A Practical Way to Summarize Uncertainty for Blockwise Resource Classification

Julián M. Ortiz (jmo1@ualberta.ca) and Clayton V. Deutsch (cdeutsch@ualberta.ca)
Department of Civil & Environmental Engineering
University of Alberta

Abstract

Geostatistical simulation allows computing multiple realizations that provide a measure of uncertainty; however, for large domains, the computer time and storage can be a problem. The use of multi-Gaussian kriging is illustrated as an alternative to stochastic simulation for quantification of local uncertainty on block grades. A probabilistic approach is proposed to quantify and visualize the block grade uncertainty. It is based on calculating the expected width of the confidence interval at a given probability level, for the average block grade. This probabilistic assessment of the uncertainty can be used for classification of the resources and reserves.

The proposed method accounts for the proportional effect of the data and therefore provides a more realistic way to quantify uncertainty in the resources on a probabilistic framework. The methodology is applied to an oil sands deposit in Northern Alberta.

Introduction

Heterogeneity and uncertainty are inherent aspects of mining exploration and production. In reality there is a single true distribution of rock types and grades; however, in presence of relatively sparse sampling we have no access to that true distribution. We know that the spatial distribution of petrophysical properties is controlled by depositional processes combined with subsequent migration and deposition. These processes cannot be modelled deterministically since we have no access to the complex details of these processes or to the initial and boundary conditions. Geostatistical techniques are used to construct models of uncertainty in rock types and petrophysical properties. These models quantify our state of incomplete knowledge and are useful for mine planning, decisions regarding future data collection, and reporting [2, 3, 6, 14].

Data often show more variability in the high grades than in the low ones, which is known as the proportional effect [14]. In these instances, it is incorrect to use the kriging variance for uncertainty assessment [7, 14]; a measure that accounts for the different variability of highs and lows is needed.

Geostatistical simulation overcomes this problem in one of two ways: (1) accounting for the proportional effect via a transformation of the data (Gaussian simulation), or (2) characterizing differently high and low grades (indicator simulation). Multiple realizations are generated to honor the spatial continuity of the grades. Then, for a given block, the uncertainty in the predicted value can be evaluated more accurately, since the local distribution of block grades is fully known (**Figure 1**).

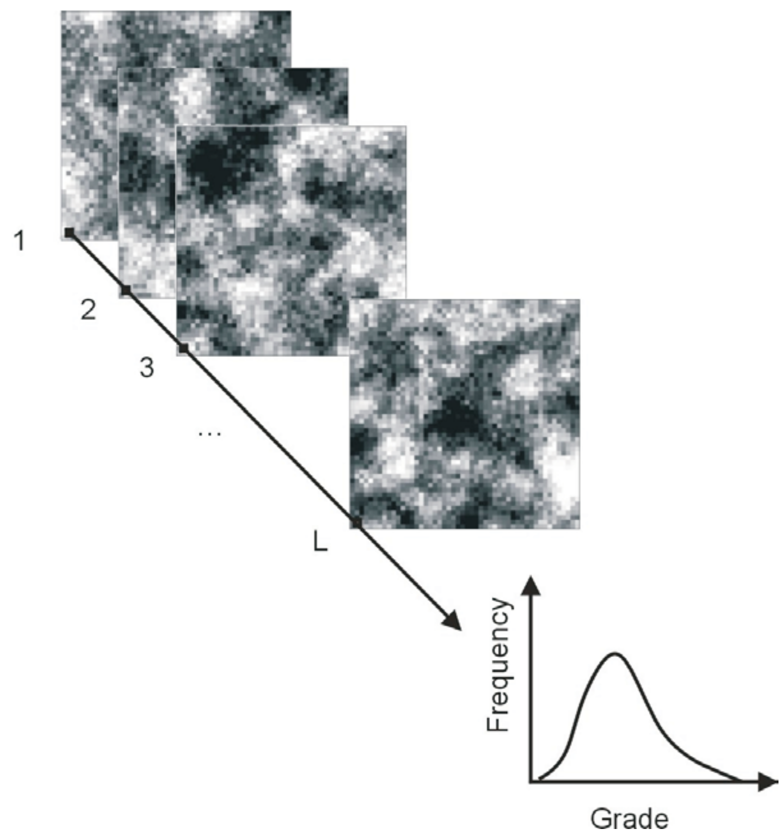

Figure 1: The local block grade distribution is obtained by pooling the simulated block grades at a given location on each one of the L simulated realizations.

The use of stochastic simulation may be prohibitive if a very large number of blocks is to be simulated. Generating and storing multiple realizations of large models are both time consuming and limited by available hardware. Multi-Gaussian kriging is proposed to locally evaluate the block grade uncertainty [23, 24]. Similar to the Gaussian simulation methods, multi-Gaussian kriging accounts for the proportional effect of the grades by means of a transformation from the original values. It is more efficient than generating multiple realizations using a simulation method, since only the kriging estimate and kriging variance of the transformed values must be stored to fully characterize the local block grade uncertainty. This is simpler than generating and storing multiple realizations of the attribute using a simulation algorithm.

The use of multi-Gaussian kriging for block uncertainty assessment requires the assumption that the transformed grades average linearly in Gaussian units, which is a reasonable assumption if the distribution of grades is not highly skewed.

Knowing the distribution of uncertainty of the block grades allows classifying the resources. This requires summarizing this uncertainty. Classification into measured, indicated, and inferred resources must be in accordance with local legal and professional guidelines [1, 9, 18, 21]. Reporting of the uncertainty in the block grade as a confidence interval that the true grade falls within a given probability is an advantage. It is a repeatable measure of the quality of the estimate. This measure of quality takes into account the distance to the samples that were used to calculate the estimate, the continuity of the grade in space, and the variability of the grade as a function of the local mean, or proportional effect as previously discussed. It also considers the geological continuity dictated by the rock type model.

We present a method to report the confidence in the estimates using geostatistical estimation or simulation. Uncertainty in the resource is reported at the 90% and 50% confidence level, that is, we report the proportions corresponding to the width of the confidence interval relative to the mean where the true value falls 90% and 50% of the time (**Figure 2**).

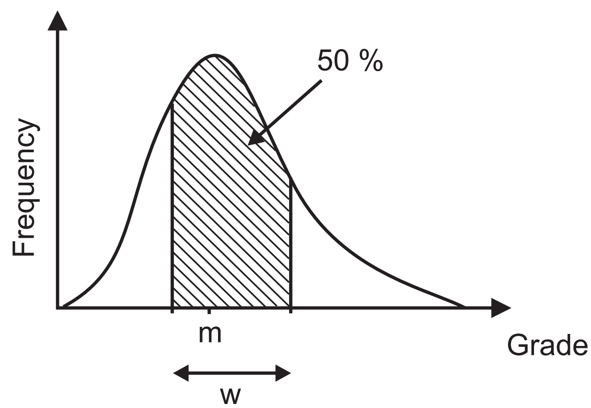

Figure 2: The mean of the distribution (estimate m) and width of the confidence interval w are used to summarize and report block grade uncertainty at a given confidence level.

Methodology

The methodology can be divided into three main steps:

1. Determination of the local block grade distribution.
2. Calculation of the proportion of the width of the confidence interval at a given significance level relative to the expected grade.
3. Classification of the blocks by defining threshold proportions to differentiate measured, indicated, and inferred resources.

Calculation of the Local Block Grade Distributions

The block grade values are characterized by their distribution of uncertainty, which reveals the range of possible values that it can take with the corresponding probabilities to be that value. To quantify the uncertainty in the resources, it is not necessary to jointly simulate the domain of interest, the uncertainty in a block grade can be assessed one block at a time. Therefore, it is enough to assess local distributions of uncertainty. Conventionally, if the goal is to assess the grade of a large volume and its uncertainty, then all blocks included in that larger volume must be simulated simultaneously, so that the average grade accounts for the spatial correlation between them.

As mentioned above, in cases where the domain to be simulated is large, the generation of multiple realizations is impractical, because of computing time and storage considerations.

Given the distribution of the samples and a measure of spatial continuity such as the variogram of the grade of interest, the expected mean and variance of each block distribution can be calculated by conventional geostatistical techniques [3, 6, 7, 14]. However, the key unknown is the shape of the local block grade distributions. Some assumption is then required [4]:

- The easiest assumption is to consider that the block distribution has the same shape as the point grade distribution. This can be obtained via an affine correction, which modifies the variance of the distribution, keeping shape and mean constant, since the linear average of grades from point to block support does not change the mean value. This assumption is often unrealistic, since in practice a symmetrization of the distribution can be seen [7].

- Using the hypothesis of permanence of normality, a multi-Gaussian approach can be taken to evaluate the local uncertainty [13, 23, 24]. In this case, the data are transformed to normal scores and then kriging of these transformed values is performed. The estimate is calculated in original units by back-transforming the local distribution and performing numerical integration. Different quantiles can also be back-transformed to assess low and high bounds of a confidence interval.

- Similar to the multi-Gaussian case, if the global distribution is close to a lognormal distribution, then under the hypothesis of permanence of lognormality, all conditional distributions are also lognormal, and so the mean and any quantiles of interest can be back-calculated [10, 19, 22].

- Disjunctive kriging, along with the discrete Gaussian model for change of support, can also be used to assess the local uncertainty [8, 15, 16, 17, 20].

- Indicator kriging allows the estimation of the full conditional distribution at every location, by kriging the indicator coded samples at different thresholds [11, 12, 13].

- Finally, stochastic simulation methods allow the characterization of the local distribution by generating multiple realizations of the variable at the location of interest. The local distribution can then be constructed by pooling together, in a histogram, the L simulated values at a given location (**Figure 1**). Confidence intervals, mean values, and the probability of exceeding a threshold, are then straightforward to calculate [3, 5, 6].

No matter what method is used, the parameters required to calculate the statistic that summarizes the uncertainty in the block grade are: the expected grade of the block, and a confidence interval defined by a low bound and a high bound, such that the grade is within these bounds with a given probability, say 90 % of the time.

If a transformation of the distribution is required, such as in the multi-Gaussian case, the back-transformed confidence interval is actually centered around the mean of the distribution in transformed (Gaussian) units, but it is not centered around the mean after back-transformation, that is, in original units.

Summarizing Uncertainty in Block Grades

Once the full distribution of block grades has been characterized, it must be summarized for reporting purposes. A confidence interval at a given probability level can be calculated by finding the grade values at percentiles of interest. This may require

Uncertainty in the block grades can be summarized by the ratio of a given confidence interval over the mean value (**Figure 2**). The proportion relative to the mean where the true value falls with a chance given by the significance level is calculated as (**Figure 3**, bottom left):

$$\pm P\% = \frac{1}{2} \cdot \frac{z_{high} - z_{low}}{\bar{z}} \cdot 100\%$$

where z_{high} and z_{low} are the grade values corresponding to the upper and lower bound of the confidence interval at some probability level $(1 - \alpha)$, $\bar{z}$ is the expected block grade, and P is such that the interval $[z_{low}, z_{high}]$ contains the true value $(1 - \alpha) \cdot 100\%$ of the time. For example, if we consider $\alpha = 0.10$, the true value will be within the confidence interval defined by $[z_{low}, z_{high}]$ 90% of the time.

Classification of Resources

Classification can be defined based on the relative uncertainty of the block grades at a given significance level. Interpretation of the results depending on the significance level may not be straightforward since there has not been enough practice regarding this type of summary statistic for uncertainty in block grades. Current regulations require the reporting of the uncertainty in the resources. Moving from a geometric definition of the resources, for example using distance to nearest drillhole or drillhole density for classification, to a probabilistic approach such as the one proposed in this paper, requires adjusting the confidence level to ensure that the uncertainty associated to each category of resources is consistent with the previous definition used for classification. The mine staff must be comfortable with the definition of the categories and have an understanding of the meaning of a given significance level.

Uncertainty Quantification at an Oil Sand Deposit in Northern Alberta

The methodology described is applied to an Oil Sand deposit in Northern Alberta, where a model is built for an area larger than 170 km^2. The objectives of the study are to assess uncertainty in the bitumen content, and to summarize the uncertainty in the predicted resource.

Modelling is done on 25 x 25 x 3 m^3 blocks. This size is defined by the selectivity of the mining method, which allows extraction of slices of a minimum of 3 m. Three main depositional environments are considered in the rock type model: fluvial, marine, and estuarine. The domain modelled includes over ten million blocks in total; however, just over a quarter of these blocks are within the three rock types of interest.

Information Available

Samples containing grade information and rock type code are available from a number of drillholes. Composites are calculated at 3 m length for modelling, coinciding with the bench definition in the block model.

Three main depositional settings exist: marine, estuarine, and fluvial. They are the facies or rock types that will be used for modelling. A deterministic rock type model is used to constrain the estimation of local distributions of uncertainty. Only data within the same rock type will be used to infer the conditional distribution at an unsampled location within that rock type.

Exploratory Statistics

There are more drillholes in the area that will be mined in the short term; this corresponds to a higher bitumen portion of the study area. The data are all equally valid, but should receive different weights or influence going into the global distribution to be representative of the entire area of interest. Declustering is performed to assign those weights. Cell declustering is the standard approach [3, 6]; this method was applied to the bitumen content.

Uncertainty Quantification

Given the large number of blocks in the model, multi-Gaussian kriging is used. There are significant trends in both the vertical and horizontal direction. For this reason, ordinary kriging with an appropriate search neighborhood has been chosen.

Since bitumen grade is not normally distributed, a transformation is required as schematically illustrated in **Figure 4**. Every sample value is assigned a corresponding normal score. Kriging is performed to compute the estimate and estimation variance of the normal score value at every location. Since the shape of this local distribution of uncertainty is assumed to be Gaussian, then any probability interval can be retrieved from it (**Figure 3**), by means of the relationship between the samples and their normal scores. The mean can also be computed from it by numerical integration (**Figure 5**)

Variograms are required within each facies. Under the requirement of multi-Gaussianity, variograms of the normal scores of the bitumen grades are calculated and fit in the vertical

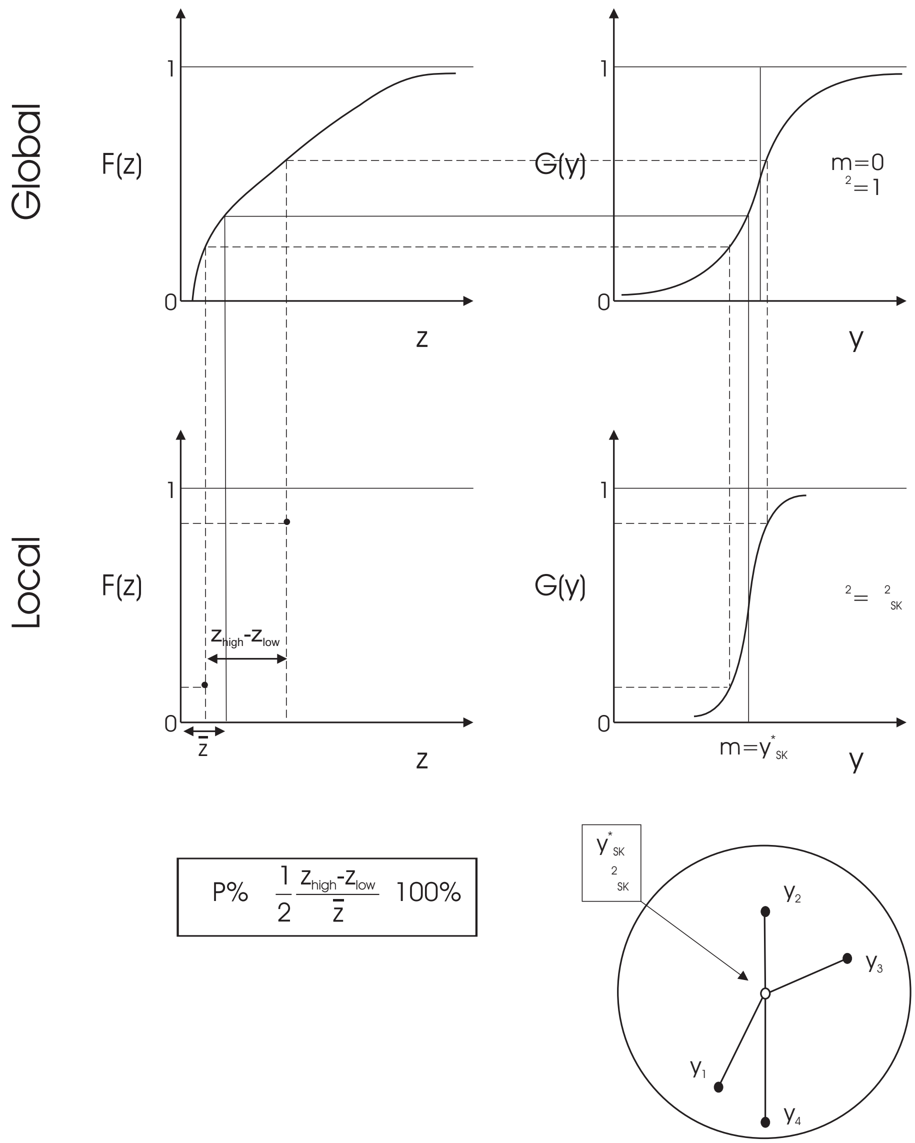

Figure 3: Calculation of the uncertainty in block grade. Kriging provides the mean and variance. Confidence limits are back-transformed to the original units. The uncertainty in the estimate is quantified using the width of the confidence interval over its expected value.

and horizontal directions. In general, the variograms are well defined and suitable for all required geostatistical calculations. These are shown in **Figure 6** and tabulated in **Table 1**. Some remarks:

- There is a slight zonal anisotropy (the experimentally observed sill variance is less than the full variance of 1.0) in almost all of the horizontal variograms, which is due to persistent long-range correlation in the horizontal direction. The vertical trend is part of this same phenomena.

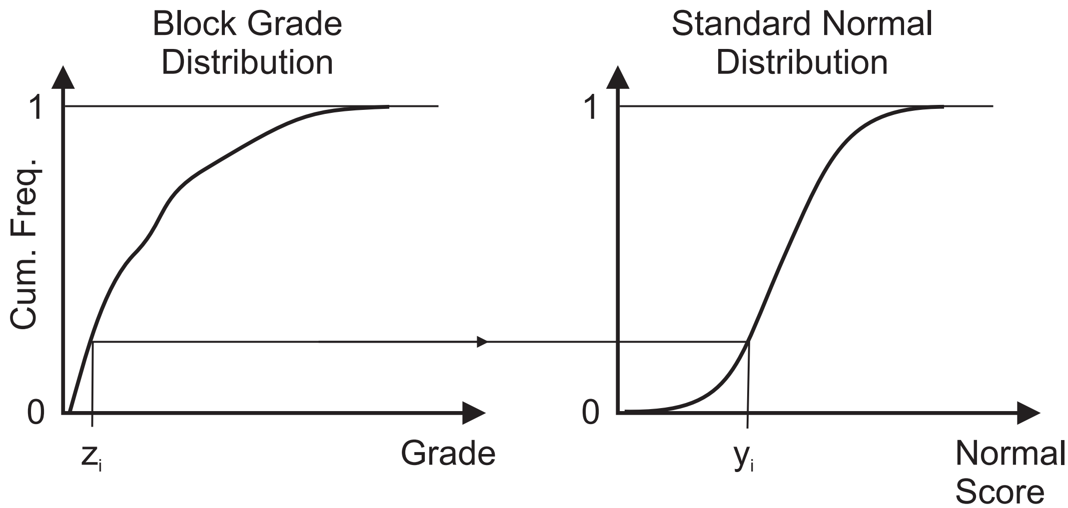

Figure 4: The normal score transformation is illustrated for a data z_i. The cumulative frequency is read in the original block grade distribution and the corresponding value y_i of a standard normal distribution, that is a Gaussian distribution with mean 0 and variance 1, is assigned to the data location.

- The vertical variograms in the marine facies are poorly defined because of the relatively thin marine layer and few data for reliable variograms.
- The vertical variograms for the estuarine and fluvial facies tend to go above the sill variance because of vertical trends.

			Structure 1: spherical			Structure 2: spherical			Structure 3: spherical		
Layer	Variable	Nugget Effect	Sill Contr.	Range Horiz.	Range Vertical	Sill Contr.	Range Horiz.	Range Vertical	Sill Contr.	Range Horiz.	Range Vertical
MARINE	BIT	0.05	0.30	20	20	0.45	2300	20	0.20	∞	20
ESTUARINE	BIT	0.05	0.25	20	28	0.45	2200	28	0.25	∞	28
FLUVIAL	BIT	0.02	0.38	18	18	0.30	1200	18	0.30	∞	18

Table 1: Variogram model parameters of the normal score transform of bitumen within each facies. All range values are in metres.

The data for kriging are the normal scores (standard Gaussian) transform of bitumen. The data are essentially at a point support relative to the 25 by 25 by 3 m block size. Block kriging with a discretization of 3 x 3 x 3 points per block is used to account for the fact that variability at a block support is significantly less than at the data-scale point support. Ordinary kriging was used because it is a robust version of estimation that accounts for local departures from stationarity, that is, areas that are systematically lower or higher in grade. The estimation variance is slightly higher than the common alternative of simple kriging; however, simple kriging depends on a stationary or statistically homogeneous domain, which is unrealistic for the area.

A relatively large horizontal search radius of 4000 m was chosen so that estimates would be made even in sparsely drilled areas. Of course, estimates far from available drill holes

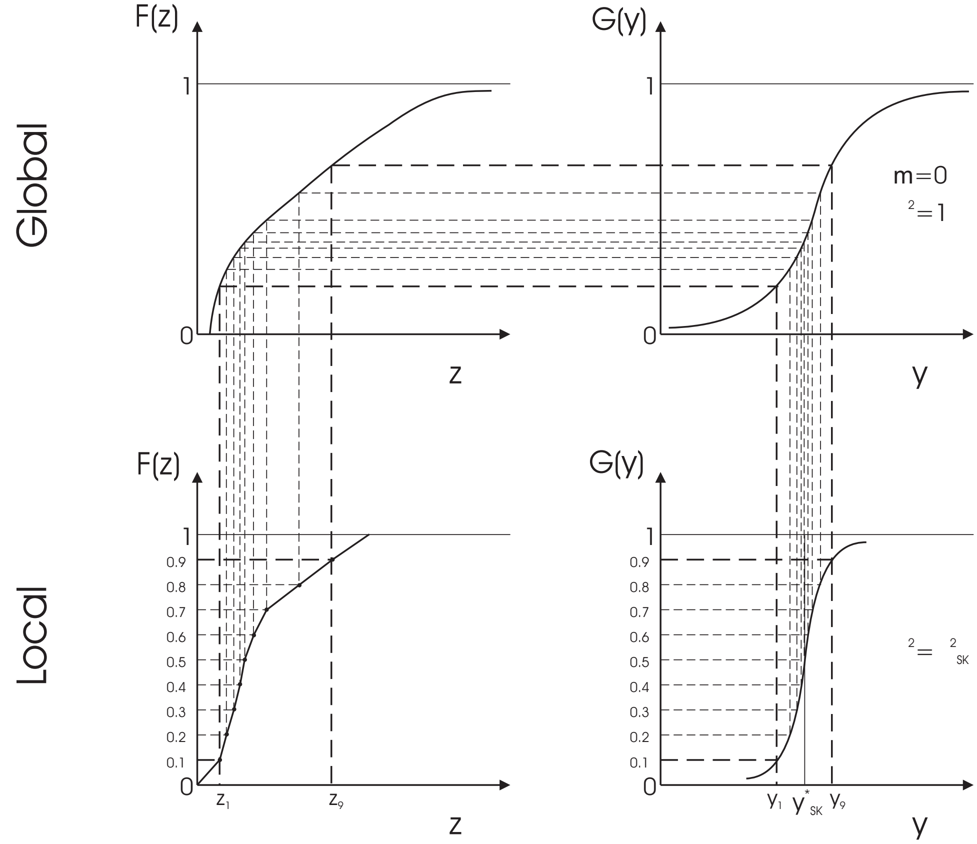

Figure 5: Calculation of the mean by numerical integration. The local uncertainty distribution is given by the kriging estimate and variance and the assumption that the shape is normal (bottom right). Several quantiles are calculated. In the illustration, the nine deciles of the distribution, $y_1, ..., y_9$, are back-transformed (top) and the corresponding values, $z_1, ..., z_9$, are used to calculate the mean as their average (bottom left).

will have large uncertainty. Most of the drill holes are vertical, therefore, a limited vertical search radius of 6m was used.

A maximum of 24 samples were used to estimate each block. A smaller number of samples would result in faster computational speed, but could result in estimates with too large uncertainty. A larger number of samples could result in estimates that use data too far away from the block being estimated. This number is a conservative balance between these considerations. A minimum of 20 samples was considered; this, too, is conservative. Accepting fewer samples within the 8000 by 12 m search ellipsoid would result in more estimates, but the uncertainty would be so large as to make the estimates unreliable.

The `kt3d` program from GSLIB was used for kriging [3]. Each kriging run leads to two values of importance: the best estimate of the variable and the kriging variance in transformed units. These values are used for uncertainty reporting. The three kriged models are assembled into a single block model using the available deterministic rock types

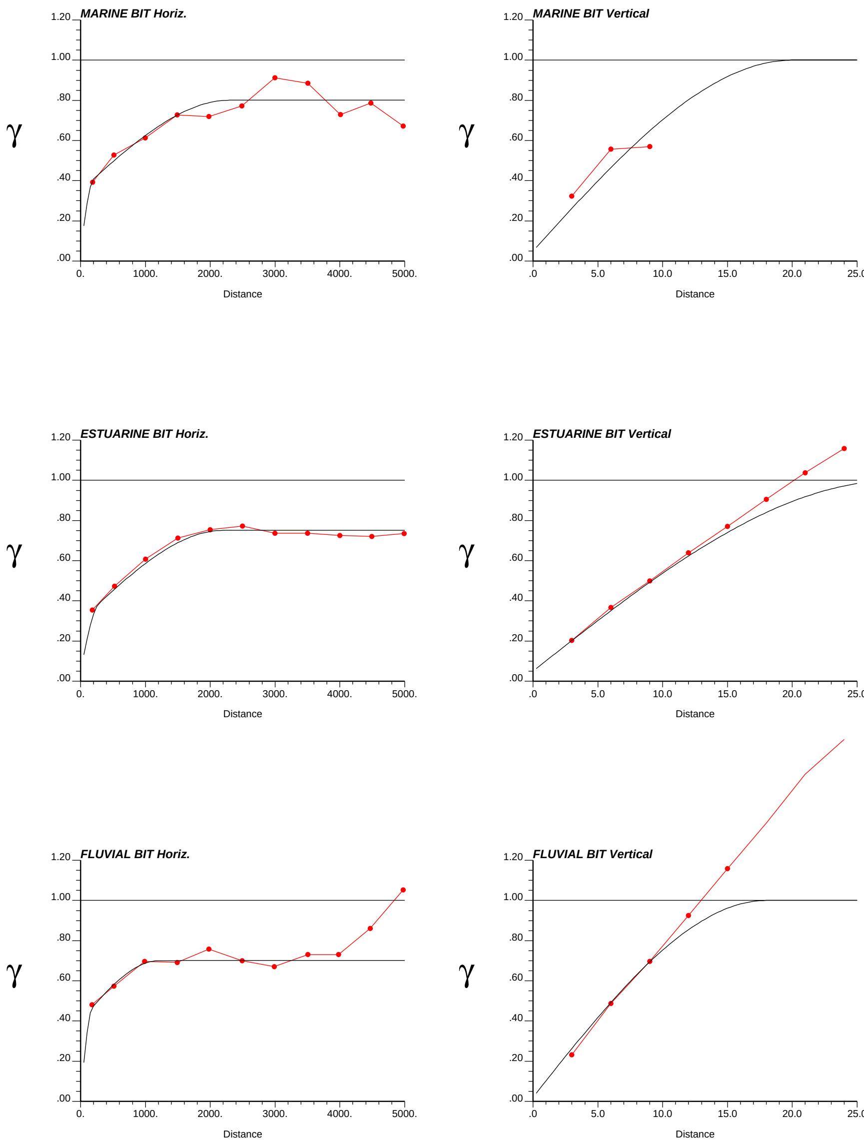

Figure 6: Experimental variograms and models for bitumen grade in the three depositional environments.

model.

The following steps summarize the procedure:

- Transform the data to normal scores of a standard Gaussian distribution (mean=0.0 and variance=1.0) as illustrated on **Figure 4**. This is done within each facies of interest;
- Calculate and model the variogram of the normal scores for each facies (**Figure 6**);
- Perform multi-Gaussian kriging, that is, calculate the kriging estimate and variance

of the normal score transform of the block grade value, y^*_{SK}, σ^2_{SK}, using the normal score transform values of the samples within a search neighborhood, $y_1, ..., y_4$ in the illustration (**Figure 3**, bottom right);

- Under the multi-Gaussian assumption, the shape of the conditional distribution is normal, and fully defined by its mean and variance (**Figure 3**, middle right);

- The estimate can be obtained by numerical integration, that is, several percentiles are back-transformed and then the values are averaged in the original units of the variable (**Figure 5**).

Summarizing Uncertainty

Uncertainty in the block grade is summarized by calculating the width defined by the bounds of a confidence interval at a given significance level. This is done by back-transforming an upper and lower limit from the Gaussian distribution of uncertainty obtained by multi-Gaussian kriging, and calculating the width of that interval in the original units of the variable. The mean, as shown previously, can be back-calculated by numerical integration. The steps are summarized next (see also **Figure 3** for a schematic illustration of the entire methodology):

- The values that represent the upper and lower limits of the confidence interval at some significance level (dotted lines) are calculated and back-transformed using the global distribution (**Figure 3**, top);

- The corresponding back-transformed values represent the same quantiles in the local block grade distribution in the original units of the variables (**Figure 3**, middle left); and

- The statistic defined by the ratio of the width of the interval over the expected block grade is calculated. It represents the relative proportion of the confidence interval width with respect to the mean (expected) value (**Figure 3**, bottom left) and gives a summary measure of the uncertainty as a function of the spread of the distribution and its local mean.

Uncertainty Reporting in Resources

Based on the bitumen grades, the blocks can be classified by the confidence in the estimate. The resources within a particular small area are summarized in **Tables 2 and 3** for a 90% and 50% significance level, respectively. These tables also give the average distance to the nearest drill hole and average drill hole density.

The conventional geometric approaches of using the distance to the nearest drill hole and drill hole density at the location of interest only account for the data location without considering the data values. High grade zones are more variable. The proposed method gives a probabilistic assessment of the confidence with consideration of the data values. Differences assessing uncertainty are due to the presence of different facies and the proportional effect (not accounted for by geometric methods).

$\pm P\%$	Percent of Blocks	Bitumen (%)	Distance to Nearest DH (m)	DH Density ($1/km^2$)
0 - 10	0.06	15.5	73	3.24
10 - 20	0.37	14.6	190	2.88
20 - 30	1.08	13.7	225	3.05
30 - 40	2.10	12.5	236	3.23
40 - 50	3.07	11.5	270	3.04
50 - 60	3.47	10.6	324	2.67
60 - 70	3.92	9.5	374	2.55
70 - 80	3.59	8.4	371	2.44
80 - 90	2.79	7.4	358	2.42
90 - 100	2.23	6.5	363	2.53
> 100	6.50	4.6	356	2.68
Unestimated	70.84			

Table 2: Summary of block classification at 90 % probability.

$\pm P\%$	Percent of Blocks	Bitumen (%)	Distance to Nearest DH (m)	DH Density ($1/km^2$)
0 - 10	1.23	14.2	220	2.84
10 - 20	3.89	12.1	248	3.19
20 - 30	5.95	10.2	293	2.92
30 - 40	6.77	8.7	359	2.64
40 - 50	5.34	7.1	360	2.49
50 - 60	3.41	5.6	377	2.38
60 - 70	1.17	4.5	414	2.18
70 - 80	0.65	4.0	390	2.23
80 - 90	0.53	3.8	422	2.50
90 - 100	0.23	3.3	493	2.35
> 100	0.00	3.2	698	0.85
Unestimated	70.84			

Table 3: Summary of block classification at 50 % probability.

Although one expects to have a narrower confidence interval in low grades because of their lower variance, the effect of more drillholes in the high grade early production times generates the opposite result. In the higher grade zones the estimates are more reliable than in the lower grade zones, because of the sample density. It is worth noting that the methodology proposed merges the drillhole density with the proportional effect to give an unbiased estimator and a minimum estimation variance.

Classification of Resources

Classification is done by defining two threshold values for the statistic used to summarize uncertainty in the block grades. The resources obtained through the proposed methodology should be reconciled with the current definition of resources. The advantage of the proposed method is that the classification is repeatable and it is based on uncertainty assessment of the block grades, taking into account the proportional effect and spatial continuity of the variables within specific rock types.

Conclusions

Quantifying uncertainty in resources is a difficult task that must account for the spatial configuration of the data and the presence of the proportional effect. Multi-Gaussian kriging allows the calculation of the distribution of uncertainty accounting for the data configuration and the proportional effect of the variable. Uncertainty can be summarized by calculating the ratio of the confidence interval at some probability level over the expected value at that location. This measure can be used to classify the resources on a repeatable manner. The case study showed that even in high grade zones, where the variance is expected to be higher, the denser sampling reduces the uncertainty, providing a narrower confidence interval. This is reflected in the parameter P, which can be used to classify resources based on their confidence. The method compares well with traditional geometric measures of quality, but gives better results since it considers the correlation between the data and the local mean.

References

[1] Canadian Institute of Mining, Metallurgy and Petroleum. CIM standards on mineral resources and reserves - definitions and guidelines. Prepared by CIM Standing Committee on Reserve Definitions, August 2000.

[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] I. M. Glacken. Change of support by direct conditional block simulation. Master's thesis, Stanford University, Stanford, CA, 1996.

[5] P. Goovaerts. Comparative performance of indicator algorithms for modeling conditional probability distribution functions. *Math Geology*, 26(3):385–410, 1994.

[6] P. Goovaerts. *Geostatistics for Natural Resources Evaluation*. Oxford University Press, New York, 1997.

[7] E. H. Isaaks and R. M. Srivastava. *An Introduction to Applied Geostatistics*. Oxford University Press, New York, 1989.

[8] M. Jackson and A. Marechal. Recoverable reserves estimated by disjunctive kriging: A case study. In *Proceedings of the 16th International APCOM Symposium*, pages 240–249, Tucson, AZ, October 1979. Society of Mining Engineers.

[9] Joint Ore Reserves Committee (The JORC Code). Australasian code for reporting of mineral resources and ore reserves. Prepared by The Australasian Institute of Mining and Metallurgy, Australian Institute of Geoscientists and Minerals Council of Australia, September 1999.

[10] A. G. Journel. The lognormal approach to predicting local distributions of selective mining unit grades. *Mathematical Geology*, 12(4):285–303, 1980.

[11] A. G. Journel. The indicator approach to estimation of spatial distributions. In *Proceedings of the 17th International APCOM Symposium*, pages 793–806. Society of Mining Engineers, 1982.

[12] A. G. Journel. Nonparametric estimation of spatial distribution. *Mathematical Geology*, 15:445–468, 1983.

[13] A. G. Journel. Recoverable reserves estimation - the geostatistical approach. *Mining Engineering*, pages 563–568, June 1985.

[14] A. G. Journel and C. J. Huijbregts. *Mining Geostatistics*. Academic Press, New York, 1978.

[15] A. Marechal. The practice of transfer functions: Numerical methods and their applications. In Guarascio et al., editors, *Advanced Geostatistics in the Mining Industry*, pages 253–276, Dordrecht, Holland, 1976. Reidel.

[16] G. Matheron. Forecasting block grade distributions: The tranfer function. In Guarascio et al., editors, *Advanced Geostatistics in the Mining Industry*, pages 237–251, Dordrecht, Holland, 1976. Reidel.

[17] G. Matheron. A simple substitute for conditional expectation: the disjunctive kriging. In Guarascio et al., editors, *Advanced Geostatistics in the Mining Industry*, pages 221–236, Dordrecht, Holland, 1976. Reidel.

[18] National Instrument 43-101. Standards of disclosure for mineral projects. Canadian Securities Administrators, February 2001.

[19] J. M. Rendu. Normal and lognormal estimation. *Mathematical Geology*, 11(4):407–422, 1979.

[20] J. Rivoirard. *Introduction to Disjunctive Kriging and Non-Linear Geostatistics*. Oxford University Press, New York, 1994.

[21] South African Mineral Resource Committee (SAMREC). South African code for reporting of mineral resources and mineral reserves (the SAMREC Code). Prepared under the auspices of the South African Institute of Mining and Mettalurgy, March 2000.

[22] P. Switzer and H. M. Parker. The problem of ore versus waste discrimination for individual blocks: The lognormal model. In Guarascio et al., editors, *Advanced Geostatistics in the Mining Industry*, pages 203–218, Dordrecht, Holland, 1976. Reidel.

[23] G. Verly. The multigaussian approach and its applications to the estimation of local reserves. *Mathematical Geology*, 15(2):259–286, 1983.

[24] G. Verly. The block distribution given a point multivariate normal distribution. In G. Verly et al., editors, *Geostatistics for natural resources characterization*, volume 1, pages 495–515. Reidel, Dordrecht, Holland, 1984.