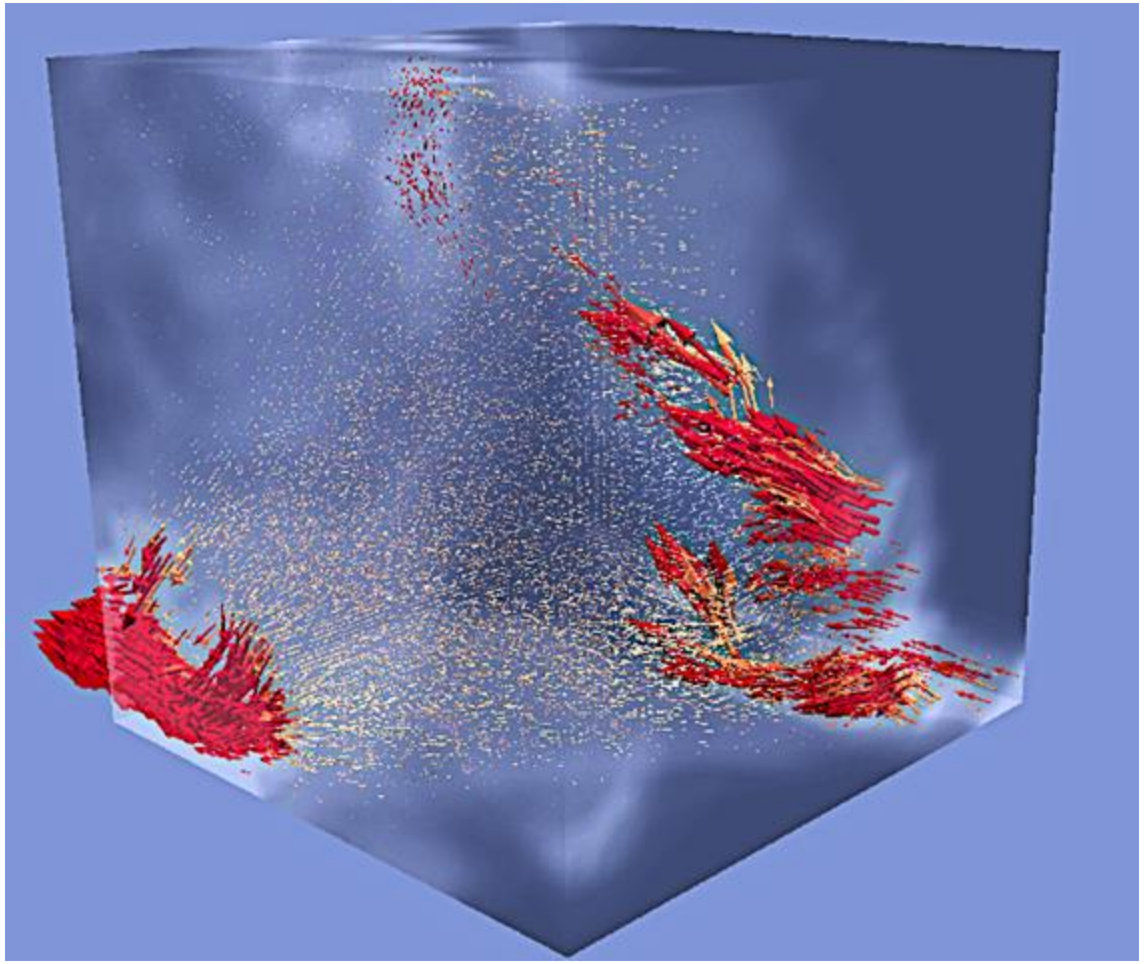


Figure 1 A 60^3 voxel sub volume where flow is computed using the DBS equation. Flow through the primary porosity (red) is connected by flow through the microporous matrix (yellow). Voxel porosity is shown in grayscale.

- GeoChemFoam: Highly versatile and open source multiphase reactive transport solver bulit using the OPENFOAM platform by Dr Julien Maes at Institute for GeoEnergy Engineering at Heriot-Watt University: https://www.julienmaes.com/geochemfoam.
- Each sub volume was solved for permeability in the X, Y, and Z directions using the Darcy-Stokes-Brinkman solver in GeoChemFoam.
- Average computational time per subvolume was 2 mins for the 60^3 volumes and 15 mins for the 120^3 voxels on 24 processors.
- The solved permeabilities become our Ground Truth to train the decision trees against the feature sets extracted with image analysis.

DBS Equation:

$$0 = -\nabla p + \nabla(\mu_e \nabla u) + \mu K^{-1} u$$

Darcy-Stokes-Brinkman Numerically Solved Permeability for 30,000 sub volumes of 60^3 voxels

- **Non-linear relationship between porosity and permeability that does not easily fit onto any power law or exponential model**
- **Estaillades is a reasonably simple microporous carbonate. What would a more complex one look like?**

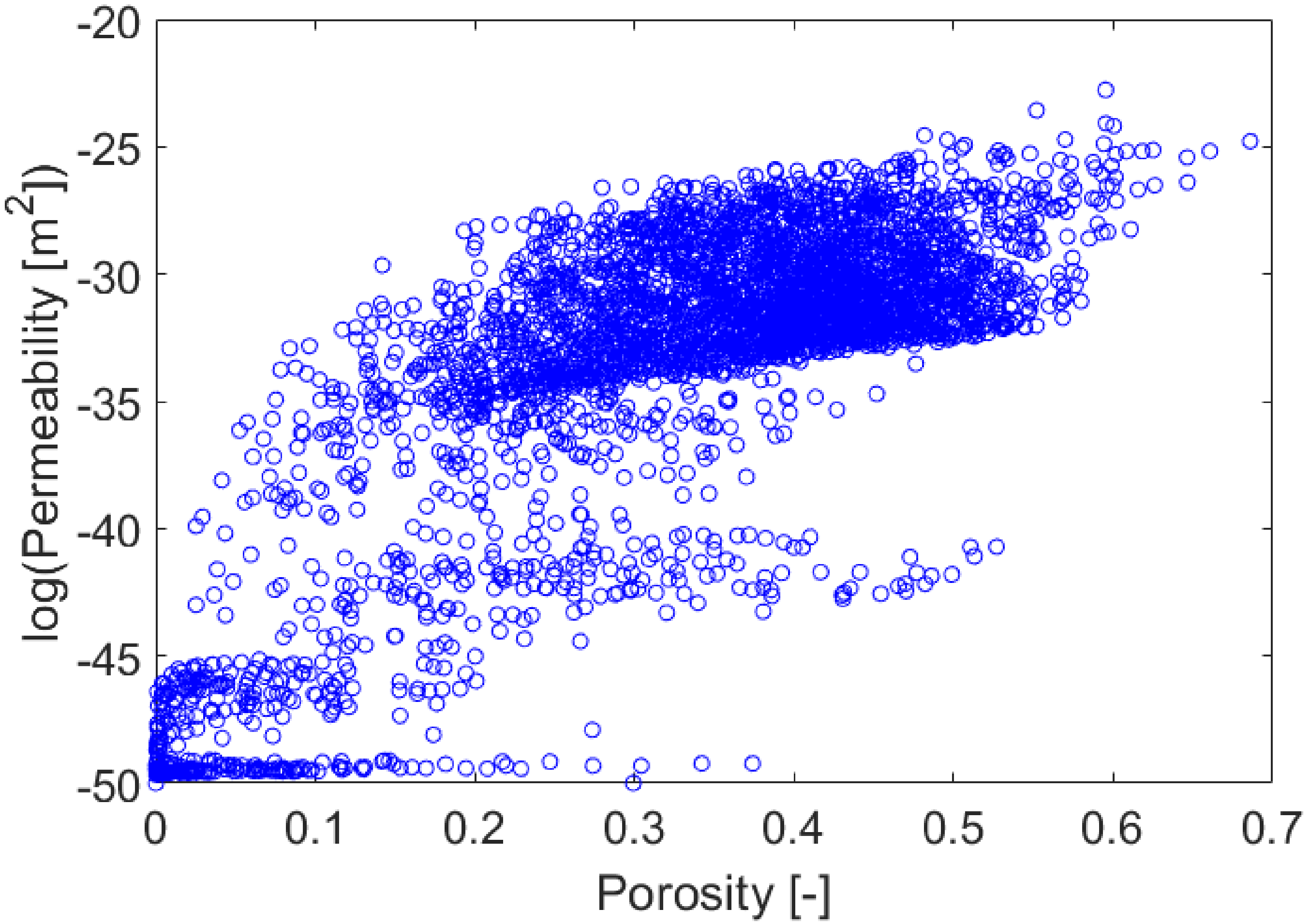


Figure 1 The porosity and numerically solved permeability for each of the 30,000 sub volumes of 60^3 voxels

Extra Randomised Trees Ensemble - SciKitLearn

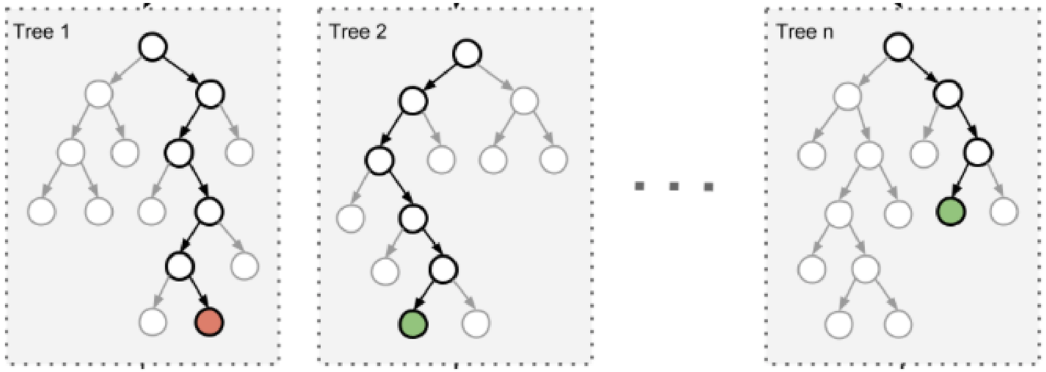


Figure 1 Decision Tree Schematic of an Extremely Randomized Forest.

Figure Credit: https://blog.statsbot.co/ensemble-learning-d1dcd548e936

```
from sklearn.ensemble import ExtraTreesRegressor
est = ExtraTreesRegressor(n_estimators=50,
                          max_features='auto',
                          max_depth=None,
                          min_samples_split=2,
                          min_samples_leaf=1,
                          min_weight_fraction_leaf=0,
                          max_leaf_nodes=None,
                          n_jobs=-1, bootstrap=True,
                          oob_score=True)

est.fit(IP_training_attributedata, SB_training_DNSresults)
```

Figure 2 Code for the Extra Trees Regressor in SciKitLearn

Regression Model Training:

- **Input**:
  - feature set of 18 variables (15 vol fractions, 3 phase connectivity values)
  - Ground Truth numerically solved permeability values (X,Y,Z)
- **Output**: Feature weights, oob score, R^2 value, RMSE
- **Computational Time**: ~2 seconds on 24 processors

Regression Model Testing:

- **Input**: feature set of 18 variables (15 vol fractions, 3 phase connectivity values)
- **Output**: Predicted Permeability
- **Computational Time**: <0.1 second on 24 processors

Question 1a: Does multivariate regression predict permeability accurately?

- **The features from 1000 60^3 sub volumes <u>not used in the model training</u> were input into the trained regression model.**

- **Root Mean Squared Error was 4.3%**

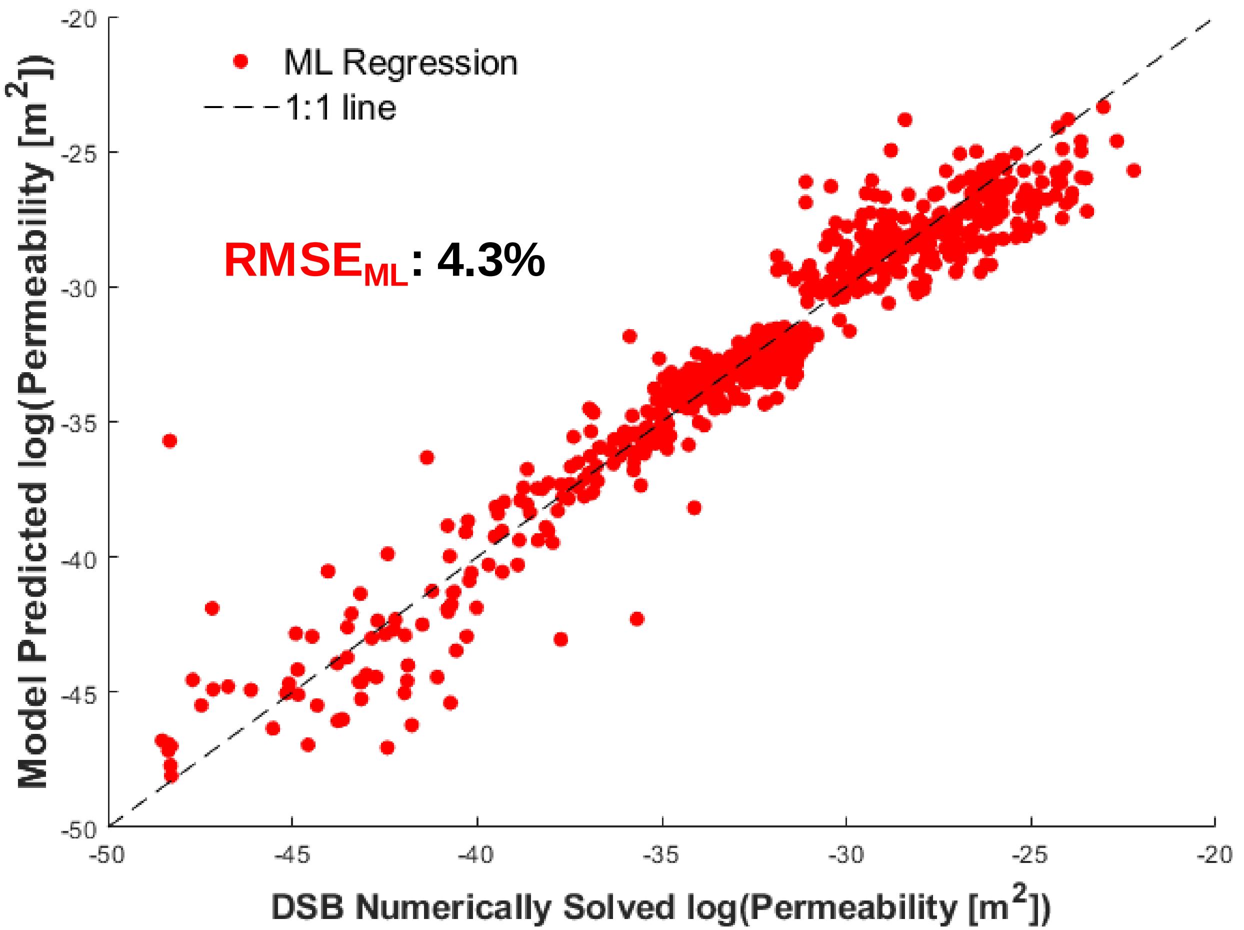


Figure 1 The machine learning regression model predicted permeability plotted against the Darcy-Stokes-Brinkman solved permeability for 1,000 sub volumes of 60^3 voxels.

Question 1b: Is this prediction better than using the traditional Kozeny-Carman approach?

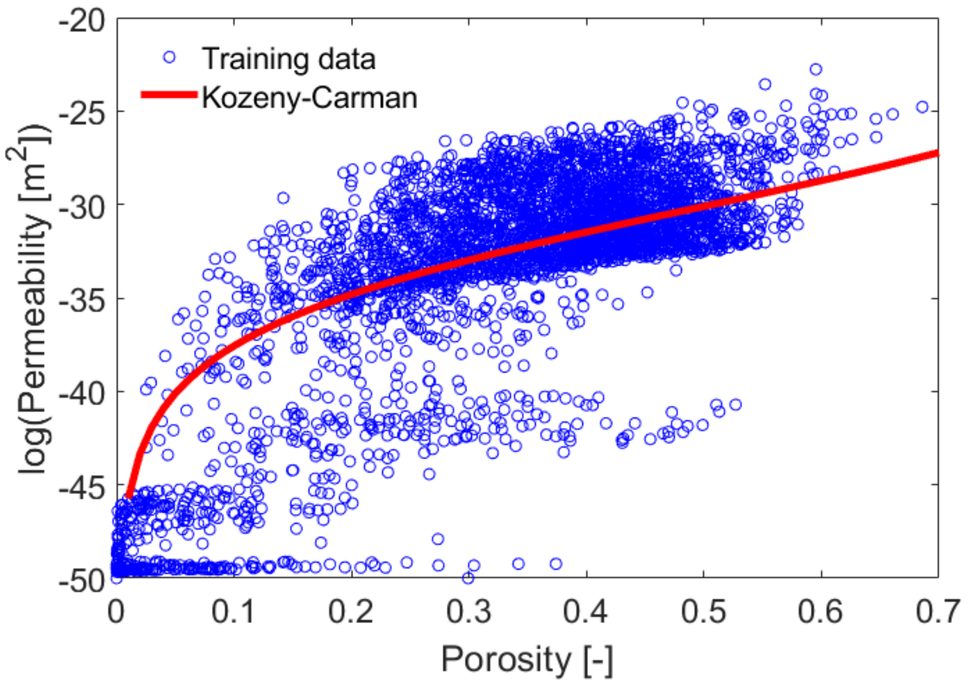


Figure 1 The porosity and log(Permeability) of 29,000 sub volumes of 60^3 voxels (blue) and the best fit power law (red) used to estimate the Kozeny-Carman model parameters.

Kozeny-Carman Fit:

$$K = 8.47 \times 10^{-14} \frac{\varphi}{1-\varphi}^{3.4}$$

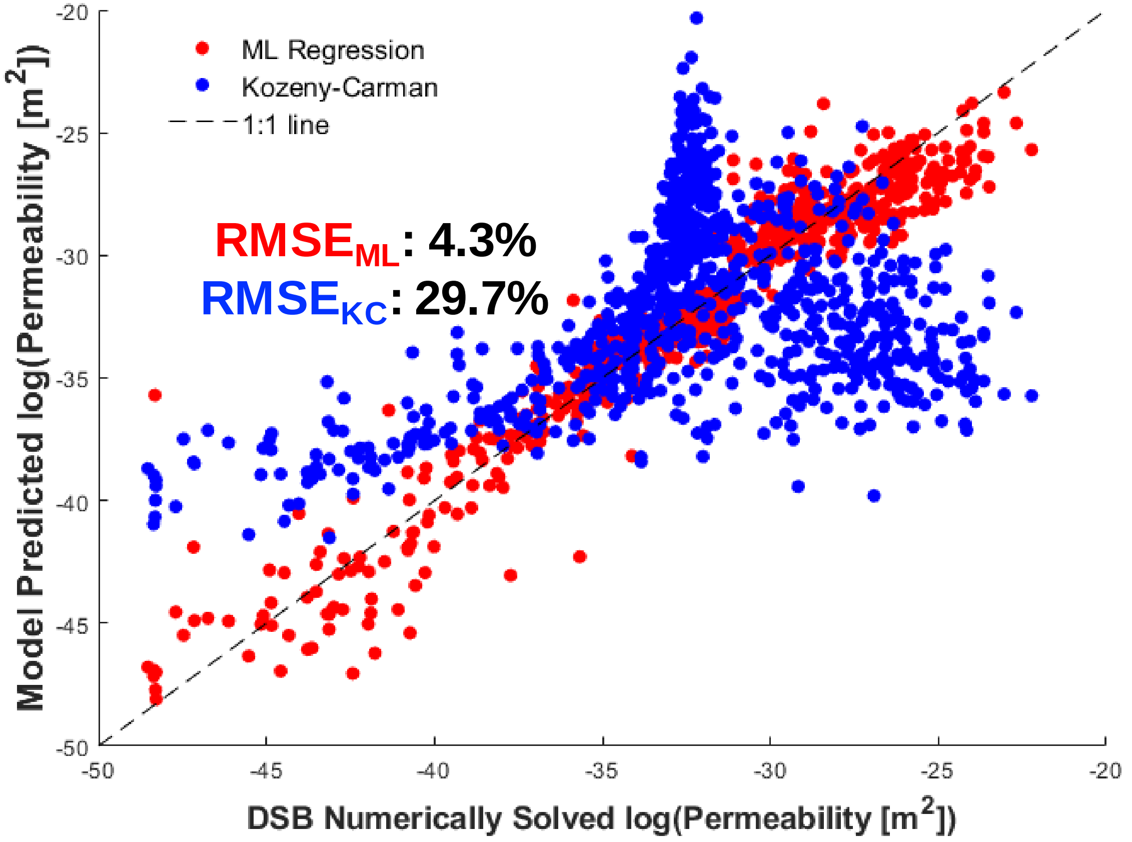


Figure 2 The machine learning regression model predicted permeability and the Kozeny-Carman permeability plotted against the Darcy-Stokes-Brinkman solved permeability for 1,000 sub volumes of 60^3 voxels.

Question 2a: How does the choice of features affect the outcome?

- **Two feature sets:**
  - **15 Phase Volume Fraction features**
  - **3 Connectivity features (one for each X,Y,Z direction)**
- **Connectivity information matters.**

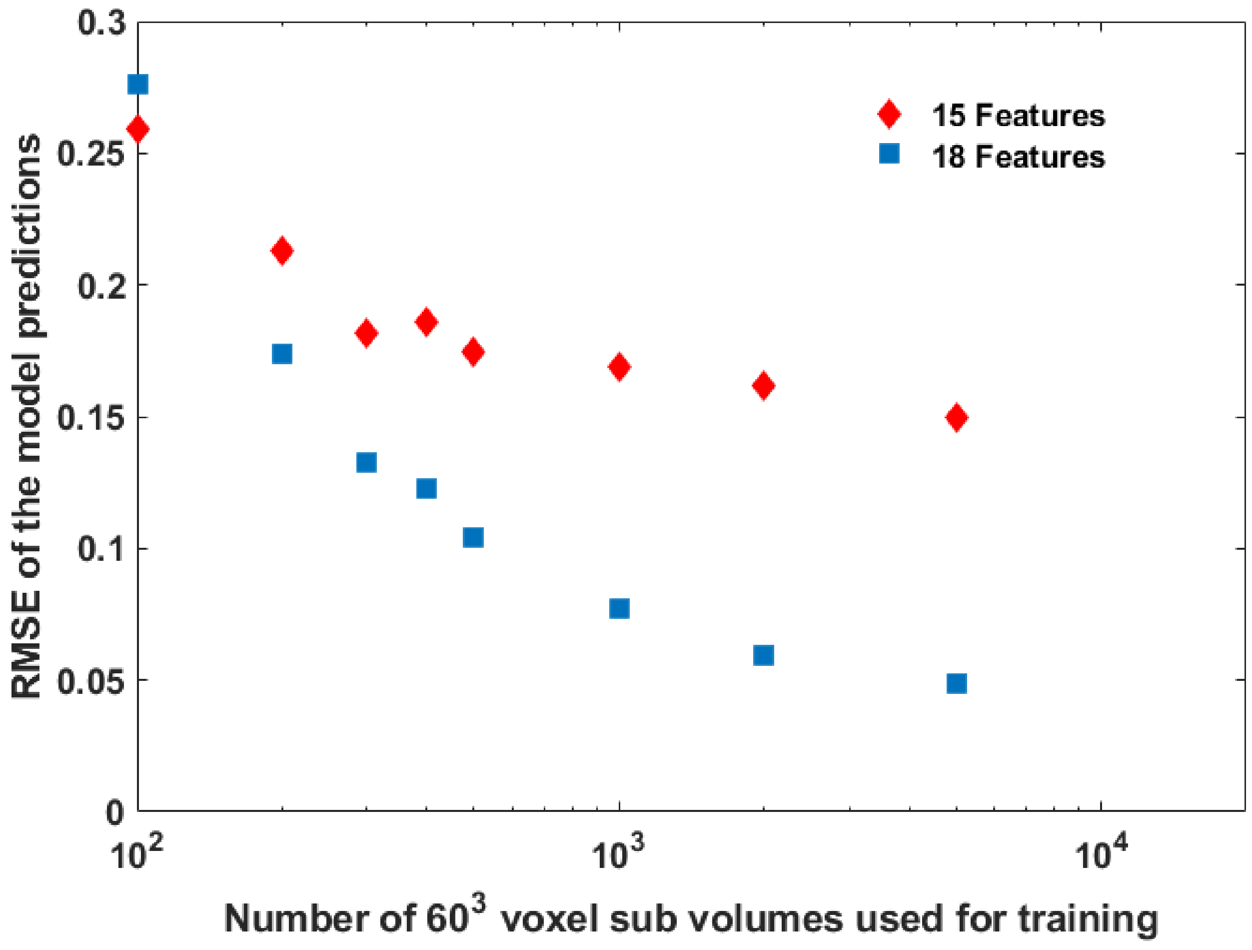
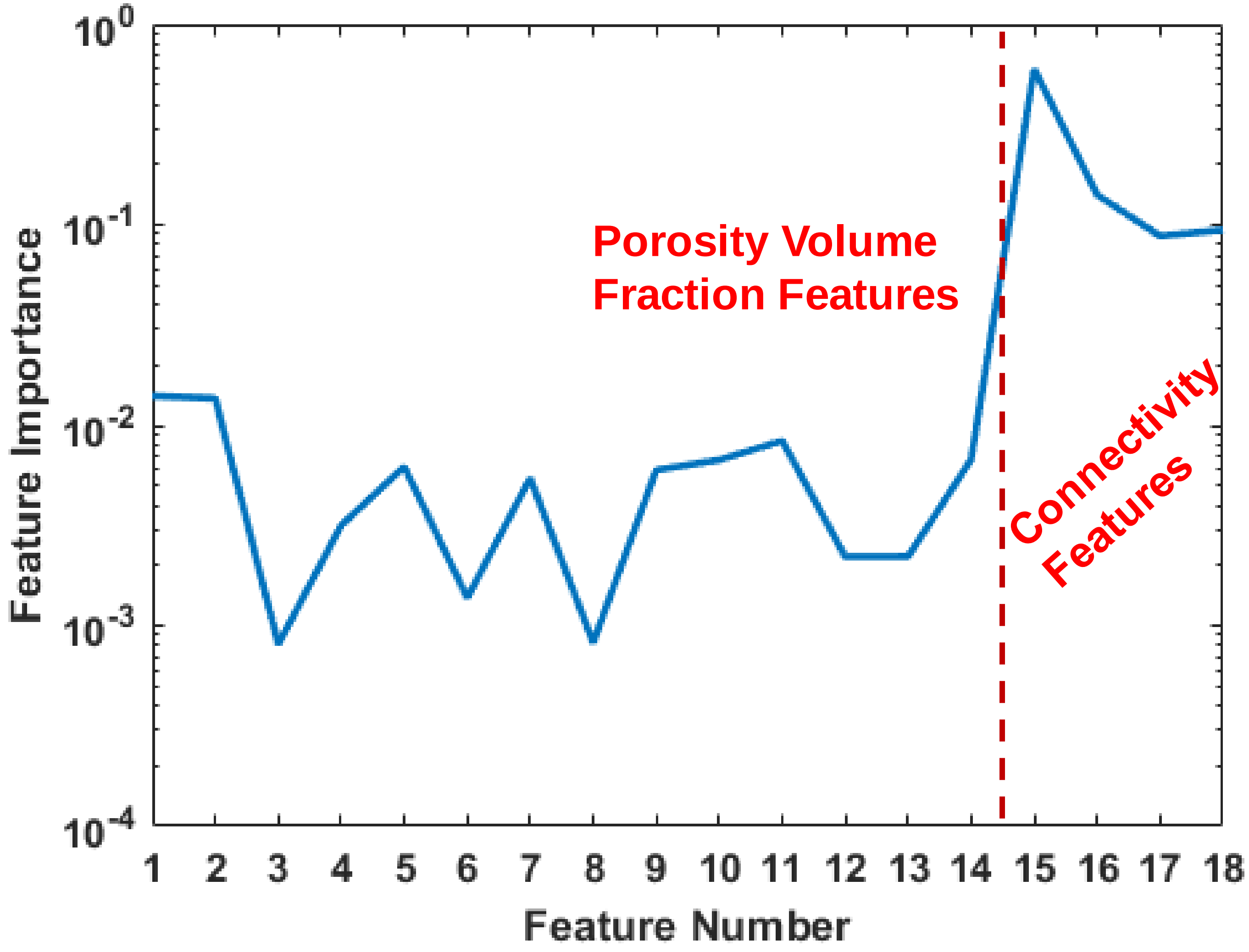


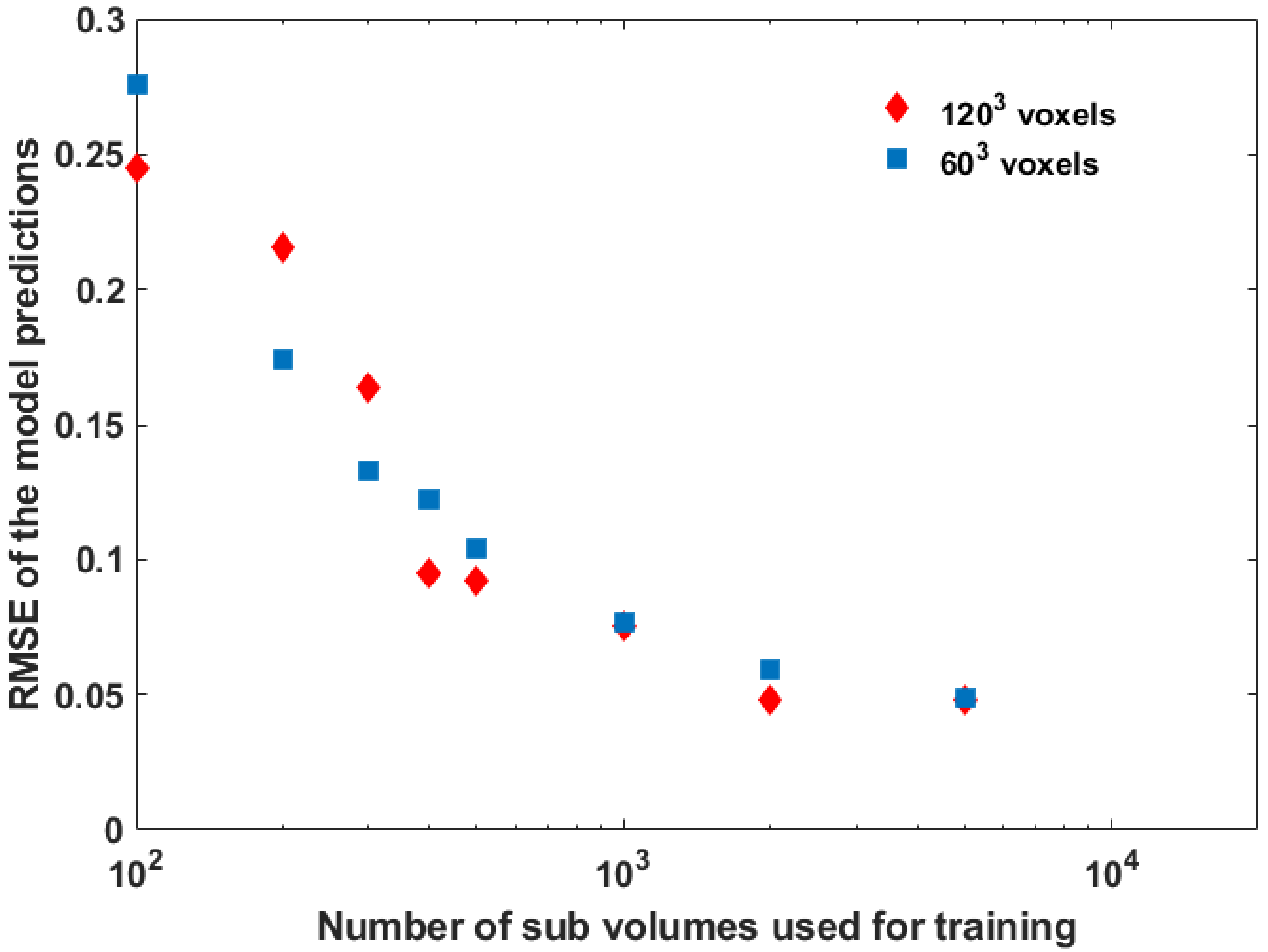
Figure 1 A comparison between the number of features used, the number of training images, and the RMSE of the model

Question 2b: Which features are the most important for the regression?



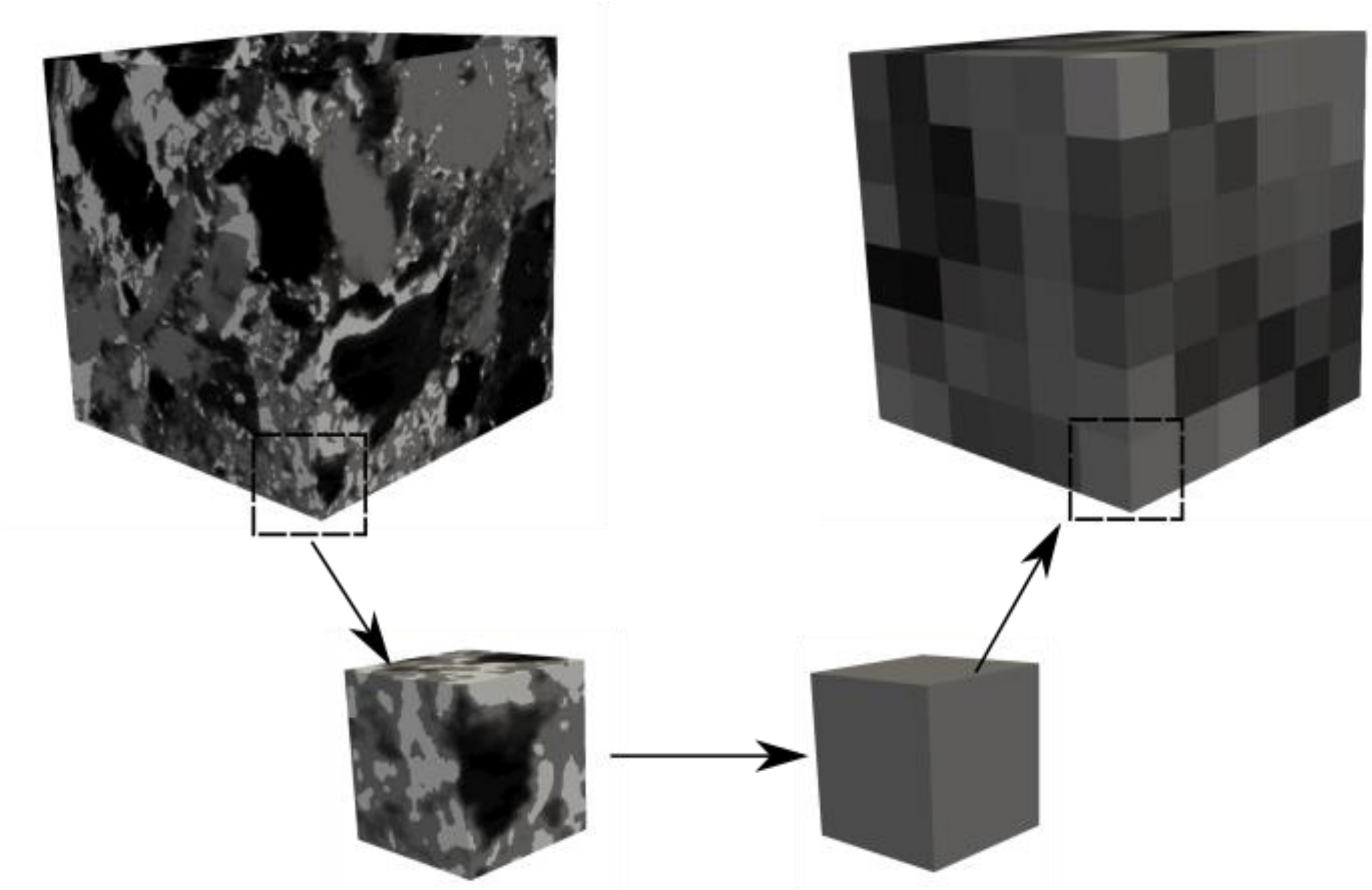
Question 3: Does the size of the sub volume change the model performance?

- **Answer: not for these sizes**
- **Could investigate 180^3 but the computational cost is much higher.**



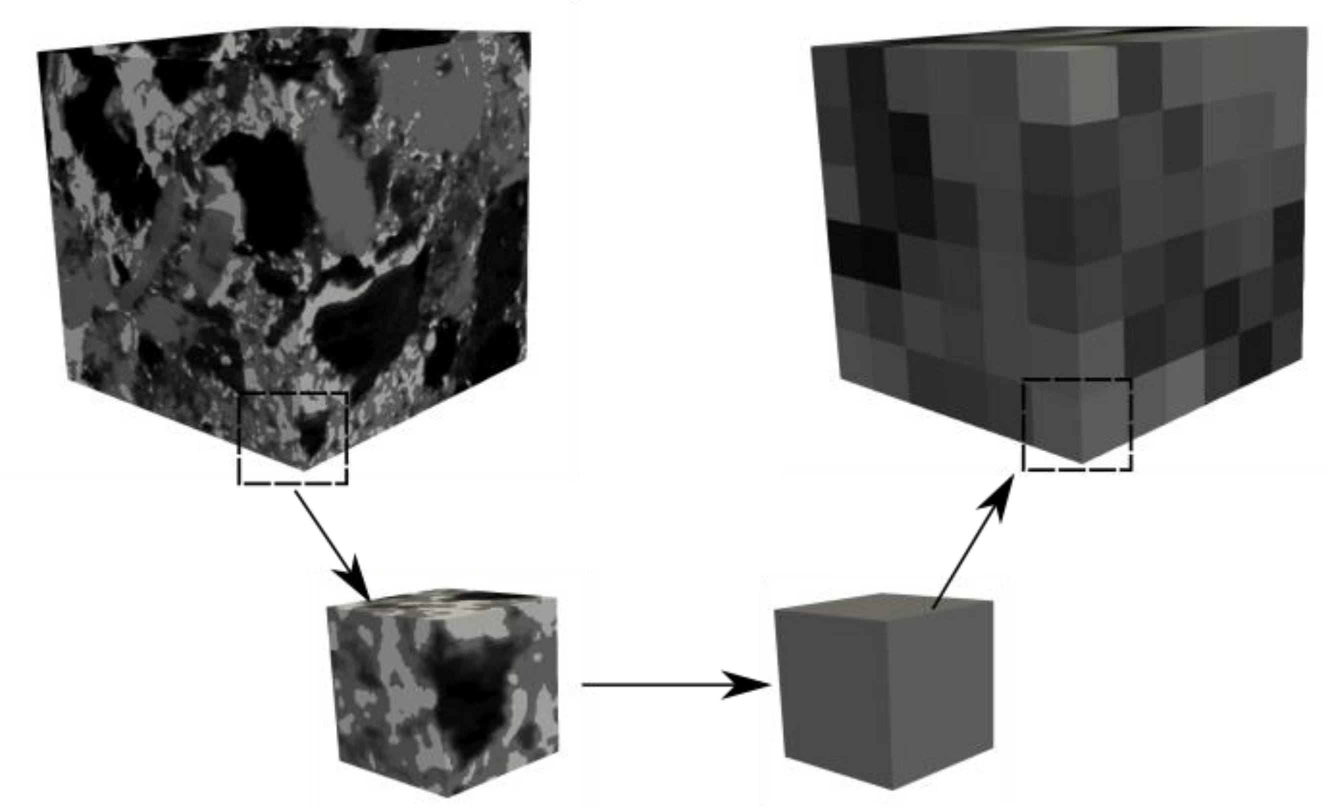
Question 4: How can this MVR model be used for upscaling at the Darcy scale?

- **Cut out three 360^3 blocks and divided them into:**
  a) **6x6x6 matrices of 60^3 sub volumes**
  b) **3x3x3 matrices of 120^3 sub volumes**



Question 4: How can this MVR model be used for upscaling at the Darcy scale?

1. Numerically solved the 360^3 with DBS
2. Numerically solved both the 60^3 and 120^3 sub volumes with DBS and used the output permeability to solve a Darcy simulation
3. Used the features of the 60^3 and 120^3 sub volumes as input into the ML regression and then used the output permeability to solve a Darcy simulation.
4. Used the porosity of the 60^3 and 120^3 sub volumes as input into the Kozeny-Carman model and then used the output permeability to solve a Darcy simulation.



Question 4: Can we upscale?

Size [voxels^3]		Volume 1	Volume 2	Volume 3	Total Run Time [min]*
360	Porosity	0.36	0.43	0.35	-
	Darcy Stokes Brinkman Permeability [m^2]	6.59×10^{-14}	9.47×10^{-15}	7.67×10^{-14}	480
60	Numerical DBS upscaled Permeability [m^2]	6.25×10^{-14} -5%	8.13×10^{-15} -16%	7.19×10^{-14} -7%	80
	Machine Learning upscaled Permeability [m^2]	6.66×10^{-14} 1%	9.99×10^{-15} 5%	5.62×10^{-14} -36%	1
	KC upscaled Permeability [m^2]	3.40×10^{-14} -94%	4.15×10^{-14} 77%	2.40×10^{-14} **-220%**	1
120	Numerical DBS upscaled Permeability [m^2]	4.50×10^{-14} -46%	2.05×10^{-14} 54%	1.01×10^{-13} 24%	80
	ML upscaled Permeability [m^2]	7.22×10^{-14} 9%	1.30×10^{-14} 27%	5.71×10^{-14} -34%	1
	Kozeny-Carman upscaled Permeability [m^2]	1.80×10^{-14} **-266%**	3.40×10^{-14} 72%	1.76×10^{-14} **-336%**	1

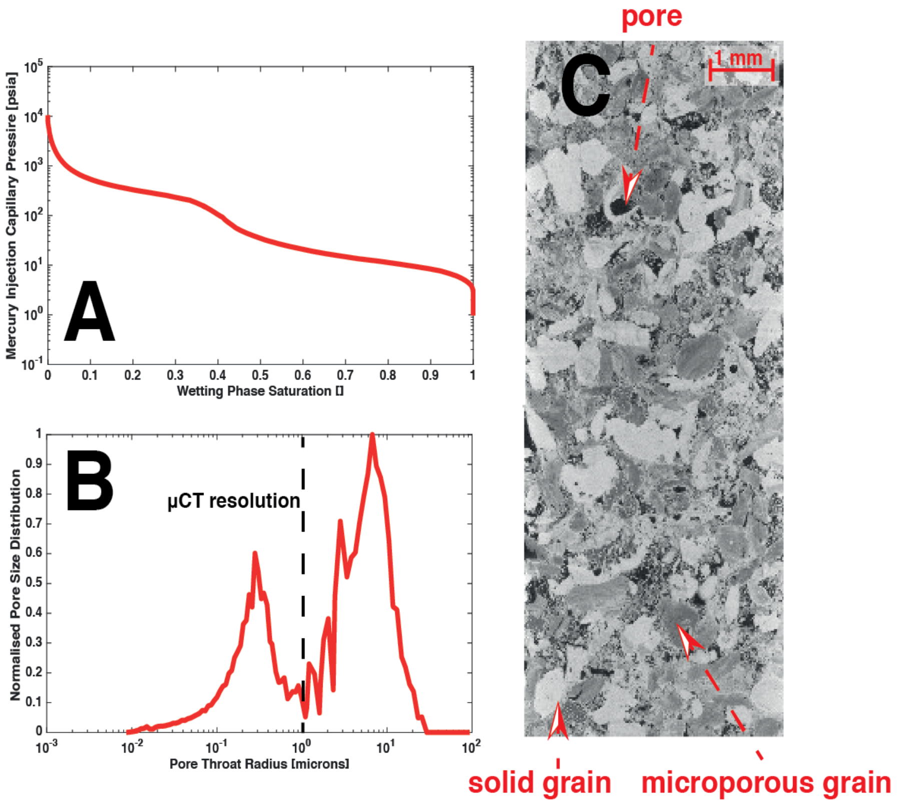
*All model run times are for a 24 CPU workstation and summed across all volumes

Conclusions

- **ML Regression Models can be trained to high accuracy with surprisingly little data (~1-5K sub volumes).**
  - **Note: this will probably change with rock complexity**
- **Increasing sub volume size had little effect on model predictions**
- **The ML Regression Model outperformed the Kozeny-Carman model by over 20% (in log space!) for both same scale prediction and upscaled Darcy simulations.**
- **The ML Regression Model had similar accuracy to the full DBS simulation with a fraction of the computations cost (1/500th)**

Perspective

- **Estalliades is a relatively simple carbonate with a bimodal pore-size distribution**
- **Absolute permeability is static in this system (rather than dynamically changing)**
- **I was choosing and extracting the features based on expert knowledge of the system (instead of automation)**
- **Trained model is specific to a single system**

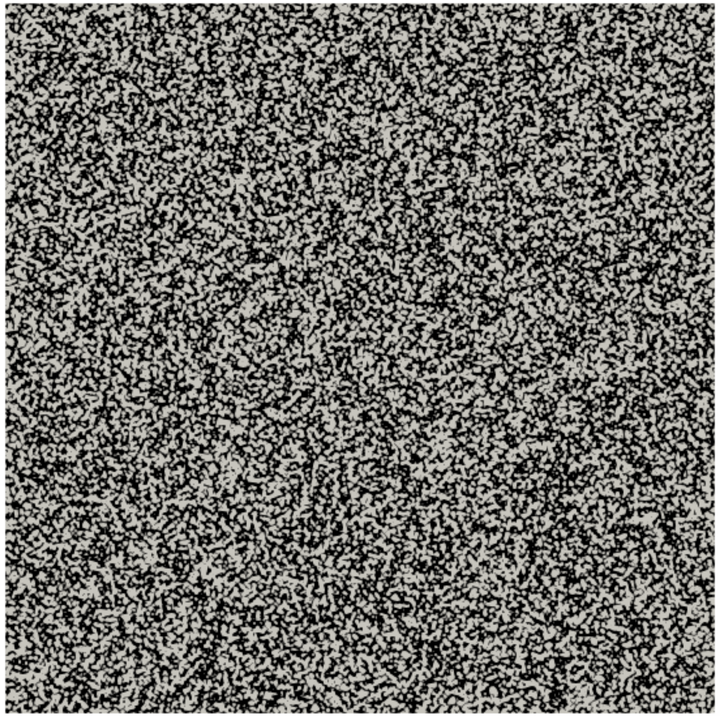


Outlook

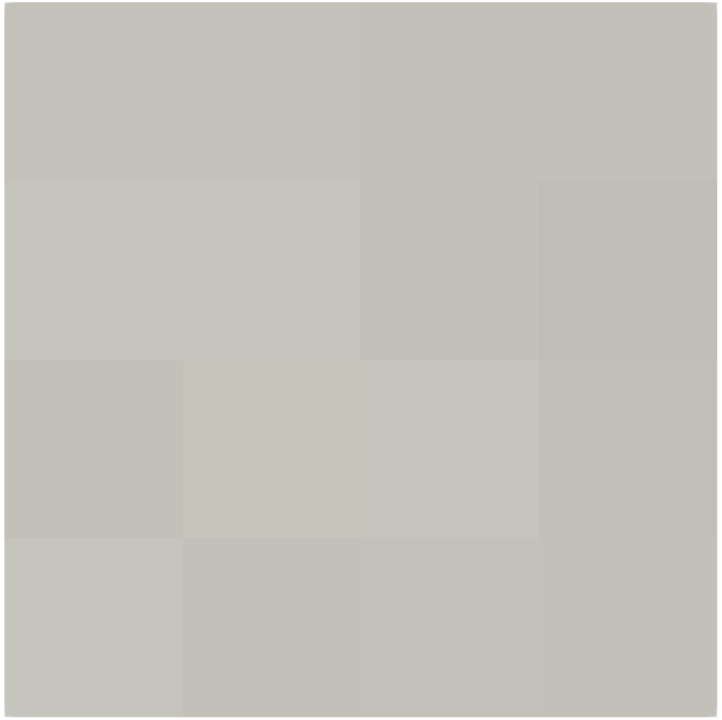
There is untapped potential to use machine learning with numerical modelling and imaging for upscaling flow and transport processes

Upscaling reactive dissolution

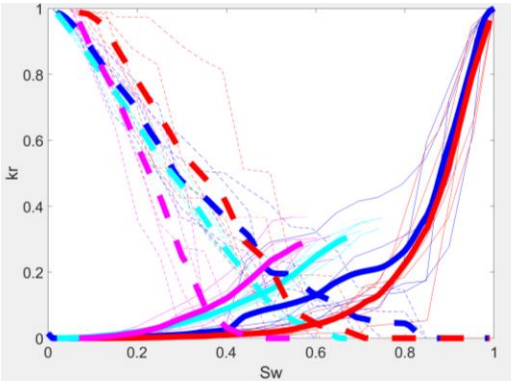
Pore-scale DBS Model Pe=0.3 K=10

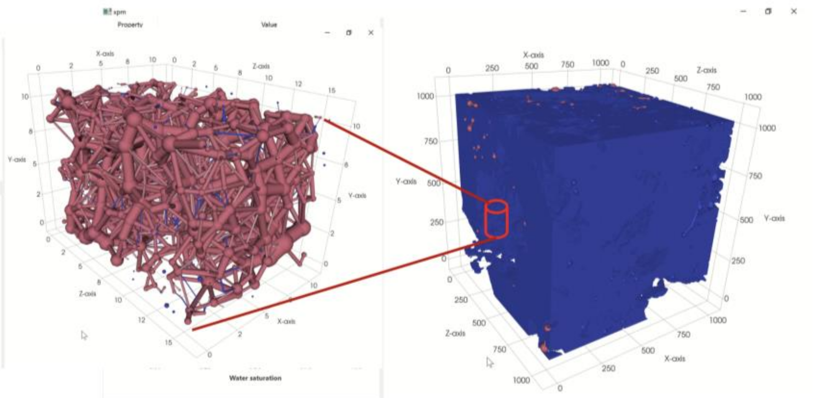


Machine learning Darcy-scale Model Pe=0.3 K=10



Upscaling multiphase flow



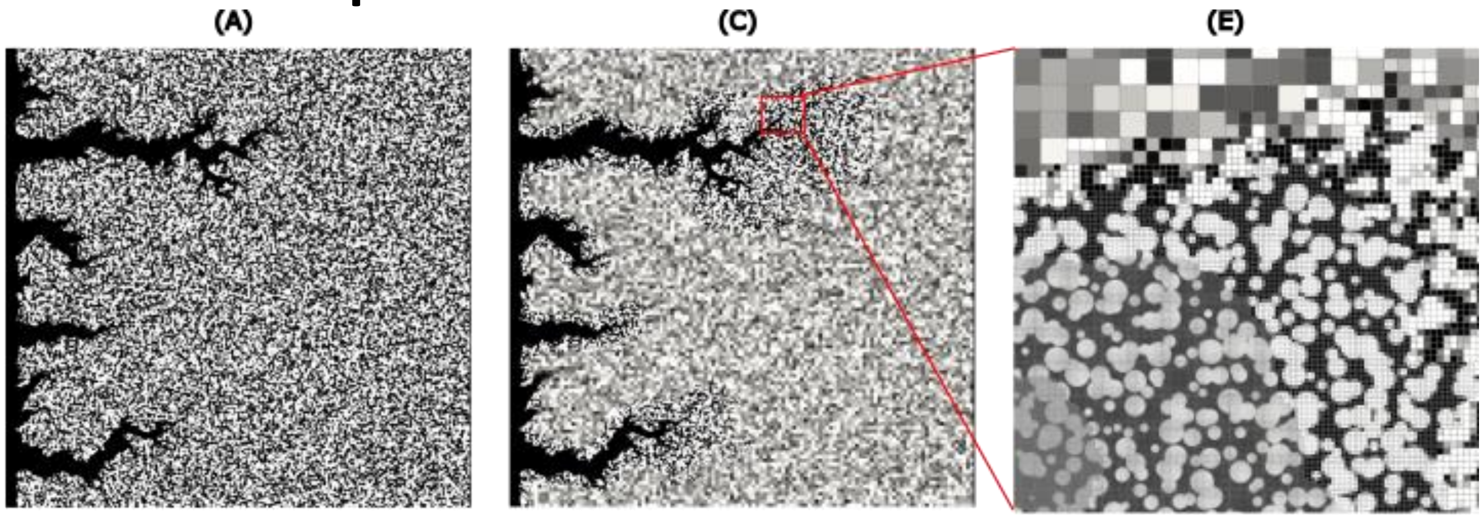


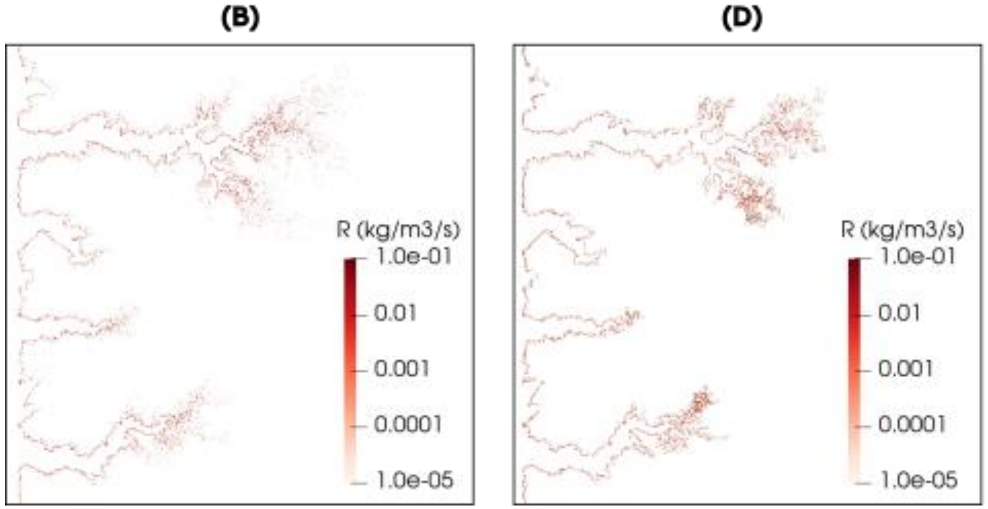
(lead by Kamaljit Singh)

Outlook

There is untapped potential to use machine learning for increasing numerical solver speed and efficiency

Adaptive mesh refinement with ML

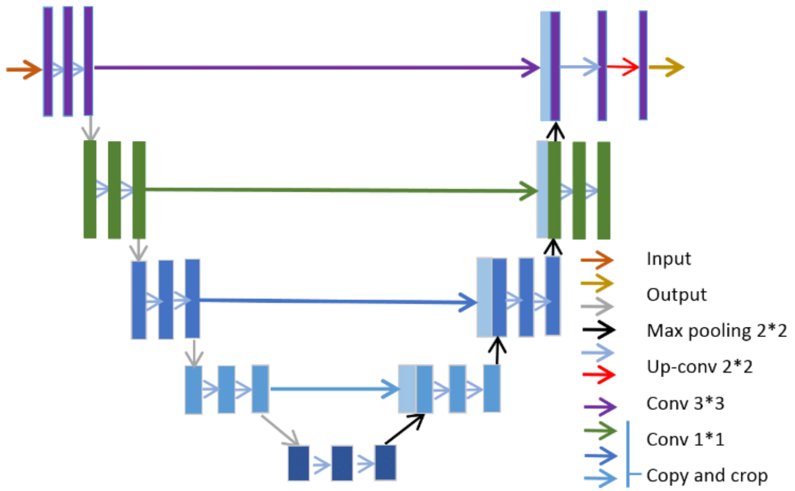




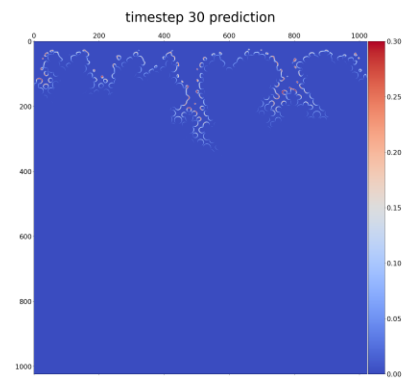
	Fine mesh	Coarse mesh
Number of cells	16000000	577689
Porosity	0.49	0.49
Permeability (m^2)	9.8×10^{-10}	9.8×10^{-10}
Reaction rate (kg/s)	2.8×10^{-11}	2.7×10^{-11}

(lead by Julien Maes)

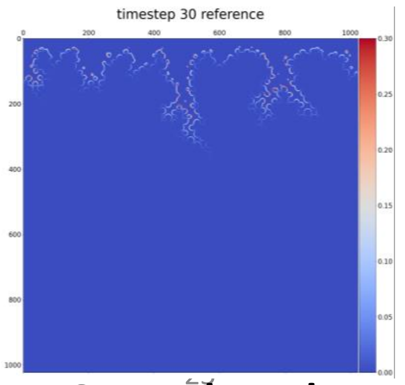
Improving model speed with ML



UNet Architecture



Predicted Dissolution



Ground Truth

EP/Y006143/1, EP/Y005732/1

Acknowledgements

- **This work was generously funded by the Energi Simulation at the Institute for GeoEnergy Engineering at Heriot-Watt University and benefited from the numerical advances in the GeoChemFoam Project partially funded by EPSRC.**
- **Special thanks to Dr Matthew Andrew (Zeiss Microscopy) for insightful conversations.**

Questions?