KRIGING COORDINATES: WHAT DOES THAT MEAN?

R. Tolosana-Delgado and V. Pawlowsky-Glahn
tolosana@ima.udg.es, Departament d'Informàtica i Matemàtica Aplicada
Universitat de Girona, E-17071, Girona, Spain

1 Introduction

Kriging is an interpolation technique whose optimality criteria are based on normality assumptions either for observed or for transformed data. This is the case of normal, lognormal and multigaussian kriging. When kriging is applied to transformed scores, optimality of obtained estimators becomes a cumbersome concept: back-transformed optimal interpolations in transformed scores are not optimal in the original sample space, and vice-versa. This lack of compatible criteria of optimality induces a variety of problems in both point and block estimates. For instance, lognormal kriging, widely used to interpolate positive variables, has no straightforward way to build consistent and optimal confidence intervals for estimates. These problems are ultimately linked to the assumed space structure of the data support: for instance, positive values, when modelled with lognormal distributions, are assumed to be embedded in the whole real space, with the usual real space structure and Lebesgue measure.

Here we present an alternative approach to the usual transformation strategy, which we call coordinate kriging: instead of working with the original space of observations as a subset or real space, and to consider transformations as functions going from real space to itself, we propose to use whenever possible a sensible (from the geostatistician's point of view) Euclidean space structure of the original sample space, and to work with the coordinates of the original observations in an orthonormal basis of this space, as well as with a Lebesgue measure compatible with them. That way we obtain a stay-in-the-support geostatistical technique, which means that we do not transform our observations from the original sample space into another space. We simply work with different representations in the same space.

We illustrate our approach with an alternative model for lognormal kriging, explained in section 3. It is based on the finite-dimensional Hilbert, *i.e.* Euclidean, space structure of the positive real line and on the normal distribution on $\boldsymbol{R}_+$ detailed in section 4. The most visible and intuitive distinction between our approach and the classic one–lognormal kriging–lays in the metric used to compute differences between samples. The most striking advantage of our approach is that it allows the estimation of the local conditional expectation, and the straightforward computation of probabilities and shortest-length confidence intervals, as explained in section 5. Block estimation is outlined in section 6, where it becomes clear that coordinate kriging has a far deeper effect, since in the $\boldsymbol{R}_+$ case it advocates the geometric mean instead of the arithmetic mean as the optimal way to estimate the expected value in a block. Finally, some conclusions and further implications close this work.

2 Brief geostatistical introduction

The theory of the regionalized variable, or Geostatistics (Matheron, 1963), was born as an inference methodology to evaluate ore grades, and extended rapidly to other spatial observations where basic assumptions of classical statistics did not work. Recall that classical statistical inference assumes independent replications of a single (unknown) random variable, while Geostatistics assumes a single replication of a set of different (unknown) variables that manifest spatial relationships. These characteristics (single replication, multiple variables, dependency) hinder the direct application of usual inference techniques to geostatistical problems. To overcome these difficulties from a theoretical point of view, data are supposed to be the realization of a random field $Z(\mathbf{x})$. A random field is a continuous collection of equally distributed random variables $Z(\mathbf{x}_i)$, each one located at each possible location $\mathbf{x}_i$ inside the study domain $\boldsymbol{D}$. Usually, it is assumed that the expectations and variances of these random variables, as well as their mutual covariances, do not depend on the location, a condition that is called *second-order stationarity*. This condition implies that the covariance $\mathrm{Cov}(Z(\mathbf{x}_i),Z(\mathbf{x}_j))$ between two random variables $Z(\mathbf{x}_i)$ and $Z(\mathbf{x}_j)$ located at $\mathbf{x}_i$ and $\mathbf{x}_j$ does not depend on these exact positions, but only on the distance between them, $h=|\mathbf{x}_i-\mathbf{x}_i|$, and we can write $\mathrm{Cov}(Z(\mathbf{x}_i),Z(\mathbf{x}_j))=\mathrm{C}(h)$. A common further

assumption is that this collection of random variables is jointly distributed as a multivariate normal. Thus, if it is second-order stationary, it is completely specified: it is then called a *Gaussian field*. However, usually the second order stationarity is weakened to the *intrinsic hypothesis*: only the increments of $Z(\mathbf{x})$ are assumed to fulfill stationarity conditions, and covariance functions (which may not exist) are replaced by variograms, defined as $\gamma(h)=\mathrm{Var}[Z(\mathbf{x}+h)-Z(\mathbf{x})]$. This option is usually taken because it needs the mean to be only locally stationary (which is almost always approximately true), while second-order stationarity needs a constant mean in all the domain $\boldsymbol{D}$. Table 1 summarizes the geostatistical notation used throughout this article.

Basic geostatistical concepts	
$\boldsymbol{D}$	study domain
$\mathbf{x}_i$	location in $\boldsymbol{D}$
$\mathbf{x}_0$	unsampled location
V	block in $\boldsymbol{D}$
$h=\lvert\mathbf{x}_i-\mathbf{x}_j\rvert$	distance between samples
$Z(\mathbf{x})$	general random field
$Y(\mathbf{x})$	Gaussian random field
$Z(\mathbf{x}_i)=Z_i$	random variable at $\mathbf{x}_i$
$C(h)$	covariance function
$\gamma(h)$	variogram
$\overline{Z}(V)$	block mean value

Sampling and inference	
$\mathbf{x}_1,\ldots,\mathbf{x}_n$	sampling points
$z_1,\ldots,z_n$	observed values of $Z(\mathbf{x})$
$y_1,\ldots,y_n$	transformed values
Y^*	kriging estimate of $Y(\mathbf{x}_0)$
Z^*	estimate of $Z(\mathbf{x}_0)$
σ_{yk}^2	kriging variance of Y^*
σ_{zk}^2	variance of Z^*

Parameters and statistics	
μ	mean parameter
σ^2	variance parameter
m	median
$\mathrm{E}[\cdot]$	expectation
$\mathrm{Var}[\cdot]$	variance

Table 1. Summary of notation used in this article.

Kriging is the name of a family of techniques devised to optimally estimate $Z(\mathbf{x}_0)$ and its estimation variance at an unsampled location $\mathbf{x}_0$ conditionally to the information given by all the available samples. In a Gaussian field, the estimation of the expected value at $\mathbf{x}_0$ and its variance will also univocally determine the conditional distribution of $Z(\mathbf{x}_0)$, something that only happens when the mean of the random field is known. In that case, a special technique called *simple kriging* can be applied.

Another usual application of kriging techniques is the estimation of some spatially averaged quantity. For instance, when mining an ore reserve, it is very useful to know the average ore grade in the blocks that will be processed. A point estimation would be unrealistic, since we will not be able to extract the richest veins and points from their (usually much) poorer frame rock. Thus it is better to estimate directly the mean grade on the block V, usually defined as

$$\overline{Z}(V)=\frac{1}{V}\int_V Z(\mathbf{x})\mathrm{d}\mathbf{x}, \tag{1}$$

by using kriging techniques adapted to blocks. It is important to notice that expression (1) defines a new random variable, different from every variable in the random field. However, it follows a normal distribution in the very special case of $Z(\mathbf{x})$ a Gaussian random field, which implies that we will be able to estimate it through the *same* kriging techniques as for point supports. This is another reason for the widespread assumption of normality.

For illustration, we use along the paper a mining data set: thickness of a mineralized vein in the Geevor Tin mine, Cornwall, extensively studied by Clark (1979). From this data set we randomly select a study subset with 500 stations represented in Figure 1, which are used for cross-validation: the left figure represents the empirical marginal distribution of the realizations of $Z(\mathbf{x})$, which in this case is clearly not normal. Instead, a lognormal assumption seems more reasonable, which means that the transformation $Y(\mathbf{x})=\log Z(\mathbf{x})$ should be approximately Gaussian. This is the reason why the variogram is computed on the transformed Y values. Sensibility to variogram modelling is assessed fitting two models to the data and comparing cross-validation results of the classical lognormal estimator with those of our proposed approach.

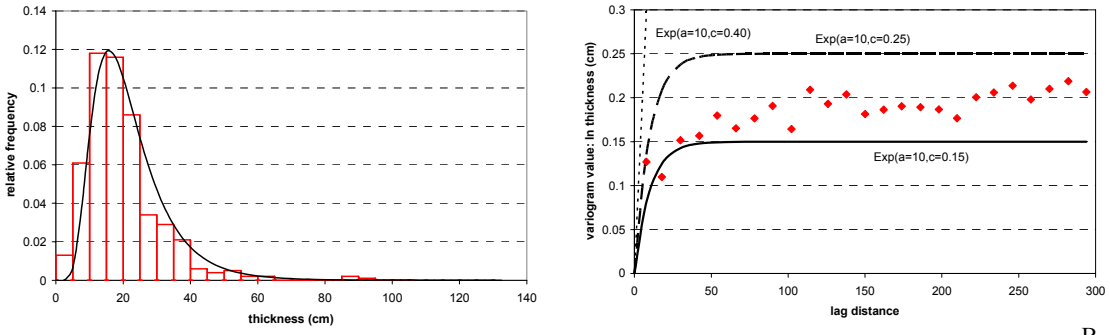

Figure 1. Conventional histogram (A, left) and omnidirectional logarithmic variogram (B, right) of thickness of mineralized vein, Geevor Tin mine. The experimental variogram is displayed with red diamonds in the right figure, as well as three different theoretical variogram models: they are used to assess robustness of discussed methods against variations of the sill.

3 Lognormal kriging

When a data set is constraint to positive values and regionalized, and specially when positively skewed, a logarithmic transformation is usually expected to symmetrize it. If the geostatistician assumes the transformed scores to be a realization of a Gaussian field, any classic kriging technique can be fully applied to them. The problem is then: how should the obtained results be back-transformed to map the original variable? This question is not a trivial one, to judge on the large amount of different lognormal kriging estimators found in the literature. Rivoirard (1990) offers a good summary of most of them, both for point and block support.

Following Chilès and Delfiner (1999), if $Y(\mathbf{x})=\log Z(\mathbf{x}) \sim N(\mu,\sigma^2)$, where $Z(\mathbf{x})$ is observed at $\mathbf{x}_1,\ldots,\mathbf{x}_n$ locations to obtain a sample $z_1,\ldots,z_n$, then the optimal estimator Z^* of $Z_0=Z(\mathbf{x}_0)$ is

$$Z^* = \mathrm{E}[Z_0|y_1,\ldots,y_n] = \exp\left(Y^* + \tfrac{1}{2}\,\sigma_{sk}^2 \right), \qquad (2)$$

where $y_1,\ldots,y_n$ is the logarithmic transformed sample, Y^* the usual *simple kriging* estimator of $Y_0=Y(\mathbf{x}_0)$ given this sample, and σ_{sk}^2 is the associated simple kriging variance. The conditional variance linked to Equation (2) is computed as

$$\mathrm{Var}[Z^* - Z_0 \mid y_1,\ldots,y_n] = (Z^*)^2 \cdot \left(\exp(\sigma_{sk}^2) - 1 \right). \qquad (3)$$

However, simple kriging is seldom applicable, since it needs the theoretical global mean to be known. A possible solution consists in substituting the simple kriging estimator by the *ordinary* kriging or *universal* kriging estimators, which are based on assuming the mean to be respectively constant or a polynomial function of location $\mathbf{x}$. However, it is unknown to which extent expressions (2) and (3) keep their meaning of conditionally unbiased estimators of Z_0. Also, it is unclear whether it is better to replace in Equation (2) the simple kriging variance by its ordinary/universal estimate –Cressie (1991), according to Clark and Harper (2000, p 323)– or to leave it unchanged –as proposed by Journel and Huijbregts (1978)–.

All these authors notice the extreme sensitivity of estimator (2), since any error in the kriging or variance estimations become exponentially magnified. The uncertain fitting of the sill is particularly important, as it has a direct effect on the kriging variance. Moreover, deviations from lognormality of the data may dramatically invalidate its use (Clark and Harper, 2000).

In addition, Journel and Huijbregts (1978, p. 572) show that Equation (2) is usually locally biased: given μ_z, known expected value of $Z(\mathbf{x})$, they propose to modify (2) by a multiplicative factor $K_0 = \mu_z / \Sigma\, Z^*$. This correction was proposed for block kriging and, according to Chilès and Delfiner (1999, p. 192), should be avoided in point kriging problems, since the obtained estimator is no longer an exact interpolator.

Likewise, Equation (2) was shown to be a bad estimator of the local conditional expectation (Roth, 1998). His example shows a case with final estimates surpassing the convex envelope of the observed data, being all kriging weights strictly positive. He blames the presence of σ_{sk}^2 for this behaviour. Thus he advocates for the estimation of the median, instead of the mean.

In agreement with this proposal, Chilès and Delfiner (1999, p. 191) explain that good results are obtained working with quantiles, since they simply follow the logarithmic transformation. In other words, the direct relationship $m_Y = \log m_Z$ between the medians m_Y and m_Z allows an easy estimation of the median of $Z(\mathbf{x})$ as a central tendency estimator. This approach offers a way to compute confidence intervals around the median, like

$$Z^* \in [\exp(Y^* - 2\,\sigma_{sk}) \,,\, \exp(Y^* + 2\,\sigma_{sk})] \,, \qquad (4)$$

defining a 95% confidence interval. However, confidence intervals (4) are not considered to be optimal, as they do not have minimal Euclidean length in R. According to Clark and Harper (2000), this issue was studied in the non-geostatistical case by Sichel (1971), who obtained a set of multiplicative coefficients intended to compute unbiased means and optimal confidence intervals for the lognormal distribution. These coefficients depend on the sample size and the logarithmic variance, but they do not capture explicitly spatial correlation, and their application must be restricted to low variance data. Another option is to use expressions (2) and (3) to compute directly a conventional confidence interval with them, based on Gaussian assumption, although its physical meaning is dubious (it might include negative values). The main features of these different computations of confidence intervals for the data considered are summarized in Table 2. Figure 2 compares a single case.

confidence intervals	conventional intervals			median intervals			Sichel intervals		
level	90%	95%	97.5%	90%	95%	97.5%	90%	95%	97.5%
%int. with true value	95%	97%	97%	93%	96%	98%	7%	9.7%	12%
%int. with negative values	0%	74%	89%	0%	0%	0%	0%	0%	0%
median Euclidean length	38.49	46.00	52.81	35.60	44.52	53.55	11.11	14.51	17.66
median logarithmic length	2.62	∞	∞	1.62	1.94	2.23	0.27	0.36	0.43

Table 2. Characterization of central tendency confidence intervals: proportions of estimated intervals containing the true value (cross-validation) or negative values, and two median interval lengths

If the interest is simulation, then Equations (2) and (3) are seldom used (Deutsch and Journel, 1992, p. 76). Instead, kriging estimates and kriging variances are used to simulate normal scores of Y, that afterwards are transformed to simulations of $Z(\mathbf{x})$ through $z_{sim}=\exp(y_{sim})$. This procedure is consistent with the median estimates and confidence intervals given in expression (4).

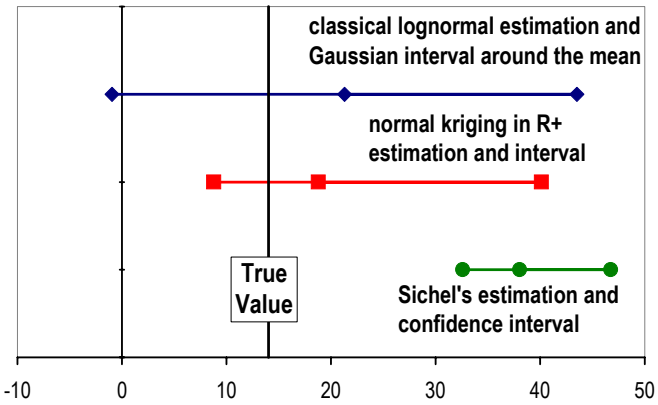

Figure 2. Confidence intervals compared.

Block lognormal kriging is a far more complex issue, which involves many theoretical and practical considerations, being the most important one the so called *permanence of lognormality*: having lognormal point values, any weighted arithmetic mean, e.g. any kriging estimator or a spatial average like (1), will

never be lognormal. This well-known result formally invalidates the rigorous application of block kriging to lognormal sets, although Rendu (1979) shows that it gives fairly good empirical results. Journel (1980) presents an approximate but elegant solution based on point lognormal kriging, specially suited to describe detailed local variability through block distribution functions instead of by a detailed kriged map. Marcotte (1997) tries a complex solution, assuming joint lognormality of the estimator and the real value, and obtains an estimator very similar to (2). Dowd (1982) develops an alternative, which requires the numerical solution of a system of integral equations, without assuming permanence of lognormality.

4 Normal on $\boldsymbol{R}_+$ distribution

The lognormal distribution is assumed to model strictly positive random variables when uncertainty follows a multiplicative scheme. Its estimation problems, even in a non-geostatistical environment, have generated a large amount of statistical literature, some of them summarized by Clark and Harper (2000). They are still today an open issue. According to Mateu-Figueras, Pawlowsky-Glahn and Martín-Fernández (2002), these problems might be caused by the assumption that the sample space –the positive real line as a subset of real vector space $\boldsymbol{R}$– has the Euclidean structure induced from $\boldsymbol{R}$.

issue	general	$\boldsymbol{R}$ space	$\boldsymbol{R}_+$ space
element	$\mathbf{v}$	$v \in \boldsymbol{R}$	$\mathbf{v} \in \boldsymbol{R}_+$, $v \in \boldsymbol{R}_+ \subset \boldsymbol{R}$
Abelian group operation	$\mathbf{v} \oplus \mathbf{w}$	$v+w$	$v \cdot w$
external product	$\lambda \otimes \mathbf{v}$	λv	v^λ
scalar product	$\langle \mathbf{v} \mid \mathbf{w} \rangle$	$v \cdot w$	$\log v \cdot \log w$
norm	$\|\mathbf{v}\|$	$\|v\|$	$\|\log v\|$
distance	$d(\mathbf{v},\mathbf{w})$	$\|v-w\|$	$\|\log v - \log w\|$
canonical basis	$\mathbf{u}$	1	$e^1=e$
neutral element	$\varnothing$	0	$e^0=1$
inverse element	$-\mathbf{v}$	$-v$	$1/v$
unit-norm basis coefficient	$\langle \mathbf{v} \mid \mathbf{u} \rangle$	$v \in \boldsymbol{R}$	$\log v \in \boldsymbol{R}$
coefficient expression	$\mathbf{v} = \langle \mathbf{v} \mid \mathbf{u} \rangle \otimes \mathbf{u}$	$\mathbf{v} = (v \cdot 1)1$	$\mathbf{v} = e^{\log v \cdot \log e}$

Table 3. Operations inducing a Euclidean space structure in $\boldsymbol{R}$ and in $\boldsymbol{R}_+$. Any element can be expressed as a unique linear combination of the unit-norm or canonical basis, whose coefficient is a scalar product. Notice that boldface letters are used when considering the values as vectors, while italic style shows a real value. Greek letters are used for scalar values.

An alternative approach can be set up considering the positive real line as a Euclidean space itself (Table 3). This idea, first used by Pawlowsky-Glahn and Egozcue (2001, p. 387) to illustrate the concepts of metric expectation and metric variance, allowed Mateu-Figueras, Pawlowsky-Glahn and Martín-Fernández (2002) to introduce a normal distribution on $\boldsymbol{R}_+$. With this space structure in $\boldsymbol{R}_+$, it is straightforward to define a normal distribution in $\boldsymbol{R}_+$ as a normal distribution on the coefficients

$$Z \sim N_+(\mu,\sigma^2) \Leftrightarrow f_+(z) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2} \frac{(\log z - \mu)^2}{\sigma} \right), \quad z \in \boldsymbol{R}_+. \tag{5}$$

Properties of the normal distribution on $\boldsymbol{R}_+$ were studied by Mateu-Figueras, Pawlowsky-Glahn and Martín-Fernández (2002), which include the coincidence of mean, mode and median of the distribution in $\exp(\mu)$. Its appearance is clearly different from the classical lognormal distribution (Figure 3), although their expressions are quite similar. This is due to the metric assumed for their support space, the positive real line: while the lognormal distribution is built up on the usual Euclidian metric on $\boldsymbol{R}$, the normal distribution on $\boldsymbol{R}_+$ is based on a logarithmic metric. This fact is implicitly shown in Figure 1 and in Figure 5, where the histogram intervals have a constant length with respect to the metric considered in each case. Furthermore, parameter estimation in $\boldsymbol{R}_+$ becomes as easy as usual parameter estimation in $\boldsymbol{R}$: given a sample $z_1, \ldots, z_n$ of Z, there is an unbiased estimator of its expected value,

$$\bar{z} = \left(\prod_{i=1}^{n} z_i \right)^{1/n}, \tag{6}$$

and of the parameters of the distribution, given by

$$\hat{\mu} = \log \bar{z} = \frac{1}{n}\sum_{i=1}^{n} \log z_i \quad ; \quad \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n} \log^2 \frac{z_i}{z_j} = \frac{1}{n}\sum_{i=1}^{n} \log^2 \frac{z_i}{\bar{z}}. \tag{7}$$

Properties of these estimators where studied by Pawlowsky-Glahn and Egozcue (2001, p. 388-392). It is worth notice that the mean value [Eq. (6)] is computed as the geometric mean of the sample, hence it is an element of the same $\boldsymbol{R}_+$ space, while the estimate of the parameter μ of the distribution [Eq. (7), left] is its coordinate according to Table 3. The σ^2 parameter [Eq. (7), right] has no direct interpretation as an element of the space; nevertheless, it can be used to compute exact and optimal confidence intervals around the mean identical to those built around the median [Eq. (4)] (Mateu-Figueras, Pawlowsky-Glahn and Martín-Fernández, 2002). Figure 3 compares one of these confidence intervals with a standard normal on $\boldsymbol{R}_+$ distribution and its equivalent lognormal distribution, and shows that in the first case they are isodensity intervals, a condition not fulfilled in the second case.

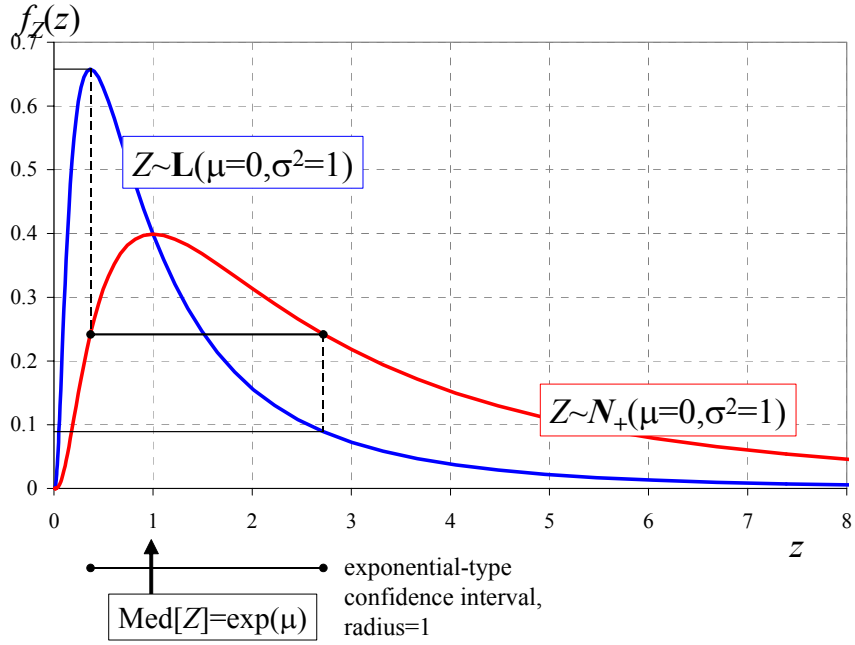

Figure 3. Probability density functions of a standard lognormal and a standard normal in $\boldsymbol{R}_+$ distribution, with a probability interval of exponential type. Interval bounds have the same density value for the normal in $\boldsymbol{R}_+$ distribution, while this is not true for the lognormal distribution.

5 Normal point kriging in $\boldsymbol{R}_+$

Adapting the usual definition of kriging, given a realization $z_1, \ldots, z_n$ of a random function $Z(\mathbf{x})$ in $\boldsymbol{R}_+$ observed at n locations $\mathbf{x}_1, \ldots, \mathbf{x}_n$, the simple kriging predictor at location $\mathbf{x}_0$ will be a weighted linear combination of observations, according to the space structure of $\boldsymbol{R}_+$,

$$Z^* = \mathrm{E}\left[Z_0 \middle| y_1, \ldots, y_n\right] = z \cdot \prod_{i=1}^{n} \left(\frac{z_i}{\bar{z}}\right)^{b_i} = z^{1-\sum_{i=1}^{n} b_i} \cdot \prod_{i=1}^{n} z_i^{b_i}, \tag{8}$$

where z is the expected value of $Z(\mathbf{x})$ in this space structure (Table 3). Moreover, the estimation variance associated to (8) is estimated as

$$\mathrm{Var}\left[Z^* - Z_0 \middle| y_1, \ldots, y_n\right] = \sigma_+(x_0, x_0) + \sum_{i=1}^{n}\sum_{j=1}^{n} b_i b_j \sigma_+(x_i, x_j) - 2\sum_{i=1}^{n} b_i \sigma_+(x_i, x_i). \tag{9}$$

In this expression, σ_+ stands for the autocovariance function of the variable $Z(\mathbf{x})$ defined on its coefficients, i.e. its logarithms. Thus, covariance functions (and variograms) used in lognormal kriging keep their meaning and usefulness.

Expressing Equation (8) as a function of its coefficients in a canonical basis of $\boldsymbol{R}_+$ (Table 3), the simple kriging estimator becomes

$$\log Z^* = \left(1 - \sum_{i=1}^{n} b_i\right) \cdot \log \bar{z} + \sum_{i=1}^{n} b_i \log z_i = \left(1 - \sum_{i=1}^{n} b_i\right) \cdot \mu + \sum_{i=1}^{n} b_i y_i = Y^*,$$

equivalent to the conventional lognormal estimator. Hence, the system to be solved will be the usual one for lognormal kriging, leading to the same kriging estimate of Y_0 and the same variance. However, here $Y(\mathbf{x})$ is the coordinate of $Z(\mathbf{x})$ in the canonical basis of $\boldsymbol{R}_+$, instead of its transformation, which means that the simple kriging estimator of Z_0 is

$$Z^* = \mathrm{E}[Z_0 | y_1, \ldots, y_n] = \exp(Y^*). \tag{10}$$

An important difference between the lognormal estimator (2) and the normal estimator in $\boldsymbol{R}_+$ (10) is the presence of the kriging variance. Skipping it, Equation (10) avoids most of the problems of Equation (2): it is robust with respect to deviations from normality on $\boldsymbol{R}_+$, it is also less sensitive to variogram fitting (Figure 4) and it offers locally consistent estimations. Obviously, it is identical to the widely used median estimator. Notice that, with expression (10) and the logarithmic kriging variance [Eq. (9)] in fact we estimate the local conditional distribution of $Z(\mathbf{x}_0)$. As a heuristic evidence of this statement, kriging error standardized with respect to kriging variance is well approximated by a standard normal on $\boldsymbol{R}_+$ (Figure 5). This implies that confidence intervals for the expected value computed with expression (4) will be exact and optimal, being minimal their logarithmic length, as was recorded in Table 2. Furthermore, if simple kriging is not applicable, ordinary or universal kriging estimates and variances can be regarded as good approximations to the real conditional expectation parameters, at least in the same sense they are in the Gaussian case.

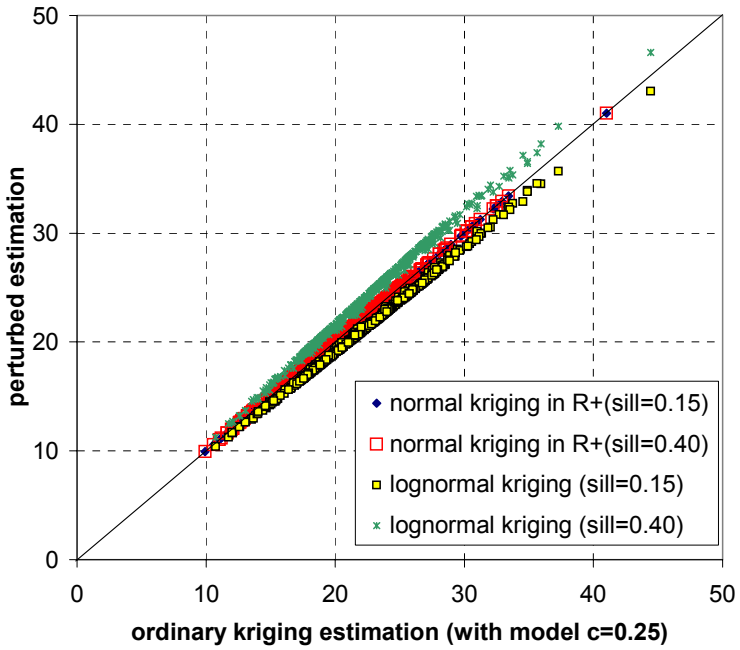

Figure 4. Scatterplots of predicted values with ordinary kriging under assumption of a variogram model with c=0.25 against the same predictions after perturbations of the sill of the variogram, with c=0.15 and c=0.40.

6 Block kriging revisited

According to Chilès and Delfiner (1999, p. 193), kriging was originally developed not for estimation of point values but of average grades over mining panels. Its main goal was to avoid the systematic bias detected when estimating a panel average grade solely on the basis of its internal samples: since richer areas are surrounded by poorer areas and vice-versa, selected blocks estimated as rich were poorer than expected, and discarded blocks estimated as poor contained more ore than expected. This effect, called *support effect*, increases with the size of the block, and implies the under-exploitation of an ore deposit. To handle it, special *support effect models* have been developed.

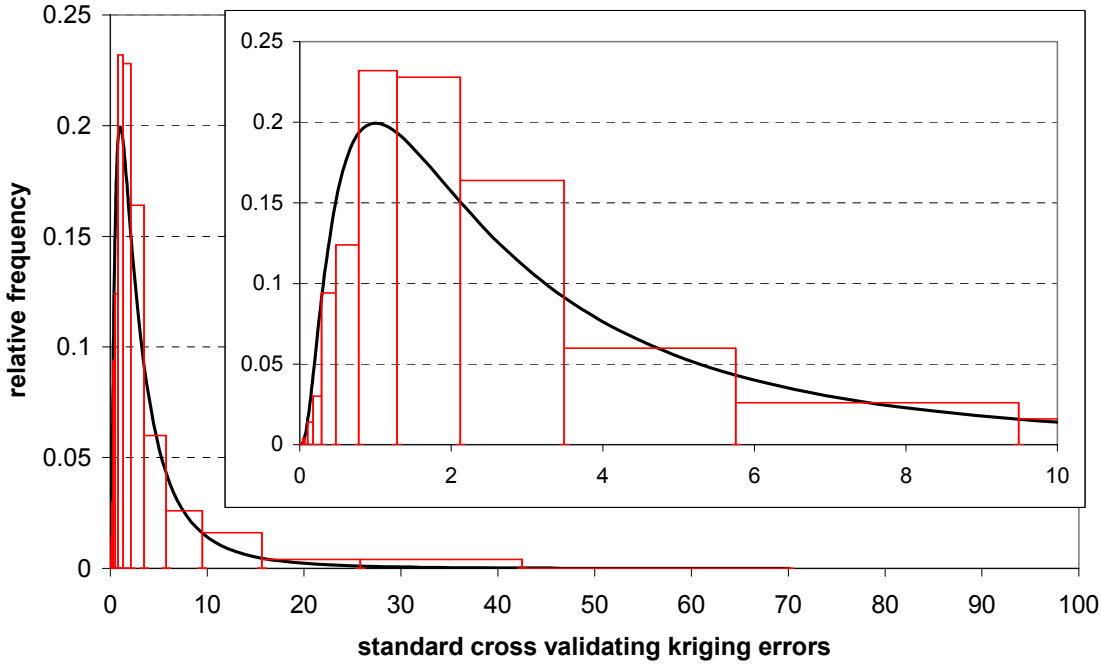

Figure 5. Histogram of standardized ordinary kriging cross-validation errors, compared with a standard normal on $\boldsymbol{R}_+$. The nested histogram shows a detail near the origin. Note that histogram classes have the same logarithmic length.

Under additional restrictive conditions, mainly ergodicity, Gaussian random fields with mean μ, variance σ^2 and covariance C(h) present also Gaussian distributions of the average grade in a block V, with the same mean μ, but with a smaller variance σ_V^2, which must be equal to the average value of C(h) for every possible pair of points inside V (Chilès and Delfiner, 1999). This is the simplest *support effect* model, called *affine correction*.

When the available data are modelled by a lognormal random field, the usual *support effect* models (discrete Gaussian model and multi-Gaussian model) assume the block distribution to be jointly lognormal with the data (in fact, bi-variate and multivariate lognormal respectively). Thus, both models imply *permanence of lognormality*: the mean value of all the random lognormal variables $Z(\mathbf{x}\in V)$ inside the block V is assumed lognormally distributed. It is well-known that this permanence is not true, although it is empirically a good approximation. Based on it, there are several proposed estimators (Rivoirard, 1990) for mean block values [Eq. (1)], although all of them have a similar form to the point lognormal estimator (2). If this hypothesis is not assumed, then one can still estimate expression (1) by solving an integral system (Dowd, 1982).

In contrast, assuming a Euclidean geometry on $\boldsymbol{R}_+$ simplifies the *support model*, since using the coordinate approach, the mean value in the block is now the mean value in the coordinates applied to the basis,

$$\overline{Z}(V) = \left(\frac{1}{V} \int_V \log Z(\mathbf{x}) d\mathbf{x} \right) \otimes \mathbf{e} = \exp\left(\frac{1}{V} \int_V \log Z(\mathbf{x}) d\mathbf{x} \right). \tag{11}$$

This definition of the mean value solves the permanence of lognormality problem. Since log $Z(\mathbf{x})$ is a Gaussian random field, we can apply to it the affine correction model, and this mean value [Eq. (11)] will automatically be normally on $\boldsymbol{R}_+$ distributed. Hence, when applying kriging to the estimation of expression (11), block normal kriging on $\boldsymbol{R}_+$ will provide optimal estimates, at least under the usual conditions for Gaussian random fields (stationarity and ergodicity). In this sense, normal on $\boldsymbol{R}_+$ random fields becomes a good alternative to lognormal random fields when the objective is block kriging.

7 Conclusions

Kriging, as a statistical interpolation technique, is based on a given optimality criterion: minimisation of an error variance. However, the capability of some real data sets to be treated according to several Euclidean space structures opens the door to different kriging techniques, each one developed to be optimal with a different space structure.

This idea is here applied to the case of positive data, which can be studied as values of a subset of $\boldsymbol{R}$ or as elements (vectors) of a one-dimensional Hilbert space on $\boldsymbol{R}_+$. In the first case, the natural choice is to model them with lognormal distributions, although then kriging estimates present some problems: they are biased, they are bad local estimators of the conditional expectation, and also, there is no generally accepted way to build optimal confidence intervals for the expected value. Thus, several partial solutions are found in the references around this topic: one of these solutions, suggested by many authors, is to work with median estimators and confidence intervals, since they seem experimentally to be far more robust and meaningful. If the normal on $\boldsymbol{R}_+$ distribution is chosen to model positive data, then kriging can be expressed on the canonical coordinates of the data in $\boldsymbol{R}_+$: the obtained estimator is equivalent to usual median estimator, but it can be shown to have all expected properties of a Gaussian estimator, including estimation of confidence intervals and probabilities, even conditional distribution. Also, it does not present any of the problems linked to lognormal kriging. Thus, the normal on $\boldsymbol{R}_+$ approach supports theoretically the heuristic solution of working with the median.

Regarding block kriging, it is worth notice that the classical lognormal approach relies on a (false though heuristically good) assumption of permanence of lognormality for block mean values. On the contrary, normal on $\boldsymbol{R}_+$ data naturally give normal on $\boldsymbol{R}_+$ block mean values, when computed as geometric means according to the geometry of $\boldsymbol{R}_+$. However, this implies that the mean of the grade in a block is the geometric mean, in contrast to the generally accepted arithmetic mean of ore grade.

Generalizing these results, canonical coordinate kriging allows a direct interpolation of observations in any space (for instance, permeability in $\boldsymbol{R}_+$, probabilities and proportions in (0,1) or compositions in the simplex considered as regionalized variables or vectors) honouring both the sensible metric of their support space and the BLU character of kriging estimates. Furthermore, it has deep implications on those optimality criteria, as well as on the interpretation of block kriging, which depend on the geometry chosen for the support.

Acknowledgements

This research has been funded by the *Dirección General de Enseñanza Superior e Investigación Científica* (DGESIC) of the Ministry of Education and Culture through the project BFM2000-0540, and from the *Direcció General de Recerca* of the *Departament d'Universitats, Recerca i Societat de la Informació* of the *Generalitat de Catalunya* through the project 2001XT 00057.

References

Chilès, J.P. and Delfiner, P., 1999. Geostatistics - modeling spatial uncertainty. Series in Probability and Statistics. John Wiley and Sons, Inc., New York, NY (USA). 695 p.

Clark, I, 1979; Geostatistical Reserve Estimation in Cornish Tin Mining, unpublished doctoral disertation, University of London.

Clark, I. and Harper, W.V., 2000. Practical Geostatistics 2000. Ecosse North America Llc, Columbus Ohio (USA). 342 p.

Cressie, N., 1991. Statistics for Spatial Data. John Wiley and Sons, New York, NY (USA). 900 p. Cited in Clark and Harper (2000).

Deutsch, C. and Journel, A., 1992. GSLIB - geostatistical software library and user's guide. Oxford University Press, New York, NY (USA). 340 p. and 2 diskettes.

Dowd, P.A., 1982. Lognormal kriging - the general case. Math. Geol. 14 (5), p. 475-499.

Journel, A.G., 1980. The lognormal approach to predicting local distributions of selective mining unit grades. Math. Geol. 12 (4), p. 285-303.

Journel, A.G. and Huijbregts, C.J., 1978. Mining geostatistics. Academic Press, London (UK). 600 p.

Marcotte, D. and Groleau, P., 1997. A simple and robust lognormal estimator. Math. Geol. 29 (8), p. 993-1009.

Mateu-Figueras, G., Pawlowsky-Glahn, V. and Martín-Fernández, J.A., 2002. Normal in R_+ vs lognormal in R. *in* Bayer, U., Burger, H. and Skala, W., eds., Proceedings of IAMG'98, Vol I, p. 305-310. Selbstverlag der Alfred-Wegener-Stiftung, Berlin.

Pawlowsky-Glahn, V. and Egozcue, J.J., 2001. Geometric approach to statistical analysis on the simplex. SERRA 15 (5), p. 384-398.

Rendu, J.M., 1979. Kriging, logarithmic kriging, and conditional expectation: comparison of theory with actual results. Proceedings of the 16th APCOM, p. 199-212.

Rivoirard, J., 1990. A review of lognormal estimators for *in situ* reserves (teacher's aid). Math. Geol. 22 (2), p. 213-221.

Roth, C., 1998. Is lognormal kriging suitable for local estimation? Math. Geol. 30 (8), p. 999-1009.

Sichel, H., 1971. On a family of discrete distributions particularly suited to represent long-tailed frequency data, *in* Laubscher, N.F., ed., Third Sym. Mathem. Statist. Pretoria South Africa, p. 51-97. Cited in Clark and Harper (2000).