https://doi.org/10.56295/1ACEG32

DEVELOPMENT OF A GEOTECHNICAL ROCK MASS MODEL USING GEOCHEMICAL DATA AND MACHINE LEARNING

Megan Baker, Sue Chan, Christine Lion and Daniel Strang
PSM

ABSTRACT

In mining, the development of a realistic and well-informed engineering geological model can be challenging. Mining projects typically present with sparse geotechnical data and variable ground conditions, resulting in challenges with the interpretation of geotechnical unit boundaries, and the development of a three-dimensional model. This paper presents a case study where geochemical data collected for orebody knowledge definition was used to develop a rock mass model to inform pit slope design. During model development, borehole domaining assisted by machine learning was trialed to improve model reliability and enhance efficiency. This paper has demonstrated that an automated approach to rock mass classification using geochemical data is an effective and useful method which can be incorporated into the geotechnical model development process.

The methodology adopted for rock mass classification and model development is summarised below:

- Rock mass classification of geotechnical boreholes based on logging data and core photographs.
- Selection of 'twin' reverse circulation (RC) boreholes (adjacent to geotechnical boreholes).
- Comparison of downhole geochemical data against the logged rock mass unit intervals.
- Identification of geochemical signatures associated with each rock mass unit.
- Trial of machine learning methods to identify and classify rock mass units using geochemical data.
- Apply the results of the machine learning to a wider dataset with the objective of producing a 3D rock mass model.

1 INTRODUCTION

The concept of using 'indirect' data sources (i.e. geochemical and geophysical data) to interpret the boundaries of geotechnical detrital units in the Pilbara was introduced by Baxter (2016). The methodology described uses secondary data sources (collected during geological investigations) to supplement geotechnical site investigation data and in turn increase the confidence of the geotechnical model. This methodology has also been used by Noble, Hemraj and Eggers (2021), where geochemical signatures for Cenozoic cover sediments were identified and applied to a series of 2D cross sections which were then used to interpret the geotechnical model.

This paper aims to expand on the work completed by Baxter (2016) and Noble et al (2021), by using machine learning methods to assign rock mass unit boundaries based on geochemical signatures. The current methodology for assigning rock mass units using geochemical data has only been applied in a 2-dimensional sense, with boreholes being characterized along key geotechnical sections. Although the development of geotechnical sections in areas of elevated geotechnical risk is crucial for analysis and slope design, the current method for assigning domains is time consuming, relies heavily on interpretation and assumes that the areas of highest geotechnical risk are known. Machine learning methods however, are expected to provide a quick approach to assigning rock mass domains and will likely require minimal manual interpretation of the data, providing the human geologist/engineer more time for detailed interpretation. Machine learning methods will also allow for the processing of large datasets which can be used to develop a 3-dimensional model. This paper presents the results of a machine learning trial where machine learning methods have been used to interpret geotechnical rock mass unit boundaries using geochemical data.

2 BACKGROUND

The dataset available for the machine learning trial is from an iron ore deposit in the Pilbara Region of Western Australia. The projects mineralisation is hosted in the Marra Mamba Iron Formation (of the Hamersley Group) which is buried beneath Cenozoic aged detrital sediments. These Cenozoic detrital sediments are geotechnically challenging due the high variability in the rock mass conditions and the uncertainty surrounding the spatial distribution of units.

The project itself is proposed to be a large open pit operation with the construction of relatively shallow (~120 m deep) satellite pits which will be spread over a large area (approx. 200 km^2). The project is currently in the early stages of design and has limited geotechnical data with only eight (8) geotechnical boreholes drilled across the proposed development

area. Therefore, defining the spatial distribution of the geotechnical units is key to the geotechnical model and the overall slope design.

3 GEOLOGICAL OVERVIEW

The project is located in the Hamersley Province in the Pilbara region of Western Australia. Archean to lower-Proterozoic bedrock sequences in the area comprise the Marra Mamba Iron Formation, the lower most unit of the Hamersley Group. Dominant lithologies of the Hamersley Group include banded iron formation (BIF) interbedded with shale, dolomite, and acid volcanics which are intruded by dolerite sills and dykes.

Iron enrichment for the deposit is concentrated within the Detrital sequence and the BIF of the Marra Mamba Iron Formation. The Marra Mamba Iron Formation outcrops in the north along the Chichester range and is buried under Cenozoic aged sediments towards the south. These Cenozoic sediments contain Detrital Iron Deposits (DID) which are hematite-goethite rich colluvial/alluvial deposits derived from the erosion of banded iron ores hosted in the Marra Mamba Iron Formation. A geology map of the Pilbara Region showing the Hamersley group rocks and known occurrences of channel iron deposits is shown in Figure 1.

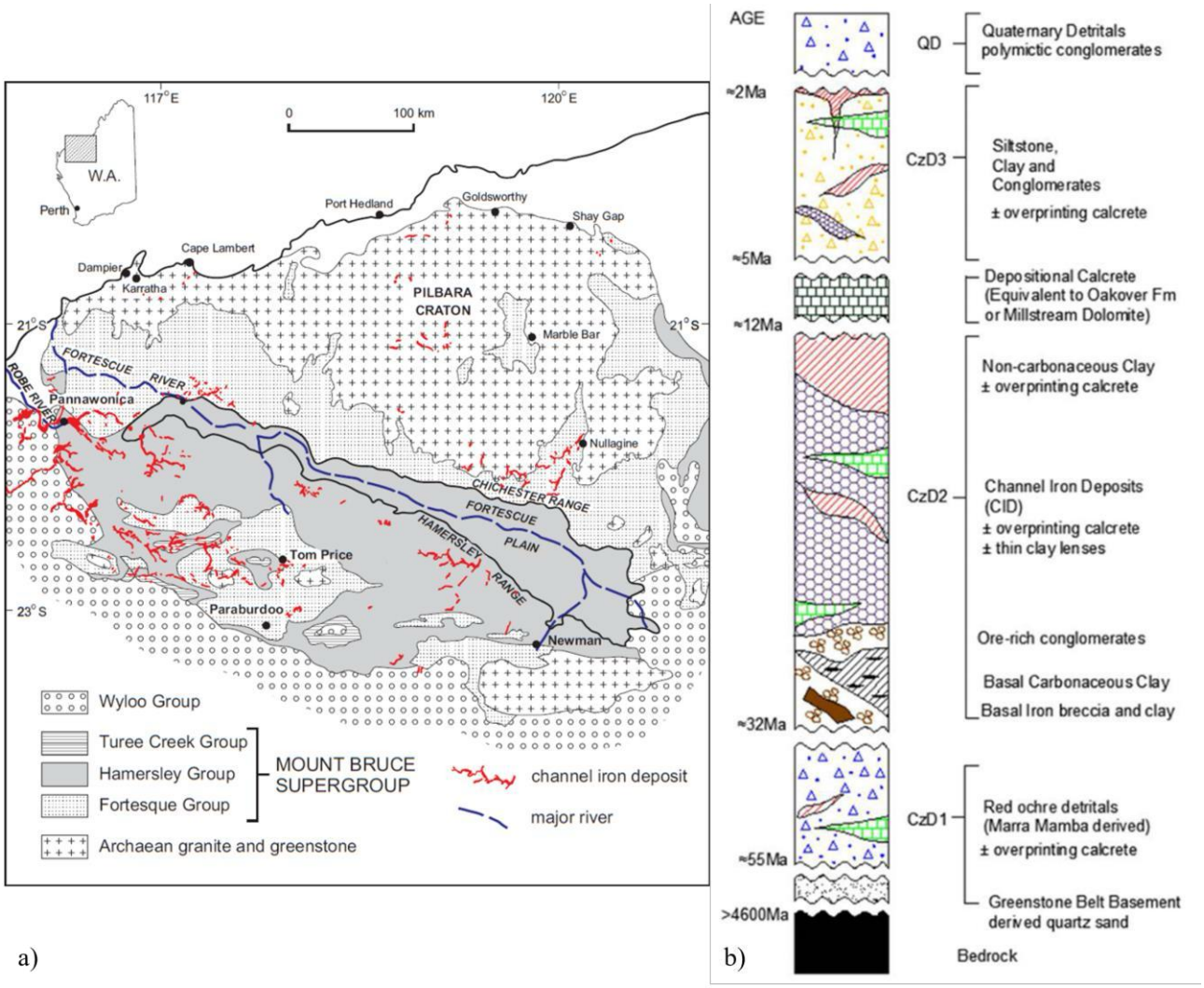


Figure 1: (a) Geology and Channel Iron Deposits of the Pilbara Region (Stone, 2005); (b) Suggested Stratigraphic column for Pilbara Cenozoic Detrital sequences (Baxter, 2016).

3.1 GEOLOGICAL HISTORY OF CENOZOIC DETRITALS IN THE PILBARA

The DID settings are typically hosted in Marra Mamba to Brockman Iron Formation Strike Valley (MBSV) environments. Figure 1 shows a suggested stratigraphic column for the Cenozoic detrital sequences produced by Baxter (2016). The DID sequences have been grouped based on their geological nature and depositional character. These three main detrital groups are described below.

- CzD1 (Cenozoic Detritals 1), oldest
  - Red ochre detritals, including a hematite conglomerate, basal conglomerates and a cemented vitreous hardcap
- CzD2 (Cenozoic Detritals 2)
  - Iron rich pisolites, lacustrine clays, overprinting calcretes and iron rich conglomerates with varying degrees of cementation
- CzD3 (Cenozoic Detritals 3), youngest.
  - Alluvial sediments such as siltstones and conglomerates.

4 AVAILABLE DATA

Limited geotechnical data was available for the project. The data available included:

- Eight (8) geotechnically logged diamond boreholes with core photographs
- Seventy-Eight (78) historic diamond boreholes with geological logging and core photographs. Three (3) of which were used in this assessment.
- Eleven (11) neighbouring Reverse Circulation boreholes (RC) with the following data:
  - Geological logging
  - Geochemical assay results including weight percentages of the elements Fe, SiO_2, Al_2O_3, P, Mn, LOI, S, TiO_2, CaO, MgO, K_2O and Zr
  - Geophysical data including density, gamma, resistivity and susceptibility for 8 of the 11 RC holes
- An additional ~9,500 RC boreholes were available across the project with geological logging and geochemical assay data. Geophysical data was also available for most of these boreholes (~8,970 holes).
- A geological model with the following surfaces:
  - Base of CzD3
  - Base of CzD2
  - Base of Detritals
  - Base of Archean units, i.e. Mount Newman, MacLeod and Nammuldi Members.

The diamond boreholes available were primarily drilled for geometallurgical purposes and were often not cored through the detrital sequence. Therefore, only three of the historic diamond boreholes were used in this assessment as rock mass classification could not be completed without core photographs.

Machine learning requires a complete dataset to train the model and make reliable predictions. An incomplete dataset may disrupt the learning process and generate inaccurate results. As geophysical data was not available for all of the neighbouring twin RC holes, the geophysical data has been excluded from the machine learning model.

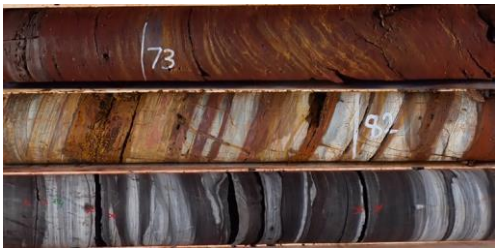
5 GEOTECHNICAL ROCKMASS UNITS

Rock mass units were assigned based on the interpreted stratigraphy of the geotechnically logged boreholes. The lithology, strength, geotechnical conditions and variability of each unit were considered when grouping rock mass types with similar geotechnical characteristics. Several geotechnical rock mass units occur within the detrital sequence. The typical characteristics and strengths of each rock mass unit are described in Table 1. The field estimated strengths are based on AS 1726:2017 Geotechnical site investigations (Standards Australia, 2017) and the CANMET 'Pit Slope Manual' (Herget, 1977). Typical photos of each rock mass unit are shown in Table 1.

Table 1: Observed Geotechnical Rock Mass Units

Geotechnical Rock Mass Unit	Field Estimated Strength	Example Photograph
Quaternary Detritals A polymictic conglomerate of BIF, hematite and shale clasts in a moderately consolidated silty matrix. Clasts are typically sub-angular and range from fine gravel to cobble.	Very stiff soil to low rock strength (S4 to R2)	

Geotechnical Rock Mass Unit	Field Estimated Strength	Example Photograph
Calcrete Typically observed as orange-brown and white overprinting of the shallow detritals and clayey units. Vigorous reaction with HCL. The presence of calcrete is indicative of saline groundwater.	Low to medium rock strength (R2 to R3) ~10% Extremely low rock strength (R0)	
Gravelly Siltstone Fine to coarse gravels of sub-rounded to sub-angular BIF and hematite in a silt matrix. Commonly moderately consolidated and matrix dominated. Pisolitic zones occur towards the base of the unit.	Very stiff soil to very low rock strength (S4 to R1)	
Pisolite Fine to medium pisoid nodules in a silt matrix. Pisoids are rounded to sub-rounded and are strongly magnetic. Moderately consolidated to unconsolidated in parts.	Very stiff soil to very low rock strength (S5 to R1), cemented in parts (~20% R2 to R4)	
Lacustrine Clay A clay-rich unit comprising ~60% clay and ~40% gravelly clay and breccia. Typically, massive with occasional relict bedding. Mottled texture in areas due to iron staining and manganese alteration.	Very stiff soil to extremely low rock strength (S4 to R0)	
Hematite Conglomerate Fine to medium gravels of sub-rounded to sub-angular hematite clasts in a silt matrix. Predominantly clast supported with a silt matrix. Moderately consolidated and indurated in parts.	Bimodal strength, typically low to high rock strength (R2 to R4), with ~10% very stiff to hard soil (S4 to S5)	
Cemented Detritals Iron rich gravelly, cemented and hydrated detritals. Overprinting of the clastic texture occurs in parts as the rock has a 'welded' appearance. When hydrated it is difficult to distinguish from in-situ Canga/Hardcap units. Vughy in parts.	Low to high rock strength (R2 to R4), ~5% very high rock strength (R5)	
Clayey Breccia Colluvium Highly variable in appearance from a gravel dominated breccia to cobbles and boulders of Canga interspersed with laterally discontinuous clays. Clast sizes and proportions are highly variable.	Bimodal strength, typically very stiff soil to very low rock strength (S4 to R1), ~15% low to high rock strength (R2 to R4)	

Geotechnical Rock Mass Unit	Field Estimated Strength	Example Photograph
Cangarised BIF Highly to extremely weathered BIF, non-magnetic and steely in parts. Vughs often elongated parallel to bedding and occasionally infilled with clay. Bedding typically indistinct and overprinted.	Very low to high rock strength (R1 to R4) with pockets of (~10%) very stiff to hard soil (S4 to S5)	
Archean Bedrock Interbedded Shale and Banded Iron Formation (BIF). Occurs as hydrated, highly to moderately weathered, mineralised and Fresh. Hematite and goethite rich zones occur when mineralised. Fresh BIF is present at depth.	**Mineralised BIF**: Friable to medium rock strength (S5-R3) **Weathered BIF:** Medium to high rock strength (R3 to R4) **Fresh BIF:** medium to very high rock strength (R2 to R5)	

6 TWIN ASSESSMENT

To distinguish geochemical signatures for each geotechnical unit, a comparison of diamond core and RC twin boreholes was undertaken. The twin holes are pairs of cored and RC holes located as close as possible to each other (up to 20 m) and are usually positioned on the same drill pad. All diamond holes with logged geotechnical rock mass units were reviewed for the twin assessment. Assay data from each twin hole was compared with the rock mass unit boundaries of the corresponding geotechnical borehole to identify marker signatures in the unit boundaries. A typical comparison log showing the logged geotechnical units of the diamond borehole and the geochemical data of the RC hole is shown in Figure 2. The example presented shows obvious geochemical variations downhole which can be correlated back to the boundaries of the different detrital units.

Geotechnical Borehole Rock Mass Units	Depth (m)	RC Borehole Geochemical Assay Data											
		Fe (%)	SiO2 (%)	Al2O3 (%)	P (%)	Mn (%)	LOI (%)	S (%)	TiO2 (%)	CaO (%)	MgO (%)	K2O (%)	Zr (%)
Quaternary Detritals	0 - 10	17.16	58.1	10.55	0.025	0.04	4.88	0.019	0.593	0.11	0.3	0.487	0.018
Gravelly Siltstone	10 - 12	17.98	48.95	16.11	0.022	0.03	6.49	0.019	1.201	0.1	0.36	0.568	0.031
	12 - 14	18.91	47.39	16.36	0.023	0.03	6.62	0.016	1.238	0.1	0.31	0.46	0.033
	14 - 16	18.08	47.75	16.86	0.022	0.02	7.05	0.016	1.215	0.12	0.29	0.412	0.03
	16 - 18	27.03	41.76	11.95	0.021	0.02	5.57	0.023	1.046	0.07	0.22	0.278	0.027
	18 - 20	37.65	27.78	10	0.025	0.03	6.7	0.03	0.785	0.05	0.16	0.14	0.02
Cemented Detritals	20 - 22	54.68	6.37	4.72	0.027	0.005	9.83	0.037	0.198	0.03	0.09	0.018	0.005
	22 - 24	56.74	5.14	4.36	0.023	0.01	9.1	0.028	0.113	0.02	0.08	0.016	0.005
	24 - 26	42.25	14.48	13.02	0.025	0.04	10.57	0.025	0.54	0.05	0.16	0.024	0.015
Lacustrine Clay	26 - 28	12.38	37.36	30.69	0.013	0.01	12.5	0.009	0.79	0.11	0.31	0.029	0.019
	28 - 30	6.44	41.1	34.09	0.009	0.02	12.8	0.006	1.957	0.1	0.31	0.037	0.035
	30 - 32	3.28	42.43	36.05	0.007	0.06	13.16	0.002	2.594	0.07	0.27	0.023	0.041
	32 - 34	3.36	42.12	36.29	0.008	0.04	13.18	0.002	2.68	0.04	0.24	0.025	0.04
	34 - 36	3.75	38.89	37.43	0.01	0.02	14.57	0.006	2.758	0.04	0.23	0.024	0.049
	36 - 38	2.8	41.5	36.09	0.019	0.1	13.36	0.008	3.843	0.06	0.23	0.022	0.056
	38 - 40	1.23	45.23	35.63	0.036	0.005	11.21	0.001	1.956	0.08	0.84	2.702	0.027
	40 - 42	25.01	27.11	21.24	0.079	0.36	11.29	0.005	1.077	0.08	0.64	1.603	0.026
Banded Iron Formation	42 - 44	37.62	17.92	13.77	0.072	0.83	11.73	0.013	0.729	0.06	0.28	0.048	0.017
	44 - 46	59.46	3.42	1.84	0.063	0.19	8.75	0.006	0.098	0.01	0.09	0.022	0.002
	46 - 48	59.31	4.47	3.37	0.047	0.14	6.55	0.006	0.259	0.01	0.08	0.012	0.007
	48 - 50	59.75	3.86	2.21	0.065	0.21	7.4	0.008	0.118	0.01	0.07	0.017	0.003
	50 - 52	60.3	3.71	1.91	0.072	0.2	7.25	0.008	0.096	0.01	0.08	0.016	0.003
	52 - 54	46.04	27.15	1	0.063	0.2	5.03	0.005	0.042	0.01	0.09	0.022	0.003
	54 - 56	45.13	28.67	0.95	0.064	0.24	5.01	0.006	0.043	0.01	0.1	0.024	0.003
	56 - 58	40.37	33.06	2.09	0.069	0.36	5.77	0.009	0.118	0.02	0.13	0.033	0.005
	58 - 60	55.67	9.27	2.23	0.123	0.36	7.51	0.008	0.088	0.02	0.12	0.02	0.001
	60 - 62	55.09	5.4	3.63	0.141	0.84	9.81	0.007	0.159	0.02	0.1	0.126	0.003
	62 - 64	58.98	3.7	1.59	0.155	0.33	9.2	0.006	0.066	0.02	0.07	0.057	0.002
	64 - 66	50.62	17.85	0.87	0.143	0.16	7.84	0.005	0.059	0.02	0.07	0.02	0.003

Figure 2: Example comparison log showing inferred rock mass units in the geotechnical borehole (left) and the assay data of the nearby RC twin (right)

7 GEOCHEMICAL SIGNATURES

Geochemical signatures were derived for each rock mass unit during the twin assessment. A series of histograms were produced for each unit to compare element percentages. Using the histograms, grade cut off values were determined based on the spread of the data. Rock mass units which indicate strong geochemical signatures include, lacustrine clay and calcrete. Iron rich units such as hematite conglomerate, pisolite and cemented detritals have very similar geochemical signatures and are difficult to differentiate. Due to the nature of the clayey breccia colluvium, a mixed breccia and clay, there are large variations in the geochemistry. These variations make the unit difficult to distinguish.

The geochemical signatures identified for the project are presented in Table 2.

Table 2 – Geochemical signatures identified for the project

Unit	Typical Assay Value (%) (1)										
	Fe	SiO2	Al2O3	LOI	Mn	CaO	MgO	TiO2	K2O	S	P
Quaternary Detritals	<30	>40	8-18	3-12	<1	<6	<3	0.3-1.3	0.4-1.9	<0.3	0.02 - 0.04
Gravelly Siltstone	10-50	10-60	8-22	4-12	<1	<1	<1	0.7-1.5	<1.5	<0.2	0.01-0.08
Pisolite	25-50	<35	10-25	3-13	<1	<1	<1.5	1-1.7	<0.3	<0.1	<0.04
Calcrete	<15	30-40	7-12	>10	<1	>10	>4	<0.5	<1	<0.5	<0.03
Cemented Detritals	35-60	<25	4-15	5-13	<1	<1	<1	<1.2	<0.2	<0.1	<0.15
Hematite Conglomerate	30-50	<40	<15	<12	<3	<1	<1	<0.6	<0.2	<0.1	<0.07
Lacustrine Clay	<35	10-50	16-40	10-15	<1	<1	<1	<3	<0.2	<0.1	<0.1
Clayey Breccia Colluvium	10-50 (varies)	5-75	5-25	<12	<1	<1	<1	<1.2	<0.2	<0.1	<0.06
Canga	>35	<30	2-10	5-15	<2	<3	<2.5	<0.8	<0.5	<0.2	0.01-0.03
Bedrock	varies	varies	<5	<15	<2	<1	<1	<0.5	<0.2	<0.4	<0.15

(1) The highlighted cells indicate elements which have diagnostic signatures for specific units

8 MACHINE LEARNING

In order to use the geochemical signatures identified and apply the signatures to a large dataset, a fast and reasonably accurate algorithm is required to domain boreholes based on the geochemistry. Machine learning is a form of artificial intelligence that can be used to analyse large datasets. It identifies and learns patterns from data to make decisions with minimal human intervention. Machine learning has fast processing capabilities and once the model is trained, it requires minimal input to generate results. Because of these factors, machine learning has been trialled to domain the boreholes with geochemical data.

During the machine learning trial, many different algorithms were tested and used to develop the classification model. In this paper we present the final results obtained with the best performing models.

8.1 DATA STATISTICS

The primary case study for this paper has a relatively small dataset with only 11 geotechnically logged diamond boreholes with RC twins. The twin RC boreholes have geochemical assay data in 2 m intervals and occasionally 10 m intervals. Some assay intervals used for training contained multiple inferred geotechnical units. In this instance, the dominant detrital unit (i.e., the unit covering over half of the assayed interval) was adopted for the entire interval. The available assay data for the project comprises weight percentages of the following elements, Fe, SiO_2, Al_2O_3, P, Mn, LOI, S, TiO_2, CaO, MgO, K_2O and Zr. Machine learning relies on data consistency. Therefore, elements without a complete dataset (e.g. Zr) were not used to build and train the model.

Table 3 summarises the available samples of geochemical data for each geotechnical unit. The Marra Mamba Iron Formation and gravelly siltstone units represent 58% of the total population of data. Some detrital units such as hematite conglomerate, calcrete and silcrete comprise less than 2% of the whole dataset. These units with low sample numbers tend to have specific geochemical or geotechnical characteristics and have therefore been kept independent to avoid noise in the data.

Table 3 – Number of Samples for Each Geotechnical Rock Mass Unit

Unit	Number of Samples	Percentage (%)
Quaternary Detritals	19	6
Gravelly Siltstone	82	24
Pisolite	33	10
Calcrete	2	<1
Cemented Detritals	31	9
Hematite Conglomerate	6	2
Lacustrine Clay	9	3
Clayey Breccia Colluvium	13	4
Canga	21	6
Silcrete	2	<1
Bedrock	123	36
Total	**341**	**100**

8.2 MACHINE LEARNING MODEL

Due to the small size of the available dataset, a 10-fold cross validation method was used to build and evaluate the model. For each iteration 90% of the data was randomly selected to train the model, and the remaining 10% was used to test the model. The representation of each rock mass unit in the dataset was imbalanced, with some units, such as the Marra Mamba Iron Formation, representing 36% of the dataset and the lacustrine clay only representing 3%. Due to the data imbalance, the model will not perform well at identifying lacustrine clay unless it has strong distinguishable geochemical characteristic. A two-step method for model development was adopted to overcome this issue:

1 Build a model to identify two units: Marra Mamba Formation versus 'the rest'
2 Build a model that classifies the remaining detrital units.

8.3 MODEL PERFORMANCE

After testing different methods, which we will not detail here, the Random Forest Classifier (RF) algorithm for both steps was adopted, as it provided an overall accuracy of 82% in assigning rock mass units. In machine learning and data analysis, the common metrics for evaluating a models performance include sensitivity/recall, specificity, precision and F-score. F-score is a measure of a model's accuracy and is the harmonic mean of precision and recall. A high F-score indicates high precision and recall while a low F-score suggests poor model performance. When working with imbalanced datasets, the F-score is an important metric to consider as it can indicate whether the model is overpredicting or underpredicting certain classes. Table 4 presents the sensitivity, specificity, precision and F-score results for all units classified by the RF model algorithm.

In terms of the model performance, we observe that:

- Marra Mamba Formation, gravelly siltstone, lacustrine clay, Quaternary detritals, cemented detritals and pisolite all have an F-score greater than 70%
- The F-score could not be calculated for hematite conglomerate as none of the samples were predicted correctly in the cross-validation
- Comprising more than a half of the total population, both Marra Mamba Formation and gravelly siltstone can be predicted with high accuracy by the machine
- Lacustrine clay has the third-highest F-score with a precision of 100% and a sensitivity of 67%. This indicates that no other lithology samples are classified as lacustrine clay mistakenly, however, some lacustrine clay samples are incorrectly predicted as other units

- Both Quaternary detritals and cemented detritals have a sensitivity higher than the precision as false positive results are greater than false negative results for these lithologies. This means samples tend to be predicted into these lithologies mistakenly
- Pisolite, calcrete, clayey breccia colluvium, and lacustrine clay are classified with higher precision than sensitivity. The false negative results for these lithologies are greater than the false positive results, indicating that their samples tend to be predicted as other lithologies.

Table 4: Specificity, sensitivity, precision and F-score results for all units

Unit	Sensitivity (average %)	Specificity (average %)	Precision (average %)	F-score (average %)
Quaternary Detritals	89	98	71	79
Gravelly Siltstone	87	92	78	82
Pisolite	64	98	81	71
Calcrete	50	100	100	67
Cemented Detritals	84	95	65	73
Hematite Conglomerate	0 (1)	99	0	Undefined (2)
Clayey Breccia Colluvium	38	99	56	45
Lacustrine Clay	67	100	100	80
Canga	57	97	52	55
Bedrock	86	95	91	88

(1) The highlighted cells indicate results with <50% sensitivity, specificity, precision or F-score
(2) Metrics are undefined when the calculations result in a division by 0 and there are no positive predictions.

8.4 RESULTS

The data imbalance and the size of the dataset are the limiting factors for this case study. The rock mass unit with the highest weighting was classified first to reduce the data imbalance when classifying the remaining units. The model developed can classify Quaternary detritals, gravelly siltstone, pisolite, cemented detritals, lacustrine clay and the Marra Mamba Iron Formation (bedrock) with >70% confidence. Overall, the RF method has proven successful in classifying rock mass unit boundaries using the case study data. However, additional samples are required to train the model and improve the accuracy when identifying the calcrete, hematite conglomerate, clayey breccia colluvium, and canga.

9 MACHINE LEARNING TRIAL OF A SECONDARY CASE STUDY

The original focus of this work was to trial machine learning methods using data from one project only (the project discussed in previous sections). As this project has limited geotechnical information, data from another project (case study 2) has also been trialled to compare models built from small and medium-sized datasets.

The secondary case study (case study 2) is another iron ore deposit in the Pilbara Region of Western Australia. The project has a similar geological setting with Cenozoic detrital sequences overlying Banded Iron Formation. The geotechnical rock mass units are also very similar except for a few additional units including calcareous lacustrine clay and calcrete (>20% CaO).

9.1 DATA STATISTICS – CASE STUDY 2

The second case study has a dataset of 249 RC boreholes with previously interpreted detrital units and accompanying assay data. The available assay data comprised Fe, SiO_2, Al_2O_3, P, Mn, LOI, S, TiO_2, CaO, MgO and K_2O measurements. Boreholes along two geotechnical sections were selected to test the model. The testing dataset comprises 1,839 samples across 11 units. Table 5 summarises the available samples of geochemical data for each geotechnical unit.

Table 5 – Number of Samples for Each Geotechnical Rock Mass Unit (Case Study 2)

Unit	Training		Testing		Total	
	No. of Samples	Percentage (%)	No. of Samples	Percentage (%)	No. of Samples	Percentage (%)
Quaternary Detritals	1447	23	384	21	1831	22
Gravelly Siltstone	742	12	262	14	1004	12
Calcareous Lacustrine Clay	4	0.063	7	0.38	11	0.13
Calcrete <20%	151	2.4	24	1.3	175	2.1
Calcrete >20%	154	2.4	14	0.76	168	2.0
Cemented Detritals	208	3.3	507	28	715	8.7
Silcrete	12	0.19	0	0	12	0.15
Hematite Conglomerate	300	4.7	55	3.0	355	4.3
Clayey Breccia Colluvium	467	7.3	26	1.4	493	6.0
Lacustrine Clay	72	1.1	19	1.0	91	1.1
Canga	506	7.9	4	0.22	510	6.2
Bedrock	2330	36	537	29	2867	34.8
Total	**6393**	**100**	**1839**	**100**	**8232**	**100**

9.2 MACHINE LEARNING MODEL

The Quaternary detritals and bedrock units represent 22% and 35% of the dataset respectively. As these units constitute a large proportion of the total dataset, they outweigh the other classes. To balance the weighting, we developed a two-step method using the RF classifier with the following steps:

1. Build a model to identify three units: Quaternary detritals, bedrock and 'the rest'
2. Build a model that classifies the remaining detrital units.

For this case study, we defined geological 'rules' to attempt to provide stratigraphic control when defining rock mass domains. The geological rules applied include defining the boundaries of lithologies in a bottom-up and top-down manner. The geological rules include:

- Setting the boundary bottom-up for bedrock and canga, i.e. canga can only occur immediately above bedrock
- Setting the boundary top-down for Quaternary detritals and gravelly siltstone, i.e. gravelly siltstone occurs directly beneath Quaternary detritals
- The top and bottom predicted calcrete intervals in a borehole defined the calcrete boundaries.

9.3 MODEL PERFORMANCE

In conjunction with the geological rules, the RF classifier algorithm provided the best results with an overall model accuracy of 83%. Table 6 presents the sensitivity, specificity, precision and F-score results for all units. We observe the following regarding model performance:

- Calcrete >20%, calcrete <20%, Quaternary detritals, bedrock, cemented detritals and gravelly siltstone all have an F-score greater than 70%:
    - The units constituting around 70% of the whole dataset, Quaternary detritals, bedrock and gravelly siltstone, can be predicted with high accuracy by the model
    - With the distinct geochemical signature for calcrete, the model can identify calcrete with high accuracy
- The F-score could not be calculated for calcareous lacustrine clay and canga because none of the samples were predicted correctly. Calcareous lacustrine clay has insignificant training and testing data (i.e. less than 10 samples); canga also has a small testing dataset (i.e. 4 samples)
- For units with F-scores lower than 70%, the units were typically classified incorrectly and mistaken for units with similar geochemistry:
    - Lacustrine clay had an F-score of 59%, a precision of 100% and a sensitivity of 42%. Therefore, no other units' samples were mistakenly identified as lacustrine clay, however, some lacustrine clay

samples were mispredicted. lacustrine clay was mistaken as either clayey colluvium or gravelly siltstone.

- Hematite Conglomerate was incorrectly classified as clayey colluvium and cemented detritals. Those units are basal detrital units with similar lithologies (i.e. clay and breccia).

Table 6: Specificity, sensitivity, precision and F-score results for all units (Case Study 2)

Unit	Sensitivity (average %)	Specificity (average %)	Precision (average %)	F-score (average %)
Quaternary Detritals	86	97	89	88
Gravelly Siltstone	82	95	75	79
Calcareous Lacustrine Clay	0	100	Undefined (2)	Undefined
Calcrete <20%	100	100	96	98
Calcrete >20%	100	100	100	100
Cemented Detritals	86	94	85	85
Hematite Conglomerate	33	98	37	35
Clayey Breccia Colluvium	42	99	41	42
Lacustrine Clay	42	100	100	59
Canga	0	100	0	Undefined
Bedrock	87	95	87	87

(1) The highlighted cells indicate results with <50% sensitivity, specificity, precision or F-score
(2) Metrics are undefined when the calculations result in a division by 0 and there are no positive predictions.

9.4 RESULTS

To summarise, the model developed using the data from the secondary case study and using geological rules had very similar results to the primary case study. The model can classify the units Quaternary detritals, gravelly siltstone, calcrete >20%, calcrete <20%, cemented detritals and bedrock with >70% confidence. However, the model performs poorly (<50% F-score) when identifying the calcareous lacustrine clay, lacustrine clay, clayey breccia colluvium, hematite conglomerate and canga units.

Figure 3 compares the downhole model prediction and the human-interpreted lithology boundaries for a geotechnical cross-section developed for the project. A good correlation exists between the model-predicted units and the human-interpreted lithology boundaries.

10 MACHINE LEARNING TRIAL OF A BALANCED DATASET

In machine learning, the model's accuracy depends on the data's quality and availability. The training and testing datasets selected in the secondary case study were selected by grouping boreholes along geotechnical sections. Using this method of grouping, the training and testing datasets have different weights for each detrital unit. For instance, cemented detritals comprised 3.3% of the training dataset and 28% of the testing dataset. Machine learning relies on a balanced dataset to ensure sufficient training and testing of the data.

In this section, a balanced dataset has been used for training and testing of the secondary case study data. Table 7 shows the number of training and testing samples for each identified unit in the balanced dataset. Samples were randomly picked for testing and training, and the data was not kept in stratigraphic order. Therefore, geological rules were not applied.

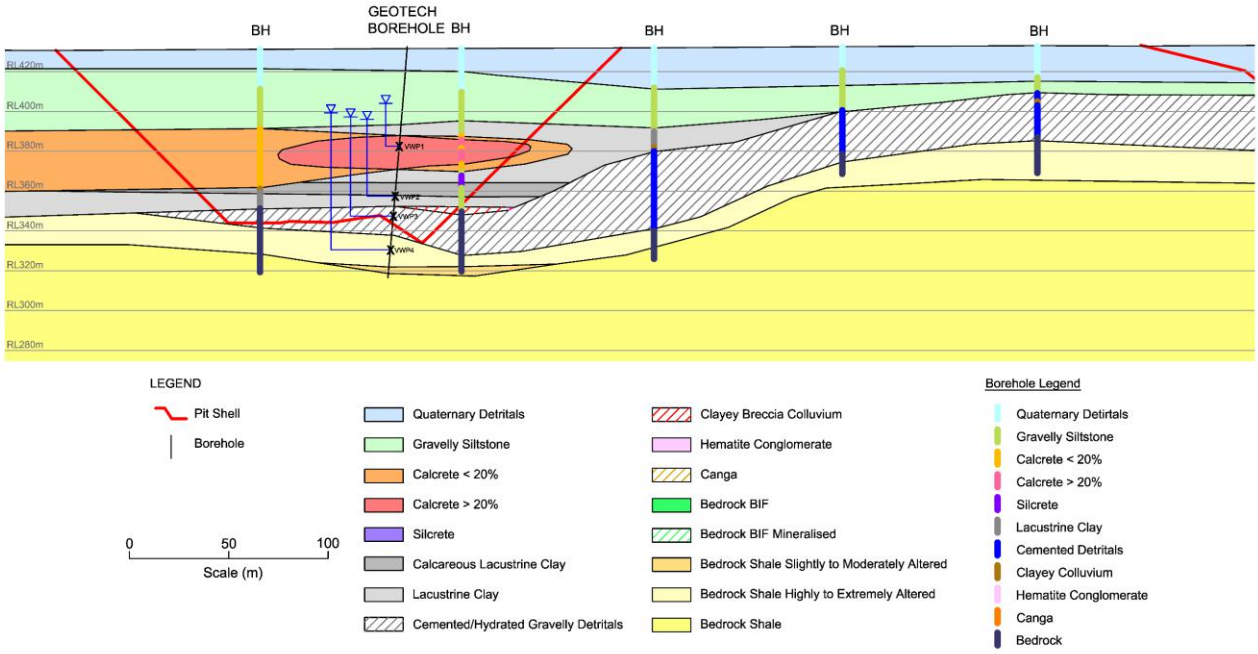


Figure 3: Comparison of Model predicted borehole intervals and the manually interpreted unit boundaries

Table 7 – Number of Samples for a Balance Dataset (Case Study 2)

Unit	Training		Testing	
	No. of Samples	Percentage (%)	No. of Samples	Percentage (%)
Quaternary Detritals	1464	25	367	22
Gravelly Siltstone	803	14	201	12
Calcareous Lacustrine Clay	8	0.14	3	0.18
Calcrete <20%	140	2.4	35	2.1
Calcrete >20%	133	2.3	35	2.1
Cemented Detritals	571	9.8	144	8.7
Silcrete	9	0.15	3	0.18
Hematite Conglomerate	285	4.9	70	4.2
Clayey Breccia Colluvium	394	6.7	99	6.0
Lacustrine Clay	72	1.2	19	1.1
Canga	408	7.0	102	6.2
Bedrock	1568	27	575	35
Total	**5855**	**100**	**1653**	**100**

10.1 MODEL PERFORMANCE

The RF classifier algorithm with the balanced dataset returned the best results with 92% overall model accuracy. Table 8 presents the sensitivity, specificity, precision and F-score results for all the units. Except for lithologies with a very limited sample size, namely calcareous lacustrine clay and silcrete, all other lithologies have an F-score greater than 70%. If the proportion of each lithology is similar between the training and testing dataset, the accuracy in identifying the units increases significantly. Therefore, the use of a balanced dataset provides the best results for training the model.

Table 8: Specificity, sensitivity, precision and F-score results for the balanced dataset (Case Study 2)

Unit	Sensitivity (average %)	Specificity (average %)	Precision (average %)	F-score (average %)
Quaternary Detritals	96	98	95	95
Gravelly Siltstone	87	98	87	87
Calcareous Lacustrine Clay	33	100	100	50
Calcrete <20%	97	100	92	94
Calcrete >20%	100	100	100	100
Cemented Detritals	90	99	87	88
Silcrete	0	100	Undefined (2)	Undefined
Hematite Conglomerate	83	100	89	86
Clayey Breccia Colluvium	81	99	86	83
Lacustrine Clay	68	100	72	70
Canga	92	100	94	93
Bedrock	95	97	94	95

(1) The highlighted cells indicate results with <50% sensitivity, specificity, precision or F-score
(2) Metrics are undefined when the calculations result in a division by 0 and there are no positive predictions. units

11 CONCLUSION

The inclusion of machine learning into the model development process has proved successful in this trial. The random forest (RF) classifier algorithm demonstrated high accuracy when domaining the boreholes and is a useful method which can be used for future projects. Borehole domaining of the primary dataset was achieved with 82% overall model accuracy using the RF method. When training and testing the model built from the secondary dataset, 92% model accuracy was achieved with the inclusion of a balanced dataset.

The use of a balanced dataset when training and testing the model has proven to significantly increase the model accuracy. Rock mass units with small sample sizes such as calcareous lacustrine clay and silcrete do not have sufficient data to train the model and classify the units correctly. To improve the accuracy when predicting these units, additional twin geotechnical boreholes and RC boreholes are required to increase the data available to train the model. The success of the model largely depends on the size of the dataset and how balanced the dataset is. Therefore, collecting adequate/sufficient data to train the model is key for model accuracy and confidence.

This paper has demonstrated that an automated approach to identification of detrital units using geochemical data is an effective method for classifying boreholes. The approach to borehole domaining using machine learning can be successfully applied to a large borehole dataset and in turn can be used to develop a 3D model. This methodology will significantly reduce the time required for manual interpretation and allow the geologist/engineer greater time for detailed interpretation.

12 REFERENCES

Baxter, H. (2016). Geophysics, geochemistry and engineering geology: how disciplines combine to improve mine slope design in the Pilbara detrital valleys of Western Australia. Geological Society, London, Engineering Geology Special Publications, 27(1), pp.81-92.

Herget, G. (1977). Pit Slope Manual Chapter 2 – Structural Geology. CANMET (Canada Centre for Mineral and Energy Technology, formerly Mines Branch, Energy, Mines and Resources Canada), CANMET REPORT 77-41; Appendix A (pp 88).

Noble, D.P., Hemraj, D.J. and Eggers, M.J. (2021). Improving reliability of the geotechnical model for Cenozoic cover sequence sediments in the Pilbara, Western Australia (pp. 373-386). Australian Centre for Geomechanics.

Standards Australia. (2017). Geotechnical Site Investigations. AS 1726:2017.

Stone, M.S. (2005). Depositional history and mineralisation of Tertiary channel iron deposits at Yandi, Eastern Pilbara, Australia. University of Western Australia.