Multivariate Change of Support

John G. Manchuk and Clayton V. Deutsch

Statistics depend on volume and this concept is significant for correctly describing natural resource properties. The mean, variance, distribution shape and variogram all depend on scale and this is described by change of support models. They are utilized to acquire the statistics at a scale that has not been sampled from a set of data at an often significantly smaller scale. Methods concentrate on the univariate statistics, but as more data is incorporated into geostatistical modeling projects, multivariate workflows are being considered. However, change of support for multivariate distribution models and linear models of coregionalization has not been developed. This work develops an approach for multivariate change of support for variables that average arithmetically. Results show the method preserves the mean as with existing univariate models and also achieves good results in terms of variance reduction for the multivariate covariance matrix.

Introduction

An important consideration in any modeling application is the size or scale of modeling elements or blocks. This scale sometimes referred to as the selective mining unit (SMU) is much larger than the samples from exploration and delineation, which may be in the form of chip samples, rock core, trench samples, or other forms. Generating a model at the scale of samples is unreasonable for computational reasons and selectivity reasons – surface mining equipment and underground stopes are substantially larger. However, generating models at the SMU scale requires statistics from data at the same scale and this is never available. Change of support models were developed for this purpose, to access data at the SMU scale based on data available at the much smaller sample scale.

Existing change of support models are useful for one variable, but many application involve multiple variables and these models do not necessarily honour the relationships across scales. Multivariate modeling applications involve correlations and spatial covariance structures that rarely exhibit independence rendering univariate change of support unusable. A method has been developed to construct a multivariate change of support model that is formulated to depend on the modeling procedure used. If the modeling procedure and all associated statistics are deemed correct, then the resulting change of support model will be equally acceptable. The general workflow involves constructing high resolution multivariate models, block averaging results to the SMU scale, and creating a vector field from the results that defines how the distribution changes from the sample scale to the SMU scale. This model can then be used to correct sample scale realizations generated on the SMU scale grid.

Background

Univariate change of support models attempt to define the distribution of a property such as mineral grade at a block scale v based on sample data having a much smaller scale. For practical purposes, the sample scale is assumed to be a point rather than a volume. Assuming they form a stationary random function (SRF), the samples $Z(\mathbf{x})$ follow a distribution function $f(z)$, but what is required is the distribution function, $f_v(z)$, of a SRF, $Z_v(\mathbf{x})$, at scale v, with no prior knowledge of what the properties of the larger scale data and distribution are. There is a set of conditions (Chiles and Delfiner, 1999) that control $f_v(z)$ for SRF's that scale arithmetically:

1. The mean does not change with scale: $E\{Z\} = E\{Z_v\}$
2. The variance is defined by (1), where x and y are spatial locations and C is the covariance function that describes Z at point support.

$$\sigma_v^2 = \frac{1}{|v|^2} \iint_v C(x-y)\,dx\,dy \tag{1}$$

3. $f_v(z)$ is less selective than $f(z)$, i.e. the variance does not increase, the range decreases, and the distribution undergoes some degree of symmetrisation.

Using these rules, various change of support techniques can be applied to derive a function that relates point scale values to their block scale equivalent. The usefulness of such a function is found in cases where point scale samples cannot be upscaled to the block scale, which occurs for two main reasons: the number of upscaled samples is inadequate to infer important statistics like the variogram; and/or point scale samples do not effectively discretize blocks resulting in unreliable upscaled values. Point scale data are instead used directly in estimation and simulation applications and results are transformed to reflect the appropriate block scale using a change of support model. Some models used in practice include affine correction, indirect lognormal correction, which are described by Isaaks and Srivastava (1989) on pages 471-476, and the discrete Gaussian model, described by Chilès and Delfiner (1999) on pages 432-433.

Change of support models have not been developed for multivariate data, but similar rules and concepts apply. The SRF is now a vector valued function, $\mathbf{Z}(\mathbf{x})$, with a density function defined by $f(z_1,z_2,\ldots z_m)$ where m is the number of variables being considered. Mapping point scale values $\mathbf{z}(\mathbf{x})$ to block scale values $\mathbf{z}_v(\mathbf{x})$, a set of rules similar to point scale applies:

1. The mean of all marginal distributions does not change with scale: $E\left\{Z^k\right\}=E\left\{Z_v^k\right\}, k=1,\ldots,m$
2. The covariance matrix is defined by (2), where x and y are spatial locations and C_{ij} is the covariance function between Z_i and Z_j at a point scale. If a valid linear model of coregionalization (LMC) is used, Σ will be positive definite
$$\Sigma_{ij}=\frac{1}{|v|^2}\iint_v C_{ij}(x-y)dxdy \tag{2}$$
3. $f_v(z_1,z_2,\ldots,z_m)$ is less selective than $f(z_1,z_2,\ldots,z_m)$
4. The correlation matrix is defined by (3). For an LMC that uses identical covariance functions and when all ratios between variance contributions and respective sills are constant, the correlation coefficients are independent of scale. Otherwise, they depend on differences between the covariance models.
$$R_{ij}=\Sigma_{ij}/\left(\Sigma_{ii}^{1/2}\cdot\Sigma_{jj}^{1/2}\right) \tag{3}$$

These rules provide only targets for statistics that describe the global distribution of a variable or set of variables. Also important is the distribution shape, and this is controlled by the sample data and their spatial orientation. In the previously mentioned existing methods for change of support, assumptions must be made about the distribution shape. Affine correction assumes the shape of $f(z)$ and $f_v(z)$ are identical; Indirect lognormal correction assumed both distributions are approximately lognormal; for the discrete Gaussian model, it is assumed that the distribution between Z and Z_v is bivariate Gaussian.

Applying the affine correction to multivariate data is accomplished by (4), where $\boldsymbol{\mu}$ is the mean vector, $\mathbf{U}^{1/2}$ is a diagonal matrix of standard deviations derived from Σ, and $\mathbf{V}^{1/2}$ is a diagonal matrix of the standard deviations of Z. Through this mapping, the desired variances are achieved, the marginal distribution shape is retained, and the correlation matrix does not change with scale. The last point is proved below in (5), where $\boldsymbol{\rho}$ is the correlation matrix of $\mathbf{Z}$, and the covariance of $\mathbf{Z}_v$ was derived from (6). For LMC specifications that do not result in a change in correlation coefficients, this model is acceptable; in all other cases (Oz and Deutsch, 2000) this method will depart from the target correlations from (3).

$$\begin{aligned}\mathbf{z}_v(\mathbf{x})&=\boldsymbol{\mu}+\mathbf{U}^{1/2}\mathbf{z}'(\mathbf{x})\\ \mathbf{z}'(\mathbf{x})&=\mathbf{V}^{-1/2}\left(\mathbf{z}(\mathbf{x})-\boldsymbol{\mu}\right)\end{aligned} \tag{4}$$

$$\begin{aligned}\mathrm{Corr}(\mathbf{Z}_v)&=\mathbf{U}^{-1/2}\,\mathrm{Cov}(\mathbf{Z}_v)\mathbf{U}^{-1/2}\\ &=\mathbf{U}^{-1/2}\left(\mathbf{U}^{1/2}\boldsymbol{\rho}\mathbf{U}^{1/2}\right)\mathbf{U}^{-1/2}\\ &=\boldsymbol{\rho}=\mathrm{Corr}(\mathbf{Z}) \qquad \square\end{aligned} \tag{5}$$

$$\begin{aligned}Cov(\mathbf{Z}_v)&=\mathbf{U}^{1/2}Cov(\mathbf{Z}')\mathbf{U}^{1/2}\\ &=\mathbf{U}^{1/2}\left(V^{-1/2}\Sigma V^{-1/2}\right)\mathbf{U}^{1/2}\\ &=\mathbf{U}^{1/2}\boldsymbol{\rho}\mathbf{U}^{1/2}\end{aligned} \tag{6}$$

Results from a simple example using two variables and random samples are shown in Figure 1. Variable one was generated with a mean of 3 and a variance of 2 and variable two was generated with a mean of 5 and a variance of 1 and a covariance between them of 1.1. Affine correction was used to reduce the variance by a factor of 4. Statistics for the sample data from before and after are provided in Table 1.

Table 1: Statistics for affine correction example

Statistic	Before	After
Mean	$\begin{bmatrix} 2.992 & 5.018 \end{bmatrix}$	$\begin{bmatrix} 2.996 & 5.009 \end{bmatrix}$
Covariance	$\begin{bmatrix} 2.084 & 1.153 \\ 1.153 & 1.031 \end{bmatrix}$	$\begin{bmatrix} 0.521 & 0.288 \\ 0.288 & 0.258 \end{bmatrix}$
Correlation	$\begin{bmatrix} 1 & 0.787 \\ 0.787 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0.787 \\ 0.787 & 1 \end{bmatrix}$

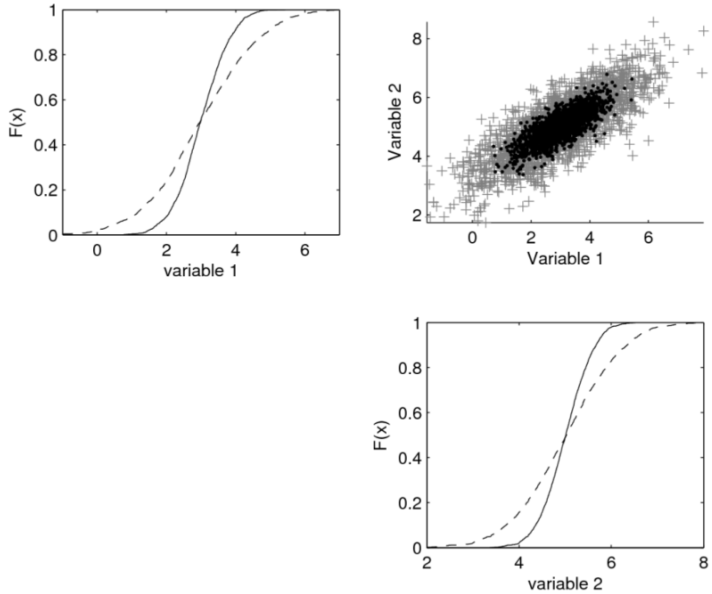

Figure 1: Results of affine correction: dashed line – original distribution; solid line – affine corrected distribution; gray plus signs – original sample pairs; black bullets – affine corrected sample pairs.

The simplicity of affine correction allows a straightforward extension to multivariate applications; however, the goal of incorporating more variables into modeling applications is to achieve more accurate models and improved estimates of uncertainty. The affine correction may do just the opposite if the incorrect multivariate relationships are achieved with its implementation. A more advanced method is required to account for the statistics and for the spatial orientation of the data.

Methodology

Working with multivariate data having potentially complex relationships and LMC specifications, development of theoretical change of support techniques becomes a challenge. A somewhat brute force approach to the problem is to use the same workflow as the modeling process will use; therefore all sample data, input parameters and the spatial context will be problem specific. Resulting change of support models will be applicable to the problem at hand. In summery, the approach uses simulation to generate point scale fields that are used in an upscaling process to acquire block support data from which a change of support model can be derived. Although in the past this approach has not been advocated, today's computational resources make it feasible.

Advantages of using simulation include: the domain of interest is used directly; the distribution and spatial orientation of sample data are used; and the variogram or LMC are incorporated. Programs and code already exists for multivariate simulation. The problem is generating a mapping between the point scale data and block averaged data from simulation, especially when the dimensionality of the data becomes high. Resulting change of support models must also meet the requirements laid out in the previous section: constant mean; covariance matrix defined by (2); correlations defined by (3); and less selectivity. Methods for computing the targeted covariance and correlation matrix already exist, but only for a single covariance function. Extending the basic discretization technique to incorporate an LMC is straightforward, there are simply more variogram models.

In practice, change of support models are applied to point scale values simulated at block centers; therefore deriving a model is done via the same process. The targeted scale is the volume of blocks in the simulation grid, which will be referred to as the coarse grid. These must be discretized into a set of fine scale blocks for which point scale simulation is carried out. To be consistent with the application of change of support models, the set of fine scale blocks should be defined so that block centers of both grids align, see Figure 2.

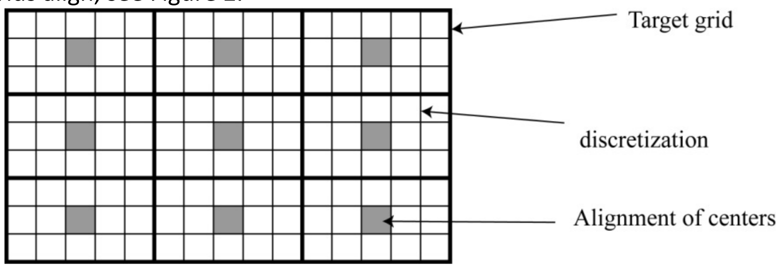

Figure 2: Target grid and discretization

Performing simulation on the discretization provides point support values and arithmetically averaged block scale values derived from them. A link between the two scales is made using difference vectors calculated from $\hat{\mathbf{z}}_v(\mathbf{x}) - \mathbf{z}(\mathbf{x}) = \Delta\mathbf{z} = [\Delta\mathbf{z}_{v,1}\ \Delta\mathbf{z}_{v,2} \cdots \Delta\mathbf{z}_{v,m}]$, where a hat indicates the block averaged value is computed from samples and is an approximation. There will be as many vectors as coarse grid blocks for a single realization; however, it is possible to obtain additional vector sets through the use of grid offsets. Keeping the fine grid fixed and offsetting the coarse grid yields another set of point support values and block averages, see Figure 3.

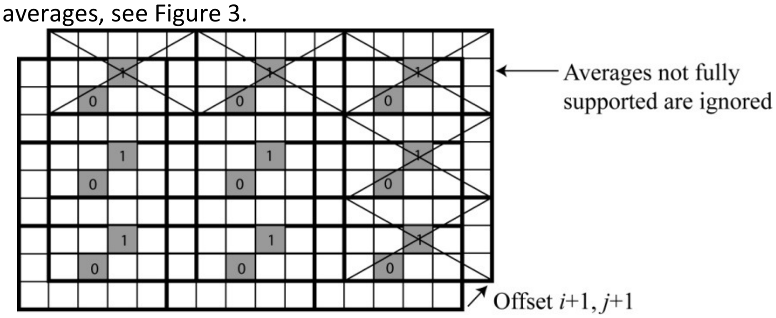

Figure 3: Example of grid offsetting

Difference vectors can be considered samples of the gradient from some underlying function *g*(**z**). By developing a model for *g*(**z**) and ensuring that it is differentiable, or for $\partial g / \partial \mathbf{z}$ directly, any point scale value **z** can be mapped to the appropriate block scale value via $\mathbf{z}_v = \mathbf{z} + \partial g / \partial \mathbf{z}$. Instead of attempting to solve the differential equation problem for *g*(**z**), this work will only model the gradient data directly using simple interpolation techniques. The former may be considered in future research.

The problem is to solve for a gradient field that maps any point scale realization to a block scale realization such that the mean is constant and the appropriate change in variance is achieved. The gradient field is the change of support model. Simulation and block averaging provides data, $\Delta\mathbf{z}$, from which the field is derived. Gaussian weighting is used to interpolate the value of $\Delta\mathbf{z}$ vector sets to a set of lattice points that covers the multivariate distribution *f*(**z**). In practice, the limits of a distribution are known, for example mineral grades are non negative and less than some sampled or extrapolated maximum grade. These define the extents of the lattice. The interpolated value for a particular lattice point is given by (7), where w_i are the weights, W is the sum of the weights, Ω_k is the local neighbourhood of $\mathbf{z}_k$, d_i is the distance between $\mathbf{z}_k$ and $\mathbf{z}_i$, $\left\|\mathbf{z}_k - \mathbf{z}_i\right\|_2$, and α_k and β_k are shape parameters.

$$\Delta\hat{\mathbf{z}}_k = \frac{1}{W}\sum_{i\in\Omega_k} w_i \Delta\mathbf{z}_i \qquad (7)$$
$$w_i = \exp\left(-\left(d_i - \alpha_k\right)^2 / \beta_k^{\,2}\right)$$

Each fine scale realization will yield a realization of difference vectors and therefore a different change of support model. In geostatistics, one purpose of multiple realizations is to obtain the most likely value, or the expected value, of some property. The same principle applies to developing the most likely change of support model for the data, domain, and modeling parameters. Multiple realizations of change of support models are averaged to obtain the one that will provide the correct mean and covariance of future realizations. Denoting the lattice of difference vectors L_l, where l is a multidimensional index, the expected change of support model is given by (8) with r being the number of realizations. If n_r multiple grid offsets are used, the number of accessible realizations is $r \cdot n_r$.

$$E\{L_l\} = \frac{1}{r}\sum_{k=1}^{r} L_{l,r} = \bar{L}_l \qquad (8)$$

Future realizations generated at a point scale on the coarse grid block centers are mapped to the appropriate scale using $\bar{L}_l$ and the same interpolation method as in (7).

Implementation

The above methodology is tested on two cases: 1 – bivariate unconditional Gaussian realizations with single structure direct and cross variograms; 2 – the same as previous but involving data transformations. Both cases are applied on a two dimensional Cartesian grid and the following items are checked: theoretical variances from evaluating (2); the change of support model; realization maps; and point and block scale distributions and associated statistics.

Case 1

Bivariate LMC parameters are defined in Table 2. Both variables are assumed Gaussian with zero mean and unit variance and a correlation of 0.7. The change of support model was generated using a small 10 by 10 coarse scale grid and the corresponding fine scale grid was designed to give 11 by 11 points per coarse scale block, see Table 3. 10 realizations were used with no grid offsets. Evaluating (2) numerically using the same discretization, and using 51 by 51 points for comparison, yields the symmetric covariance matrices in Table 4. In this case, there is no change in the correlation coefficient with scale.

Table 2: LMC for case 1

Structure	Head variable	Tail variable	Variance	Azimuth	Major range	Minor range
Nugget			0			
Spherical	1	1	1	45	50	15
	1	2	0.7	45	50	15
	2	2	1	45	50	15

Table 3: Grid specifications for case 1

Grid	Axis	Number	Origin	Size
Coarse	*x*	10	5	10
	y	10	5	10
Fine	*x*	110	0.45…	0.90…
	y	110	0.45…	0.90…

Table 4: Theoretical covariance matrices for case 1

11 by 11 points	51 by 51 points
$\begin{bmatrix} 0.654812 & 0.458369 \\ & 0.654812 \end{bmatrix}$	$\begin{bmatrix} 0.653556 & 0.457489 \\ & 0.653556 \end{bmatrix}$

The bivariate distribution was discretized into a 20 by 20 grid. A few iterations were carried out to optimize the parameters for the neighbourhood size and Gaussian kernel variance by hand. Using the nearest 15 points to define Ω_k and a variance of $r_k^2 / 4$, the error between the theoretical covariance matrix and that achieved using the change of support model was small. Figure 4 shows the resulting change of support model as a vector field. All vectors converge to the mean and the field appears symmetric as should be the case for a bivariate Gaussian distribution.

Testing of the model was done using a 50 by 50 grid with the same block dimensions as the coarse grid used to generate it. 50 realizations of point scale values were generated and transformed according to $\hat{\mathbf{z}}_v(\mathbf{x}) = \mathbf{z}(\mathbf{x}) + \Delta\mathbf{z}$, where $\Delta\mathbf{z}$ is interpolated from the change of support field. Resulting statistics averaged over all realizations at the point scale and block scale are provided in Table 5. A realization for both variables and scales is in **Error! Reference source not found.**. Considering that the change of support model was built without any constraints on the mean and covariance matrix, results are encouraging: the mean for both variables stayed very close to zero and the reduction in variance is close to 0.65 as dictated by the theoretical values calculated from Table 4. The covariance between variables contains more error with a reduction of only 0.58 rather than 0.65. This has the effect of a lower correlation of 0.638 instead of 0.7.

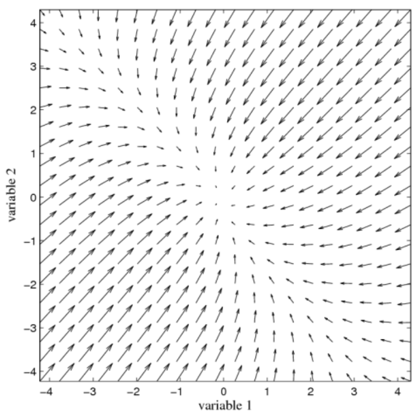

Figure 4: Bivariate change of support model for case 1

Table 5: Case 1 point and block scale statistics

Statistic	Point scale	Block scale
Mean	$\begin{bmatrix} 0.0094 & 0.0072 \end{bmatrix}$	$\begin{bmatrix} -0.0032 & 0.0051 \end{bmatrix}$
Covariance	$\begin{bmatrix} 0.99649 & 0.69693 \\ & 0.99222 \end{bmatrix}$	$\begin{bmatrix} 0.64197 & 0.40790 \\ & 0.63561 \end{bmatrix}$
Covariance reduction		$\begin{bmatrix} 0.64423 & 0.58528 \\ & 0.64059 \end{bmatrix}$
Correlation	0.70073	0.63832

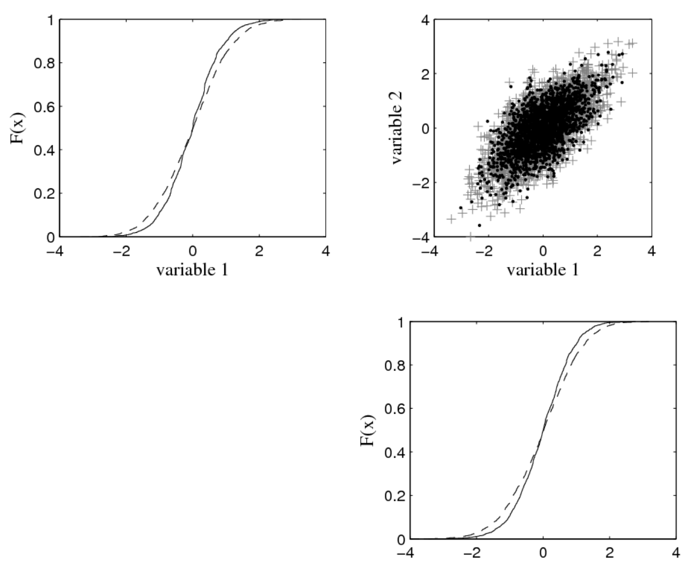

Figure 5: Results for case 1 realization 5 change of support: dashed line – original distribution; solid line – corrected distribution; gray plus signs – original sample pairs; black bullets –corrected sample pairs.

Case 2

Variables are often never standard Gaussian so this case involves data transformations to convert a set of Gaussian realizations to those of a more realistic distribution. For example, low valued positive variables like gold grade might appear to follow a lognormal distribution. The same realizations used in Case 1 were transformed into lognormal distributions using (9), where **Y** represents the Gaussian values. The mean and variance of both variables are the same and are determined by (10) and (11) and the covariance is defined by (12). To evaluate the theoretical covariance matrix at the block scale the LMC for the lognormal data are required; however, only the Gaussian variograms have been defined. These are converted according to (12) as well, but Cov(**Z**) and Cov(**Y**) are replaced by covariance functions that are defined by the LMC. The correlation coefficient is no longer constant and is defined by (13) at the point scale; it is computed numerically for the block scale along with the covariance matrix in Table 6. The mean should remain constant since arithmetic averaging is used.

$$\mathbf{Z} = \exp\left(\alpha + \beta \cdot \mathbf{Y}\right) \qquad \alpha = 0,\ \beta = 1/2 \tag{9}$$

$$m_{\mathbf{Z}} = \exp\left(\alpha + \beta^2/2\right) = 1.13314 \tag{10}$$

$$\sigma_{\mathbf{Z}}^2 = m_{\mathbf{Z}}^2\left(\exp\left(\beta^2\right) - 1\right) = 0.36470 \tag{11}$$

$$\mathrm{Cov}(\mathbf{Z}) = m_{Z1} m_{Z2}\left(\exp\left(\beta^2\,\mathrm{Cov}(\mathbf{Y})\right) - 1\right) = 0.245561 \tag{12}$$

$$\rho_{Y1,Y2} = \frac{\exp\left(\beta^2 \sigma_{Y1,Y2}\right) - 1}{\sqrt{\exp\left(\beta^2 \sigma_{Y1}^2\right) - 1}\sqrt{\exp\left(\beta^2 \sigma_{Y2}^2\right) - 1}} = 0.673342 \tag{13}$$

Table 6: Theoretical covariance matrices and correlations for case 2

	11 by 11 points	**51 by 51 points**
Covariance matrix	$\begin{bmatrix} 0.230215 & 0.156758 \\ & 0.230215 \end{bmatrix}$	$\begin{bmatrix} 0.229747 & 0.156445 \\ & 0.229747 \end{bmatrix}$
Correlation coefficient	0.680922	0.680945

As before, a few iterations were used to optimize the parameters and it was found that using the nearest 12 points to define Ω_k and a variance of $r_k^2/8$, the error was roughly minimized. Figure shows the resulting change of support model, which is roughly symmetric and shows convergence to the mean. Gaussian realizations on the 50 by 50 grid were converted to lognormal using (9) and transformed according to the change of support model. Resulting statistics are given in Table 7 showing almost no change in the mean and the covariance is similar to the results from Table 6. The covariance reduction calculated using Table 6 and the values from equations (11) and (12) is roughly 0.63 for all terms. As in case 1, the largest departure is observed with the off diagonal term having a reduction factor of 0.60. Resulting distributions are shown in Figure . Realizations are identical to those of **Error! Reference source not found.** when back-transformed using (9) and are not shown.

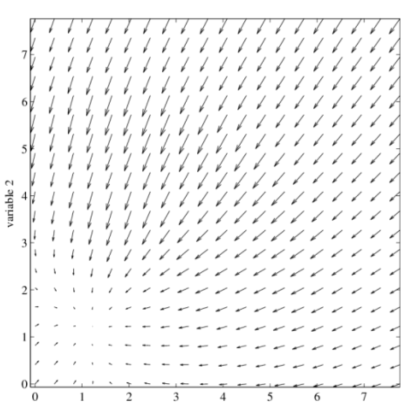

Figure 6: Bivariate change of support model for case 2

Table 7: Case 2 point and block scale statistics

Statistic	**Point scale**	**Block scale**
Mean	$\begin{bmatrix} 1.1387 & 1.1363 \end{bmatrix}$	$\begin{bmatrix} 1.1309 & 1.1394 \end{bmatrix}$
Covariance	$\begin{bmatrix} 0.37035 & 0.24781 \\ & 0.36195 \end{bmatrix}$	$\begin{bmatrix} 0.23505 & 0.14870 \\ & 0.22216 \end{bmatrix}$
Covariance reduction		$\begin{bmatrix} 0.63467 & 0.60006 \\ & 0.61379 \end{bmatrix}$
Correlation	0.67668	0.65039

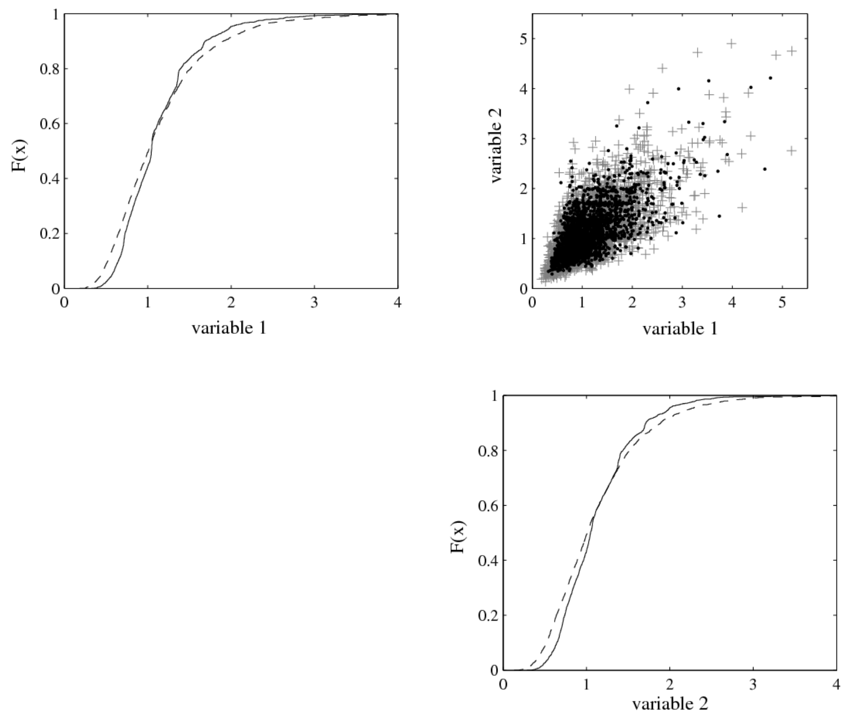

Figure 7: Results for case 2 realization 5 change of support: dashed line – original distribution; solid line – corrected distribution; gray plus signs – original sample pairs; black bullets –corrected sample pairs.

Conclusions

Multivariate change of support is an important problem in mining applications when several variables are involved and they show correlation structure that cannot be overlooked. For example, different variables may impact ore processing in various ways – incorrectly categorizing SMU's because change of support was not implemented to account for correlation could lead to adverse effects or monetary losses. Correctly scaling all targeted variables for production is important for resource valuation purposes as well. Presented in this work was a fairly general technique for multivariate change of support. It was shown to be applicable in the Gaussian case and the lognormal case. If any distribution model is available for use in multivariate simulation this method can be implemented. One potential issue is dimensionality – the size of the change of support model grows rapidly with the number of variables, for example, 6 variables and a grid with 20 nodes per dimension requires 64 million vectors. More advanced modeling techniques to represent the multivariate distribution and the vector field will be required for such cases.

References

Isaaks E.H. and Srivastava R.M. (1989) An introduction to Applied Geostatistics. Oxford University Press, 592

Chiles J.-P. and Delfiner P. (1999) Geostatistics: Modeling Spatial Uncertainty. John Wiley & Sons, 641

Oz B. and Deutsch C.V. (2000) Size scaling of cross-correlation between multiple variables. Centre for Computational Geostatistics, Annual Report 3, 1-25