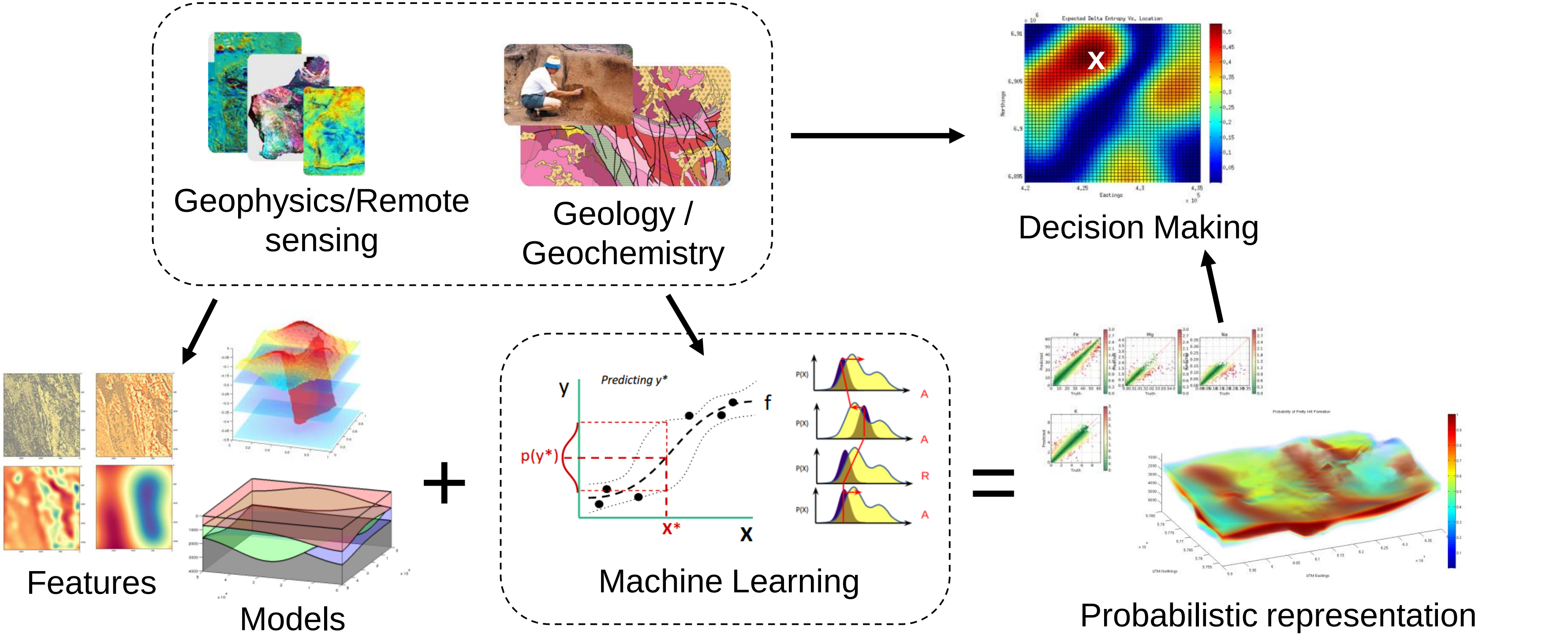
Data-Driven Mineral Exploration and Geological Mapping in the North West Mineral Province

Data61 & CSIRO Mineral Resources

David Cole, Lachlan McCalman, Vasek Metelka, Alexander Otto, Jess Robertson, Andrew Rodger, and Daniel Steinberg

www.csiro.au

Machine learning for geoscience applications



Machine learning

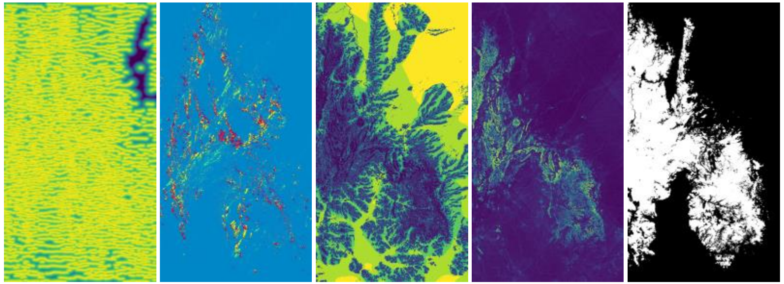
1. Data preparation
2. Feature extraction
3. Model selection
4. Train/test/validate
5. Prediction
6. Communication

- Gather, ingest, and clean data
- Data formats
- Missing data
- Detection limits
- Consistent units
- Measurement techniques
- Coordinate systems
- Consult experts

Machine learning

1. Data preparation
2. Feature extraction
3. Model selection
4. Train/test/validate
5. Prediction
6. Communication

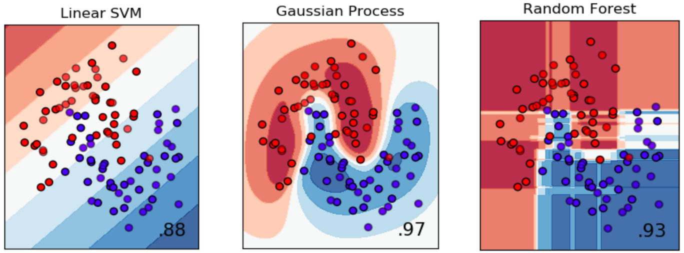
- What secondary features or interpretations are relevant?
- Transforms (e.g. wavelets), textures, distances (e.g. to faults)
- Dimensionality reduction



Machine learning

1. Data preparation
2. Feature extraction
3. Model selection
4. Train/test/validate
5. Prediction
6. Communication

- Classification, regression, unsupervised?
- Model structure

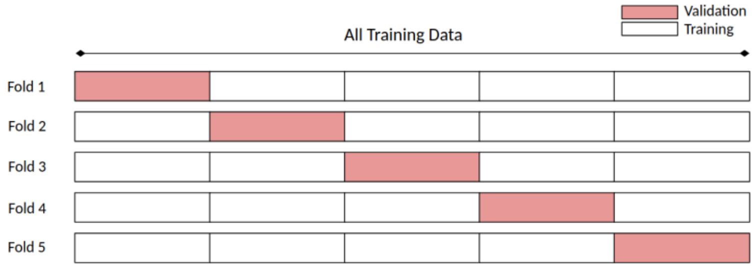


- Probabilistic?
- Hyperparameters
- Iterate as part of validation

Machine learning

1. Data preparation
2. Feature extraction
3. Model selection
4. Train/test/validate
5. Prediction
6. Communication

- Learn a generalisable model
- Split dataset into train and test/validation dataset

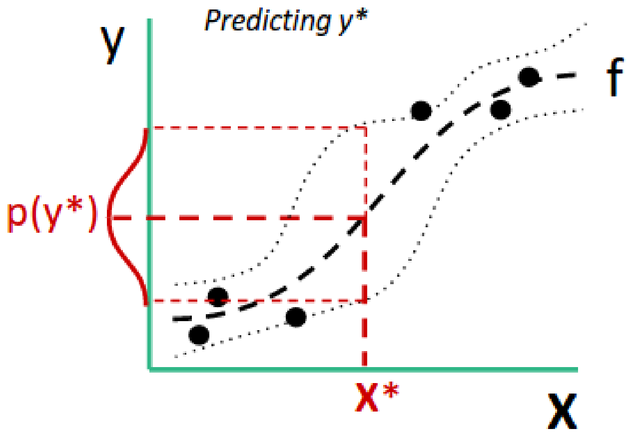


- Measure performance (e.g. MSE, R^2, precision, recall, f1, log loss)

Machine learning

1. Data preparation
2. Feature extraction
3. Model selection
4. Train/test/validate
5. Prediction
6. Communication

- Process query data
- Run model on unseen data

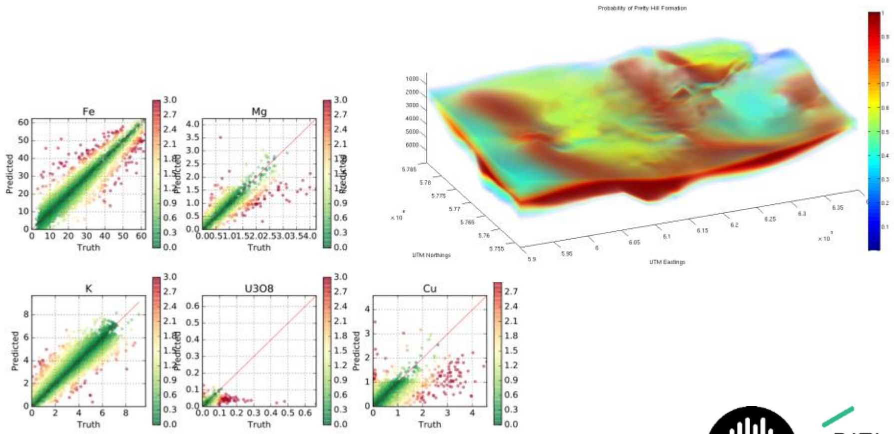


- Covariate shift?

Machine learning

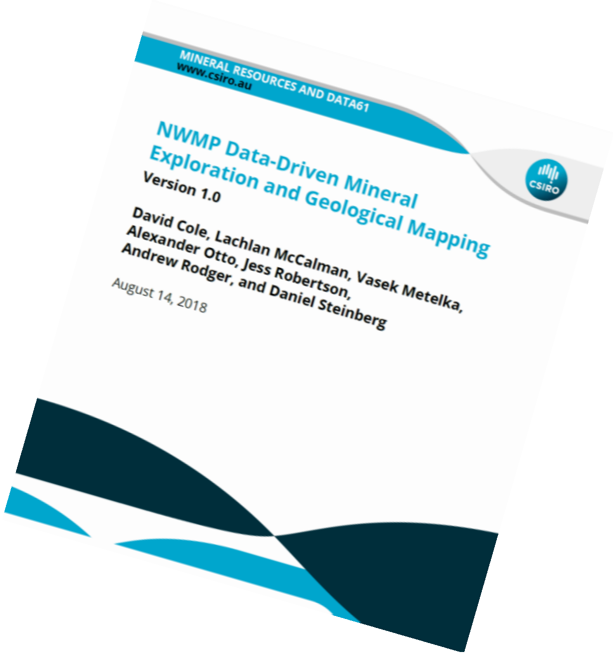
1. Data preparation
2. Feature extraction
3. Model selection
4. Train/test/validate
5. Prediction
6. Communication

- Model assumptions and limitations
- Visualisation
- Interpretation/insights



Applications of data-driven modelling

1. Mineral occurrence mapping
2. Automated geological classification
3. Interpreted geology anomaly detection
4. Modelling geochemistry


MINERAL RESOURCES AND DATA61
www.csiro.au

NWMP Data-Driven Mineral Exploration and Geological Mapping

Version 1.0

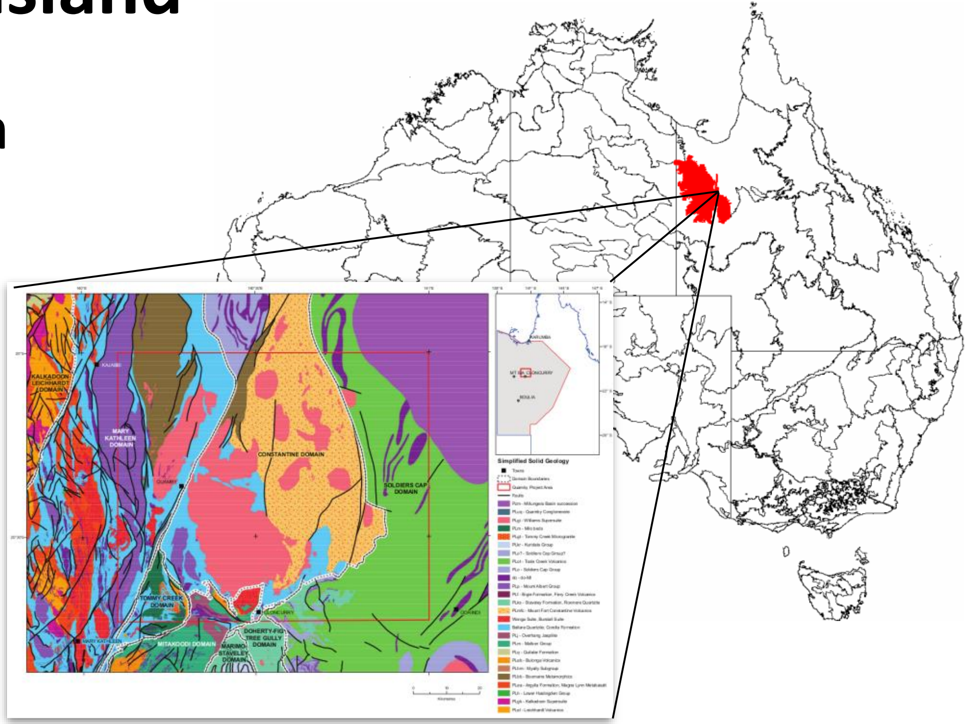
David Cole, Lachlan McCalman, Vasek Metelka, Alexander Otto, Jess Robertson, Andrew Rodger, and Daniel Steinberg

August 14, 2018

Mt Isa region of Queensland

- Existing mines and known mineral deposits
- E.g. Cu-Au, Ag-Pb-Zn, U

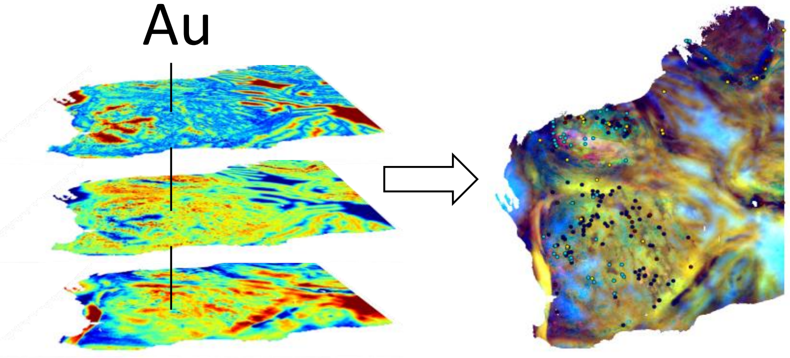
- Geochemical surveys
- Geophysics
- ASTER, DEM, etc.
- Interpreted geology



GREENWOOD, M.L. & DHNARAM, C.R., 2013: 3D mineral potential of the Quamby area. Queensland Minerals and Energy Review Series, Department of Natural Resources and Mines, Queensland.

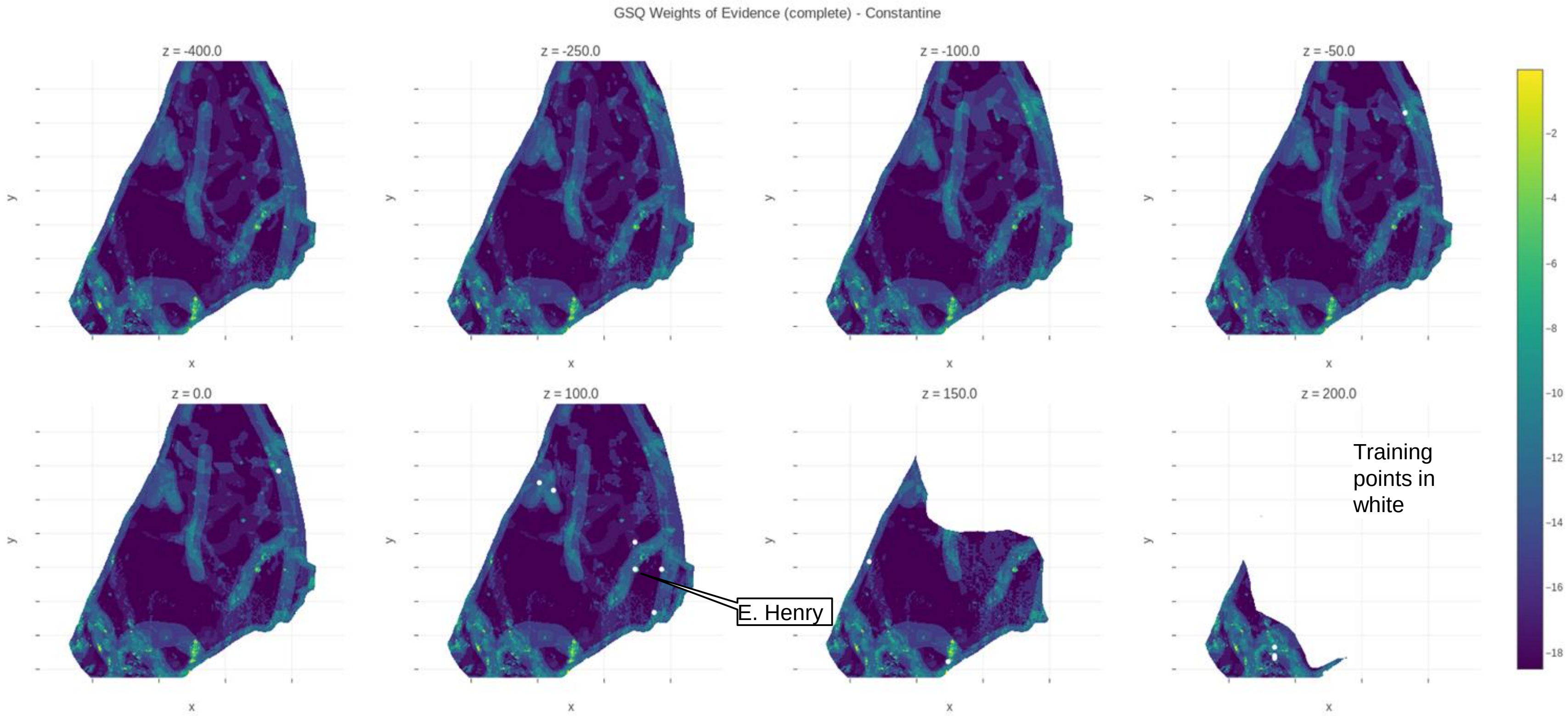
Mineral occurrence mapping

- Generate a map of mineral "prospectivity" using:
  - Geospatial feature data
  - Known deposit locations

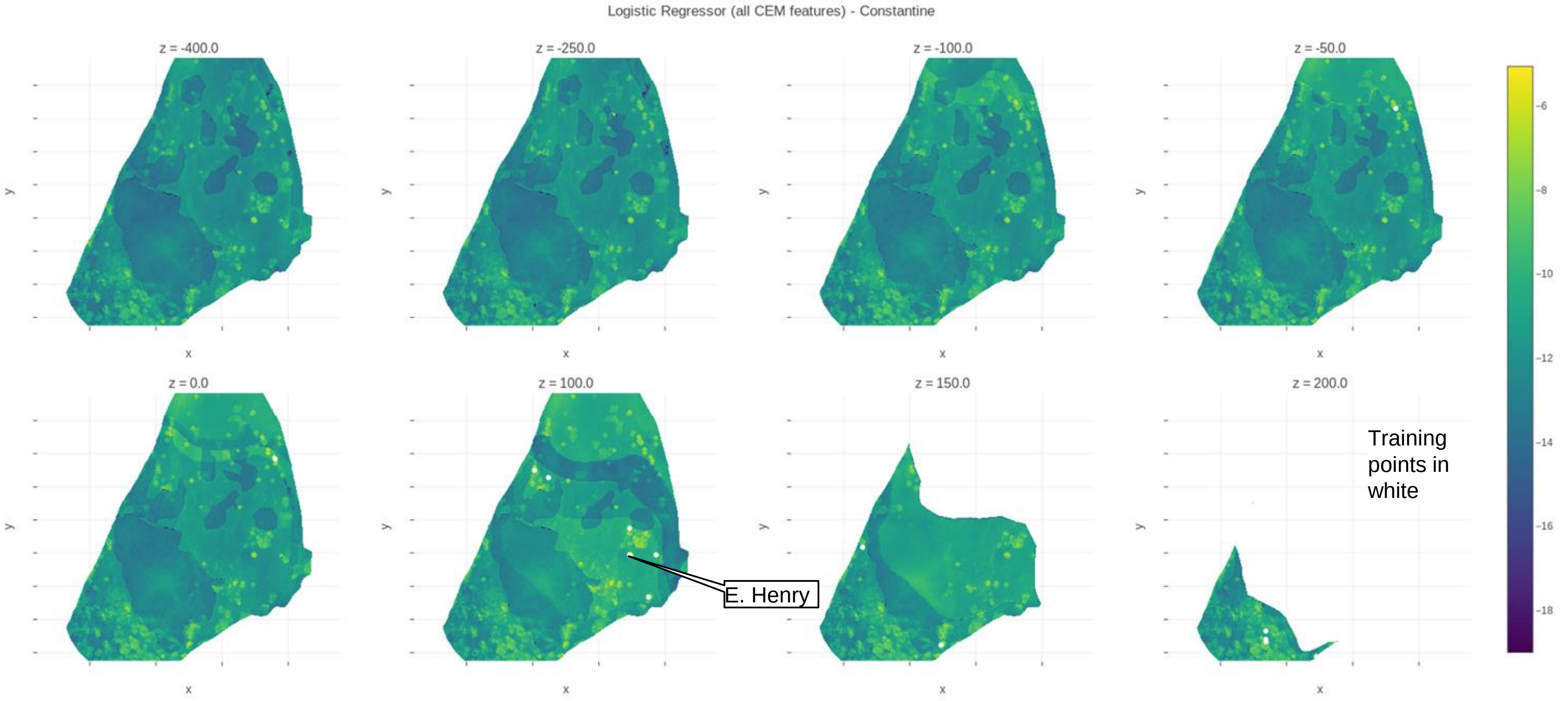


- *Weights of Evidence* vs *Machine learning*
  - Comparison with classic ML approach shows ML has less constraints, better cross-validation performance
  - Difficult problem - small and biased training data set

Mineral occurrence mapping



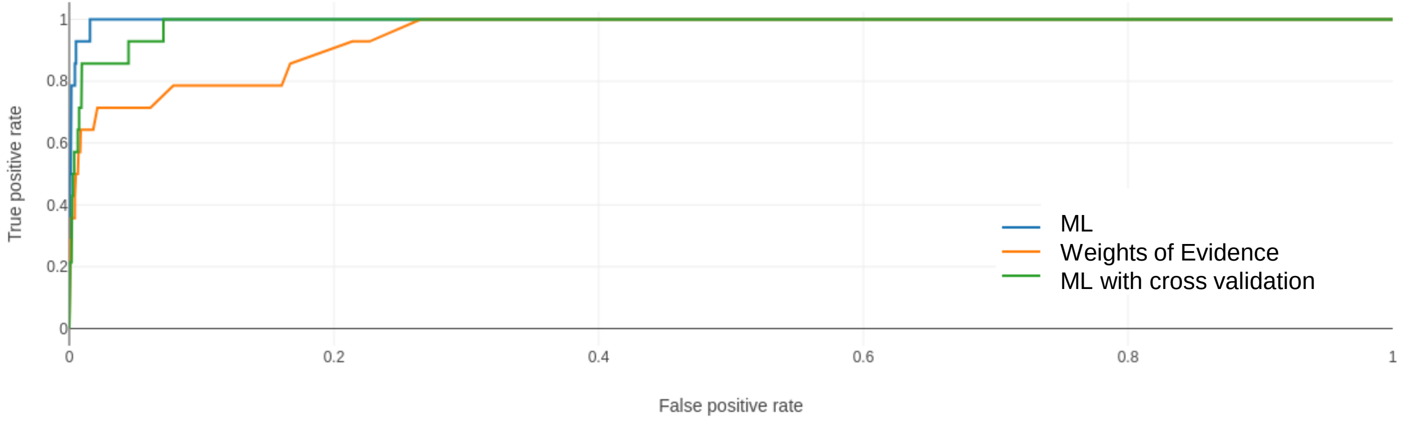
Mineral occurrence mapping



Mineral occurrence mapping

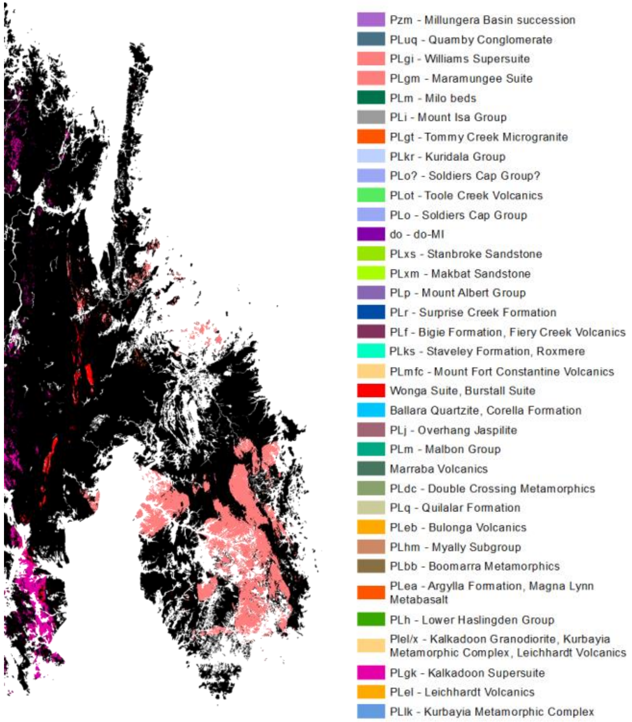
- Classifier performance
  - Log loss
  - ROC curve

classifier	log loss
logistic regressor (with CV)	0.000039
weights of evidence	0.000337

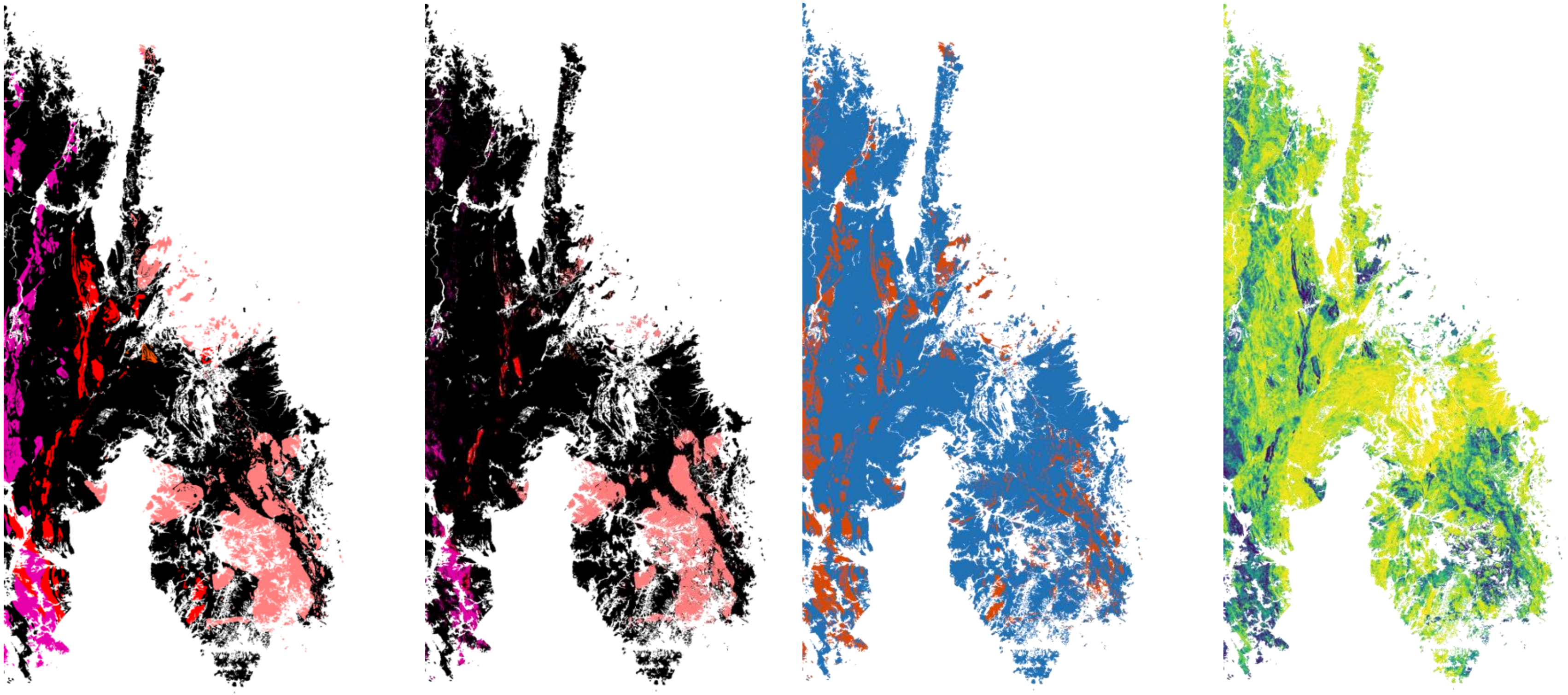


Classification of outcropping granite

- 6000 surface observations
  - 6 classes (5 granites + None)
- Surface + geophysical covariates
  - ASTER, DEM, MrVBF, Gamma, Gravity, Magnetics
- Random forest classifier



Classification of outcropping granite



Interpreted geology | Model prediction | Difference | Model probability

Classification of outcropping granite

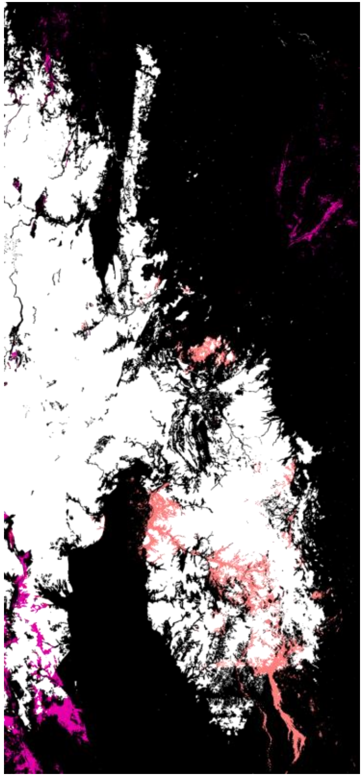
- Comparison to interpreted geology

label	precision	recall	f1-score	% of data points
None	0.86	0.98	0.92	76.6
PLgk – Kalkadoon Supersuite	0.69	0.21	0.32	6.9
PLgm – Maramungee Suite	1	0	0	0.2
PLgt – Tommy Creek Microgranite	0.91	0.08	0.14	0.1
Plgi – Williams Supersuite	0.86	0.74	0.8	11.5
PLgw – Wonga Suite, Burstal Suite	0.86	0.1	0.18	4.7

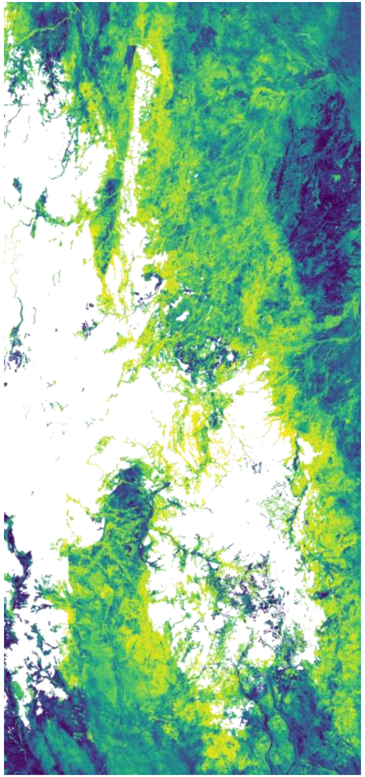
Classification of outcropping granite

- Predicting outcrops outside interpreted geology model

Interpreted geology



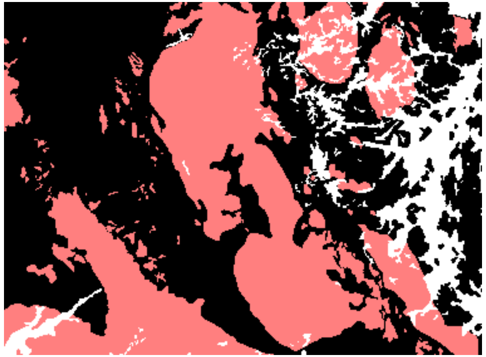
Model prediction



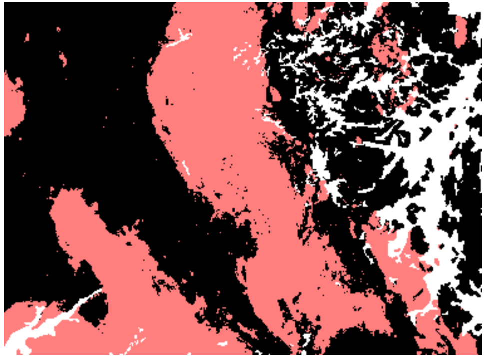
Model probability

Classification of outcropping granite

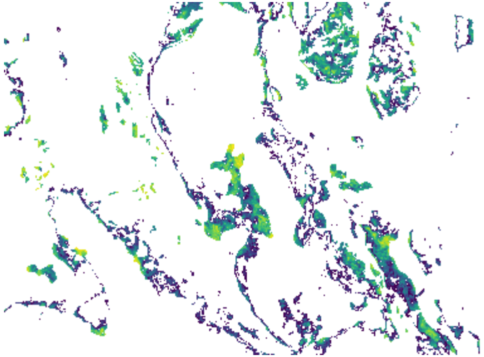
- Identify areas of disagreement with interpreted geology



Interpreted geology



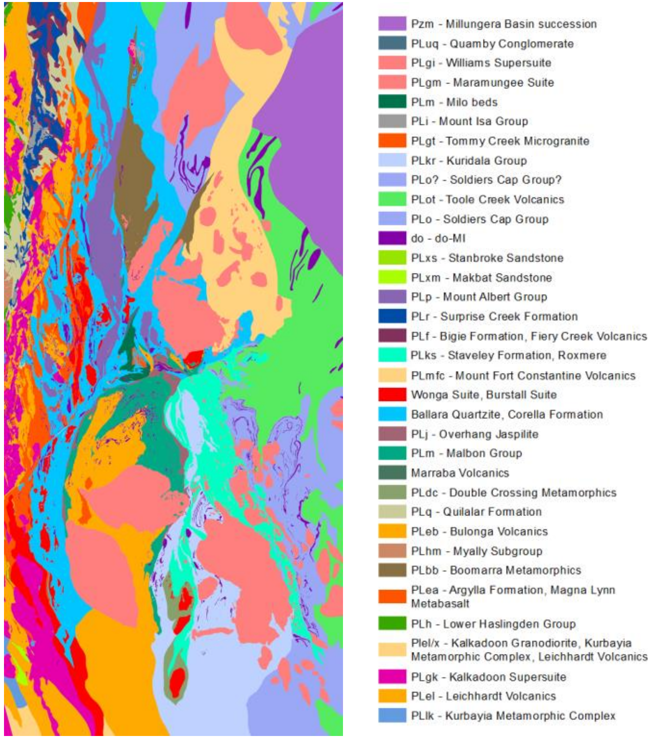
Model prediction



Model probability

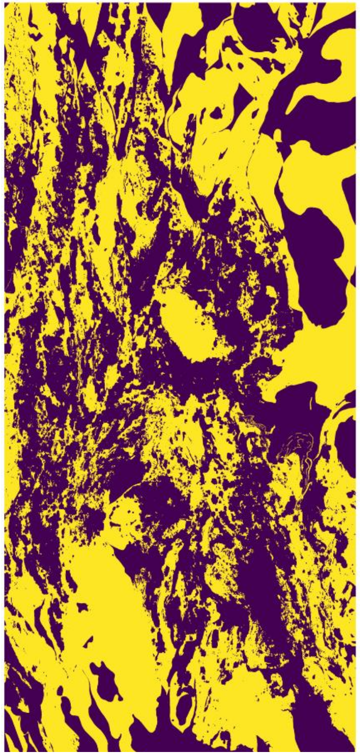
Anomaly detection of interpreted geology

- Interpreted geological mapping is an involved manual process
- Can we “audit” this by finding areas of potential disagreement
- Utilise anomaly detection algorithms such as ‘*One Class SVM*’
- Per-class model using geophysical datasets

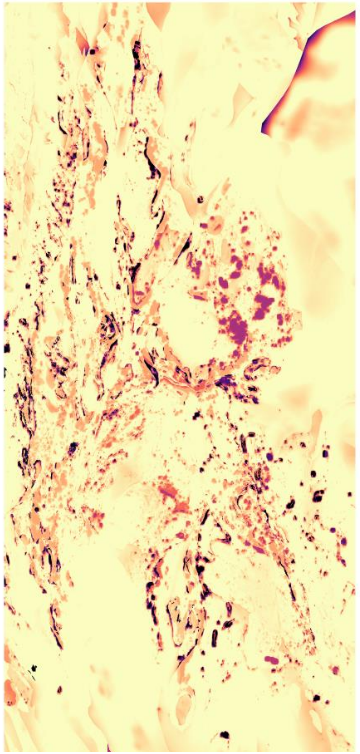


Anomaly detection of interpreted geology

- up to 50% anomalous (nu = 0.5)
- Anomalies (purple)
- Distance to hyperplane
- Mean anomalous distance per class



Anomalies



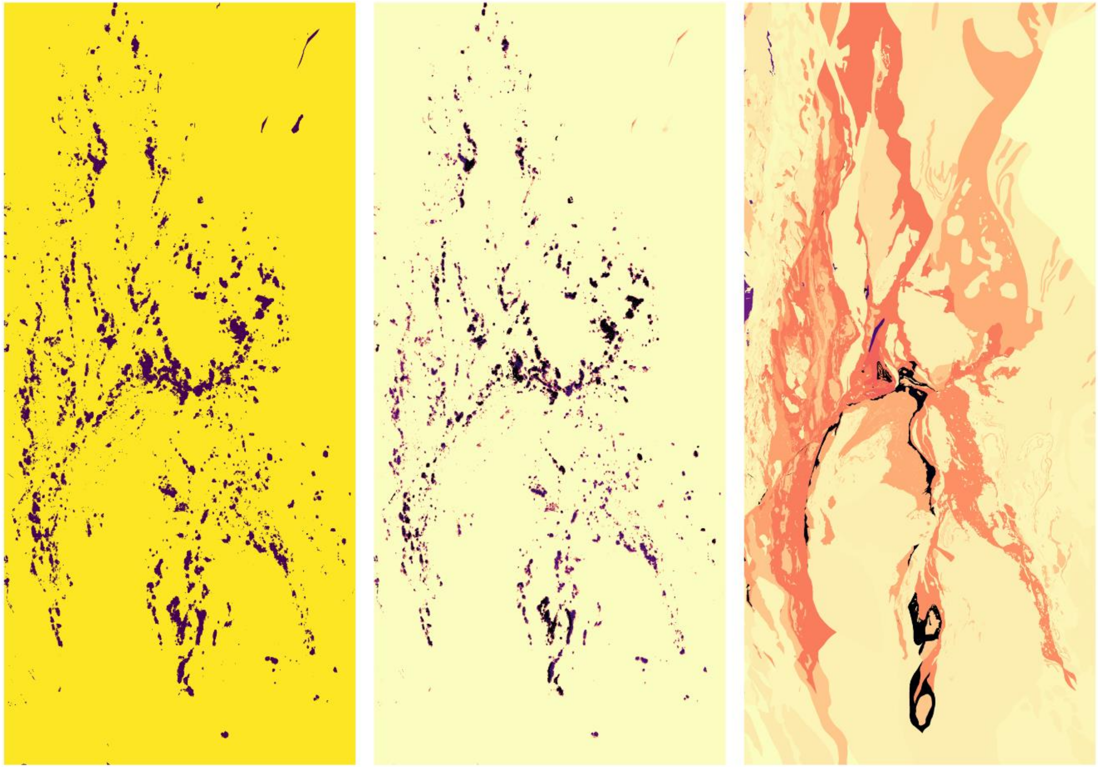
Anomalous distance

Mean anomalous distance

Anomaly detection of interpreted geology

- 5% anomalous

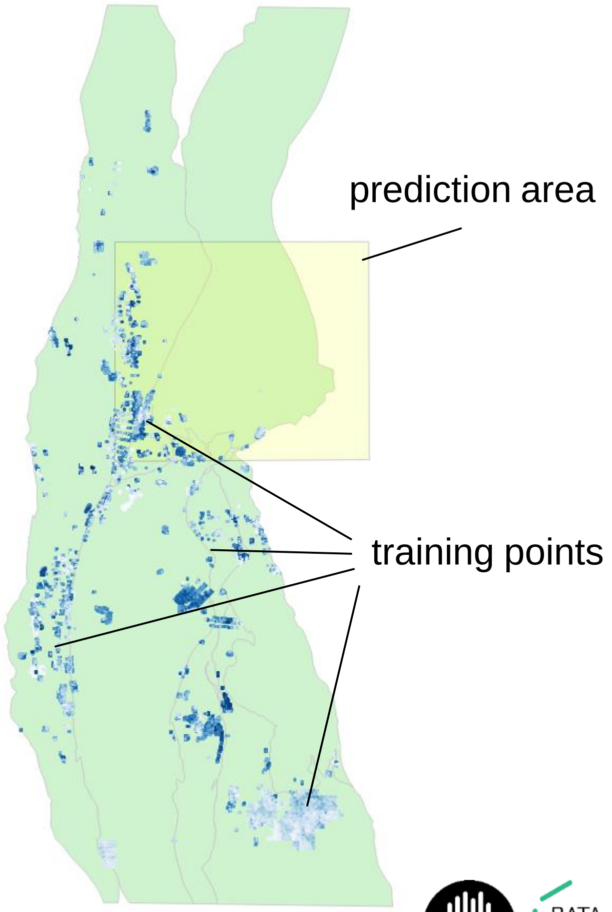
- Anomalies (purple)
- Distance to hyperplane
- Mean anomalous distance per class



Anomalies　Anomalous distance　Mean anomalous distance

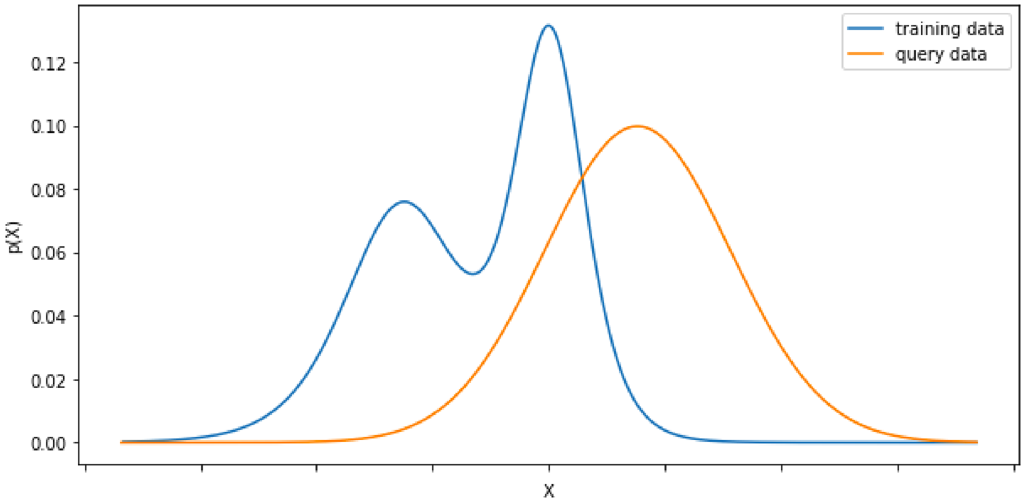
Modelling of geochemistry

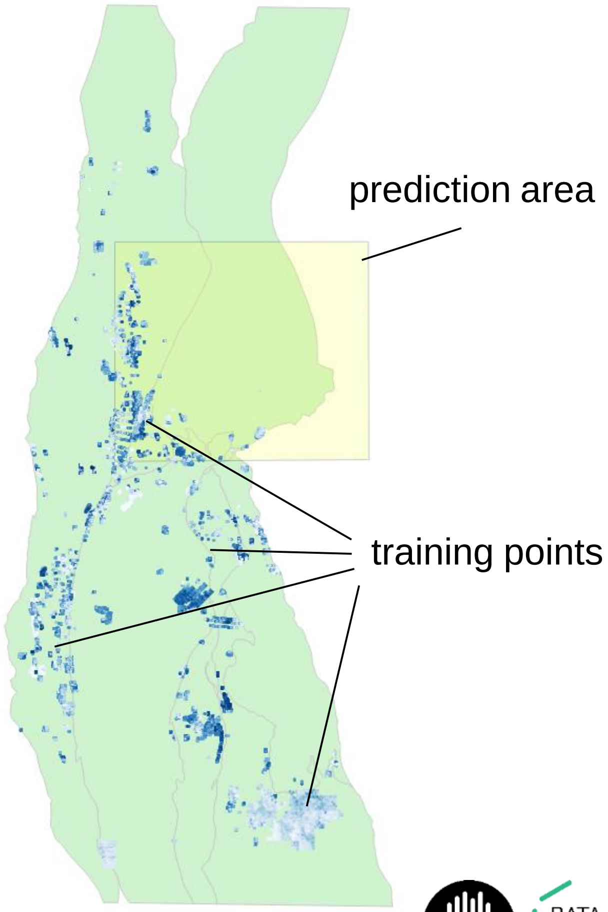
- Spatial model of soil geochemistry
- Covariates: geophysics
- Targets: chemical concentrations
- Training data: survey samples from out-cropping areas
- Model
  - Random forest regressor
  - Classifier to model covariate shift



Covariate shift

- Training data different from query data
- Biased sampling (e.g. fixed area, outcropping, known/expected properties)
- Train a classifier to distinguish between them





Modelling of geochemistry

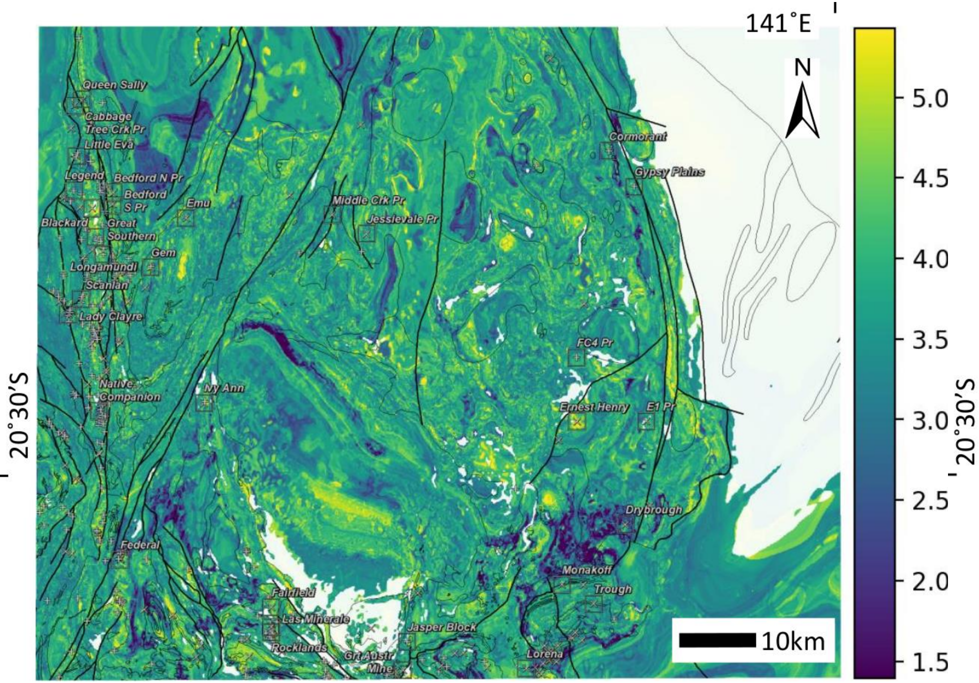
- Cross-validation results

Element	R^2 score (magnetics and gravity)	R^2 score (magnetics and gravity + wavelets)
Au	0.44	0.51
Cu	0.65	0.72
Pb	0.69	0.79
Zn	0.61	0.68

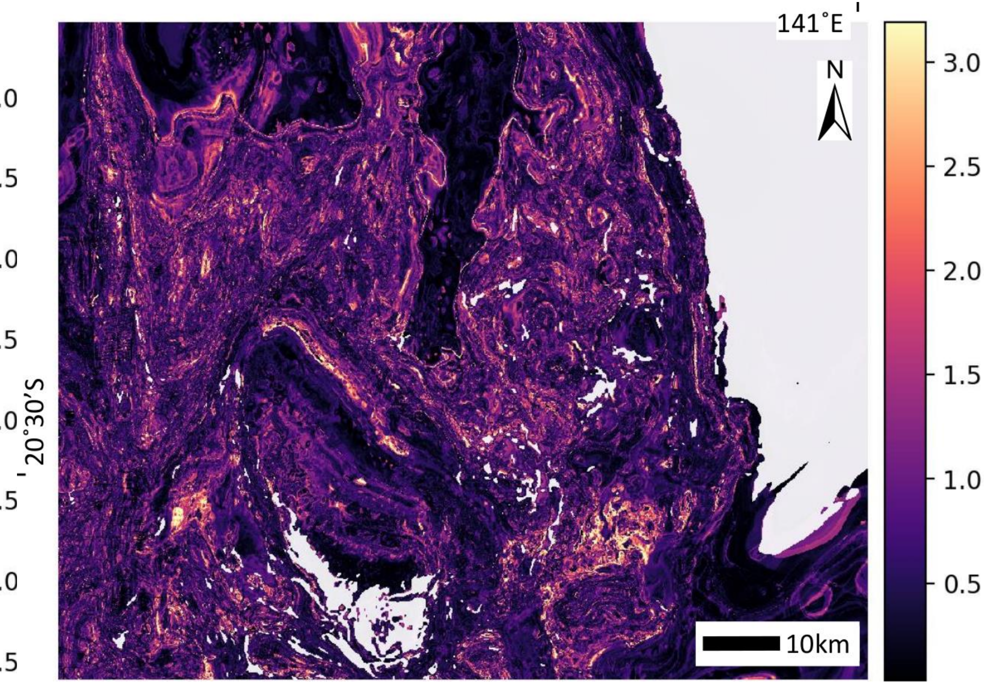
Covariate shift classifier accuracy	55%	95%

Modelling of geochemistry

- log(Cu) prediction - gravity and magnetics



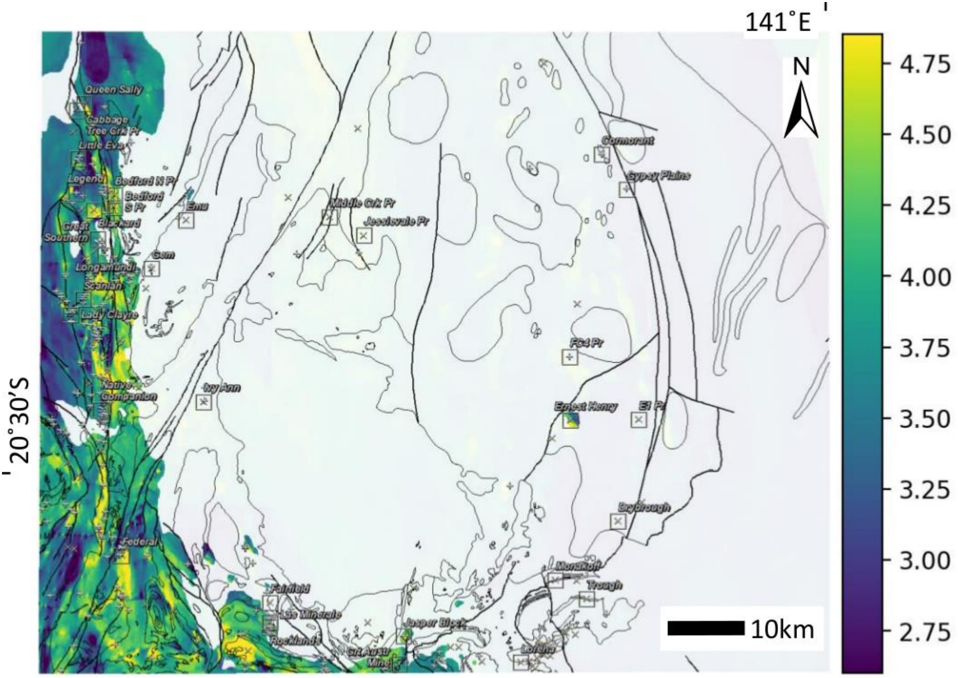
Model prediction



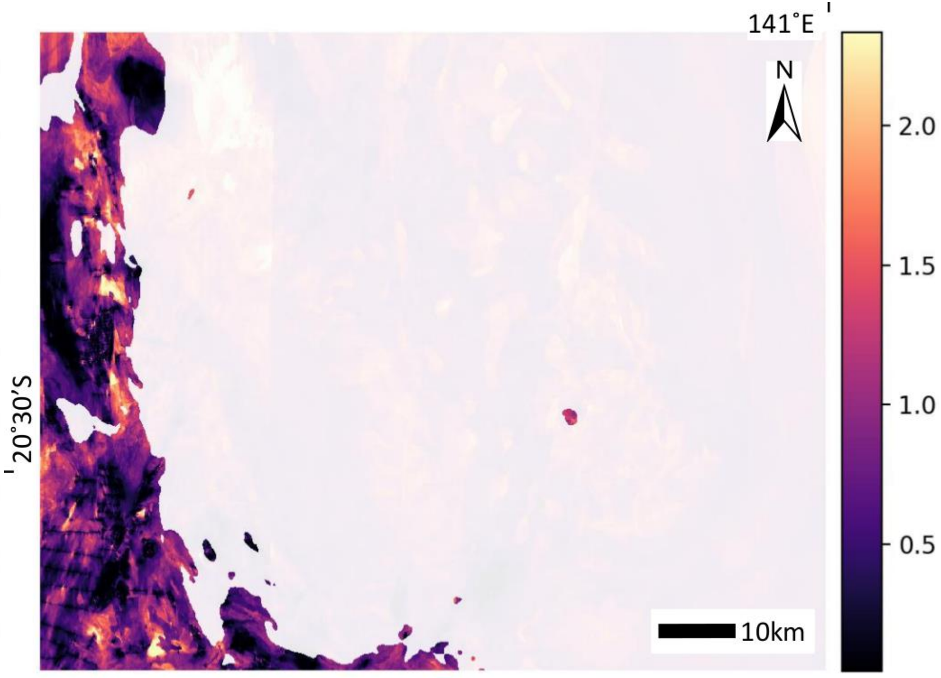
Model uncertainty

Modelling of geochemistry

- log(Cu) prediction - gravity and magnetics + wavelets



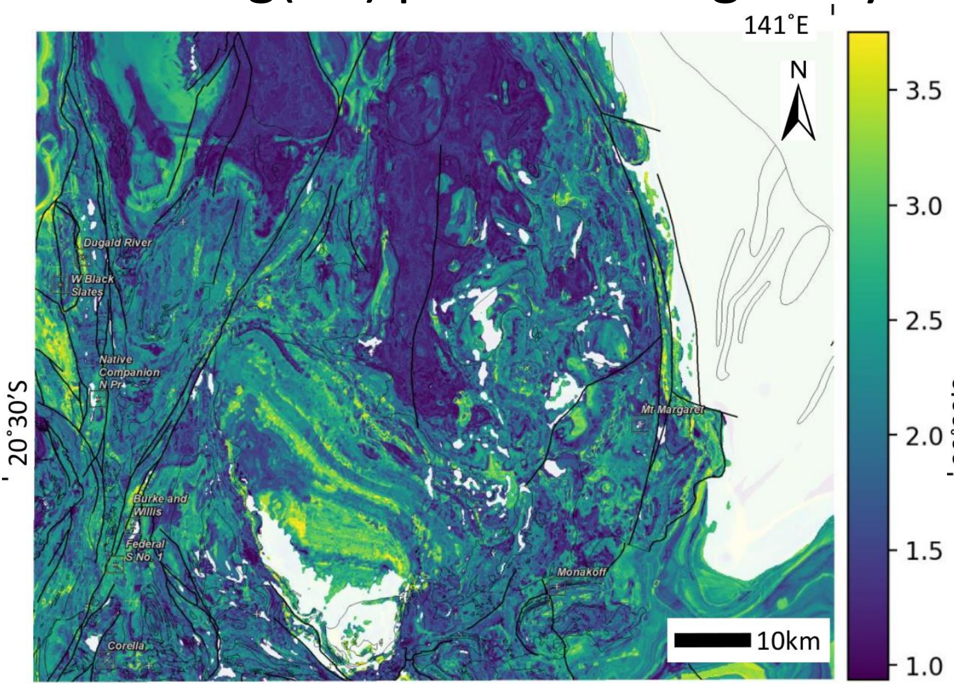
Model prediction



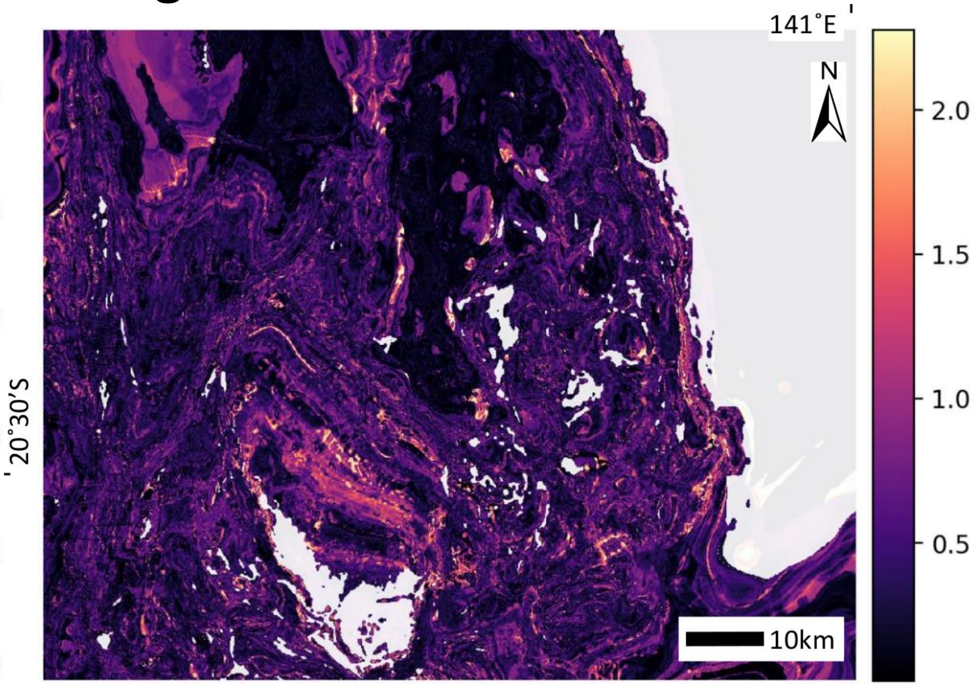
Model uncertainty

Geochemical modelling

- log(Pb) prediction - gravity and magnetics



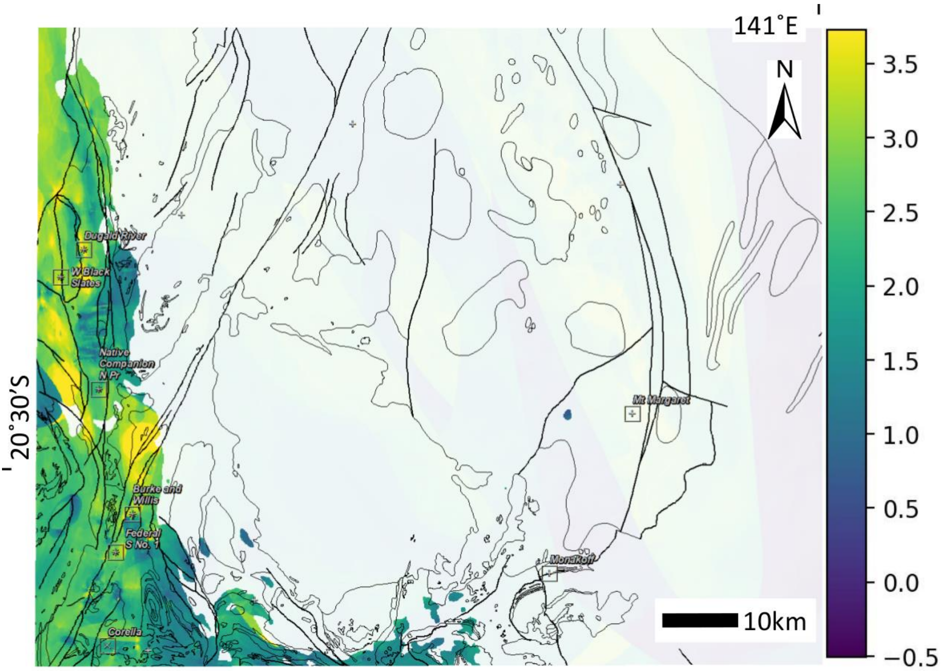
Model prediction



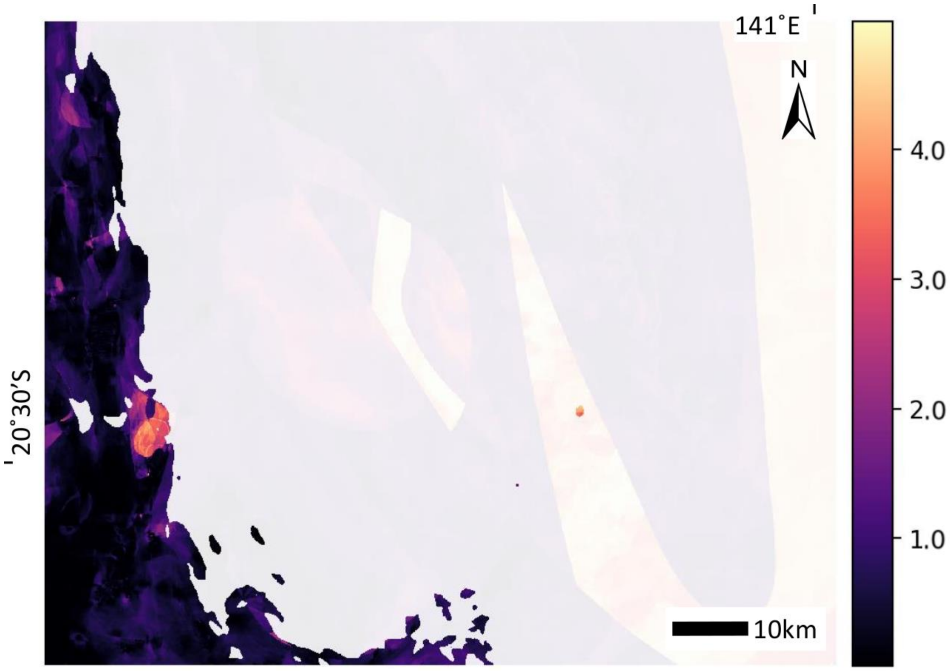
Model uncertainty

Modelling of geochemistry

- log(Pb) prediction - gravity and magnetics + wavelets



Model prediction



Model uncertainty

Conclusions & Recommendations

- Machine learning approaches in geological mapping and exploration targeting:
  - Augment existing mapping processes (e.g. geology classification, anomaly detection, outcrop prediction)
  - Provide geospatial predictions from point observations (e.g. geochemistry - better gridding/can serve as alternative to prospectivity mapping)
  - Maximise data use while estimating uncertainty in prediction
- Considerations for GSQ data practises:
  - Validate data-driven models with new/unseen data
  - Collect new datasets (based on model predictions / uncertainty)
  - Awareness of sampling bias (e.g. negative samples, unexplored regions)
  - Data management / integration to support data-driven modelling

THANK YOU

CSIRO Mineral Resources
Vasek Metelka
Research scientist

t +61 8 6436 8827
e vaclav.metelka@csiro.au

www.csiro.au