Adequacy of Sampling and Volume Estimation for Pre-mining Evaluation of Potentially Acid Forming Waste: Statistical and Geostatistical Methods

D. Kentwell[1], A. Garvie[2], J. Chapman[3]

1. Principal Consultant (Resource Evaluation) FAusIMM, SRK Consulting, Level 8, 365 Queen St, Melbourne, VIC 3000; email: dkentwell@srk.com.au
2. Principal Consultant (Water and Environment), SRK Consulting, Level 2, 44 Market Street, Sydney, NSW 2000; email: agarvie@srk.com.au
3. Principal Consultant (Water and Environment), SRK Consulting, Suite 7-9, Level 6, 141 Queen St, Brisbane, QLD 4000; email: jchapman@srk.com.au

This paper first appeared in a modified form in Proceedings of Life-of-Mine Conference, 10-12 July 2012, Brisbane, Australia (The Australasian Institute of Mining and Metallurgy: Melbourne).

Abstract

To assess the risk of acid and metalliferous mine drainage (AMD), most new mining projects require an initial assessment of the locations and volumes of potentially acid forming (PAF) waste that is likely to be moved and dumped in waste stockpiles. This requires sampling of the waste rock zones and submitting the samples for a series of tests. Key to the outcome of the waste characterisation process is obtaining sufficient samples that are representative of the various rock types, both in number and spatial distribution.

Various regulatory bodies throughout Australia and the rest of the world are currently struggling with how to define minimum amounts of sampling to achieve realistic estimates without imposing excessive costs and time constraints on a project.

This paper addresses two aspects of the process:

1. What constitutes a 'sufficient' number of samples
2. How to use those samples in determining volumes of likely PAF material

Geostatistical methods of characterising ore bodies with large numbers of samples are routinely used in the mining industry by geologists and geostatisticians to estimate and classify mineral resources. In theory, the same methods can be used to estimate PAF waste locations and volumes.

In practice, the number and spatial distribution of dedicated PAF waste samples at feasibility stage is typically so small that it is almost impossible to estimate the spatial distribution of PAF material with any degree of confidence. Consequently, the common approach is to characterise a few samples from a lithological unit and then assume that they represent the unit as a whole. This approach, however, can be fraught with risk.

This paper describes some of the findings of SRK's evaluation of several coal projects with respect to PAF waste evaluation and highlights some common misconceptions.

Introduction

There is a general recognition among regulators that sampling waste rock for characterising the potential for acid and metalliferous mine drainage (AMD) should be undertaken in such a way that the levels, volumes and variability of parameters of interest (such as sulphur content and acid neutralising capacity (ANC)) can be quantified with some level of confidence.

The former Queensland Department of Environment and Resource Management (DERM) (now Department of Environment and Heritage Protection (DEHP)) Guidelines – Assessment and Management of Acid Drainage (DERM 2012) states:

"The prediction of acid drainage typically requires an iterative process of preparing geological models, examining comparative information, sampling, analysis and classification, similar to the process used to determine other geological characteristics such as ore reserves.

[Sampling] is a critical phase of the prediction program. Sufficient samples should be taken to accurately characterise the nature, distribution and variability of critical parameters in each material type, including waste rock and ore samples. Ideally this should be based on accepted statistical procedures."

The US Environmental Protection Agency (USEPA) - Technical Document, Acid Mine Drainage Prediction (EPA 530-R-94-036 NTIS PB94-201829) December 1994 (USEPA 1994) states:

"Selection of samples has important implications for subsequent acid prediction testing. The purpose of testing rock material is to allow classification and planning for waste disposal based on the predicted drainage quality from that material. Samples must be selected to characterize both the type and volume of rock materials and also account for the variability of materials that will be exposed during mining.

Sampling techniques used to evaluate recoverable mineral resources (assay samples) are similar to those required for prediction of acid generation potential."

In the following discussion, we address sampling and estimation of the proportion of PAF within a single rock type and single weathering state. Modelling the overall rock type volume, or issues of mixing rock types, and / or weathering states are outside the scope of this paper.

Why geostatistics?

At this point, we need to distinguish between the different stages of a project. At the early planning and pre-feasibility stages, typically very few waste samples are selected for characterisation. In some cases, the number of samples characterised increases as the project moves into feasibility. However, the number of samples remains small in comparison to the resource database, even though the mass of waste to be mined often is several times that of the resource. As a project moves into production, many closely-spaced samples from blast holes are often available (assuming no visual controls such as colour are used), and the potential exists for much better waste characterisation and definition; often, however, this opportunity is not used. In this paper, we are specifically referring to pre-mining evaluations.

Classical statistical tests such as confidence intervals on the mean, and analysis of variance (ANOVA and variants), can be used to determine adequacy of sampling where parameters (such as total sulphur) are either well above, or well below, thresholds of interest. The total unit or volume of that rock type is then classified as PAF or non-acid forming (NAF) and can then be dealt with as required. Geostatistically, this is known as having a 'global' understanding of a sample distribution.

Determining the volume of rock within each rock type with parameter values above or below specific thresholds is more problematic, in part because characterising the spatial distribution of parameter values within the rock category requires more samples than determining simple averages for the whole rock type. Geostatistical methods can help determine the number of samples required. Understanding the spatial distribution is known within the field of geostatistics as having a 'local' knowledge of the distribution of the parameter value.

A good summary of the use of geostatistics for waste characterisation is provided in Kleinmann (2000).

How many samples?

There are two questions that need to be answered in order to determine how many samples will be sufficient:

1. What is the required level of confidence of the global mean and global variability (variance) of the parameters of interest for the rock type?
2. How important is the local (spatial) variability of the parameters of interest across the entire volume of each rock type?

The required level of confidence for the global mean and variability depends on many factors including the stage of the project and the sensitivity of the receiving environment to AMD. Generally, the required confidence increases as the project moves from conceptual to pre-feasibility to feasibility. For example, should acid generating materials be present, and should specific management or treatment methods be required to prevent acid generation, the confidence level of the assessment needs to be commensurate with the engineering and costing level. For a pre-feasibility assessment, this might be ± 25%, for example.

The second question relates to the significance of spatial variation to the waste management strategy. A good understanding of the spatial variability can be critical to the development of management strategies. For example, a small proportion of acid generating waste within a dump could lead to acidic seepage from the dump as a whole, where the balance of the waste rock either has no ANC or the ANC is not available for acid neutralisation. In this case, segregation of acid generating materials may be essential for the success of the waste management strategy. Consequently, a high level of confidence in the spatial variability is required so that the practicality and feasibility of segregation at the mining scale can be evaluated. Where short range variability within a production unit is high, a management strategy relying on segregation may not be successful.

AMD literature contains various simplistic tonnage / volume-based methods and formulae that attempt to generically address the question of the minimum number of samples required to adequately characterise the acid generating properties of rock types. There is also acknowledgement in the literature that every project and every rock type will differ, and each case should be examined individually and iteratively.

For example, the guidelines provided by DERM (DEHP) are as follows:

"*Given the variation in mineral types and locations it is important to establish what is an adequate sampling intensity. The usual procedure is to collect a large number of samples and to undertake a simple screening process. The number of samples required and the sampling intensity will depend on a number of factors including:*

- *geological variability and complexity in rock types, information and experience from adjacent or geologically comparable mine sites,*

- *potential for significant environmental or health impacts,*
- *the size of the operation,*
- *relative costs,*
- *statistical requirements which ensure samples are representative,*
- *the volume of each class of waste rock type, and*
- *the level of confidence in predictive ability.*

The minimum number of samples collected from each rock/overburden type during initial sampling should generally be as follows (Table 1):

Table 1: Sample guidelines

Mass of Each Separate Rock Type (tonnes)	*Minimum Number of Samples*
<10,000	*3*
<100,000	*8*
<1,000,000	*26*
<10,000,000	*80*

The actual number of samples will be project specific."

The USEPA guidelines recommend the following:

"*There are many opinions concerning the number of samples to be collected in a fixed-frequency sampling program. One mining consulting firm recommends about 8 - 12 samples of each significant rock type or 1 sample for each 1 million tons, at a minimum (Schafer 1993). In this case a significant rock type represents one or two percent of the total mine rock volume. Gene Farmer of the U.S. Forest Service suggests that one sample (about 1,500 grams) be collected per 20,000 tons of waste rock, or about 50 samples for each 1 million tons (USDA Forest Service 1992).*

The British Columbia AMD Task Force recommends a minimum number of samples based on the mass of the geologic unit. Their recommended minimum sample number is 25 for a 1 million ton geologic unit, or one sample for every 40,000 tons. Using the British Columbia method, as waste volume increases, the number of samples decreases."

In all cases, the basis for the number of samples recommended, be it statistical or other, is not provided (USEPA).

"*There are therefore reservations to prescribing a fixed number of samples for collection per volume of material. Fixed frequency sampling does not encourage the use of best judgment on the part of the sample collector (typically a mining company)*".

In comparison to resource estimation datasets, usually consisting of thousands to millions of samples, early stage waste characterisation datasets often contain less than 100 samples spread across several rock types. In many cases, this is not even sufficient to provide confidence in the global statistics let alone assess the spatial variability.

Assessing global statistics

There are various parametric (classical) and non-parametric (Monte Carlo) methods of calculating confidence intervals on the mean of an experimental distribution at any given confidence level. This requires a decision to be made on what is an appropriate interval (e.g. 75%, 95%, 99%) for acceptance of what constitutes a 'sufficient' number of samples. The dashed blue lines in Figure 1 show an example of a 95% confidence interval around the mean. As the confidence interval reduces, from that shown in the figure, the confidence level decreases. As the confidence interval increases so does the confidence level that the mean lies within the interval.

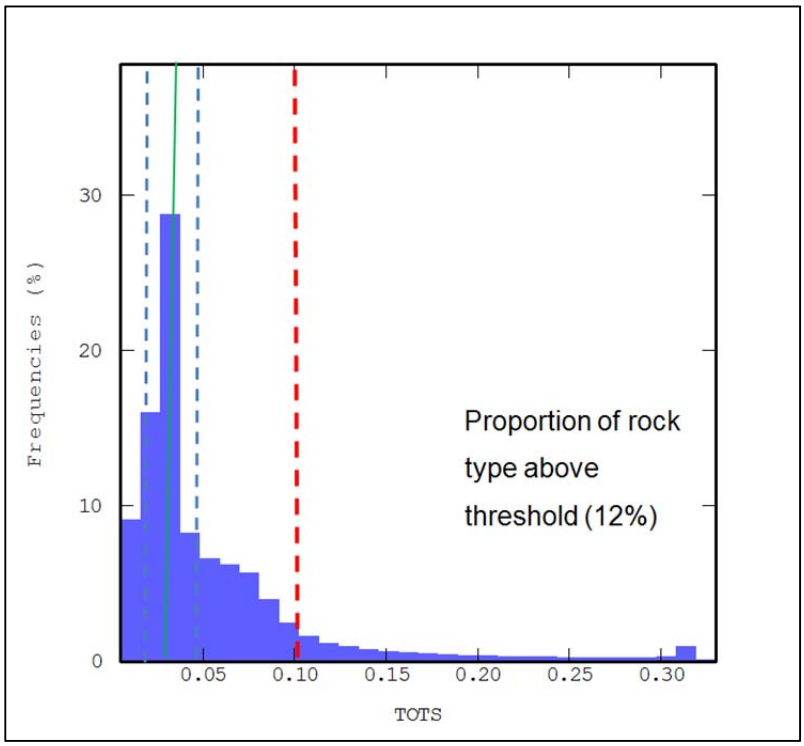

Figure 1: Histogram of sandstone unit

If the threshold of interest for a parameter lies well outside the confidence interval on the mean, then the number of samples is adequate for global, whole of rock type, decisions to be made. For total sulphur, for example, if the threshold of interest is 0.1% and the mean of the sample distribution is 2.7%, with a 95% confidence interval between 1.2% and 4.9%, then the rock type can confidently be classified as PAF overall (in the absence of any ANC). If the mean total sulphur content of the sample distribution was 0.025%, with a 95% confidence interval from 0.015% to 0.040% (as for the sandstone unit in Figure 2), then the rock type could confidently be classified as non PAF overall.

If the mean total sulphur of the sample distribution is 0.082%, with a 95% confidence interval from 0.04% to 0.16% (as for the mudstone unit in Figure 2), then there is insufficient confidence to make a global PAF / NAF decision, and more samples would be required to narrow the confidence interval.

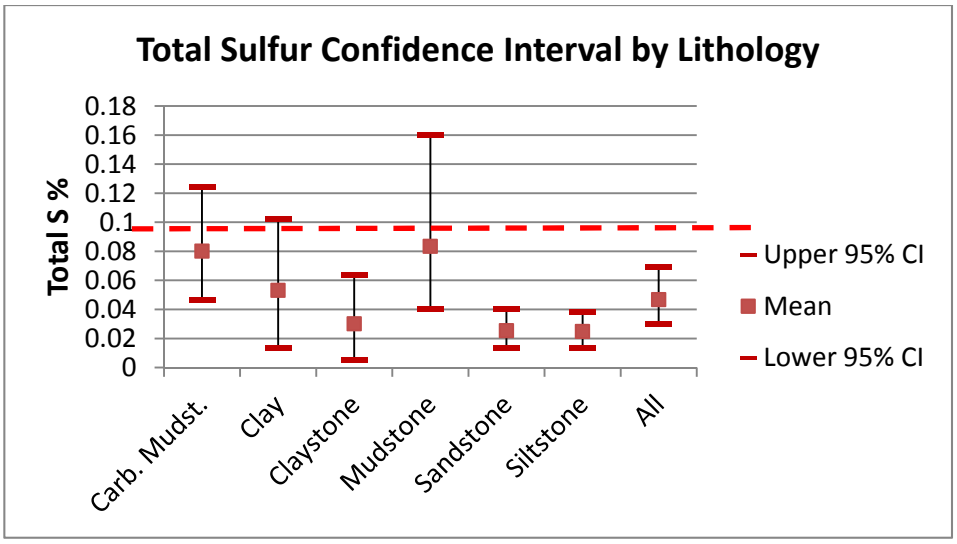

Figure 1: Example of 95% confidence intervals for the waste lithologies of a coal deposit

The actual number of samples required will be dependent on the required confidence level, the threshold of interest and the specific chemical and geological characteristics of the individual rock type (or in mathematical terms, dependent on the required confidence interval, mean, variance and skewness of the distribution). Refer to the note at the end of this paper on skewed distributions and non-parametric tests.

In the case where the confidence interval on the mean is below the threshold of interest, even if the confidence level has been met, there may still be a proportion of the sample distribution that is above the threshold. That is, the maximum value of the actual rock type distribution and some portion of the upper tail of that distribution may still be above the threshold. This portion of the rock type is PAF. Figure 1 shows the histogram for the sandstone unit from Figure 2 with the mean (green line) and the threshold of interest (red dashed line). The proportion of the distribution above the threshold of interest is 12%.

Support / scale issues

To help understand the distribution of parameter values such as total sulphur within rock types, the probability of a parameter value associated with a sample being greater than any chosen threshold can be calculated from the available samples.

Experimental probabilities of a sample being above any selected threshold can be obtained from the inverse cumulative distribution of the samples in each group or lithology. A larger the number of samples produces a larger the confidence in the probabilities. Probabilities can be equated to proportions in a volumetric sense.

These probabilities are usually derived from sample volumes obtained from drillholes. Probabilities and proportions based on sample size volumes can be somewhat misleading, as the waste will not be mined with equipment with the selectivity as small as that of a drill rig (i.e. the size of the drillcore that is generated and sampled). Mining selectivity is more likely to take place at the scale of an excavator bucket or truckload. The size of the selective mining unit should be taken into consideration as this will change the variability of the parameter under consideration.

In geostatistical terms, the size of the unit at which final selection takes place is referred to as ‘support’.

As the selective unit size (block size) increases, any distribution of a material property will lose the highest highs and the lowest lows due to a “mixing” effect. This compresses the histogram of the distribution as the

support size increases. The effect is termed the 'volume variance relationship' and often is referred to as the support effect (Figure 3).

Reducing the selectivity to mining unit size changes the probability of a parameter value exceeding any particular threshold value.

The jump from sample scale support to some likely block size is the most critical as it produces the largest change in distribution shape. For AMD purposes, the difference between distributions of parameters such as total sulphur at different block sizes is likely to be less critical. The change in histogram shape between different block sizes is much less pronounced than the change in histogram shape between sample size and block size, as the proportional change in volume is much less.

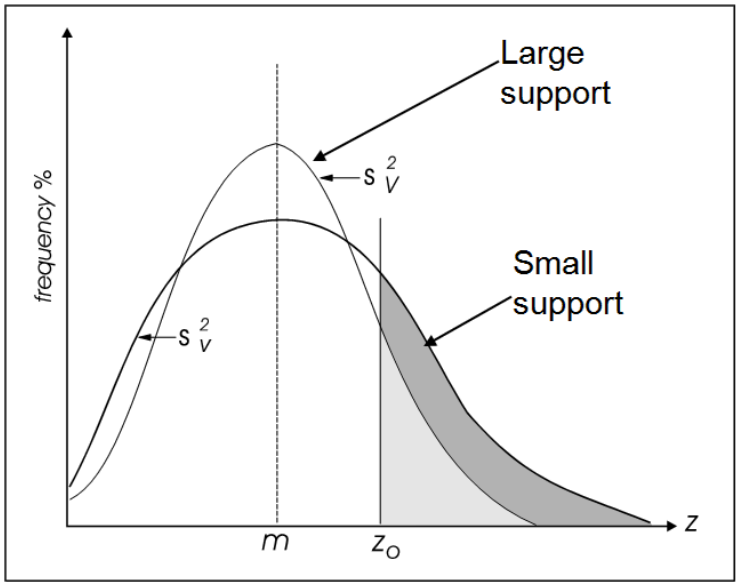

Figure 3: Effect of volume size (support size) on histogram

Note: (m=mean, z_0 = threshold, s^2_v = small support, s^2_V = large support)

Example of the effect of change of support

Consider the example below (Figure 4, Table 2) where we have 16 samples each representing 1 cubic metre of rock, analogous to sample sized support.

Using the 1.0% as a threshold of interest, the results indicate that 7 m^3 of waste would be below the threshold (say, NAF), whereas the other 9 m^3 would need to be managed as acid generating materials.

If the volumes are amalgamated so that the size of the support is larger, i.e. to represent 4 m^3 analogous to block support, the outcome would be as shown in Figure 5 and Table 3. On this basis, one of the units (blocks) representing 4 m^3 would fall below the threshold, and 12 m^3 of the waste would need to be managed as net acid forming (i.e. the volume of material classified as NAF would decrease to only 4 m^3, leading to a loss in opportunity for selective management of the waste; conversely a proportion of the NAF material has been 'contaminated' so that it had to be reclassified).

Consider now the example (Figure 6, Table 4) showing the extreme case where all material is amalgamated and only a total average is the basis for the PAF / non-PAF decision. All of the material would be classified as acid generating.

In the first case where only sample sized selectivity is involved, the proportion of PAF material is underestimated at block selectivity scale. In the last case, the opportunity of minimising waste tonnages that need to be managed as PAF is lost. However, this needs to be balanced with the cost involved in obtaining and implementing selectivity at a block scale.

Figure 4: Example block grid and assay results

0.2	0.7	2.2	1.4
0.4	1.3	0.6	0.3
1.1	0.6	1.3	1.2
2.3	0.9	1.7	1.9

Note: (m=mean, z_0 = threshold,
s^2_v = small support, s^2_V = large support)

Figure 5: Amalgamated grid

0.65	1.125
1.225	1.525

Figure 6: Global average

1.131

Table 2: Volume and average above thresholds for 1 m^3 blocks

Threshold	Volume	Average grade
0.5	13	1.32
0.75	10	1.53
1	9	1.60
1.25	7	1.73
1.5	4	2.03
1.75	3	2.13
2	2	2.25

Table 3: Volume and average above thresholds for 4 m^3 blocks

Threshold	Volume	Average grade
0.5	16	1.13
0.75	12	1.29
1	12	1.29
1.25	4	1.53
1.5	4	1.53
1.75	0	
2	0	

Table 4: Volume and average above thresholds for global selection

Threshold	Volume	Average grade
0.5	13	1.32
0.75	10	1.53
1	9	1.60
1.25	7	1.73
1.5	4	2.03
1.75	3	2.13
2	2	2.25

Application of change of support

The application of the concept of change of support requires some careful thought in relation to AMD.

In the case of ore mining, we have a sample support volume determined by the drillcore at a particular length of sample. Via the variogram and some estimation methodology, the sample populations can be converted to a block population based on an assumed mining selectivity volume, for example, a 20 x 20 x 5 m volume. The statistical aspect of the grade of that block of material assumes that when that volume reports to the processing plant, the entire block is processed and (subject to the plants recovery limits), and all of the elements of interest is extracted.

This is not necessarily the case for potential AMD material in a dump. For example, if an *in situ* block of waste before extraction and transport to a dump contains both high and low AMD parameter values at drillcore (or sample) scale, it could take only one of those high drillcore-sized samples within the entire block volume to cause a potential AMD problem within the dump (assuming, for now, that there are no neutralising components).

How much mixing, and therefore dilution, should be assumed to have been caused within the same rock type, of the sample scale material, by extraction, transport and placement? This can be further expanded to whether or not the support size should be the entire dump, or may be based on the size of each 'water catchment area' under the dump?

The question is important because, where a rock type contains values both above and below critical thresholds, the scale of 'selectivity' (mixing) equates to the amount of dilution, and changes the volumes / proportion of the population, above any particular threshold.

Of course if a rock type has a minimum sample parameter value above the critical threshold, or a maximum sample parameter value below the critical threshold, the scale of selectivity is irrelevant.

The support size chosen therefore depends on the proposed management plan. The outcome will be different for the case where all waste, regardless of whether it is PAF or not, will be placed in one dump and the drainage water managed, compared to the case where the PAF waste is selectively mined to be taken to a specific location.

If the waste is selectively mined, the scenario is the same as that for ore mining, and the appropriate support size should match the selectivity of the mining equipment being used and the geometry of the waste volume being excavated. The support volume therefore could be in the range of a truckload to a mining block.

Either way, estimation of volumes above thresholds should not be done at sample scale.

This is all very well if exhaustive close-spaced sampling is available. However, this is seldom the case for waste characterisation at the pre-mining stage of a project. In fact, it may be that only two or three samples would have been available. Assuming, for argument's sake, (in the example cited in the grid block shown in Figure 4 using a 1% sulphur threshold) that the three samples coincidently comprised the 0.2, 0.3 and 0.6% sulphur block samples, the global average would be 0.37% and the 16 m^3 would be classified as NAF. Since the material would require no special handling or management, there would be no PAF material management costs assigned to the mass. The risk is that with the misclassification, the materials will be improperly handled and AMD will occur during operations, or even after closure. Costs associated with the

mitigation of AMD can be substantially higher than the management costs to prevent AMD in the first instance, not to mention the risk to the environment.

During production with better sampling intensity (e.g. utilising blasthole samples), the material characteristics could be correctly identified with higher resolution. The actual proportion of AMD waste rock could be more accurately identified and the costs associated with the special management requirements estimated. In addition, the risk of AMD would be mitigated and the potential costs that may have been incurred in closure or treatment reduced. Similarly, samples may equally have represented the high end numbers, in which case, the outcome would have been that the entire 16 m^3 would attract PAF material management costs. The potential benefits from the proportion of NAF materials would have been missed. There is therefore a clear case for understanding the representativeness of samples and spatial variability in the waste.

Variography is a geostatistical tool that characterises the spatial variability and can be used to predict the change in the shape of a distribution with a given change of support, utilising a relatively small dataset. Global change of support calculations, which apply to the overall dataset, can be done utilising a much smaller dataset than is required for local block or point estimations which can be calculated using methods such as Kriging. Block Kriging is the typical method used for mineral resource estimation. If there are enough samples to enable a variogram to be modelled, we can begin to model the effects of change of support and attempt to model the likely volumes of individual rock types above thresholds of interest.

Variography

The average value and variance of a population can often be estimated to a prescribed level of confidence globally from a relatively small number of samples. Confidence in a global average at a particular level is not always sufficient, and knowledge of the localised values of a parameter and its local variability are often required for waste characterisation and mine planning purposes.

The fundamental basis of most geostatistics is the experimental semi-variogram, $\tilde{\gamma}(h)$ (hereafter referred to as the variogram). The variogram is the basic diagnostic tool for spatially characterising a property (z), and measures the average dissimilarity between data separated by a distance, h. It is also central to geostatistical estimation or interpolation methods (Kriging), and the more advanced methods of change of support and conditional simulation.

An experimental variogram is computed as half the average squared difference between the components of every data pair. That is $\tilde{\gamma}(h)$ is given by:

$$\tilde{\gamma}(\boldsymbol{h}) = \frac{1}{2N(h)} \sum_{\alpha=1}^{N(h)} [z(\boldsymbol{u}_\alpha) - z(\boldsymbol{u}_\alpha + \boldsymbol{h})]^2$$

Here $z(u_\alpha)$ and $z(u_\alpha+h)$ are the values of the pairs of points and the summation is over all N pairs separated by distance, (h). In general, h is a vector, having both an amplitude and direction.

Another way of saying this is that the experimental variogram is the plot of the separation distance (h) of paired points against the variance of the property value for all pairs of samples separated by distance, h. An idealised variogram is shown in Figure 7.

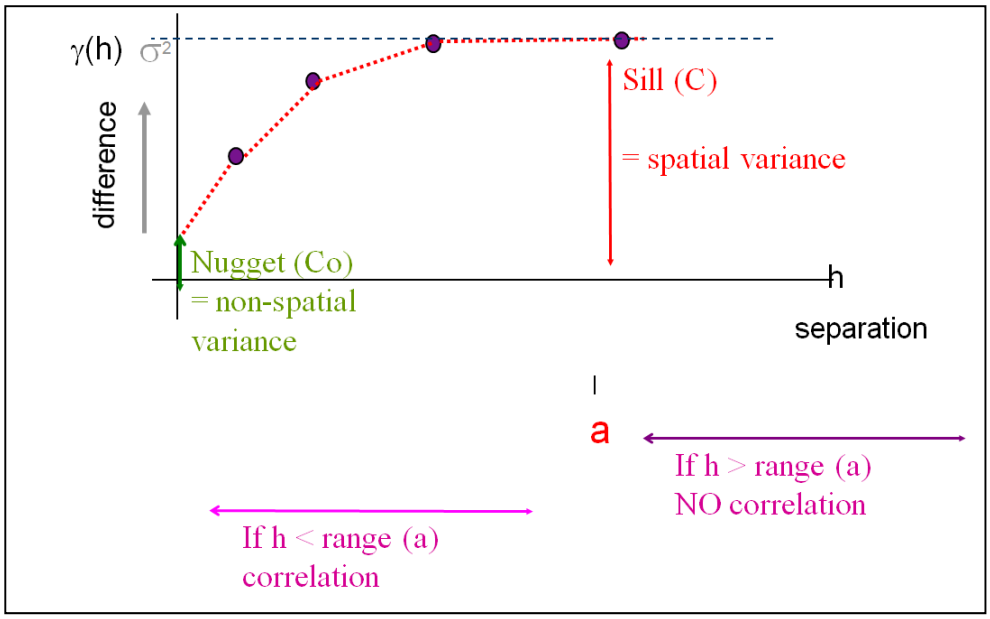

Figure 7: Variogram components

Spatial correlation exists when the average variance of pairs of data at low separation distances (compared to the dimensions of the area sampled) is smaller than the variance of the entire population. In simple terms, things close together tend to be similar; things far apart tend to be different.

At some sufficiently large separation distance, known as range (a), the average variance of the pairs will be approximately the same as that of the overall population σ^2, indicating that spatial correlation no longer exists beyond that distance.

For example, coal resources typically have coal quality ranges in the order of hundreds to thousands of metres, iron ore resource chemistry will typically have ranges in the order of tens to hundreds of metres and gold resources typically have gold content ranges in the order of metres to tens of metres.

These characteristic ranges are in line with our intuitive understanding of the continuity of geology and commodity content / quality distributions within various different resources. They also reflect the typical drill spacings used to estimate the resources in these various types of deposits.

A variogram that rises consistently from low separation distances and then flattens is a typical form of spatial correlation. A variogram that is flat or very irregular from low separation distances indicates lack of spatial correlation or insufficient data.

If experimental variography does not show any structure or ranges beyond the smallest sample spacing, then no spatial correlation exists between samples at or above the smallest sample spacing. This means that values of parameters at unknown locations cannot be interpolated or extrapolated with any confidence, and the mean of all of the samples in the domain is the best estimate at any unknown location.

How many samples for spatial characterisation?

To be able to model variograms, both sufficient numbers of samples and appropriately-spaced samples are required. If a variogram cannot be modelled, then more samples and closer-spaced (infill) samples, are required to enable understanding of the spatial distribution of the parameters within the rock type, and change of support calculations to be made.

The actual number of samples will be dependent on the specific chemical and geological characteristics of the individual rock type.

In order to minimise the initial sampling requirements, it is best to have wide-spaced coverage over the whole extent of the rock type to capture the full variability of the parameters of interest, as well as some closer-spaced sampling in one or two local areas to try and capture the short-scale variability and to assess the range of the variogram(s).

Assessment of the ranges of variability also enables us to calculate the requirements for amounts of waste sampling during production. Cost benefit trade-offs for selective or non-selective handling of PAF material can then be examined.

SRK is in the process of building up its knowledge of typical variogram ranges for different AMD parameters in different waste rock types normally associated with various commodities. These can then be used as initial sampling densities on new projects where selective waste handling may be required.

Coal case study

Change of support results from Hancock Prospecting's Kevin's Corner and Alpha projects are shown in Table 5 to Table 8. These demonstrate the potential differences in volumes (probabilities can be equated to proportions / volumes on a global scale) above thresholds that can occur when implementing the change of support from sample size to block size. For example, at the drillcore sample size, 54% of the Fresh Carbonaceous waste is expected to have a total sulphur content above 0.10%; whereas, at the assumed mining block size of 100 x 100 x 2 m, 100% of the Fresh Carbonaceous waste is expected to have a total sulphur content greater than 0.10%.

Figure 8 shows actual ANC and total sulphur variograms for coal overburden indicating ranges in the order of 2,000 m.

The full Alpha and Kevin's Corner reports can be found on the Hancock Prospecting website (SRK, 2010; SRK, 2011). These include variography and statistical summaries.

Table 5: Kevin's Corner - Probabilities of the total sulphur content of an individual drillcore sample being above specified thresholds

Group / lithology	Number of samples	Sulphur thresholds (%)				
		0.10	0.20	0.30	0.50	1
Weath all ex coal	134	8	3	2	2	0
Fresh all ex coal	138	16	7	6	2	0
Fresh Carbonaceous*	17	54	24	22	16	0
Fresh REM	120	11	5	4	0	0
Fresh Coal *	10	66	29	22	8	0

Table 6: Kevin's Corner - Probabilities (%) of the total sulphur content of a 100 x 100 x 2 m block being above specified thresholds

Group / lithology	Number of samples	Sulphur thresholds (%)				
		0.10	0.20	0.30	0.50	1
Weath all ex coal	134	7	4	3	2	1
Fresh all ex coal	138	23	7	3	0	0
Fresh Carbonaceous*	17	100	79	3	0	0
Fresh REM	120	14	5	2	0	0
Fresh Coal *	10	100	52	22	2	0

Table 7: Alpha - Probabilities (%) of the total sulphur content of an individual drillcore sample being above specified thresholds

Group / lithology	Number of samples	Sulphur thresholds (%)			
		0.10	0.30	0.50	1
Weath all ex coal	105	9	0	0	0
Fresh all ex coal	117	17	6	1	1
Fresh Carbonaceous*	27	41	27	8	3
Fresh REM	90	12	1	0	0
Fresh Coal	55	80	50	14	0

Table 8: Alpha - Probabilities (%) of the total sulphur content of a 100 x 100 x 2 m block being above specified thresholds

Group / lithology	Number of samples	Sulphur Thresholds (%)			
		0.10	0.30	0.50	1
Weath all ex coal	105	0	0	0	0
Fresh all ex coal	117	23	1	0	0
Fresh Carbonaceous*	27	92	6	0	0
Fresh REM	90	0	0	0	0
Fresh Coal	55	100	39	0	0

*insufficient samples for 95% confidence that the mean is above 0.3%S or below 0.2% S

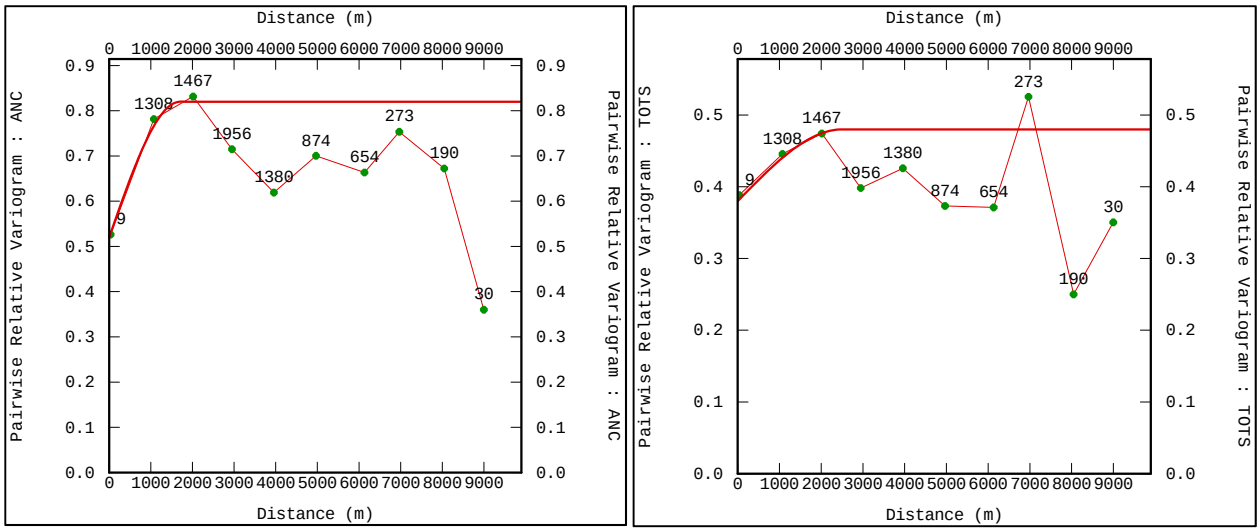

Figure 8: Example ANC and total sulphur (TOTS) experimental (showing number of pairs at each distance) and model (smooth line) variograms

Skewed and Multi-modal Distributions

The classical (parametric) method of calculating confidence intervals assumes a near 'normal' distribution. Sulphur distributions, in particular, are often highly skewed or multi-modal; therefore it is inappropriate to assume a 'normal' distribution. An incorrect confidence interval on the mean can result from using classical tests on highly skewed or multi-modal distributions (**Error! Reference source not found.**). The multi-modal distribution may be indicative of the presence of a sub-category of material types that could be distinguished at the mining scale.

'Bootstrapping' is a non-parametric technique that is more appropriate under these conditions; it also gives correct confidence intervals for normally distributed data. Bootstrapping is a term that covers a number of methods that use random sub-sampling to build multiple possible distributions from the initial full dataset. These are then used to build up a probability distribution for the mean of the data and assign confidence intervals at any desired level of confidence (90%, 95%, 99%, etc.).

An MS Excel plug-in called Datapilot for calculation of nonparametric statistical tests is freely available on the internet.

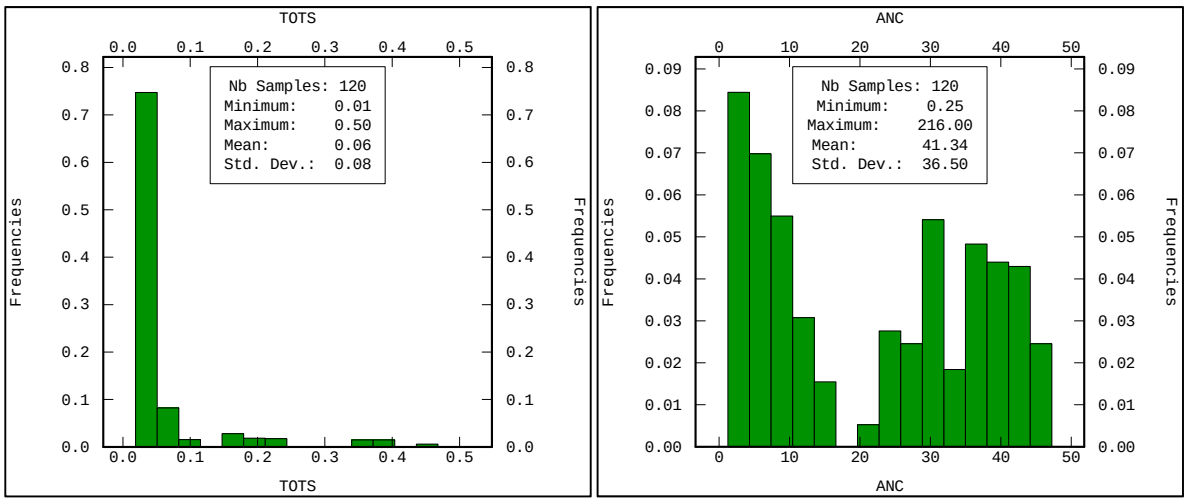

Figure 9: Left – Highly skewed histogram (total sulphur), right – multi-modal histogram (ANC)

Non-additive variables

It is not valid to calculate certain statistics and variograms on parameters that are non-additive, such as pH and neutralisation potential ratio (NPR). For example, if you physically combine two 1 m-long samples, one of which has a pH of 7 and the other, a pH of 3, the resultant pH will not be:

$$(7 + 3)/2 = 5$$

As pH measurements come from log scale, the correct result (in terms of hydrogen ion concentration H+) is:

$$(0.0000001 + 0.001)/2 = 0.00050005$$ which equates to a pH of 3.5

Another example of a non-additive parameter is the addition of NPR values. NPR is a ratio, and summing or averaging ratios does not give a correct total ratio unless appropriate weights are applied. For example, where we have two 1 m-long samples, one with an ANC of 27 kg(H_2SO_4/t) and a total S of 0.2% and the other having ANC of 87 kg(H_2SO_4/t) and Total S of 0.03%, using:

$$NPR = ANC/(Total\ S * 30.6)$$

gives NPRs of 4.41 and 94.77 respectively; while averaging these gives a value of 45.59.

However, if the two samples are physically combined, then:

$$ANC = (27+87)/2 = 57$$ kg(H_2SO_4/t); and

$$Total\ S = (0.2+0.03)/2 = 0.115\%$$

The NPR then becomes $(57/(0.115*30.6)) = 16.20$

Conclusions

Sample selection for AMD is a dynamic and iterative process and the extent of sampling at the pre-mining stage is dependent on many factors. Minimum sampling requirements should not be prescribed in terms of numbers of samples per volume or tonnage of rock, as key parameter distributions will vary with rock type, weathering state, project stage and PAF management technique employed.

In order to minimise the initial sampling requirements, it is best to have wide-spaced coverage over the whole extent of the rock type to capture the full variability of the parameters of interest, as well as some closer-spaced sampling in one or two local areas to try and capture the short-scale variability and assess the range of the variogram(s). Appropriate spacings for both wide and close sampling are dependent on the commodity type and host rock formation type and will vary from several kilometres to tens of metres.

For deposit types where there is no pre-existing information or practitioner experience, the decision on waste sample spacing will need to be an iterative process. A guide to the initial wide-spaced sampling dimensions is a factor of four times the existing resource drill spacing. Initial close-spaced sampling should be at the existing resource drilling spacing or less. Once the initial samples are assessed, more sampling may be required before an acceptable minimum level of information is reached to characterise the waste globally, and / or spatially.

Assessment of the ranges of variability also enables the requirements for amounts of waste sampling during production to be calculated. Cost benefit trade-offs for selective or non-selective handling of PAF material can then be examined. The concepts of spatial variability change of support need to be understood when making decisions about PAF material management strategies, and geostatistics can be used to assist with these decisions.

References

Kleinmann, R L P (ed), 2000. Prediction of water quality at surface coal mines: National Mine Land Reclamation Centre, West Virginia University.

Queensland Department of Environment and Resource Management (DERM) Guidelines – Assessment and Management of Acid Drainage (Current as of Jan 1995). Available from: http://www.derm.qld.gov.au/register/p01206am.pdf>[Accessed 01 February 2012].

Schafer, W M, 1993. Design of Geochemical Sampling Programs, presented at Mine Operation and Closure Short Course, Helena, Montana, USA, April 27-29.

SRK, 2010, Geochemical Characterisation of the Alpha Project. Available from: http://hancockcoal.com.au/files/alpha-coal-eis-files/Volume_5_-_Coal_Mine_Appendices/Appendix%20J1%20Geochemical%20Report.pdf

SRK, 2011. Kevin's Corner Geochemical Characterisation and Assessment. Available from: http://www.qikpress.com.au/KevinsCorner_2/484a9d8b-b6bd-453d-8286-9ec2009a20df?chapterId=a0376e05-d149-40cc-af1d-9ec200c3bd31

US Environmental Protection Agency (USEPA) - Technical Document, Acid Mine Drainage Prediction (EPA 530-R-94-036 NTIS PB94-201829) December 1994. Available from: http://water.epa.gov/polwaste/nps/upload/amd.pdf>[Accessed: 01 February 2012].

USDA Forest Service, 1992. A Conceptual Waste Rock Sampling Program for Mines Operating in Metallic Sulfide Ores with a Potential for Acid Rock Drainage (author G Farmer, Department of Agriculture, Forest Service, Ogden, Utah, USA).