INDICATOR SIMULATIONS OF CATEGORICAL VARIABLES

Mario E. Rossi
Principal Geostatistician
GeoSystems International
15998 Mataro Bay Ct.
Delray Beach, FL, 33446, USA
Phone: 1-561-495-8797
Fax: 1-561-498-1262
Email: mrossi@geosysint.com

August 2004

ABSTRACT
This paper describes an implementation of the Sequential Indicator Simulation method used to characterize the uncertainty of a geologic model. An interpreted, deterministic geologic model is routinely used in the context of ore resource evaluation to aid in the estimation of grade and tonnages above cutoff at different stages of project development. This is particularly consequential for those cases where the modeled geologic attribute is shown to be a significant grade control.

Simulation models are constructed to assess the range of possible values at any given location. In the case of a geologic model, the variables of interest are categorical (non continuous), which characterizes the simulation models constructed. It is argued that geologic attributes are used to help define the grade stationarity (estimation) domains, which often results in the modelled geology becoming the most important variable in the definition of resource and reserve tonnage above cut-off.

An example of the uncertainty of the deterministic lithology model used in support of grade estimation at the Minera Michilla mine is presented. The lithology uncertainty modeled was then used to assess resource risk and its impact on the Mine Plan (reserve risk), as well as the predicted uncertainty of the official resource and reserve statement of the mine.

1 Introduction

The Lince-Estefanía mine is located in the Coastal Cordillera of Chile, at 900m elevation above sea level, in a district where Cu has been mined for several decades. The closest major city is Antofagasta, approximately 120 km to the south. The mine is currently operated by Minera Michilla S.A.

The Lince-Estefanía deposit is subdivided into 17 main zones for both the open pit and the underground portions. These sub zones correspond to areas with somewhat different geologic characteristics (open pit) and/or according to the mining extraction sequence (underground).

The geology of the area is characterized by a very thick volcanic sequence regionally dipping 30° to the NW, and composed of a stratified series of andesites and volcanic breccias of different characteristics. The andesitic sequence varies from afanitic to porphyritic, with intercalated volcanic breccias. Lithology is the single-most important control for Cu mineralization, since it is hosted in this volcanic sequence, where breccias are the most porous rocks, and therefore the more favorable hosts (Ferraris and Di Biase, 1978).

Copper mineralization has been mostly deposited in small high-grade bodies, preferentially in volcanic breccias. However, significant Cu mineralization can also occur in andesites, which are less permeable rocks. The small mineralized bodies (“mantos”) are generally ellipsoidal in shape, 4 to 5m thick, with length and widths of up to 40 or 50m, but many less than 25m. It is difficult to predict the existence and the size of each manto, as well as its grade distribution. Mineralization is also strongly controlled by a series of important structures, principally the Muelle fault, striking N45°E to N60°E, and its conjugate set. Given this geologic setting and the drill hole grid currently used, traditional methods used to interpret and model the lithology sequence implicitly carry a high degree of uncertainty.

Figure 1 is a schematic view from underneath, showing the breccia mantos (in light grey), the Muelle fault zone and associated high-angle structures and dikes (in dark grey), and tectonic breccias and an intrusive body (in dark grey). Also, surface topography and the open pit and underground workings are shown in light brown.

In early 1999, after completing an infill drilling campaign, the geologic model was updated interpreting lithology units on section, followed by interpretation and validation on several control benches. A new resource block model was obtained, into which copper grades were estimated using the multiple indicator kriging method. Following the completion of the Resource Model and its classification into Measured, Indicated, and Inferred, a conditional simulation model was implemented to assess uncertainty of both the geologic and grade models, and corresponding risk on the resource and reserve statement of the mine. This conditional simulation model was also based on indicator kriging techniques, so that both the kriged and the simulation models would be based on the same basic Random Function model.

The simulation model was built in the volume corresponding to some key sectors and mining phases for the following 5 years of production, both from the open pit mine and for the underground mine. The lithology units were simulated at nodes on a small grid using the indicator sequential algorithm (Alabert, 1987), as implemented in a modified version of the GSLIB code (Deutsch and Journel, 1992). The simulated geology was then used to condition the uncertainty model for Cu grade simulation.

2 DATABASE AND GEOLOGIC MODEL

A total of 1,901 drill holes were available at the time of the conditional simulation study presented here. The database includes different types of drill holes, including conventional rotary (which are older holes), reverse circulation, and diamond drill holes. The quality of geologic mapping and sampling is different depending on drill hole type considered. In particular, mapping cuttings from reverse circulation and conventional rotary drill holes is less accurate than mapping core from diamond drill holes.

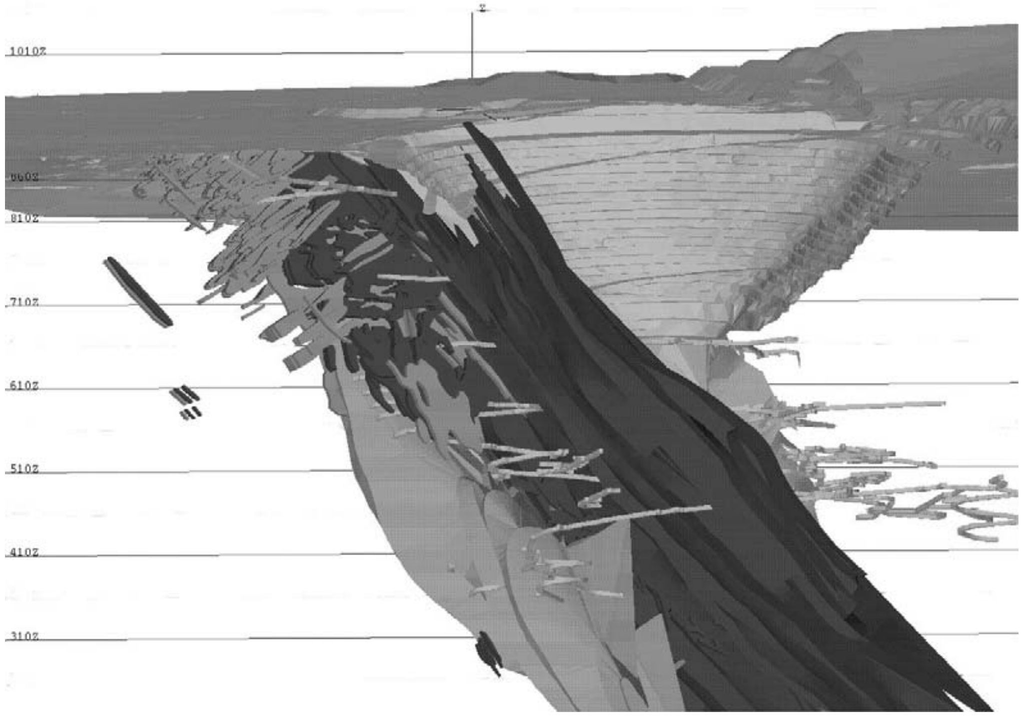

Figure 1: Schematic view of the Lince-Estefanía Ore body mantos (light grey) and tectonic breccias (dark grey). Surface Topography, Open Pit, and Underground workings are shown in light grey. The Muelle Fault Zone is shown in dark grey. Elevations to the left are meters above sea level and provide scale.

Traditionally, the sequence of steps required to obtain the lithology model is as follows:

1. The geologic codes incorporated into the database are checked and validated. These checks include identification of cases where the Cu grade for any given interval falls within an expected range. For example, it is extremely unusual to have grades in afanitic andesites above 2% TCu. These unusual intervals are flagged and checked.
2. The lithology units are then modeled in cross sections spaced every 25m, orientated along the main axis of mineralization according to a local, rotated coordinate system.
3. A simple solid modeling method is used to obtain a three-dimensional representation and volumes of each lithology.
4. This process is repeated to define a mineralized envelope, which encompasses all intervals with a TCu greater than 0.1%. The 0.1% TCu value is in fact a geochemical "mineralization threshold" value, as determined by several geologic studies. It indicates the volume within which Cu mineralization may occur, although it may not be economic grade. Outside the envelope there is no probability of any significant Cu mineralization occurring.
5. Several plan views of the modeled solids are obtained, which are used to assess continuity of the lithology units across sections, and validate the sectional interpretation.

The block model used to estimate resources utilizes a basic block size of 12.5x12.5x5m, with 3.125x3.125x2.5m sub-blocks at contacts. The solids representing the lithology units are used to flag each sub-block and to calculate the majority lithology for each main block. Cu grade is then estimated into the block model using different indicator variogram models and different kriging plans according to each lithology.

The model described does not include a measure of uncertainty, assuming that the polygons on section accurately depict the limits of the lithology units.

3 EXPLORATORY DATA ANALYSIS AND VARIOGRAPHY

3.1 Basic Statistics

The initial step in preparation for simulating the geologic sequence at Michilla is a statistical analysis of the drill hole database that allows understanding of the distribution of copper grades and the geologic controls imposed by the pseudo-stratigraphic sequence and rock permeability.

The original sample intervals were composited to 5m because this length is representative of the minable unit at the underground mine, which uses a cut-and-fill method on 5m lifts, and also at the Open Pit, where the bench height is 10m, but with ore selection occurring on half benches (5m) where necessary. The 5m composites were obtained by lithology units and truncating the composite when a transition to a different lithology occurs. This method avoids dilution from one geologic unit to the other, although a significant number of composites less than 5m in length occur.

The volcanic breccia composites have significantly higher average grade than other units, but yet the full grade range (from no mineralization to very high grade zones) is observed. A similar comment can be made about the other units, although their average grade is lower.

3.2 Relative Indicator Variograms for Lithology Units

Indicator variograms were obtained and modeled for each lithology for different sectors within the deposit. It was found that most areas have similar anisotropies and ranges, and therefore a single, deposit-wide model was used in obtaining the simulation models.

The main directions of anisotropy for the Andesites and Volcanic Breccias correspond to the orientations of the stratigraphic sequence. The anisotropy ratio for the ranges in the dipping plane is about 3:1 for the second variogram structure, with ranges between 100 and 250m. The main directions of anisotropy for the Intrusives and Tectonic Breccias also are as expected, orientated North-South, and dipping 60º.

The relative nugget effect and the variance factor of the first structure are high, and the second structure generally represents less than 20% of the total variance (intrusives) or less than 10% of the total variance (andesites and volcanic breccias). Indeed, the mine uses a 12.5x12.5m infill grid to confirm the presence and location of mineralized mantos to obtain short-term models before bringing into production certain areas.

The indicator variograms models were analyzed with mine geologists to compare with existing geologic knowledge. This is an important validation step, in that the variogram models should always be checked against geologic models.

4 SIMULATING GEOLOGIC UNITS

The purpose of simulating lithology units is to obtain a more representative model of resource and reserve uncertainty, which in turn can provide more realistic risk models. The simulated geologic model is defined by its application, and is not necessarily intended to faithfully reproduce geologic details. In mining, it is generally a major component of the overall uncertainty of the resource and reserve model.

4.1 Sequential Indicator Simulation

Sequential Indicator Simulation (SIS) technique is the most commonly used categorical simulation technique in mining applications. The SIS technique is well suited for simulating lithology units at Michilla because structural displacements of the lithology units (faults and dikes) are significant. They incorporate a high degree of variability in the sequence, adding significant short-scale variability, and making the transition much less predictable. In addition, the Lince-Estefanía resource model uses multiple indicator kriging to estimate grades into the block model. This is also an important consideration because each component of the overall uncertainty model should use the same Random Function (RF) model.

All uncertainty models are model-dependent, implying that choices made about the algorithms used to obtain the simulations, the selection of conditioning data, and other model-specific parameters will impact the final uncertainty model (Goovaerts, 1998, Rossi, 2003). Therefore, if an uncertainty model is built from several components (simulated geology plus simulated grades, for example, then all its components must share with the base case (resource model) as many of the basic assumptions as possible.

4.2 Definition of the Simulated Volumes

The lithology units and sub-zones modeled were defined according to knowledge gained from prior statistical studies that demonstrate the relationships between lithology and copper grades. Sub-zones of interest within the deposit where defined based on the mining sequence, rather than by geology. These are the volumes of interest for assessing the mine plan risk for future production years.

The simulated geologic model describes the main mineralization controls, which in this case are: a) Volcanic Breccias, including a porous andesite; b) Andesites and poorly mineralized breccias; and c) Intrusives and tectonic breccias.

In addition, there is a geochemical envelope that is used to define the limits of the deposit. The underlying assumption is that there is 0% probability of finding mineralization outside the envelope, although if inside the geochemical boundary there is no assurance that significant mineralization will be found. This can be expressed as:

$$Ind_{Cu=0.1\%} = \begin{cases} 1, \text{ if Cu} \geq 0.1\%, P_{\min} > 0\% \\ 0, \text{ if Cu} < 0.1\%, P_{\min} = 0\% \end{cases}$$

The geochemical Cu indicator takes a value of 1 if the expected grade is greater or equal than 0.1%, implying that the probability of finding mineralization (P_{min}) is greater than 0%, and takes a value of 0 if the expected grade is less than 0% Cu, and thus P_{min} is equal to 0%.

4.3 Definition of the Simulation Grid and Plan

The simulated geology and the later grade simulation both were done on the same grid. This grid has the same origin coordinate and rotation than the Resource Block Model, so that comparisons and evaluations on a block by block can be easily done.

The geometry and other specifications of the simulated models at Michilla are such that they coincide with the grid used to define the Resource Block model. Grid node Spacing is 5x5x5m, mostly because of practical limitations, with the total number of simulated nodes being 27.874.000. However, not all these nodes receive a simulated value, because two conditions were imposed on the simulation: a) that the node existed inside the 0.1% TCu grade envelope (discussed above), and b) that at least two composites (hard data) exist within the search neighbourhood. Given these conditions, the actual number of simulated nodes was close to 7M.

The simulated models were obtained using a modified version of the SIS routine of the GSLIB software (Deutsch and Journel, 1992). The modifications introduced relate to the handling of the input and output data, the addition of simulation parameters and options, and the definition of the nodes to be simulated. The algorithm itself was not modified.

The simulation search neighborhood was defined according to the main orientation of the stratigraphic sequence, using a 25m-radius ellipse in the main plane of anisotropy, which is dipping 30º (with an azimuth N30ºW). The search distance in the plane normal to dip was 18.75m. In order to simulate a node, a minimum of two composites were required, with a maximum of 10 composites (hard data) and 10 previously simulated nodes. Also, an octant condition was imposed, accepting a maximum of 3 conditioning data per octant (composites plus nodes).

5 RESULTS

A total of 20 lithology simulations were obtained. These simulations were analyzed statistically and graphically, and compared to the lithology information of the conditioning dataset (5m composites). Three were discarded because they did not reproduce adequately the proportions of each lithology, or their arrangement in space was not "realistic".

There are several alternatives that can be used to incorporate the simulated lithology model into a simulated grade model, and eventually a risk analysis study:

1. The most laborious option is to use each simulated lithology model as direct conditioning information on the simulated grade model, much like a traditional, deterministic lithology model is used to define domains and constrain grade interpolation. This requires that a simulated lithology model be obtained and validated for each simulated grade model desired. Although this is the preferred alternative, in this case there are insufficient hardware resources to complete the study.
2. A second alternative is to obtain the most likely lithology using, for example, a majority rule to determine which of the three lithologies is more likely at any given point in space. Then, this "secondary" simulated model would be used to constrain all simulated grade models. This option could be appropriate if the model is built using local varying means (LVM).
3. A different alternative, chosen in this case, is to use the simulation to obtain a prior cumulative distribution function of grade at each node (Alabert, 1987). Figure 2 shows the simulation chosen the set of 17 that reproduced well the overall lithology proportions and spatial distributions. The process requires definition of the lithology-grade relationship, based on original drill hole data, and done on a local basis (by sectors or sub-zones). Table 1 shows an example for the "Lince-low" sub zone.

The probability functions derived are in fact a quantification of subjective statements often heard from the mine geologists, such as "volcanic breccias are more mineralized than andesites". Note from Table 1 how, for the grade indicators defined, the volcanic breccias show a higher probability of being mineralized: if considering the 1.0% TCu indicator for example, there is just over 24% probability that the grade will be higher if the lithology is a breccia, while the same probability is 7% for andesites and 5.6% for intrusives. Also, there is a 4.6% probability that the grade in volcanic breccias will be higher than 5% TCu (very high grade), but only a 0.5% probability if lithology is either andesites or intrusives.

6 APPLICATIONS OF THE SIMULATED GEOLOGY MODEL

The uncertainty model will be more realistic if all possible sources of uncertainty are accounted for. The geology model is typically of the most significant sources of errors, in a resource estimate. Therefore, incorporating the simulated geology into the uncertainty model will result in a better assessment of possible grade ranges.

A natural follow-up to the uncertainty model is to transform it into a risk model, usually through the application of a Transfer Function (Journel, 1989). In a mining setting, the Transfer Function is a pit optimization and scheduling algorithm, or, more generally, a mine plan.

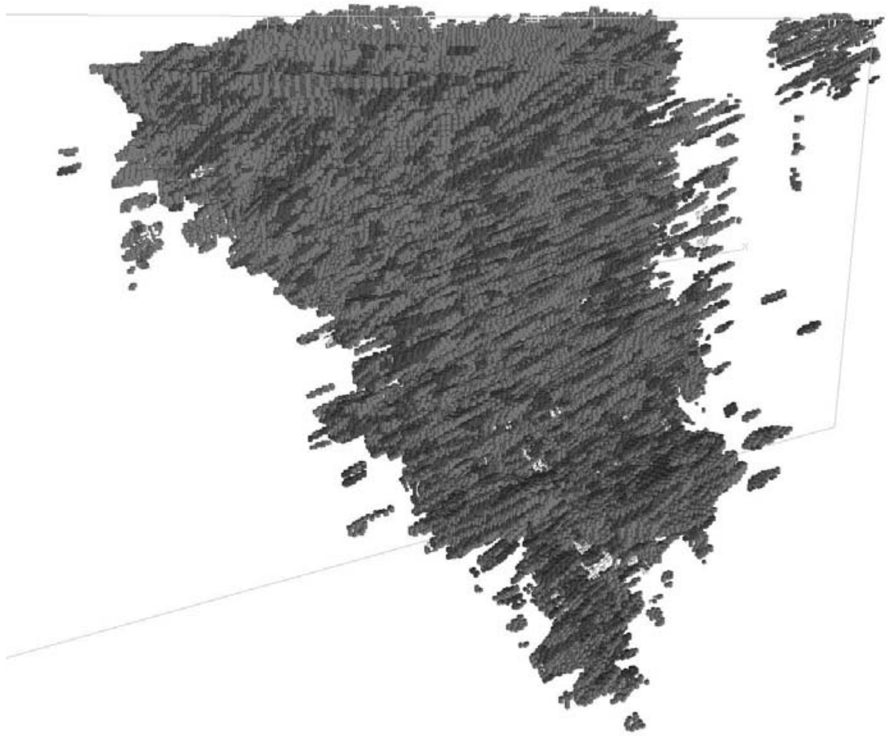

Figure 2: Perspective View of the Final Output Lithology Simulation. Note the gently dipping pseudo-stratigraphic sequence. Dark Grey: Volcanic Breccias; Light Grey: Andesites; Medium Grey: Intrusives and Tectonic Breccias.

The geology simulation model described above was used to provide a series of simulated models, regularized to match the exact same geometry and block size of the resource block model, representing possible grade ranges on a block-by-block basis. This uncertainty model was used to assess possible biases of the resource model.

Additionally, the simulated geology and grade models were used to analyze the risk of achieving the expected tonnage and grade for certain volumes of the deposit, and according to an existing mine plan and schedule.

7 SIMULATION MODEL VALIDATION

Several methods were used to validate the lithology simulations described. These included statistical comparisons of the resulting global (deposit-wide) and local (by sub-zones) proportions of each lithology against the existing lithology proportions in the database, and also comparing with the proportions resulting from the deterministic model.

There is always a variance (but hopefully not a bias) between the information as provided by the drill hole data and the geologic model constructed, deterministic or stochastic. This bias can be generated by clustering, but also by the process of interpreting and assigning spatial influence to each composite. For a geologically complicated deposit such as Michilla, a Modeling Error Rate (MER[1]) of 10% or less is acceptable. This level of "acceptable" variance also provides a guideline for validating the simulated model. However, because the simulation always honor the conditioning data point, this comparison has to be done on a larger volume, regularizing the original cells to a block (typically the same size as the block of the resource model), and assigning its lithology based on a majority rule, for example.

Table I: Soft indicators used, sub zone Lince-Low. The probabilities shown are taken from the cumulative distribution function of the composites for each lithology within the sub zone.

Grade Indicators ➔	***0.2% TCu***	***0.5% TCu***	***0.8% TCu***	***1.0% TCu***	***1.2% TCu***	***1.5% TCu***	***2.0% TCu***	***3.0% TCu***	***5.0% TCu***	***10.0% TCu***
Volcanic Breccias	0.525	0.675	0.73	0.758	0.782	0.811	0.849	0.902	0.954	0.992
Andesites	0.605	0.837	0.922	0.930	0.946	0.964	0.979	0.992	0.995	0.998
Intrusives / Tectonic Breccias	0.735	0.890	0.932	0.944	0.961	0.973	0.982	0.991	0.995	0.999

[1] The "modelling error rate" is defined as the percentage of the original data that are misclassified in the lithology model, that is, the composite is located within a modelled lithology that is different from the one originally logged by the geologist.

Also, it is important to visually compare the interpreted (deterministic) lithology units on cross sections and plan views, with the various simulated images. This comparison provides a subjective assessment of the reasonableness of the simulated models, and, if needed, is also an important component on convincing the local geologists as to the merits and characteristics of the simulated model. In particular, the characteristics of the simulated lithology should be analyzed in problematic areas, where specific concerns may need to be addressed.

Finally, the lithology model should be validated against actual production information, such as blast hole and underground mapping. If this information has not been used to obtain the simulation itself, it can be used to obtain the error rates (misclassification) of lithology codes on a block-size level compared to observed (mapped) data. If, as is commonly the case, production data is to be used as conditioning data in the simulation model itself, then it is advisable to run a smaller lithology simulation model on a sub-zone where production data is available, using only exploration drill hole data. This area can be used to calibrate the simulation model under the assumption that, if acceptable error rates are achieved in the calibration area, then the simulation model should behave similarly well elsewhere within the deposit.

8 CONCLUSIONS

This case study illustrates the application of indicator variables to the simulation of geological properties, or, more generally, categorical variables.

Some of the key aspects that motivate the simulation of lithology at Minera Michilla include the degree of control that different lithologies have on copper grade distribution, and the inherent difficulties in modeling deterministically the pseudo-stratigraphic sequence given the drill hole information available.

The Sequential Indicator method has been the most common approach used to date in a mining setting mostly because it is reasonably simple and easily understood by practitioners. Additionally, it provides several alternatives to input the geologic uncertainty model into the grade uncertainty model, principally through the indicator formalism.

The use of soft indicators at Michilla allowed representing the lithology control on copper grades on a local basis (by sub-zones), ensuring that the posterior grade uncertainty model incorporates the uncertainty related to the deterministic lithology modeling.

Indicator-based methodologies allow also incorporating uncertainty related to the sampling and assaying of the original exploration drill hole data, which is another common source of uncertainty. Further, if a complete “Chain-of-mining” study is done, simulated blast holes could be use to assess levels of planned and unplanned mining dilution and ore loss.

9 REFERENCES

- X Alabert, F.G., 1987, *Stochastic Imaging of Spatial Distributions using Hard and Soft Information*, MSc. Thesis, Stanford University, Stanford, California, USA.
- X Deutsch, C.V., and Journel A.G., 1992, *GSLIB: Geostatistical Software Library and User's Guide*, Oxford University Press, New York, 340 p.
- X Ferraris, F., and Di Biase, 1978, *Hoja Antofagasta*, Carta Geológica de Chile, Instituto de Investigaciones Geológicas, Santiago, Chile.
- X GeoSystems International, July 1999, *Informe Simulaciones Condicionales, Lince-Estefanía*, Cía. Minera Michilla, S.A, 105p. Unpublished Internal Report.
- X GeoSystems International, May 1999, *Informe Actualización Cálculo de Recursos, Lince-Estefanía*, Cía. Minera Michilla, S.A., 85p. Unpublished Internal Report.
- X Goovaerts, P., 1997, *Geostatistics for Natural Resources Evaluation*, Oxford University Press, 483p.
- X Journel, A.G., 1989, *Fundamentals of Geostatistics in Five Lessons*, Volume 8, Short Course in Geology, American Geophysical Union, Washington D.C.
- X Rossi, M.E**.**, 2003, *Practical Aspects of Large-Scale Conditional Simulations*, Proceedings of the 31st International Symposium on Applications of Computers to the Mineral Industry (APCOM), Cape Town, South Africa.