A PRACTITIONERS IMPLEMENTATION OF INDICATOR KRIGING

Ian Glacken and Paul Blackney
Snowden Associates

Abstract

Indicator Kriging (IK) was introduced by Journel in 1983, and since that time has grown to become one of the most widely-applied grade estimation techniques in the minerals industry. Its appeal lies in the fact that it makes no assumptions about the distribution underlying the sample data, and indeed that it can handle moderate mixing of diverse sample populations. However, despite the elegant and simple theoretical basis for IK, there are many practical implementation issues which affect its application and which require serious consideration. These include aspects of order relations and their correction, the change of support, issues associated with highly skewed data, and the treatment of the extremes of the sample distribution when deriving estimates.

This paper discusses the theoretical and practical bases for these considerations, and illustrates through examples and case studies how the issues associated with the daily application of the IK algorithm are addressed. Finally, some less commonly-used IK applications are presented, and the limitations of IK are discussed, along with proposed alternatives.

Key Words: *geostatistics, indicator kriging, minerals industry, categorical kriging, soft kriging.*

Introduction

Indicator Kriging (IK) as a technique in resource estimation is over fifteen years old. Since its introduction in the geostatistical sphere by Journel in 1983, many authors have worked on the IK algorithm or its derivatives. The original intention of Journel, based on the work of Switzer (1977) and others, was the estimation of local uncertainty by the process of derivation of a local cumulative distribution function (cdf). The original appeal of IK was that it was non-parametric – it did not rely upon the assumption of a particular distribution model for its results. From slow beginnings in the early eighties as a technique in mineral resource estimation, and in many other natural resource mapping applications, IK has grown to become one of the most widely-used algorithms, despite the relative difficulty in its application. It is the prime non-linear geostatistical technique used today in the minerals industry.

This paper presents an overview of the theory of IK, followed by some discussion of practical applications. A number of practical aspects concerning the implementation of the IK algorithm and its variants are then discussed, including various ways to overcome some of the shortcomings of the technique. Finally, some of the less common applications of the indicator approach are presented, and an approach which is the successor to IK is proffered.

Overview of theory of Indicator Kriging

The concept of indicator coding of data is not new to science, but has only been proposed in the estimation of spatial distributions since the work of Journel (1983). The essence of the indicator approach is the binomial coding of data into either 1 or 0 depending upon its relationship to a cut-off value, z_k. For a given value z(x),

$$i(x;z_k) = \begin{cases} 1 & if \quad z(x) \geq z_k \\ 0 & if \quad z(x) < z_k \end{cases}$$

This is a non-linear transformation of the data value, into either a 1 or a 0. Values which are much greater than a given cut-off, z_k, will receive the same indicator value as those values which are only slightly greater than that cut-off. Thus indicator transformation of data is an effective way of limiting the effect of very high values. Simple or Ordinary kriging of a set of indicator-transformed values will provide a resultant value between 0 and 1 for each point estimate. This is in effect an estimate of the proportion of the values in the neighbourhood which are greater than the indicator or threshold value.

Carried out over a larger area for a series of blocks (putting aside for a moment the change of support issue), IK has the potential to generate recoverable resources – in other words, the proportion of a block theoretically available above a given cut-off grade (indicator threshold).

The outcome of IK is a conditional cumulative distribution function (ccdf) – in effect a distribution of local uncertainty or possible values conditional to data in the neighbourhood of the block to be estimated. This distribution of grades can be used for many purposes, in addition to simply deriving the average (or 'expected') value. Any relevant criteria may be used to derive the estimate required, not simply the arithmetic mean of the local distribution.

The practice of IK involves calculating and modelling indicator variograms (that is, variograms of indicator-transformed data) at a range of cut-offs or thresholds which should cover the range of the input data. This approach is termed Multiple Indicator Kriging (MIK). Until recently, this has been a somewhat time-consuming exercise. One approximation is to simply infer the variogram for the median of the input data and to use this for all cut-offs. This so-called Median IK approach is very fast, since the kriging weights do not depend on the cut-off being considered. Median IK also necessitates the solution of only one kriging system per block in contrast to the multiple systems required for MIK. However, Median IK has its own assumptions and drawbacks, as discussed below.

Since IK generates at each point or block a cumulative distribution, this should be non-decreasing and valued between zero and one. These two requirements are sometimes not met, leading to so-called order relations violations. Many methods have been proposed to counteract the order relations issue – the most commonly-used involves direct correction of the indicator values (eg. Deutsch and Journel, 1998, p82). Another approach, proposed by Dagbert and Dimitrakopoulos (1992), involves the use of nested indicator variables – in other words, indicator variables which are defined by successively halving the data set to define the thresholds. This nested indicator kriging approach eliminates any problems associated with order relations, but suffers from data deficiency problems, especially at high thresholds.

The IK algorithm has been extended to not only include the indicator transform of the data, but also the data itself. This approach, first postulated by Isaaks (1984) and termed probability kriging, is essentially indicator co-kriging between the indicator-transformed data and the *uniform* ($0 \rightarrow 1$) transform of the sample data. As with most co-kriging, the downside is the calculation and modelling of the cross-variograms between the two data types in addition to the univariate indicator variograms.

Rivoirard (1993) used the relationships between the cross-variograms of indicators at adjacent cut-offs to draw conclusions about the nature of the processes influencing the distribution of data values. This work led to the definition of the so-called mosaic and diffusion models, among others, which lead to a particular style of IK or other non-linear estimation algorithm.

Despite much theoretical development, in practice it is the straightforward implementation of MIK, using non-nested indicator transforms of data at multiple cut-offs leading to the definition of local distributions of grade, which has proved most popular and enduring. The next section examines the advantages and motivation behind the choice of IK as an estimation technique.

Why use IK?

The primary motivation behind the use of IK in most earth science applications, and one of the main reasons for its introduction, is that it is non-parametric. Moreover, it is one of the few techniques that addresses mixed data populations. Figure 1 shows a typical scenario; a single, clearly-defined mineralogical/geological domain inside an orebody which is currently being mined, yet where the lead grades depicted show clearly mixed populations. In this instance, the geology and mineralogy of the orebody precludes further domaining. Since IK actually partitions the overall sample distribution by a number of thresholds, there is no need to fit or assume a particular analytically-derived distribution model for the data.

MIK requires the inference of a variogram model at each cut-off, and to a moderate degree, can handle different anisotropies at different cut-offs. Figures 2 and 3, from a gold project, show a single normal scores (Gaussian) variogram contour fan which suggests two directions of anisotropy, and two (non-adjacent) indicator variogram fans at a lower and an upper threshold, which demonstrate the partitioning of the anisotropy.

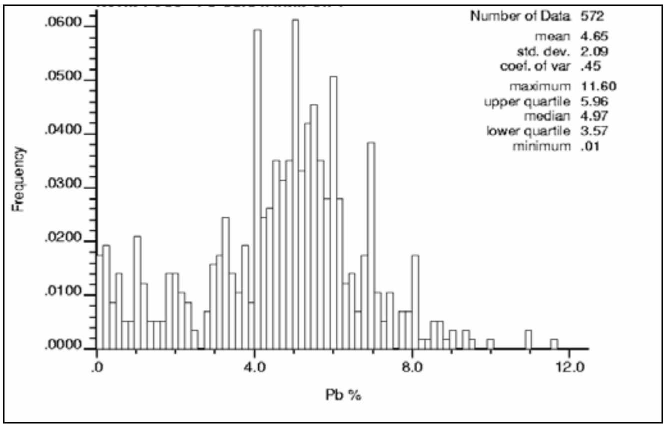

Figure 1. Histogram of lead data showing typical mixed populations

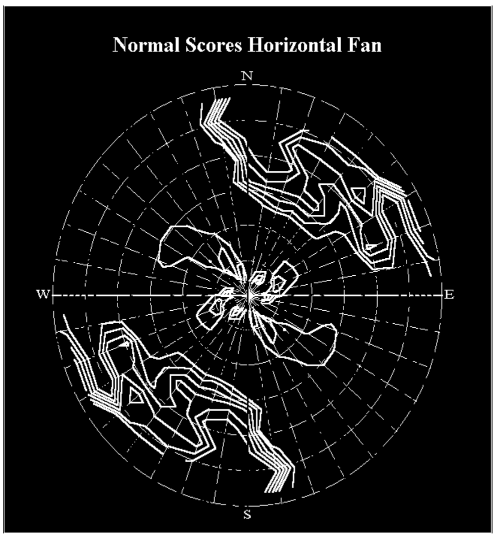

Figure 2. Normal scores variogram contours in horizontal plane showing two anisotropy directions

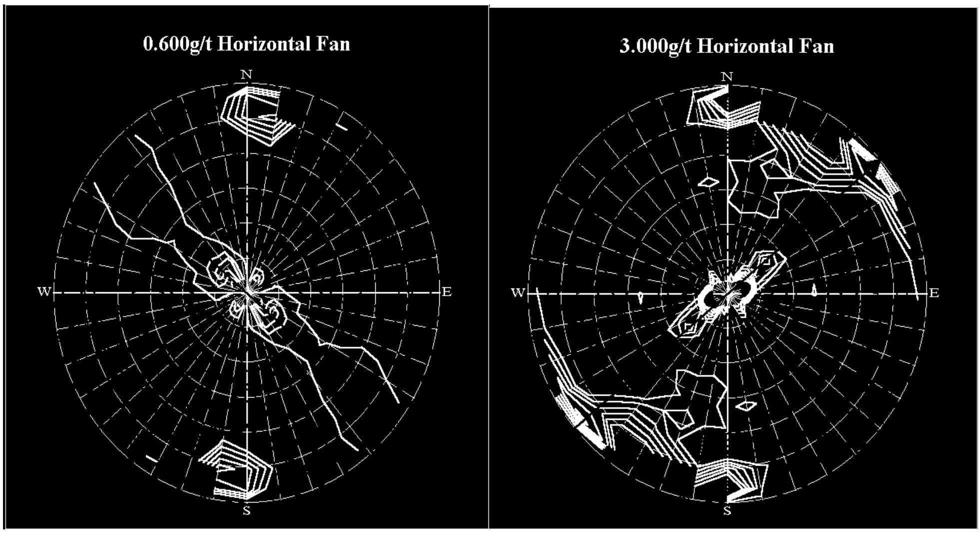

Figure 3. Equivalent indicator variogram fans for a lower (left) and an upper (right) threshold showing resolution of the two major anisotropy directions

Note that if the anisotropy changes too much between adjacent thresholds, the order relations violations become prohibitively large, but if the changes are gradual then the situation depicted can easily be handled.

IK is a favoured option for highly-skewed data sets, as it offers a practical way of treating the upper tail of the distribution which does not depend entirely on an arbitrary upper cut value. This is discussed in more detail below, but IK allows the practitioner to use features of the actual data for defining any upper tail treatment.
In many situations the level of knowledge of a deposit precludes the definition of clear geological or mineralogical domains. While unconstrained estimation is not advocated wherever there is the opportunity of defining constraining domains, if it is a necessity then MIK is an approach which can minimise the smoothing under certain conditions.

One of the great benefits of the indicator transformation is that it allows common coding of diverse data types and their integration into the single process. Since all data is transformed to $0 \rightarrow 1$ space, other, secondary data types can easily be accommodated by the same coding scheme. This approach is discussed in some detail by Journel and Alabert (1989) and by Journel and Deutsch (1996). A practical example is given below.

Finally, practitioners of IK are able to make use of one of the main motivations behind its development, that is the provision of an *estimate of uncertainty* at unsampled locations, via the inference of a distribution of values. This data can be used to derive an expected value, but also to yield risk-qualified outcomes, such as the probability of exceeding a given grade – the cut-off grade, or to map a given percentile of the distribution.

Practical implementation of IK

Treatment of upper and lower tails

An indicator kriging program will, in its most basic form, provide an estimate of the proportion of a model block above each of the indicator grades or thresholds assessed. To reduce this data into an estimate of mean block grade or grade above a cut-off, it is a requirement that each indicator class interval be assigned a grade. A number of sensitivities must be considered when undertaking the task of class interval grade assignment.

If indicator grades have been carefully selected with adequate regard to the input grade distribution, then the distribution of grades within many classes will be nearly linear. The average grade of the input data or of the bounding indicator grades will normally suffice for the assignment of grade in these classes.

The distribution of sample grades in the uppermost and lowermost grade classes of the distribution will not normally be linear and therefore require special treatment. In the case of a positively-skewed grade distribution, such as gold, the greatest estimation sensitivities relate to the grade assigned to the uppermost class. Distribution skew and grade outliers both influence the grade distribution in this class, which requires a more sophisticated method of mean grade selection if grade over-estimation or under-estimation is to be avoided.

Deutsch and Journel (1998) propose a modelling method based upon a hyperbolic distribution for representing the grade distribution above the uppermost indicator grade (Figure 4). The class mean grade calculated by this method is dependent upon the rate of decay of the hyperbolic function and the upper grade limit applied to the grade distribution.

Both of these variables may be judged from the sample grade distribution. The choice of the decay rate parameter is generally less subjective than the selection of an upper grade limit for the class. If a slow decay rate is selected the class mean is sensitive to the magnitude of the class upper grade limit. If a high decay rate is chosen, the class mean is relatively insensitive to the upper grade limit.

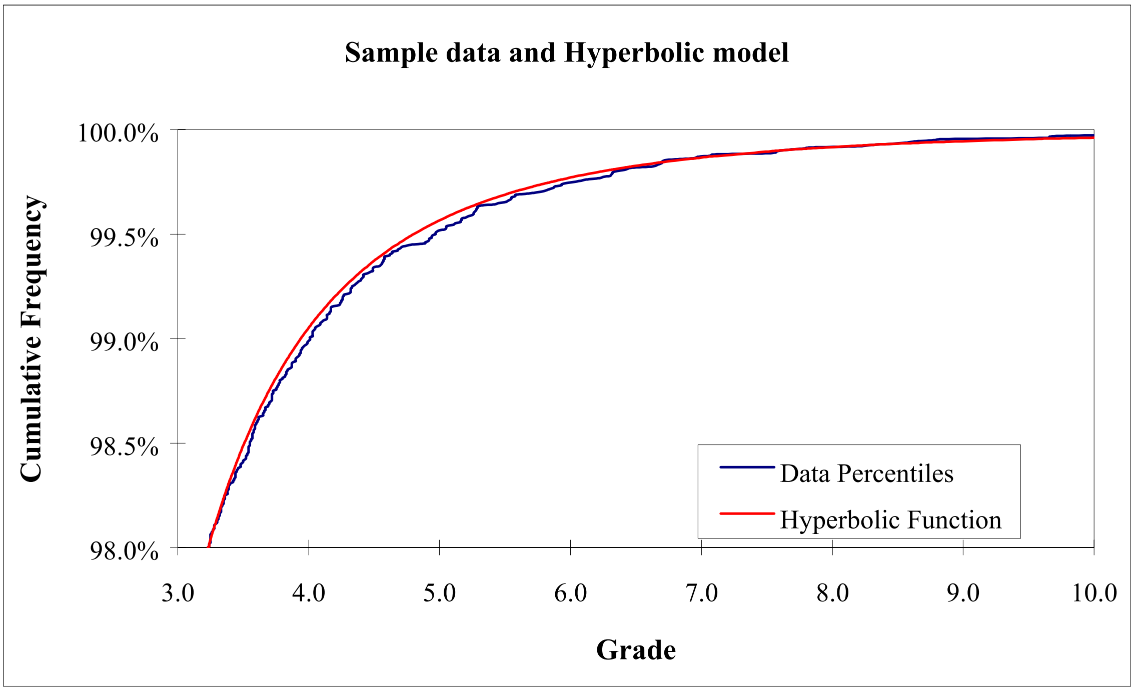

Figure 4. Example of a hyperbolic distribution fitted to the upper tail of a data set

In either case, the mean grade assigned to the uppermost class must be carefully considered if a good overall grade estimate is to result.

The data dilemma

When applying multiple indicator kriging to resource estimation it is necessary to settle on a finite number of thresholds that adequately represent the input data distribution shape. There is always a trade-off between the number of thresholds selected and the time available for the required analysis.

In a typical gold mineralisation environment, the grade distribution of the mineralisation within a geological domain might, in the simplest case, be divided into decile classes based on grade or perhaps metal content. Additional indicators may be included to discriminate between components of mixed population distributions. The disadvantage of grade-based deciles is that many of the indicator grades will be concentrated at the lower grade end of the positively skewed grade distribution. Fewer indicator thresholds will represent the higher grades, which contain most of the metal. The advantage of metal-based indicator grade selection is the reduced emphasis on the lower grade end of the distribution and a better selection of indicators at grades associated with significant amounts of metal.

Metal content indicators also provide the advantage of revealing the amount of metal that is associated with the higher grades. Metal indicators for gold projects often reveal that 30 percent of the metal is derived from grades above the 90^{th} grade percentile. In some cases, up to 70 percent of the metal may be above the 90^{th} grade percentile. The recognition of the amount of metal that is dependent upon so few samples in an exploration drillhole database is quite often an eye-opener!

The disadvantage of metal-based indicators is that, in practical terms, it is often impossible to define the continuity of grades for indicators set at or above the 90^{th} grade percentile, due to the low numbers of high grade samples, and their sparse spatial distribution. The outcome is often that the spatial distribution of higher grade mineralisation, and thus the most significant part of the distribution (in terms of metal content) is estimated using poorly-defined or assumed grade continuity conditions. This in itself is not a fatal flaw, but requires some judgement and experience.

This resource estimation problem is revealed by indicator analysis, but is not unique to the method. All resource estimation methods suffer from this weakness; the indicator approach only serves to highlight the issue.

The only real answer to this dilemma is to collect more data. Enough close-spaced data needs to be collected from representative areas to allow the better definition of the high-grade continuity and the short-range continuity of lower-grade indicators. The other possibility that exists is to borrow continuity models from other sources such as in-pit grade control data located in similar geological environments with similar overall grade distributions.

Unfortunately, a well-defined model of grade continuity is only part of the answer. If sampling density in the area being modelled is at typical exploration data levels, then the high grade continuity is likely to be shorter than the sample spacing. Unfortunately there is no easy solution to this dilemma, although conditional simulation is an approach which offers some promise.

Median IK

Median indicator kriging (Median IK) uses the median indicator variogram to define the continuity conditions for all indicators. This method is a simplified form of MIK that might be considered in the early stages of a resource project, when sample data is sparse and it is difficult or impossible to define grade continuity for a full range of indicators. The median indicator variogram is typically the most robust of all indicators, it tends to have the greatest range of continuity, and it is the easiest to define with some confidence from sparse data.

The application of variograms from a single indicator to all thresholds reveals the main assumption associated with the median indicator method. This is that the direction and range of continuity does not vary with changing thresholds. Experience from full indicator variography studies shows that grade continuity almost always varies with indicator grade, and invariably declines with increasing indicator grade. This will tend to result in an overestimation of the quantiles of the upper grade classes with Median IK, resulting in a higher-than-normal expected grade. In practical terms, Median IK is not a recommended technique where the data permits full estimation of a set of indicator variograms.

Use of a mineralisation indicator and choice of cut-offs

As mentioned above, in many geological estimation problems there may be no way to clearly define mineralogical or geological domains within which the spatial continuity and grade behaviour is more consistent than elsewhere. A typical scenario is an advanced exploration prospect with demonstrated geological and grade continuity, yet in which the detailed structures controlling and constraining mineralisation are unknown. Ideally in such a situation, the preference is to allocate some form of domain within which local controls on mineralisation may be applied, even if this is only a grade envelope. However, it is sometimes physically impossible to clearly separate ore-grade material from low-grade background material. In such a situation, descriptive statistics and geostatistics are often biased by the multitude of trace to low-grade assays, and this can also affect the choice of indicator cut-offs.

One solution to this is to apply a mineralisation cut-off grade to the bottom end of the grade distribution. This will separate possibly-mineralised from clearly-background material. Values below this cut-off are rejected for statistical and geostatistical analysis. This leads to better-defined distributions, and assists with the selection of appropriate indicator cut-offs.

Change of support

The most trenchant criticism of IK by its detractors (for instance Dowd, 1992, or Matheron, 1982) is the issue of change of support – that is, the generation of a distribution of grades for non-point (block, *sensu stricto*) data. Unlike ordinary kriging, inverse distance, and other linear estimation methods, the indicator transform is non-linear, as is the logarithmic transform, the uniform transform, and the normal scores transform. The consequence of this is that one cannot average indicators linearly, and thus cannot obtain a block distribution by averaging a series of point distributions derived at a smaller scale. The Ordinary Kriging (or inverse distance) corollary is to discretise a block by a series of points, estimate the grade at each point, and carry out the arithmetic mean to derive the block grade. However, if the statistic required from the IK distribution at each point (or sub-block) is known and fixed, such as the median or the mean value, there is nothing to stop the practitioner subdividing the block into a sub-blocks (actually points), estimating the ccdf by MIK at each point, deriving the desired statistic, and then averaging the result. This is shown diagrammatically in Figure 5. One downside of this approach is the extra computational effort.

A short cut would be to use the same data configuration for each sub-block, thus yielding only one matrix inversion on the left-hand side of the kriging system; the difference between the sub-blocks would be in the different weights obtained due to the differing positions of each sub-block in the parent block. Note that the objective of actually obtaining a distribution of uncertainty for each larger block (as shown in the top illustration of Figure 5) is a much more difficult task than simply correcting a series of mean values.

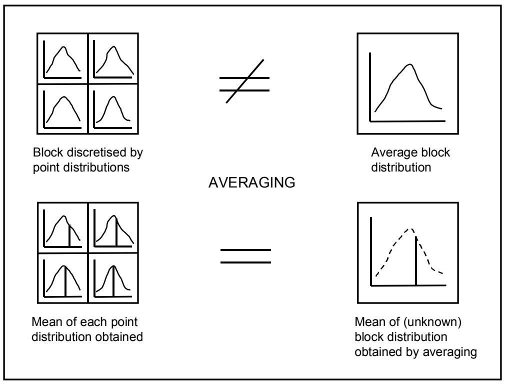

Figure 5. Diagrammatic illustration of permitted and non-permitted changes of support during Multiple indicator kriging

The traditional approach to the change of support in MIK has been to apply a variance correction factor on a global basis to the point statistics, exactly as one might do with point kriged data. There are a number of common techniques for achieving this. Perhaps the most widely used is the affine correction, a simple factoring of variance from point to (theoretical) block variance. Another approach is to use the indirect lognormal correction as described in Isaaks and Srivastava (1989, p 472).

The most elegant and theoretically correct solution to deriving block values is to move beyond the realm of estimation to that of sequential indicator simulation - this approach provides a truly local change of support, conditional only upon values in the neighbourhood. However, this approach is beyond the limits of the topic of indicator kriging addressed in this paper.

Variogram inference

Another criticism of indicator kriging, which paradoxically is a direct consequence of one of its great benefits, is the need to calculate a variogram and develop a model for each threshold value. For multiple (10 to 12) thresholds this was once a major exercise. However, advances in both computer speed and memory capacity and in software technology have produced new generations of fast and efficient variogram generation and modelling software. Such software allows the calculation of a full set of indicator variograms over all thresholds in one pass; this allows the practitioner to iterate between the various cut-offs and ensure a smooth variation in the variogram parameters. This is turn will serve to minimise the number of order relations corrections required. The combination of modern software and hardware has all but eliminated the time penalty of variogram inference, and a full set of thresholds may now be generated and modelled in a few hours for a moderately-large data set, say up to a few thousand data. Other benefits of variogram calculation and modelling software include the provision of *relative* indicator variograms, scaled to a sill of one, which also eases the variogram modelling burden.

Other applications of IK

Categorical Kriging

The indicator transform also lends itself to the estimation of *categorical* data – in other words, variables which are not continuous but which have discrete values. Some examples of categorical data are the presence or absence of a rock type (direct binary data requiring no transform), or a series of lithological or facies codes, or mineral sands hardness values. In this case, instead of kriging indicators at a set of thresholds, categorical IK will produce the probability of a given rock type or domain code occurring at a given location. Thus it is possible to produce probability maps for given lithologies based upon actual rock code data. This may be combined with indicator estimation or simulation of grade data, as described in Dowd (1996).

Soft IK

Soft indicator kriging, or soft kriging, is an area which holds great promise. Because of the common coding of all data types, both 'hard' data values (i.e. those of the variable to be estimated) and 'soft' data values (those correlated to the variable to be estimated) may both be used in the same indicator framework. The critical step is to 'calibrate' the soft data from a scatter plot of soft and hard data.

In a recent case study, the objective was to estimate cyanide-soluble copper values from a combination of few cyanide-soluble assays (the hard data) and many total copper assays (the soft data). Figure 6 shows a smoothed scatter plot between the cyanide–soluble copper data (on the Y axis) and the copper data (on the X axis), with shading which represents the probability. This was used to derive 'soft' indicator values for the copper data which were valued anywhere between 0 and 1, not just 0 or 1 as with the hard cyanide-soluble copper values. These two data sets were then integrated into a multiple indicator kriging approach for the estimation of cyanide-soluble copper. The result was a significantly-improved definition of the cyanide-soluble copper than if only the limited hard data were used.

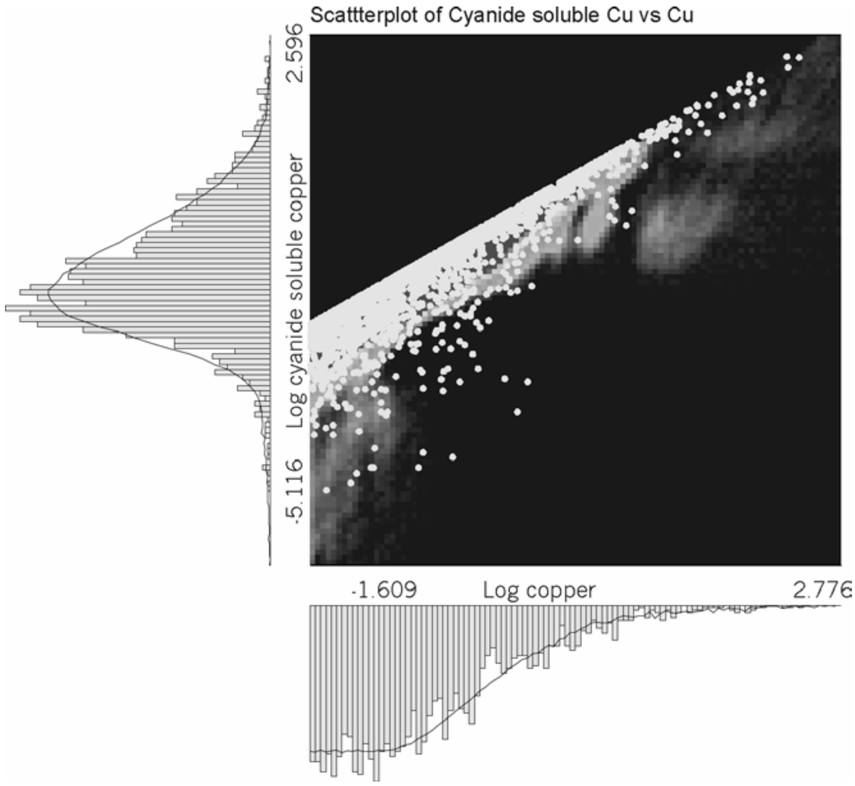

Figure 6. Smoothed bivariate scatter plot between cyanide-soluble copper (Y axis) and copper, used to derive soft indicator values for the copper samples.

Conclusions

Indicator kriging is now widely used in the mining industry as an estimation technique over a wide range of deposits and environments, because it offers practical solutions to some common estimation problems. In particular the issue of mixed or poorly-domained distributions, and the general trend away from the so-called parametric techniques (particularly after some spectacular disasters), has probably enhanced the acceptance of IK. The appeal of being able to generate (at least in theory) 'recoverable' resources has undoubtedly contributed to the popularity of the approach.

However, the use of IK is has its downside. Particular criticisms have been the relative difficulty of deriving true block distributions, the nuisance of order relations, and the sheer work involved in inferring variogram models at multiple thresholds. The practitioners in the industry, aided and abetted by academic research, have come up with solutions to almost all of the perceived problems, some more elegant than others.

Notwithstanding these successful developments, the IK theory has now largely been superseded and improved by the conditional simulation paradigm. Conditional simulation offers all of the advantages of IK and more. The single drawback – still a major issue at most sites – is the quantum leap in processing time and computing power required for the successful implementation of a simulation approach. However, even this is becoming a diminishing problem as computers exponentially increase in speed and memory capacity. It is predicted that within five years, indicator simulation and other simulation algorithms will be as commonplace in practical situations as indicator kriging is today. The extra dimension of simulation is hard to resist.

Acknowledgements

The management of Snowden Associates is thanked for its permission to publish this paper. The paper benefited from a review by Fleur Dyer.

References

Dimitrakopoulos, R, 1997. Indicator kriging course notes, W H Bryan Centre, University of Queensland

Dimitrakopoulos, R, and Dagbert, M, 1992. Sequential modelling of relative indicator variables: dealing with multiple lithology types, in Geostatistics Troia '92, Soares, A Editor, Kluwer 1993.

Dowd, P A, 1996. Structural controls in the geostatistical simulation of mineral deposits, in Baafi E Y, and Schofield, N A (Eds), Geostatistics Wollongong '96, pp 647-657.

Dowd, P A, 1992. A review of recent developments in geostatistics, Computers and Geosciences Vol. 17, No. 10, p. 1481-1500.

Isaaks E H, 1984. Risk qualified mappings for hazardous waste sites: a case study in distribution-free geostatistics. Master's thesis, Stanford University, Stanford, California, USA.

Isaaks E H and Srivastava, R M, 1989. An introduction to applied geostatistics. Oxford University Press, New York.

Journel , A G, 1983. Nonparametric estimation of spatial distributions, Math. Geol Vol 15, No. 3, p. 445-468.

Journel A G, and Deutsch C V, 1996. Rank order geostatistics: a proposal for a unique coding and common processing of diverse data, in Baafi E Y, and Schofield, N A (Eds), Geostatistics Wollongong '96, pp 174-187.

Journel A G, and Alabert F, 1989. Non-Gaussian data expansion in the earth sciences, Terra Nova 1: 123-34.

Journel, A G, and Deutsch, C V, 1998. GSLIB Geostatistical software library and users guide, Second Edition, Oxford University Press, New York.

Matheron, G, 1982. La destructuration des hautes teneurs et le krigeage des indicatrices, Centre de Geostatistique et de Morphologie Mathematique, Note N-761, 33p.

Rivoirard, J, 1992.. Relations between the indicators related to a regionalised variable, in Geostatistics Troia '92, Soares, A Editor, Kluwer 1993.

Switzer, P, 1977. Estimation of distribution functions from correlated data. Bull. Inter. Stat. Inst., Vol 47, No. 2, p. 123-137.