



The Value of Adaptive Mineral Processing based on spatially varying Ore Fabric Parameters

K. Gerald VAN DEN BOOGAART¹, Christin WEIßFLOG¹, Jens GUTZMER³

¹ Institut für Stochastik, TU Bergakademie Freiberg, Germany,
boogaart@math.tu-freiberg.de

² Institut für Mineralogie, TU Bergakademie Freiberg, Germany,
christin.weissflog@geosamm.tu-freiberg.de

³ Helmholtz Institute for Resource Technology, Freiberg, Germany,
jens.gutzmer@mineral.tu-freiberg.de

Peer-reviewed IAMG 2011 publication

[doi:10.5242/iamg.2011.0265](https://doi.org/10.5242/iamg.2011.0265)

Abstract

We show that adapting the mineral processing to the local ore fabric (mineralogy and micro-structure) can substantially improve the profitability of the mine, however, only a proper geometrical methodology using conditional expectations of profits rather than direct measurements can avoid losses. Based on geometallurgical exploration data and processing models potential gain and actual gain of this approach can be quantified before exploitation commences. Rules for optimal decisions, estimators for the gain of this approach from exploration data, and a method to compute the optimal geometallurgical sampling density are presented.

1 Context

We investigate the possibility of adaptive mineral processing in an ore deposit based on a spatial dataset of Mineral Liberation Analyzer (MLA) measurements.

Unlike a classical grade measurement the MLA provides very detailed information on the mineralogy and fabric of the ore, e.g. information on

- mineral composition of the ore
- grain sizes distributions of individual minerals
- phase interfaces
- mineral association

Depending on the micro fabric of the ore different processing parameters and different processing paths might lead to different results. E.g. if the target element is bound in a different mineral with a different density or if we have empty minerals with densities similar to those of the target minerals in the ore we will get different gravity separation results. Different grain sizes will lead to different particle diameter below which milling will ensure the liberation of the target mineral. Depending on phase interfaces separation during milling might occur along grain boundaries or by cracking within one or the other mineral. I.e. we might have well separated minerals with a clearly different density or contaminated grains with not so clearly different densities and a not so good separation. The general intergrowth of the target mineral might ease or complicate its liberation e.g. by enclosing the target mineral in another hard to break mineral or by having it located along grain boundaries. The presence of difficult to mill minerals e.g. very ductile material might change the energy consumption of milling to a specific size. Thus even material with the same grade subjected to the same type of processing will result in a different recovery.

The general idea of geometallurgy is to optimize the processing through such detailed information on the mineralogy and fabric of the ore. The aim of this paper is to develop and illustrate some geomathematical basics for this approach.

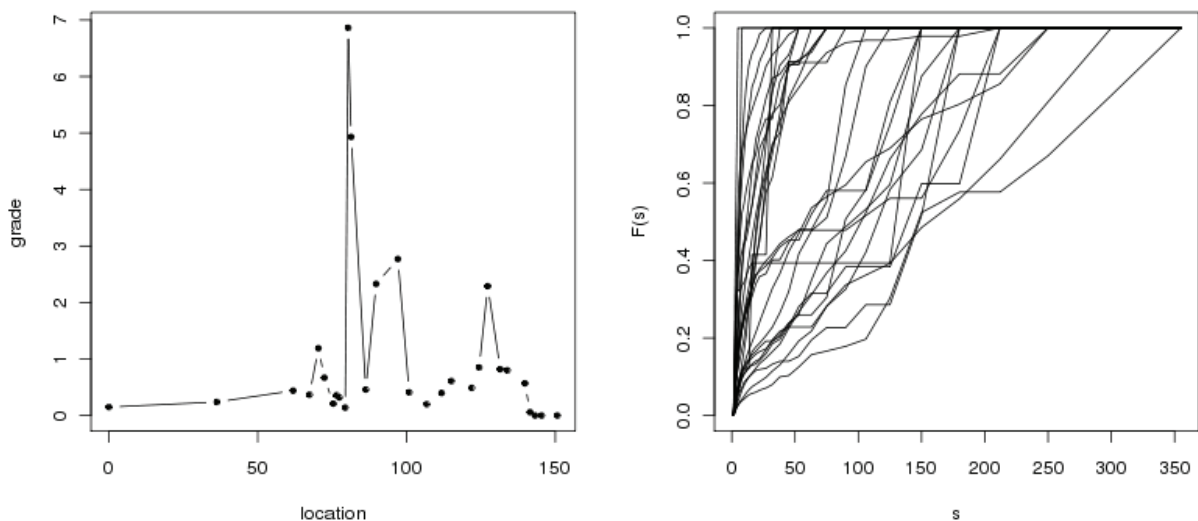


Figure 1: The left panel shows the measured mineral abundance of the value mineral along the transect. The right panel shows the observed mass percentages below given sizes observed for these samples. We can see a substantial variation of grades and grain size distribution along the transect in this deposit.

2 Example Dataset and Processing model

As an illustration example we will use 29 datasets of MLA measurements from an MLA 600F from FEI done in XBSE mode on polished thin section surfaces of mineralized Zechstein Kupferschiefer and sandstone samples received from KupferschieferLausitz GmbH (KSL), which were taken along a vertical transect through a Cu-Ag deposit. The Permian Kupferschiefer is a thin unit of marine bituminous marl that conformably overlies sandstones of the Pre-Zechstein and grades upward into dolomitic limestones (Speczik et al., 1995; Vaughan et al., 1989). It occurs from England through the Netherlands, Germany into Poland, the Baltic states, and Belarus and covers an area of approximately 600 000 km² (Kopp et al., 2008). Stratabound polymetallic base metal mineralization is locally developed in this sedimentary succession, especially along the southern and southwestern perimeter of the distribution of the Kupferschiefer. Mineralization is structurally - and lithologically - controlled. The metallogenesis of the Kupferschiefer-type mineralization remains a contentious issue, with various models advocated by previous authors. Excellent summaries of the available models have been provided by Vaughan et al. (1989) and, more recently, by Hitzman et al. (2010). In the Spremberg-Graustein area of Lower Lusatia, Germany, Kupferschiefer is situated in depths between 800 and 1500 m. Economically significant mineralization is evenly distributed in the Pre-Zechstein sandstone (31.4 %), in the marine bituminous marl (45.7 %) and in the dolomitic limestone (22.9 %) and has an average thickness of 2.4 m. The main copper ore minerals are chalcocite/digenite, bornite and chalcopyrite (Kopp et al., 2008).

Twenty-nine polished thin sections were analyzed with an MLA 600F from FEI, which is a scanning electron microscope equipped with a field emission source and two energy dispersive Bruker X-ray spectrometers. The instrument is used to acquire backscattered electron (BSE) images and compositional information from a large number of particles. The MLA 600F is able to characterize mineral grains $\leq 0.1 \mu\text{m}$ in size. The MLA software offers different measurement modes, varying from purely BSE-based area-analysis to an X-ray analysis point counting technique (XMOD) (Fandrich et al., 2007). The XBSE measurement mode was used in this investigation analysis of mineralogically and texturally complex surface areas of polished thin section surfaces.

The mass portion of one ore mineral and its corresponding grain size distribution functions as provided by the MLA are shown in Figure 1. For the sake of simplicity and confidence we will focus on one mineral only.

Since in this paper we are focused on mathematical aspects of the optimization of mineral processing and not on the optimization of a specific mining operation we will use a very simplified one parametric processing model:

- The target element is in only one of the minerals.
- Each grain would be liberated (i.e. separated from other minerals and thus accessible to processes like flotation) if milling to the apparent grain size in the 2D thin section as reported by the MLA. This is obviously an oversimplification of the relationship, because a complex stereological relation involving properties of the milling process relate the apparent 2D grain sizes with the actual liberation sizes in an actual milling process.
- The processing costs are dominated by the energy consumption of milling, which is modeled as proportional to $1/\text{diameter}$, which corresponds to the increase in inner surface for decreasing diameter.
- The extracted material and thus the monetary gain is proportional to the amount of the value mineral liberated by milling.

In this simple model the monetary gain per unit mass only depends on the mass distribution function F of the grain sizes on that target mineral and on the grade m (for mass) of the value mineral:

$$V(s) = a_1 m (1 - F(s)) - \frac{a_2}{s} - a_3$$

where a_1 is the market price of the mineral, m is the mass percentage, $F(s)$ the portion not liberated if milling until size s , s is the chosen “size” to which we mill the material, a_2 is the proportionality factor for milling costs and a_3 the other costs of mineral processing (here set to 0). We assume that parameters of the function (here a_1 , a_2 , a_3) are known, that ore specific quantities (here m and $F(s)$) can be estimated from the MLA measurements and that process parameters (here s) can be chosen. The precise form of this value function is not important for the general aspects of the paper, it is just the function used in the example. It is only important that such a function is known for the processing problem under consideration. Our assumption also implies that the average gain for mixed material processed at the same s is the average of gains computed for the components. This is however only true if different sorts of ore do not interact in the sense that the processing effects on each portion of material is not changed by further added different material. Without this basic assumption the mathematical treatment gets more complicated, whenever mining blocks are considered.

In this setting optimizing the mineral processing is reduced to finding a good value for the process parameters (here s for the grain size to which we mill) for each portion of material processed together.

Figure 2 shows the value functions $V(s)$ computed for the parameters observed in each of the 29 samples for the illustrated example. The red crosses show the optimal particle size and the corresponding maximum achievable monetary gain. The blue line and the blue cross show the average gain if all types of ore would be processed with the same parameter s and the same corresponding optimal joint particle diameter. The red horizontal line represents the average $V(s)$ if always the optimal value would be taken.

The value of different types of ore along the transect vary substantially. Ore with higher grades provide much higher gain. For higher grade ore with the same grain size distribution it is typically worth to invest more into the milling, since the Maximum of $V(s)$ is characterized by a zero derivative $V'(s) = \frac{a_2}{s^2} - a_1 m f(s)$, where a larger value of m automatically shifts this value to the left, since the derivative $\frac{a_2}{s^2}$ of the milling energy is a decreasing function of the milling target size. The red line is about 20% higher than the blue line, which would correspond to a 20% increase of the gross yield for a mining operation using always the best milling parameter with respect to a mining operation always using the best joint milling parameter. Every individual mining block would provide a larger yield.

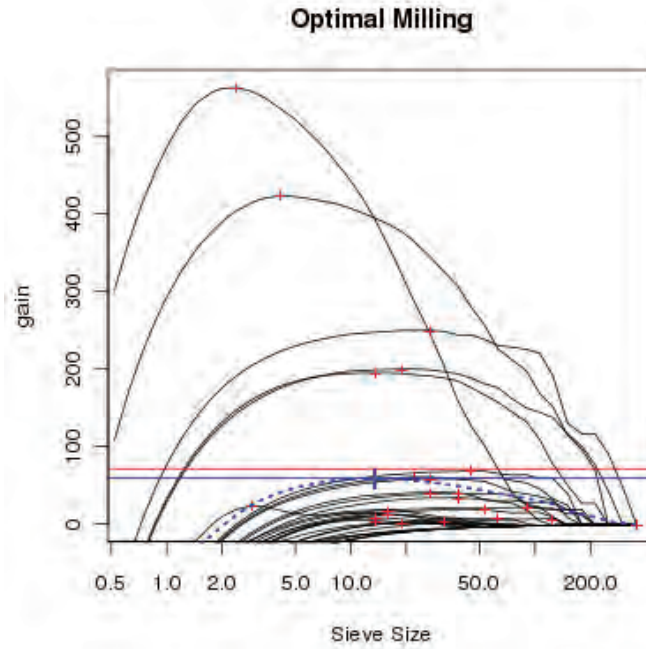


Figure 2: The black curves correspond to the $V(s)$ computed for the observed mineralogy and microfabric of the specimen. The blue curve corresponds to the average. The red crosses locate the optimal milling parameter s for each specimen blue cross the optimal milling parameter for the average. The blue cross locates the optimal milling for the average, i.e. the optimal milling parameter, when everything is processed in the same way.

This corresponds to a mathematical theorem:

Theorem 1:

For an objective function $G(s_1, \dots, s_n) := \sum_{i=1}^n V_i(s_i)$ it holds:

$$\max_s G(s, \dots, s) \leq \max_{s_1, \dots, s_n} G(s_1, \dots, s_n)$$

with equality if and only the maxima are the same for all i, j

$$\operatorname{argmax}_s V_i(s) = \operatorname{argmax}_s V_j(s)$$

Proof: A maximum over a larger set of possibilities is always greater equal.

I.e. one can show mathematically that the optimal processing always provides a better output than applying the same mineral processing to the complete deposit. Clearly in practice the first theorem does not consider several problems:

- There are the higher investment costs for processing plants capable of adaptive processing.
- The microfabric of the ore is never completely known.
- There are measurement costs for obtaining the microstructural information.

We will come back to these problems.

We call the expected value of the difference $U := E[\max_s V(s)] - \max_s E[V(s)]$ the upside potential of adaptive processing, because it is the value that would be possible in case of perfect information at no additional cost.

3 Pitfalls of the “measured microfabric” approach

To apply adaptive processing sounds at first like a no brainer as, e.g., an increase of the gross yield by 20% before the subtraction of the other costs of say 80% could easily correspond to a doubling of the net gain. However reality is not so simple. In reality we do not know the true mineralogy and microfabric of the ore processed. We only know the mineralogy and microfabric of a very small thin section of a sample somewhere near the processed material.

To model this effect of only having a nearby measurement, we will apply in our example model the processing parameters optimized for the prior measurement along the transect instead of the optimal choice for the sample itself. We quantify the effect using the following general theorem allowing estimating the effective value of any adaptive processing decision rule from geometallurgical exploration data.

Theorem 2:

Let denote $\hat{V}_i(s)$ an unbiased estimator of $V_i(s)$ the value estimated from an observed microfabric, i.e. $E[\hat{V}(s)|M] = V(s)$ for every microfabric M and \hat{s} be any random variable providing a choice on the processing parameters s , where $\hat{V}(s)$ is independent to \hat{s} conditional to the microfabric (I.e. \hat{s} does not contain any information about the prediction error not already contained in the actual sample). Then it holds that

$$E[\hat{V}_j(s)] = E[V_j(\hat{s})]$$

Proof: $E[V_j(\hat{s})] = E[E[V_j(\hat{s})|M]] = E[V_j(\hat{s})]$

I.e. the computed value of the second sample processed based on a choice on the processing given \hat{s} computed from some observations of different samples is an unbiased estimator of the actual value that would be achieved by such a procedure. In our case the approach would thus give an unbiased estimate of the actual average value we would get, when processing the material at the observed locations with choice from a similar distance distribution. A paired sampling strategy like in Boogaart 2010 is thus a useful tool to construct unbiased estimators for the performance of decision strategies in mining operations.

The result of applying this to our example dataset is shown in Figure 3. This more realistic strategy would not provide a gain but a loss reducing the yield by approximately 25%. In conclusion a naive: “measure and use the apparent optimal choice approach” can lead to a great loss in the mining operation, with respect to a traditional “treat all the same” approach.

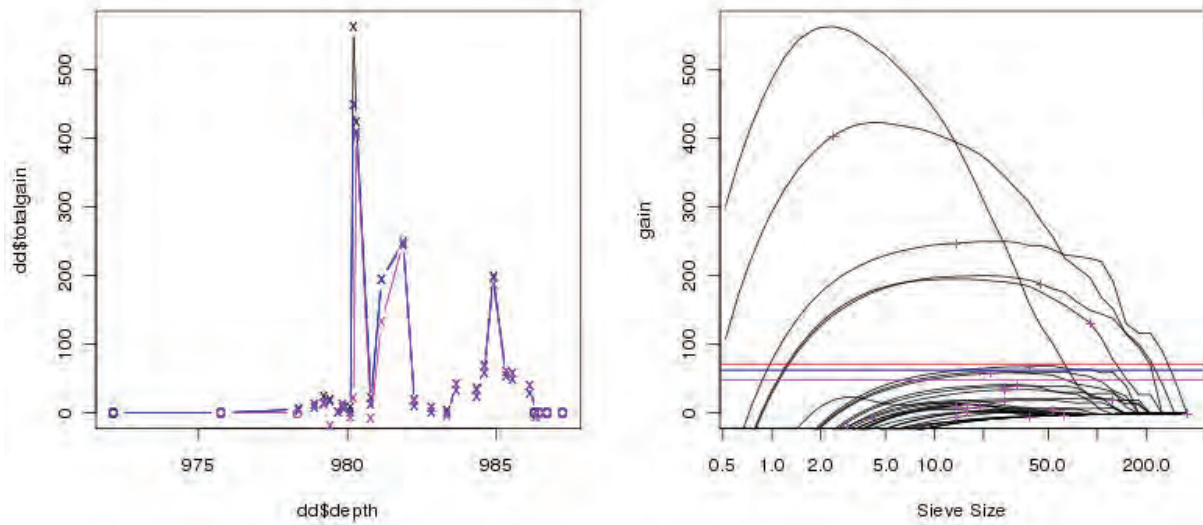


Figure 3: The left panel shows the value as a function of space. The right panel shows the value as a function of the processing parameter. Crosses mark the effective gain we would get with three different strategies: The black lines (mostly covered by blue) correspond to the optimal choice, the blue line to the same choice for every location. The magenta line corresponds to a choice based on the previous measurement. For several cases this third choice is bad because the prior measurement is substantially different from the actual material. The horizontal lines show the average result.

4 Problems with Estimated choices

This last observation might seem very surprising, since the measured microfabric and mineralogy provide an unbiased estimator of a true microfabric function $M(s) := m(1 - F(s))$ of a block, like an ordinary kriging predictor of the microfabric would yield an unbiased prediction and the gain $V(s)$ depends linearly on $M(s)$, i.e. $E[\hat{V}(s)] = V(s)$. Should this not imply that the choice is at least somehow on average the same as the optimal one? However the precise theorem is:

Theorem 3:

For the conditional expectation $\hat{V}(s) := E[V(s)|Data]$ given the data it holds:

$$\hat{s} := \max_s \hat{V}(s)$$

provides the best possible choice on s in the sense that \hat{s} gives a better choice than all alternative Estimators \check{s} which can be written as a function of the data:

$$E[\hat{V}(\hat{s})] \geq E[\hat{V}(\check{s})]$$

Proof: $E[V(\hat{s})] \geq E[E[V(\check{s})|Data]] = E[V(\check{s})]$

This equivalence to the conditional expectation is however something completely different than unbiasedness for a given $V(s)$ considered as a parameter.

The problem is illustrated in figure 4.

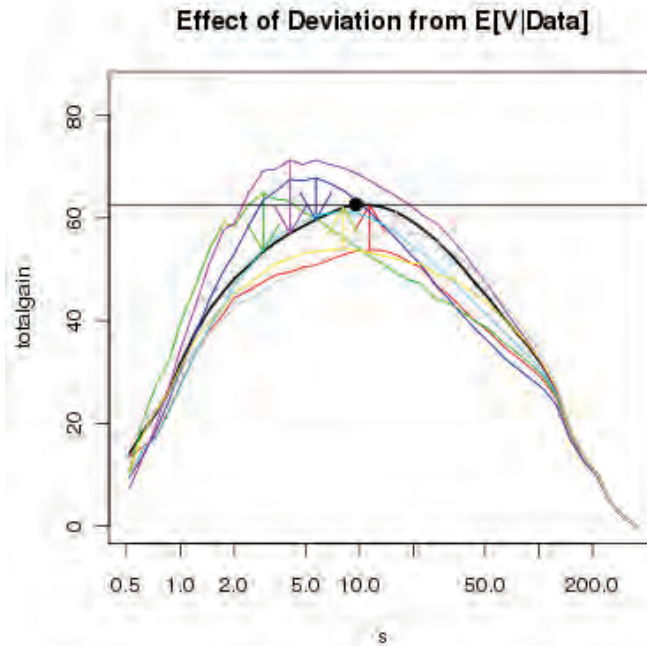


Figure 4: The figure shows one true conditional expected gain and curves simulated unbiasedly around this curve, i.e. the expectation of all these curves was the given curve. The black line represents the conditional expected gain.

Due to the variation around the conditional expectation, the location of the maximum varies. Since always those s are preferred, for which the function is high by chance. However a wrong s will always imply a lower effective yield, i.e. the actual value of the black curve given the choice is always below the optimal choice. These effects thus do not cancel out. Any variation around the conditional expectation will lead to a decrease in performance of the choice.

On the other hand there is an interesting effect: The actual prediction error for the true performance is irrelevant, only the conditional expectation itself matters. The theorem allows the explicit construction of the optimal processing choice. The theorem also shows, which microfabric parameter should be predicted and how it should be predicted: Since the final choice needs the conditional expectation of the value of a processing choice, the value of the choice itself should be the target of the estimation.

5 Practical choices

In our example case we do not have an explicit way of constructing the conditional expectation and restrain ourselves to a linear approximation to it, through a regression model, in this case done for every choice separately. Figure 5 shows the unclear dependence on the value for the previous location along the transect with a corresponding regression line in two examples:

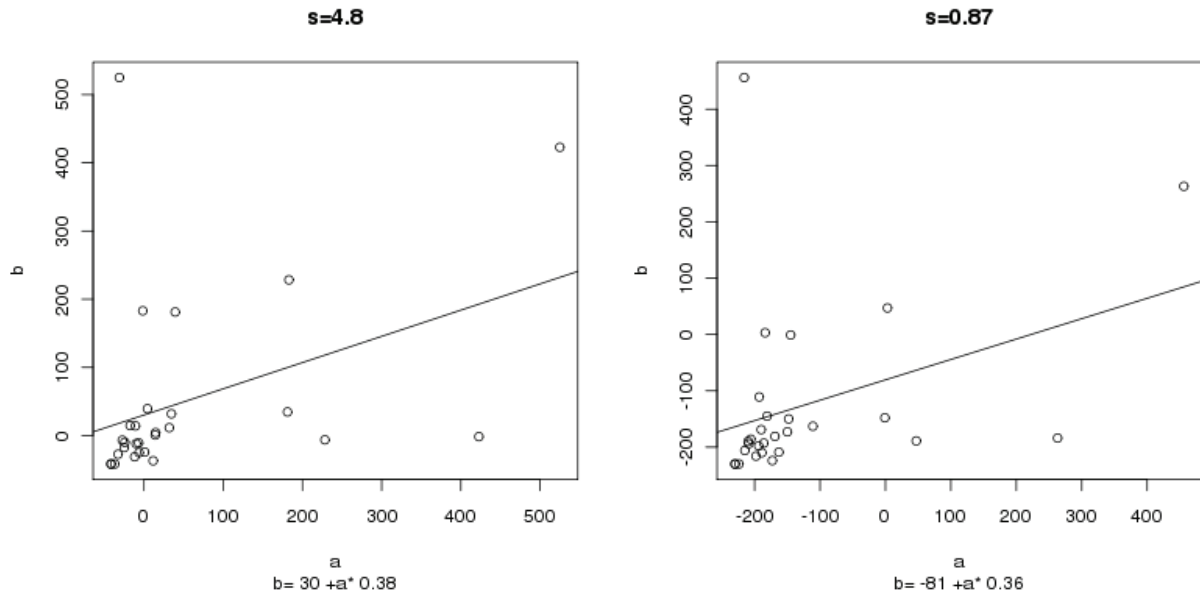


Figure 5: The figure shows the expected value for a sample (a) and its subsequent sample (b) for two different possible choices of s . The regression lines are not on the $y = x$ line because in regression analysis the errors are asymmetrically assigned to the dependent variable only. The prediction gain $\hat{V}(s)$ is thus a somehow weighted average of the “measured” $V(s)$ and the average $V(s)$. This effect is known as the “regression to the mean effect”.

We have applied this simple strategy to a discretized set of possible s . Since the “observations” a and the “results” b have nearly the same mean for a linear regression equation $b = A + Ba$ we always have $A = \bar{b} - B\bar{a} \approx \bar{b} - B\bar{b} = \bar{b}(1 - B)$. The regression is thus a convex combination of the overall mean and the measurement. The regression coefficient B can be interpreted as a number between 0 and 1 showing the degree of relying on the data. If the B is 0 the data is not used at all and the predicted gain and thus the choice would be the same as for the mean prediction. If $B = 1$ the predicted gain would be the measurement and we would get the same result as in the perfect information situation. Assuming that both gain curves look approximately like a parabola in the area of interest a convex combination of their linear derivatives would move the point of choice linearly. Since the actual value is then taken from a parabola the errors are approximately proportional to the squared error. And thus the measure of determination R^2 measures the portion of the maximum upside potential we might gain, which is approximately 14% in both cases. According to these simplified models we would expect an increase in average processing gain of $\hat{U} * R^2 = 20\% * 14\% = 2.8\%$.

Figure 6 shows the corresponding predicted curves along with the corresponding choices on the milling size.

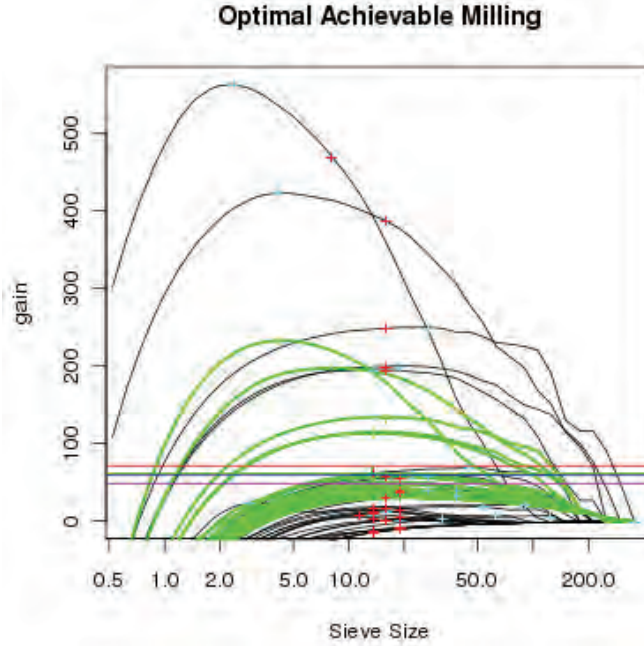


Figure 6: Green lines show the decision values predicted by the linear model and the corresponding milling choices as green crosses. Accordingly the actually achieved gains are represented by the cyan crosses. The green horizontal line shows the overall gain which is only slightly above the blue “treat everything in the same way” line.

Applying the estimation based on theorem 2 to this approach we estimate an improvement of 3.9% in gross yield with respect to the classical ‘treat everything the same’ approach. The dramatic gain of 20% provided by the first model was not realistic because it wrongly assumed that we know the microfabric perfectly well. These 20% are somehow the limit of what increase we can reach with a perfect adaptive processing strategy based on perfect information. However, our information only allows us to exploit $\frac{3.9\%}{20\%} = 19\%$ of this upside potential, which is slightly more than the 14% expected from the simplified approach with R^2 .

6 Finding optimal sampling distances

If the current sampling distance only allows to exploit 19% of the potential gain, because R^2 is small corresponding to a bad prediction of $V(s)$, the problem might be that we do not have enough measurements of the fabric. The question must be, what might be the right number of measurements. To compute this we assume to predict the conditional expectation of (x, s) , where x is a spatial location and s the processing parameter with a simple kriging predictor (Cressie 1993) based on values $V(x_i, s)$ computed based on MLA measurements at locations s . If $\hat{V}(x_i, s)$ is an unbiased estimator of the gain using processing parameter s on the measured data at location x_i as in theorem 3 the upside potential U of mean gain using an adaptive mineral processing described by a value function $V(s)$ can be estimated by

$$\hat{U} = \frac{1}{n} \sum_{i=1}^n \max_s (\hat{V}(x_i, s)) - \max_s \frac{1}{n} \sum_{i=1}^n \hat{V}(x_i, s).$$

\hat{U} is biased but consistent. The upside potential is the mean increase in gain if everything is processed optimally with respect to perfect information. Chapter 5 suggests that the portion of the upside potential we can realistically achieve with a given sampling density corresponds to the

mean R^2 of the prediction. In this case the prediction R^2 corresponds to the kriging predictor variance (I.e. variance of the field – kriging error variance) divided by the variance of the field (I.e. of the predicted value). This however would in a simple kriging situation only depend on the correlogram of $V(x, s)$ and the sampling plan. For the sake of simplicity we assume that we have a single correlogram $c(h)$ for all $V(x, s)$ independent of s in the area of interesting values of processing parameters and a square sampling plan. Obviously a nugget effect N would prevent the correlation of the prediction on the actual value to grow above $(1 - N)$. The practical upside potential would thus be limited by $U(1 - N)$ in case of complete sampling. Having such a correlogram would allow doing a numerical computation of $G(d)$ defined as the spatial average of the prediction R^2 as a function of the measurements per area unit. Unfortunately our example dataset was not large enough to estimate a correlogram. Figure 7 shows these calculations for an example with a spherical correlogram with range 1 and nugget 0.1.

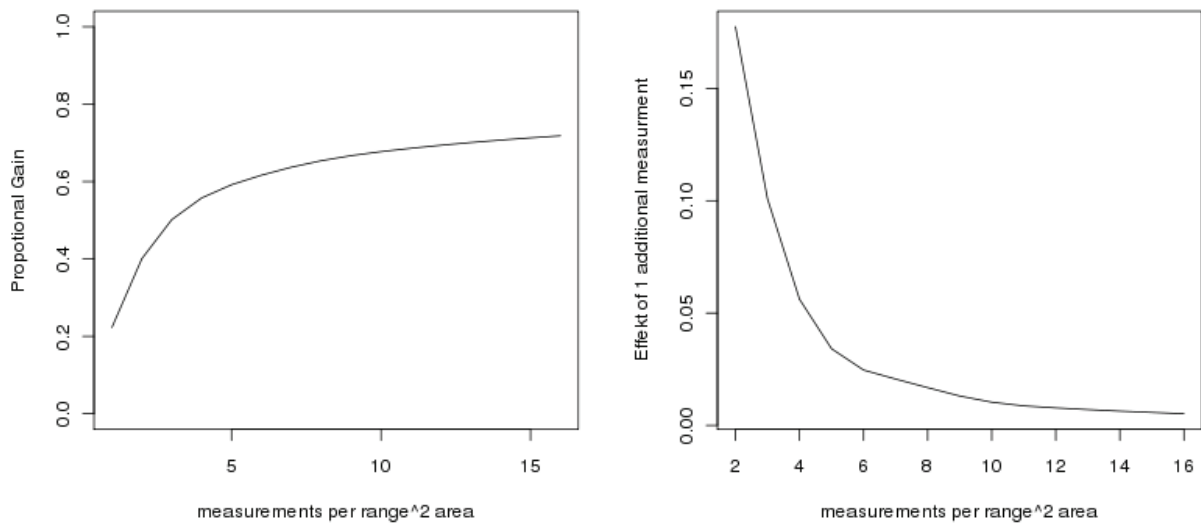


Figure 7: The left panel shows a numerical computation of the spatial average $G(d)$ of the prediction R^2 for a spherical correlogram with range 1 and nugget 0.1 using a rectangular measurement grid. The right panel shows the numerical derivative $E(d) = G'(d)$ of that function.

Let's consider the density d of measurements per unit area. Adding one measurement would cost a price p and provide additional gain, which would in our simplified first model estimate by $E(d) * \hat{U}$. An increase of the measurement density is thus reasonable as long as $E(d) > \frac{p}{\hat{U}}$ and thus the apparently optimal measurement density is given by $E^{-1}\left(\frac{p}{\hat{U}}\right)$, depending on the price of the measurements, the estimated upside potential \hat{U} and the correlogram's of the $V(x, s)$. The actually achievable gain would be estimated by $\hat{V}_{adapted} := G\left(E^{-1}\left(\frac{p}{\hat{U}}\right)\right)$. If such value exceeds the additional investment costs I for plants capable of adaptive processing, adaptive processing increases the value of the mine.

Clearly this last step is based on many assumptions and needs to be tested against larger datasets in which the correlograms can be estimated and the performance checked against a performance estimate based on theorem 3.

7 Conclusions

The loss of 25% through a naive “use the measurements directly” approach might explain, why adaptive processing is not yet an established method. Only a proper theoretical foundation combined with informative predictive models for local microfabric and mineralogical characteristics as well as corresponding processing results can provide sufficient information for an adaptive approach to mineral processing. However, if these ingredients are present along with a sufficient dense measurement grid a huge upside potential of e.g. 10% of gross yield in our example deposit might be possible, which would correspond to a dramatic increase in capital return. Optimization of mineral processing through such geometallurgical methods should thus gain more interest in the geomathematical community.

Acknowledgments

The authors thank Kupferschiefer Lausitz GmbH (KSL) for the kind support.

References

- Brongersma-Sanders, M. (1966): Metals of Kupferschiefer supplied by normal sea water: *Geologische Rundschau*, v. 55, p. 365-375.
- Boogaart, K.G., Wellmer, F.-W. Cutoffs and selectivity (2010): How to estimate the processable content of a deposit in the exploration phase, *Proceedings of IAMG 2010, Budapest*
- Cressie, N (1993): *Statistics for spatial data*, Wiley
- Fandrich, R., Gu, Y., Burrows, D., and Moeller, K. (2007): Modern SEM-based mineral liberation analysis: *International Journal of Mineral Processing*, v. 84, p. 310-320.
- Gu, Y. (2003): Automated scanning electron microscope based mineral liberation analysis: an introduction to JKMRC/FEI mineral liberation analyser: *Journal of Minerals & Materials Characterization & Engineering*, v. 2, p. 33–41.
- Hitzman, M.W., Selley, D., and Bull, S. (2010): Formation of Sedimentary Rock-Hosted Stratiform Copper Deposits through Earth History: *Economic Geology*, v. 105, p. 627-639.
- Kopp, J., Herrmann, S., Höding, T., Simon, A. and Ullrich, B. (2008): Die Kupfer-Silber-Lagerstätte Spremberg-Graustein (Lausitz, Bundesrepublik Deutschland) – Buntmetallanreicherungen an der Zechsteinbasis zwischen Spremberg und Weißwasser: *Zeitschrift der geologischen Wissenschaften*, v. 36, p. 75 – 114
- Jowett, E.C. (1986): Genesis of Kupferschiefer Cu-Ag deposits by convective flow of Rotliegendes brines during Triassic rifting: *Economic Geology*, v. 81, p. 1823-1837.
- Rentzsch, J. (1974): The “Kupferschiefer” in comparison with deposits of the Zambian copperbelt. In: P. Bartholomé, Editor, *Gisements Stratiformes et Provinces Cuprifères*, Soc. Géol. Belgique, Liège (1974), p. 403–418.
- Speczik, S., Bechtel, A., Sun, Y.Z., and Püttmann, W. (1995): A stable isotope and organic geochemical study of the relationship between the Anthracosia shale and Kupferschiefer mineralization (SE Poland): *Chemical Geology*, v. 123, p. 133-151.
- Vaughan, D.J., Sweeney, M.A., Friedrich, G., Diedel, R., and Haranczyk, C. (1989): The Kupferschiefer; an overview with an appraisal of the different types of mineralization: *Economic Geology*, v. 84, p. 1003-1027.