A Step by Step Guide to Bi-Gaussian Disjunctive Kriging

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

Abstract

The Disjunctive Kriging formalism has been implemented for a number of tasks in geostatistics. Despite the advantages of this formalism, application has been hindered by presentations that assume a fairly advanced level of mathematical familiarity. This paper will go through the steps to perform Disjunctive Kriging in a simple case.

The global stationary distribution of the variable under consideration is fit by Hermite polynomials. Disjunctive Kriging amounts to simple kriging to estimate the polynomial values at an unsampled location. These estimated values completely define the local distribution of uncertainty. It is straightforward to implement this formalism in computer code, but a clear exposition of the theoretical details is required for confident application and any future development.

Introduction

Disjunctive Kriging (DK) has been available for more than 25 years, however the seemingly complex theory makes it unappealing for most practitioners. DK is a technique that provides advantages in many applications. It can be used to estimate the value of any function of the variable of interest, making it useful to assess truncated statistics for recoverable reserves. DK provides a solution space larger than the conventional kriging techniques that only rely on linear combinations of the data. DK is more practical than the conditional expectation, since it only requires knowledge of the bivariate law, instead of the full multivariate probability law of the data locations and location being estimated [7, 9, 11, 12, 14].

The theoretical basis of DK is sound, internally consistent, and has been extensively developed and expanded, among geostatisticians [1, 2, 4, 8, 10, 13]. In practice, those developments have not been applied to their full potential. DK has been applied mainly with the use of Hermite polynomials and the bivariate Gaussian assumption [6, 15]. The discrete Gaussian model for change of support has been used in practice. Still, relatively few practitioners have mastered DK. The discomfort of many practitioners is due in part to the difficult literature focussed on theory rather than applications. The available theoretical work uses complicated notation and the steps are explained for those very comfortable with mathematics. There is a need for detailed documentation of DK with emphasis on

implementation and practical details. This work aims to present DK in a rigorous manner, with greater focus on its practical aspects.

We start by presenting some background on Hermite polynomials, the bivariate Gaussian assumption, and then introduce DK with an example. More extensive theory can be found in Chilès and Delfiner [3], Emery [5], and Rivoirard [14].

Hermite Polynomials

Before getting into DK, we need to define and review some of the properties of Hermite polynomials. This family of polynomials is important because it will help us parameterize conditional distributions later on.

Hermite polynomials are defined by Rodrigues' formula:

$$H_n(y) = \frac{1}{\sqrt{n!} \cdot g(y)} \cdot \frac{d^n g(y)}{dy^n} \qquad \forall n \geq 0 \tag{1}$$

where n is the degree of the polynomial, $\sqrt{n!}$ is a normalization factor, y is a Gaussian or normal value, and $g(y)$ is the standard Gaussian probability distribution function (pdf) defined by

$$g(y) = \frac{1}{\sqrt{2\pi}} \cdot e^{-y^2/2}$$

For a given value of y the polynomial of degree n, $H_n(y)$, can easily be calculated and will result in a specific number (**Equation 1**).

A useful recursive expression exists to calculate polynomials of higher orders. This recursion is important for computer implementation:

$$H_{n+1}(y) = -\frac{1}{\sqrt{n+1}} \cdot y \cdot H_n(y) - \sqrt{\frac{n}{n+1}} \cdot H_{n-1}(y) \qquad \forall n \geq 1 \tag{2}$$

This expression along with the knowledge of the first two polynomials is enough for fast calculation up to any order. The first three polynomials are given as an example:

$$\begin{aligned} H_0(y) &= 1 \\ H_1(y) &= -y \\ H_2(y) &= \frac{1}{\sqrt{2}} \cdot (y^2 - 1) \end{aligned} \tag{3}$$

These polynomials have the following properties:

1. Mean of $H_n(Y)$ is 0, except for the polynomial of order 0, which has a mean of 1.
2. Variance of $H_n(Y)$ is 1, except again for the polynomial of order 0 which is constant and therefore its variance is 0.
3. Covariance between $H_n(Y)$ and $H_p(Y)$ is 0 if $n \neq p$. This property is known as *orthogonality* and can be understood in the same manner as the factors and principal components in multivariate statistical analysis; they correspond to uncorrelated components of a function of Y. Of course, if $n = p$ the covariance becomes the variance of $H_n(Y)$. A covariance of zero is sufficient for full independence if the multivariate distribution is Gaussian.

Hermite polynomials form an orthonormal basis with respect to the standard normal distribution, other polynomials families can be considered if a different transformation of the original variable is performed [3, 5].

Bivariate Gaussian Assumption

Consider the variable Y distributed in space. We can define the random function model $\{Y(\mathbf{u}), \forall \mathbf{u} \in Domain\}$, where $\mathbf{u}$ is a location vector in the three-dimensional space. Taking a pair of random variables $Y(\mathbf{u})$ and $Y(\mathbf{u}+\mathbf{h})$, we say they are bivariate Gaussian if their joint distribution is normal with mean vector μ and variance-covariance matrix $\mathbf{\Sigma}$. In this case, we can write:

$$(Y(\mathbf{u}), Y(\mathbf{u}+\mathbf{h})) \sim N_2(\mu, \mathbf{\Sigma})$$

$$\mu = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \qquad \mathbf{\Sigma} = \begin{pmatrix} 1 & \rho(\mathbf{h}) \\ \rho(\mathbf{h}) & 1 \end{pmatrix}$$

Notice that these two terms, the mean vector and variance-covariance matrix, fully define the bivariate Gaussian distribution of $Y(\mathbf{u})$ and $Y(\mathbf{u}+\mathbf{h})$. The correlogram $\rho(\mathbf{h})$ gives all the structural information of the bivariate relationship.

Under this assumption, one additional property of Hermite polynomials is of interest. The covariance between polynomials of different order is always 0, and if the order is the same, it identifies the correlation raised to the polynomial's degree power, that is:

$$Cov\{H_n(Y(\mathbf{u})), H_p(Y(\mathbf{u}+\mathbf{h}))\} = \begin{cases} (\rho(\mathbf{h}))^n & \text{if } n = p \\ 0 & \text{if } n \neq p \end{cases} \tag{4}$$

The only term that is left is the covariance between polynomial values of the same degree for locations separated by a vector $\mathbf{h}$. Since $\rho(\mathbf{h}) < \rho(0) = 1$, this spatial correlation tends rapidly to zero as the power n increases, that is, the structure tends to pure nugget. Also, notice that there is no spatial correlation between polynomials of different orders.

Fitting a Function with Hermite Polynomials

Any function with finite variance can be fitted by an infinite expansion of Hermite polynomials in the same way that this could be done using a Taylor or Fourier expansion. Hermite polynomials are used instead of using the derivatives of the function or sines and cosines. The idea is to express the function of y as an infinite sum of weighted polynomial values.

$$\begin{aligned} f(y(\mathbf{u})) &= f_0 + f_1 \cdot H_1(y(\mathbf{u})) + f_2 \cdot H_2(y(\mathbf{u})) + \cdots \\ &= \textstyle\sum_{n=0}^{\infty} f_n \cdot H_n(y(\mathbf{u})) \end{aligned} \tag{5}$$

The only question that remains is how to find the coefficients f_n, $\forall n$. This can be done by calculating the expected value of the product of the function and the polynomial of degree n:

$$\begin{aligned} E\{f(Y(\mathbf{u})) \cdot H_n(Y(\mathbf{u}))\} &= E\{\textstyle\sum_{p=0}^{\infty} f_p \cdot H_p(Y(\mathbf{u})) \cdot H_n(Y(\mathbf{u}))\} \\ &= \textstyle\sum_{p=0}^{\infty} f_p \cdot E\{H_p(Y(\mathbf{u})) \cdot H_n(Y(\mathbf{u}))\} \end{aligned} \tag{6}$$

The expected value can be taken inside the summation, since it is a linear operator and the coefficients f_p are constants.

Notice that the expected value of the product of polynomials of different degrees corresponds to their covariance. The property of orthogonality comes in so that all terms where $p \neq n$ equal zero and we only have the ones where $p = n$. In these cases, the covariance becomes the variance that equals 1. We can then simplify Equation 6 and get an expression to calculate the coefficients f_n:

$$f_n = E\{f(Y(\mathbf{u})) \cdot H_n(Y(\mathbf{u}))\}$$

We can rewrite this expression and write the expected value in its integral form, which will later be discretized for numerical calculation:

$$f_n = E\{f(Y(\mathbf{u})) \cdot H_n(Y(\mathbf{u}))\} = \int f(y(\mathbf{u})) \cdot H_n(y(\mathbf{u})) \cdot g(y(\mathbf{u})) \cdot dy(\mathbf{u}) \tag{7}$$

It is worth noting that the coefficient of 0 degree corresponds to the mean of the function of the random variable. This can be seen directly from Equation 7, since, if $n = 0$, then $H_n(Y(\mathbf{u})) = 1$ and the integral becomes the definition of the expected value of $f(Y(\mathbf{u}))$. A second point of interest is that the variance of the function of $Y(\mathbf{u})$ can also be calculated (although we are not going to show it here) and corresponds to the infinite sum of squared coefficients [14]. In summary:

$$E\{f(Y(\mathbf{u}))\} = f_0 \tag{8}$$

$$Var\{f(Y(\mathbf{u}))\} = \sum_{n=1}^{\infty} (f_n)^2 \tag{9}$$

The practical implementation of this expansion calls for some simplifications: the infinite expansion is truncated at a given degree P. The truncation causes some minor problems, such as generating values outside the range of the data. These values can simply be reset to a minimum or maximum values. If the number of polynomials used is large enough, these problems are of limited impact.

Practical Implementation

Fitting a Global Distribution

Consider the normal score transform (also known as the Gaussian anamorphosis in most DK literature) of a variable Z with N available samples at locations $\mathbf{u}_\alpha$, $\alpha = 1, ..., N$. The cumulative distribution function (cdf) of Z is denoted $F_Z(z)$:

$$y = G^{-1}(F_Z(z)) \quad \Leftrightarrow \quad z = F_Z^{-1}(G(y)) = \phi(y)$$

where $\phi = F_Z^{-1} \circ G$, in shortened notation, is the anamorphosis function.

This normal score transformation is often done through a one to one table lookup that relates the sample data $z(\mathbf{u}_\alpha)$ with corresponding quantiles of the standard normal distribution $G(y)$ (Figure 1). As an approximation to this quantile matching procedure, we can parametrize the relationship with a finite number of coefficients. The Hermitian expansion

of this function up to a degree P is used to give a close approximation to the shape of the function ϕ.

$$Z(\mathbf{u}) = \phi(Y(\mathbf{u})) \approx \sum_{p=0}^{P} \phi_p \cdot H_p(Y(\mathbf{u}))$$

Notice how we use ϕ as a function, like we used f in the previous section. We know how to calculate the coefficients:

$$\begin{array}{rcl} \phi_0 & = & E\{\phi(Y(\mathbf{u}))\} = E\{Z(\mathbf{u})\} \\ \phi_p & = & E\{Z(\mathbf{u}) \cdot H_p(Y(\mathbf{u}))\} = \int \phi(y(\mathbf{u})) \cdot H_p(y(\mathbf{u})) \cdot g(y(\mathbf{u})) \cdot dy(\mathbf{u}) \end{array} \tag{10}$$

The last expression can be approximated with the data at hand, as a finite summation. Consider the data values $y(\mathbf{u}_1), y(\mathbf{u}_2), ..., y(\mathbf{u}_N)$ sorted increasingly. The terms $\phi(y(\mathbf{u}))$ become $z(\mathbf{u})$. Since we have N data, we approximate the integral as a sum of N elements:

$$\begin{array}{rcl} \phi_p & = & \displaystyle\sum_{\alpha=1}^{N} \int_{y(\mathbf{u}_\alpha)}^{y(\mathbf{u}_{\alpha+1})} z(\mathbf{u}_\alpha) \cdot H_p(y(\mathbf{u})) \cdot g(y(\mathbf{u})) \cdot dy(\mathbf{u}) \\ & = & \displaystyle\sum_{\alpha=1}^{N} \int_{y(\mathbf{u}_\alpha)}^{y(\mathbf{u}_{\alpha+1})} z(\mathbf{u}_\alpha) \cdot \frac{1}{\sqrt{p!} \cdot g(y(\mathbf{u}))} \cdot \frac{d^p g(y(\mathbf{u}))}{dy(\mathbf{u})^p} \cdot g(y(\mathbf{u})) \cdot dy(\mathbf{u}) \\ & = & \displaystyle\sum_{\alpha=1}^{N} z(\mathbf{u}_\alpha) \cdot \frac{1}{\sqrt{p!}} \cdot \int_{y(\mathbf{u}_\alpha)}^{y(\mathbf{u}_{\alpha+1})} \frac{d}{dy(\mathbf{u})} \left(\frac{d^{p-1} g(y(\mathbf{u}))}{dy(\mathbf{u})^{p-1}} \right) \cdot dy(\mathbf{u}) \\ & = & \displaystyle\sum_{\alpha=1}^{N} z(\mathbf{u}_\alpha) \cdot \frac{1}{\sqrt{p!}} \cdot \int_{y(\mathbf{u}_\alpha)}^{y(\mathbf{u}_{\alpha+1})} \frac{d}{dy(\mathbf{u})} \left(H_{p-1}(y(\mathbf{u})) \cdot \sqrt{(p-1)!} \cdot g(y(\mathbf{u})) \right) \cdot dy(\mathbf{u}) \\ & = & \displaystyle\sum_{\alpha=1}^{N} z(\mathbf{u}_\alpha) \cdot \frac{1}{\sqrt{p}} \cdot \Big[H_{p-1}(y(\mathbf{u})) \cdot g(y(\mathbf{u})) \Big]_{y(\mathbf{u}_\alpha)}^{y(\mathbf{u}_{\alpha+1})} \\ & = & \displaystyle\sum_{\alpha=1}^{N} z(\mathbf{u}_\alpha) \cdot \left(\frac{1}{\sqrt{p}} H_{p-1}(y(\mathbf{u}_{\alpha+1})) \cdot g(y(\mathbf{u}_{\alpha+1})) - \frac{1}{\sqrt{p}} H_{p-1}(y(\mathbf{u}_\alpha)) \cdot g(y(\mathbf{u}_\alpha)) \right) \end{array}$$

with $g(y(\mathbf{u}_1)) = g(-\infty) = 0$ and $g(y(\mathbf{u}_{N+1})) = g(\infty) = 0$.

This expression can be further simplified by cancelling out the term:

$$\frac{1}{\sqrt{p}} \cdot H_{p-1}(y(\mathbf{u}_{\alpha+1})) \cdot g(y(\mathbf{u}_{\alpha+1}))$$

of the α element of the summation, with the term:

$$-\frac{1}{\sqrt{p}} \cdot H_{p-1}(y(\mathbf{u}_\alpha)) \cdot g(y(\mathbf{u}_\alpha))$$

of the $(\alpha + 1)$ element of the summation, resulting in the following concise result:

$$\phi_p = \sum_{\alpha=2}^{N} (z(\mathbf{u}_{\alpha-1}) - z(\mathbf{u}_\alpha)) \cdot \frac{1}{\sqrt{p}} H_{p-1}(y(\mathbf{u}_\alpha)) \cdot g(y(\mathbf{u}_\alpha)) \tag{11}$$

In summary, all we need to fit a global distribution with Hermite polynomials is:

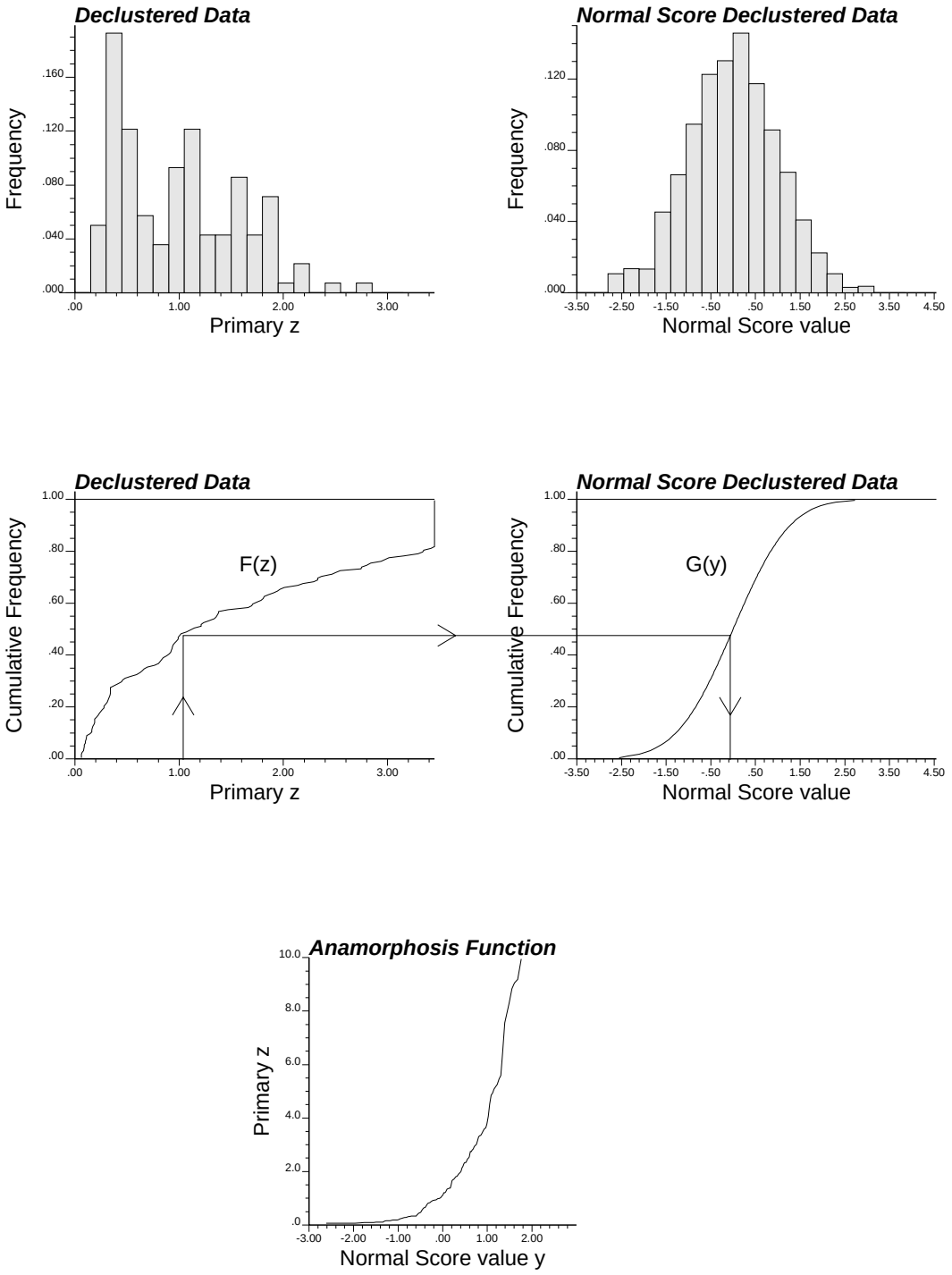

Figure 1: Graphical transformation to normal scores and anamorphosis function for a global distribution.

1. The sample data set $z(\mathbf{u}_\alpha), \alpha = 1, ..., N$.
2. The Hermite polynomial values computed for each one of the corresponding normal score transforms of the data $y(\mathbf{u}_\alpha), \alpha = 1, ..., N$.
3. The pdf value of a standard normal distribution of all the normal score transforms of the data $g(y(\mathbf{u}_\alpha)), \alpha = 1, ..., N$.

Fitting an Indicator Function

A second application is to use Hermite polynomials to fit an indicator function. Consider a threshold z_c and its corresponding normal score transform $y_c = G^{-1}(F_Z(z_c))$

We can define the indicator function for the variable Z, which is equivalent to another indicator function defined for the normal scores:

$$I_Z(\mathbf{u}; z_c) = Prob\{z(\mathbf{u}) \leq z_c\} = Prob\{y(\mathbf{u}) \leq y_c\} = I_Y(\mathbf{u}; y_c)$$

Again, we can consider expanding this function of Y up to a degree P:

$$I_Y(\mathbf{u}; y_c) = \sum_{p=0}^{P} \psi_p \cdot H_p(Y(\mathbf{u}))$$

The coefficients ψ_p are the unique coefficients defined for the expansion using Hermite polynomials for the indicator function, as the f_p were the coefficients for the expansion of the general function f in Equation 5.

The coefficients can be calculated as:

$$\begin{aligned}
\psi_0 &= G(y_c) \\
\psi_p &= E\{I_Y(\mathbf{u}; y_c) \cdot H_p(Y(\mathbf{u}))\} \\
&= \int_{-\infty}^{\infty} I_Y(\mathbf{u}; y_c) \cdot H_p(y(\mathbf{u})) \cdot g(y(\mathbf{u})) \cdot dy(\mathbf{u}) \\
&= \int_{-\infty}^{y_c} H_p(y(\mathbf{u})) \cdot g(y(\mathbf{u})) \cdot dy(\mathbf{u}) \\
&= \frac{1}{\sqrt{p}} \cdot H_{p-1}(y_c) \cdot g(y_c)
\end{aligned}$$

For clarity, let us see how this last step was done. First, recall the definition of the Hermite polynomial presented in Equation 1 and replace it in the previous equation:

$$\begin{aligned}
\psi_p &= \int_{-\infty}^{y_c} H_p(y(\mathbf{u})) \cdot g(y(\mathbf{u})) \cdot dy(\mathbf{u}) \\
&= \int_{-\infty}^{y_c} \frac{1}{\sqrt{p!} \cdot g(y(\mathbf{u}))} \cdot \frac{d^p g(y(\mathbf{u}))}{dy(\mathbf{u})^p} \cdot g(y(\mathbf{u})) \cdot dy(\mathbf{u}) \\
&= \frac{1}{\sqrt{p!}} \cdot \left(\frac{d^{p-1} g(y(\mathbf{u}))}{dy(\mathbf{u})^{p-1}} \right)_{y(\mathbf{u})=y_c} \\
&= \frac{1}{\sqrt{p!}} \cdot H_{p-1}(y_c) \cdot \sqrt{(p-1)!} \cdot g(y_c) \\
&= \frac{1}{\sqrt{p}} \cdot H_{p-1}(y_c) \cdot g(y_c)
\end{aligned}$$

It is important to emphasize that the Hermite coefficients for the indicator function are not the same as the ones shown for the continuous function. The idea behind the expansion is that for each function a different set of coefficients is found that allows us to express the function as a linear combination of Hermite polynomials. This expansion fits the function under consideration.

Application to Uniform and Lognormal Distributions

Two experimental distributions are used to illustrate the procedure. The number of Hermite polynomials used is important to get a close match between the original and fitted distributions. The general procedure to fitting a continuous function with a finite expansion of Hermite polynomials is:

- Transform the original data $z_i, i = 1, \ldots, N$ to normal scores

$$y_i = G^{-1}(F(z_i)), \qquad i = 1, \ldots, N$$

- Calculate the Hermite polynomials for all the Gaussian transformed y_i data values using Equations 2 and 3.
- Calculate the coefficients using Equations 10 and 11.
- Generate the approximate function as the linear combination of the Hermite polynomials weighted by the corresponding coefficients, as in Equation 5.

Data from a uniform distribution are presented in Figure 2, along with the approximated histogram obtained by Hermite expansion of 100 polynomials. Recall that the expansion matches the function if an infinite number of polynomials is used, therefore the larger the number of polynomials, the closer the expansion to reproducing the original distribution. It was found that any expansion over the order of 25 gave satisfactory results, matching the statistics up to the fourth significant digit.

Data from a lognormal distribution were fitted with expansions of 10, 50, and 100 polynomials, giving very good match with an expansion of order 100 (Figure 3). Notice that the approximation for high values seen in the q-q plot does not perform well. This is due to the truncation of the infinite expansion. The mean was reproduced up to the fifth significant digit, while the variance was reproduced up to the third significant digit.

Disjunctive Kriging

Disjunctive Kriging (DK) allows the estimation of any function of $Z(\mathbf{u})$, based on a bivariate probability model. A bivariate Gaussian distribution of the normal scores of the data is almost always chosen.

DK provides the solution that minimizes the estimation variance among all linear combinations of functions of one point at a time.

In simple words, DK relies on the decomposition of the variable (or a function of it) into a sum of factors. These factors are orthogonal random variables, uncorrelated with each other, and therefore the optimum estimate can be found by simple kriging each component.

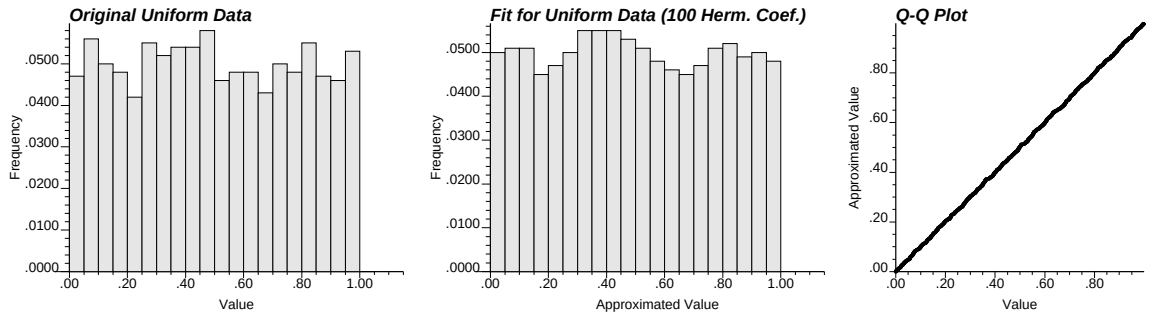

Figure 2: Original uniform distribution and corresponding reproduction of uniform distribution with 100 Hermite polynomials. A q-q plot of the original and approximated global distributions shows the excellent match.

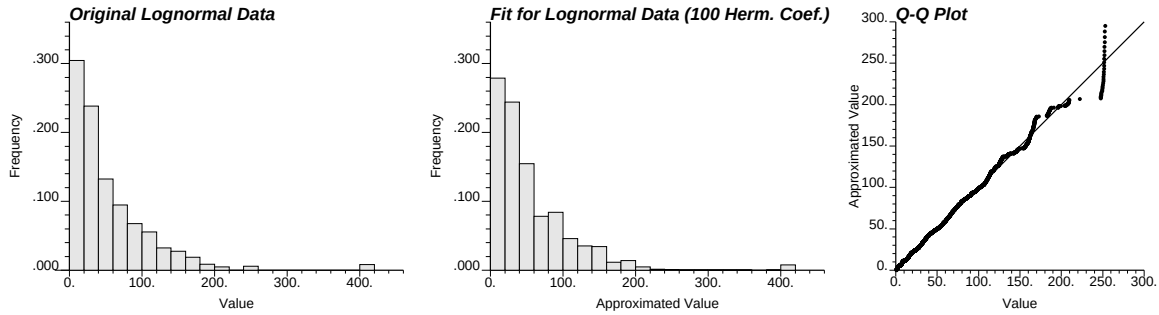

Figure 3: Original lognormal distribution and corresponding reproduction of lognormal distribution with 100 Hermite polynomials. The q-q plot shows the problem of reproduction of the upper tail due to the truncation of the infinite expansion.

Consider a random variable Z and a transformed random variable Y, in general its Gaussian transform. The disjunctive kriging estimate finds the family of functions of Y that minimizes the estimation variance. Under a particular bivariate assumption, an isofactorial family of functions can be found. Under the bivariate Gaussian assumption, this family is the Hermite polynomials. However, other transformations can be done and different orthogonal polynomials must be used. For a uniform transformation, Legendre polynomials are used. If the transformation is exponential, Laguerre polynomials are appropriate. They are called isofactorial families because they decompose the function of the random variable into factors that are independent. Although in the general case the DK estimate is obtained by simple cokriging of the functions of different order, if these are independent from each other, just a simple kriging of the functions of the same order and their posterior linear combination suffices to obtain the best estimate.

The DK estimate is presented next under the bivariate Gaussian assumption using Hermite polynomials:

$$[f(Y(\mathbf{u}))]^*_{DK} = \sum_{p=0}^{\infty} f_p \cdot [H_p(Y(\mathbf{u}))]^*_{SK}$$

The expansion is normally truncated at a degree P, usually under 100.

To calculate the DK estimate the normal score transformation of the data is necessary:

$$y(\mathbf{u}_\alpha) = G^{-1}(F_Z(z(\mathbf{u}_\alpha))) \qquad \alpha = 1, ..., N$$

Then, the spatial covariance of the transformed variable $\rho(\mathbf{h})$ is calculated and modelled. The covariance function is the correlogram because Y has unit variance.

The Hermite polynomials are computed for all the transformed data up to a degree P using Equations 2 and 3. Finally, the coefficients of the Hermitian expansion, up to a degree P, can be calculated by Equations 10 and 11.

Simple kriging is performed P times. The estimate of the Hermite polynomial at an unsampled location $\mathbf{u}_0$ is calculated as:

$$[H_p(y(\mathbf{u}_0))]^*_{SK} = \sum_{i=1}^{n(\mathbf{u}_0)} \lambda_{p,i} \cdot H_p(y(\mathbf{u}_i)) \qquad \forall p > 0$$

where $\lambda_{p,i}$ is the simple kriging weight for the data $y(\mathbf{u}_i)$ and the degree p; $n(\mathbf{u}_0)$ is the number of samples found in the search neighborhood used for kriging. Notice that the term for the mean is not present, since the mean value of the Hermite polynomial is 0, for all $p > 0$. Also, note that the SK estimate for the polynomial of degree 0 is 1, since this is its value by definition (**Equation 3**).

The weights are obtained by solving the following system of equations:

$$\begin{bmatrix} (\rho_{1,1})^p & \cdots & (\rho_{1,n(\mathbf{u}_0)})^p \\ \vdots & \ddots & \vdots \\ (\rho_{n(\mathbf{u}_0),1})^p & \cdots & (\rho_{n(\mathbf{u}_0),n(\mathbf{u}_0)^p}) \end{bmatrix} \cdot \begin{bmatrix} \lambda_{p,1} \\ \vdots \\ \lambda_{p,n(\mathbf{u}_0)} \end{bmatrix} = \begin{bmatrix} (\rho_{1,0})^p \\ \vdots \\ (\rho_{n(\mathbf{u}_0),0})^p \end{bmatrix} \tag{12}$$

We can now rewrite the DK estimate as:

$$[f(Y(\mathbf{u}_0))]^*_{DK} = \sum_{p=0}^{P} f_p \cdot \left(\sum_{i=1}^{n(\mathbf{u}_0)} \lambda_{p,i} \cdot H_p(y(\mathbf{u}_i)) \right)$$

The estimation variance is:

$$\sigma^2_{DK}(\mathbf{u}_0) = Var\{f(Y(\mathbf{u}_0)) - [f(Y(\mathbf{u}_0))]^*_{DK}\} = \sum_{p=1}^{P} (f_p)^2 \cdot \sigma^2_{SK,p}$$

where $\sigma^2_{SK,p}$ is the estimation variance from SK of the Hermite polynomials for the system of order p, that is:

$$\sigma^2_{SK,p} = \sigma^2_{H_p(Y)} - \sum_{i=1}^{n(\mathbf{u}_0)} \lambda_{p,i} \cdot (\rho_{i,0})^p$$

Since $\sigma^2_{H_p(Y)} = 1, \forall p > 0$, the estimation variance of DK can be rewritten as:

$$\sigma^2_{DK}(\mathbf{u}_0) = \sum_{p=1}^{P} (f_p)^2 \cdot \left(1 - \sum_{i=1}^{n(\mathbf{u}_0)} \lambda_{p,i} \cdot (\rho_{i,0})^p \right)$$

Example

The following example is presented to illustrate the implementation of DK. Ten sample values are considered. The sample statistics are:

$$Mean = 7.278$$
$$Variance = 24.280$$

To proceed with DK, the global distribution must be fitted with Hermite polynomials. The normal score transforms have been calculated considering only these ten values as the global distribution. Table 1 shows the original values, their transforms, the corresponding probability densities and the values for the Hermite polynomials up to the tenth degree.

Once the Hermite polynomials have been calculated, the next step is to calculate the coefficients that will be used to fit the global distribution. The coefficients in Equation 11 are calculated in Table 2. The values f_p are obtained by summing the elements in the corresponding columns. Recall that $f_0 = E\{Z\}$. This expected value is estimated with the sample mean, that is, $f_0 = 7.278$.

The global distribution is fitted with Hermite polynomials up to the tenth degree. For comparison, the fitting using 100 Hermite polynomials is also shown on Figure 4.

We are interested in estimating the value of variable Z at location $\mathbf{u}_0 = (0, 0)$ by Disjunctive Kriging. Considering a spherical covariance with range of 40 for the Gaussian transform, we can calculate the estimate at that location given the location of the nearby samples, as shown on Figure 5.

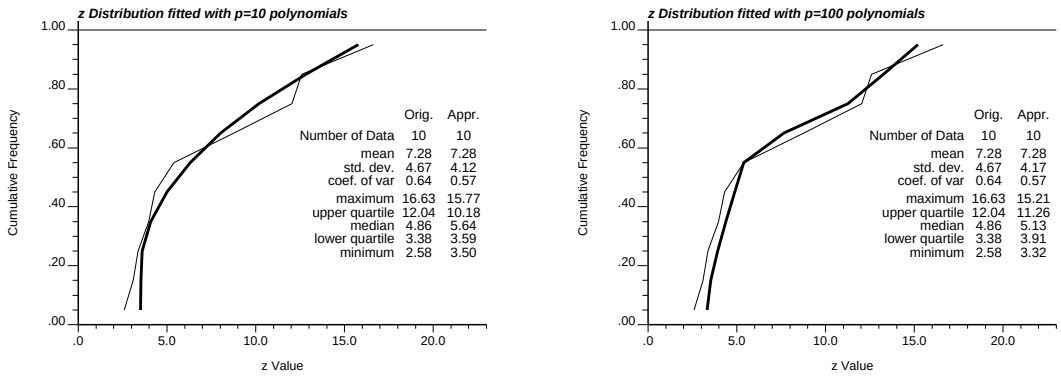

Figure 4: Fitting of the global distribution using the ten polynomials. For comparison, a fitting using 100 polynomials is shown. The fitted distribution is shown as a thick line.

Only three samples are found in the search neighborhood: $z(\mathbf{u}_1) = 3.377$ located in $\mathbf{u}_1 = (-2, 0)$; $z(\mathbf{u}_2) = 12.586$ located in $\mathbf{u}_2 = (4, 0)$; and $z(\mathbf{u}_3) = 5.398$ located in $\mathbf{u}_3 = (0, 4)$. A matrix D of absolute distance between the samples, and a vector d with the distance between the samples and the point to be estimated are built.

$$D = \begin{pmatrix} 0.00 & 6.00 & 4.47 \\ 6.00 & 0.00 & 5.66 \\ 4.47 & 5.66 & 0.00 \end{pmatrix} \qquad d = \begin{pmatrix} 2.00 \\ 4.00 \\ 4.007 \end{pmatrix}$$

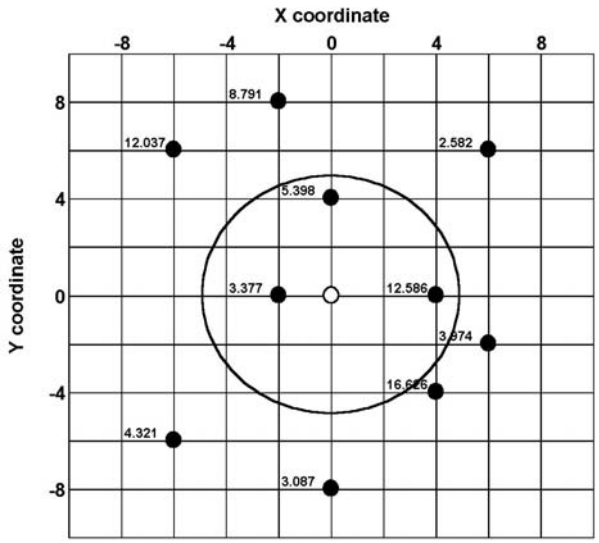

Figure 5: Data configuration. The figure shows the original z data. A search radius is defined to find the nearby samples to estimate by DK the value of $z(\mathbf{u})$ at location $\mathbf{u} = (0, 0)$. Only three samples are found in the search neighborhood.

The DK estimate is obtained by solving $P = 10$ simple kriging systems to estimate the polynomial values that will be combined using the coefficients calculated to fit the global distribution. Recall that to estimate the polynomial of order p, the covariance must be exponentiated to the corresponding degree.

Consider for example the simple kriging systems and estimates for the first three polynomials:

$$[H_1(y(\mathbf{u}_0))]^*_{SK} = \lambda_1^{p=1} \cdot H_1(y(\mathbf{u}_1)) + \lambda_2^{p=1} \cdot H_1(y(\mathbf{u}_2)) + \lambda_3^{p=1} \cdot H_1(y(\mathbf{u}_3))$$

$$p = 1: \quad \begin{pmatrix} 1.000 & 0.777 & 0.833 \\ 0.777 & 1.000 & 0.789 \\ 0.833 & 0.789 & 1.000 \end{pmatrix} \cdot \begin{pmatrix} \lambda_1^{p=1} \\ \lambda_2^{p=1} \\ \lambda_3^{p=1} \end{pmatrix} = \begin{pmatrix} 0.925 \\ 0.851 \\ 0.851 \end{pmatrix}$$

$$[H_2(y(\mathbf{u}_0))]^*_{SK} = \lambda_1^{p=2} \cdot H_2(y(\mathbf{u}_1)) + \lambda_2^{p=2} \cdot H_2(y(\mathbf{u}_2)) + \lambda_3^{p=2} \cdot H_2(y(\mathbf{u}_3))$$

$$p = 2: \quad \begin{pmatrix} (1.000)^2 & (0.777)^2 & (0.833)^2 \\ (0.777)^2 & (1.000)^2 & (0.789)^2 \\ (0.833)^2 & (0.789)^2 & (1.000)^2 \end{pmatrix} \cdot \begin{pmatrix} \lambda_1^{p=2} \\ \lambda_2^{p=2} \\ \lambda_3^{p=2} \end{pmatrix} = \begin{pmatrix} (0.925)^2 \\ (0.851)^2 \\ (0.851)^2 \end{pmatrix}$$

$$[H_3(y(\mathbf{u}_0))]^*_{SK} = \lambda_1^{p=3} \cdot H_3(y(\mathbf{u}_1)) + \lambda_2^{p=3} \cdot H_3(y(\mathbf{u}_2)) + \lambda_3^{p=3} \cdot H_3(y(\mathbf{u}_3))$$

$$p = 3: \quad \begin{pmatrix} (1.000)^3 & (0.777)^3 & (0.833)^3 \\ (0.777)^3 & (1.000)^3 & (0.789)^3 \\ (0.833)^3 & (0.789)^3 & (1.000)^3 \end{pmatrix} \cdot \begin{pmatrix} \lambda_1^{p=3} \\ \lambda_2^{p=3} \\ \lambda_3^{p=3} \end{pmatrix} = \begin{pmatrix} (0.925)^3 \\ (0.851)^3 \\ (0.851)^3 \end{pmatrix}$$

Recall that for $p = 0$ the polynomial is equal to 1 (**Equation 3**).

z value	y value	$g(y)$	Hermite polynomial $H_p(y)$										
			$p=0$	$p=1$	$p=2$	$p=3$	$p=4$	$p=5$	$p=6$	$p=7$	$p=8$	$p=9$	$p=10$
2.582	-1.645	0.103	1.000	1.645	1.206	-0.198	-1.207	-0.711	0.624	1.046	0.024	-0.973	-0.529
3.087	-1.036	0.233	1.000	1.036	0.052	-0.815	-0.468	0.512	0.644	-0.222	-0.683	-0.027	0.639
3.377	-0.674	0.318	1.000	0.674	-0.385	-0.701	0.097	0.656	0.092	-0.584	-0.225	0.500	0.320
3.974	-0.385	0.370	1.000	0.385	-0.602	-0.449	0.435	0.476	-0.322	-0.488	0.235	0.490	-0.163
4.321	-0.126	0.396	1.000	0.126	-0.696	-0.153	0.593	0.170	-0.533	-0.183	0.490	0.193	-0.457
5.398	0.126	0.396	1.000	-0.126	-0.696	0.153	0.593	-0.170	-0.533	0.183	0.490	-0.193	-0.457
8.791	0.385	0.370	1.000	-0.385	-0.602	0.449	0.435	-0.476	-0.322	0.488	0.235	-0.490	-0.163
12.037	0.674	0.318	1.000	-0.674	-0.385	0.701	0.097	-0.656	0.092	0.584	-0.225	-0.500	0.320
12.586	1.036	0.233	1.000	-1.036	0.052	0.815	-0.468	-0.512	0.644	0.222	-0.683	0.027	0.639
16.626	1.645	0.103	1.000	-1.645	1.206	0.198	-1.207	0.711	0.624	-1.046	0.024	0.973	-0.529

Table 1: Hermite polynomial values for the data up to the tenth degree

z value	y value	$g(y)$	Element of the summation $(z(\mathbf{u}_{\alpha-1}) - z(\mathbf{u}_\alpha)) \cdot \frac{1}{\sqrt{p}} H_{p-1}(y(\mathbf{u}_\alpha)) \cdot g(y(\mathbf{u}_\alpha))$ to obtain f_p									
			$p=1$	$p=2$	$p=3$	$p=4$	$p=5$	$p=6$	$p=7$	$p=8$	$p=9$	$p=10$
2.582	-1.645	0.103	—	—	—	—	—	—	—	—	—	—
3.087	-1.036	0.233	-0.118	-0.086	-0.004	0.048	0.025	-0.025	-0.029	0.009	0.027	0.001
3.377	-0.674	0.318	-0.092	-0.044	0.021	0.032	-0.004	-0.025	-0.003	0.019	0.007	-0.015
3.974	-0.385	0.370	-0.221	-0.060	0.077	0.050	-0.043	-0.043	0.027	0.038	-0.017	-0.034
4.321	-0.126	0.396	-0.137	-0.012	0.055	0.011	-0.036	-0.010	0.028	0.009	-0.022	-0.008
5.398	0.126	0.396	-0.426	0.038	0.171	-0.033	-0.113	0.030	0.086	-0.028	-0.070	0.026
8.791	0.385	0.370	-1.257	0.342	0.437	-0.282	-0.244	0.244	0.153	-0.217	-0.098	0.195
12.037	0.674	0.318	-1.032	0.492	0.230	-0.361	-0.045	0.276	-0.036	-0.213	0.077	0.163
12.586	1.036	0.233	-0.128	0.094	-0.004	-0.052	0.027	0.027	-0.031	-0.010	0.029	-0.001
16.626	1.645	0.103	-0.417	0.485	-0.290	-0.041	0.225	-0.121	-0.098	0.154	-0.003	-0.128
		f_p	-3.828	1.248	0.693	-0.629	-0.210	0.354	0.096	-0.238	-0.071	0.198

Table 2: Calculation of coefficient for fitting of global distribution

	$p=1$	$p=2$	$p=3$	$p=4$	$p=5$	$p=6$	$p=7$	$p=8$	$p=9$	$p=10$
$H_p(y(\mathbf{u}_1))$	0.674	-0.385	-0.701	0.097	0.656	0.092	-0.584	-0.225	0.500	0.320
$H_p(y(\mathbf{u}_2))$	-1.036	0.052	0.815	-0.468	-0.512	0.644	0.222	-0.683	0.027	0.639
$H_p(y(\mathbf{u}_3))$	-0.126	-0.696	0.153	0.593	-0.170	-0.533	0.183	0.490	-0.193	-0.457
λ_1^p	0.596	0.590	0.580	0.566	0.548	0.528	0.505	0.480	0.535	0.428
λ_2^p	0.287	0.281	0.272	0.259	0.244	0.227	0.209	0.190	0.234	0.153
λ_3^p	0.128	0.139	0.147	0.150	0.150	0.147	0.142	0.134	0.148	0.115
$[H_p(y(\mathbf{u}_0))]^*_{SK}$	0.088	-0.309	-0.162	0.023	0.209	0.116	-0.223	-0.172	0.245	0.183

Table 3: Simple kriging weights and hermite polynomial values. The estimated Hermite polynomials for location $\mathbf{u}_0$ are shown at the bottom row.

We solve for the weights $\lambda_\alpha^p, \alpha = 1, ..., n, p = 1, ..., P$ and estimate each one of the polynomials for location $\mathbf{u}_0$. Table 3 shows the weights, Hermite polynomials for $\mathbf{u}_1$, $\mathbf{u}_2$, and $\mathbf{u}_3$, and the estimated Hermite polynomial values for $\mathbf{u}_0$.

Finally, the DK estimate of $z(\mathbf{u}_0)$ is obtained by combining these estimated Hermite polynomial values with the coefficients for the global transformation.

This gives:

$$[z(\mathbf{u}_0)]^*_{DK} = [f(y(\mathbf{u}_0))]^*_{DK} = \sum_{p=0}^{P} f_p \cdot [H_p(y(\mathbf{u}_0))]^*_{SK} = 6.462$$

Conclusion

This paper presents the methodology to estimate the value of a regionalized variable at an unsampled location by Disjunctive Kriging. The use of the Hermite polynomials as an isofactorial family was discussed and the most fundamental equations were presented. A simple example where a single point was estimated was presented, showing all calculations required to obtain the estimated value.

DK can be applied under a variety of assumptions regarding the bivariate spatial law. These extensions have not been presented in this paper. Implementation of DK under the biGaussian assumption and with other isofactorial families of polynomials could be considered in the future.

References

[1] M. Armstrong and G. Matheron. Disjunctive kriging revisited - part I. *Mathematical Geology*, 18(8):711–728, 1986.

[2] M. Armstrong and G. Matheron. Disjunctive kriging revisited - part II. *Mathematical Geology*, 18(8):729–742, 1986.

[3] J. P. Chilès and P. Delfiner. *Geostatistics Modeling Spatial Uncertainty*. John Wiley & Sons, New York, 1999.

[4] X. Emery. Conditional simulation of non-Gaussian random functions. *Mathematical Geology*, 34(1):79–100, 2002.

[5] X. Emery. *Simulación Estocástica y Geoestadística No Lineal*. To Be Published by Departamento de Ingeniería de Minas - Universidad de Chile, 487 pages, 2003.

[6] D. Guibal and A. Remacre. Local estimation of the recoverable reserves: Comparing various methods with the reality of a porphyry copper deposit. In G. Verly, M. David, A. G. Journel, and A. Marechal, editors, *Geostatistics for natural resources characterization*, volume 1, pages 435–448. Reidel, Dordrecht, Holland, 1984.

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

[8] A. Marechal. Recovery estimation: A review of models and methods. In G. Verly, M. David, A. G. Journel, and A. Marechal, editors, *Geostatistics for natural resources characterization*, volume 1, pages 385–420. Reidel, Dordrecht, Holland, 1984.

[9] G. Matheron. Le krigeage disjonctif. Internal note N-360, Centre de Géostatistique, Fontainebleau, 40 pages, 1973.

[10] G. Matheron. Les fonctions de transfert des petits panneaux. Internal note N-395, Centre de Géostatistique, Fontainebleau, 73 pages, 1974.

[11] G. Matheron. Forecasting block grade distributions: The tranfer function. In M. Guarascio, M. David, and C. Huijbregts, editors, *Advanced Geostatistics in the Mining Industry*, pages 237–251, Dordrecht, Holland, 1976. Reidel.

[12] G. Matheron. A simple substitute for conditional expectation: the disjunctive kriging. In M. Guarascio, M. David, and C. Huijbregts, editors, *Advanced Geostatistics in the Mining Industry*, pages 221–236, Dordrecht, Holland, 1976. Reidel.

[13] G. Matheron. Isofactorial models and change of support. In G. Verly, M. David, A. G. Journel, and A. Marechal, editors, *Geostatistics for natural resources characterization*, volume 1, pages 449–467. Reidel, Dordrecht, Holland, 1984.

[14] J. Rivoirard. *Introduction to Disjunctive Kriging and Non-Linear Geostatistics*. Oxford University Press, 181 pages, New York, 1994.

[15] R. Webster and M. A. Oliver. Disjunctive kriging in agriculture. In M. Armstrong, editor, *Geostatistics*, volume 1, pages 421–432. Kluwer, 1989.