The Influence of Sample Numbers and Distribution on the Assessment of AMD Potential

A.M. Garvie , D. Kentwell

This paper was first presented at the 9th Australian Workshop on Acid and Metalliferous Drainage on 20-23 November 2017.

Abstract

The number and distribution of waste rock samples geochemically characterised before and during mine operation impacts the ability to accurately represent the waste characteristics and to predict the potential for acid and metalliferous drainage (AMD). Numerous regulatory and industry bodies recommend the number of samples that are to be characterised. Commonly, the practitioner is advised to take account of the complexity of the geology and the scale of the mine. However, in many instances, the scientific and statistical basis for the recommended numbers is either not provided or is not clear. Consequently, there is ambiguity as to how the recommendations should be applied to a particular mine at a particular stage of development and no quantitative method of gauging the accuracy of conclusions drawn, based on the approach taken to sampling is recommended.

This paper demonstrates how conclusions regarding the potential for AMD production at a mine can depend on the number of samples characterised and the specific samples selected for characterisation. Data from three mines are used to illustrate the impact of various sample numbers on preliminary conclusions related to the potential for AMD.

1. Introduction

Various Australian and international regulatory bodies require an initial assessment of the geochemistry of waste rock at the early stages of mine development. This often includes estimates of the locations and masses of potentially acid forming (PAF) materials and requires sampling of the waste rock zones and submitting the samples for a series of tests.

Government and industry guideline documents (BCAMDTF,1989; INAP GARD Guide 2017; US EPA, 1994; Aus. Gov. 2016, WA Gov. 2016, MEND 2009) provide guidance on sampling approaches to be taken so that the volumes and variability of parameters of interest (such as sulfur content and acid-neutralizing capacity (ANC)) can be quantified with some level of confidence. Perhaps the most comprehensive discussion is provided by MEND (2009), which states that it is important to provide good spatial, geological and geochemical representation, because contaminant discharge may be produced by only a portion of the geological material. Further, it states that sufficient numbers of samples should be taken to accurately characterize the variability and central tendency of the different waste materials, project components and geological units.

All the above guidelines indicate that the number of samples tested will depend on the stage of mine development and will be site specific. However, it not necessarily clear when enough samples have been tested. As an indication of the number of samples required, Aus. Gov. (2016) recommends testing up to several hundred samples at the pre-feasibility stage of mine development, whilst WA Gov. (2016) recommends that, as a starting point, 8 to 26 samples are used for disturbed waste units of mass between 100,000 and 1,000,000 tonnes, with the number of samples varying with the mass of waste. For a comprehensive presentation of the recommendations, the original documents should be consulted.

This paper presents assessments of the average values and variability of total sulfur content for nine geological units across three mines of different commodities in different geological settings to illustrate how these can vary with the number of samples. The units cover low, moderate and high sulfur scenarios.

2. Background

Many exploration and resource evaluation drill hole datasets include assays of total sulfur content. In preliminary assessments for the potential for acid and/or metalliferous drainage (AMD), it is common practice to use the distribution of sulfur as an indicator of potentially acid forming (PAF) material, because the total sulfur content is an indication of the maximum potential acidity. Often, a sulfur cut-off threshold is used to classify materials, whereby materials with sulfur contents below the threshold are considered to represent a low risk of acid generation. In this paper, we use a total S cut-off of 0.1%; however, identification of defendable sulfur cut-off requires site-specific assessment of the availability of acid neutralising capacity (ANC). Final AMD assessments are therefore supported by more rigorous geochemical characterisation that includes an examination of ANC and sulfur speciation within the mined materials.

Statistical methods were applied to large exploration drill hole data sets from nine units across three ore deposits (Table 1) to illustrate the relationship between sample numbers, sample locations, and levels of confidence on AMD assessment outcomes based on sulfur content.

Table 1. Types of ore deposits assessed

Ore Deposit	Type
Ore deposit A	Iron ore
Ore deposit B	Gold-copper
Ore deposit C	Copper gold porphyry

3. Statistical Analysis

The drill hole data sets were reviewed, and samples were grouped into units based on their lithology, alteration, and degree of oxidation. The number of drills holes per unit available for random sampling varied between 40 to more than 1000. Nine of these groups were selected to examine changes in the average sulfur content with increasing numbers of samples.

A sub sampling procedure utilising a random number generator was used to select samples randomly from the full data set (population) for each unit. Initially, four hundred samples were randomly selected, and sets of different sizes were selected from the 400 samples (Figure 1). The first set consisted of the first ten samples, the second set consisted of the first 20 samples and other sets were constructed in a similar manner for totals of 50, 100 and 200 samples. Note that the sets are therefore dependent sets in that the larger sets contain all the samples from the smaller sets. The average total S content was calculated for each set, the set of 400 samples and the population. This selection and calculation process, which generated a representation of the sulfur distribution, was performed four more times. Each representation was designated as Rx, where x had the values 1 to 5. The five representations are only a small fraction of the many thousands that are possible.

The construction of the larger sample sets performed by adding samples to smaller sample sets was expected to be representative of the manner that larger sample sets would be produced by practitioners. This approach can be contrasted with collecting a small sample set and independently collecting a larger sample set, which was not expected to be the approach taken by practitioners.

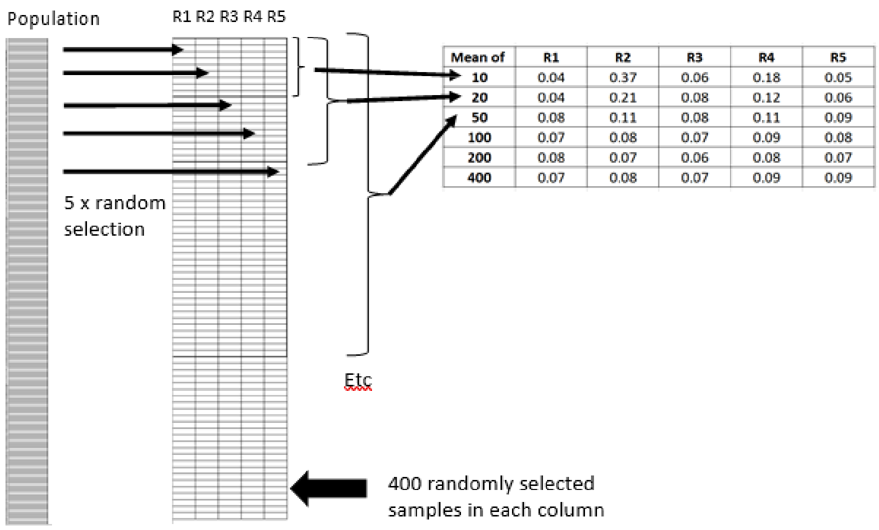


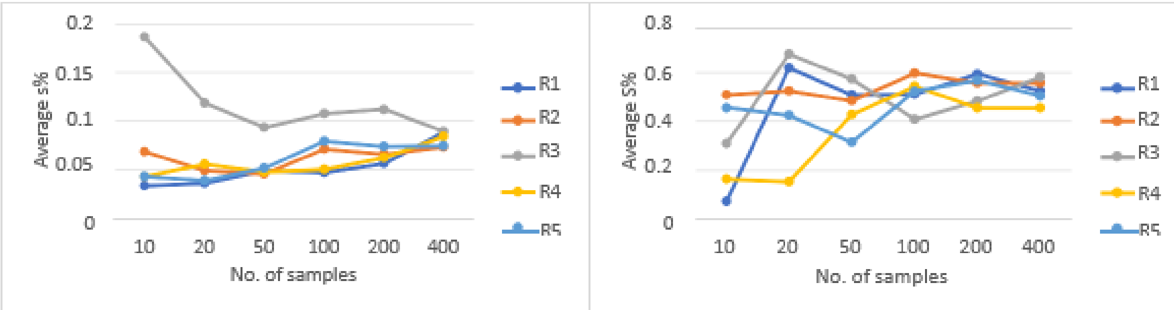
Fig. 1. Procedure for generation of realisations and calculation of means

Figure 2 presents results of five representations for one rock type for each of two ore deposits for sample numbers 10, 20, 50, 100, 200 and 400.

Figure 2a shows that, for Realisation R3, 50 samples are required to indicate that the rock type is NAF. It also shows that it is possible at higher sample numbers (R3) to conclude that the waste may be PAF, when the average total S value for the population (0.08) indicates that on average the rock type is NAF. Figure 2b indicates that, by characterizing only 10 samples, it may be concluded, based on realization R5, that the rock type is NAF, whereas the full average total S for the population indicates that the rock type may be PAF.

These examples illustrate the possibility that a rock type, based on an assessment of a small number of samples, may be inappropriately classified. This outcome may detrimentally influence early planning for waste management, for example, the volume of waste that needs to be managed as PAF may be over or underestimated.

Figure 2b also shows that the assessment of 20 samples produced upper and lower limits of the average total S content of 0.16 and 0.68% across the five realisations, a variation of a factor of more than four. Although these values are above the threshold, the lower value is less than one-third of the average total S content for the population (0.52). This may be significant, for example, when making water quality predictions, as the maximum potential acidity used in calculations would be less than a third of that of the population. In comparison, at more than 50 samples, the difference from the population average for all realizations is 20% or less, potentially leading to more reliable predictions.



a) Ore deposit A, rock type Lith3, 0.08 S%

b) Ore deposit C, rock type LITH06 1 O, 0.52 S%

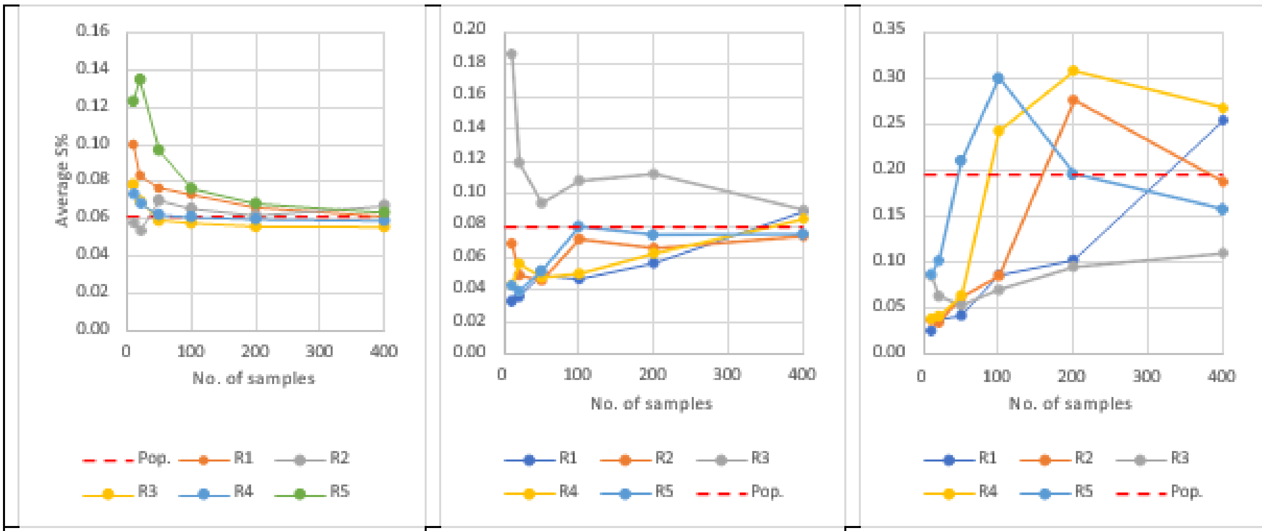
Fig. 2. Five representations for one rock type from ore deposits A and C

Note: The sub caption indicates the average total S content for the population.

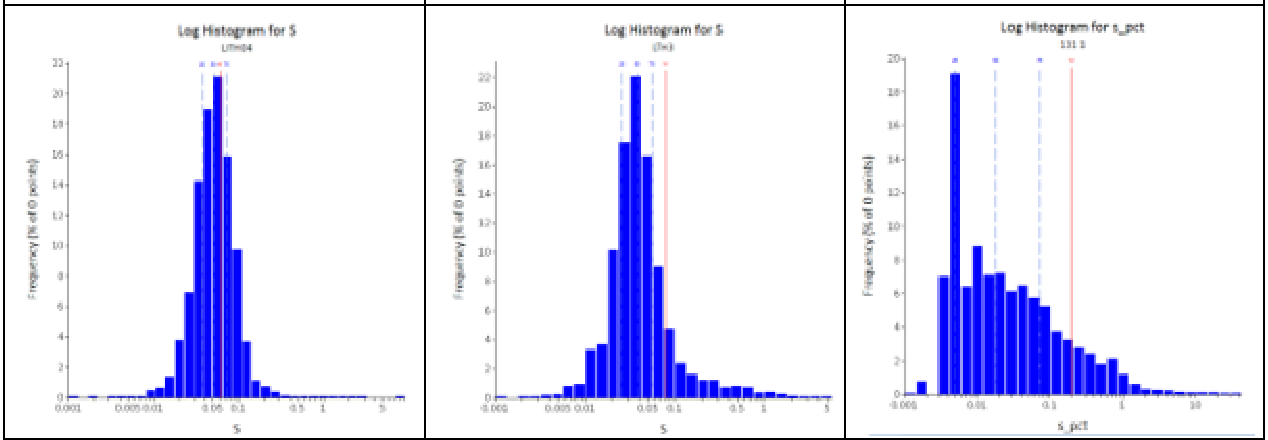
Figure 3 to 5 present statistics for five realisations of each of seven sample sizes for nine units. Eight of the nine plots indicate that, as the number of samples increases, the average values, minima and maxima tend towards the average total S content of the population, which is shown as a red dashed line. This is the expected trend; however, it is not exhibited in Figure 3c), and the causes are discussed below.

Figure 3a shows that, for one realization at 10 and 20 samples, the material is possibly PAF on average, whereas the average total S content for the population indicates that on average the rock type is NAF. This is also the case for at least one realization of 10, 20, 100 and 200 samples of Figure 3b. (Note that Figure 3b is an alternative display of data presented in Figure 2a.) In contrast, most realisations of Figure 3c for 1, 20 and 50 samples, and for three realizations at 100 samples, indicate the rock type 131 1 is NAF, when it is possibly PAF.

Figure 3d to Figure 3f present the total S distributions of the population for the rock types presented in Figure 3a to Figure 3c. The average total S content is provided in the captions. The total S distribution on the logarithm plot in Figure 3d is symmetric, centred approximately on the average value of the distribution and has limited low and high tails. For this case, the average total S contents of the realizations converge smoothly to the average S content of the population. The sulfur distribution of Figure 3f is markedly different. It has a high frequency of samples of total S content of about 0.005% and a long high total S content tail. These two factors would have contributed to the disparate and widening spreads in average total S content with increasing sample numbers of Figure 3c.



a) Ore Deposit A, Lith04 | b) Ore Deposit A, Lith03 | c) Ore Deposit B, 131 1



d) Ore Deposit A, Lith04, 0.06 S% | e) Ore Deposit A, Lith03, 0.080 S% | f) Ore Deposit B, 131 1, 0.196 S%

Fig. 3. Summary of realization statistics for low S rock types and total S distribution plots

A review of the spatial distribution of the total S contents of Figure 3f using the 3D geological modelling software Leapfrog Geo provided insight to the origins of the total S distribution. The high number of 0.005 S% values is real and for samples from a rock type of large extent, whilst the high S content values generally occurred in a specific part of the unit controlled by a series of interacting folds and faults that also host the high Au / Cu values.

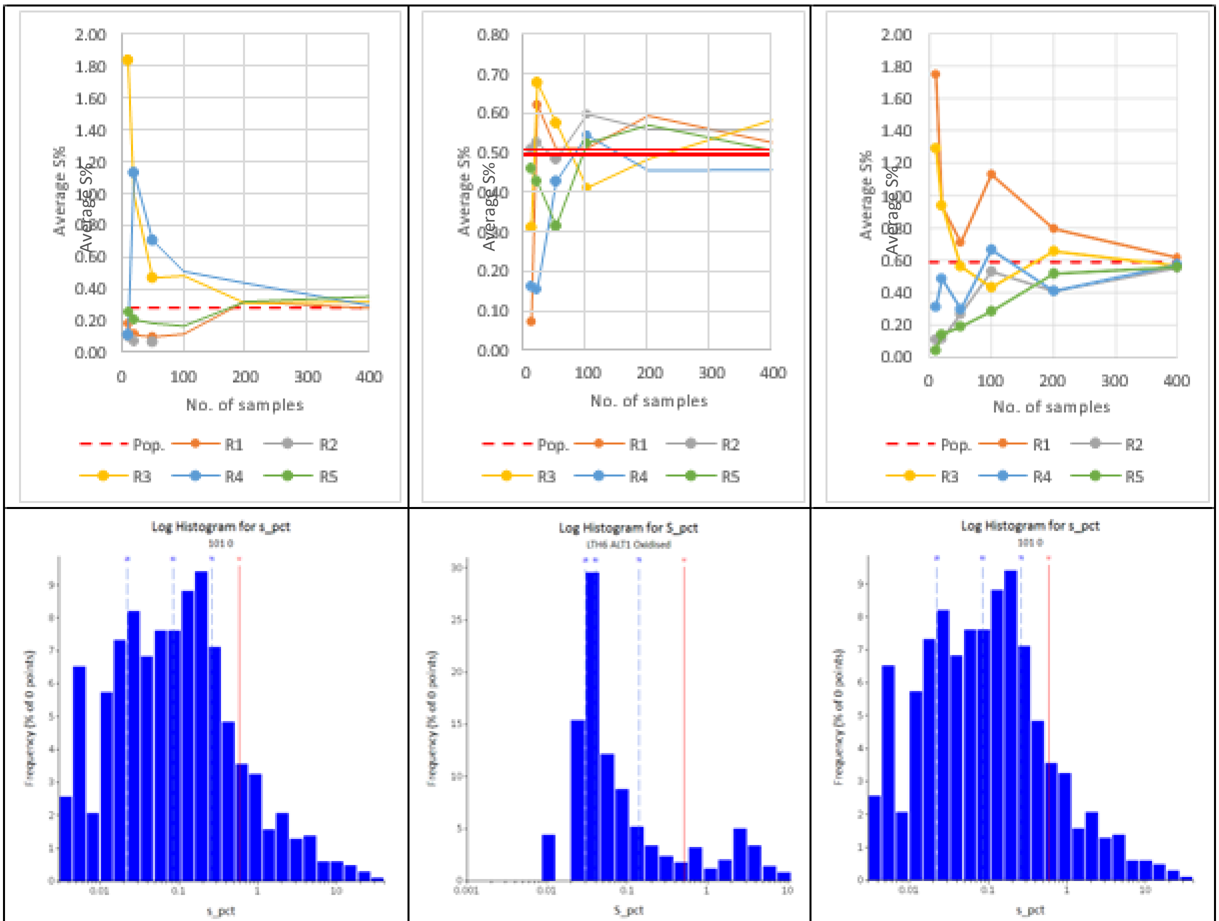
The potential for inappropriate conclusions regarding the acid generating capacity of this rock type was investigated further by generating 60 realisations as opposed to the 5 initially examined. Table 2 shows that for 50 samples, 35% of the realisations produced average total S contents less than threshold and that indicate that the rock type is NAF, whereas the population average indicates the rock type may be PAF.

Table 2. Ore Deposit B, 131 1 - results from 60 realisations

No. samples	10	20	50	100	200	400
No. realisations with avge S% <0.1	19	25	21	12	5	2
% realisations with avge S% <0.1	32	42	35	20	8	3

Thorough interpretation of sulfur data for this rock type would involve the production of sub domains and independent analysis of these to provide insight as to how differences in likelihoods of AMD production might be managed. Kentwell et al., 2016, upon investigating anomalies in total sulfur content realisations of a waste lithology at an iron ore deposit, showed that high total S samples originated in a spatially localized region. Having identified the presence and location of the high S samples meant that high S content waste could be selectively mined and managed separately. Similar considerations might be justified for unit 313 1.

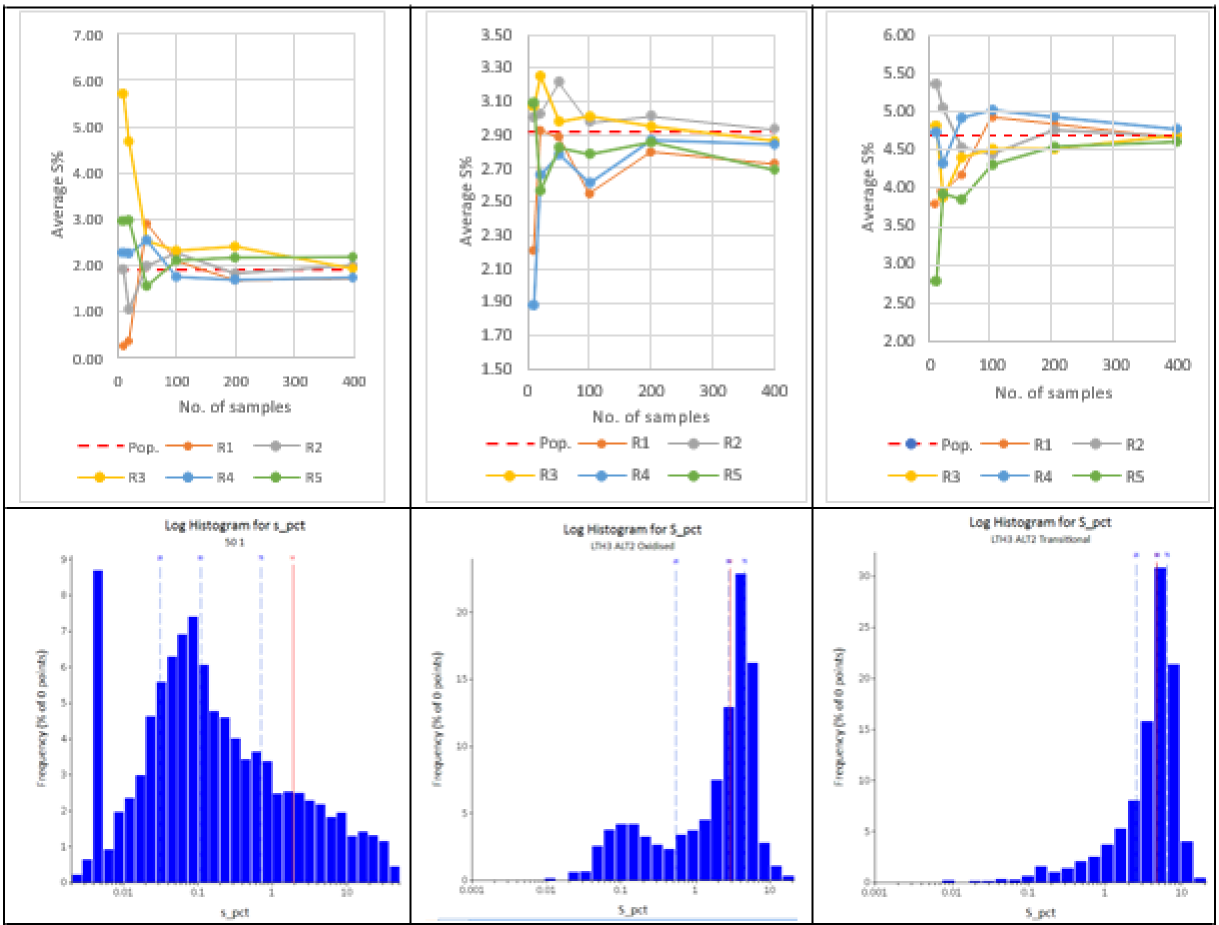
Figure 4 presents the summaries of realization statistics and total S distributions for units with medium level sulfur contents with average values for the populations above the 0.1 S% threshold. Figure 4a shows that, even at 100 samples, there is a realization with an average total S content less than the threshold. Figure 4b and 4c show that, although the units have a population average total S content more than five times the threshold, there can be realizations with small sample numbers that are below the threshold. Such realisations are likely to lead to the conclusion that the unit is NAF when it is potentially PAF.



a) Ore Deposit B, 131 0, 0.279 S%	b) Ore Deposit C, Lith6 Alt 1 O, 0.515 S%	c) Ore Deposit B, 101 0, 0.590 S%

Fig. 4. Medium S content units - summary of realization statistics and total S distributions

Realisation results are presented for high sulfur content units with three differently shaped distributions in Figure 5. The range of realisations at low sample numbers indicate that accuracy with which the maximum potential acidity can be estimated may be unacceptable, and 50 samples or more may be required depending on the accuracy required of AMD characterisation. The spread of the five realisations of Ore Deposit C, Lith3 Alt 2 Oxidised in Figure 5b is relatively narrow, even though the sulfur distribution is bi-modal. The results of further investigation of this unit using a total of 60 realisations are presented in Figure 6 and Table 1. The results indicate that the potential spread of average total S contents is larger than indicated in Figure 5b. For example, for 50 samples, the average total S content of 22% of the realizations differed from the mean value by more than 15%.



a) Ore Deposit B, 50 1, 1.914 S%

b) Ore Deposit C, Lith3 Alt 2 Oxidised, 2.916 S%

c) Ore Deposit C, Lith3 Alt 2 Transitional, 4.69 S%

Fig. 5. High S content units - summary of realization statistics and total S distributions

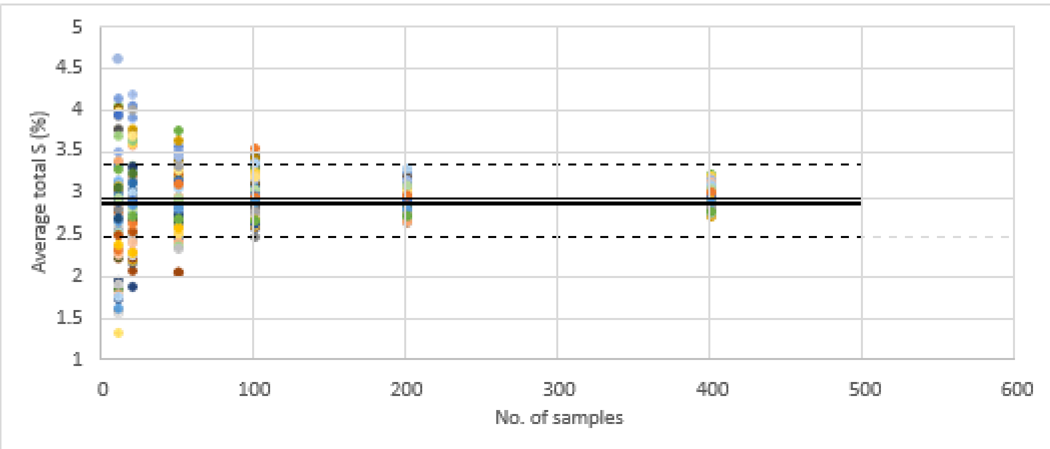


Fig. 6. Average total S content for 60 realisations for Ore Deposit C, Lith3 Alt 2 Oxidised

Note: The solid and dashed lines indicate average total S content of the population and +/-15%.

Table 3. Ore Deposit C, Lith3 Alt 2 Oxidised - Average total S spread

No. samples	10	20	50	100	200	400
Minimum	1.3	1.9	2.0	2.5	2.6	2.7
Maximum	5.0	4.2	3.7	3.5	3.3	3.2
No. of average total S contents within +/-15%	29	32	47	52	60	60
Percentage of average total S contents within +/-15%	48	53	78	87	100	100

Note: Results are for a total of 60 realisations, average total S content of population is 2.916%.

4. Conclusions

- Decisions based on a small number of samples are more likely to be inappropriate. For example, misclassifications are more likely for units with an average total S content near a threshold value used to distinguish a unit as NAF or PAF. This type of wrong decision may initiate the development of plans that could lead to mismanagement of the waste.
- Based on the nine units assessed in this paper, 400 samples or more may be required to estimate the average total sulfur content of a unit to within +/- 10%. From the nine units and five realisations, 45 individual tests were completed. The success rate (+/- 10% of population average) by number of samples for these 45 tests was as shown in the Table 4. Over a range of different geological conditions, even with 400 samples, +/- 10 % of the true average value was only achieved 71% of the time. It should be noted that this is still a very limited overall data set and that the conclusions drawn are representative of this data only. However, it shows that the minimum required samples for characterization of sulfur data averages from drillhole sized samples may be in the order of hundreds rather than tens.

Table 4. Success rate for given sample numbers

Number of samples	Success rate (%)
10	20
20	13
50	29
100	38
200	42
400	71

Note: Success means percentage of realizations with average total S within +/-10% of population value.

- The nature of the sulfur populations examined is such that their experimental average values and variability are dependent not only on the number of samples taken but on the locations of those samples.
- The exercise has shown that, below a certain number of samples, the same number of samples chosen at five different sets of random locations, can give an unacceptably wide range of averages and standard deviations and that the sample averages can be markedly different to the population average. The number of samples at which the sample statistics converge on the population statistics to an acceptable level varies depending on the geological characteristics of each particular unit and on how well the unit has been defined.
- In high sulfur cases, where all samples are above the selected PAF threshold and a PAF outcome is not in doubt, their remains the assessment of the quantity of potential acid production which also requires a level of confidence in the average and variability of a unit.
- As the number of samples increases, the average total S content tends towards the average total S content of the population; however, the trend may not be smooth and consistent for some units and sample sets.
- Generally, the number of samples required to estimate the average sulfur content of a unit depends on the specific samples selected, the sulfur distribution within the waste, and the required accuracy of the estimate.
- Provision of a detailed guideline for determining the number of samples to be used for a geochemical

assessment in specific geological settings is beyond the scope of this paper. However, it is recommended that practitioners identify the accuracy required in any assessment, develop an understanding of the total S distribution (possibly via the resource database) and estimate the likely level of accuracy that could be expected when interpreting data from the generally relatively small number of samples submitted for geochemical characterisation. Comparison of the likely and required accuracies could then guide the need to modify the number of samples to be geochemically characterised.

5. References

Aus. Gov. (2016) Preventing Acid and Metalliferous Drainage, Leading Practice Sustainable Development Program for the Mining Industry, September 2016, Commonwealth of Australia.

BCAMDTF (1989) Draft Acid Rock Drainage Technical Guide, Vol. 1, Prepared for the British Columbian Acid Drainage Task Force by Steffen Robertson and Kirsten (BC) Inc.

INAP GARD Guide, http://www.gardguide.com/index.php?title=Main_Page, accessed 25/08/2017.

Kentwell D, Garvie A and Linklater C. (2016) Geostatistics and Sample Numbers from and Acid and Metalliferous Drainage Perspective, Proceeding of the Life of Mine Conference, 28-30 September. Brisbane, AusIMM SMI/CLMR.

MEND (2009) Price, W.A., Prediction Manual for Drainage Chemistry from SulphidicGeologic Materials, MEND Report 1.20.1, Prepared by CANMET – Mining and Mineral Sciences Laboratories, Smithers, British Columbia.

US EPA (1994) Acid Mine Drainage Prediction, US Environmental Protection Agency (EPA) Report number EPA-530R-94-036.

WA Gov. (2016) Government of Western Australia, Draft Guidance: Materials Characterisation Baseline Data Requirements for Mining Proposals, March 2016, Western Australian Government, Department of Mines and Energy.