GEOSTATISTICAL MODELING OF MULTIPLE VARIABLES IN PRESENCE OF COMPLEX TRENDS AND MINERALOGICAL CONSTRAINTS

G. Lyall
Anglo American Chile Ltda, Santiago, CHILE (glyall@mantos.cl)

C. V. Deutsch
University of Alberta, Edmonton, Alberta, CANADA (cdeutsch@civil.ualberta.ca)

Detailed models of mineral grades are critical to mine planning, that is, estimation of recoverable reserves, production scheduling, and assessment of blending requirements. Geostatistical simulation is being used increasingly to create such models. The tendency toward simulation may be attributed to two significant advantages. Firstly, simulated realizations reproduce the correct patterns of variability, that is, the histograms, scatterplots, and variograms of the mineral grades. Secondly, multiple realizations may be created to provide a measure of uncertainty.

Notwithstanding recent advances in simulation methodology, there remain practical issues that must be addressed; we discuss some in this paper. The spatial distribution of mineral grades is partly deterministic (vertical or areal trends) and partly stochastic. Trends must be entered explicitly in geostatistical simulation; otherwise, they will not be captured in the final models. The stochastic, or more random, component of the grade variation is suited to simulation. There may exist, however, local variations in the nature of the randomness. For example, the correlation between multiple variables may change in a predictable manner. We demonstrate how such complex trends may be used in geostatistical modeling, in particular, in Gaussian simulation.

Data from a nickel laterite deposit are used for illustration. We demonstrate the geostatistical modeling procedure to create realizations of nickel, iron, magnesia and silica with the correct patterns of spatial variability. The data exhibit clear vertical trends due to weathering, areal trends, vertically varying correlation between the grades, and mineralogical constraints on the grades.

INTRODUCTION

Procedures to construct geostatistical models of mineral deposits are changing. Historically, ordinary kriging (OK) was used extensively. Recently, conditional simulation methods are gaining popularity since they (1) avoid the smoothing or "information effect" of kriging, (2) provide an unbiased estimate of recoverable reserves, (3) permit an assessment of uncertainty, and (4) make it possible to build models of multiple variables with the correct relationships between the variables. Kriging is still used at the heart of conditional simulation methods to construct the needed distributions of uncertainty. Nevertheless, experience gained in application of kriging methods is not directly transferable to simulation. This is particularly true when considering trends. OK within restricted search neighborhoods is remarkably robust at capturing trends and other local variations in the mineral grades; however, the use of OK in simulation is not as robust because of a greater reliance on

WJ Kleingeld and DG Krige (eds.), Geostats 2000 Cape Town, Volume 2, 837-846

the kriging variance and, implicitly, on the decision of stationarity. The books by Journel and Huijbregts, 1978, Isaaks and Srivastava, 1987, and Goovaerts, 1997 provide good background references.

For simulation purposes, trends in average grades are commonly dealt with by deterministic modeling of locally varying trends followed by stochastic simulation of residuals of the trend. Real simulated values are obtained by adding trend back to the simulated residuals. On the other hand, trends in variance or covariance relationships can be handled by allowing local adjustment of specific parameters in the simulation algorithm.

Correlations between multiple variables are affected by physical constraints. Some variables show negative correlation because they compete for space within the constant mass. In a similar way, constraints are present due to mineralogical relations and geochemical transformations. Where possible, these relationships should be imposed explicitly on geostatistical models by transforming the original variables into multivariate re-expression variables (Aitchison 1982; Butler, 1981; Mosteller and Tukey, 1977; Rosenblatt, 1952). Such re-expression is uncommon in current geostatistical practice; we show an example with a Nickel laterite deposit.

Nickel laterite deposits are formed by sub-surface weathering and, as a result, show characteristic vertical trends in grades and mineralogy. Models that honor average grades, variability, trends and the multivariate relationships, are necessary for assessing required blending strategies for a pyrometallurgical process. The case study example shows how such characteristics can be reproduced using simulation methods.

RE-EXPRESSION OF VARIABLES FOR MINERALOGICAL CONSTRAINTS

The first important step in geostatistical modeling is to establish the correct "variable" to model. In many cases the choice of variables is evident; however, mineralogical constraints require that original data variables be re-expressed or transformed. For example, there is often a *sum constraint*, that is, the grades may not sum to more than 100%. Moreover, there may be *ratio constraints*, that is, the ratio of certain minerals may not exceed a certain value for stoichiometric reasons.

The constant-sum constraint becomes important when the chemical assays constitute a "whole rock analysis" or the assays sum to a substantial proportion of the total rock mass. In such a case, certain variables compete for space within the rock mass inducing negative correlations.

Mineralogical constraints are induced when the variables are present in specific minerals. Consider for example a rock containing two mineral species $(X)_2Y$ and XY where the constituents X and Y are assumed to have identical atomic masses. Assays are performed for variables X and Y. A sample containing the mineral $(X)_2Y$ will have a X to Y ratio of 2. A sample containing the mineral XY will have a ratio of 1. It follows therefore that any rock containing a mixture of minerals XY and $(X)_2Y$ will have ratios of X to Y constrained between 1 and 2.

To illustrate the modeling of such multivariate constraints we will consider an example shown in Luster, 1985. This example shows an interesting link to real mineralogy. In practice, as in the case study described later on, the actual mineralogical constraints may not be known, however the relationships can be deduced from scatterplots. In Luster's example, analyses for $MgCO_3$ and $CaCO_3$ are required for limestone quality control; the only two mineral species which contain these compounds being dolomite ($CaMg(CO_3)_2$) and calcite ($CaCO_3$). By stoichiometric calculation it can be shown that a rock composed entirely of dolomite contains 46% $MgCO_3$ and 54% $CaCO_3$. It follows that for a limestone composed of dolomite plus some calcite, the $\%MgCO_3$ will be less than 46%. These constraints are illustrated on Figure 1, below, and can be represented by the following expressions.

$$0\% < CaCO3+MgCO_3 < 100\% \qquad \text{and} \qquad 0\% < MgCO_3/(MgCO_3+CaCO_3)<0.46$$

The original variables are therefore re-expressed as the following two nested variables:

$$U_1 = CaCO3+MgCO_3 \qquad \text{and} \qquad U_2 = MgCO_3/(MgCO_3+CaCO_3) = MgCO_3/U_1$$

$CaCO_3$ and $MgCO_3$ contents are calculated from U_1 and U_2 by inverting the above expressions:

$$MgCO_3 = U_2.U_1 \qquad \text{and} \qquad CaCO3 = U_1 - MgCO_3$$

From Figure 1 it can be seen that sum variable U_1 is plotted on the *x*-axis and is constrained by the maximum of 100%. U_2 is the gradient of a line passing through the origin. Note that all data fall below the constraint line with gradient 0.46. Imprecision of the analyses or erroneous data may cause some data to plot outside the constraint boundaries.

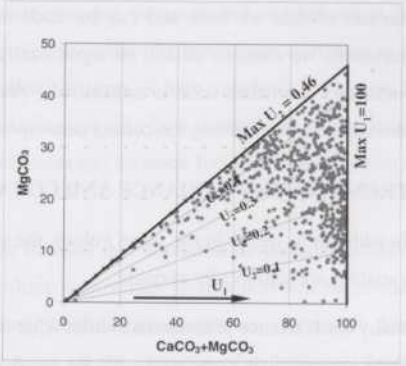

Figure 1: The relationship between the re-expressions, U_1 and U_2, and the original $CaCO_3$ and $MgCO_3$ contents.

There are a number of alternative nested re-expressions that can be devised to deal with more than two variables. We encounter this below.

The re-expressed variables may be simulated with a variety of geostatistical algorithms. A straightforward alternative consists of (1) normal score transformation, (2) sequential gaussian co-simulation, and (3) back transformation. A warning: some of the "U" re-expression variables are non-linear transforms of the original variables; therefore, a bias will be introduced if U-averages are directly back transformed. This is not a problem in simulation because the full variability is preserved.

TRENDS IN THE MEAN

Perhaps the most important step in geostatistics is ensuring that the variable being modeled is stationary over the study area. Geologic modelers typically adopt one of two attitudes toward stationarity: (1) any decision of stationarity is unacceptable because the variation patterns are too complex for statistical modeling, or (2) group the data by major rock type and let local variation patterns be controlled by the data themselves. Both views have validity, but may be too extreme. There are complex patterns of variation that are important and yet the data must be grouped together somehow to proceed with any numerical modeling.

A practical approach is to group the data together by major rock types and yet explicitly model deterministic trends within each rock type. The Z-grade variable is split into a mean or trend component m and a stochastic residual component R, that is, at every location $\mathbf{u}$: $Z(\mathbf{u}) = m(\mathbf{u}) + R(\mathbf{u})$. The locally varying mean $m(\mathbf{u})$ captures our geologic understanding of trends, which are no more than departures from stationarity. We show later an example of a clear vertical trend due to surface weathering; the grades decrease with depth. In areas of less data, simulation would not honor this trend without our explicit modeling.

Clearly, we should not overdo trend modeling. Our understanding of the mean $m(\mathbf{u})$ changes with the amount of data we have and can be made more precise with many data. The scale of the mean variations we consider should be significantly larger than the data spacing. Explicit accounting of variations at smaller scale is unjustified. Practical implementation and cross validation may be of assistance in determining the correct decomposition of the grades into mean and residual.

TRENDS IN THE VARIANCE AND COVARIANCE

In addition to trends in the mean values, the variability or covariance relationships between multiple variables are also locally variable.

Many earth science phenomena exhibit what is commonly termed as the "proportional effect," that is, local variability is correlated with the local average. Although normal-score transformation often mitigates the proportional effect, the variability remains greater in high grade areas and lesser in low grade areas. Trends can become even more complicated when dealing with multiple variables. In such cases, it is likely that correlation characteristics also show trends. The most common of these trends is related to the proportional effect. Areas showing lower grades and variability also tend to exhibit lower correlation coefficients.

In some cases, not only will the intensity of correlation show trends, but the nature (or sign) of correlation may also change. Such cases are usually indicative of different populations and a geological explanation should be sought. We may be able to separate the study area into different populations; however, a gradual change in the correlation coefficient is more difficult to handle. The nickel laterite case study described later shows such a trend.

Local variations in correlation could be handled by separating into different units; however, there may then be too few data for reliable statistical inference. In addition, hard boundary zoning would lead to unrealistic artifacts. Transitional boundaries could be modeled by considering correlations between adjacent units and simulating each unit, correlated to the adjacent unit, in a sequential fashion. The simulation procedure becomes complex; we have to consider the multivariate correlations within each individual unit and cross-correlations between adjacent units. A straightforward alternative is to make parameters locally variable in simulation program.

SIMULATION WITH TRENDS

Sequential Gaussian simulation is a popular geostatistical simulation algorithm. A detailed description and program can be found in GSLIB (Deutsch and Journel, 1997). The procedure consists of sequentially visiting all grid nodes. A (co) kriging is performed at each grid node to define the normal distribution(s) of uncertainty. Simulated values are obtained by Monte Carlo simulation from the normal distribution(s). A full "linear model of coregionization" or a simplified Markov model of coregionization may be used for cokriging. Increasingly, the simpler Markov model or collocated cokriging is being used.

Trends in the mean may be handled by working with residuals from locally variable mean values. Trends in variance and covariance are imposed by locally varying variances and correlations in the simulation process. Variance correction factors are applied to tighten the conditional distributions in areas showing low variability and expand the distributions in areas of high variability. The use of variance reduction factors is equivalent to scaling the variogram sill. In a similar way, whilst using the collocated simulation approach, the correlation coefficient can be made locally variable, which is equivalent to scaling the cross-variogram sill.

In carrying out cosimulation exercises one should bear in mind that normal-score transformation of residual data variables does not guarantee that the resulting multivariate distribution will be normal. Bivariate normality is indicated by a scatterplot with elliptical Gaussian contours. Departure from multivariate normality leads to artifacts such as distortions of the univariate distributions. Luster (1985) describes checks for multivariate normality.

CASE STUDY ON A NICKEL LATERITE DEPOSIT

Nickel laterite deposits form by surface weathering and leaching processes in tropical and sub-tropical climates. These phenomena have created three main mineralized units in the case study, which can be pictured on the cross section in Figure 2. Typical vertical trends in nickel and iron grades are also shown. The study area, containing a fairly regular drill hole spacing of 25x50 meters is shown in Figure 3. Drilling in this area aimed at intersecting the whole of the laterite and saprolite units, which generally does not reach more than 30 meters below the surface.

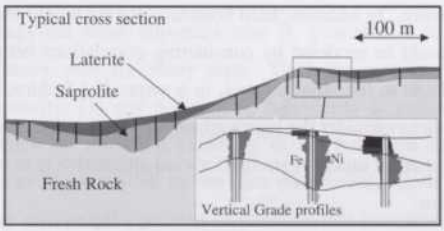

Figure 2: Typical cross section and example vertical grade profiles. The grades of iron and nickel decrease with depth through the saprolite.

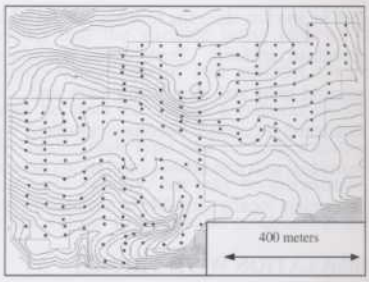

Figure 3: Topographic map of the study area with the drillhole locations

The unweathered fresh rock at the base has a dunitic to peridotitic composition, of which the principal constituents are approximately 40% SiO_2, 35% MgO and 8% Fe. Note that these constituents may be present in more than one mineral. Nickel grades in this unit are sub-economic. The laterite unit is characterized by Fe enrichment (>30%) and SiO_2 and MgO depletion (generally both < 10%). The highest Ni grades are encountered within the saprolite zone, which shows compositions between fresh rock and laterite.

Nickel will be recovered from the higher grade saprolite unit by a direct smelting process, which is sensitive to the composition of the feed material, particularly Fe content and the SiO_2/MgO ratio. These issues impose a need for blending of the ore prior to smelting. Simulating the spatial multivariate relationships between Fe, SiO_2, MgO and Ni is required as a first step in determining the appropriate blending strategy for the plant. It is essential that the simulated realizations honor vertical and areal trends in grades and correlation characteristics between them.

Whole rock chemical assays of drill hole cuttings include analyses for %Ni, %Fe, %MgO and %SiO_2. Histograms and correlation plots included in Figure 4 summarize the multivariate statistics for Ni, Fe, SiO_2 and MgO within the saprolite unit. The scatterplots between Ni and the other variables suggest the presence of two populations. This is most likely a geological distinction that can be attributed to transitions of the saprolite unit with the overlying laterite unit. Some consideration will be given to this in the stochastic simulation exercises. The scatterplots between Fe, SiO_2 and MgO show strong correlations, and characteristics that suggest mineralogical constraints. It should also be noted that the sum of SiO_2 + MgO + Fe amounts to approximately 80% of the total rock mass.

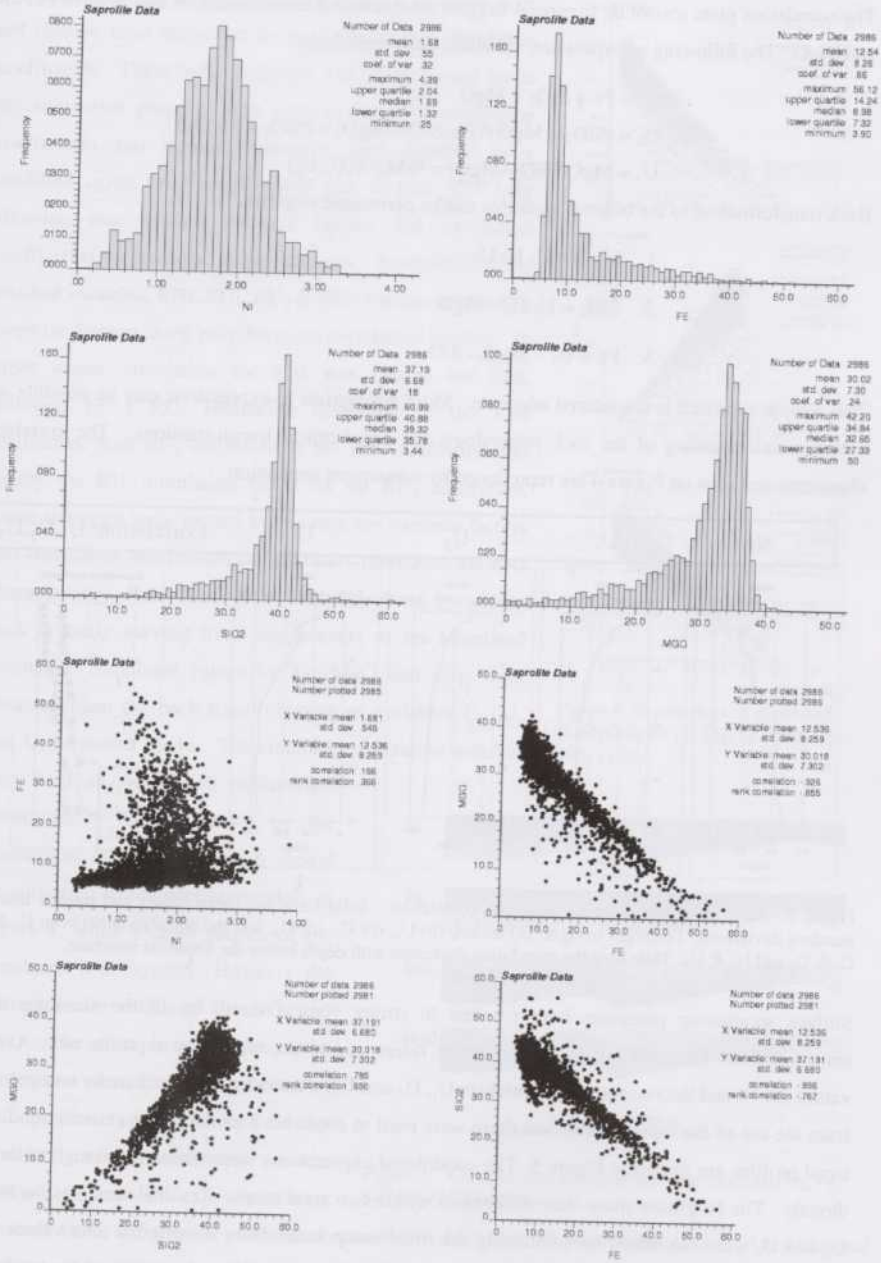

Figure 4: Histograms and scatterplots of the grades in the saprolite unit. The histograms are of Ni, Fe, SiO_2, and MgO. Scatterplots reveal the mineralogical constraints.

The correlation plots shown on Figures 4 suggest mineralogical transformations that involve Fe, SiO_2 and MgO. The following re-expression variables were considered:

$$U_1 = Fe + SiO_2 + MgO$$
$$U_2 = (SiO_2 + MgO)/(Fe+SiO_2+MgO) = (SiO_2 + MgO)/U_1$$
$$U_3 = MgO/(SiO_2+MgO) = \%MgO/(U_1.U_2)$$

Back transformation to the original variables can be performed stepwise:

1. $MgO = U_1.U_2.U_3$
2. $SiO_2 = U_1.U_2 - MgO$
3. $Fe = U_1 - MgO - SiO_2$

This simple approach is considered adequate. More appropriate re-expressions may be possible with a better understanding of the rock mineralogy and geochemical transformations. The correlation characteristics seen on Figure 4 are reproduced by subsequent simulation.

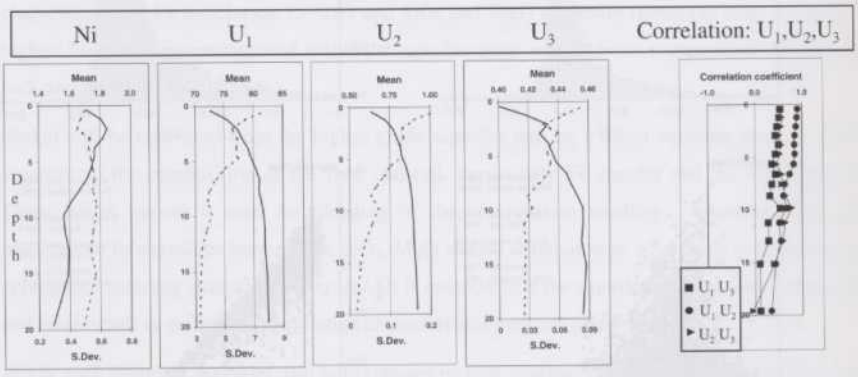

Figure 5: vertical trends in mean value and correlation. Solid lines are mean values and dashed lines are standard deviations. From left to right: (a) nickel, (b) U_1, (c) U_2, (d) U_3, and the correlation between U_1 & U_2, U_1 & U_3 and U_2 & U_3. Note how the correlation decreases with depth below the Saprolite interface.

Surface weathering processes have resulted in strong vertical trends in all the variables under consideration. The drill hole elevations were referenced to the top of the saprolite unit. Average values for Ni and the re-expression variables U_1, U_2 and U_3, were calculated for 2 meter vertical slices from the top of the saprolite unit and these were used to construct a model of the vertical trend. The trend profiles are shown in Figure 5. The conditional expectations were "smooth" enough to be used directly. The Ni grades show clear differences within two areal zones. Residual variables for Ni, U_1, U_2 and U_3 were calculated by subtracting the trend components from the original data values. The residuals have a stationary mean of zero. The trend component explains about 40% of the total grade variability.

The transformation of the grades and removal of trends does not remove local variations in the variance and correlation coefficients. These local variations will be accounted for in the simulation process. The vertical trends in correlation coefficients and residual variances were handled in a modified *sgsim* program (Deutsch and Journel, 1997) by allowing user defined variance factors and correlation coefficients for specified depth intervals. Simulation of the residual variables, RNi, RU_1, RU_2 & RU_3 was carried out in a stepwise fashion using only bivariate correlation models. In other words, simulation for RNi was carried out first, followed by a RU_1 simulation correlated to the RNi realization, then RU_2 correlated to the RU_1 simulation and finally the RU_3 simulation based on the RU_2 realization. Some iterations were needed to fine tune the variance factors and correlation coefficients to accurately reproduce the data characteristics. Real data values were calculated by adding back spatially varying trend components to the simulated residuals. Simulated values for Fe, MgO and SiO_2 were obtained from the back transformation of variables U_1, U_2 and U_3 as shown above. The simulated realization matches very well all trends and multivariate constraints. This is shown by the scatterplots in Figure 6, which should be compared to the data scatterplots in Figure 4, while a cross section of the simulation is in Figure 7. However, the histograms for simulated Fe, SiO_2 and MgO suffer some distortion due to the assumption of multivariate gaussianity. This problem could be resolved by modeling, as an alternative, nested bivariate gaussian re-expressions obtained from a "stepwise-conditional transform" described by Rosenblatt (1952). An additional problem arises from the independent modeling of trend and residuals, which leads to a small percentage of values that need to be corrected to the data minimum and maximum values.

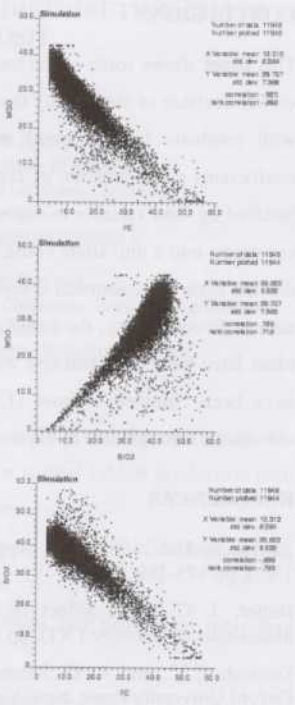

Figure 6: Scatterplots of simulated (1)MgO vs Fe (2) MgO vs SiO_2 (3) SiO_2 vs Fe

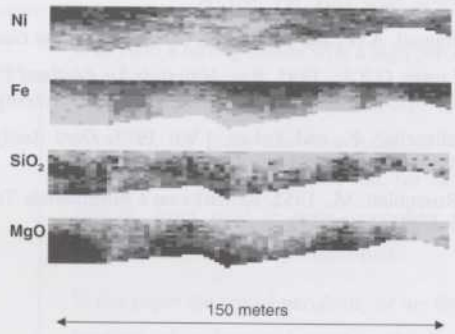

Figure 7: Cross section example showing simulated Ni, Fe, SiO_2 and MgO realization

CONCLUSIONS

This paper shows some of the tradecraft necessary to obtain reliable models. Simulation relies more on our decision of stationarity than the historic workhorse of geostatistics: ordinary kriging. Working with residuals from a trend model and considering locally varying variance and correlation coefficients allow creation of complex, yet realistic, orebody models. Mineralogical constraints are handled by data variable re-expression. In essense, re-expression amounts to converting a bivariate constraint into a univariate constraint. Honoring the histograms of the re-expressed variables ensures that the bivariate constraint is reproduced. Normal-score transformation of the re-expressed variables adds little complexity; the transformations must be reversed in the correct order. These two important ideas have been demonstrated with data from a Nickel Laterite deposit. Reliable models could not have been obtained without (1) complex trend modeling, and (2) accounting for mineralogical constraints through data re-expression.

REFERENCES

Aitchison, J. A, 1981, New Approach to Null Correlations of Proportions, *Mathematical Geology* 13(2), pp 175-189.

Butler, J. C, 1981, Effect of Various Transformations on the Analysis of Percentage Data, *Mathematical Geology* 13(1), pp 53-68.

Deutsch, C.V., and A. G. Journel, 1997, *GSLIB: Geostatistical Software Library*, Second Edition, Oxford University Press, New York, NY, 369 pp.

Goovaerts, P., 1997, *Geostatistics for Natural Resources Evaluation*, Oxford University Press, New York, NY, 483 pp.

Isaaks, E.I., and Srivastava, R.M., 1987, *An Introduction to Applied Geostatistics*, Oxford University Press, New York, NY, 561 pp.

Journel, A.G., and Huijbregts, C.J.,1978, *Mining Geostatistics*, Academic Press, London, 600 pp.

Luster, G.R.L., 1985, *Raw Materials for Portland Cement: Applications of Conditional Simulation of Coregionalization*, Ph.D. Thesis, Stanford University, 531 pp.

Mosteller, F., and Tukey, J.W., 1977, *Data Analysis and Regression*, Addison-Wesley, Reading, Massachusetts, 588 pp.

Rosenblatt, M., 1952, Remarks on a Multivariate Transformation, Annals of Mathematical Statistics, vol. 23, pp 470-472.