Multivariate Geostatistical Simulation at Red Dog Mine, Alaska, USA

Oy Leuangthong (oy@ualberta.ca)

Department of Civil and Environmental Engineering, University of Alberta

Abstract

Red Dog mine is the world's largest Zn producer. The objective of the case study is to characterized seven different minerals, Zn, Pb, Fe, Ba, sPb, Ag and TOC, within eight different rock types. Geostatistical models were constructed for each variable within the eight rock types, and subsequently assembled to give 40 realizations for six 25ft benches. The stepwise conditional transformation was used to account for the complex multivariate relations. In addition to reproducing the input data, histogram and variogram, the resulting models also respect multivariate relations locally and globally. This paper documents the methodology to construct these models.

A thorough validation procedure was implemented. Comparisons with blasthole data show that the conditional simulations and the long-term models agree in expected value. The main advantages of the conditional simulations are (1) the explicit accounting of multivariate relations between the data in model construction, and (2) the ability to quantify variability and uncertainty at any scale.

A small exercise on the value of simulation is also examined. Results show that the simulation approach showed a 3% increase in profit, relative to the conventional estimation approach.

1   Introduction

The Red Dog mine is located 90 miles north of Kotzebue, Alaska, USA. It is owned and operated by Teck Cominco Limited. The deposit consists of sulphide ore zones in sedimentary exhalative (sedex) deposits, and is characterized by the presence of multiple metals and multiple ore types. The mine assays for as many as ten variables; the four primary ones being Zn, Pb, Fe and Ba.

A key issue is the variability within the deposit and the effect of this variability on Zn recovery. Recovery is adversely affected by the presence of high barite and other deleterious minerals and ore textures. The existing long term resource model was constructed by independently kriging the four main variables. Improved multivariate modeling of the different elements and ore types should improve the reliability of the long-term resource model and therefore the prediction of Zn recovery.

Direction	Minimum (ft)	Maximum (ft)	Number of Cells	Size of Cells (ft)
Easting	585000	589500	360	12.5
Northing	141500	146000	360	12.5
Elevation	800	950	12	12.5

Table 1: Red Dog model coordinate limits.

2 Background

The Red Dog Main Pit consists of three geological plates: Upper, Median and Lower. There are a total of 31 geology codes, of which only eight will be modeled. These eight geological rock types correspond to four different ore type units in two separate plates. These were chosen because they correspond to a volume that includes both recently mined material and material that will be mined in the near future.

The existing grade models were independently kriged at a 25ft × 25ft × 25ft resolution. For this case study, the geostatistical models will be simulated at 12.5ft × 12.5ft × 12.5ft resolution, and will later be upscaled to $(25\text{ft})^3$ for comparison purposes. There are some good reasons to model at a finer scale than will be required later. Firstly, the 12.5ft composite data are a good compromise between retaining some of the variability of the smaller drillhole sample data and the faster simulation of larger, and hence fewer blocks. Secondly, the simulation is essentially a "point"-scale simulation; current implementations do not explicitly account for volume-variance relations. Thus, simulating at a finer resolution and then averaging to larger blocks will show the variability of the block grades more accurately.

Six benches were modeled to allow for model reconciliation with blast hole samples. Table 1 lists the coordinate limits of the conditional simulation model. These limits essentially cover the entire areal extents of the Main pit and the vertical extents of the six benches of interest. This model consisted of a total of 1,555,200 cells.

The simulations were constructed on a by rock type basis, and all figures shown correspond to one particular rock type. Once all rock types were simulated, the realizations were merged. All global comparisons consist of all rock types taken together.

3 Available Data

Three types of data were made available by Teck Cominco: drillhole data, composited drillhole data and blasthole data. Multivariate geostatistical modeling considered the 12.5ft composites, while the blasthole data were used to test the predictive ability of the resulting models.

There were a total of 9847 12.5ft composites available for the eight rock types of interest. The term drill hole (DH) refers to the 12.5ft composites. DH data are at a nominal 100ft × 100ft spacing.

For these same rock types, there were 58566 blast hole (BH) data available for model validation. BH data are more closely spaced than DH data at 10ft × 12ft spacing with 25ft vertical extent (the vertical span one bench). Figure 1 shows the projection of the available
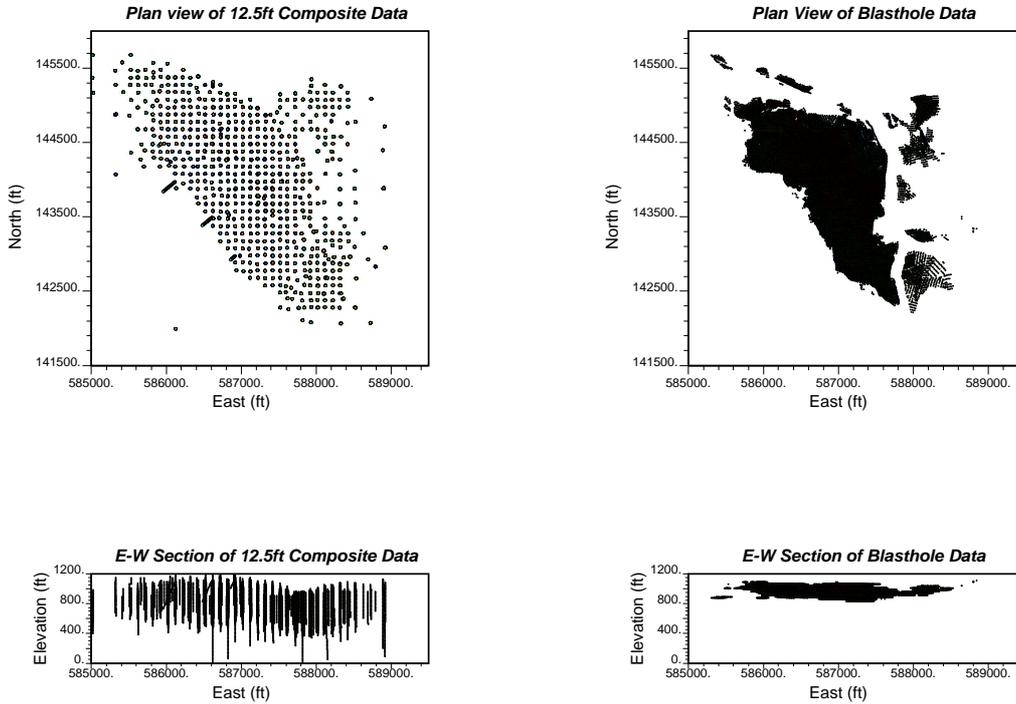
Figure 1: All available data projected onto a horizontal plan (plan view (top)) and onto a vertical plane (E-W sectional view (bottom)): 12.5ft composites (left) and blasthole data (right).

data onto horizontal and vertical plane for both data types separately. Note that in the plan and section for BH data, the data density is high, and the distance between the BH samples is very small relative to the size of the field.

A geology model at $(25ft)^3$ resolution was also available. For consistency with the simulation models, the $(25ft)^3$ geology model was reformatted into a $(12.5ft)^3$ model.

4 Multivariate Geostatistical Simulation

Conditional simulations were performed for seven variables: Zn, Pb, Fe, Ba, sPb (soluble Pb), Ag, and TOC (total organic content). These seven variables were modeled for each rock type, using Gaussian simulation with stepwise conditionally transformed variables. The main steps of the simulation are:

1. Transform data in a stepwise conditional [7] manner to obtain independent Gaussian variables.

2. Calculate and model the directional variograms for each of the transformed variables within each rock type.

3. Simulate transformed variables via sequential Gaussian simulation [4].

Transform No.	Variable	Conditioning Variable(s)
1	Zn	
2	Pb	Zn
3	Fe	Zn, Pb
4	Ba	Zn, Fe
5	sPb	Zn, Pb
6	Ag	Zn, Pb
7	TOC	Zn, Fe

Table 2: Transformation ordering for stepwise conditional transformation.

4. Back transform simulated values in a stepwise conditional manner to return values to original units.

Once all variables within all rock types were modeled, all block models were merged to form multiple realizations of the study area for uncertainty assessment and post-processing. All simulation related tasks were performed using GSLIB [3] and related GSLIB-compatible tools.

The need to consider seven variables simultaneously for any one rock type poses a problem in practice. The multivariate stepwise conditional transform would require 10^7 composites in order to have a minimum of 10 data per probability class. This is impractical. A *nested* application of the stepwise conditional transformation is proposed to overcome this problem. Accounting for a lower-dimensional multivariate distribution was considered. Inference of a trivariate distribution would require approximately 10^3 or 1000 data to define the conditional distributions with a minimum of 10 data. This is more reasonable given the number of composites available.

The transformation ordering for the stepwise conditional transform will affect the reproduction of the variogram from simulation. Thus, the most important variable or the most continuous variable should be chosen as the primary variable [7]. For Red Dog, Zn is the most important variable, and so all others will be conditioned to it. To account for the other six variables, the following sets (Table 2) of transformations were proposed:

The transformation order reflects the significance Teck Cominco staff attribute to each variable. Zn is considered to be the most important, and so all other variables are transformed conditional to Zn. In all cases, Fe or Pb act as secondary variables, and all remaining variables are then transformed conditional to either Zn and Pb or Zn and Fe.

Data Declustering. An important aspect of geostatistical simulation is to assemble representative distributions for each variable. Given the multivariate nature of this dataset and the intended application of a multivariate transformation technique, declustering must be consistent between all variables. This consistency involves respecting the multivariate relations and the manner in which simulation will account for them. Multivariate dependency between all seven variables is a direct consequence of the transformation order that will be imposed (see Table 2).

The representative distribution of Zn must be established by a declustering procedure
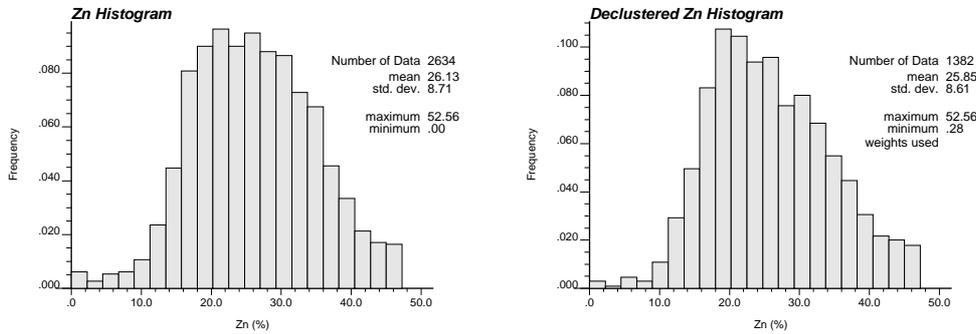
Figure 2: Comparison of equally weighted Zn distribution and representative Zn distribution.

using the data configuration and the volume of a particular rock type. For this purpose, kriging within a rock type was performed; the kriging weights given to each data were accumulated, and these weights were then used as the declustering weights. This approach not only respects the rock type being populated (much like nearest neighbour declustering), but it also respects the spatial variability of the data and hence their area of influence within this rock type.

Declustering of the secondary variables (say Pb) must respect the bivariate relations since these will be transformed conditional to Zn. The stepwise conditional transform considers only those secondary data where primary data is available (that is, at locations where there are both Zn and Pb data). As a result, the distribution that will be reproduced in the back transformation is the isotopically sampled values of Pb with Zn. For this reason, a bivariate calibration of the Pb distribution was performed using both the representative distribution of Zn and the relationship between Zn and Pb. For all tertiary variables, the same rationale was applicable, and the representative histograms for Fe through TOC were determined using the representative histograms for the two dependent variables plus the trivariate calibration data.

Figure 2 shows the comparison of the histogram of the Zn data within one rock type and the corresponding representative histogram of Zn grade. Note there are 2634 data for the equal weighted histogram on the right. As discussed above, not all of these are used to assemble the representative histogram of Zn; in fact, only about half the total data are used to decluster. Declustering with kriging weights was applied to assemble the representative Zn histogram.

Only minor changes were apparent. There was no change in the data values, they were only weighted differently. The mean and standard deviation have decreased slightly. The scatterplot of Pb given Zn was used to calibrate the marginal distribution of the Pb to give the representative Pb distributions (see Figure 3). Figure 4 shows a comparison between the equal weighted histogram of all Pb data (using 2634 Pb data) and the representative Pb histogram. Given the positive correlation between Pb and Zn and the slight decrease in the mean of Zn, it was expected that the representative histogram of Pb should have a slightly lower mean than the 7.02% reported on the equal weighted histogram; and it has indeed decreased to 6.91 %.
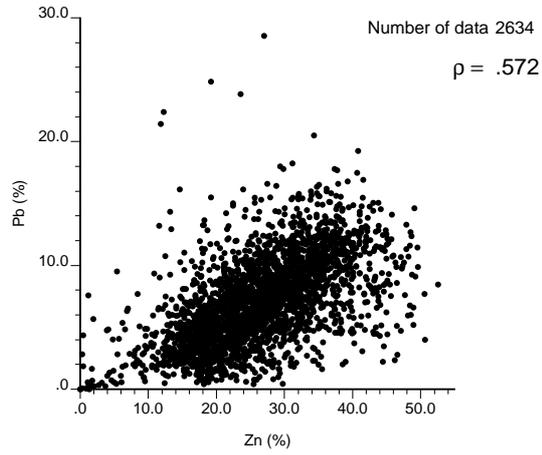
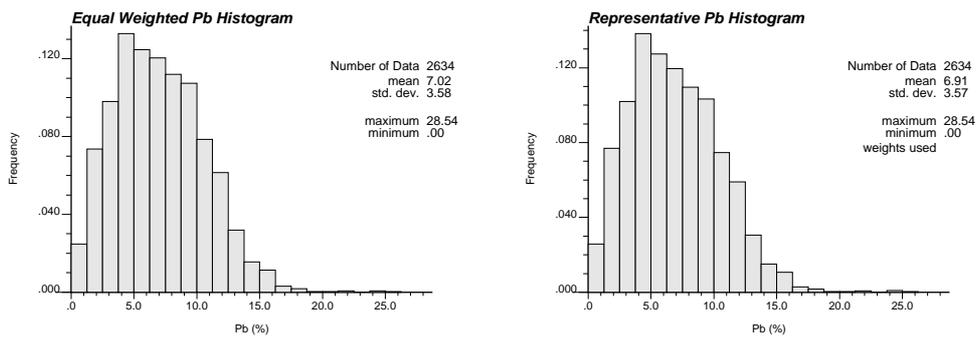Figure 3: Calibration crossplot of Pb given Zn.

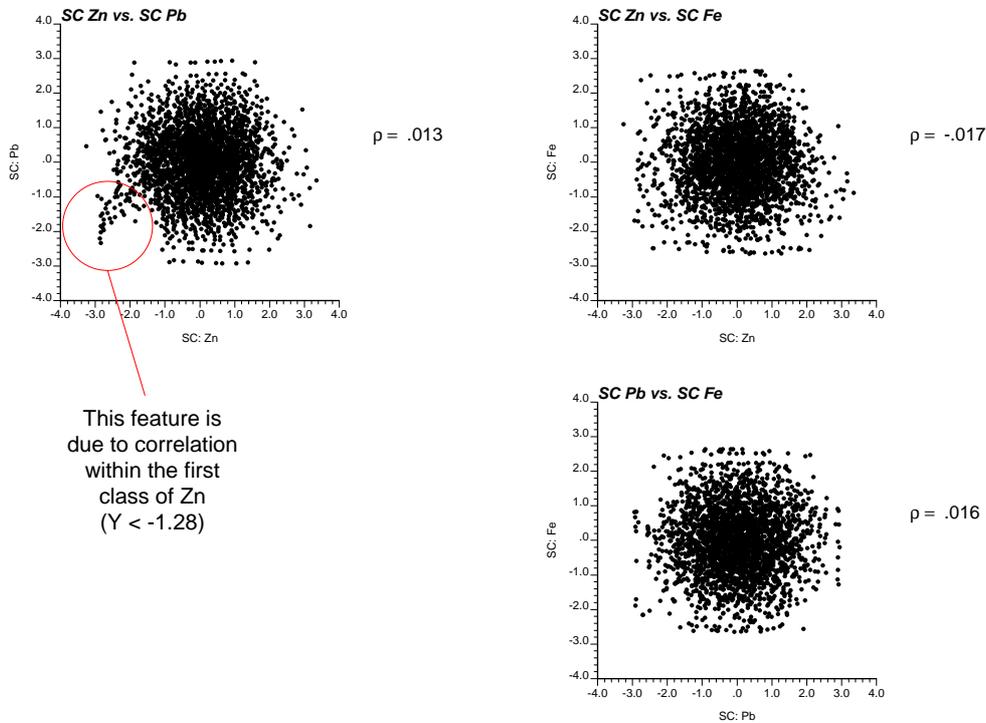Figure 4: Comparison of equally weighted Pb distribution and representative Pb distribution.

6

Figure 5: Crossplot between stepwise conditionally transformed variables for Zn, Pb and Fe. Zn was transformed first, then Pb was transformed conditional to Zn, and finally Fe was transformed conditional to both Zn and Pb.

Stepwise Conditional Transform of Red Dog Data. Figure 5 shows the scatterplots of the variables resulting from the first transform sequence of Zn, Pb and Fe (see Table 2).

The transformed variables are independent and multiGaussian, which translates to a circular shape in the crossplot. From Figure 5, the crossplot between the first two variables (Zn and Pb) appears approximately circular. Crossplots with the third variable (Fe, in this case) show some banding; however this is simply a numerical artefact of having many classes and consequently fewer data within each class [7].

Independence of the transformed variables means that each variable can be simulated independently [7].

Variogram Analysis. Spatial statistics can now be calculated and modeled for the transformed data.

Figure 6 shows an example of variogram maps, experimental variograms and the variogram model for Zn. Both the variogram model and the experimental variogram points are shown in the principal directions. Overall, the experimental variograms are fairly stable and the corresponding models fit the experimental points well.

Variogram maps are a common tool used to pick directions of continuity. The maps are

7

in radial coordinates; location $\mathbf{h} = 0$ is at the centre of the map. Basically, the maps are used to visualize large scale continuity directions and distances. These directions and distances are then further refined during the calculation and modeling of experimental variograms.

Simulation. Sequential Gaussian simulation (SGS) was independently performed for each of the seven transformed variables on a by rock type basis: Zn, Pb, Fe, Ba, sPb, Ag and TOC. A total of 40 realizations were generated for each variable within each rock type. For greater computational efficiency, only those blocks belonging to the specific rock type were simulated (as controlled by the geology model).

Back Transformation. The simulation results must be back transformed to the original units of the data. Similar to the forward transformation that relied on conditioning one variable to another, the back transformation for each simulated realization must be performed in a conditional fashion. For example, the back transform of Fe will be conditional to the simulated values for Zn and Pb. The key to back transformation is to maintain consistency with the forward transformation.

Validation of Simulation Models

A number of basic checks must be performed prior to using these models for decision making.

An important validation is the reproduction of the input data and the variogram. Statistical fluctuations are inherent in stochastic simulation; however, these fluctuations should be reasonable and unbiased. Simulation produces simulated values that are approximately standard normal in *expected* value. For any one realization, minor fluctuations from a zero mean and unit variance are expected; however, when these values are back transformed to original units a slight shift of the mean in normal space may translate to a more significant shift of the mean in original units. Similarly, the combined fluctuation of the mean and variance in normal space may translate to more noticeable shifts in original space. This is particularly true for skewed distributions, which is the case for most variables in the Red Dog data.

Deviations from the limit standard normal distribution could be due to a number of factors. Firstly, the algorithms employed are based on an assumption of stationarity. Nonstationary data can lead to shifts in the mean and/or variance of simulated values in normal space. Secondly, Gaussian simulation techniques assume the data are multiGaussian in a spatial context. There are no techniques to ensure multiGaussianity in the spatial domain. To mitigate the effects of fluctuations in normal space and its translation to original space of the data, a standard transform is applied to the simulated values to ensure reproduction of the histogram and its corresponding summary statistics [5]:

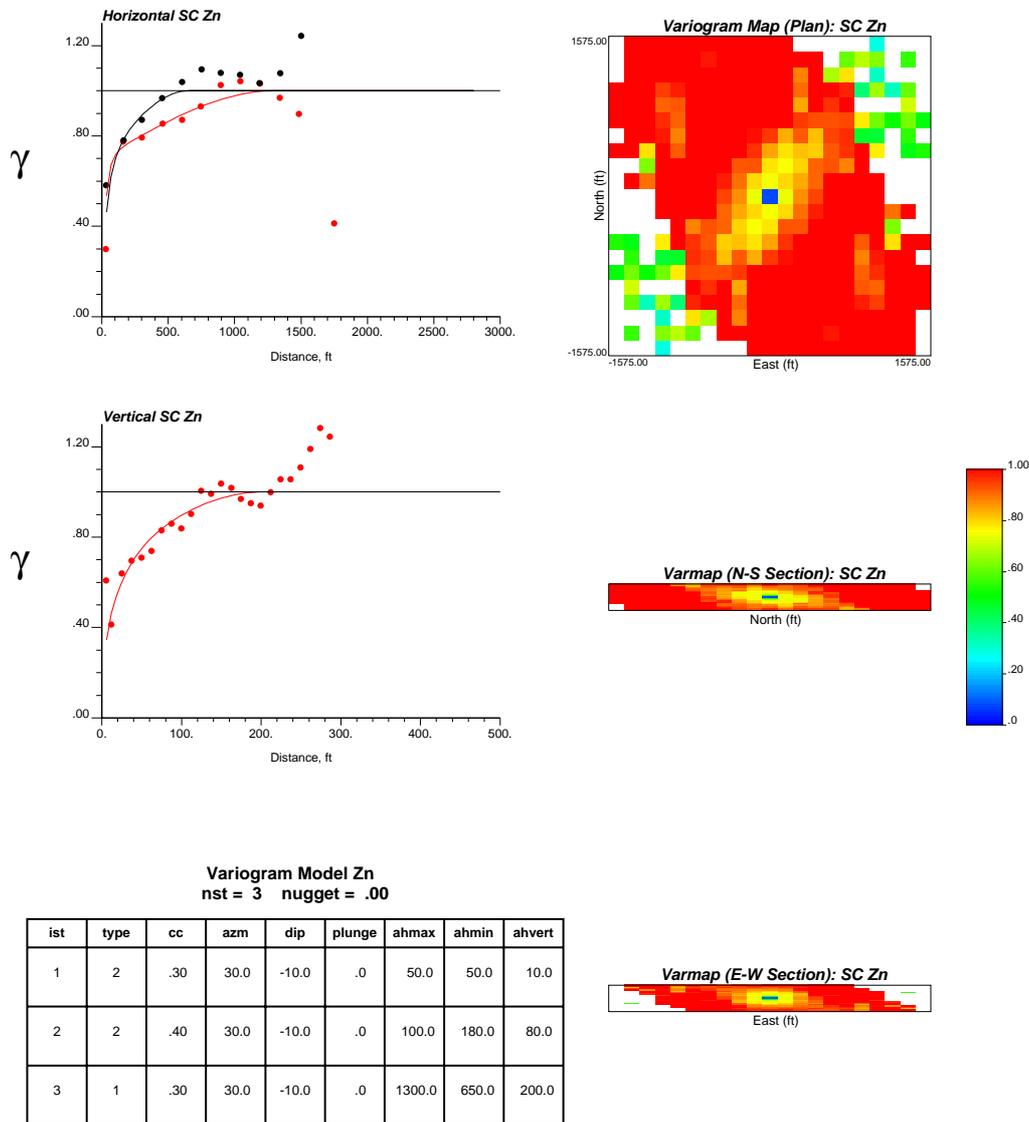$$z_2(\mathbf{u}) = z_0(\mathbf{u}) + \lambda(\mathbf{u})[z_1(\mathbf{u}) - z_0(\mathbf{u})] \tag{1}$$

8

Horizontal SC Zn

Variogram Map (Plan): SC Zn

Vertical SC Zn

Varmap (N-S Section): SC Zn

Variogram Model Zn
nst = 3   nugget = .00

ist	type	cc	azm	dip	plunge	ahmax	ahmin	ahvert
1	2	.30	30.0	-10.0	.0	50.0	50.0	10.0
2	2	.40	30.0	-10.0	.0	100.0	180.0	80.0
3	1	.30	30.0	-10.0	.0	1300.0	650.0	200.0

Varmap (E-W Section): SC Zn

Figure 6: Variogram model results for Zn. Horizontal variograms (top left) show maximum continuity direction (N30E) in grey and minimum continuity direction in black. Vertical variogram (middle left) and corresponding variogram model parameters (bottom left). Variogram maps are shown on the right: plan view (top), N-S section looking west (middle), and E-W section looking north (bottom).

where

$$
\begin{aligned}
F_0(z) &= \text{cdf of the simulated values} \\
F_1(z) &= \text{target cdf} \\
z_0(\mathbf{u}) &= \text{set of originally simulated values, } \mathbf{u} \in A \\
z_1(\mathbf{u}) &= \text{corrected value based on quantile transformation alone} \\
&= F_1^{-1}(F_0(z_0(\mathbf{u}))) \\
z_2(\mathbf{u}) &= \text{corrected value based on quantile transform and kriging variance} \\
\lambda(\mathbf{u}) &= \text{correction factor defined as } [\sigma_K(\mathbf{u})/\sigma_{max}]^{\omega} \\
\sigma_K(\mathbf{u}) &= \text{kriging variance at location } \mathbf{u} \\
\sigma_{max}) &= \max\{\sigma_K(\mathbf{u}), \mathbf{u} \in A\} \\
\omega &= \text{correction level parameter, must be} > 0
\end{aligned}
$$

Use of the kriging variance in Equation 1 ensures that the values at data locations are reproduced. The transform is applied over individual realizations to ensure reproduction of the global histogram for each realization. Alternatively, "sets" of realizations can also be transformed and data would still be honoured; however, this does not guarantee that the global histogram per realization is reproduced. For this case study, the transform is applied over individual realizations.

Data Reproduction. The goal is to verify that the corresponding simulated values reproduce the assigned composite values. For each variable, a crossplot showing the DH values and their corresponding simulated values demonstrates whether input data were reproduced within the numerical precision of the storage and the transformation table being used. Figure 7 shows an example of the crossplot for Zn. Over the multiple rock types and multiple variables, the composite data were reproduced exactly in almost all cases. Deviations from the true value were a result of the numerical precision reported in the transformation table. In general, minor fluctuations at the high and low ends of the units were attributed to the number of quantiles reported in the transformation table. Overall, DH data were reproduced at their respective locations from the simulation models.

Histogram Reproduction

Another important check is the histogram of the simulated values after back transformation. These distributions should be similar to the representative histograms, with comparable statistics. Figure 8 shows an example of one such comparison for Zn. Overall, the histograms were reproduced within reasonable statistical fluctuations in the summary statistics by construction (see Equation 1).

Over the 40 realizations (or the ensemble), the summary statistics can be checked for reproduction. For each realization, there is a mean and a variance associated to the resulting global distribution. A histogram of the mean from all realizations will show if the mean is reproduced. Similarly, a histogram of the variance will show if the global variance is reproduced. Figure 9 shows these two histograms for Zn. The mean of the representative distribution was reproduced; differences in magnitude lie in the second decimal place. As well, the distribution of variances showed only minor differences in magnitude. This reproduction was a consequence of the transform applied (Equation 1).
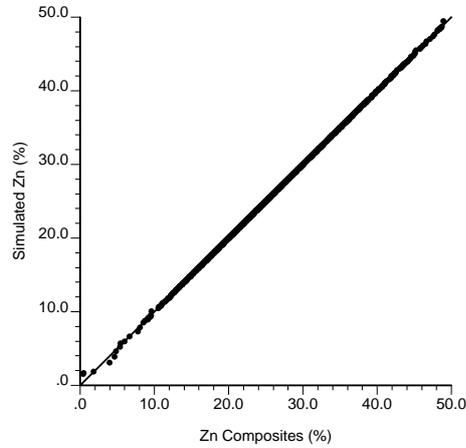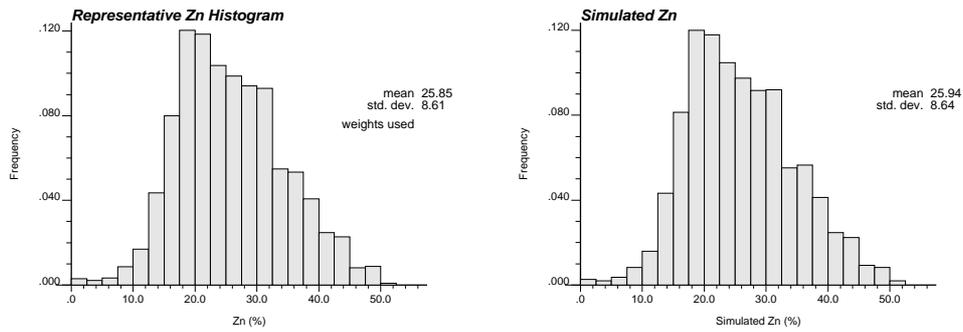
Figure 7: Composites data reproduction for Zn.

Figure 8: Histogram reproduction for Zn: representative Zn histogram (left) compared to simulated Zn histogram (right).
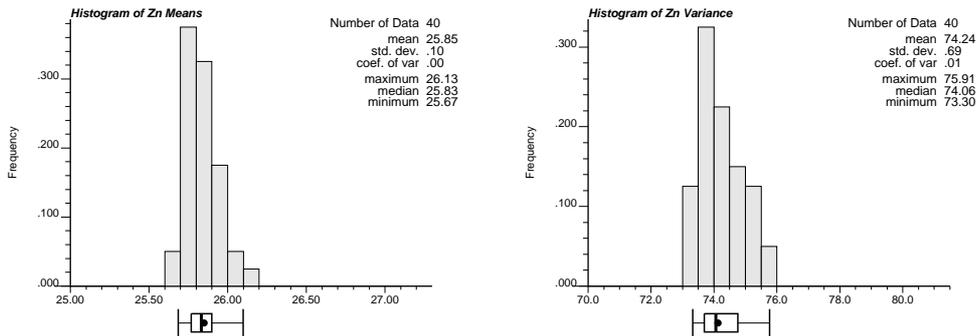

Figure 9: Reproduction of summary statistics for Zn: histogram of means (left) and variances (right) from multiple simulated realizations. Box plots on x-axis shows the 95% probability interval (outside lines), 50% probability interval (box) and the median (vertical bold line inside box). The dot indicates the mean value of the summary statistic from the declustered distributions.

Variogram Reproduction. After verification that the first order statistics were satisfactorily reproduced, the next check involved the variogram. It is important to note that this check was performed in normal or transformed space (prior to back transformation), since only the Gaussian variogram will be reproduced. Figure 10 shows the results of performing this check for Zn. The variogram directions calculated from the simulated models correspond to the standard North-South (N-S), East-West (E-W), and vertical directions. These variograms are shown as grey dashed lines. The variogram models were calculated in these same three directions (although principal directions may differ) and are shown as black solid lines. The experimental variogram points calculated from the 12.5ft composite data are also shown, and correspond to the black dots on these figures.

Reproduction of Multivariate Features. The multivariate relations are important and must be checked. Teck Cominco was also interested to see how their existing models performed for the same type of check.

In order to allow for direct comparisons between a simulation and the existing long term model, the simulations must first be upscaled to the same volume as the existing model. This required upscaling from $(12.5\text{ft})^3$ models to $(25\text{ft})^3$ models. This upscaling is discussed in more detail in the section on Validation.

Figure 11 shows a comparison of the crossplot reproduction from simulation to those crossplots from the 25ft composites and the existing long term model. In general, the simulated realizations reproduce the trivariate relations with comparable variability to the 25ft composites; the corresponding plots from the existing long term model shows similar bivariate relations but with noticeably reduced variability.

Recall that the relations between Zn and Ba was the most important for Zn recovery. Comparing the Zn-Ba crossplot from all three sources (composites, simulation, and long term model) shows the existing model reproduced neither the bivariate relations nor the inherent variability of the data. This result and its potential impact on production supports the use of multivariate geostatistics in model construction.

General Comments on Conditional Simulation Models

Once all simulated models were generated and validated on a by rock type basis, a single realization for each variable was obtained by merging the simulated properties from each rock type. Figure 12 shows a few of the simulated realizations for Zn at the 12.5ft grid resolution.

The modeling methodology implemented in this project was quite complex. Conventional approaches are sufficient for straightforward problems; however, for the complexity of the Red Dog data, these common approaches are inadequate. The availability of multiple metal grades within multiple rock types warrants some consideration of the relationship between these grades and how these relationships change from one rock type to the next. The approach documented in this section was designed to explicitly address this key issue. Consequently, the resulting models not only reproduce the univariate data and its spatial variability, but taken together, they also honour the multivariate relations between the different metals/minerals within the different rock types.
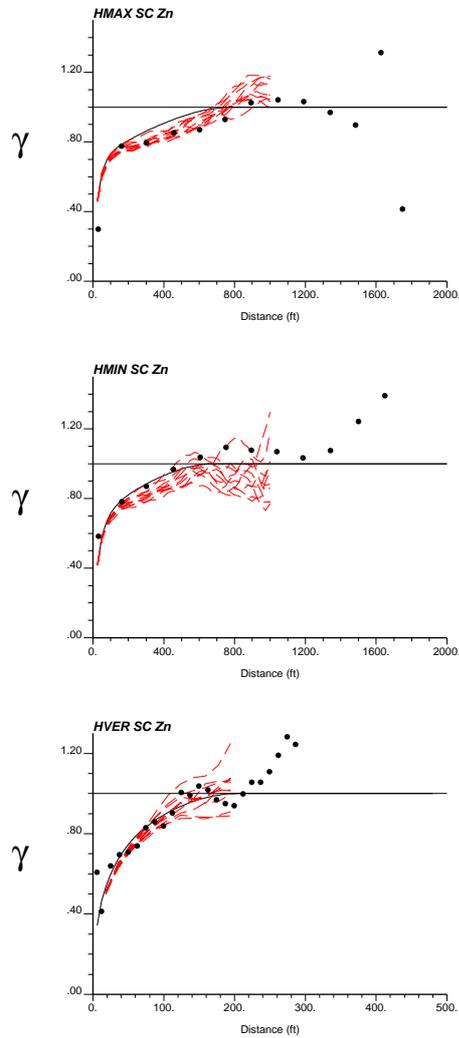
Variogram Model Zn
nst = 3   nugget = .00

ist	type	cc	azm	dip	plunge	ahmax	ahmin	ahvert
1	2	.30	30.0	-10.0	.0	50.0	50.0	10.0
2	2	.40	30.0	-10.0	.0	100.0	180.0	80.0
3	1	.30	30.0	-10.0	.0	1300.0	650.0	200.0

Figure 10: Variogram reproduction for Zn: NS direction (top), EW direction (second), vertical direction (third), and table listing variogram model. Black solid line represents the variogram model, and dashed lines represent the variogram of the simulated models.

13
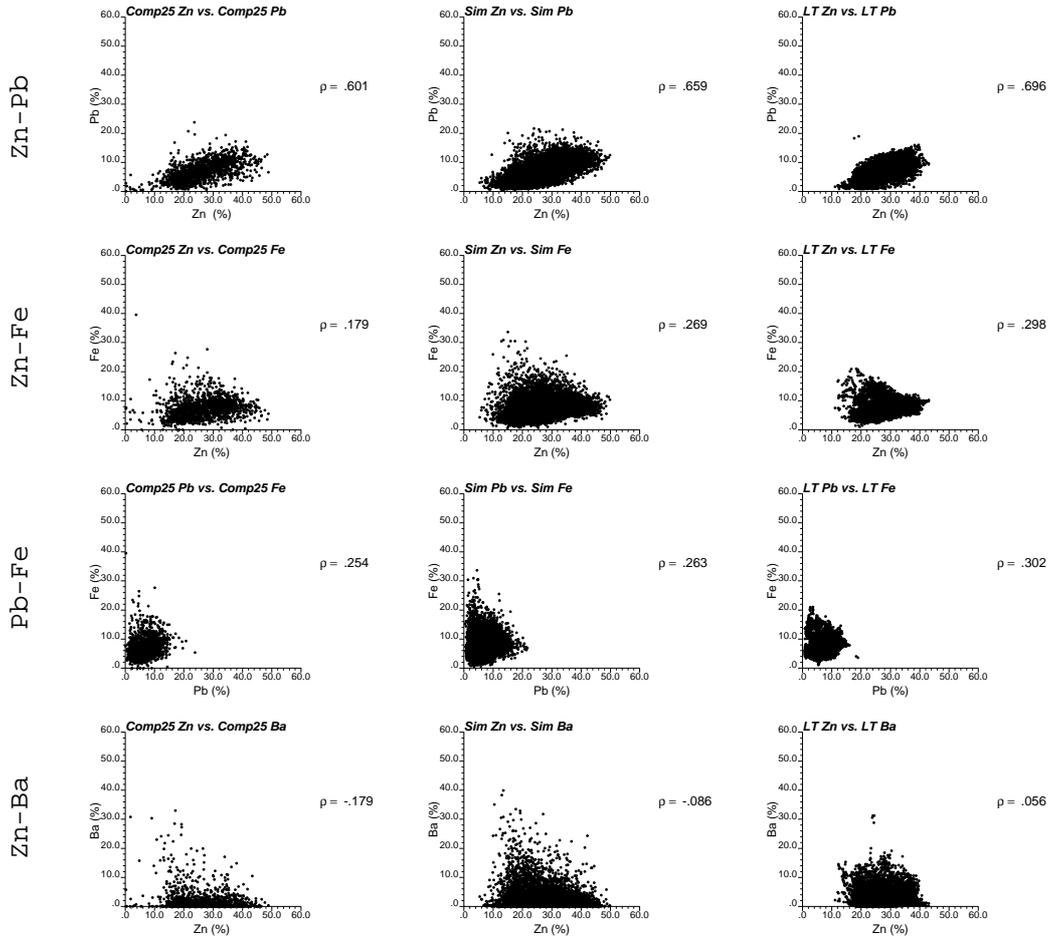
Figure 11: Comparison of multivariate features reproduction for Zn-Pb (top row), Zn-Fe (second row), Pb-Fe (third row), and Zn-Ba (bottom row). Crossplots using the 25ft composites are shown on the left column, from the upscaled simulations are shown in the middle column and those from the existing long term model are shown in the right column.
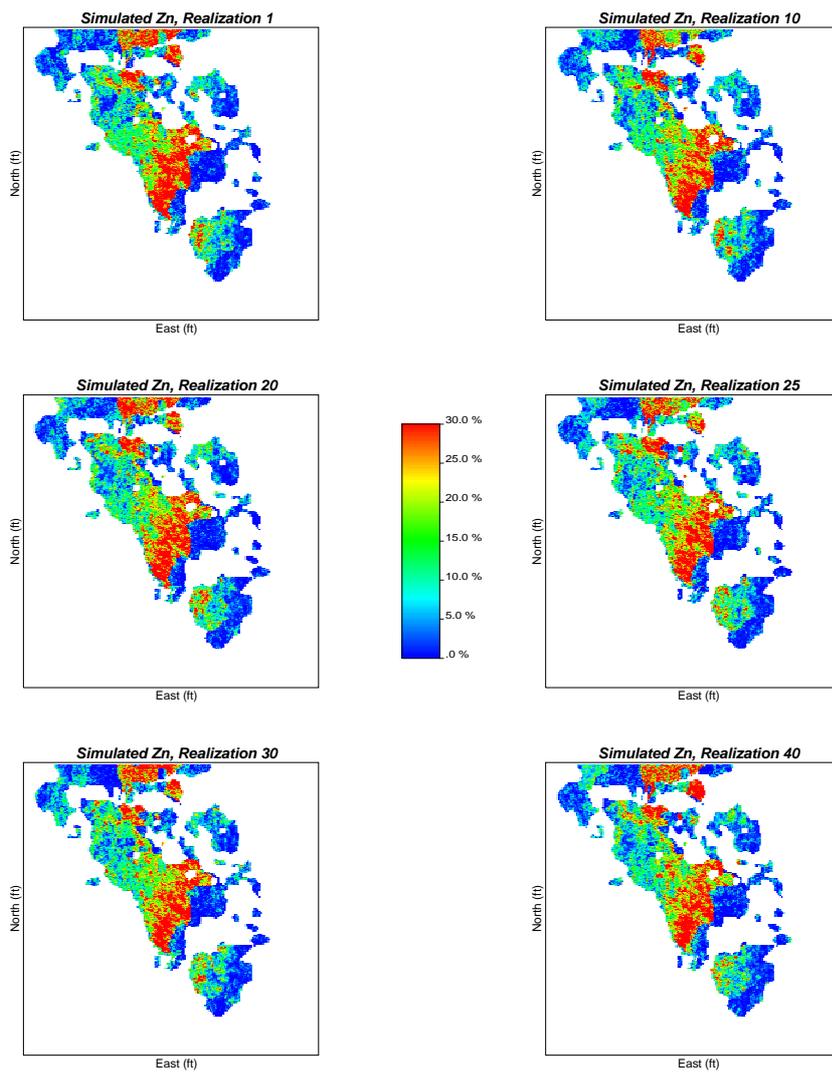
14

Figure 12: Simulated realizations of Zn at 12.5ft grid resolution. This section spans elevations 862.5 to 875ft.

5 Validation with Additional Data

Blasthole (BH) data was intentionally excluded from the input data used for model construction. The idea was to assess the predictive ability of the conditional simulation models using the BH data. In particular, four variables will be compared: Zn, Pb, Fe and Ba.

Comparing BH and DH Data. The BH data were paired with the closest DH data within a tolerance of 50ft. A total of 9846 DH composites were checked against 58560 BH data. Figure 13 shows the results of pairing up the BH with the DH data. Significant banding was expected as a result of pairing the DH data with multiple BH data.

This nearest-neighbour pairing of the BH and DH data presented the lower bound on the expected correlation. Simulation depends on kriging which accounts for data redundancy, closeness, and the surrounding data values. Having considered the spatial correlations in addition to the data values, the simulated values should be better correlated to BH data. Of course, the variogram used in simulation will have an impact on how much better the expected correlation should be. For instance, a higher nugget effect would reduce the correlation between the simulated value and the BH data. Alternatively, a variogram model with a long range and low nugget effect should increase the correlation between the simulated value and the BH data.

Comparing BH and Conditional Simulation Models. Prior to any type of comparative studies, the $(12.5\text{ft})^3$ models must first be upscaled to a resolution comparable to the BH data, which are at 10 x 12ft areal spacing with a length of 25ft. The models were upscaled to 12.5 x 12.5 x 25ft for consistency with BH data. The block averaging method was a weighted average based on specific gravity equations, provided by Teck Cominco. Figure 14 shows a few of the realizations of Zn at this resolution. As a result of upscaling in the vertical direction, these realizations appear slightly smoother than those in Figure 12.

Given that there were 40 realizations for the conditional simulation models for each variable, there was the issue of which realization should be compared against the BH data. Rather than choose any arbitrary realization (since all are equally likely to be chosen), the E-type estimate was compared. The E-type estimate refers to the expected value at each location calculated based on the local distribution constructed using the 40 realizations. This is similar to a kriged model, since it is a model of expected values. The effect of the variability inherent in any one realization will be mitigated by choosing the E-type estimate.

Figure 15 shows the comparison between the BH data and the E-type estimate. The crossplots show fairly strong positive correlations between the model and the BH data, ranging from 0.62 to 0.86. Two distinct populations were apparent from the Ba crossplot, which represented the two rock types. A comparison with Figure 13 shows that for all four variables, the correlation between the BH data and the simulated values were higher than the lower bound represented by the pairing of BH to DH data.

Determining Expected Correlations. It is possible to determine the expected correlation coefficient between the simulation models and the BH data, given the known DH and BH spacing, and the variogram model used to construct the simulations. Figure 16 shows a schematic illustration of the relationship between the BH data spacing, variogram model
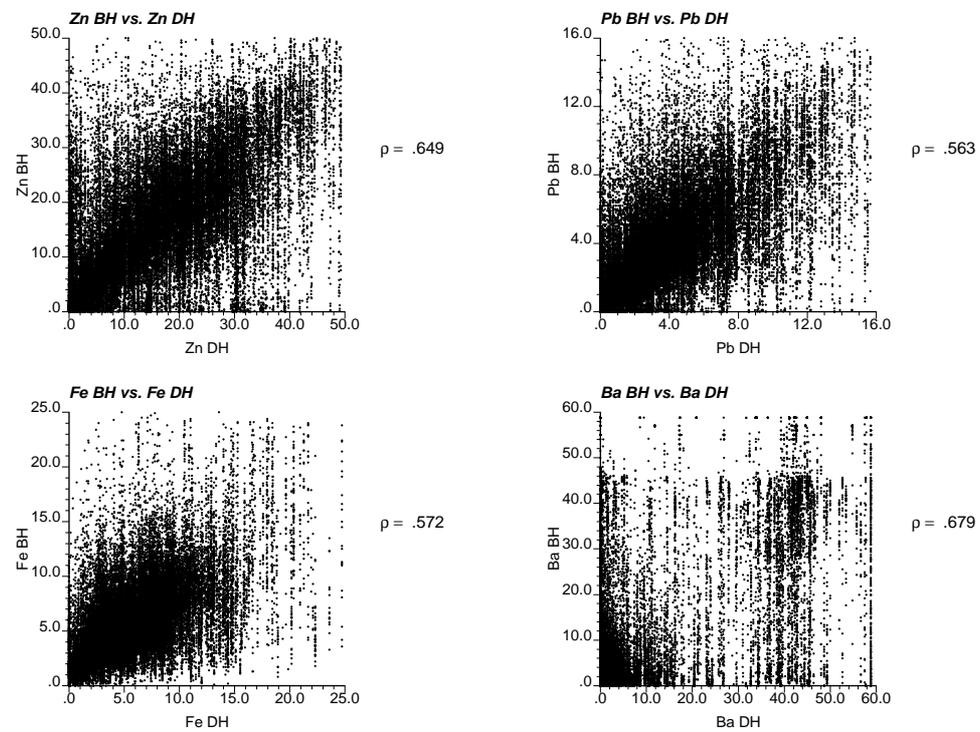
Figure 13: Crossplot of BH data against nearest neighbour DH data for Zn, Pb, Fe and Ba over all eight rock types. Significant banding is a result of pairing DH data to multiple BH data.
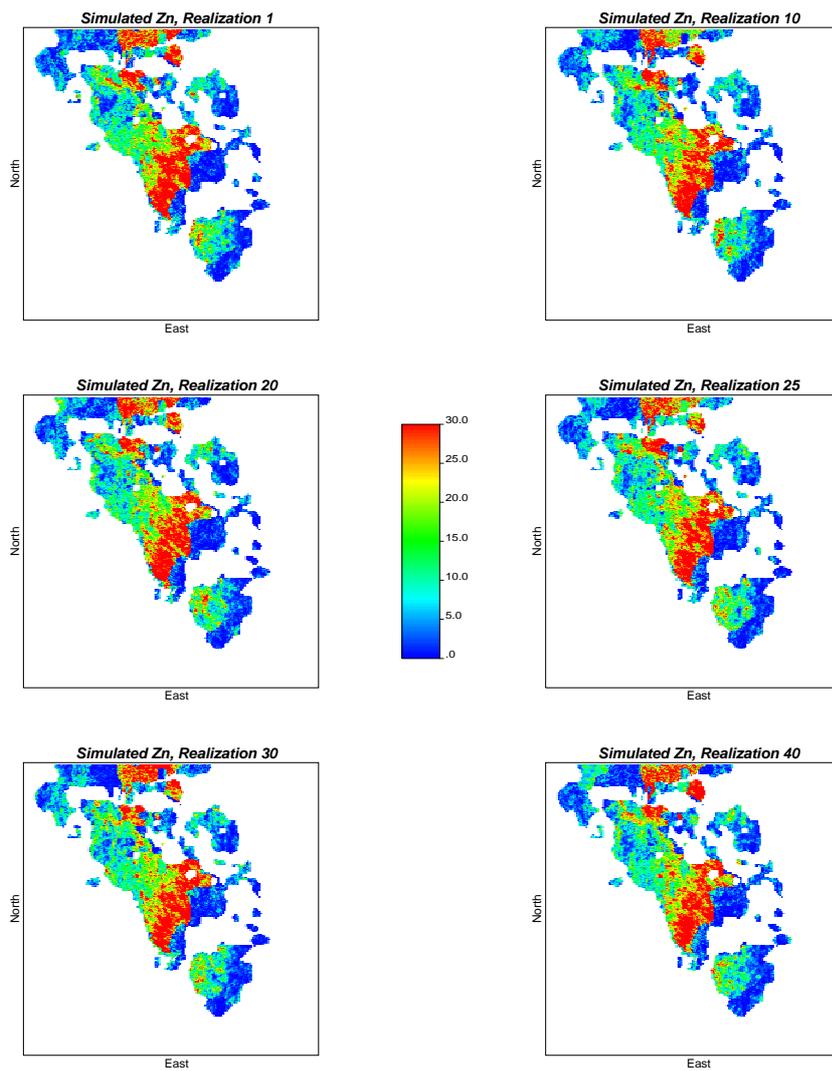
Figure 14: Simulated realizations of Zn at 12.5 x 12.5 x 25ft grid resolution, consistent with BH data. The section shown spans elevations 850 to 875ft.
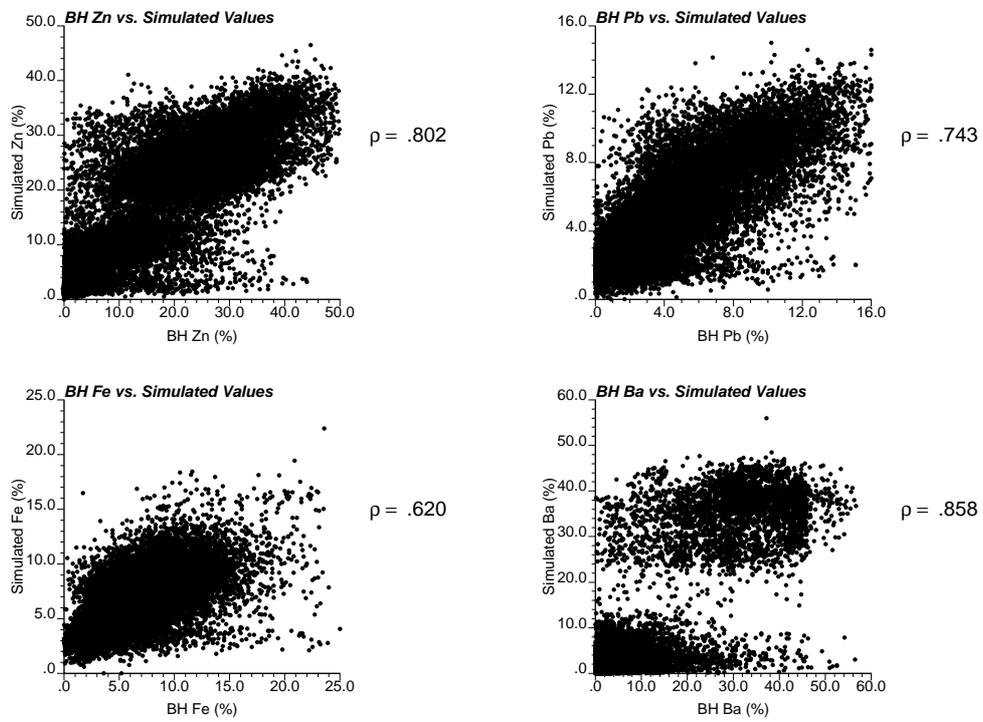
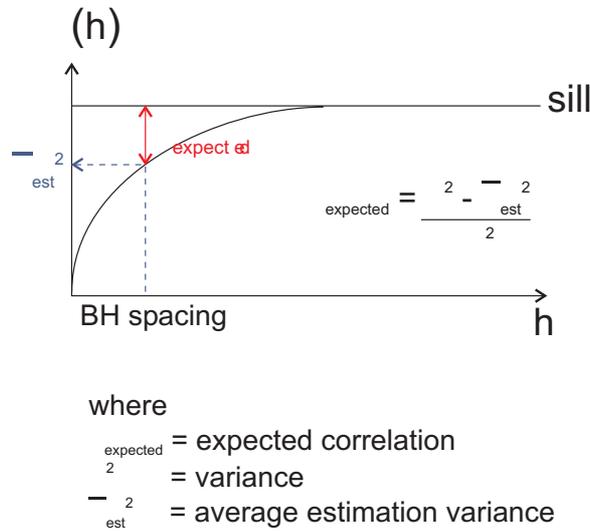Figure 15: Crossplots of BH data and E-type estimate from Conditional Simulation Models.

Figure 16: Schematic illustration of relation between expected correlation and estimation variance at BH spacing. This is calculated for multiple BH locations, and then averaged to obtain the average estimation variance.

and the expected correlation. Use of this approach assumed that the BH data and the DH data have the same variogram model, that is, the spatial variability of the BH data was the same as that of the DH data. As well, the BH data were assumed to have similar statistical properties as the DH data.

From Figure 16, the key to determining the expected correlation was to first determine the average value of the estimation variance at the BH data spacing. Recall that the DH data are at 100 x 100ft spacing while the BH data are at 10 x 12ft spacing. As a result, BH data are interspersed between the DH data, and so the estimation variance must be calculated at each BH location and then averaged within the 100 x 100ft block. For the purposes of approximating this value, the model's 12.5 x 12.5ft horizontal spacing was sufficient for substituting the actual 10 x 12ft BH spacing.

Since the variogram models were constructed for the normal scores, the variance is 1.0. The expected correlation is given by:

$$\rho_{expected} = 1.0 - \overline{\sigma^2_{est}}$$

Figure 17 shows the configuration of the DH data and the block model spacing. For simplicity, the BH data were assumed to be centered about a DH sample on an approximate 8 x 8 grid with 12.5 x 12.5ft blocks. Kriging can be performed on this data configuration with an appropriate variogram model to determine the estimation variance for each of the 64 blocks. Averaging these values yields the average estimation variance of the BH data surrounding a DH data. For example, kriging was performed using the Zn variogram for the configuration shown in Figure 17. Specific DH values at the 100 x 100ft spacing were not important since we were only concerned with the estimation variance (recall that the estimation variance is not dependent on the data value, but rather the data locations).
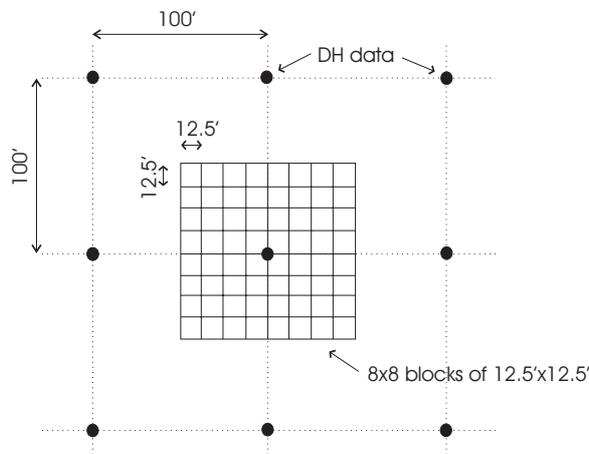
Figure 17: Configuration for determining average estimation variance. DH data are at 100 x 100ft spacing, BH data are at 10 x 12ft spacing. Set up an 8 x 8 grid centred about a DH data, with block sizes of 12.5 x 12.5 x 25ft. Perform kriging to determine the estimation variance at each block, and then average these to get the mean estimation variance.

Variable	$\bar{\sigma}^2_{est}$	$\rho_{expected}$	$\rho_{BH-E-type}$
Zn	0.658	0.342	0.522
Pb	0.647	0.353	0.695
Fe	0.577	0.423	0.648
Ba	0.834	0.166	0.375

Table 3: Summary table for determining expected correlation coefficient for one rock type: average estimation variance from kriging (second column), and expected correlation coefficient (third column), and correlation between BH and E-type estimate from simulation models (fourth column). For each variable, the actual correlation exceeds the expected correlation.

Figure 18 shows the histogram of estimation variances for the 64 blocks obtained from kriging and the crossplot of the BH data to the E-type estimate. From the histogram, the mean estimation variance is 0.658. From the above equation, the expected correlation is (1.0-0.658) or 0.342. The crossplot of the E-type estimate and the closest BH data shows a correlation of 0.522, which exceeds the expected correlation.

Table 3 summarizes the results of calculating the average estimation variance and the corresponding expected correlation for each variable within one rock type, and compares this with the correlation between the BH data and the E-type estimate of the simulation models.

A comparison of the last two columns in Table 3 shows that for all four variables, the correlation between the E-type estimates of the models was higher than the expected correlation determined from the variogram models. It was understandable that these correlations may appear unreasonably low; however, the correlations obtained from the models
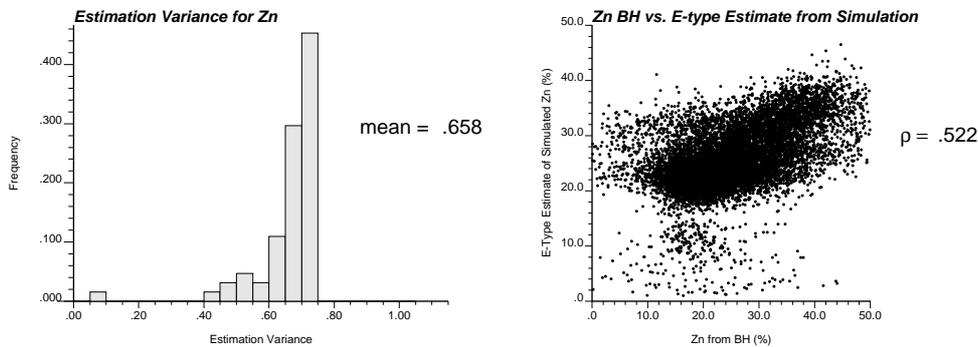
Figure 18: Calculation of expected correlation for Zn within one rock type: Histogram of estimation variance from kriging for 8 x 8 grid centered about a DH sample (left), and crossplot of BH data and E-type estimate (right).

Variable	ρ_{BH-LT}
Zn	0.417
Pb	0.439
Fe	0.559
Ba	0.298

Table 4: Correlation between BH data and Long Term Model values.

were comparable to the expected correlation.

Similar correlations were expected from the long term model, if the comparisons were carried out on a by rock type basis. Table 4 summarizes the correlation coefficients from this type of comparison using the long term model. As expected, these numbers were comparable to the fourth column in Table 3. Unfortunately, a direct comparison of these two sets of correlation coefficients would be technically incorrect since the expected correlations for the long term model were a function of the variography used to construct the simulation models. The appropriate variogam models for the long term model were not available.

Model Accuracy and Precision. Another interesting validation is to assess model accuracy and precision. In a statistical context, a model is considered accurate if for a given symmetric probability interval p, the fraction of true values falling within the p interval is greater than or equal to p for all p in $[0, 1]$. For instance, for a probability interval of 50%, an accurate model should have at least 50% of the true values falling within this interval. The term precision refers to how close this fraction of true values is to p for all p in $[0, 1]$ [1, 2]. These two closely related terms are neatly characterized by a simple crossplot of the fraction of true values against the corresponding probability interval. A model is both accurate and precise if this crossplot shows a 45 degree line. Figure 19 shows the crossplot for the E-type estimate of Zn, with the BH data taken to be the true data. For the 70% probability interval, the fraction of the true values falling within this interval is 63%. For this and all other intervals shown in the plot, the match between these two numbers was
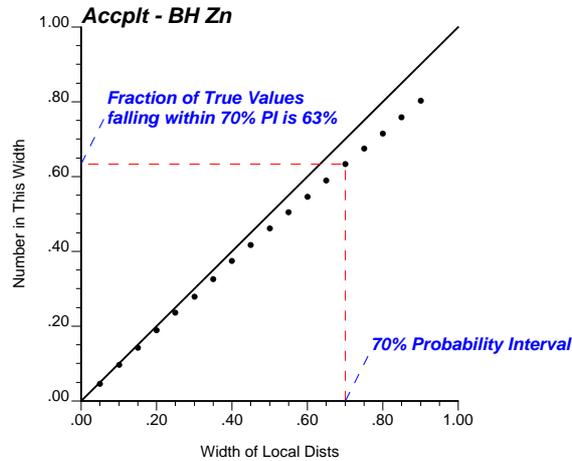
22

Figure 19: Accuracy plot for E-type estimate of Zn data.

sufficiently close to indicate that the Zn models were fairly accurate and precise. Figure 20 shows the crossplots corresponding to Zn, Pb, Fe and Ba (Zn crossplot is identical to that shown in Figure 19). For all four variables, the plots show that the models were satisfactory in their accuracy and precision, with the Fe models as the most accurate and precise.

Scaling up to 25ft blocks. For each of the merged realizations, the $(12.5\text{ft})^3$ grid is scaled up to a $(25\text{ft})^3$ grid to match the resolution of the existing grade and geology models. Similar to the upscaled models for BH comparison, upscaling to $(25\text{ft})^3$ blocks was performed by a weighted average based on specific gravity equations. Figure 21 shows a few of the simulated realizations for Zn at the 25ft grid resolution.

Comparison between Simulation to Existing Long-Term Model. Figure 22 shows a visual comparison between the E-type estimate from simulation and the existing long term model for bench 850. As expected, both maps were smooth. The E-type estimate should be similar to the kriged results because the E-type is the expected value taken over 40 realizations at each location within the block model and kriging gives the expected value at each location. Figure 23 shows the crossplot comparison of the E-type estimates from simulation to the corresponding long-term model values. The correlations between the two approaches were high in the case of all four primary variables. The comparison between Ba showed the most differences in the crossplot, with high simulated values paired with some corresponding low long-term model values, and vice versa. Despite this, there was a strong positive correlation. Overall, the high correlations between the two modeling approaches were encouraging statistics that provide validation for both the simulations and the existing models.
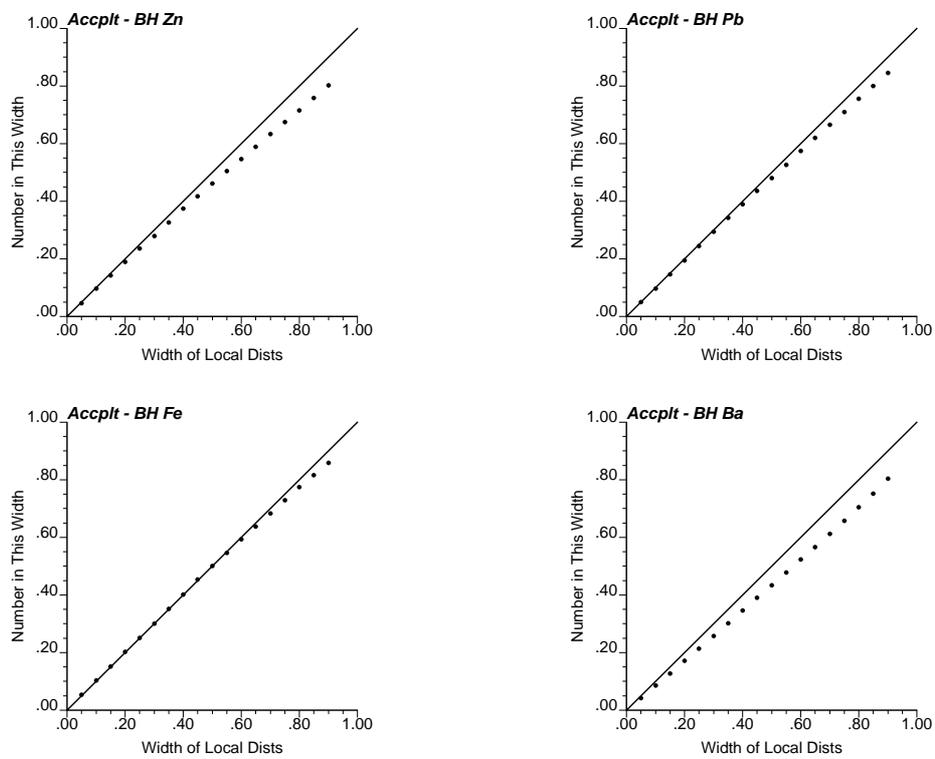
Figure 20: Accuracy plot for BH data and E-type estimate of simulation models: Zn (top left), Pb (top right), Fe (bottom left) and Ba (bottom right).
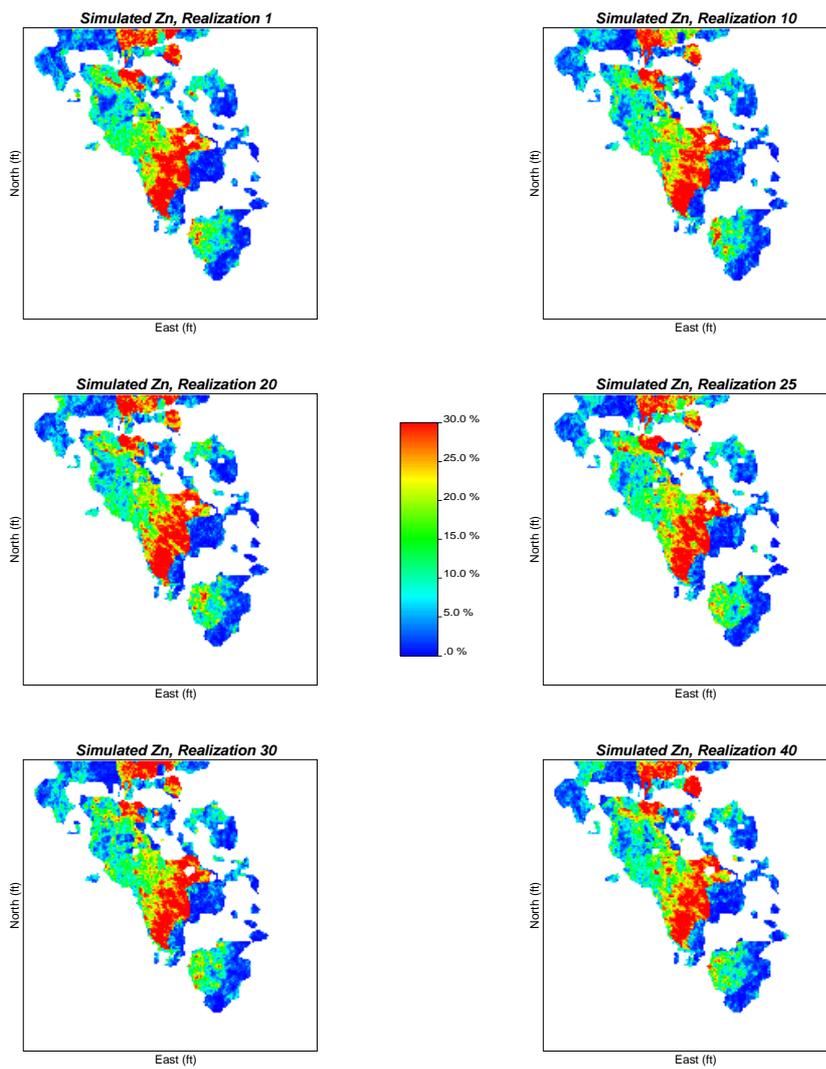
Figure 21: Simulated realizations of Zn at 25ft grid resolution. The section shown corresponds to bench 850 (spanning elevations 850 to 875ft).
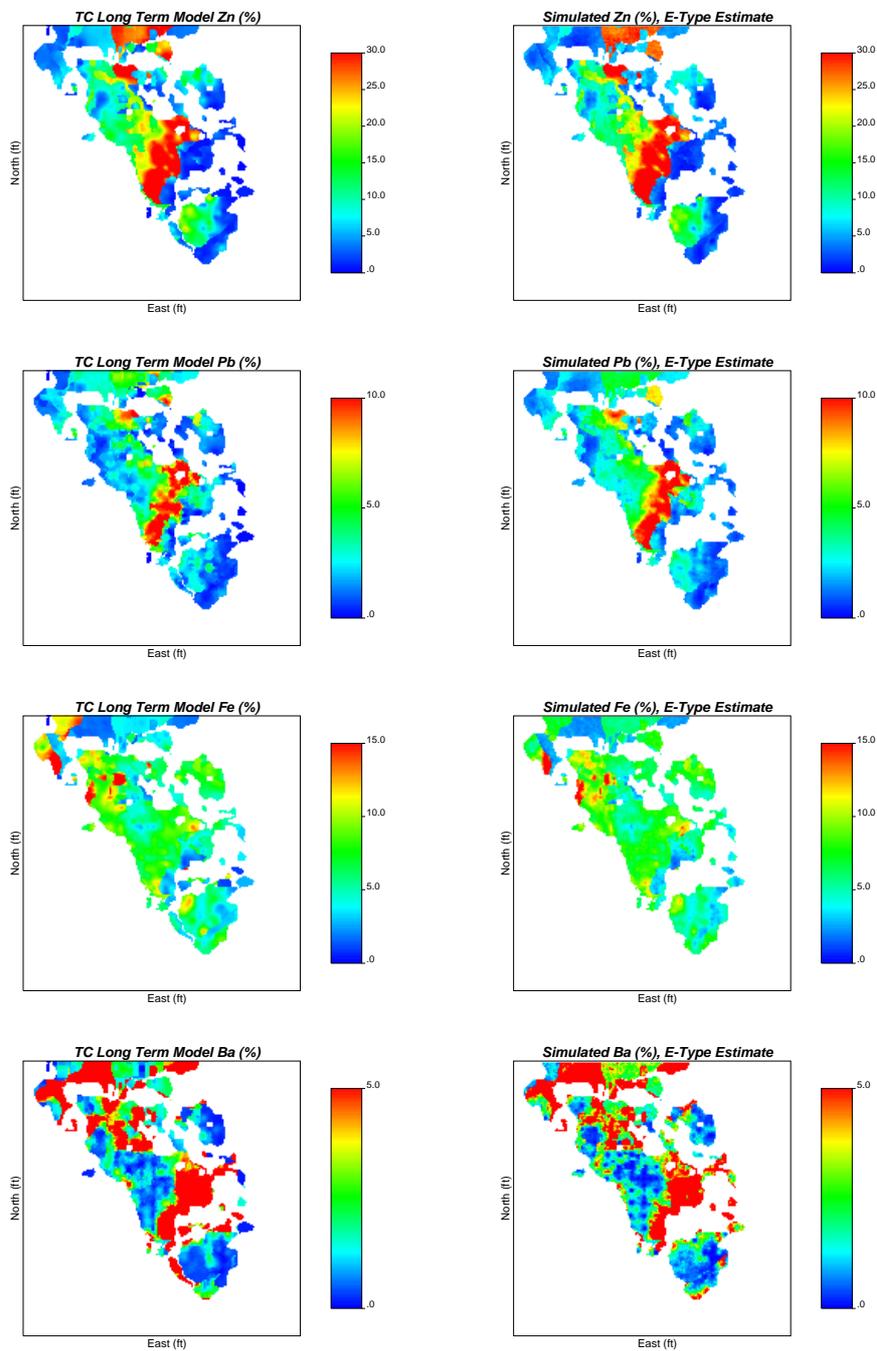
Figure 22: Comparison between existing long term model (left) with E-type estimate from simulation (right) at 25ft grid resolution. The section shown corresponds to bench 850 (spanning elevations 850 to 875ft).
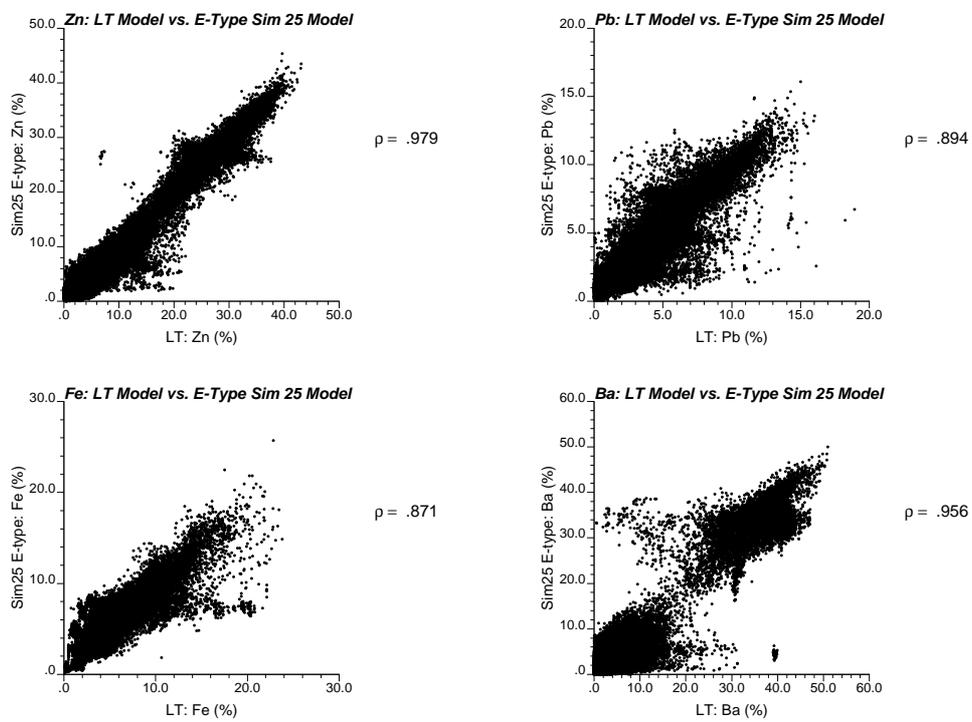
Figure 23: Crossplots of Zn, Pb, Fe and Ba from E-type estimates and existing long-term model.

Variable	BH-DH	BH-LT Model	BH-E-type	LT Model - E-Type
Zn	0.649	0.803	0.802	0.979
Pb	0.563	0.696	0.743	0.894
Fe	0.572	0.641	0.620	0.871
Ba	0.679	0.856	0.858	0.956

Table 5: Summary of correlation coefficients from all comparisons: BH to DH, BH to Long Term (LT) model, BH to E-type estimate, and Long Term model to E-type estimate.

Summary of Comparisons

This section described the comparisons between the BH and the DH data, the BH and the E-type estimate of the simulations, the BH and the existing long-term model, and the E-type estimates to the long term model. Table 5 gives the summary of the correlation coefficients for these comparisons. From Table 5, the worst correlations were given by the BH-DH comparison, which was expected given that the paired BH to DH samples may be separated by distances of up to 50ft. This type of comparison was consistent with a comparison between a simulated or E-type model and the BH data, since model values also suffer from being far away from real DH data. The comparison between the BH and both the long term model and the E-type estimate from simulation showed very similar correlations, thus indicating that both models have similar predictive abilities.

Given these comparable results, the conditional simulations were considered an improvement over the existing model in that multivariate relations were honoured. The long term model was generated using independent kriging of each variable. Consequently, there was no assurance to honour the multivariate relations between the different metals. For example, the grade of Zn at any one location has no effect on the modeled grade of Pb, Fe or Ba at that same location. In contrast, the simulations were constructed by explicitly accounting for the multivariate relations between the different variables. As a result, the grade of one variable would affect the simulated value of other metals at the same location. Furthermore, the multiple realizations from simulation allow for the assessment of uncertainty on both a local and global scale for decision making.

6 The Value of Simulation

In practice, multiple variables are estimated independently with ordinary kriging. This section addresses the impact of the multivariate simulation approach using the stepwise conditional transform relative to the conventional practice of kriging.

The idea is to compare the profit of ore from both methods with true reference data coming from Red Dog. A profit function is applied to obtain a true profit dataset. A subset of the reference data will be extracted and used to model the grades using both kriging and simulation. The profit function will be applied to these grade models. Based on the expected profit from each approach, each block within the model will be classified as either ore or waste. The true profit at each location is known, so the profit from each model can be calculated.

Profit Function. The real profit function was not available; a profit function was developed for this exercise. The following simple function was proposed:

$$profit(\mathbf{u}) = (Zn(\mathbf{u}) \cdot r_z \cdot f_1(Ba(\mathbf{u})) \cdot f_2(Fe(\mathbf{u})) \cdot p_z + Pb(\mathbf{u}) \cdot r_p \cdot p_p - c_{fix}) \cdot tons \qquad (2)$$

where

$$
\begin{array}{rcl}
\mathbf{u} &=& \text{location vector} \\
Zn &=& \text{Zn grade} \\
r_z &=& \text{Zn recovery function used to scale the maximum Zn recovery} \\
f_1(Ba) &=& \text{factor that accounts for effect of Ba on Zn recovery} \\
f_2(Fe) &=& \text{factor that accounts for effect of Fe on Zn recovery} \\
p_z &=& \text{price of Zn, in \$US/ton} \\
Pb &=& \text{Pb grade} \\
r_p &=& \text{Pb recovery} \\
p_p &=& \text{price of Pb, in \$US/ton} \\
c_{fix} &=& \text{fixed cost in \$US/ton} \\
tons &=& \text{tons of material based on specific gravity equations} \\
&& \text{provided by Teck Cominco}
\end{array}
$$

The only information available are the metal grades. All other parameters were developed or chosen to be constant. The metal recoveries for both Zn and Pb, r_z and r_p, were calculated as Red Dog's five year average recovery (1998-2002) based on Teck Cominco's financial report [8]. These were 83.6% Zn recovery and 58.7% Pb recovery. The price for Zn was chosen to be \$680/ton of Zn, and the price for Pb was chosen as \$380/ton of Pb; both prices were approximated based on the metal prices from the London Metal Exchange in 2003 [9].

Although recovery functions were provided by Teck Cominco, those functions did not actually depend on Ba grade [6]. At the time of this work, Teck Cominco was developing new functions based on extensive metallurgical testing. In light of this lack of confidence in the recovery functions, the Zn recovery function was scaled by functions that quantify the impact of Ba grade and Fe grade as a fraction of the maximum recovery. For Zn, the recovery function (see Figure 24) reaches a constant maximum recovery beyond a threshold grade of 10% Zn. Lower Zn grades than this threshold results in a fraction of the maximum recovery, to a minimum of 50%. The rate of this change was expected to be fairly gradual.

A discount function for Ba content was developed by considering that a threshold grade of 7% Ba resulted in significant impact on Zn recovery. At grades below this threshold, the discount factor was expected be fairly constant. Near the threshold grade of 7%, an inflection point in the discount function was expected, and would gradually flatten at a minimum discount of 35% since some Zn would still be recovered (Figure 25).

The discount function for Fe was based on the recovery functions provided by Teck Cominco [6], and is illustrated in Figure 26.

All that remains to determine is the fixed cost. For this, an arbitrary cost per ton mined was chosen such that the area of interest yielded approximately 50% ore and 50% waste classification. This depends on the region chosen for modeling; for this exercise and the
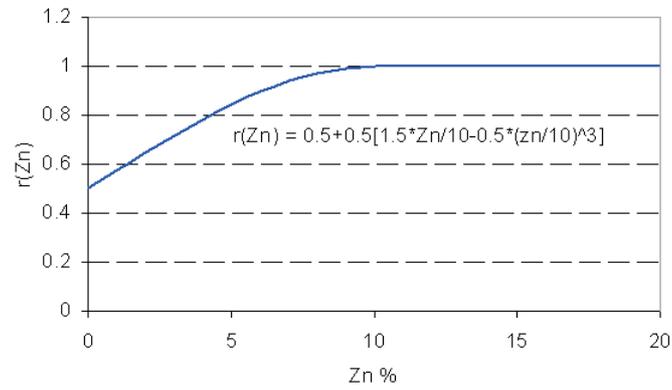
Figure 24: Zn recovery function developed for comparison of kriging and simulation. Function for grades between 0 and 10% is shown, no reduction in recovery was expected beyond 10%.
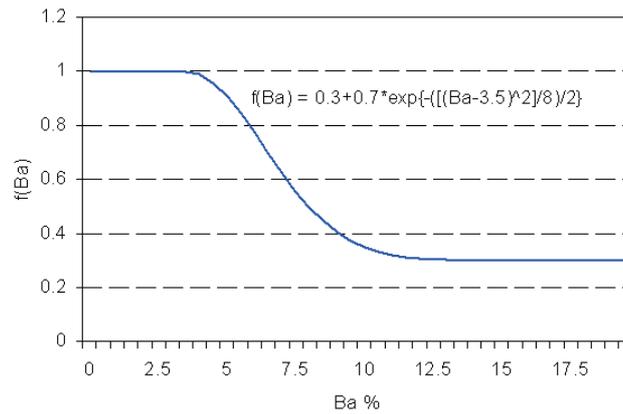

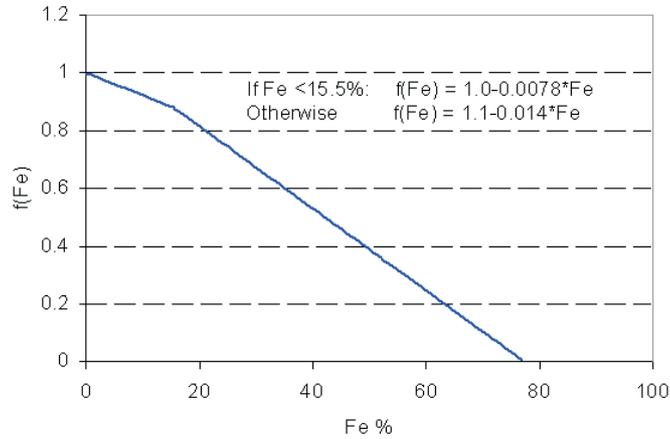Figure 25: Discount function on Zn recovery due to %Ba content.

Figure 26: Discount function on Zn recovery due to %Fe content. Two different discount functions were applied with a threshold grade of 15.5% Fe. These functions were based on the recovery functions provided by Teck Cominco Limited.

area described below, the fixed cost was set at $128/ton and accounts for all operational costs including mining and milling cost.

Reference Data. For a fair comparison to be made, real data must be used. The density and number of BH data available make it an attractive database for true data. Rather than modeling the entire area, only a small area will be modeled. The area was chosen to be in a marginal zone, where ore/waste classification based on the models would have the largest impact.

Figure 27 shows the available BH data in the chosen region of 400ft × 400ft in the 850 bench, and the subset of data extracted from this region. The available data consists of 532 BH samples of Zn, Pb, Fe and Ba. From this dataset, 25 samples separated at a nominal 100ft × 100ft spacing were chosen to act as exploration data. This spacing is consistent with the DH data available for Red Dog. This subset of data was used as conditioning data for kriging and simulation.

Model Construction. The model grid was chosen to be 10ft × 10ft × 25ft, which is similar to the 10ft × 12ft × 25ft spacing of the BH data. A total of 1600 blocks were modeled.

With only 25 samples available for modeling, variography would be very difficult. To filter out the influence of poor variogram inference, variograms for both approaches were calculated and fitted using the reference 532 BH data.

The variograms for kriging were calculated for the original data (Figure 28). The variograms for simulation were calculated and fitted for the stepwise conditionally transformed data (Figure 29). In both sets of variograms, a trend was apparent from the experimental points extending beyond the sill of 1.0. This was not surprising given that the area was purposely chosen to be in the transition zone between ore and waste material, hence a trend
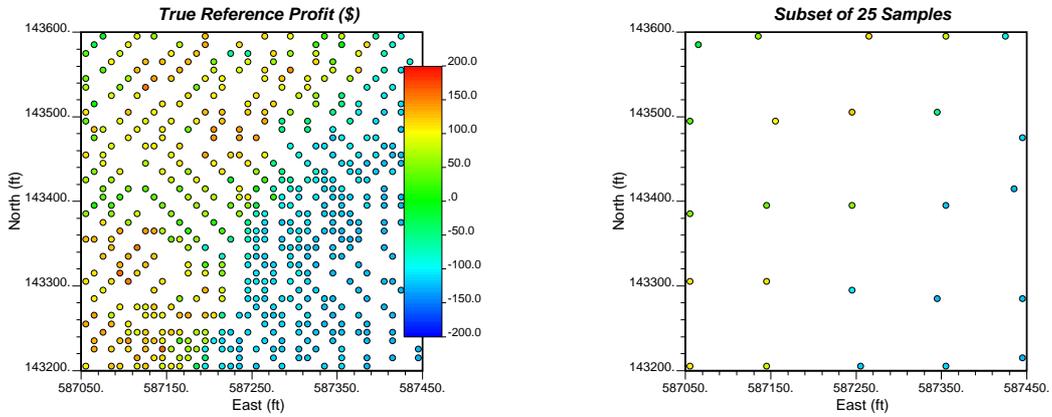
Figure 27: Location map of reference BH data (left) and sampled BH data (right) for use in comparing model approaches.

from low to high grades was expected. Trend modeling was not performed for this exercise because of the relatively small area.

For kriging, each variable was estimated independently using ordinary kriging. For simulation, the stepwise conditionally transformed variables were independently simulated using sequential Gaussian simulation to generate 100 realizations of the grades. Figure 30 shows a comparison of the estimated grades from kriging and one realization of the simulated grades. As expected, the kriged models were very smooth, while the simulated realization showed greater variability while honouring the same large scale features shown in the kriged models.

Results. These grade models were then processed by applying the profit function at each location within the model. Although, 100 realizations of profit were available from simulation, the ore/waste classification was based on the expected profit map obtained by calculating the expected value of profit at each location. Figure 31 shows the profit map obtained from simulation and kriging along with the true profit at the 532 locations where real data were available.

Although 1600 locations were modeled, only the 532 blocks corresponding to locations where true data were available can be compared. At these locations, the true profit was known. The models from kriging and simulation were used to classify the 532 locations as either ore or waste:

$$i(\mathbf{u}_\alpha; profit) = \left\{ \begin{array}{ll} ore, & \text{if profit}(\mathbf{u}_\alpha) \geq 0 \\ waste, & \text{if profit}(\mathbf{u}_\alpha) < 0 \end{array} \right.$$

Figure 32 shows the comparison of the ore/waste classification of the 532 locations from the true relative to the kriging and the simulation approaches. Overall, both approaches clearly show the waste and the ore region; relatively few blocks were misclassified.

Figure 33 shows the summary of the ore/waste classification from both kriging and simulation relative to the true classification. The tables show that the kriging approach
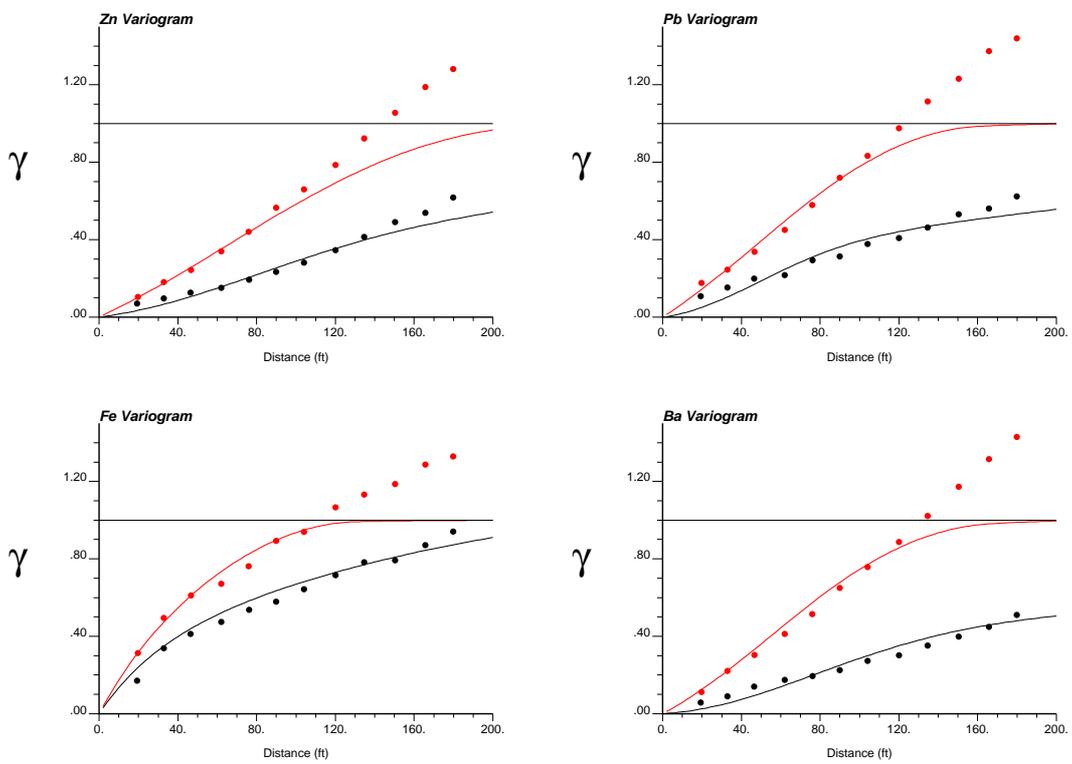
Figure 28: Variograms of direct space data for use in ordinary kriging approach: Zn (top left), Pb (top right), Fe (bottom left) and Ba (bottom right). The two directions shown correspond to the horizontal minimum and maximum directions of continuity.
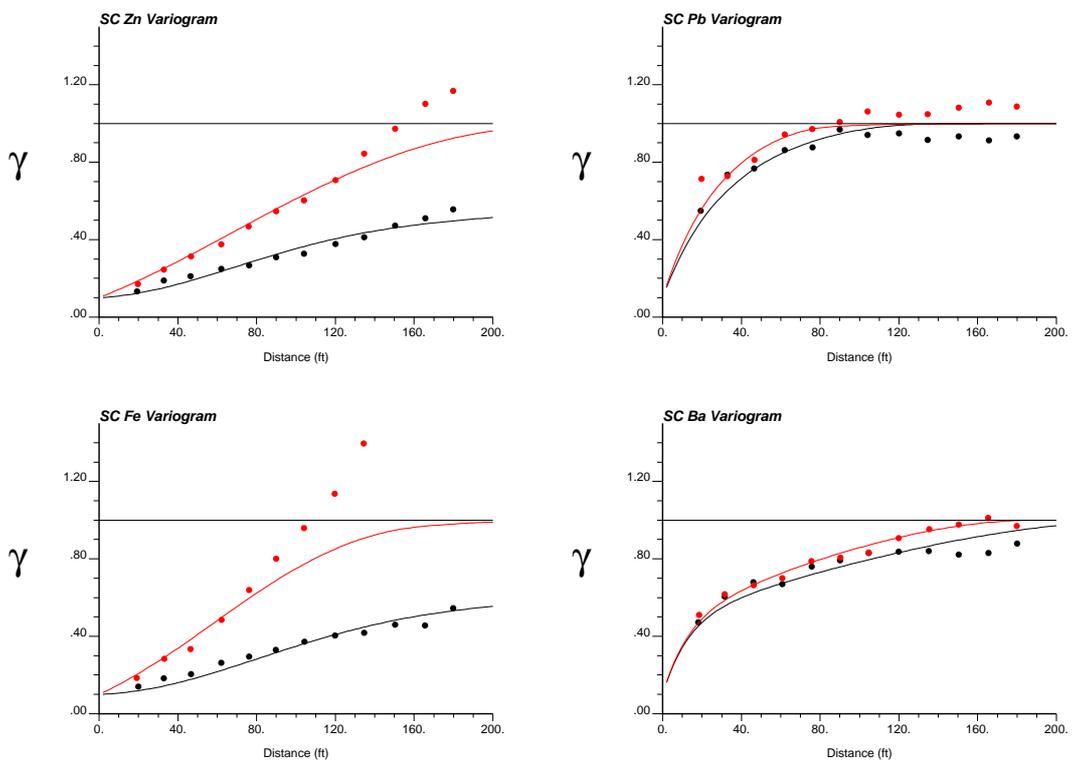
Figure 29: Variograms of stepwise conditional scores for use in simulation approach.: Zn (top left), Pb (top right), Fe (bottom left) and Ba (bottom right). The two directions shown correspond to the horizontal minimum and maximum directions of continuity.
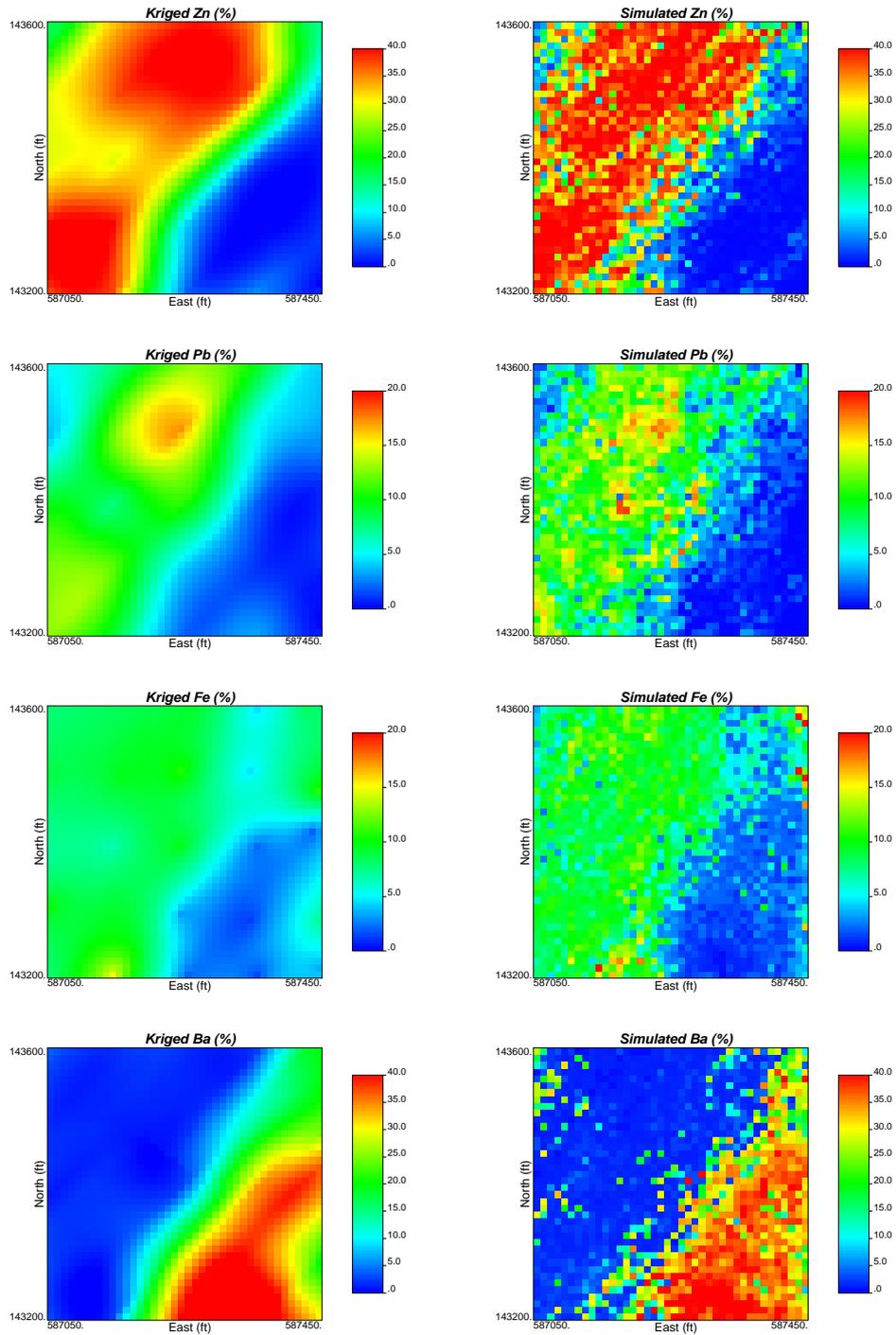
Figure 30: Comparison of kriged model (left) and one realization from simulation (right) for Zn, Pb, Fe and Ba (from top to bottom).
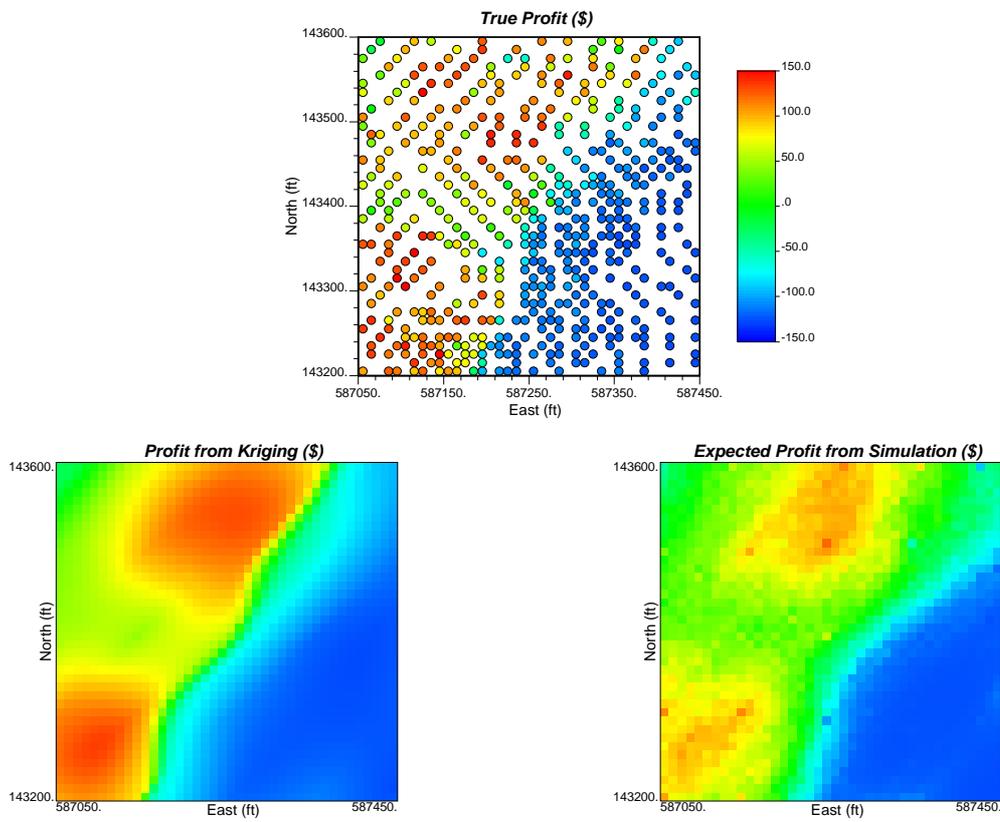
Figure 31: Comparison of true profit map at data locations (top) and the profit map for ore/waste classification from kriging (bottom left) and simulation (bottom right).
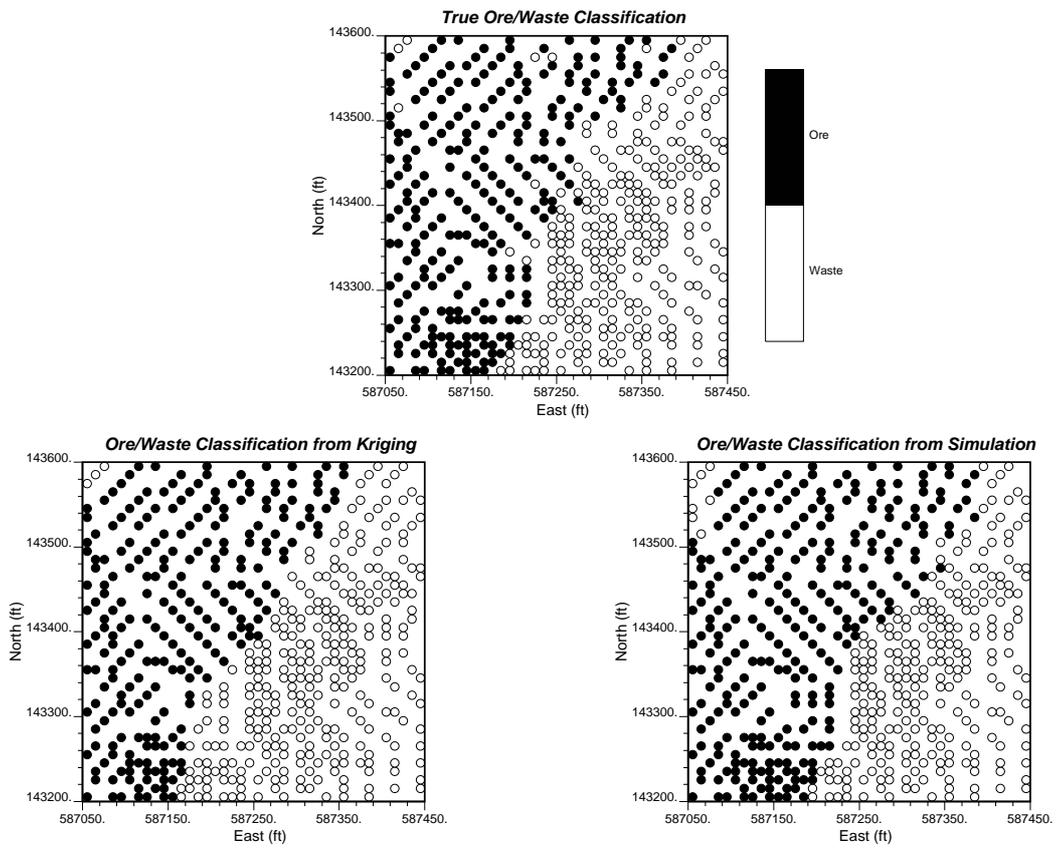
Figure 32: Comparison of true ore/waste classification (top) and the classification from kriging (bottom left) and simulation (bottom right) at data locations.

	TRUE	
	Ore	Waste
Ore	225	11
Waste	26	270

Kriging

	TRUE	
	Ore	Waste
Ore	246	27
Waste	5	254

Simulation

Figure 33: Ore/Waste classification summary of kriging (left) and simulation (right) relative to true ore/waste classification.

resulted in a total 7% of blocks that were misclassified, compared to the 6% misclassified by simulation. From the true profit, 251 blocks (47% of the true data) were classified as ore; simulation correctly classified ore for 98% of those blocks while kriging correctly classified 90% of those blocks.

For those blocks classified as ore, the profit of ore mined as a result of the classification from each method was compared with the true profit of $7.89M (million). The results from such a comparison showed that the simulation approach yielded $7.28M while kriging yielded $7.06M in profit. Although these profit values appear high for the relatively small area, the relative percentage increase in profit is the key result. Multivariate simulation resulted in 92% of the true profit relative to the 89% yielded by kriging. In practice, this 3% difference may translate to several millions of dollars in increased profit.

7   Remarks

For the seven variables within the eight rock types, conditional simulation models were constructed using the stepwise conditional transformation technique to account for the multivariate relations. 12.5ft composites and a geology model at $(25ft)^3$ resolution were used to develop these models. Each model was validated by checking reproduction of the input drillhole data, representative histogram, variogram, and the multivariate distributions.

Validation with additional data and comparisons to the existing long term model showed the conditional simulations have similar predictive abilities to the existing models. Multivariate simulation provides two significant improvements from the existing long term model. Firstly, the simulated realizations account for the complex multivariate relations inherent in the data, resulting in models that respect these relations on both a local and global scale. Secondly, the simulation models provide a basis for some interesting applications for decision making and risk assessment. These applications range from classification of ore/waste regions based on complex criteria to recovery forecasting given a clear understanding of metallurgical processes and relations.

A comparison of the multivariate simulation approach used in this case study and the common practice of kriging multiple variables independently showed that the simulation models resulted in an increase in profit of 3% over the kriging approach, yielding a total of 92% of the true profit.

References

[1] C. V. Deutsch. Direct assessment of local accuracy and precision. In E. Y. Baafi and N. A. Schofield, editors, *Geostatistics Wollongong '96*, volume 1, pages 115–125. Kluwer Academic Publishers, 1996.

[2] C. V. Deutsch. *Geostatistical Reservoir Modeling*. Oxford University Press, New York, 2002.

[3] C. V. Deutsch and A. G. Journel. *GSLIB: Geostatistical Software Library and User's Guide*. Oxford University Press, New York, 2nd edition, 1998.

[4] E. H. Isaaks. *The Application of Monte Carlo Methods to the Analysis of Spatially Correlated Data*. PhD thesis, Stanford University, Stanford, CA, 1990.

[5] A. G. Journel and W. Xu. Posterior identification of histograms conditional to local data. *Math Geology*, 26:323–359, 1994.

[6] O. Leuangthong. *Stepwise Conditional Transformation for Multivariate Geostatistical Simulation*. PhD thesis, University of Alberta, Edmonton, AB, 2003.

[7] O. Leuangthong and C. Deutsch. Stepwise conditional transformation for simulation of multiple variables. *Mathematical Geology*, 35:155–173, 2003.

[8] Web. *Teck Cominco Limited 2002 Annual Report*. Website address http://www.teckcominco.com/investors/reports/ar2002/tc-2002-operations.pdf, 2002.

[9] Web. *The Northern Miner website*. Website address http://www.northernminer.com/, 2003.