On quantifying the space of uncertainty of stochastic simulations

S.Mehrdad Heidari*and Gary Froyland†

March 15, 2011

Abstract

Distance computation as a technique to measure dissimilarity/similarity between images, objects, models and so on has received attention in recent years. This paper presents a formal measure of dissimilarity for generated realizations by adapting the Kantorovich metric to the geostatistics context. We propose a new methodology for mapping the space of uncertainty by a distance function that is based upon a physically meaningful notion of dissimilarity between pairs of realizations. We are able to quantify the dissimilarity of different realizations, and to use this information for modeling and visualizing the space of uncertainty. As a consequence, any question about sampling the space of uncertainty, evaluation of simulation algorithms, relative locations between the random realizations, how far the realizations are from local accuracy (Kriging model) and so on, can be answered by this approach. Furthermore, we can use our methodology to optimally subsample a large collection of realizations, and quantify how well this high-quality subsample represents the overall uncertainty of the collection.
Keywords : Uncertainty quantification, Mapping the space of uncertainty, Geostatistical simulation, Realization reduction, Kantorovich distance.

1 Introduction

Major sources of financial risk for mining projects include geological uncertainty and uncertainty in future commodity prices, costs, demand levels, interest rates and so on. Geological uncertainty is more difficult to model as there are complicated spatial considerations which are not present in the other sources of uncertainty. Stochastic simulations are now the common approach to assessing geological uncertainty, and one of the most common practical methods of producing realizations is sequential Gaussian simulation (SGS) (Deutsch and Journel 1998). An SGS algorithm can create multiple realizations that honor the original

*School of Mining Engineering,The University of New South Wales, Sydney NSW 2031, Australia
†School of Mathematics and Statistics,The University of New South Wales, Sydney NSW 2031, Australia

histogram and covariance matrix. By applying a transfer function to this set of realizations the space of the response uncertainty (i.e. the distribution of the output value of the transfer functions) can be made. This technique is used for a variety of purposes in the mining industry such as resource/reserve estimation (Dowd and David 1976), feasibility study and risk assessment (Dowd 1994), open pit design and production scheduling (Dimitrakopoulos 1998). In all of these applications, the multiple realizations are randomly sampled without any idea about underlying probability space (commonly named the space of uncertainty). The nonlinearity of transfer functions, and equal probability of the realizations are the most common reasons for not quantifying the space of uncertainty during the sampling. To quote P.Goovaerts (Goovaerts 1999), "In summary, contrasting with the increasing use of stochastic simulation in risk analysis, it appears that little attention has been paid to the definition of the space of uncertainty, and related issues such as the equiprobability of realizations, the equivalence of spaces of uncertainty generated by different algorithms, and the number of realizations required to sample this space".
It is obvious that two realizations can be significantly different in ways that cannot be captured by descriptive geostatistics. By visually comparing the color realization images (if they are 2D) we can easily see uncaptured spatial differences, which may affect the results of transferring functions. Although there are many different realizations, some of them are visually more similar to each other than others. For example, figures 1 and 2 present images of four realizations selected from 450 generated realizations of the Walker lake data set (Isaaks and Srivastava 1989). By visually comparing these images, we conclude that the images of figure 1 are closer to each other than the images of figure 2.

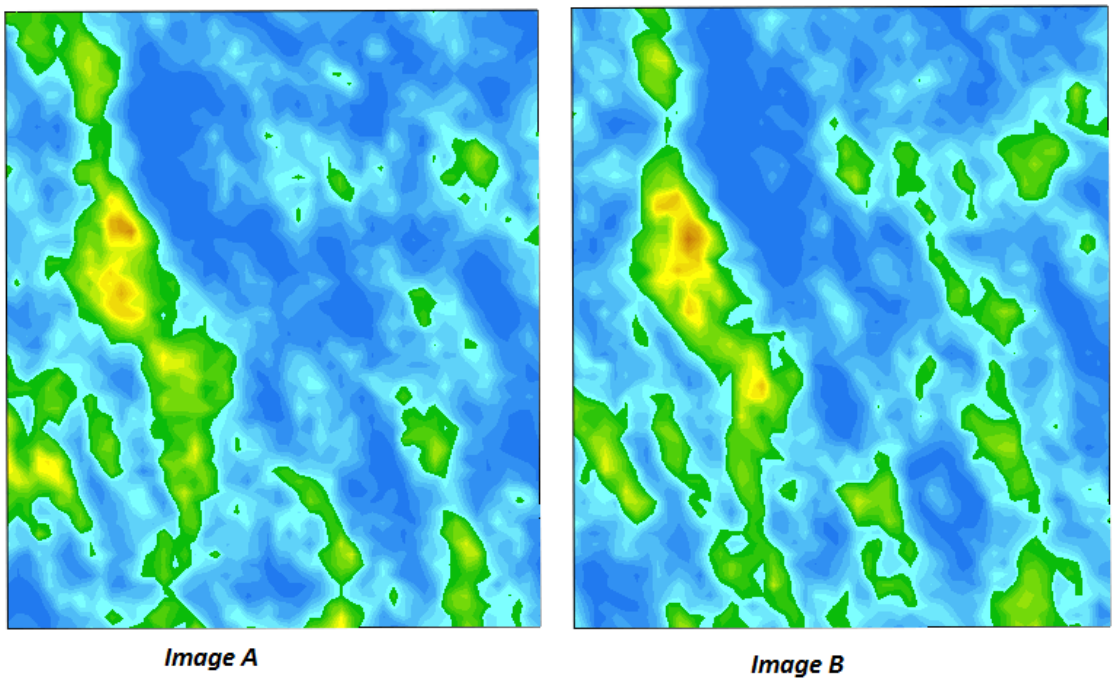

Figure 1: Selected realizations of Walker lake data set that are very close to each other.

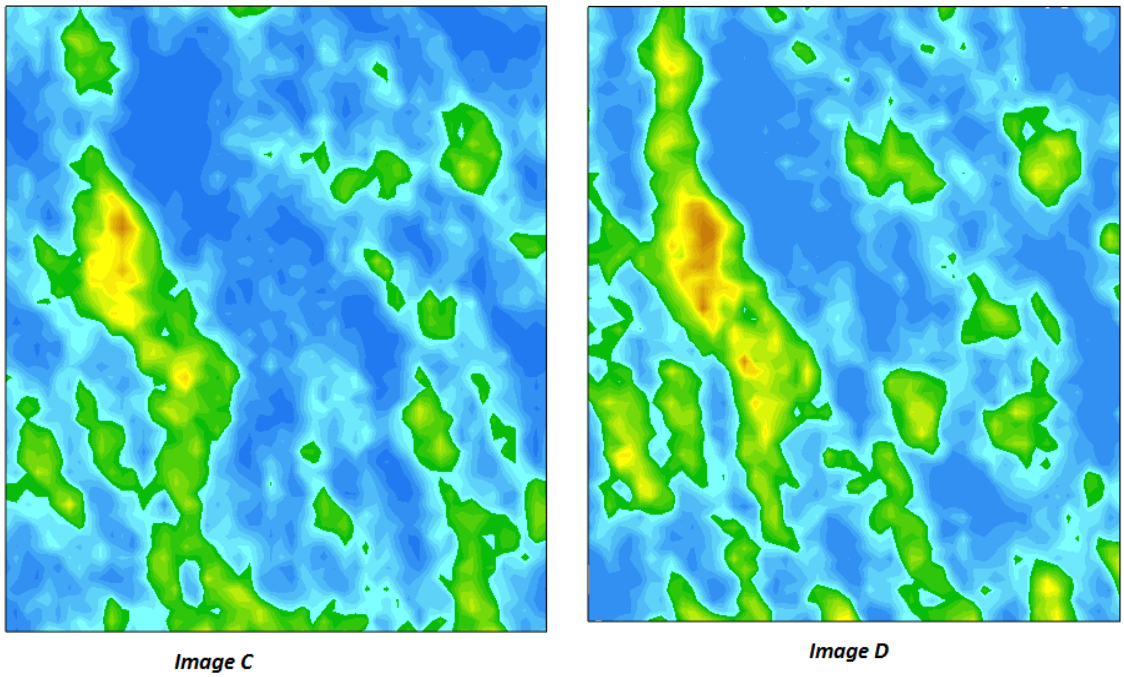

Figure 2: Selected realizations of Walker lake data set that are far from each other.

Quantifying this spatial dissimilarity can be a powerful technique to assess and map the space of uncertainty. Such quantification of the space of uncertainty makes it possible to compare the impact of changing the geostatistical parameters or even simulation algorithms on the space of uncertainty. Our methodology has the potential to consistently compare the output of different stochastic simulation algorithms such as SGS, sequential indicator simulation, probability field simulation, and so on. Previously, attempts at such comparison have been based on transfer functions responses (Gotway and Rutherford 1993, Qureshi and Dimitrakopoulos 2005), which lose a significant amount of the underlying structure of the space of uncertainty.

Furthermore, if we put any deterministic geological reserve estimation, produced for example, by Kriging, inside the space of uncertainty, the method can easily reveal how dissimilar other realizations are to the estimated model. In other words, measuring how close global accuracies (different realizations) are to the local accuracy (estimation), is now possible.

Moreover, the mining processes such as mine optimization, open pit design, long term scheduling, etc, are only able to handle relatively modest numbers of realizations. It is difficult to say how many realizations are required to achieve a prescribed level of accuracy based on a very large number of possible realizations. This method has the ability to construct a collection of realizations so that the overall uncertainty is captured in a way prescribe by the user. For example, in open pit design, because of smoothing effects, the NPV outcome from an ordinary Kriging model (estimation) is usually higher than the mean of the distribution of NPVs of conditionally simulated realizations (Dimitrakopoulos et al. 2002). This means

that selecting realizations that are not close to the Kriging model is an effective way to indicate the range of possible NPV estimates, and selecting realizations that are far from one another indicates the range of geological uncertainty affecting open pit mine design.

To quantify this spatial dissimilarity, we present the Kantorovich distance[1]as an important and intuitive measure of dissimilarity. This distance may be used to detect and identify the structural relationships between realizations and has obvious application to clustering, selecting representative realizations, and visualizing uncertainty.
In this paper, we focus on the concept and methodology of our approach rather than the applications. In a forthcoming paper we will present applications based on the concept introduced here. An outline of our paper is as follows. In section 2 we introduce the concepts of dissimilarity between realizations and specialize the application of Kantorovich metric to the geological setting. This describes a computational approach for evaluating the metric on pairs of realizations and assigns a geologically relevant "distance" between realizations. In Section 3 we modify the optimization problem constructed in Section 2 to determine the subcollection of realizations that best represent the possible outcomes. In Section 4 we apply our methodology to unconditional and conditional cases. We conclude with discussion of future work.

2 Definition of dissimilarity by Kantorovich distance

The concept of dissimilarity between multi-point geostatistical realizations has been exploited in the context of evaluating oil/gas reservoirs by Suzuki and Caers (Suzuki and Caers 2008), Celine Scheidt and Caers (Scheidt and Caers 2009). This method has been used for oil reservoirs to model the uncertainty of channel facies, or patterns modeling by them. However, different geostatistical modeling techniques were used, and the distance function of the method are not applicable in mining applications.

A notion of dissimilarity or "distance" between realizations is a key ingredient in quantifying the space of uncertainty. This notion of dissimilarity or distance must be geologically meaningful, and be able to explain what causes the dissimilarity between realizations. It is obvious that grades and their locations are the sole source of dissimilarity between realizations, therefore the "distance" has to be defined based on these. We now begin to describe the necessary background constructions.

Throughout, we assume that our block model contains N blocks. We denote the three-dimensional coordinates of the center of block i by c_i, $i = 1, \ldots, N$, and set $d_{ij} = \|c_i - c_j\|_2$ to be the Euclidean distance between the centers of blocks i and j.

Now consider a particular pair of realizations r and s. We assume that there is only one material of interest (for example, copper), although the techniques described here are trivially extendable to multiple attributes. For each block $i = 1, \ldots, N$, realization r will assign a mass of material, which we denote by m_i^r. In the same way, from realization s we have $m_i^s, i = 1, \ldots, N$. We assume that $\sum_{i=1}^N m_i^r = \sum_{i=1}^N m_i^s := M$; this is not unreasonable, in view of the law of large numbers and the fact that N is typically large.

[1]In some context, Kantorovich distance may be known as the Earth Mover's Distance (EMD).

We now begin to construct a formal measure of dissimilarity for generated realizations, adapting the Kantorovich metric to the geostatistics context. The Kantorovich metric has already been used successfully in other applications such as image processing, data mining and bioinformatics (Vershik 2006). For the interested reader, further background on the Kantorovich metric may be found in these references (Deng and Du 2009, Villani 2003).

2.1 The distance computation

In this section we describe how to set up a simple optimization problem to calculate distances between all pairs of realizations (inter-point distances) to construct the distance matrix or the space of uncertainty.

Our notion of dissimilarity for realizations r and s will be based upon the "work" required to "transform" realization r into realization s. Both realizations contain approximately the same total mass M, but the spatial distribution of mass is different. To "transform" realization r into realization s, mass (the metal content) must be moved between individual blocks. The total "work" for moving mass will be proportional both to the *mass moved* and the *distance moved*. We define variables $f_{ij}^{rs} \geq 0$, $1, \leq i, j \leq N$, to be the mass in block i that is moved to block j; the total work for this movement will be $f_{ij}^{rs} d_{ij}$.

We thus have the following minimization problem, which is a linear program and is easily solved using standard techniques[2].

$$\min_{f^{rs}} \quad \sum_{i,j=1}^{N} d_{ij} f_{ij}^{rs} \tag{1}$$

$$\text{subject to} \quad \sum_{j=1}^{N} f_{ij}^{rs} = m_i^r, \qquad 1 \leq i \leq N \tag{2}$$

$$\sum_{i=1}^{N} f_{ij}^{rs} = m_j^s, \qquad 1 \leq j \leq N \tag{3}$$

$$f_{ij}^{rs} \geq 0, \qquad 1 \leq i, j \leq N \tag{4}$$

One way to think of this transportation problem is as follows. If blocks i and j are spatially distant, the corresponding distance penalty per unit of mass moved, d_{ij}, will be large and such a move will be discouraged. However, if blocks i and j are spatially close, moving mass from i in realization s to j in realization r is comparatively attractive. The linear program (1)-(4) finds the minimal amount of *work* (mass moved multiplied by distance moved) to turn realization s into realization r. This minimum amount of work defines the distance between realizations r and s.

[2]This problem is a *transportation problem* and is easily solved using the simplex method or other specialized methods.

Definition 2.1. If $\hat{f}^{rs}$ is the minimizing array of mass flows from (1), then we define the dissimilarity or distance between realizations r and s to be

$$D_{rs} = \sum_{i,j=1}^{N} d_{ij} \hat{f}_{ij}^{rs}. \tag{5}$$

Remark 1. Note that D_{rs} satisfies all those properties expected of a notion of distance:

1. nonnegativity and zero implies identity: $D_{rs} \geq 0$ and $D_{rs} = 0 \Leftrightarrow$ realization r and realization s have identical mass distributions at the block level,
2. symmetry: $D_{rs} = D_{sr}$,
3. triangle inequality: $D_{rs} \leq D_{rt} + D_{ts}$ for all t

Thus, for a set of S realizations, the square symmetric distance matrix $\mathcal{S} \times \mathcal{S}$ (dissimilarity distance matrix) with zero diagonal elements $D_{ii} = 0$, is constructed. For instance, by calculating the pairwise distances D_{sr} between the given four images in figure 1 and 2, the following 4×4 dissimilarity distance matrix can be constructed.

$$\begin{array}{c} \\ ImageA \\ ImageB \\ ImageC \\ ImageD \end{array} \begin{array}{c} \begin{array}{cccc} ImageA & ImageB & ImageC & ImageD \end{array} \\ \begin{pmatrix} 0 & 4,548,600 & 6,798,720 & 5,564,880 \\ 4,548,600 & 0 & 6,668,630 & 4,936,460 \\ 6,798,720 & 6,668,630 & 0 & 9,995,100 \\ 5,564,880 & 4,936,460 & 9,995,100 & 0 \end{pmatrix} \end{array}$$

Furthermore, note that D_{rs} has the units of mass $\times$ Euclidean distance, so we may interpret the quantity D_{rs}/M as the average distance that 1 mass unit of material is moved to transform realization r into realization s.

$$D_{4\times4}/M = \begin{pmatrix} 0 & 4.84 & 7.24 & 5.92 \\ 4.84 & 0 & 7.10 & 5.26 \\ 7.24 & 7.10 & 0 & 10.64 \\ 5.92 & 5.26 & 10.64 & 0 \end{pmatrix}$$

As can be seen in the matrix above, the Image A and B (figure 1) are at minimum distance from each other (4.84 m) while images C and D (figure 2) are at maximum distance (10.64 m).

3 Realization reduction

We now suppose that we have constructed a large number of realizations numbered $1, \ldots, \mathcal{S}$. With increasing computer performance, it is possible to construct for example $\mathcal{S} \approx 1000$ realizations for block models of size $N \approx 10^5$ in 48 hours on a desktop PC using techniques such as SGS (it would be faster if direct block Kriging is used). We will assume that the collection $\{1, \ldots, \mathcal{S}\}$ well samples the distribution of all possible realizations and use this collection as the benchmark against which we compare subcollections.

Remark 2. We remark that in practice, N will typically be larger than $\mathcal{S}$ so this assumption is in most cases false.

Nevertheless, for many applications of realizations, many more realizations can be created than used, whether in an optimization process (Ramazan and Dimitrakopoulos 2004, Menabde. et al. 2004, Froyland et al. 2004, Boland et al. 2008, Tarhan et al. 2009, Newman et al. 2010), or as part of a small collection of possible geological outcomes interrogated by engineers or geologists. It is therefore important to choose $\mathcal{S}$ as large as is practicable, and then reproduce the variety contained in these $\mathcal{S}$ realizations as best as possible with a significantly smaller number $S \ll \mathcal{S}$ realizations, for use in optimization or other applications.

Remark 3. Sampling (realization reduction) from an original collection $\mathcal{S}$ can be done based on different distance criteria such as minimum or maximum distance between subcollection (sampled realizations) and original collection $\mathcal{S}$, or any other specified distance criteria which may be defined according to the purpose of sampling.

In Section 2 we calculated D_{rs} for every pair $1 \leq r < s \leq \mathcal{S}$. Our task now is to best represent the collection $\{1, \ldots, \mathcal{S}\}$ by a smaller subcollection $\{s_1, \ldots, s_S\}$, where $s_k \in \{1, \ldots, \mathcal{S}\}$, $k = 1, \ldots, S$. For simplicity, we assume that each realization in the original collection $\{1, \ldots, \mathcal{S}\}$ is equiprobable (probability $1/\mathcal{S}$) although the method proposed here is trivially extendable to unequal probabilities.

We will again use the idea of "transformation" to transform the larger set of $\mathcal{S}$ realizations to smaller set of S realizations with minimal "work". Let F_{rs} denote the probability flow from realization r to realization s, $1 \leq r, s \leq \mathcal{S}$. Because we are reducing $\mathcal{S}$ realizations to S realizations, there will be at most S special realizations s for which F_{rs} may be greater than zero; that is, only the S special realizations in the reduced collection are allowed to receive a positive probability flow. "Work" now will be a product of the probability flow F_{rs} and the distance D_{rs}.

The special realizations will be chosen using binary variables x_s, $s = 1, \ldots, \mathcal{S}$, with $x_s = 1$ signifying that $s \in \{s_1, \ldots, s_S\}$ (realization s is selected for the subcollection) and $x_s = 0$ otherwise. We do not insist that each realization $s_k, k = 1, \ldots, S$ have an equal probability, but that the weightings of the chosen realizations can vary. The weight for realization s will be denoted by w_s. We state the following mixed integer linear program (MILP) and then describe the effect of the various constraints.

$$z_S = \min_{F,x,w} \quad \sum_{r,s=1}^{\mathcal{S}} D_{rs} F_{rs} \tag{6}$$

$$\text{subject to} \quad \sum_{s=1}^{\mathcal{S}} F_{rs} = 1/\mathcal{S}, \qquad 1 \le r \le \mathcal{S} \tag{7}$$

$$\sum_{r=1}^{\mathcal{S}} F_{rs} = w_s, \qquad 1 \le s \le \mathcal{S} \tag{8}$$

$$\sum_{s=1}^{\mathcal{S}} x_s = S \tag{9}$$

$$\sum_{s=1}^{\mathcal{S}} w_s = 1 \tag{10}$$

$$w_s \le x_s, \qquad 1 \le s \le \mathcal{S} \tag{11}$$

$$F_{rs} \ge 0, \qquad 1 \le r, s \le \mathcal{S} \tag{12}$$

$$w_s \ge 0, x_s \in \{0, 1\}, \qquad 1 \le s \le \mathcal{S} \tag{13}$$

Equations (7) and (8) ensure that the probability flow *out of* each realization is $1/\mathcal{S}$ and that the probability flow *into* realization s is w_s, respectively. Equality (9) ensures that exactly S realizations are chosen for the subsampled collection. Equality (10) guarantees that the sum of the weights w_s in the chosen subcollection is 1. Inequality (11) forces the weight assigned to realization s to be zero unless realization s is selected for the subsample. As in the transportation problem (1)-(4), the problem (6)-(13) transports a probability of $1/\mathcal{S}$ from the original very large collection of $\mathcal{S}$ realizations onto a smaller collection of S realizations. The penalty for the probability flow from realization r to realization s is proportional to both the size of the flow F_{rs} and the distance between the realizations D_{rs}. Thus, if there is a set of several realizations with mutually small distances D_{rs}, it is likely that this set will be replaced by one realization from the set with all the probability of that set flowing to the one realization that now "represents" that set. Having solved the MILP above, one has the minimizing arrays $\hat{F}, \hat{x}$, and $\hat{w}$. The selected subcollection consists of those s with $\hat{x}_s = 1$ and the corresponding weights are given by the vector $\hat{w}_s$, which has at most S positive entries. The value z_S has the units of D_{rs} (mass×Euclidean distance) as the F_{rs} are dimensionless. Thus, z_S has the interpretation of "work" required transform all of the $\mathcal{S}$ realizations into $S \ll \mathcal{S}$ realizations. The quantity z_S/M has the units of Euclidean distance and thus may be interpreted as the average distance over which a single unit of mass is moved to effect this transformation.

4 Numerical results

In order to better present the effects of spatial continuities (variogram ranges) and variability (variogram sills) on the space of uncertainty we start with unconditional sequential Gaussian realizations. We generate three series of 3D-realizations with parameters shown in table 1. For each set of realizations, the variograms are isotropic spherical, the search neighborhood is spherical, and there are no nugget effects. The block models have a block size $25 \times 25 \times 12.5$ m $(x, y, z$ direction). Each of the block models contains 2805 blocks.

Table 1: Parameters for variogram models

Models	Number of realizations	Nugget	Sill	Range(E-W)	Range(N-S)	Range(Z)
Unconditional simulations						
Series 1	100	0	0.05	500	500	500
Series 2	200	0	0.68	250	250	250
Series 3	100	0	1	100	100	100
Conditional simulations						
Normal score model	400	0.32	0.68	100	115	77

We first calculate the pairwise distances D_{rs} to construct 2805×2805 distance matrix for series 2 in table 1 using ILOG CPLEX V12.1.

4.1 Visualizing the space of uncertainty

Each realizations consists of a 2805×3 array of three-dimensional block center coordinates and a 2805×1 vector of masses for each block. Visualizing these high dimensional objects and the distances D_{rs} between them is difficult directly. Given the $\mathcal{S} \times \mathcal{S}$ array of distances D_{rs}, the technique of *multidimensional scaling* (MDS) (Cox and Cox 2000) attempts to find $\mathcal{S}$ points $\{y_1, \ldots, y_{\mathcal{S}}\}$ in $\mathbb{R}^n$ for some specified n such that the distance array $\tilde{D}_{rs} := \|y_r - y_s\|_2$ approximates the array D_{rs}. For visualization purposes, we choose $n = 2$. Figure 3 shows a multidimensional scaling embedding of the 200 (series 2) realizations in $\mathbb{R}^2$ so that the interpoint distances in $\mathbb{R}^2$ approximate the pairwise realization distances D_{rs}.

4.2 Variation in approximation accuracy with S

Having computed the distance array D_{rs} we solve the realization reduction problem (6)-(13) to determine the subset of S weighted realizations that are the best representatives of 200 realizations. In order to know how large S should be, and how well an increasing number of optimum points can represent the full collection of realizations, we solve (6)-(13) for different values of S ranging from 1 to 200. In practice, the number S may be selected based on time or other resource costs.

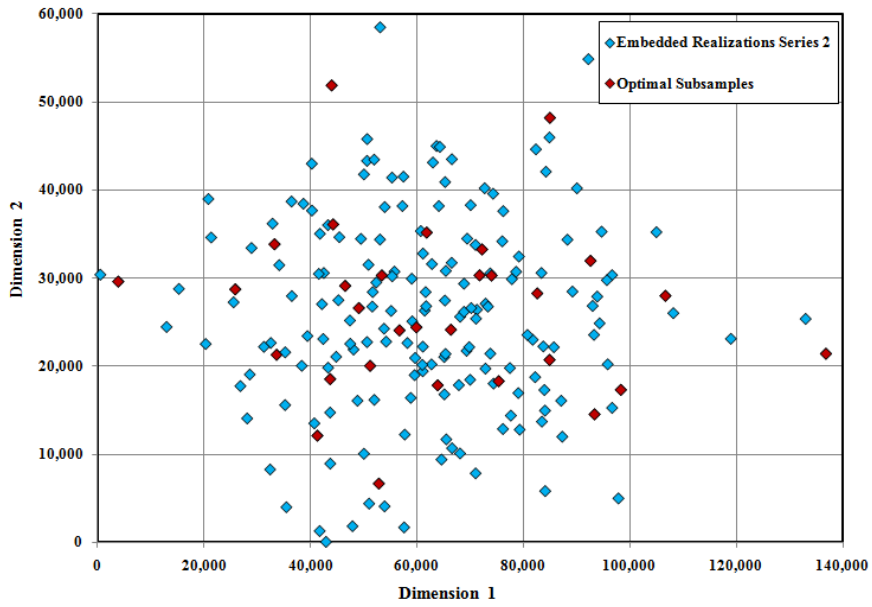

Figure 3: Multidimensional scaling embedding of 200 realizations (series 2 of table 1) in $\mathbb{R}^2$ (blue diamonds) and the 30 realizations that form the best subsample of size 30 (red diamonds).

Following (Dupacova et al. 2003), we report z_S relative to z_1, the latter representing the "base" distance between the best deterministic approximation of the collection of $\mathcal{S}$ realizations. Thus, z_S/z_1 represents the distance between the optimally subsampled S realizations and the full set of $\mathcal{S}$ realizations *relative to* the distance between the best single realization (deterministic approximation) and the full set of $\mathcal{S}$ realizations. For brevity, we call the quantity $I_S = (1 - z_S/z_1) \times 100\%$ the *relative accuracy* of the optimal S-subcollection. Clearly $I_1 = 0$ ($S = 1$ provides no relative accuracy), and $I_{\mathcal{S}} = 100$ ($S = \mathcal{S}$ provides a 100% relative accuracy).

The result of these calculations for thirty different values of S is presented in figure 4. Figure 4 shows, for example that using just $S = 30$ realizations (15% of all 200 realizations), one obtains a relative accuracy of around 52%. The red diamonds in figure 3 give a visual representation via multidimensional scaling of the relative position in $\mathbb{R}^2$ of the optimal 30 realizations. Note that the selected (red) points are distributed throughout the entire set of (blue) points to well-sample the point distribution. The relative weights assigned to the red points is not shown in figure 3, but typically those points at the periphery of the point set have the lowest possible nonzero weight of $1/\mathcal{S} = 1/200$. Points closer to the center will have a weight above the average of $1/S = 1/30$. We will demonstrate these effects in a real case study in Section 4.5.

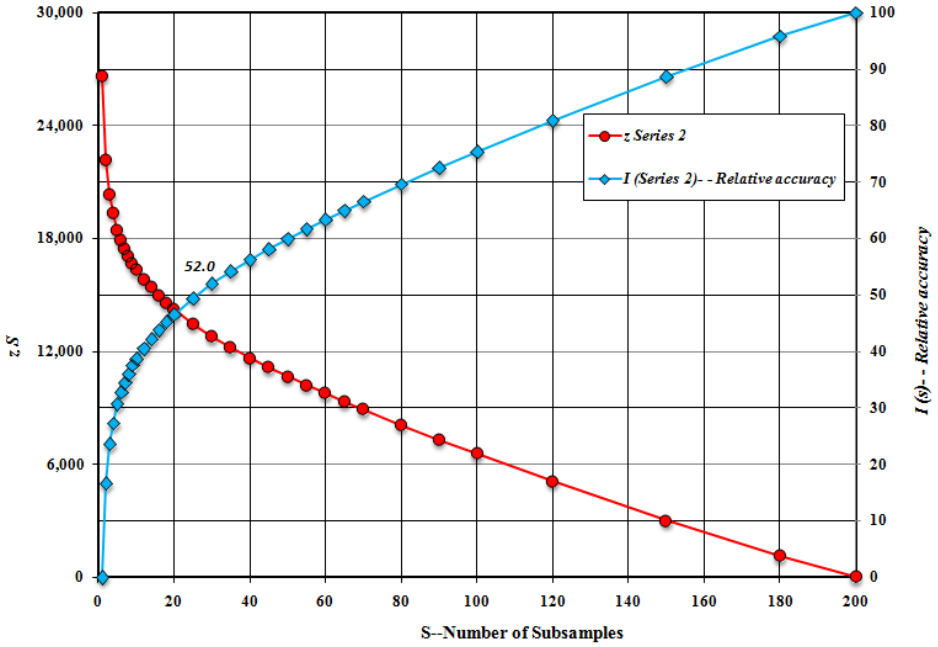

Figure 4: Distance z_S between the S-subcollection and the full collection of $\mathcal{S}$ simulations (red) and relative accuracy (blue).

Remark 4. If realizations (points) are clustered rather than being randomly distributed in the space, the approximate MDS representation of the realizations in $\mathbb{R}^2$ opens the possibility of using Euclidean-based subsampling or clustering techniques such as K-means (MacKay 2003) to select S "cluster centers", which attempt to minimize the total distance from center points to non-center points. These cluster centers might be used as the selected subsamples. Increasing the MDS embedding dimension from $n = 2$ to a larger value will increase the inter-point distance approximation of the distances D_{rs}, and if such an MDS-based approach to subsample selection is taken, we suggest using larger values of n. Nevertheless, we advocate MDS for visualization purposes only and solving (6) – (13) to determine the optimal subsamples. Our reasons for this are: (i) one finds optimal subsamples based on the original distances D_{rs} and not approximations, (ii) the objective minimized has a physical meaning in terms of "work" done, (iii) one has the flexibility of optimizing over the weights assigned to the subsamples, (iv) optimality (or solution quality) is guaranteed and the process is reproducible, and (v) in all our cases, as shown in figure 3, the realizations are not clustered.

4.3 The impact of spatial continuity and variability

We use two additional series (Series 1 and 3 in table 1) to investigate the effects of spatial continuity (variogram range) and variability (variogram sill) on the ability to subsample.

Repeating the computations used for figure 4 for series 1 and 3 we obtain the results shown in figure 5.

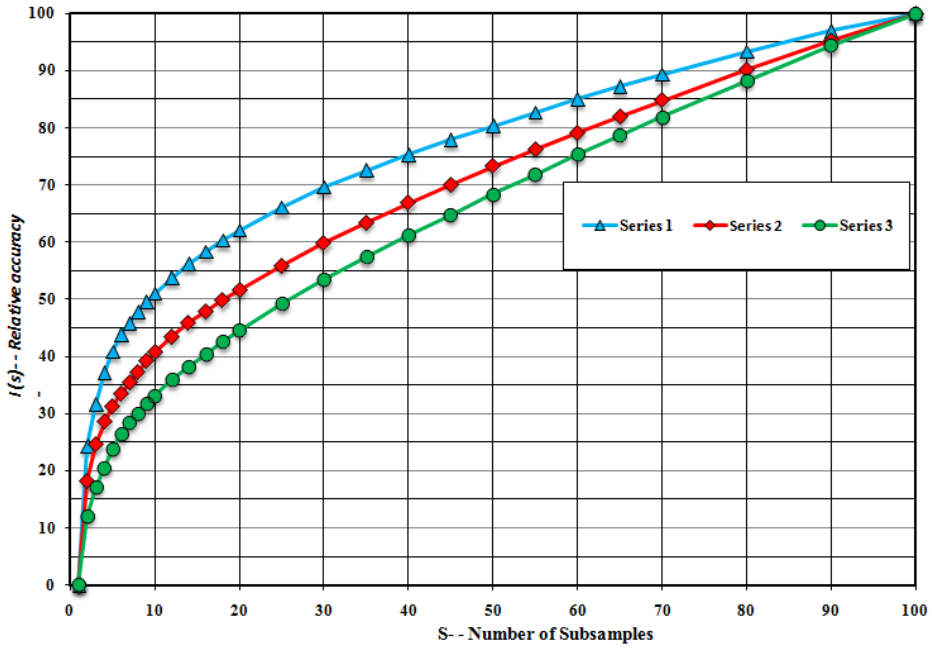

Figure 5: Illustration of the impact of spatial continuity (range) and variation (sill) on the relative accuracy of optimal subsamples.

From figure 5 we see that for a fixed S, we can achieve a greater relative accuracy for Series 1, which has a lower sill and a larger range. *Increasing sill or decreasing range leads to a reduction in relative accuracy.* These results are exactly what one would expect; longer ranges of correlations (greater spatial continuity) and lower sills (lower variability) will tend to make the collection of $\mathcal{S}$ realizations more structured and less random. This greater structure, encoded via the distances D_{rs} can be exploited by our optimal subsampling procedure. Putting this another way, *simulations with larger sill and smaller range will require a larger number of subsamples to achieve the same relative accuracy.*

4.4 Optimal vs. random subsampling

We now compare the relative accuracy of S optimal subsamples with several random selections of S subsamples to demonstrate how much additional relative accuracy can be obtained

by solving the MILP (6)-(13). Using Series 1 from table 1, in figure 6 for each $S = 5, 10, 20,$ and 30, we randomly select 20 collections of S realizations from the total of $\mathcal{S} = 100$.

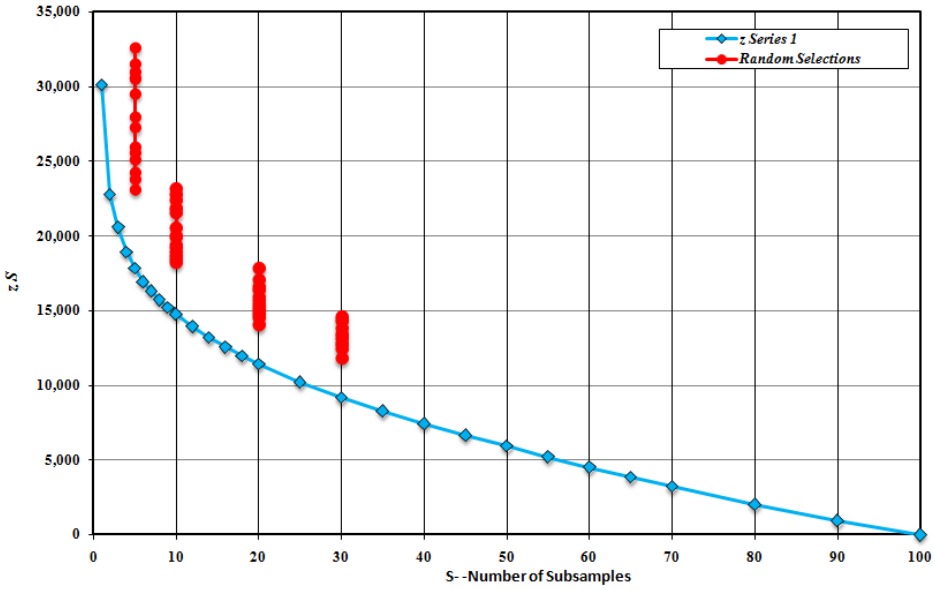

Figure 6: Comparison of distances z_S between the optimal subsamples and randomly selected subsamples.

Figure 6 clearly demonstrates the distances z_S that can be achieved by optimally subsampling rather than randomly sampling. The variability of random sampling can also be very high. For example, suppose rather than generating 100 realizations and sampling the best 5, one merely generated the first 5 realizations of the 100 and then stopped. The variability in distance z_S of the first 5 ranges from 23,128 to 32,585.

Thus, we argue it is worth investing more time to create a larger number of realizations to attempt to better represent the true resource uncertainty and to then subsample as best possible from that larger collection of realizations. Indeed, in this example, the optimal 5 realizations would have a better relative accuracy than some of the collections of 20 random realizations!

4.5 Application to conditional simulations

Having illustrated some of the basic properties and advantages of our dissimilarity quantification and subsampling approach on unconditional simulations, we now turn to a realistic set of conditional simulations of a porphyry copper deposit.

Although geological data is available for all of the lithological domains, for simplicity we focus on a single estimation domain, supergene, which has 40 Mt of measured and indicated resources at a copper grade of 0.8%. Within the supergene domain, the copper grades are

given over a set of 414 six meter (composites) core length from the drill holes. We calculated experimental variogram maps calculated using these composite samples for both the total copper grades and normal scores; the variogram map sets were calculated in the X,Y,Z orthogonal directions. One-stage spherical models were used to model the experimental variograms. This is equivalent to the variogram model parameters listed in table 1.

The variogram results revealed that the models have high nugget effects, which confirms high small scale variations over the supergene zone. Because of this we expect to have high grade variation in the realizations. Three types of block models were estimated by using a block size $25 \times 25 \times 12.5$ m $(x, y, z$ direction). The first type is estimated by Kriging, the second type by generating $\mathcal{S} = 400$ sequential Gaussian conditional simulations and the third type by averaging all 400 simulated models (we call this *E-type*). Each of these block models contains 2805 blocks that cover the entire supergene zone.

We now compute the distances D_{rs}, $1 \leq r < s \leq 402$ using (1)-(4) (to the 400 realizations, we add the Kriged and E-type block models and compute distances between these models as well). Interestingly we note that because of smoothing effects, the Kriging and E-type models are the closest pair of models, that is $min_{1\leq r,s\leq 402}\ D_{r,s} = D_{r^*,s^*}$ where r^* = Kriging and s^* =E-type. We then solve (6) – (13) with $S = 30$, which we find produces a relative accuracy of 25.64% (figure 7).

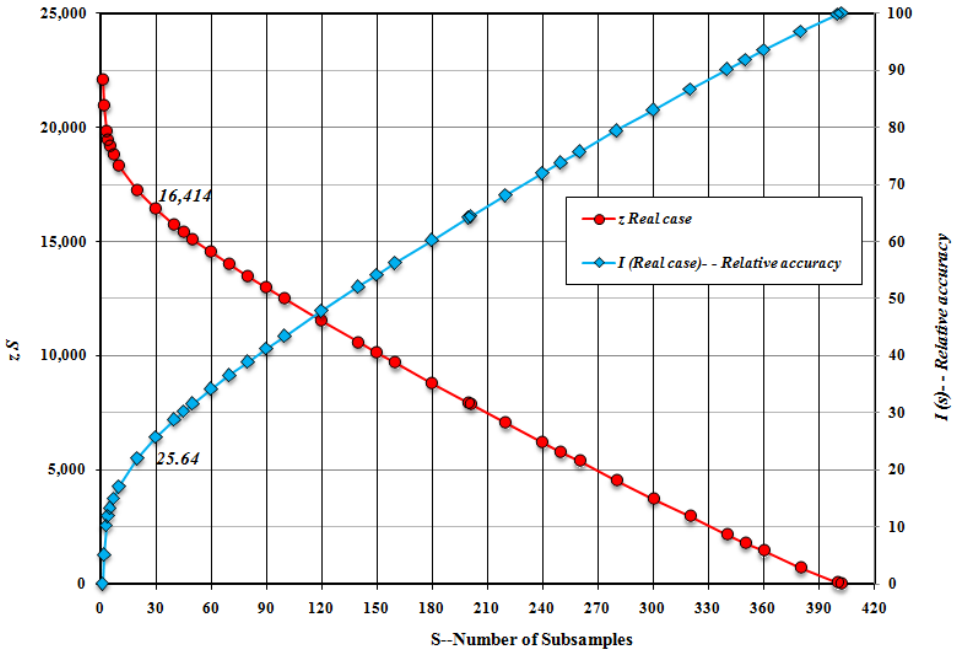

Figure 7: Distance z_S between the S-subsample and the full sample of $\mathcal{S}$ realizations (red) and relative accuracy (blue).

To visualize in an approximate way the relative positions of the realizations, we employ the MDS technique described in Section 4.1 to produce points in $\mathbb{R}^2$ for which the inter-

point distances approximate the distances D_{rs}. Figure 8 shows the MDS embedding of 400 simulations, E-type and Kriging models with the optimal weights assigned to each of the 30 subsamples shown as spike heights. The Kriged and E-type models are shown as green and black points, respectively.

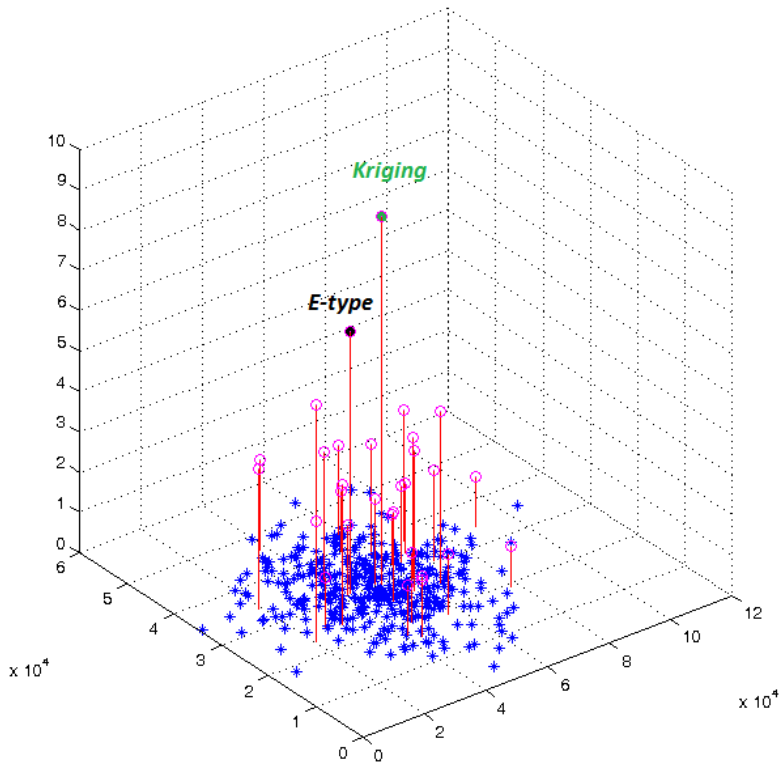

Figure 8: Multidimensional scaling embedding of 400 realizations, the E-type model, and the Kriging model in $\mathbb{R}^2$. The 30 subsamples that best represent the 402 points are shown in red, with the height of the "spike" indicating the relative weight given to each of the 30 subsamples. The Kriged and E-type models are shown in green and black, respectively.

A couple of things are apparent from figure 8. Firstly, the distances D_{rs} for $r =$ Kriged or $r =$ E-type and $s = 1, \ldots, \mathcal{S}$ are small on average, when compared to the overall average of the distances $D_{rs}, 1 \leq r < s \leq \mathcal{S}$. Formally, $(1/\mathcal{S}) \sum_{s=1}^{\mathcal{S}} D_{rs} = 23,606$ for $r =$ Kriged and $22,073$ for $r =$ E-type, while the overall average of distances is $\frac{1}{\mathcal{S}(\mathcal{S}-1)/2} \sum_{s=1}^{\mathcal{S}} \sum_{r<s} D_{rs} = 31,313$. Secondly, we see that the weights given to the selected realizations tend to be higher nearer to the center of the figure, while those selected realizations near the periphery tend to have lower weights (and often, the lowest possible nonzero weight of $1/S = 1/30$). These peripheral selections are effectively chosen to represent themselves only, while the selected realizations nearer to the center of the figure represent not only themselves but those realizations nearby,

effectively absorbing the weights from these nearby realizations. This unequal weight distribution is another advantage that our approach has over simply taking the first S realizations computed (we remark that if one requires $w_s = 1/S$ in (6)-(13), the optimization model (6)-(13) is easily modifiable to achieve this). We get to see more of the overall distribution of realizations by computing a larger number $\mathcal{S}$ and then are able to use this information to not only choose a good subsample, but also to allocate optimal weights to the subsample.

5 Conclusions

Although all conditional realizations are faithful to hard data, histogram and spatial correlation, there is no parameter that can provide further information about high order statistics for realizations. In other words, the space of uncertainty consisting of all generated realizations cannot be defined by low order statistics. In this paper we have introduced and applied the Kantorovich distance for the following two main purposes.
Firstly, we have developed a practical quantitative methodology to define and map the space of uncertainty by computing Kantorovich distance between generated realizations. This metric is a good notion of distance for measuring dissimilarity between the geological realizations. Indeed, the space of uncertainty is quantified by constructing the dissimilarity distance matrix D_{ij} (which contains all pairwise distances). Any related issues about the space of uncertainty such as sampling the space of uncertainty, evaluation of simulation algorithms, relative locations between the random realizations, impact of geostatistical parameters, and so on can now be evaluated.
Secondly, in order to select the subcollection of realizations that the best represents the possible outcomes, we have again used the concept of Kantorovich distance, and developed a simple optimization model to find the best samples. The optimization model can be used as a general tool, which is able to select any subset of representative realizations against user defined criteria.

In this paper, we focused on the concepts and methods rather than detailed applications, but the future work will include application to risk assessment, assessing the dependence of the space of uncertainty on parameters, and developing realizations with desired local and global accuracy.

Acknowledgments

We would like to acknowledge Dr. Jacques Deraisme, scientific adviser - principal consultant of Geovariance company for reviewing this paper and his valuable remarks, and Dr. Ed Malone from the school of mining engineering of UNSW for reviewing the manuscript.

References

C.V. Deutsch and A.G. Journel. *GSLIB Geostatistical Software Library and user's guide*. Oxford University Press, New York, NY, 1998.

P.A. Dowd and M. David. Planning from estimates: sensitivity of mine production schedules to estimation methods. In *Advanced Geostatistics in the Mining Industry*, pages 163–183. D. Reidel Publishing Company, 1976.

P.A. Dowd. Risk assessment in reserve estimation and open-pit planning. *Trans. Instn. Min. Metall*, 103:A148 – A154, 1994.

R. Dimitrakopoulos. Conditional simulation algorithm for modelling orebody uncertainty in open pit optimisation. *International of surface mining, Reclamation and Environment*, 12:173–178, 1998.

P. Goovaerts. Impact of the simulation algorithm, magnitude of ergodic fluctuations and number of realizations on the spaces of uncertainty of flow properties. *Stochastic Environmental Research and Risk Assessment*, 13:161–182, 1999.

E. H. Isaaks and R. M. Srivastava. *Applied geostatistics*. Oxford University Press, 1989.

C.A. Gotway and B.M. Rutherford. Stochastic simulation for imaging spatial uncertainty: comparison and evaluation of available algorithms. In *Geostatistical simulations*, pages 1–21. Kluwer Academic Publishers, 1993.

S.E. Qureshi and R. Dimitrakopoulos. Comparison of stochastic simulation algorithms in mapping spaces of uncertainty of non-linear transfer functions. In *Geostatistics Banff 2004*, volume 2, pages 959–968. Springer, 2005.

R. Dimitrakopoulos, C.T. Farrelly, and M.C. Godoy. Moving forward from traditional optimization: grade uncertainty and risk effects in open-pit design. *Transactions of the Institutions of Mining and Metallurgy,Section A: Mining Technology*, 111(1):82–88, 2002.

S. Suzuki and J. Caers. A distance-based prior model parameterization for constraning solutions of spatial invers problems. *Mathematical Geosciences*, 40:445–469, 2008.

C. Scheidt and J. Caers. Representing spatial uncertainty using distances and kernels. *Mathematical Geosciences*, 41:397–419, 2009.

A.M. Vershik. Kantorovich metric: initial history and little-known applications. *Journal of Mathematical Sciences*, 133:1410–1417, 2006.

Y. Deng and W. Du. The Kantorovich metric in computer science: A brief survey. *Electronic Notes in Theoretical Computer Science*, 253:73–82, 2009.

C. Villani. *Topics in optimal transportation*. American Mathematical Society, USA, 2003.

S. Ramazan and R. Dimitrakopoulos. Stochastic optimisation of long-term production scheduling for open pit mines with a new integer programming formulation. *Orebody Modelling and Strategic Mine Planning*, 14:353–360, 2004.

M. Menabde., G. Froyland, P. Stone, and G. Yeates. Mining schedule optimisation for conditionally simulated orebodies. *Proceedings of the international symposium on orebody modelling and strategic mine planning: uncertainty and risk management*, pages 347–352, 2004.

G. Froyland, M. Menabde, P. Stone, and D. Hodson. The value of additional drilling to open pit mining projects. *Proceedings of the International Symposium on Orebody Modelling and Strategic Mine Planning: Uncertainty and Risk Management*, pages 169–176, 2004.

N. Boland, I. Dumitrescu, and G. Froyland. A multistage stochastic programming approach to open pit mine production scheduling with uncertain geology. *Available at http://www.optimization-online.org*, 2008.

B. Tarhan, I.E. Grossmann, and V. Goel. Stochastic programming approach for the planning of offshore oil or gas field infrastructure under decision-dependent uncertainty. *Industrial & Engineering Chemistry Research*, 48(6):3078–3097, 2009.

A.M. Newman, E. Rubio, R. Caro, A. Weintraub, and K. Eurek. A review of operations research in mine planning. *Interfaces*, 40(3):222–245, 2010.

T. F. Cox and M.A.A. Cox. *Multidimensional scaling. 2nd ed.* Chapman and Hall, 2000.

J. Dupacova, N. Growe-Kuska, and W. Romisch. Scenario reduction in stochastic programming an approach using probability metrics. *Math. Program.*, A 95:493–511, 2003.

D. MacKay. *Information theory, inference, and learning algorithms.* Cambridge University Press, 2003.