RANK ORDER GEOSTATISTICS: A PROPOSAL FOR A UNIQUE CODING AND COMMON PROCESSING OF DIVERSE DATA

A.G. JOURNEL AND C.V. DEUTSCH
Department of Petroleum Engineering
Stanford University, Stanford, CA 94305-2220

ABSTRACT One of the greatest challenges in earth sciences data processing, and an area where geostatistics has contributed extensively, is the integration of data of diverse types, scales, supports, and accuracies.

We suggest that integration starts by coding the information in a common [0, 1] format of either indicator or uniform score transforms and prior cumulative probability values (or *cdf's*). This coding preserves the spatial ranks (hence structures) of the data. Kriging and stochastic simulation can be performed in that uniform space and the results back-transformed to the original data units. We also show how prior probability distributions from direct and cross **h**-scattergrams coded in the standard [0, 1] data format can be weighted to produce posterior probability distributions.

1. Introduction

A typical feature of earth sciences is the multiplicity of data types contributing to the understanding of any single phenomenon. Data originate from a variety of sources related to different attribute values measured on different volume supports or time spans, at different scales and with different accuracies. Characteristically, because of accessibility and cost, there is a shortage of "hard" data defined as the most direct or accurate measurement of the (principal) variable(s) of interest. Because of that shortage, all related secondary data must be used accounting for their respective information content and discounting information redundancy. One of the greatest challenges in earth sciences data processing is the integration of data of diverse types, scales, supports, and accuracies.

The major roadblock in data integration is the difference in the format (not only measurement units) under which each data type is presented. Some information types are interpretive in nature yet could be critical in

E.Y. Baafi and N.A. Schofield (eds.), Geostatistics Wollongong '96, Volume 1, 174-187.

where $F(z) = Prob\{Z(\mathbf{u}) \leq z\}$ is the cumulative distribution function (cdf) of the stationary random function (RF) model $Z(\mathbf{u})$, inferred from the sample cdf.

- The distribution of $V(\mathbf{u})$ is uniform in $[0,1]$. Indeed:

$$\begin{aligned} Prob\{V(\mathbf{u}) \leq v\} &= Prob\{F(Z(\mathbf{u})) \leq v\} = Prob\{Z(\mathbf{u}) \leq F^{-1}(v)\} \\ &= F(F^{-1}(v)) = v \ , \forall \ v \in [0,1] \end{aligned}$$

  as long as the cdf $F(z)$ is invertible which excludes severely discontinuous histograms presenting large spikes. Recall that ties can be broken by considering their neighborhood values.
- Consider n data values $z(\mathbf{u}_\alpha), \alpha = 1, \ldots, n$, and their rank orders $r(\mathbf{u}_\alpha)$, with $r(\cdot) = 1$ for the lowest datum and $r(\cdot) = n$ for the largest. The standardized rank $v(\cdot) = r(\cdot)/n$ is the uniform transform (1).
- The Spearman rank correlation is but the traditional linear correlation calculated on the uniform transform V. Similarly, the rank cross-covariance $C_{V_1,V_2}(\mathbf{h}) = Cov\{V_1(\mathbf{u}), V_2(\mathbf{u}+\mathbf{h})\}$ measures the degree of monotonic dependence between the two original variables $Z_1(\mathbf{u})$ and $Z_2(\mathbf{u}+\mathbf{h})$, no matter their different types or supports. This uniform score (cross)covariance is unaffected by the respective univariate distributions of Z_1 and Z_2, it depends only on the bivariate (spatial) relations between $Z_1(\mathbf{u})$ and $Z_2(\mathbf{u}+\mathbf{h})$.
- The normal score transform, Deutsch and Journel (1992, p. 138),

$$Y(\mathbf{u}) = G^{-1}(F(Z(\mathbf{u})) = G^{-1}(V(\mathbf{u}))$$

  starts by a uniform transform followed by a standard normal quantile transform (G^{-1}). That second transform is useful only if one wishes to call on the congenial but restrictive properties of the Gaussian RF model. If kriging and stochastic simulation can be performed on the uniform score transform $V(\mathbf{u})$ and the results back transformed by F^{-1}, then there is no advantage to the additional transform G^{-1}.

2.1. KRIGING OF RANKS

Uniform score values can be estimated by (co)kriging using the corresponding uniform scores (cross)covariances:

$$v^*(\mathbf{u}) = \sum_{\alpha=1}^{n} \lambda_\alpha v(\mathbf{u}_\alpha) \tag{2}$$

The result can be interpreted as the estimated rank of the unknown original value $z(\mathbf{u})$. An estimate of $z(\mathbf{u})$ is then given by the back transform:

$$z^*(\mathbf{u}) = F^{-1}(v^*(\mathbf{u})) \tag{3}$$

the understanding of the primary phenomenon, some data are categorical (e.g., facies types), some are numerical (e.g., concentration values), some appear as constraints (e.g., stoichiometric inequalities), and some are prior probability distributions valued in $[0,1]$. A standardization of formats and units that does not tamper with the information content would be an important first step towards data integration. A methodology for merging the data, in a common format, would follow.

An important aspect to remember when coding information is that information is goal-dependent: what counts in a secondary variable value, say $z_s(\mathbf{u}_\alpha)$ for a datum of type s at location $\mathbf{u}_\alpha$, is the information it carries relative to the primary variable (the goal), say $z_0(\mathbf{u})$ at possibly different locations $\mathbf{u} \neq \mathbf{u}_\alpha$. When the goal changes, the information carried by datum $z_s(\mathbf{u}_\alpha)$ changes. A piece of information may be considered "hard" for one application and "soft" or even irrelevant for other applications.

The probabilistic language and methodology is unique and universal in that they are not linked to any particular field of application or data type. Probability values are unit-free; in their cumulative form they are standardized in the ultimate standard interval $[0,1]$. The concept of prior probability distributions, say of $Z(\mathbf{u})$ given datum $Z_s(\mathbf{u}_\alpha) = z_s(\mathbf{u}_\alpha)$, allows a common coding of diverse data related to the same goal, say evaluation of $z(\mathbf{u})$. The concept of Bayesian updating provides a methodology, or at least a model for developing a methodology, for merging prior distributions into a single posterior probability distribution. All original data and the final posterior probability distribution are coded as cumulative probability values in the interval $[0,1]$. From the posterior cumulative distribution function (ccdf) probability intervals can be derived, simulated values can be drawn for the unsampled value, or a single "best" estimated value can be retained for any given optimality criterion.

2. Uniform Score Transform

The first task in standardizing continuous data of different types, supports, or scales is to get rid of their different units through an invertible transform that allows the original units to be recovered at any time. This is done traditionally by considering the standard residual $Y = (z - m)/\sigma$ where m and σ^2 are the mean and variance of the original variable Z. Indeed, the linear coefficient of correlation is the covariance of such standard residuals. As for filtering the mean and variance, we may want to filter out the entire distribution (histogram) of the variable Z by considering its standard rank transform or uniform scores defined as:

$$V(\mathbf{u}) = F\left(Z(\mathbf{u})\right) \quad \in \quad [0,1] \tag{1}$$

- That estimate is exact; indeed at any datum location: $v^*(\mathbf{u}_\alpha) \equiv v(\mathbf{u}_\alpha)$, hence $z^*(\mathbf{u}_\alpha) = z(\mathbf{u}_\alpha)$, as long as the cdf transform $F(\cdot)$ is invertible.
- The problem with the kriging (2) is that it does not ensure that the resulting estimates $v^*(\mathbf{u})$ are valued in $[0,1]$. A sufficient condition would be to ensure that no weight λ_α is negative and that they sum up to one, Rao and Journel (1996), and others.
- Typical of kriging, the rank estimates although unbiased (mean 0.5) are smooth with variance less than $1/12$, the variance of a $[0,1]$ uniform distribution. Consequently, the backtransformed values $z^*(\mathbf{u})$ are median-unbiased with a distribution different from the original cdf $F(z)$ and smaller variance. To correct for this problem, the estimates $v^*(\mathbf{u})$ can be themselves rank-transformed under the constraint of data reproduction, Journel and Xu (1994).

$$z^{**}(\mathbf{u}) = F^{-1}(v^*(\mathbf{u})) + \lambda(\mathbf{u}) \cdot \left[F^{-1}(L(v^*(\mathbf{u}))) - F^{-1}(v^*(\mathbf{u}))\right] \quad (4)$$

where $L(v)$ is the cdf of all kriged values $v^*(\mathbf{u})$, thus the $L(v^*(\mathbf{u}))$'s are uniformly distributed in $[0,1]$, $\lambda(\mathbf{u}) = [\sigma_K(\mathbf{u})/\sigma_{Kmax}]^\omega \in [0,1]$ is a relative correction factor ensuring data reproduction; at any datum location the kriging variance related to $v^*(\mathbf{u}_\alpha) = v(\mathbf{u}_\alpha)$ is $\sigma_K^2(\mathbf{u}_\alpha) = 0$, hence $\lambda(\mathbf{u}_\alpha) = 0$ and $z^{**}(\mathbf{u}_\alpha) = F^{-1}(v(\mathbf{u}_\alpha)) = z(\mathbf{u}_\alpha)$, σ_{Kmax}^2 is the largest of all kriging variances $\sigma_K^2(\mathbf{u})$, $\omega > 0$ is a correction level parameter; the larger ω the more gradual the correction away from the data locations.

- Except close to the data locations where $\lambda(\mathbf{u})$ is small, $z^{**}(\mathbf{u}) \approx F^{-1}(L(v^*(\mathbf{u})))$, hence the histogram of the z^{**}-estimates approximates the original z-cdf $F(z)$. Therefore the estimator $Z^{**}(\mathbf{u})$ is now mean-unbiased with the correct variance, that of $F(z)$.
- $Z^{**}(\mathbf{u})$ is not minimum error-variance, only $V^*(\mathbf{u})$ has that property. Although the correction (4) restitutes the correct global variance to the field of z^{**}-values it does not ensure reproduction of the Z or even V-covariance. Such reproduction calls for stochastic simulation

2.2. SIMULATING THE RANKS

Rather than retaining the smoothed kriged value $v^*(\mathbf{u})$ defined in relation (2), consider correcting for the missing variance by drawing a (simulated) value $y^{(l)}(\mathbf{u})$ from *any* distribution with mean $v^*(\mathbf{u})$ and variance equal to the kriging variance $\sigma_K^2(\mathbf{u})$. For example, a uniform distribution in the interval $[v^*(\mathbf{u}) \pm a/2]$ with $a = \sqrt{12} \cdot \sigma_K(\mathbf{u})$ might be considered. That simulated value $y^{(l)}(\mathbf{u})$ is then used as a v-datum for kriging and simulation at all subsequent nodes of the grid A. It can be shown that the field so simulated, $\{y^{(l)}(\mathbf{u}), \mathbf{u} \in A\}$ for realization # l, will have the same covariance as that used in the kriging of the $v^*(\mathbf{u})$'s, that is, the V or rank order covariance $C_V(\mathbf{h})$, Journel (1993).

There remains to enforce a uniform $[0,1]$ distribution to the simulated values $y^{(l)}(\mathbf{u})$ while preserving the two properties of data exactitude $y^{(l)}(\mathbf{u}_\alpha) = v(\mathbf{u}_\alpha)$ and covariance $C_V(\mathbf{h})$ reproduction. This is done by a rank-preserving transform similar to (4):

$$v^{(l)}(\mathbf{u}) = y^{(l)}(\mathbf{u}) + \lambda(\mathbf{u}) \cdot \left[L^{(l)}(y^{(l)}(\mathbf{u})) - y^{(l)}(\mathbf{u})\right] \tag{5}$$

where $L^{(l)}(y)$ is the cdf of all values $y^{(l)}(\mathbf{u}), \mathbf{u} \in A$, of realization l, $\lambda(\mathbf{u})$ is as defined in relation (4). Again, at any data location, $\lambda(\mathbf{u}_\alpha) = 0$, hence $v^{(l)}(\mathbf{u}_\alpha) = y^{(l)}(\mathbf{u}_\alpha) = v(\mathbf{u}_\alpha)$. Away from the data locations: $v^{(l)}(\mathbf{u}) \approx L^{(l)}(y^{(l)}(\mathbf{u}))$, hence the histogram of the $v^{(l)}(\mathbf{u})$'s approximates a uniform $[0,1]$ distribution.

Last, the simulated ranks $v^{(l)}(\mathbf{u})$ are back transformed into z-values:

$$z^{(l)}(\mathbf{u}) = F^{-1}(v^{(l)}(\mathbf{u})) \tag{6}$$

These $z^{(l)}$-simulated values are such that (1) z-data are honored $z^{(l)}(\mathbf{u}_\alpha) = F^{-1}(v(\mathbf{u}_\alpha)) = z(\mathbf{u}_\alpha)$, (2) the z-cdf $F(z)$ is reproduced, entailing unbiasedness, and (3) the z-rank covariance $C_V(\mathbf{h})$ is reproduced.

The previous kriging and simulation of uniform scores can be extended to cokriging and cosimulation of the respective uniform scores of several covariates $Z_1(\mathbf{u}), Z_2(\mathbf{u}), \ldots$. The distinct advantage of working with uniform scores, besides the ultimate standardization of all data types to the same marginal uniform $[0,1]$ distribution, is the extreme robustness of the rank order statistics involved.

3. A Small Example

Figure 1 shows a location map, histogram, and cumulative distribution of 29 data. The 29 data locations are shown on a 50 by 50 grid of reference true values (from GSLIB (Deutsch & Journel 1992)). The histogram of the 29 data values together with a smoothed histogram model are shown in the upper right of Figure 1. The cumulative distribution $F(z)$ of the 29 data values is shown in the lower right. The cumulative distribution $F(z)$ can be seen as a graphical relationship between the Z values (valued between 0.01 and 100) and the V rank transformed values (valued between 0.0 and 1.0). As illustrated, a Z-value of 2.3 is transformed to a V-value of 0.64. The 29 V rank transformed values are now considered; the Z values can be retrieved at any time.

The omnidirectional semivariogram of the V data is shown on Figure 2. The sill of this semivariogram is the variance of a uniform distribution, i.e.,

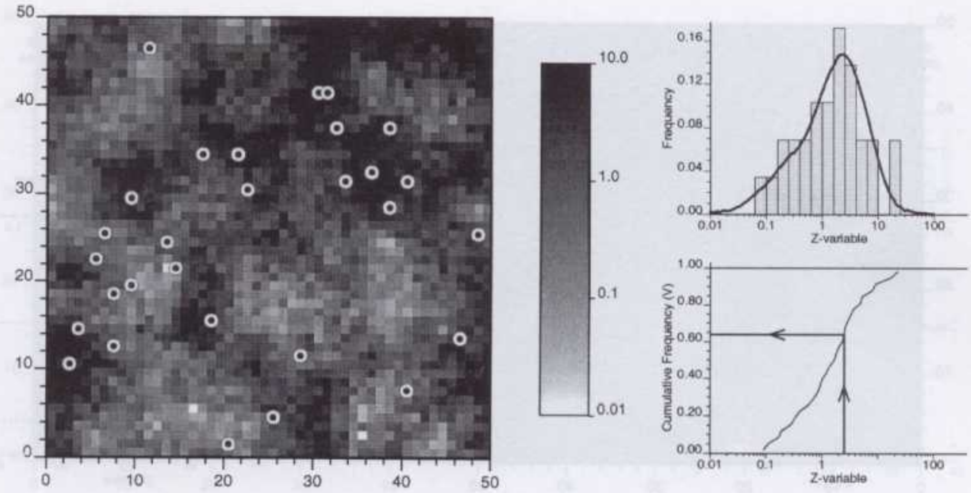

Figure 1. Gray-scale coded location map, histogram, and cumulative histogram of 29 data (see `data.dat` in GSLIB (Deutsch & Journel 1992)).

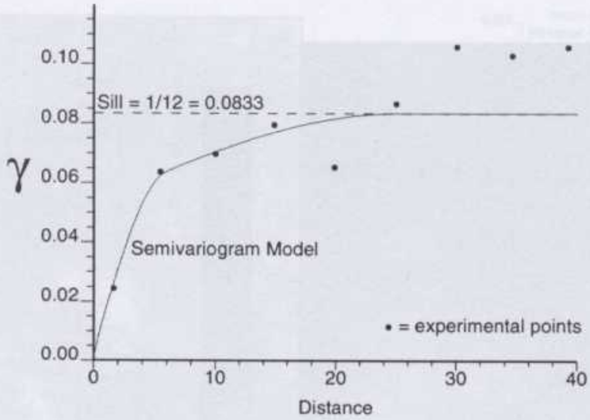

Figure 2. Isotropic semivariogram of rank transform of 29 data shown on Figure 1.

$1/12 = 0.0833$. The semivariogram model has a nugget effect of 0.0033, a first spherical structure with a sill of 0.05 and range of 6.0, and a second spherical structure with a sill of 0.03 and range of 25.0.

Simple kriging of the rank transformed data on the 50 by 50 grid using a global mean of 0.5 yields the results shown on Figure 3. The characteristic smoothing of kriging is evident on the map, histogram, and variogram. The histogram departs significantly from the input uniform distribution of the data. The variance of the kriged estimates is only 17% of the data variance of $1/12 = 0.0833$.

The kriged values are then transformed under the constraint of data

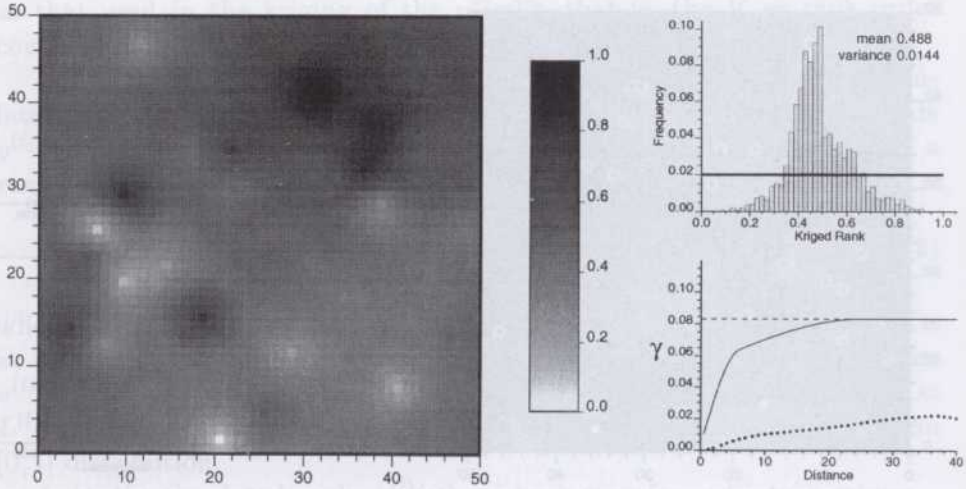

Figure 3. Gray-scale map, histogram, and semivariogram of kriged values. The solid semivariogram curve is the model and the black dots are the experimental values.

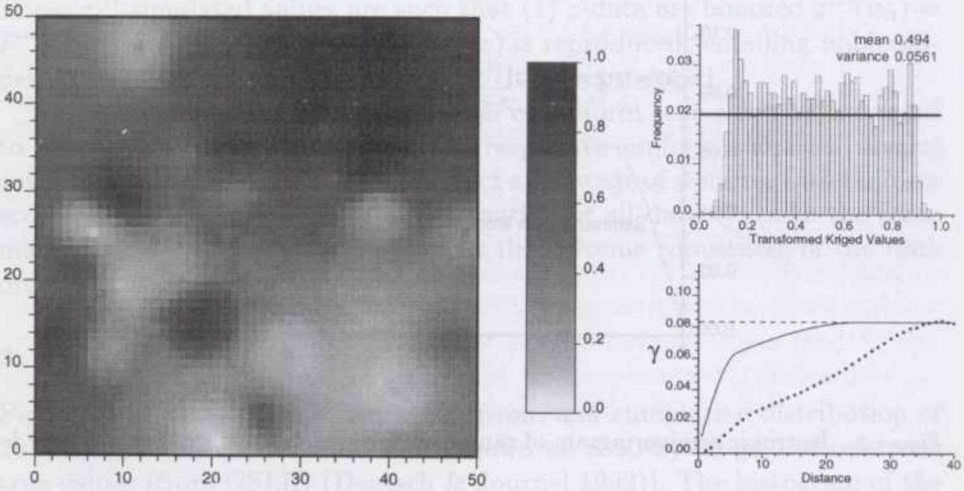

Figure 4. Gray-scale map, histogram, and semivariogram of transformed kriged values. The solid semivariogram curve is the model and the black dots are the experimental values.

reproduction, see relation (4), to yield the results illustrated on Figure 4. The map and variogram are still smooth but the histogram is much closer to the uniform V data histogram.

As described above in Section 2.2, a sequential uniform simulation can be considered where the simulated values are drawn from a uniform distribution with mean equal to the kriged rank $v^*(\mathbf{u})$ and variance equal

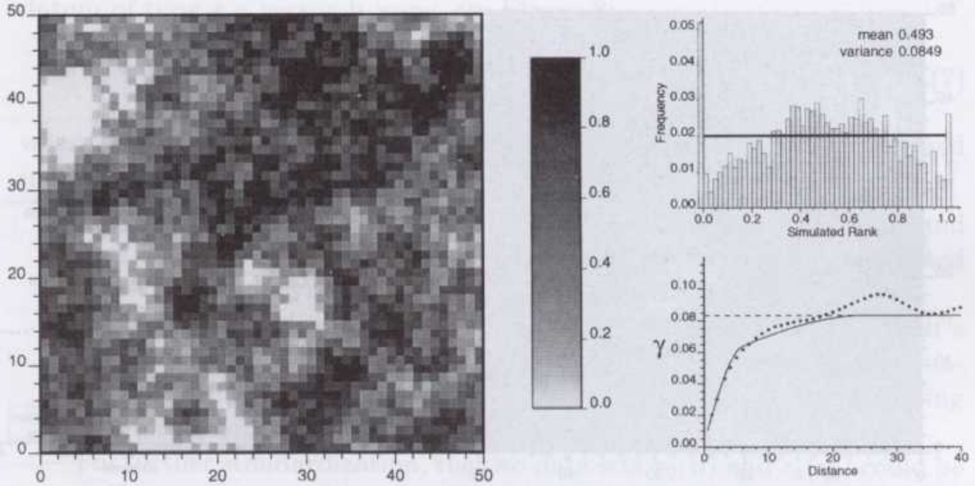

Figure 5. Gray-scale map, histogram, and semivariogram of simulated rank values. Simulated values are from sequential uniform simulation.

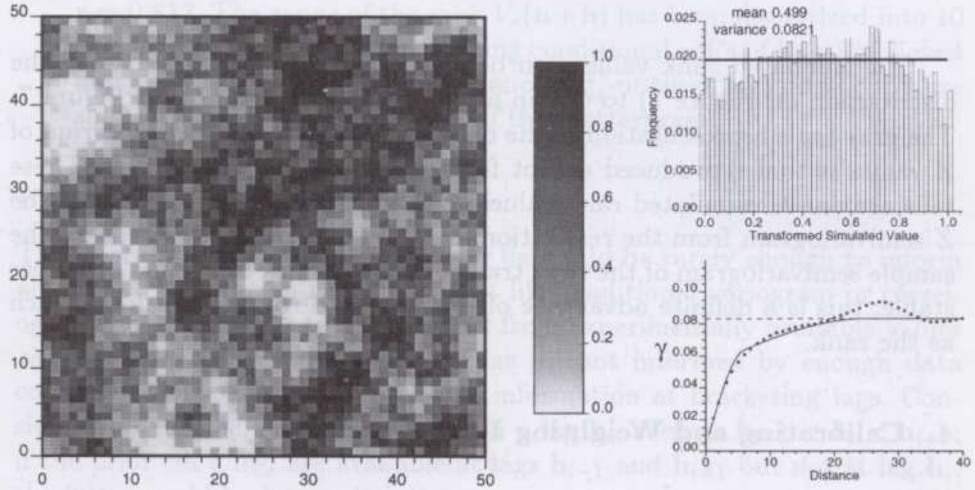

Figure 6. Gray-scale map, histogram, and semivariogram of transformed simulated rank values.

to the kriging variance $\sigma_K^2(\mathbf{u})$, i.e., from the interval $[v^*(\mathbf{u}) \pm a/2]$ with $a = \sqrt{12} \cdot \sigma_K(\mathbf{u})$. Figure 5 shows the results. As predicted by theory, (Journel 1993), the covariance and variance is reproduced. The histogram is not reproduced; there are simulated values less than 0.0 and greater than 1.0 (reset to 0.0 or 1.0 in the histogram display of Figure 5). Transforming the simulated values, under the constraint of data reproduction, yields the results illustrated on Figure 6. The map and variogram show little change but the histogram is now close to the uniform distribution.

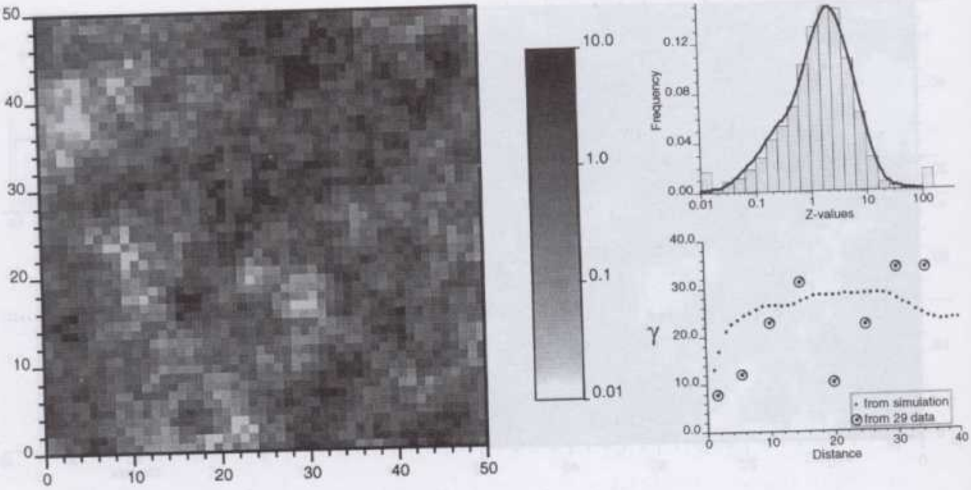

Figure 7. Gray-scale map, histogram, and semivariogram of Z simulated values.

The simulated rank values can be back transformed (according to the distribution on Figure 1) to obtain a realization of Z values; see Figure 7. The gray scale representation of the realization is similar. The histogram of Z values is well reproduced except for the tails of the distribution. These tails are due to simulated rank values outside of the interval $[0.0, 1.0]$. The Z semivariogram from the realization and the 29 data are also shown. The sample semivariogram of the rank transform, refer back to Figure 2, is more stable; this is a definite advantage of working with robust transforms such as the rank.

4. Calibrating and Weighting Information

A datum value, say $z_s(\mathbf{u}+\mathbf{h})$, no matter its type s, whether categorical or continuous, is relevant to estimation of an unknown, say $z_0(\mathbf{u})$ at location $\mathbf{u}$ a vector $\mathbf{h}$ away, *only* if its dependence to that specific unknown has been established through some prior model. Traditionally that model takes the form of a (cross)variogram or correlogram model $\rho_{0,s}(\mathbf{h})$ between the two random variables $Z_0(\mathbf{u})$ and $Z_s(\mathbf{u}+\mathbf{h})$. That correlogram model is fitted from a few available $\mathbf{h}$-scattergrams built from available pairs of data $\{z_0(\mathbf{u}_i), z_s(\mathbf{u}'_i)\}$ approximately separated by the same vector $\mathbf{h}$, see Figure 8. We suggest that there is much more valuable information to retain from any experimental $\mathbf{h}$-scattergram than the mere coefficient of correlation $\rho_{0,s}(\mathbf{h})$. For example, after proper scattergram smoothing (Deutsch, 1996), one can extract from a $\mathbf{h}$-scattergram various prior distributions for $Z_0(\mathbf{u})$ given a datum of type s a vector $\mathbf{h}$ away, see Figure 8:

$$w_{0s}(t|z_s, \mathbf{h}) = Prob\{Z_0(\mathbf{u}) \leq t | Z_s(\mathbf{u}+\mathbf{h}) = z_s\} \quad \mathbf{u} \in A \tag{7}$$

where A is the stationary area over which the $\mathbf{h}$-scattergram is deemed representative.

- $w_{0s}(\cdot)$ is a cdf, hence is valued in $[0, 1]$. It is a function of both z_s and $\mathbf{h}$. The subscript $0s$ recalls that it carries information of type s related to a variable of type 0.
- In soft indicator kriging, Deutsch and Journel (1992, p. 85), prior cdf's are also used but they are limited to $\mathbf{h} = 0$, i.e., to co-located information. The prior cdf's (7) generalize the concept of probability coding of information to all pairs $0s$ and all separation vectors $\mathbf{h}$.
- For further standardization, the two data sets $z_0(\mathbf{u})$ and $z_s(\mathbf{u}')$ could be uniform score-transformed prior to derivation of prior rank cdf's of type (7). Figure 8 shows a $\mathbf{h}$-scattergram of the standardized ranks (uniform scores) of two variables $Z_0(\mathbf{u})$ and $Z_s(\mathbf{u}+\mathbf{h})$; the rank correlation is $r = 0.812$. The range of the rank $V_s(\mathbf{u}+\mathbf{h})$ has been discretized into 10 decile classes and the corresponding conditional cdf's of $V_0(\mathbf{u})$ retrieved and shown at the right; the variances σ^2_{0s} of these two prior cdf's are shown by the bar chart on top of the scattergram.

4.1. INTERPOLATING PRIOR CDF'S

The information available on the pair $0s$ would be rarely enough to inform all possible separation vectors $\mathbf{h}$. Just like traditional variogram or correlogram models $\rho_{0s}(\mathbf{h})$ are interpolated from experimentally available values $\hat{\rho_{0s}}(\mathbf{h}_i)$, prior cdf's $w_{0s}(t|z_s, \mathbf{h})$ at lags $|\mathbf{h}|$ not informed by enough data could be interpolated from available information at bracketing lags. Consider, for example, in a given direction a lag $\mathbf{h}_i$, such as $\mathbf{h}_{i-1} < \mathbf{h}_i < \mathbf{h}_{i+1}$; if the prior cdf's w_{0s} are available at lags $\mathbf{h}_{i-1}$ and $\mathbf{h}_{i+1}$ but not at lag $\mathbf{h}_i$, the latter could be interpolated as:

$$w_{0s}(t|z_s, \mathbf{h}_i) = \frac{1}{2}\left[w_{0s}(t|z_s, \mathbf{h}_{i-1}) + w_{0s}(t|z_s, \mathbf{h}_{i+1})\right] \quad \forall t, z_s$$

4.2. UPDATING PRIOR CDF'S

Consider estimation of the unknown $z_0(\mathbf{u})$. Relevant information arises from diverse data types $s = 1, \ldots, S$ at different neighboring locations $\mathbf{u}_{\alpha_s}$. Let these data be:

$$z_s(\mathbf{u}_{\alpha_s}), \alpha_s = 1, \ldots, n_s \; ; \; s = 1, \ldots, S$$

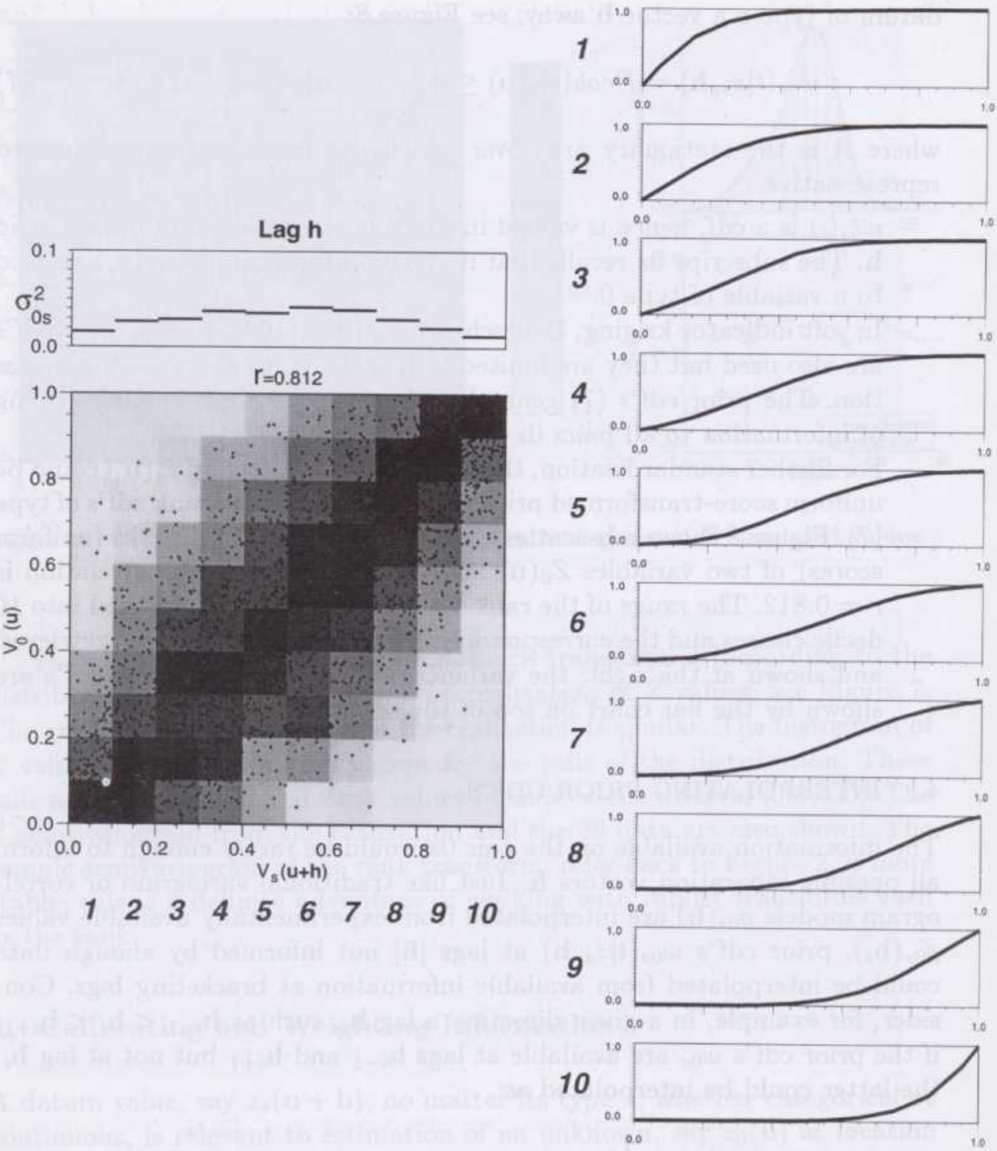

Figure 8. **h**-scattergram of $v_0(\mathbf{u})$ vs. $v_s(\mathbf{u}+\mathbf{h})$. The two variables are uniform scores of original variables $z_0(\mathbf{u})$, $z_s(\mathbf{u}+\mathbf{h})$. The range of variability of $v_s(\mathbf{u}+\mathbf{h})$ has been divided into 10 classes of equal probability. The 10 conditional cdf's of $v_0(\mathbf{u})$ given v_s are given to the right. The corresponding ten conditional variances are given on top of the scattergram.

The previous exercise of information calibration and interpolation yields the $n = \sum_{s=1}^{S} n_s$ prior cdf's:

$$w_{0s}(t|z_s(\mathbf{u}_{\alpha_s}), \mathbf{u}_{\alpha_s} - \mathbf{u}), \quad \alpha_s = 1, \ldots, n_s \; ; \; s = 1, \ldots, S$$

The problem is to update these n prior cdf's into a single "posterior" cdf for the unknown $Z_0(\mathbf{u})$. Short of a rigorous updating which would require knowledge of the joint $(S+1)$-variate distribution of $Z_0(\mathbf{u}), Z_s(\mathbf{u}_{\alpha_s})$, a weighting of the n prior cdf's should possess the following properties:

1. result in a licit posterior cdf, i.e., a non-decreasing function of the threshold t,
2. preserve hard data exactitude, i.e., at any datum location $\mathbf{u} = \mathbf{u}_{\alpha_s}$ where the prior cdf $w_{0s}(\cdot)$ has zero variance $\sigma_{0s}^2(z_s(\mathbf{u}_{\alpha_s}), \mathbf{u}_{\alpha_s} - \mathbf{u}) = 0$, then the posterior cdf should identify that prior cdf.
3. give more weight to those prior cdf's corresponding to original data $z_s(\mathbf{u}_{\alpha_s})$ most "related" to the unknown $z_0(\mathbf{u})$; a measure of that relation could be the coefficient of correlation itself $r_{0s}(\mathbf{u}_{\alpha_s} - \mathbf{u})$ directly available from the calibration stage, see Figure 8.
4. account for information redundancy: two prior cdf's might correspond to extremely redundant data, hence should not contribute both to the posterior cdf at the detriment of a third "independent" prior-cdf.

The following linear weighting scheme features the 3 first properties:

$$Prob\{Z_0(\mathbf{u}) \leq t|(n)\} = \sum_{s=1}^{S} \sum_{\alpha_s=1}^{n_s} \lambda_{\alpha_s} w_{0s}(t|z_s(\mathbf{u}_{\alpha_s}), \mathbf{h}_{\alpha_s}) \qquad (8)$$

$$\text{with : } \lambda_{\alpha_s} = \frac{1}{C} \cdot \frac{r_{0s}^2(\mathbf{h}_{\alpha_s})}{\sigma_{0s}^2(z_s(\mathbf{u}_{\alpha_s}), \mathbf{h}_{\alpha_s})} \geq 0, \mathbf{h}_{\alpha_s} = \mathbf{u}_{\alpha_s} - \mathbf{u}$$

C being a constant ensuring $\sum \lambda_{\alpha_s} = 1$

- The posterior cdf (8) being a positive linear combination of cdf's, with $\sum \lambda_{\alpha_s} = 1$, is a licit cdf, see Figure 9.
- At a hard datum location $\mathbf{u}_{\alpha_s}$ where $\sigma_{0s}^2(z_s(\mathbf{u}_{\alpha_s}), \mathbf{h}_{\alpha_s}) = 0$, $\lambda_{\alpha_s} = 1$, hence the posterior cdf is identified to the corresponding prior cdf, i.e., $Z_0(\mathbf{u}) = E\{Z_0(\mathbf{u})|z_s(\mathbf{u}_{\alpha_s})\}$ with probability one.

The prior cdf's corresponding to original data $Z_s(\mathbf{u}_{\alpha_s})$ with largest rank correlation $r_{0s}^2(\mathbf{h}_{\alpha_s})$ with $Z_0(\mathbf{u})$ receive greater weight. Note that any negative correlation between $Z_0(\mathbf{u})$ and $Z_s(\mathbf{u}_{\alpha_s})$ is already factored in the calibration of the prior cdf $w_{0s}(t|z_s(\mathbf{u}_{\alpha_s}), \mathbf{h}_{\alpha_s})$.

Unfortunately, the weighting system (8) does not account for redundancy between the n original data $z_s(\mathbf{u}_{\alpha_s})$ as would, for example, a kriging system. One could consider using the weights of cokriging of $z_0(\mathbf{u})$ using

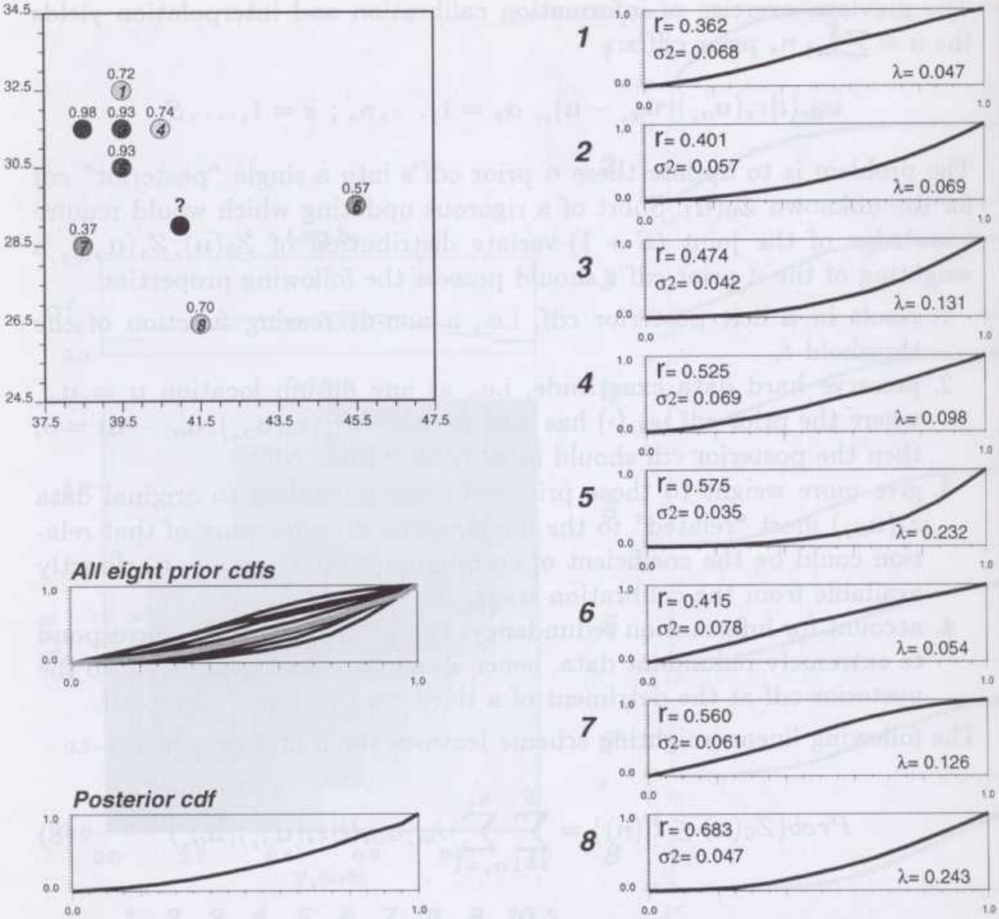

Figure 9. Weighting prior cdfs into a posterior cdf for the unsampled value at location marked by a question mark. The 8 prior cdfs are given at the right together with their variance σ^2 and rank correlation r with the unknown. The 8 original data all relate to the variable being estimated, their uniform scores are given on the location map at the top left.

the original data $z_s(\mathbf{u}_{\alpha_s})$; this would require prior modeling of the matrix of (cross)covariance functions relating the $(S+1)$ variables $Z_0(\mathbf{u}), Z_s(\mathbf{u})$. The advantage of this latter approach is that the E-type estimate (conditional expectation) resulting from the posterior cdf would identify the direct cokriging estimate of $z_0(\mathbf{u})$.

The distinct feature of the approach (8) to deriving posterior cdf's is speed: the only substantial cost is establishing the various prior cdf's $w_{0s}(\cdot)$ to be used repetitively at any node $\mathbf{u}$. This prior calibration step is similar to that of establishing a matrix of crosscovariance functions to be used for kriging at any node.

The convex weighting system (8) entails that the resulting posterior conditional variance is no lesser than the smallest prior variance. Linear weighting of prior cdf's allows weighting the influence of each datum taken one at a time, it does not account for the multivariate ($N \geq 3$) effects such as information 2 improving considerably on information 1 relative to variable 0. One could extend the concept of prior cdf's to conditioning 2 or more data values such as:

$$w_{0ss'}(t|z_s, z_{s'}, \mathbf{h}, \mathbf{h}') = Prob\{Z_0(\mathbf{u}) \leq t | Z_s(\mathbf{u}+\mathbf{h}) = z_s, Z_{s'}(\mathbf{u}+\mathbf{h}') = z_{s'}\}$$

The problem becomes then one of inference of such multiple points or multiple events probabilities, Guardiano and Srivastava (1992).

5. Conclusions

Integration of information should start by a common coding of the diverse data types. We suggest using the standardized ranks of any ordered variable, or uniform score transform, which standardizes all data types to the same univariate uniform distribution in $[0, 1]$. Kriging and/or stochastic simulation can be performed in that uniform space and the results back-transformed to the original space. Prior probability distributions as obtained from sample $\mathbf{h}$-scattergrams also represent standardized data valued in $[0, 1]$; these can be weighted to produce a posterior probability distribution for the unknown suitable for estimation or stochastic simulation. It is suggested that prior distributions be derived not only from co-located ($\mathbf{h} = 0$) data pairs but also from those $\mathbf{h}$-scattergrams ($\mathbf{h} \neq 0$) sufficiently informed. Prior cdf's at other lags $\mathbf{h}$ can be interpolated just like experimental sample variograms at specific lags $\mathbf{h}$ are fitted to yield a model $\gamma(\mathbf{h})$ valid for all $\mathbf{h}$.

References

Deutsch, C. (1994). Constrained modeling of histograms and cross plots with simulated annealing, *Accepted by Technometrics for Publication* .

Deutsch, C. & Journel, A. (1992). *GSLIB: Geostatistical Software Library and User's Guide*, Oxford University Press, New York.

Guardiano, F. & Srivastava, R. M. (1993). Multivariate geostatistics: Beyond bivariate moments, *in* A. Soares (ed.), *Geostatistics Troia 1992*, Vol. 1, Kluwer, pp. 133–144.

Journel, A. (1993). Modeling uncertainty: Some conceptual thoughts, *in* Dimitrakopoulos (ed.), *Geostatistics for the Next Century*, Kluwer, Dordrecht, Holland, pp. 30–43.

Journel, A. & Xu, W. (1994). Posterior identification of histograms conditional to local data, *Math Geology* **26**: 323–359.

Rao, S. & Journel, A. G. (1996). Deriving conditional distributions from ordinary kriging, *Fifth International Geostatistics Congress*, Wollongong.