Machine Learning Applications in Exploration and Mining

Tom Carmichael, Brenton Crawford, Liam Webb.

QEC - The role of data in discovery
28th February 2018

SOLVE GEOSOLUTIONS

Outline

Where and when should we use machine learning?

- *Why ML can be the most optimal answer but not the best answer*
- *Hurdles to the successful implementation of ML*

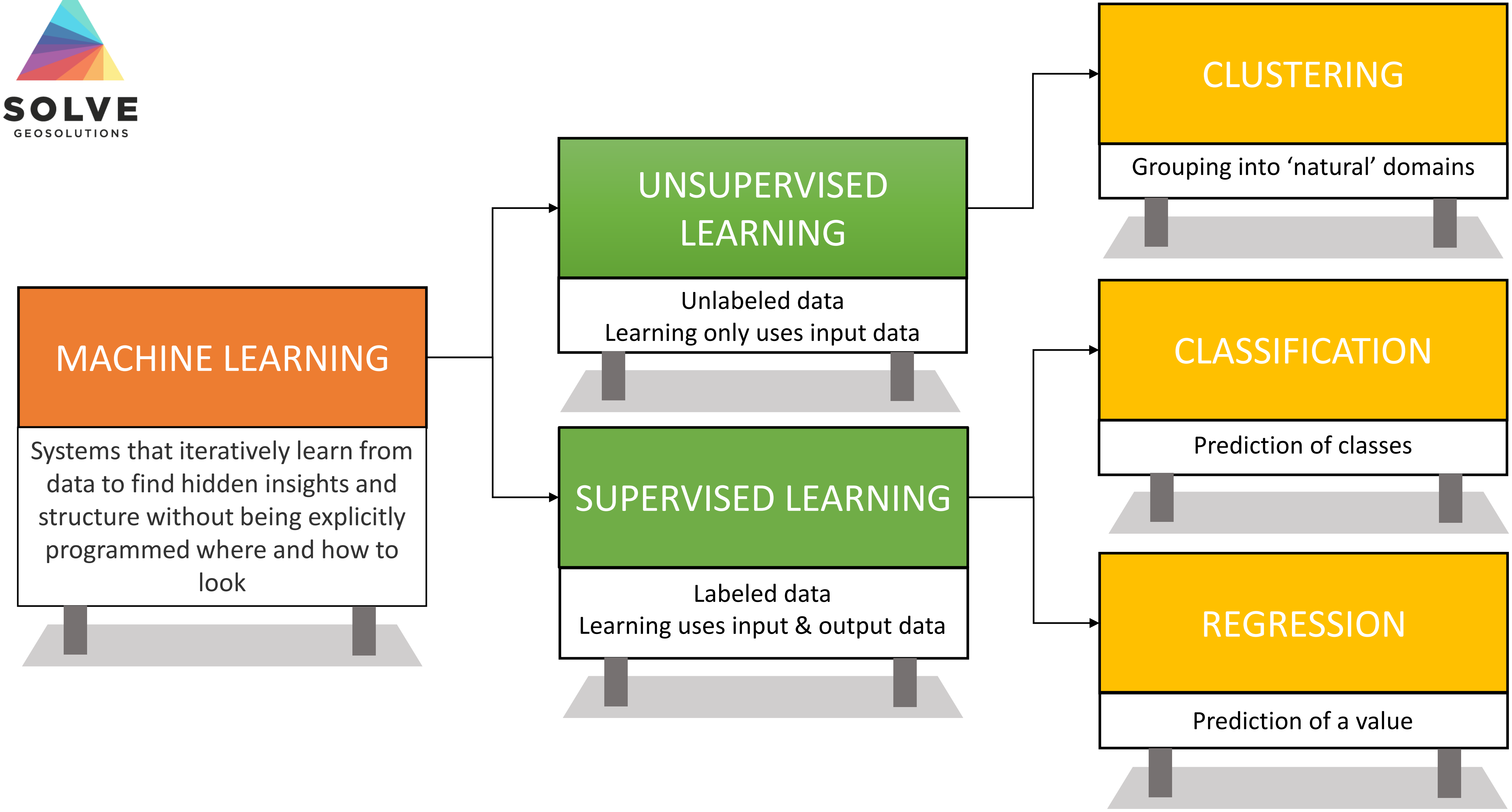
Examples of machine learning applied to exploration and mining

- *Searching for surface signatures in regional datasets*
- *Creating data-driven mineral domains using clustering*
- *Classifying Corescan mineral textures*
- *Quantifying the relationship between mineral associations and Au*

AURORA MINERALS GROUP

NEWCREST MINING LIMITED

corescan


MACHINE LEARNING
Systems that iteratively learn from data to find hidden insights and structure without being explicitly programmed where and how to look
UNSUPERVISED LEARNING
Unlabeled data
Learning only uses input data
SUPERVISED LEARNING
Labeled data
Learning uses input & output data
CLUSTERING
Grouping into 'natural' domains
CLASSIFICATION
Prediction of classes
REGRESSION
Prediction of a value

Getting it right at the start of a project

Asking the right question

Finding the appropriate data

Pre-processing inputs

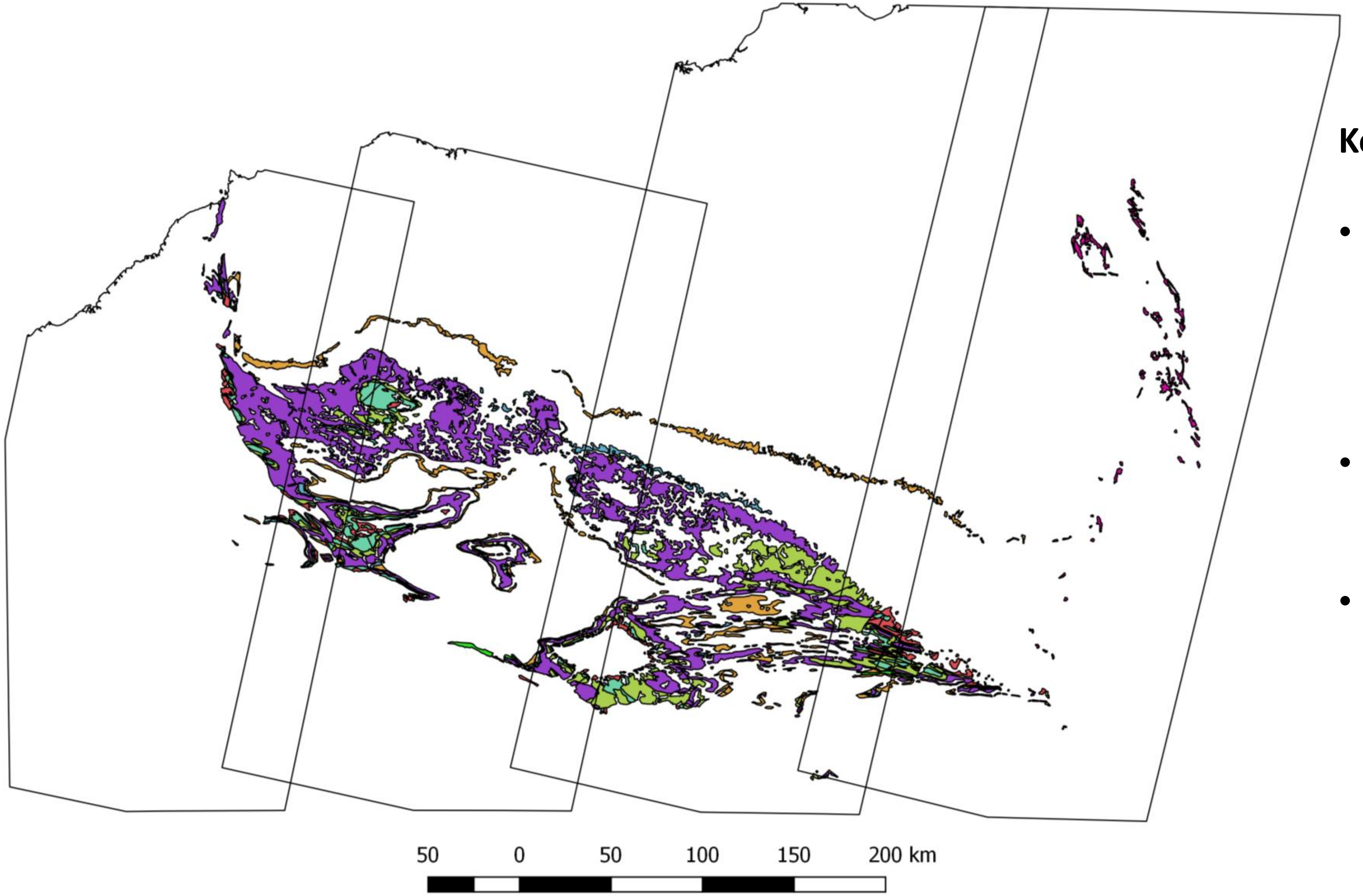
80% of the work involved is getting these parts right!

Machine Learning algorithm

The algorithm of used in this phase and the tuning of hyperparameters doesn't matter if the above steps aren't addressed first.

Example 1
Searching for surface signatures in the Pilbara using supervised learning

Searching for surface signatures in the Pilbara using supervised learning

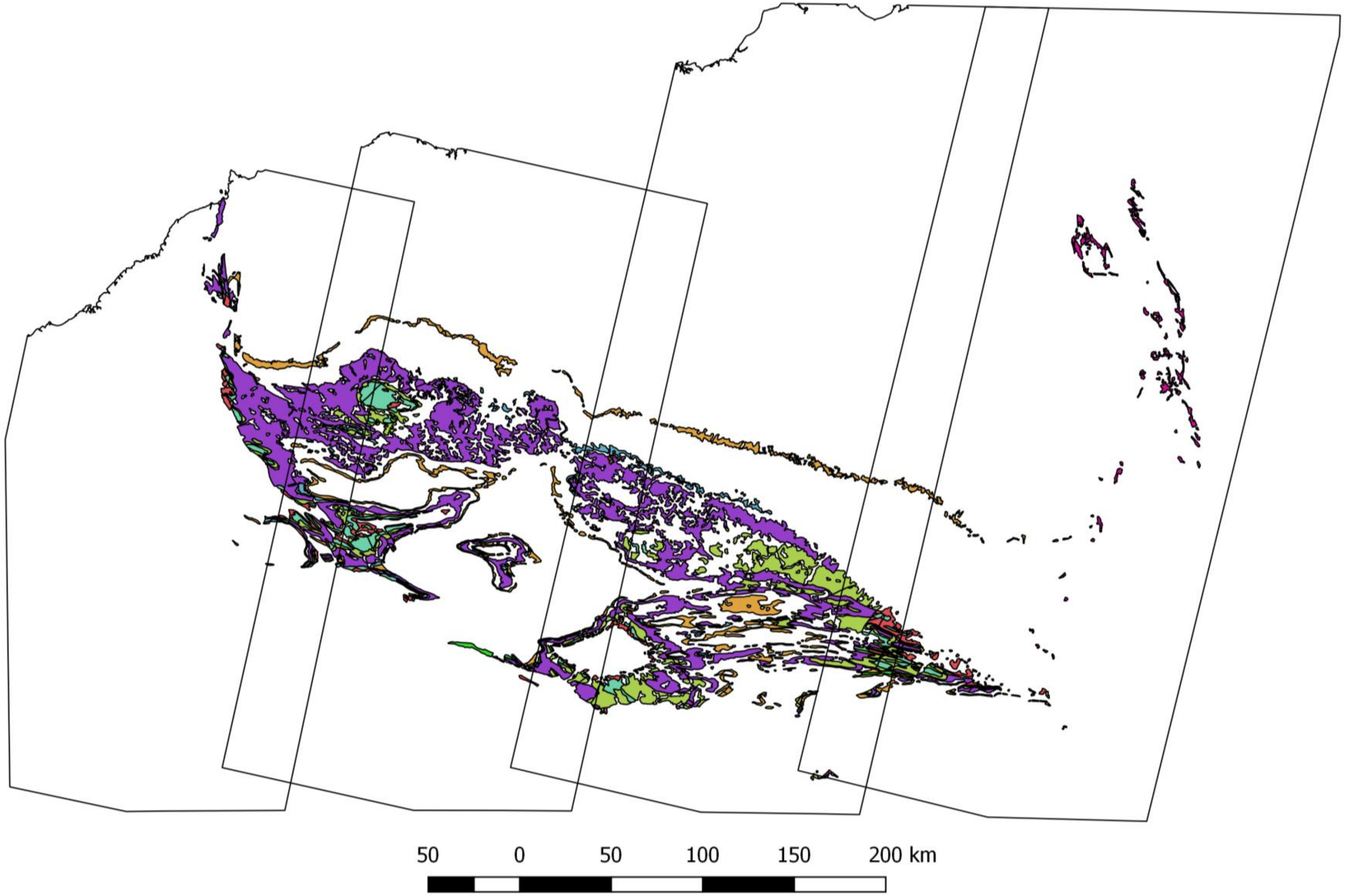


Key Questions:

- Can we find additional non-mapped exposures of economic iron-bearing lithologies either in outcrop or regolith?
- Can we find areas of the map that have potentially been misclassified?
- Where is the mapping in agreement/disagreement with the data?

AURORA MINERALS GROUP

Searching for surface signatures in the Pilbara using supervised learning

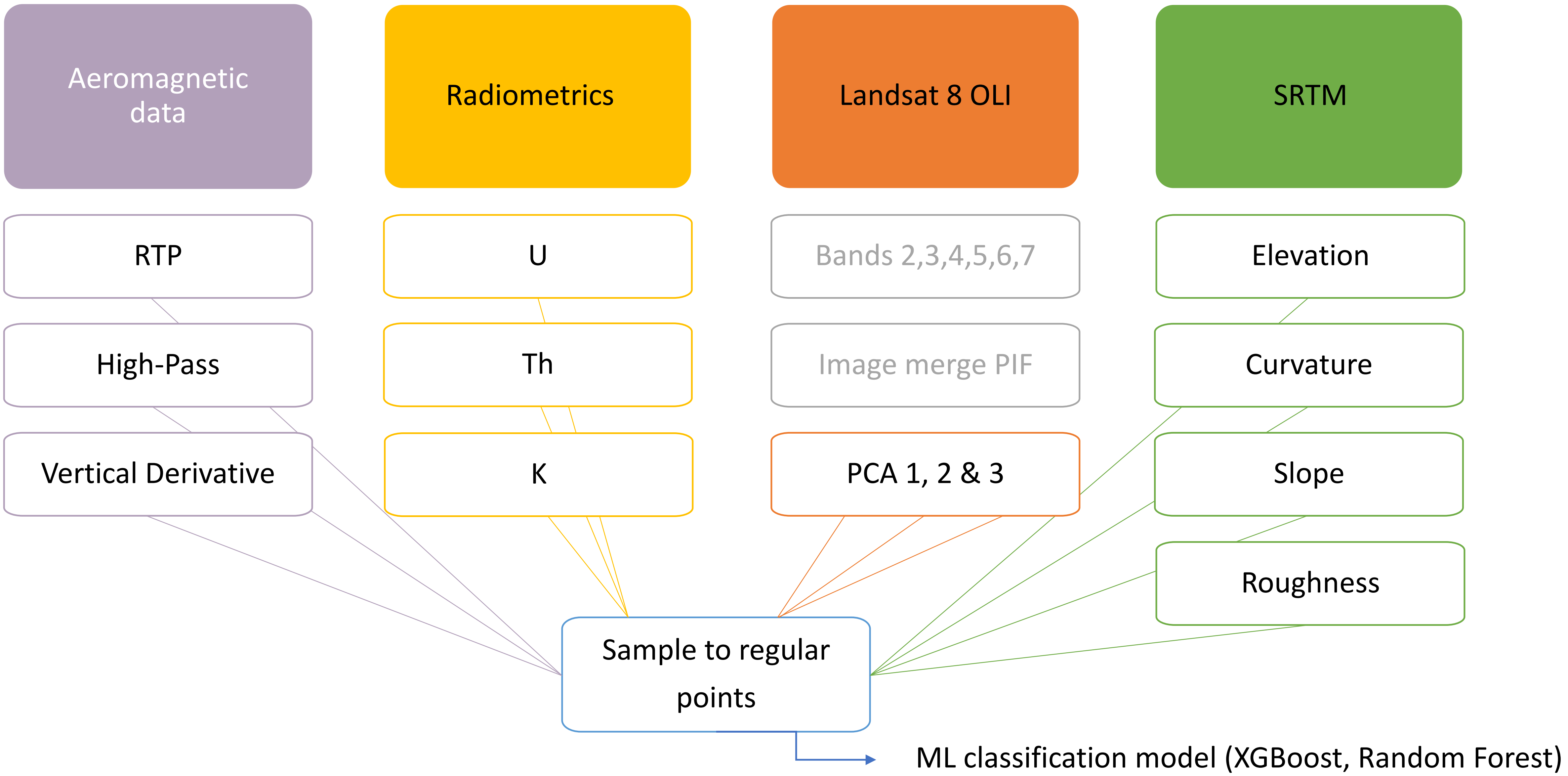


- Total model area 262,704 km²
- 300,000,000 data points
- 11-15 layers of data

- 10 Landsat 8 OLI scenes
- SRTM
- Regional radiometrics
- Regional aeromagnetic data

AURORA MINERALS GROUP

Data workflow



Cloud computing for large models

This workflow requires us to process and analyse hundreds of 40-300 million point models. To run these models we employed 4 EC2 instances.

Instances are simple to spin up and can run any software (even dongle-based licences)

AWS EC2 Instance type

m4.4x large X 4

- 16 CPU
- 64 G RAM

Total run time

- 52 hours

Total costs

- 1.65 US per hour
- $343 US

m4.4x large

- Pre-processing
- Sampling of rasters

m4.4x large

m4.4x large

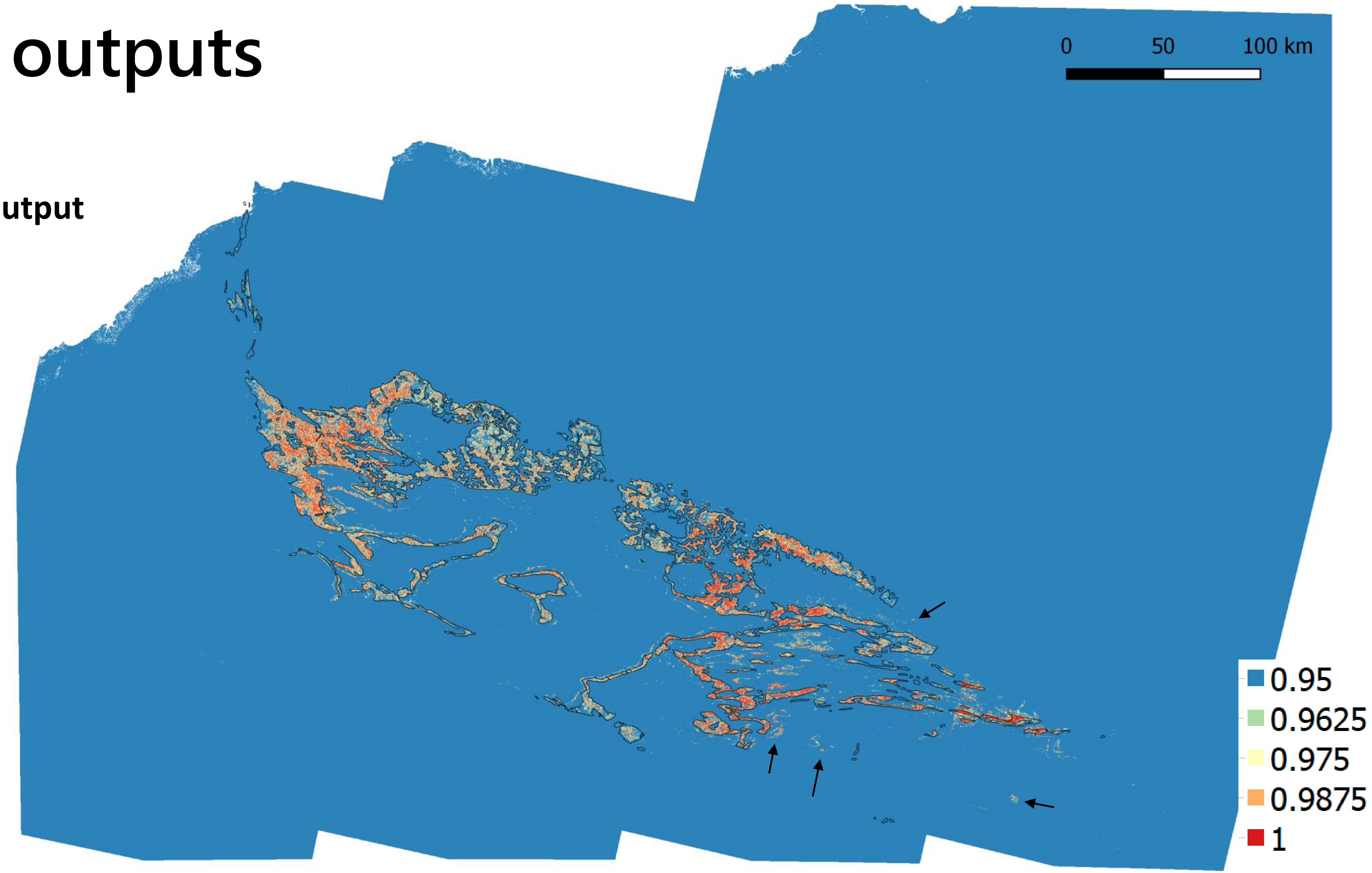
m4.4x large

- Machine learning classifier
- Variable importance
- Recombining data
- Raster creation

Example outputs

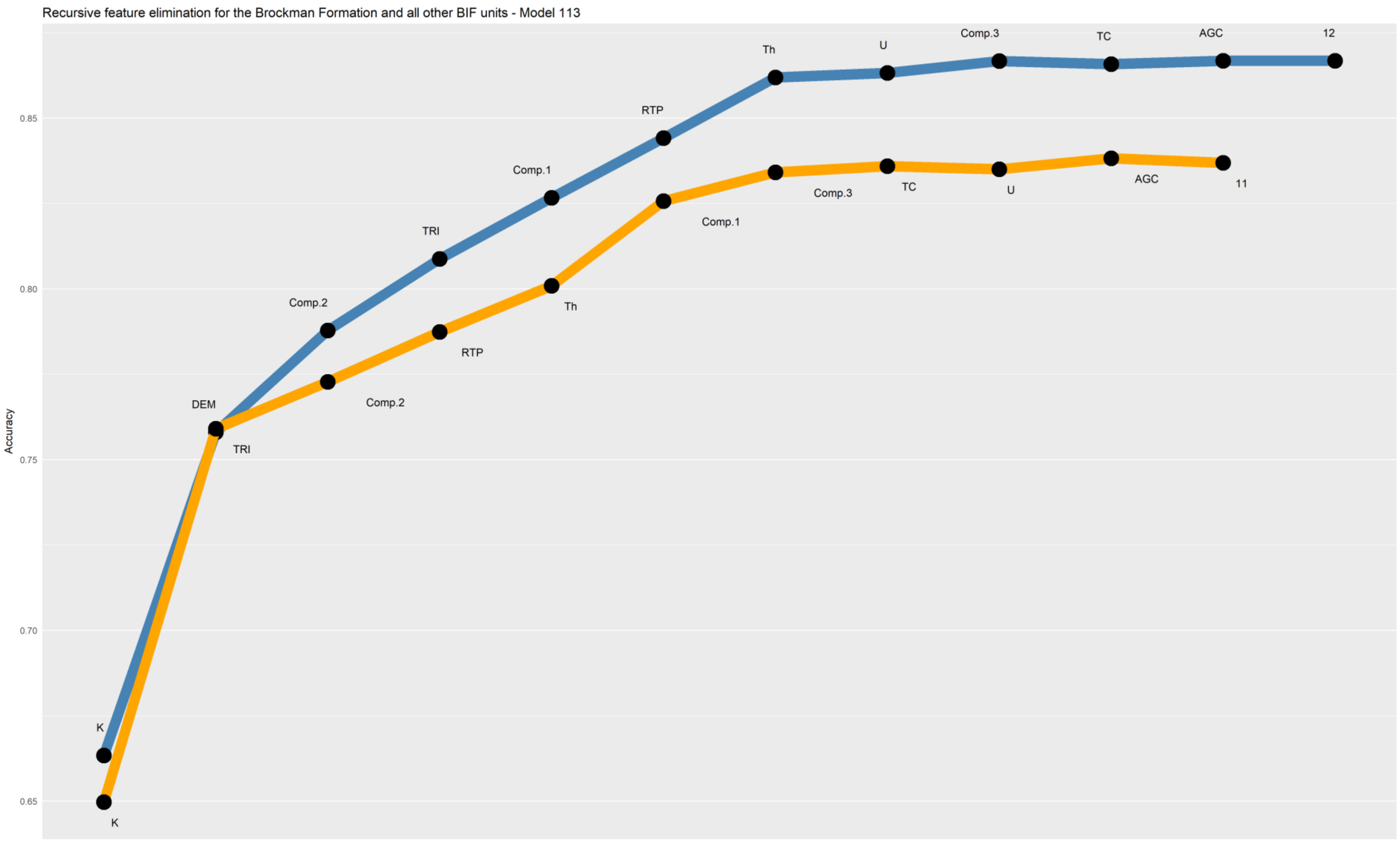
Probability raster output

The probability values in each pixel are the average of hundreds of Random Forest and XGBoost models that were made with different parameters and training sizes.



Probability grid for the Brockman Formation coloured with probability ranging from 0.95 (blue) to 1 (red) overlain with mapped extent (black)

Determining which variables are important



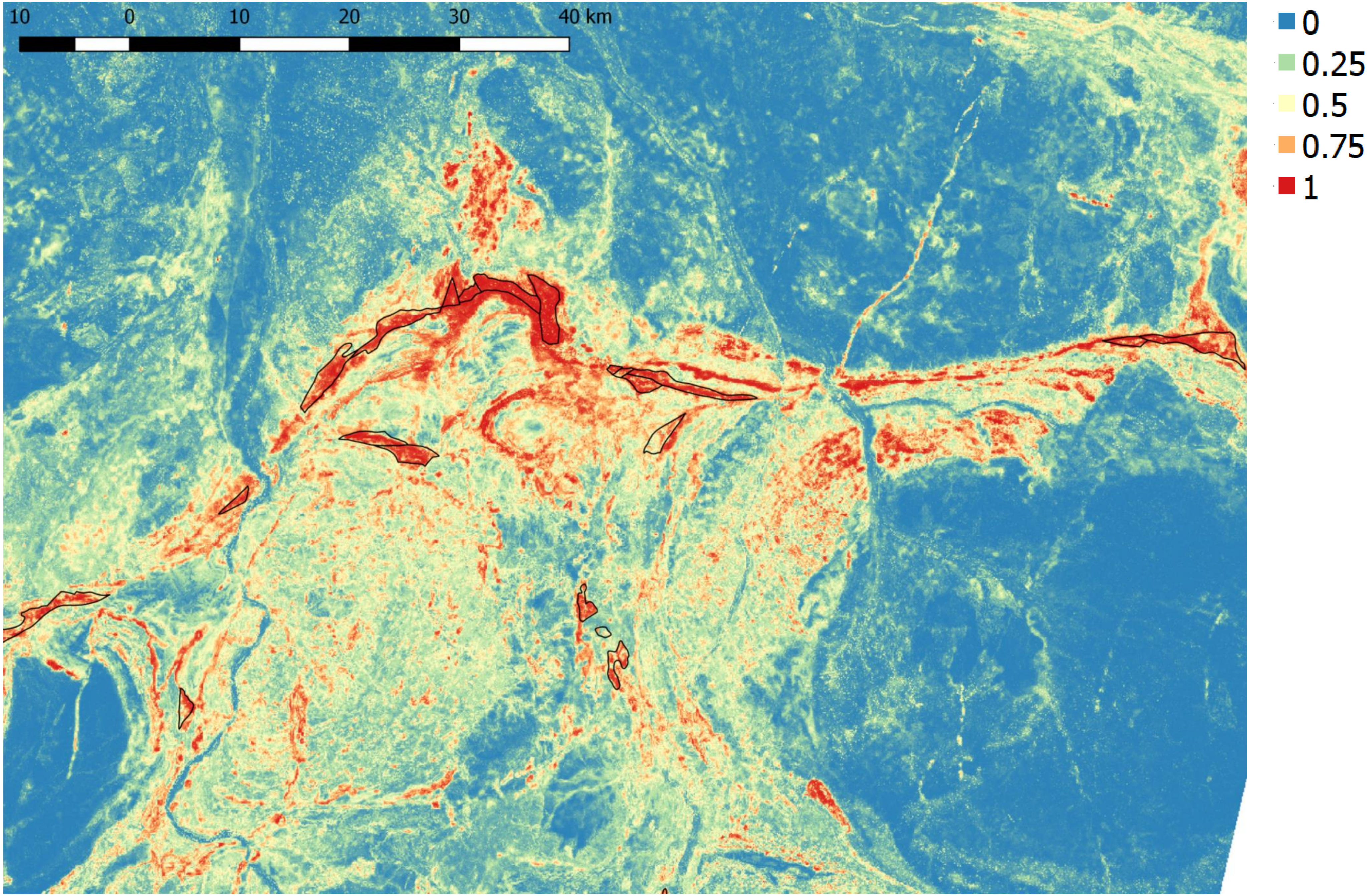
Recursive feature elimination

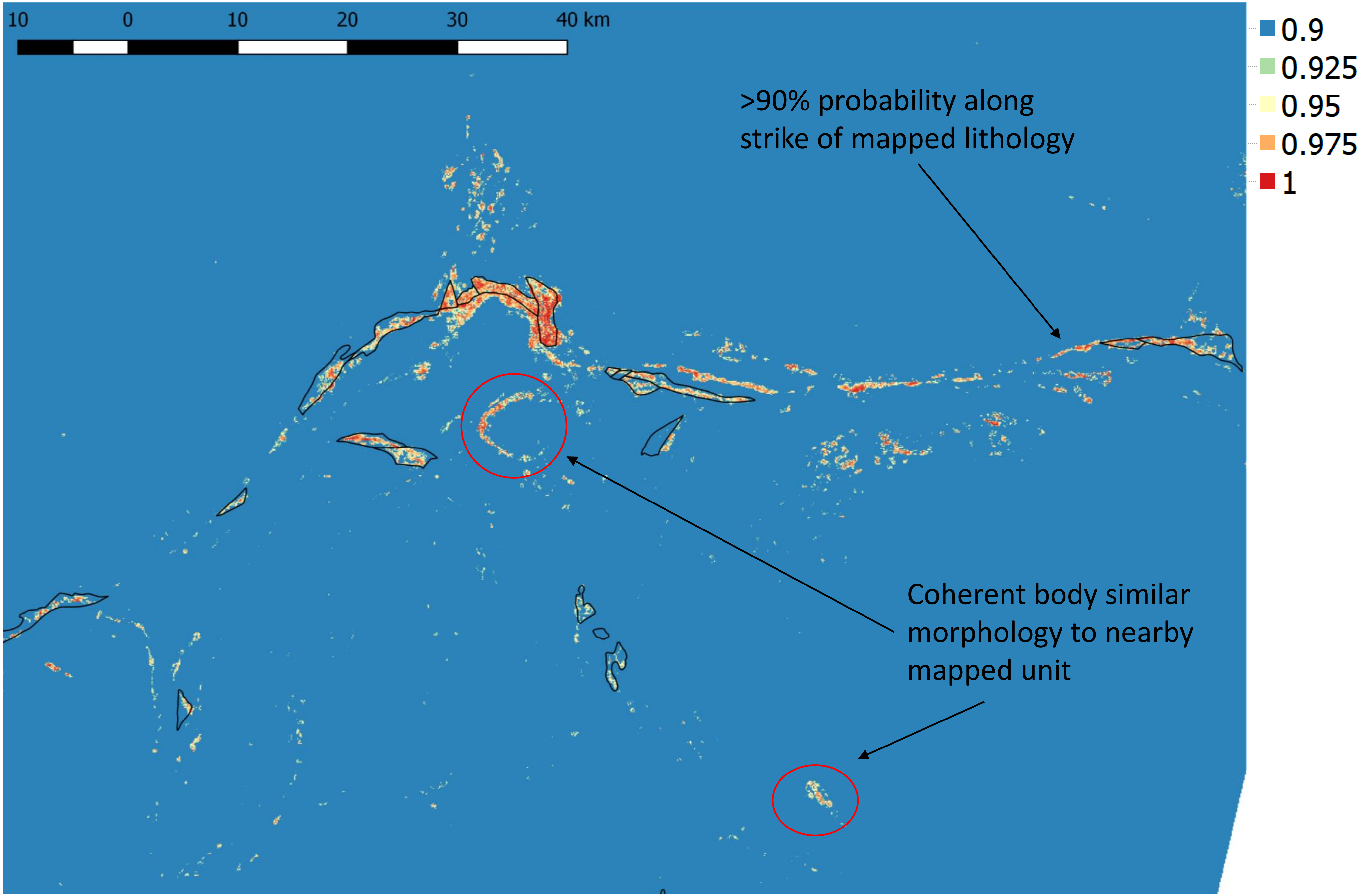
Brockman versus all other iron-bearing units

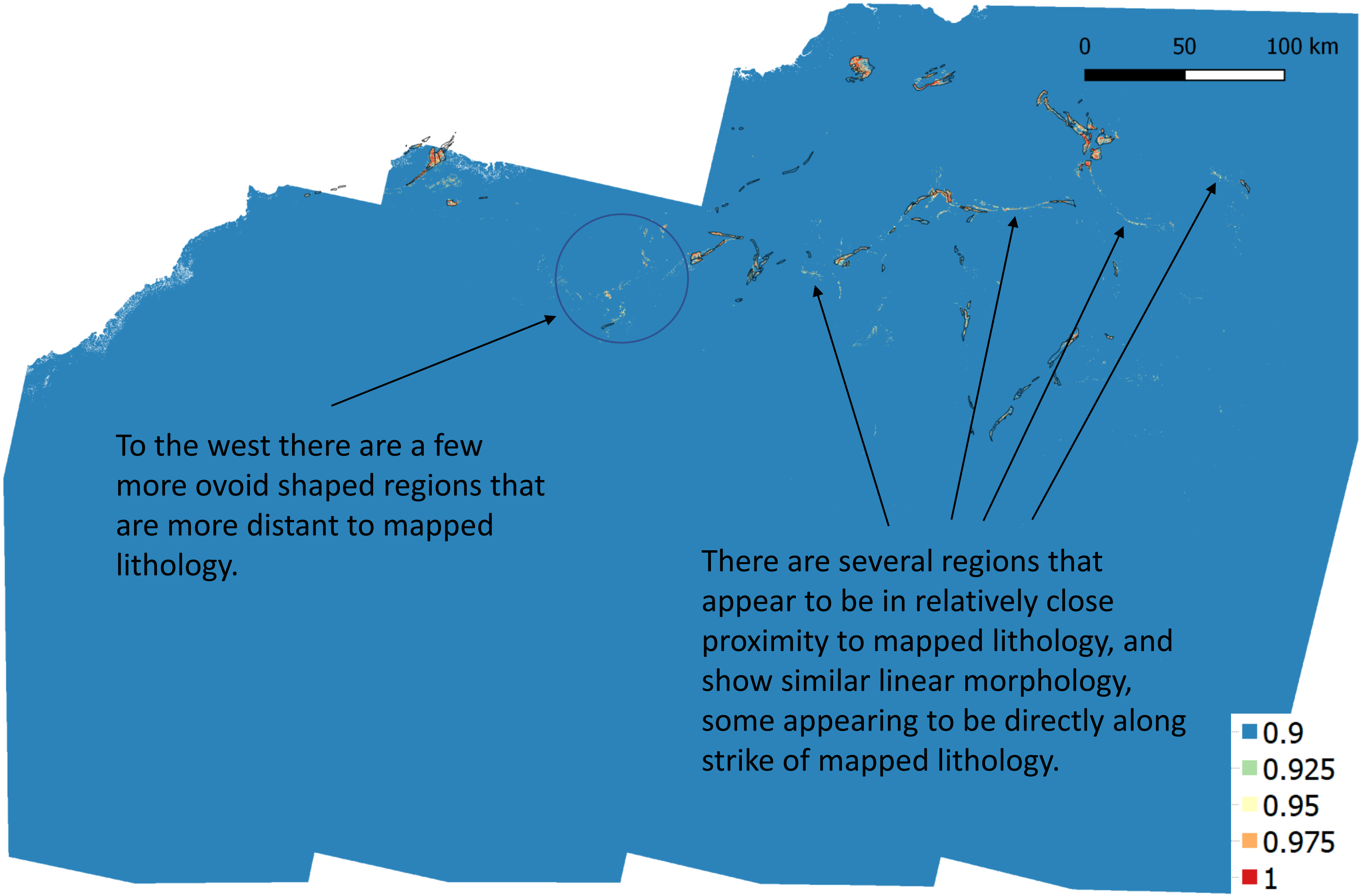
RFE analysis involves building the model recursively, each time looking at model performance, iteratively leaving out the poorest performing variables.

RFE helps understand which variables may be redundant or irrelevant.

RFE also informs on the optimal order in which variables should be used.

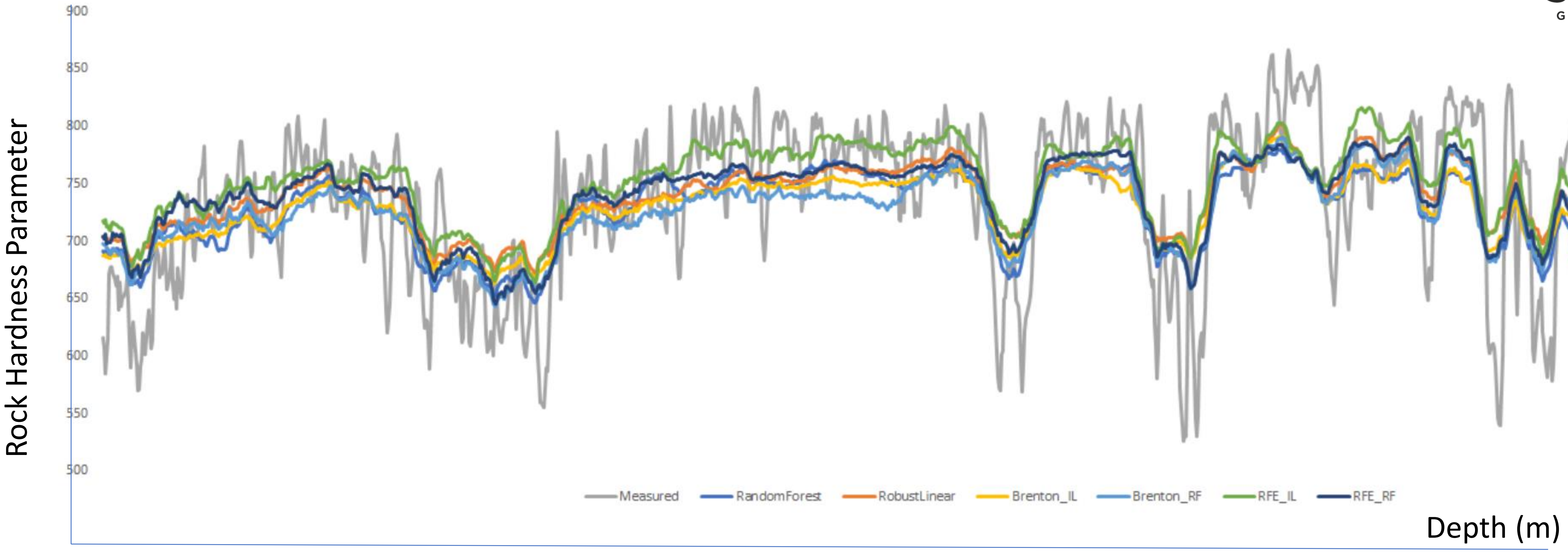

10
0
10
20
30
40 km
0
0.25
0.5
0.75
1


10
0
10
20
30
40 km
0.9
0.925
0.95
0.975
1
>90% probability along strike of mapped lithology
Coherent body similar morphology to nearby mapped unit


0
50
100 km
To the west there are a few more ovoid shaped regions that are more distant to mapped lithology.
There are several regions that appear to be in relatively close proximity to mapped lithology, and show similar linear morphology, some appearing to be directly along strike of mapped lithology.
0.9
0.925
0.95
0.975
1

Example 2
Prediction of rock hardness from Corescan mineralogy

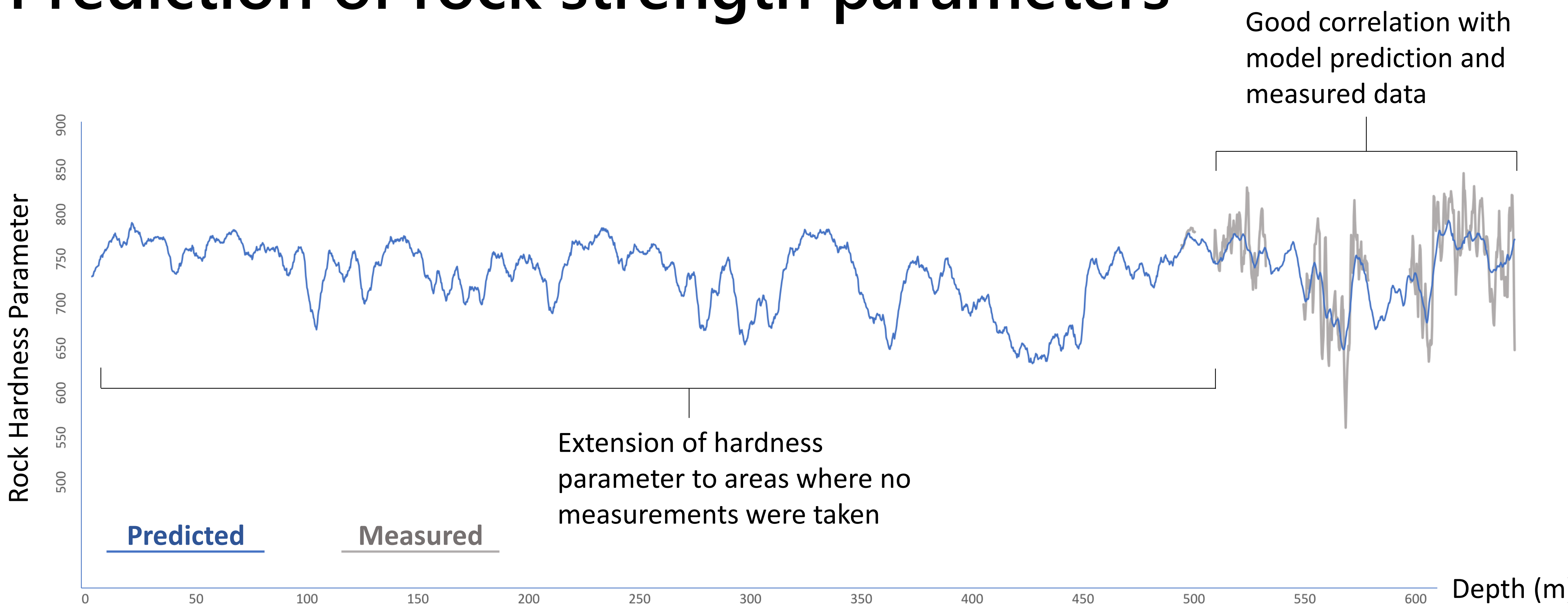
Prediction of rock strength parameters



The above graph shows a comparison between 7 different regression models (coloured lines) trained on Corescan mineralogy to predict rock hardness (grey line).

Corescan data may be used to predict datasets that are more expensive or suffer from long lead times.

Prediction of rock strength parameters



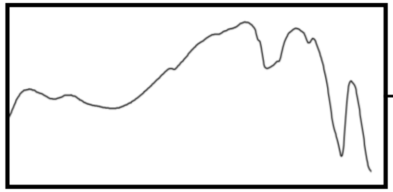
If a robust relationship between Corescan and other datasets can be identified, they can be predicted across areas where no measurements were taken.

Example 3

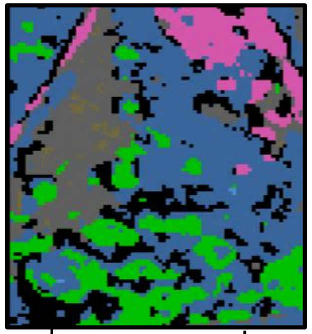
Data-driven domaining of Corescan mineral data using unsupervised learning

Upscaling and domaining Corescan data

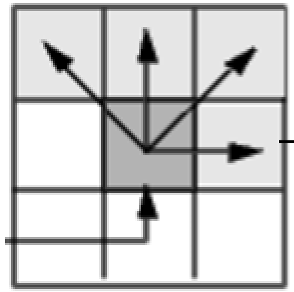
Spectra at 500µm



Mineral presence images



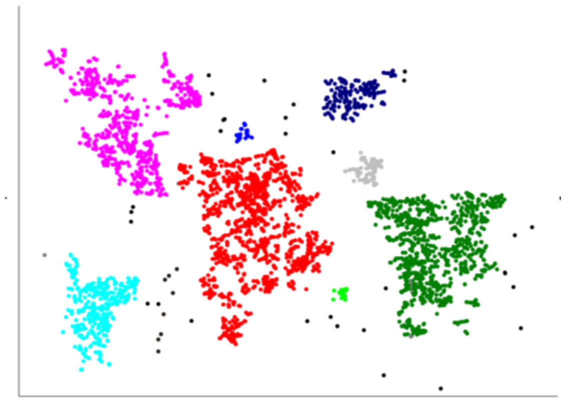
Mineral association



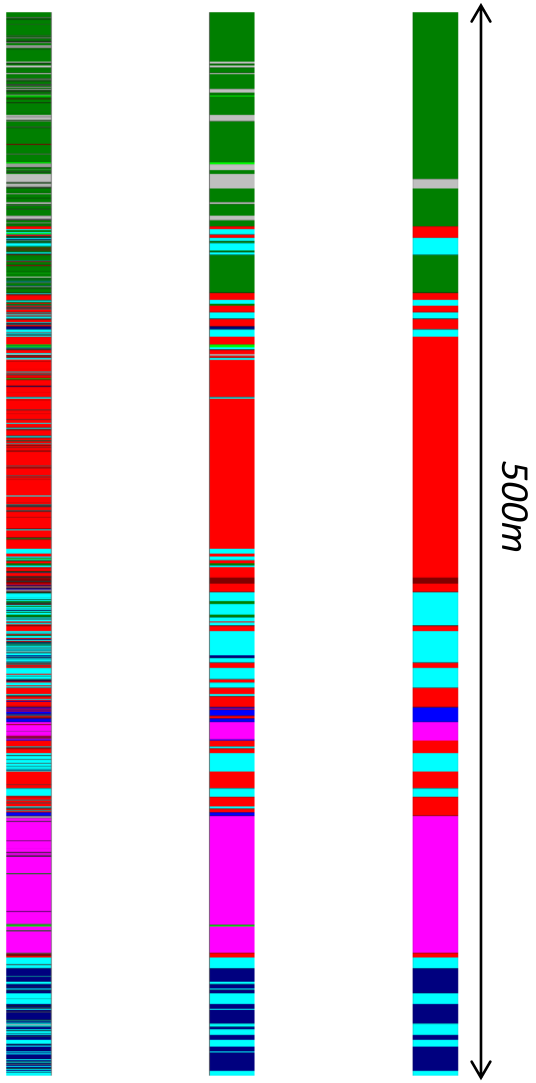
Mineral proportion

aspectral_	biotite_pxa	carbonate_pxa	chlorite_pxa	epidote_pxa
0.045	0	0.125	0.404	0.391
0.089	0.001	0.095	0.392	0.138
0.083	0.002	0.155	0.409	0.278
0.022	0.018	0.215	0.403	0.452
0.002	0.002	0.203	0.536	0.56
0.027	0.004	0.263	0.516	0.3
0.036	0	0.046	0.113	0.761

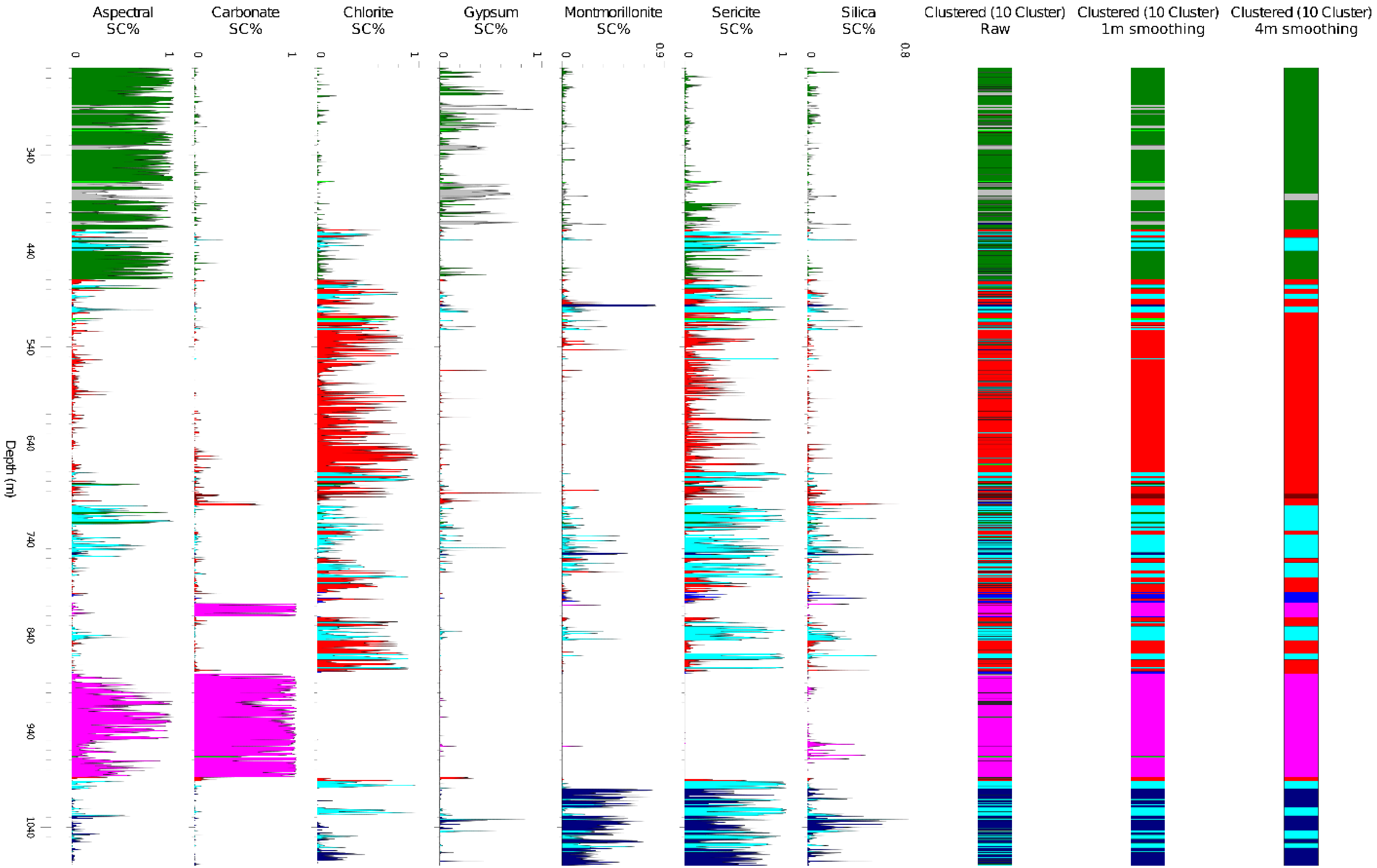
Plotting of similarity metric in low dimensional space for clustering

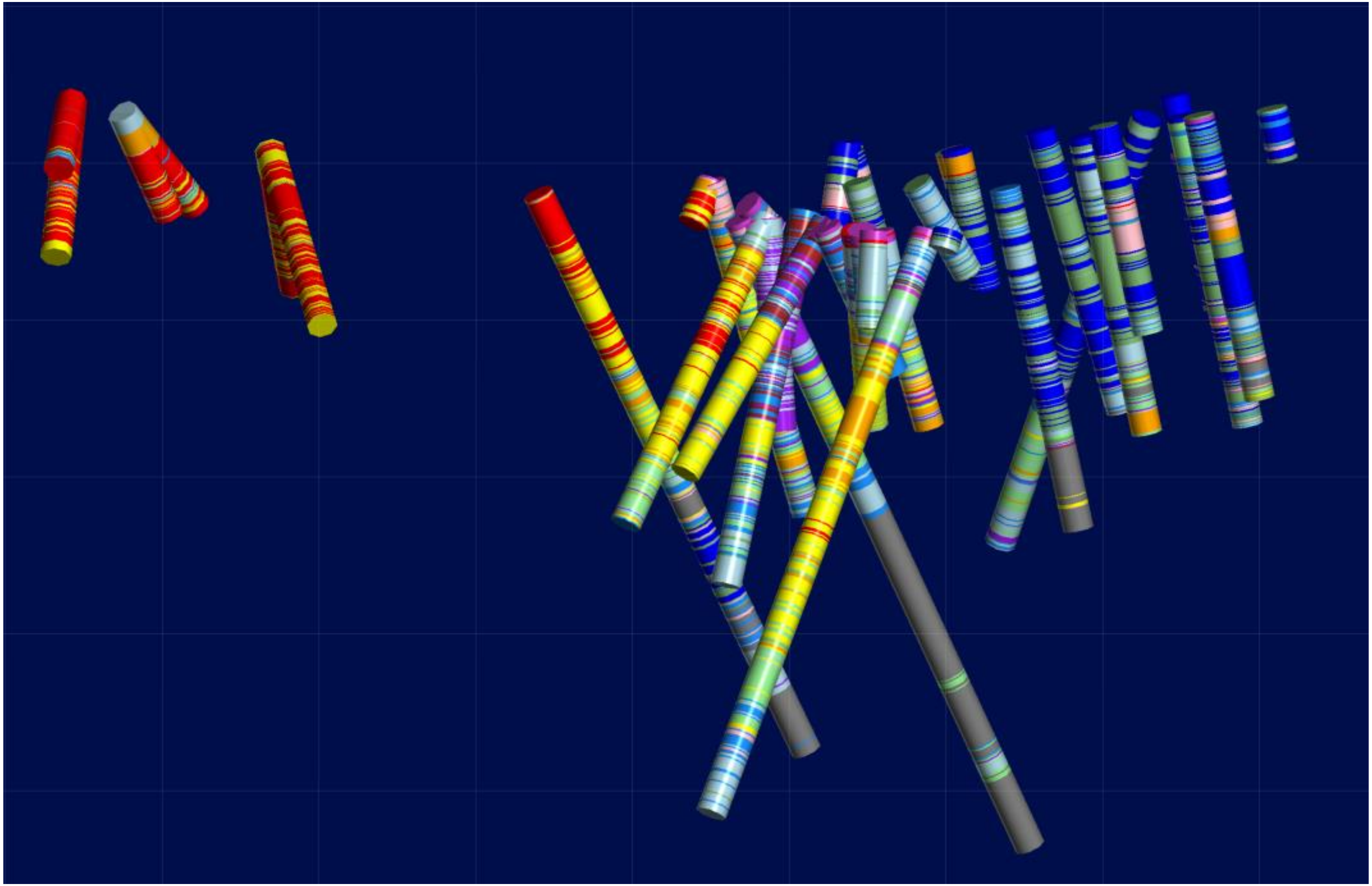


Clusters are smoothed to desired level of detail



corescan


Aspectral SC%
Carbonate SC%
Chlorite SC%
Gypsum SC%
Montmorillonite SC%
Sericite SC%
Silica SC%
Clustered (10 Cluster) Raw
Clustered (10 Cluster) 1m smoothing
Clustered (10 Cluster) 4m smoothing
0
1
0.9
0.8
340
440
540
640
740
840
940
1040
Depth (m)



Example 4

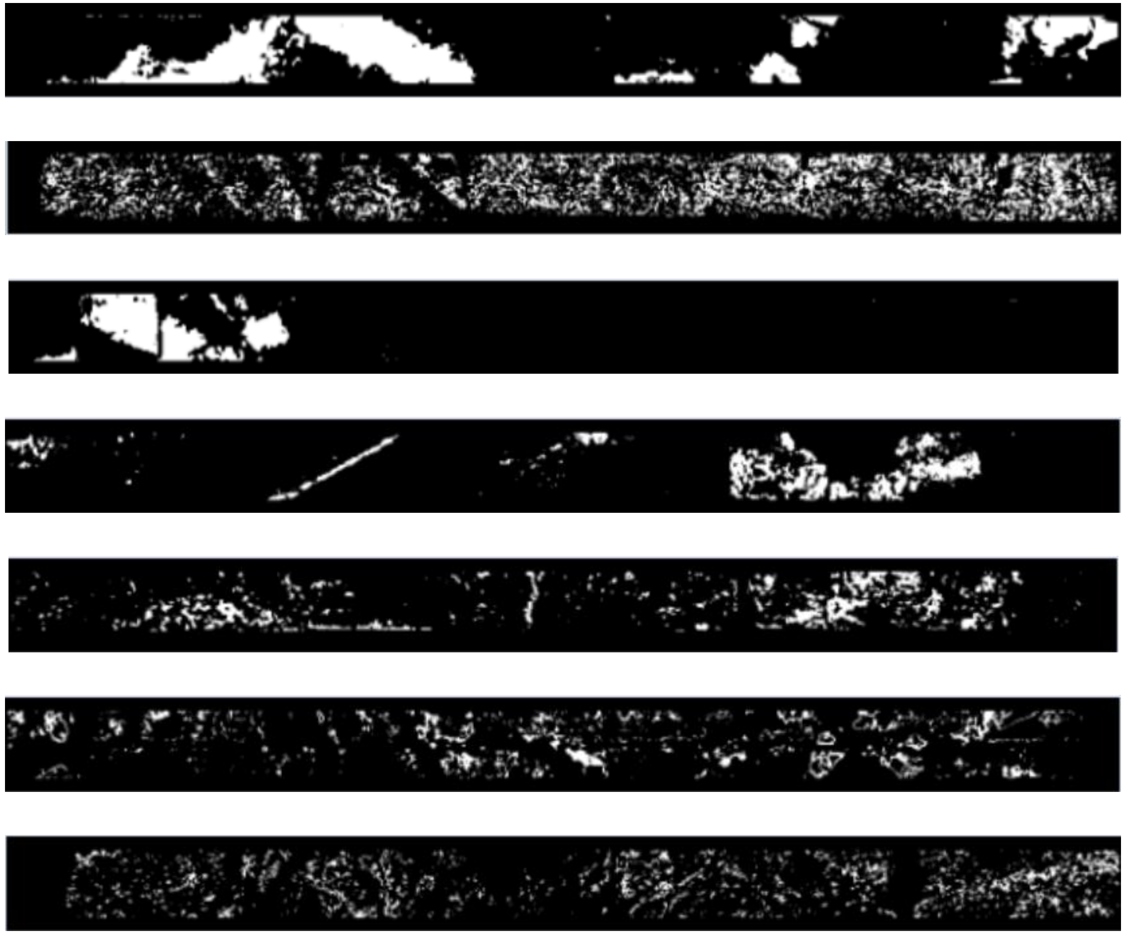
Classifying Corescan mineral textures

Extraction of texture parameters from Corescan mineralogy maps

By looking at the statistics of how pixels are connected and spatially distributed, it is possible to extract some statistical measures of mineral texture from the Corescan mineral maps.

These statistics can then be used to classify mineral texture in both supervised classification, and clustering applications.

On the next slide we show an example of how these texture parameters can be combined with the abundance data to produce texture clusters downhole.

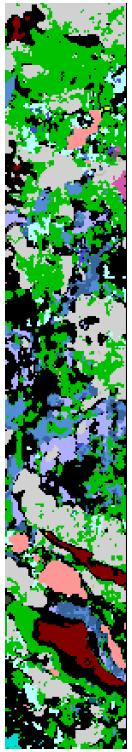


Example individual mineral images showing the diversity of texture collected by the Corescan system.

Extraction of texture parameters from Corescan mineralogy maps

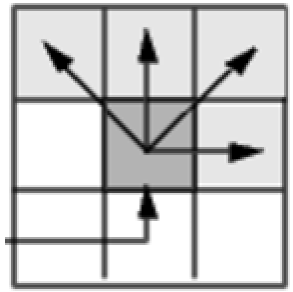
Input data

Individual Mineral Maps



Texture algorithm

Mineral Abundance GLCM statistics



The algorithm cycles through each pixel and looks at how it is connected to the pixels surrounding it in several directions.

Output variables

Mineral Correlation

Describes how connected the mineral texture is. Veins have high correlation values as they have pixels touching in a particular direction.

Mineral Complexity

Describes the complexity of the mineral texture. Disseminated or matrix textures are more complicated than massive textures.

Texture direction/strength

Describes how dominant a particular texture direction is, and returns the direction with respect to core axis

Mineral abundance

Counts pixels where mineral is present.

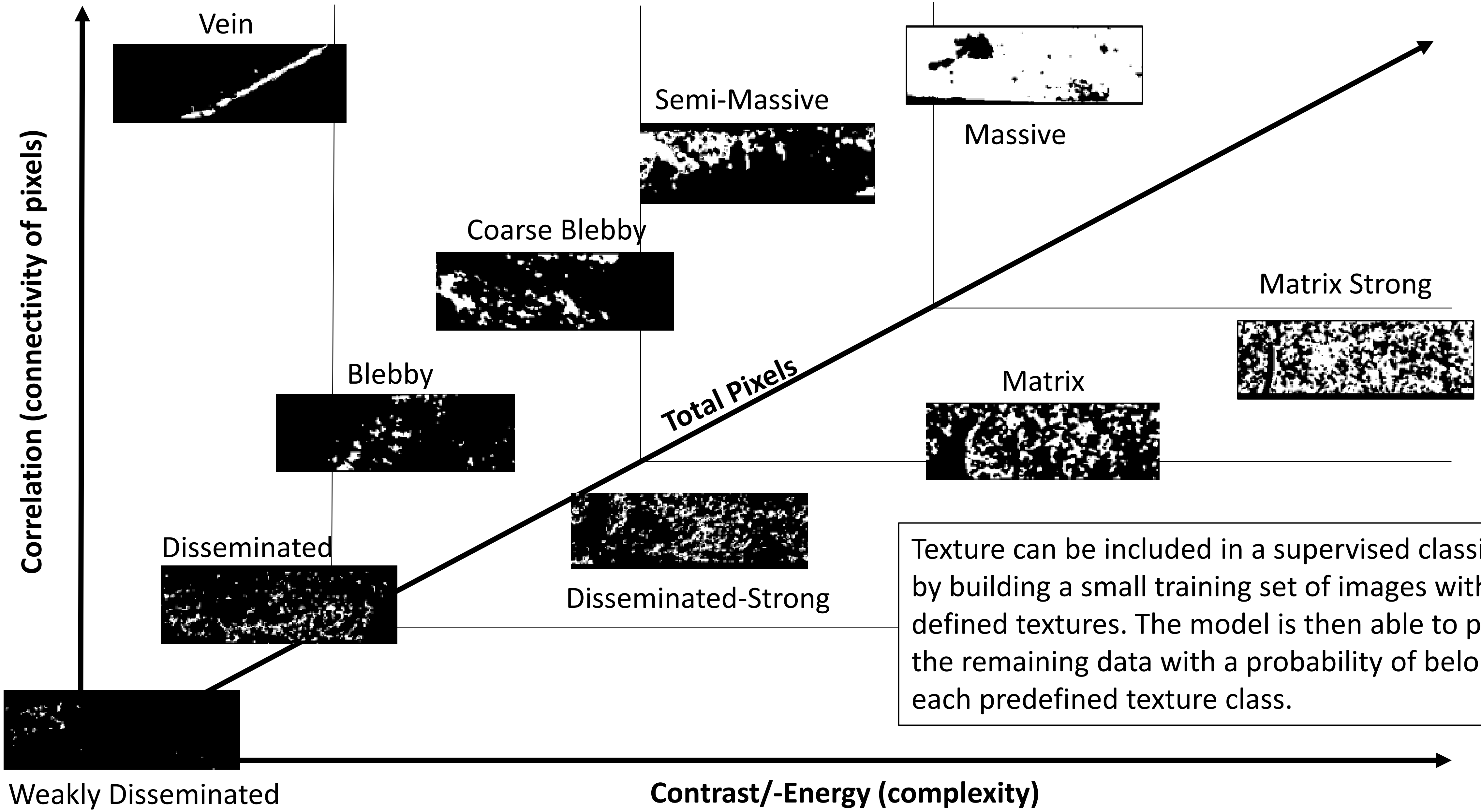
Unsupervised texture clustering

Texture parameters can be used as inputs into data driven clustering

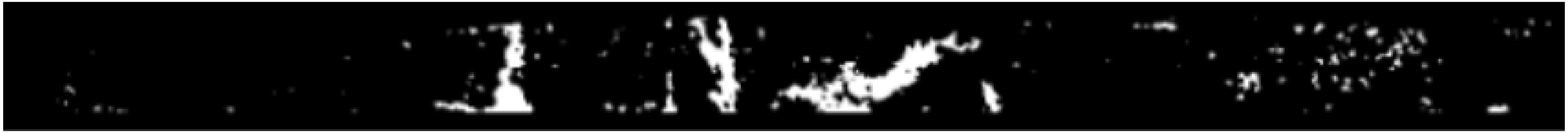
Supervised texture classification

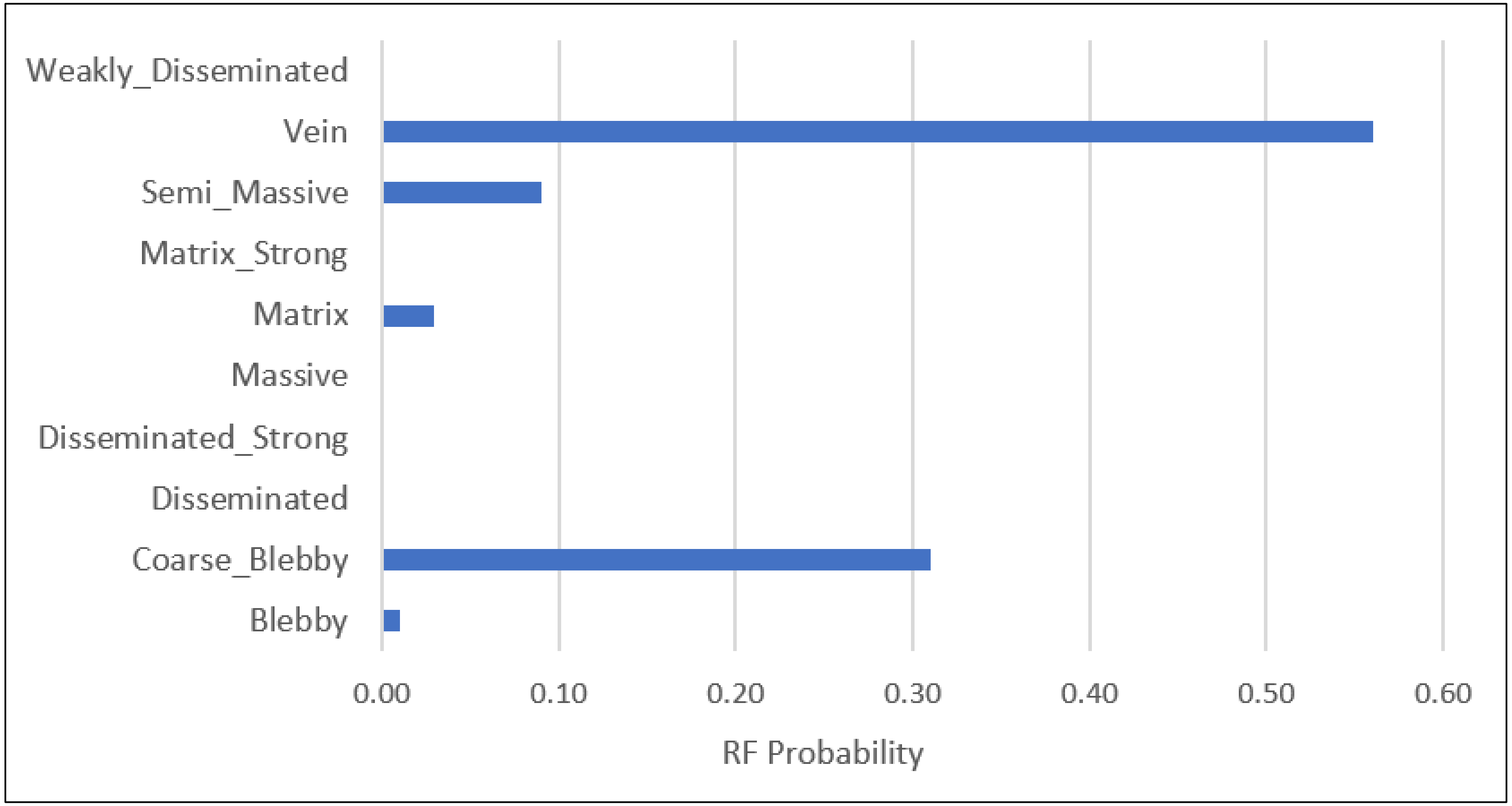
Textures of interest can be identified by the geologist and be fed into a supervised classification model (shown next)

Texture Classes



Texture can be included in a supervised classification by building a small training set of images with well defined textures. The model is then able to predict on the remaining data with a probability of belonging to each predefined texture class.

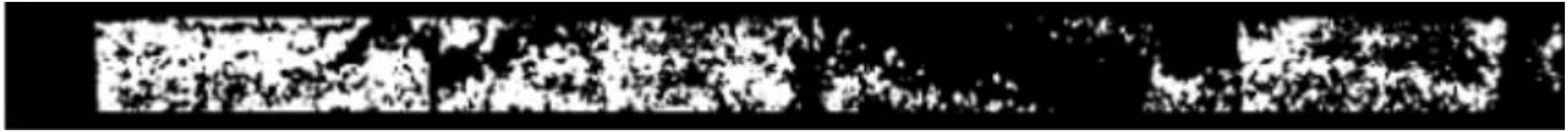


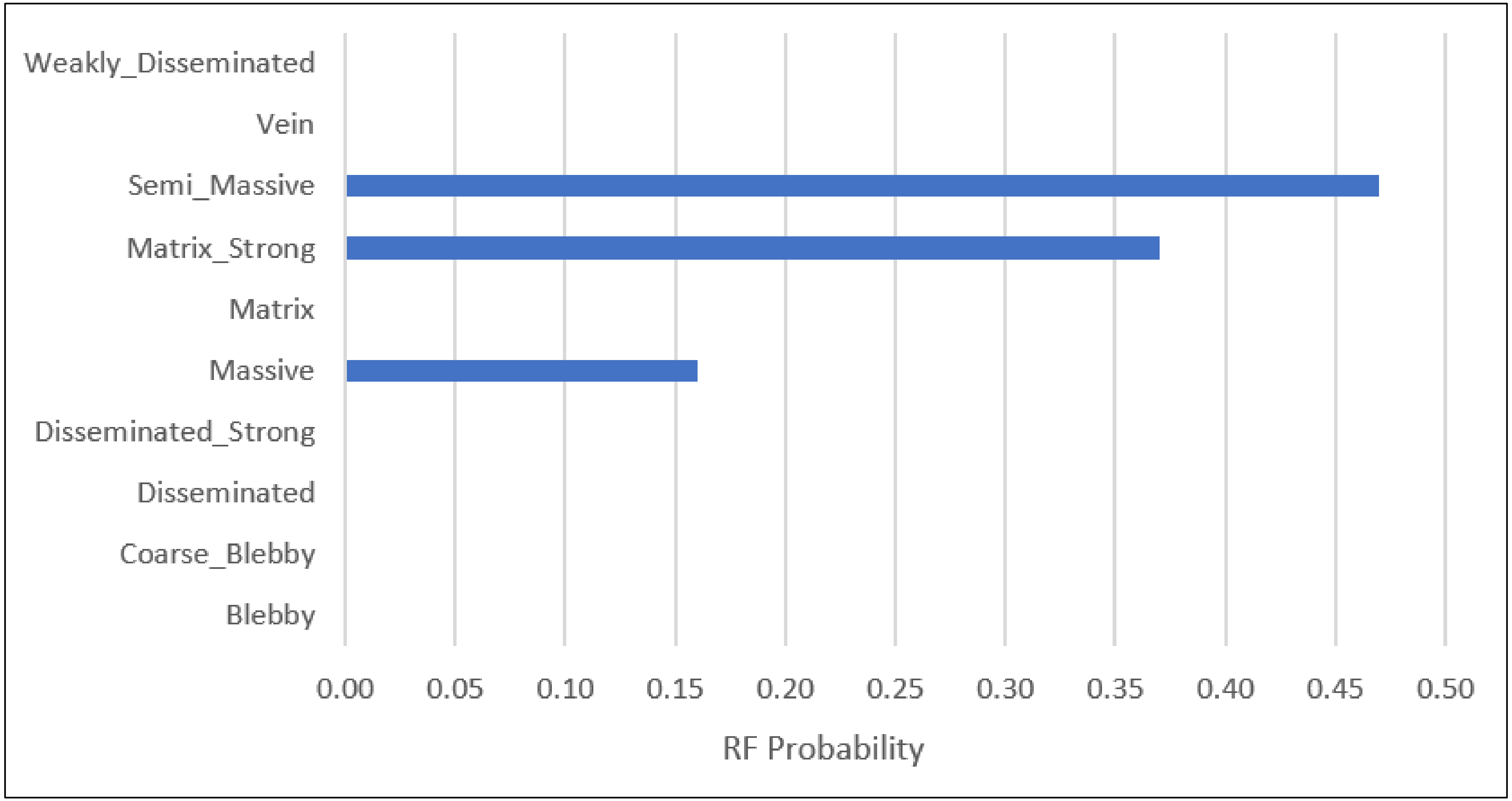


Example of outputs from the supervised texture classification.

The above image contains elements of several different end member textures including Vein (56%) and Coarse Blebby (31%), with smaller amounts of Semi-massive and Matrix.

The RF model allows for the image to display probabilities for several textural classes.

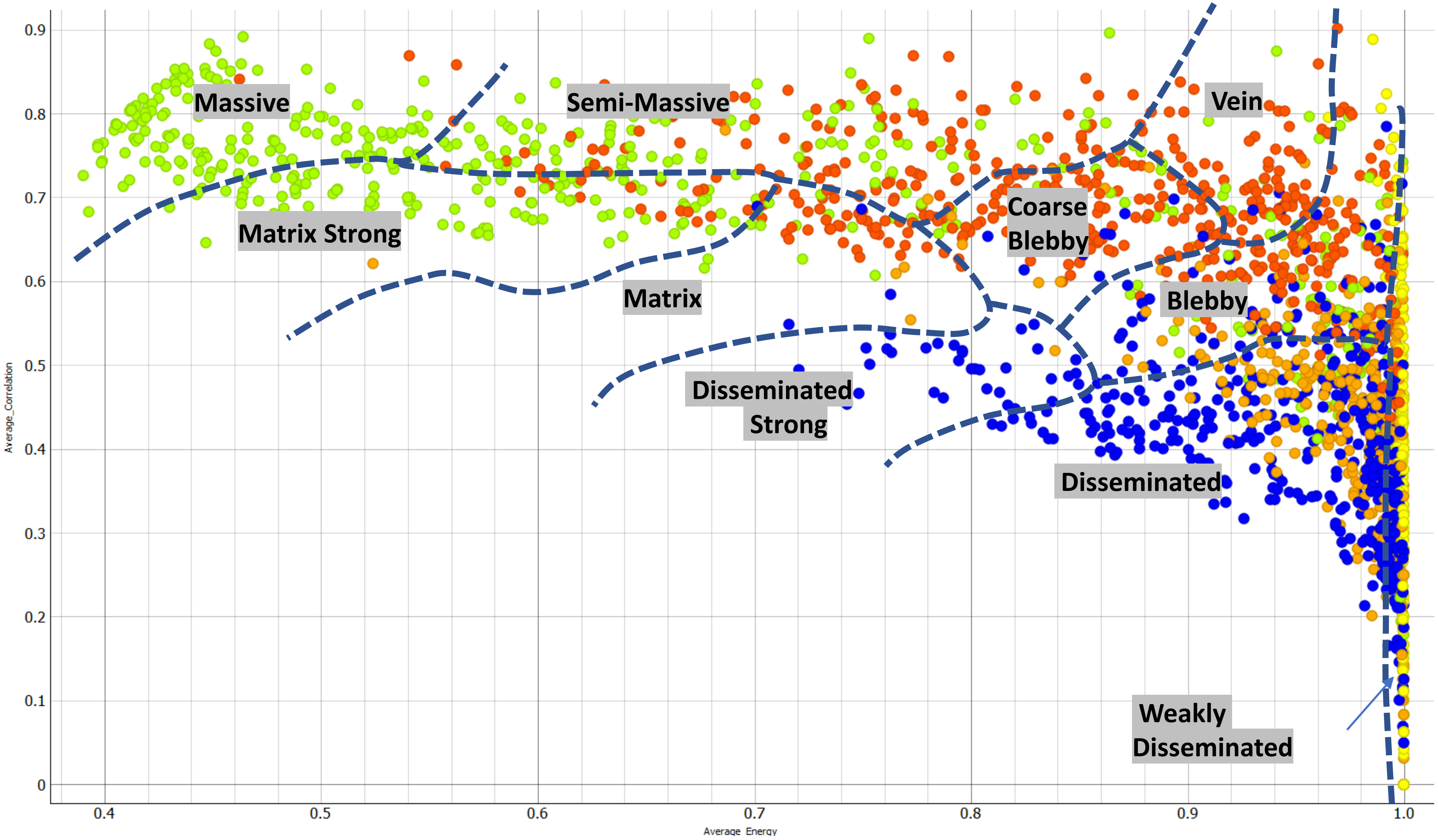




As most images contain more than one texture, giving an image a single class is simplistic and potentially misleading. The machine learning classification allows the image belong to several different texture classes.

This image contains elements of Semi Massive, Matrix Strong and Massive texture according to the classification model.

Individual minerals in the texture space



Points coloured by different mineral groups overlain with approximate boundaries with >50% probability of that texture existing