Mathematical Geology, Vol. 33, No. 8, November 2001 (© 2001)

Compositional Kriging: A Spatial Interpolation Method for Compositional Data[1]

Dennis J. J. Walvoort[2,3] and Jaap J. de Gruijter[2]

Compositional data are very common in the earth sciences. Nevertheless, little attention has been paid to the spatial interpolation of these data sets. Most interpolators do not necessarily satisfy the constant sum and nonnegativity constraints of compositional data, nor take spatial structure into account. Therefore, compositional kriging is introduced as a straightforward extension of ordinary kriging that complies with these constraints. In two case studies, the performance of compositional kriging is compared with that of the additive logratio-transform. In the first case study, compositional kriging yielded significantly more accurate predictions than the additive logratio-transform, while in the second case study the performances were comparable.

KEY WORDS: closed data, geostatistics, indicator kriging, fuzzy sets, ternary diagram.

INTRODUCTION

In the earth sciences, data are often expressed as fractions or percentages. Examples are soil textural classes, the chemical composition of rock, and the pollen and foraminiferal composition of sediments. Other, less obvious, examples are indicator data (Bierkens and Burrough, 1993a,b; Journel, 1983) and the data structures resulting from continuous classification algorithms such as fuzzy c-means (Bezdek, Ehrlich, and Full, 1984; De Gruijter and McBratney, 1988). These are all examples of a special type of data usually referred to as closed data, compositional data, or compositions (Aitchison, 1986; Davis, 1986). A regionalized composition is formally defined as a vector random function $\mathbf{z}(x_i)$ located at point x_i in a spatial domain $\mathcal{D}$ with p components $z_k(x_i)$

$$\mathbf{z}(x_i) = [z_1(x_i), z_2(x_i), \ldots, z_p(x_i)]^{\mathrm{T}}$$

[1]Received 1 June 1999; accepted 2 November 2000.

[2]Alterra–Wageningen University and Research Centre, P.O. Box 47, 6700 AC Wageningen, The Netherlands; e-mail: d.j.j.walvoort@alterra.wag-ur.nl

[3]Laboratory of Soil Science & Geology, Wageningen University and Research Centre, P.O. Box 37, 6700 AA Wageningen, The Netherlands.

that are nonnegative

$$z_k(x_i) \geq 0 \quad \text{for } k = 1, \ldots, p$$

and sum to a constant c which usually equals 100 (%) or 1:

$$\mathbf{z}^{\mathrm{T}}(x_i)\mathbf{1}^{(p)} = c$$

where $\mathbf{1}^{(p)}$ represents the p-dimensional vector of ones, and superscript T denotes transposition. Pawlowsky (1989), Pawlowsky and Burger (1992), Olea, Pawlowsky, and Davis (1993), and Pawlowsky, Olea, and Davis (1994, 1995) employed a more restrictive definition. To avoid unnecessary complications, these authors did not consider compositions with one or more zero components. In this paper, the formal definition given above is adopted.

Spatial interpolation methods come into scope when compositions at unobserved sites $x_0 \in \mathcal{D}$ are required. Obviously, a suitable spatial interpolation method should at least satisfy the constant sum and nonnegativity constraints of compositional data. In addition, the interpolator should also take the spatial covariance structure into account.

The aim of our study is to identify an interpolation method that complies with these demands. First, we give an overview of existing methods. Because none of the existing methods satisfy all the given criteria, a new interpolation method, viz. compositional kriging, is introduced. Its performance is evaluated in two case studies.

EXISTING METHODS

Most spatial interpolation methods yield predictions $\mathbf{z}(x_0)$ that are weighted linear combinations of the available data:

$$\mathbf{z}(x_0) = \operatorname{diag}(\mathbf{W}^{\mathrm{T}}\mathbf{Z})\,\mathbf{1}^{(p)}$$

where $\mathbf{Z}$ is the data matrix, given by

$$\mathbf{Z} = [z_k(x_i) \mid k = 1, \ldots, p;\ i = 1, \ldots, n] = \begin{bmatrix} z_1(x_1) & z_2(x_1) & \cdots & z_p(x_1) \\ z_1(x_2) & z_2(x_2) & \cdots & z_p(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ z_1(x_n) & z_2(x_n) & \cdots & z_p(x_n) \end{bmatrix}$$

$\mathbf{W}$ is the corresponding matrix of weights, given by

$$\mathbf{W} = [w_k(x_i) \mid k = 1, \ldots, p;\ i = 1, \ldots, n]$$

and diag(·) is the diagonal matrix formed from the diagonal elements of its argument. Appropriate interpolation methods should at least yield predictions that fully comply with the nonnegativity and constant sum constraints of compositional data. Examples of such methods are nearest neighbor interpolation, triangulation, local sample mean, and inverse distance interpolation. Isaaks and Srivastava (1989) provide more details on these methods, which always yield compositions because the weights range from zero to one, the weights sum to unity, the columns of **W** are identical, and the conditioning data are compositions. However, as will be shown later, these properties are too restrictive. Another disadvantage of these methods is that they do not take the spatial covariance structure into account. This is an important property of kriging (Journel and Huijbregts, 1978), and we will therefore focus on this geostatistical interpolation method in the remainder of this paper.

Burrough, Van Gaans, and Hootsmans (1997) and Odeh, McBratney, and Chittleborough (1992) used ordinary kriging to interpolate fuzzy membership vectors. These compositions were obtained by means of fuzzy clustering (Bezdek, Ehrlich, and Full, 1984; De Gruijter and McBratney, 1988). However, ordinary kriging does generally not yield compositions for two reasons. First, as ordinary kriging is a nonconvex interpolator, predictions are not necessarily nonnegative. Negative predictions may occur if observation points are screened by other observation points (Isaaks and Srivastava, 1989; Journel and Huijbregts, 1978). This so-called screen effect is more pronounced for semivariograms with a parabolic shape near the origin and a small relative nugget variance. Second, the constant sum constraint is only guaranteed if all columns of **W** are identical. This can be accomplished by using semivariograms γ_k that are proportional in scale:

$$\gamma_k = a_k \gamma \quad \text{for } k = 1, \ldots, p \tag{1}$$

where a_k is a positive scaling factor, and γ a permissible semivariogram. Practically speaking, this kind of spatial structure is only encountered when $p = 2$. In that case, the semivariograms are necessarily identical because of the constant sum constraint.

It can be concluded that ordinary kriging by itself is not sufficient for predicting compositions at unvisited sites. Therefore, several authors applied kriging in conjunction with a transformation method. In this paper, two of these methods are discussed, viz. the basis method (Olea, Pawlowsky, and Davis, 1993; Pawlowsky, Olea, and Davis, 1995) and the additive logratio-transform (Pawlowsky, Olea, and Davis, 1994, 1995).

The basis method transforms a composition into an associated regionalized basis, which is defined as a vector random function $\mathbf{b}(x_i)$

$$\mathbf{b}(x_i) = [b_1(x_i), b_2(x_i), \ldots, b_p(x_i)]^{\mathrm{T}}$$

with p components $b_k(x_i)$ that are all measured on the same measurement scale

and are positive (Olea, Pawlowsky, and Davis, 1993; Pawlowsky, Olea, and Davis, 1995):

$$b_k(x_i) > 0 \quad \text{for } k = 1, \ldots, p$$

The relationship between a regionalized basis and its associated regionalized composition is given by

$$\mathbf{z}(x_i) = \frac{c\mathbf{b}(x_i)}{\mathbf{b}^{\mathrm{T}}(x_i)\mathbf{1}^{(p)}} = \frac{c\mathbf{b}(x_i)}{s(x_i)} \tag{2}$$

The denominator of this expression, $s(x_i)$, is called the size of the basis (Aitchison, 1986). Each regionalized basis $\mathbf{b}(x_i)$ has a unique size $s(x_i)$ and composition $\mathbf{z}(x_i)/c$. As an example of a regionalized basis consider the equivalent thickness of oil, water, and rock observed in wells in an oil field (Olea, Pawlowsky, and Davis, 1993).

The basis method can be summarized as follows:

1. Given a regionalized composition $\mathbf{z}(x_i)$ and size $s(x_i)$, compute a regionalized basis by means of

$$\mathbf{b}(x_i) = \frac{s(x_i)\mathbf{z}(x_i)}{c}$$

2. Perform semivariogram analysis and (co-)kriging on the regionalized basis;
3. Backtransform predicted regionalized bases into regionalized compositions by means of Eq. (2).

Although the basis method meets the constant sum constraint, predictions are not necessarily nonnegative. Moreover, backtransformation does not lead to a predictor with known and desirable properties, such as unbiasedness or minimum prediction error variance (Olea, Pawlowsky, and Davis, 1993; Pawlowsky, Olea, and Davis, 1995).

The second transformation method considered in this paper is the additive logratio-transform alr$(\cdot)$ (Aitchison, 1986; Pawlowsky, Olea, and Davis, 1994, 1995). It is given by

$$\mathrm{alr}(\mathbf{z}(x_i)) = \mathbf{y}(x_i) = \left[\ln\left(\frac{z_1(x_i)}{z_p(x_i)}\right), \ln\left(\frac{z_2(x_i)}{z_p(x_i)}\right), \ldots, \ln\left(\frac{z_p(x_i)}{z_p(x_i)}\right)\right]^{\mathrm{T}}$$

After performing semivariogram analysis and kriging on the alr-transformed data, predictions are backtransformed by means of the additive generalized logistic

transformation agl(·):

$$\text{agl}(\mathbf{y}(x_0)) = \mathbf{z}(x_0) = \frac{c[\exp(y_1(x_0)), \exp(y_2(x_0)), \ldots, \exp(y_p(x_0))]}{\mathbf{1}^{\mathrm{T}}[\exp(y_1(x_0)), \exp(y_2(x_0)), \ldots, \exp(y_p(x_0))]}$$

Note that the last element of $\mathbf{y}(x_i)$ equals 0 for all $x_i \in \mathcal{D}$, including prediction point x_0. Hence, this element is excluded from semivariogram analysis and kriging.

Kriging in conjunction with the alr-transform yields predictions that fully satisfy the constant sum and nonnegativity constraints. However, unlike the transformed data, the backtransformed data do not possess desirable properties such as unbiasedness and minimum prediction error variance (Pawlowsky, Olea, and Davis, 1994, 1995). Nevertheless, backtransformation (i.e., agl-transform) is generally preferred as it makes the results easier to interpret.

Neither the basis method nor the alr-transform honor all requirements of a suitable interpolation method listed in the Introduction. Nevertheless, Pawlowsky (1989), Olea, Pawlowsky, and Davis (1993), and Pawlowsky, Olea, and Davis (1994, 1995) advocate the use of these transformation methods because they make cokriging of (transformed) regionalized compositions *as a whole* possible. In this way, all available information is integrated in the prediction process, thus leading, at least from a theoretical point of view, to better predictions. Direct application of cokriging to regionalized compositions as a whole is not possible because of spurious negative correlation between the components. Spurious correlation is induced by the constant sum constraint, and results necessarily in singular cokriging matrices (Pawlowsky, 1989). We take the view that the benefits of cokriging do not outweigh the limitations of the transformation methods for the following reasons:

- In case of cokriging, semivariogram analysis is very laborious. Not only must one model $p(p+1)/2$ semivariograms and cross-semivariograms, but all resulting covariance matrices must be positive definite;
- Although the linear model of coregionalization guarantees positive definiteness (Isaaks and Srivastava, 1989; Journel and Huijbregts, 1978), this model may be too restrictive, resulting in (cross-)semivariograms that fit the data poorly. This may deprive the cokriging system of some of its potential improvements over ordinary kriging (Isaaks and Srivastava, 1989);
- The cokriging system most commonly applied contains $p \times p$ unbiasedness conditions. Let λ_{kl}^{i} denote the weight assigned to the kth element of conditioning composition $\mathbf{z}(x_i)$ to predict the lth element of $\mathbf{z}(x_0)$. Then for each combination of k and l the weights must sum to unity if $k = l$ and to zero if $k \neq l$, i.e. $\sum_{i=1}^{n} \lambda_{kl}^{i} = \delta_{kl}$, where δ_{kl} is Kronecker delta (Myers, 1982). If $k \neq l$, positive weights should necessarily be counterbalanced by negative weights. As a result, this cokriging system is prone to

negative predictions (Isaaks and Srivastava, 1989), hampering use of the basis method;
- Cokriging assumes linear relationships between the covariates, while in practice nonlinearities may occur.

For these reasons, we will not take cross-correlations into account. Although this may imply a loss of information in the prediction process, we are no longer dependent on transformation methods. Instead, we will focus on more direct approaches that act on the compositional data vectors themselves.

Szidarovszky, Baafi, and Kim (1987) and Deutsch (1996) provide algorithms that produce nonnegative kriging weights. When applied to compositional data, these convex predictors inevitably lead to nonnegative predictions. Szidarovszky, Baafi, and Kim (1987) enforced nonnegative weights by replacing the weights in the ordinary kriging system by the square of a new variable. This modification necessarily results in nonnegative weights, because for real numbers, squares are always greater than or equal to zero. The weights of this system are found by solving the original ordinary kriging system in an iterative and systematic way for different sets of conditioning data. Deutsch (1996) provides an algorithm that corrects the kriging weights a posteriori. That is, the ordinary kriging weights are computed, and then negative weights and related compensating positive weights are reset to zero. Subsequently, all weights are restandardized to sum to one.

Both methods are unbiased predictors that quantify prediction error variance. However, although the requirement of nonnegative weights is a sufficient condition to obtain nonnegative predictions, it is not a necessary one. Even in the presence of negative weights, predictions can be nonnegative. In accordance with this observation, Barnes and You (1992) extended the ordinary kriging system by imposing upper and lower bounds on the predictions. They concluded that if the ordinary kriging prediction is between the lower and upper bounds, it is also the optimal solution for the bounded ordinary kriging system. If on the other hand, the ordinary kriging prediction violates one of the bounds, then the bounded ordinary kriging prediction is equal to the violated bound, and the prediction error variance is increased accordingly.

Although Szidarovszky, Baafi, and Kim (1987), Barnes and You (1992), and Deutsch (1996) solved the problem of nonnegative predictions, their methods only warrant the constant sum constraint when the semivariograms are proportional in scale [Eq. (1)]. Practically speaking, these methods are only applicable when $p = 2$.

In short, none of the methods considered so far complies with all requirements for an appropriate interpolation method given in the Introduction. Therefore, compositional kriging is proposed, i.e. a spatial predictor that fully complies with the properties of compositional data. Furthermore, it uses the spatial covariance structure to derive the optimal set of weights $\mathbf{W}$. Like ordinary kriging, compositional

kriging is an unbiased predictor that minimizes prediction error variance. It may be argued however, that a biased maximum likelihood predictor may be more appropriate for characterizing skewed distributions like those of compositional data. Furthermore, the prediction error variance is hard to interpret when the objective is to construct confidence intervals (Pawlowsky-Glahn and Barcelo-Vidal, 1999). Therefore, in the derivation of the compositional kriging system given below, the prediction error variance is merely considered as an objective function in a minimization problem rather than a measure for constructing confidence intervals.

COMPOSITIONAL KRIGING

Compositional kriging is a straightforward extension of ordinary kriging. Therefore ordinary kriging can be taken as a starting point for the derivation of the compositional kriging system. The aim of ordinary kriging is to minimize the prediction error variance subject to the unbiasedness constraint (Journel and Huijbregts, 1978; Isaaks and Srivastava, 1989):

$$\begin{aligned} \min_{\mathbf{w}_k} \quad & \left(\sigma_k^2 + \mathbf{w}_k^{\mathrm{T}}\mathbf{C}_k\mathbf{w}_k - 2\mathbf{w}_k^{\mathrm{T}}\mathbf{d}_k\right) \\ \text{s.t.} \quad & \mathbf{w}_k^{\mathrm{T}}\mathbf{1}^{(n)} = 1 \end{aligned}$$

where σ_k^2 is the variance of the kth component of $\mathbf{z}(x_i)$, $\mathbf{w}_k$ is the kth column of $\mathbf{W}$, $\mathbf{C}_k$ is the $n \times n$ matrix containing the covariances between the data points for component k, and $\mathbf{d}_k$ is the vector of dimension n containing the covariances between the data points and the prediction point for component k. This constrained optimization problem can be converted into an unconstrained one by adding the unbiasedness constraint with Lagrange multiplier μ_k to the objective function. The resulting objective function, i.e. the Lagrangian, can be minimized by setting its partial first derivatives with respect to the weights and the Lagrange multiplier equal to zero. This results in the ordinary kriging system:

$$\begin{cases} \mathbf{C_k}\mathbf{w}_k + \mu_k\mathbf{1}^{(n)} = \mathbf{d}_k \\ \qquad \mathbf{w}_k^{\mathrm{T}}\mathbf{1}^{(n)} = 1 \end{cases}$$

Solving this system for $\mathbf{w}_k$ and μ_k yields an optimal set of weights for each component k. Optimal in this sense refers to weights that lead to an unbiased predictor with minimum prediction error variance.

However, since the constant sum and nonnegativity constraints are not guaranteed, these weights are only optimal for each component k separately, and not

necessarily for the composition as a whole. Therefore compositional kriging considers all components simultaneously by minimizing the sum of their prediction error variances, and by taking the unbiasedness, nonnegativity, and constant sum constraints into account:

$$\begin{aligned} \min_{\mathbf{w}_k} \quad & \sum_{k=1}^{p} \left(\sigma_k^2 + \mathbf{w}_k^{\mathrm{T}} \mathbf{C}_k \mathbf{w}_k - 2\mathbf{w}_k^{\mathrm{T}} \mathbf{d}_k\right) \\ \text{s.t.} \quad & \mathbf{W}^{\mathrm{T}} \mathbf{1}^{(n)} = \mathbf{1}^{(p)} \\ & \mathbf{w}_k^{\mathrm{T}} \mathbf{z}_k \geq 0 \quad \text{for } k = 1, \ldots, p \\ & \mathrm{tr}(\mathbf{W}^{\mathrm{T}} \mathbf{Z}) = 1 \end{aligned}$$

where $\mathbf{z}_k$ represents the kth column of $\mathbf{Z}$, and tr(·) gives the trace of its argument. Since this optimization problem also contains inequality constraints, its solution is more complicated than that of ordinary kriging. Fortunately, the concept of active constraints (Wismer and Chattergy, 1978) is very useful in this respect. It can be illustrated by means of a simple univariate optimization problem:

$$\begin{aligned} \min_{x} \quad & f(x) \\ \text{s.t.} \quad & x \geq 0 \end{aligned}$$

where $f(x)$ is a convex quadratic function of x. At point x^* satisfying $\min_x f(x) = f(x^*)$ subject to $x^* \geq 0$ it follows from Figure 1 that

$$f'(x^*) = 0$$

or

$$f'(x^*) > 0 \quad \text{and} \quad x^* = 0$$

where $f'(x^*) = df(x^*)/dx^*$. In other words, the following conditions must hold at the minimum:

$$\begin{cases} f'(x^*) + \alpha^* = 0 \\ x^* \geq 0 \\ \alpha^* \leq 0 \\ \alpha^* x^* = 0 \end{cases}$$

where α^* is a Lagrange multiplier. These results are called the Kuhn–Tucker stationary conditions (Wismer and Chattergy, 1978). The inequality constraint is said

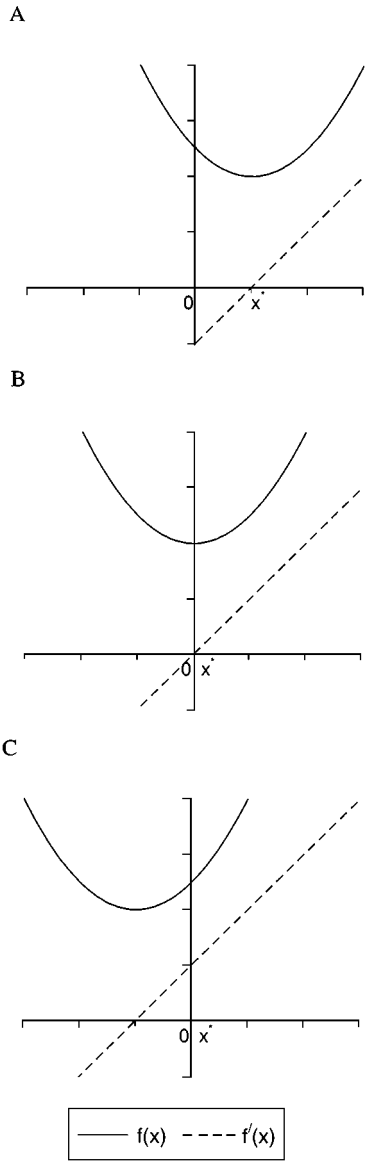

Figure 1. Three convex quadratic functions $f(x)$ (solid lines) and their first derivatives $f'(x)$ with respect to x (dashed lines).

to be *active* if $\alpha^* < 0$ and consequently $f'(x^*) > 0$ and $x^* = 0$ (Fig. 1C). On the other hand, it is *inactive* if $\alpha^* = 0$ and $f'(x^*) = 0$ (Figs. 1(A) and 1(B)). Hence, active inequality constraints can be considered as equality constraints, whereas inactive inequality constraints can be left out of consideration.

Analogously, the Kuhn–Tucker conditions for the compositional kriging optimization problem are given by

$$\begin{cases} \mathbf{C}_k\mathbf{w}_k + \mu_k\mathbf{1}^{(n)} + \alpha_k\mathbf{z}_k + \beta_k\mathbf{z}_k = \mathbf{d}_k & \text{for } k = 1, \ldots, p \\ \mathbf{W}^{\mathrm{T}}\mathbf{1}^{(n)} = \mathbf{1}^{(p)} & \\ \mathrm{tr}(\mathbf{W}^{\mathrm{T}}\mathbf{Z}) = 1 & \\ \mathbf{w}_k^{\mathrm{T}}\mathbf{z}_k \geq 0 & \text{for } k = 1, \ldots, p \\ \alpha_k \leq 0 & \text{for } k = 1, \ldots, p \\ \alpha_k\left(\mathbf{w}_k^{\mathrm{T}}\mathbf{z}_k\right) = 0 & \text{for } k = 1, \ldots, p \end{cases}$$

where α_k, β, and μ_k are Lagrange multipliers pertaining to the nonnegativity, the constant sum, and the unbiasedness constraints, respectively. The resulting set of equations is the compositional kriging system. Clearly, if the active inequality constraints were known in advance, the solution of the compositional kriging system would be rather straightforward. Wismer and Chattergy (1978) provide an effiecient iterative algorithm to find these active constraints. This algorithm, known as the method of Theil and Van de Panne, starts with solving the compositional kriging system with all inequality constraints removed. Its solution is optimal if no inequality constraints are violated. Otherwise, combinations of the violated inequality constraints are added iteratively as equality constraints to the compositional kriging system until the optimal solution is found.

Compositional Kriging was first described and applied by De Gruijter, Walvoort, and Van Gaans (1997) who used it to interpolate fuzzy membership vectors. Also, compositional kriging appears to be a promising alternative for indicator kriging, because order relation problems (Isaaks and Srivastava, 1989) are implicitly taken into account.

CASE STUDY I: WESEPE

Compositional kriging was applied to interpolate the clay, silt, and sand contents of the top 30 cm soil layer of a 1200 ha region near the village of Wesepe in the eastern part of the Netherlands. This gently undulating Pleistocene cover-sand area is characterized by east-southeast to west-northwest running parallel ridges separated by small stream valleys. Field estimates of the textural composition of 652 soil profiles were made from auger borings by an experienced soil surveyor. A total of 415 augering points were located on a regular grid of 18 by 24 nodes, with 17 points missing because they were located in a built-up area. The distance between adjacent grid rows and columns was approximately 167 m. In addition to the grid points, 137 augering points were selected randomly to improve the estimation

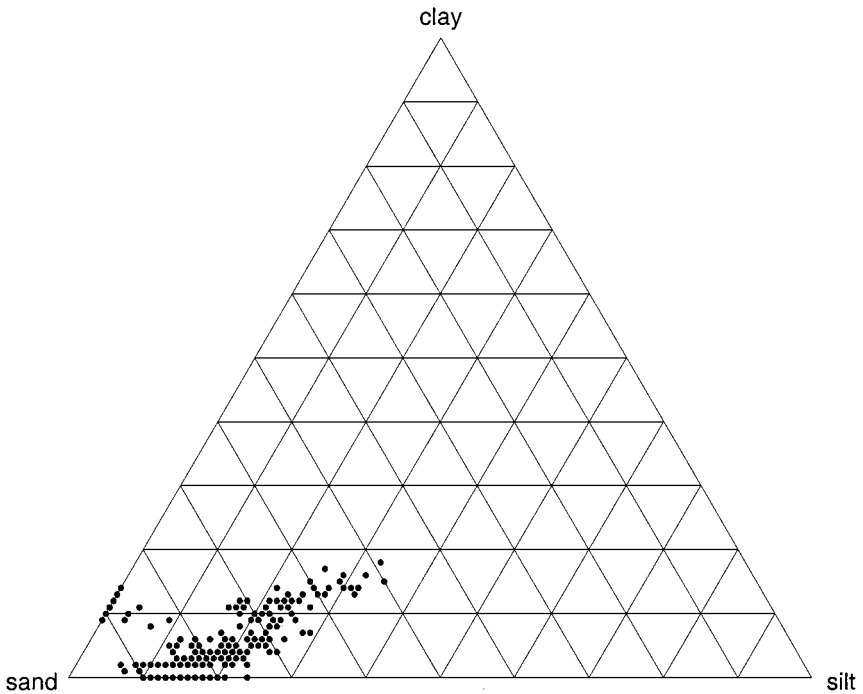

Figure 2. Ternary diagram showing the distributions of the clay, silt, and sand contents estimated in the field at the 552 grid and semivariogram points. The vertices represent 100% of clay, silt, and sand respectively in clockwise direction from the top, while the sides of the triangle represent 0% of clay, silt, and sand respectively in clockwise direction from the bottom.

of the semivariogram at small lag distances. Furthermore, augerings were made at 100 randomly selected points for validation purposes. These validation data were excluded from semivariogram analysis and kriging. The compositions estimated in the field at the grid and additional semivariogram points are given in the ternary diagram of Figure 2.

The alr-transform and compositional kriging were applied to predict compositions at the 100 validation points. The basis method could not be applied, because the regionalized size $s(x_i)$ could not be quantified. The 552 grid and semivariogram data were used as conditioning data.

The spatial structures revealed by semivariogram analysis on the untransformed data were clearly anisotropic. The direction of maximum spatial continuity coincided with the orientation of the stream valleys. However, the semivariograms computed for the alr-transformed data were isotropic. The semivariogram parameters are given in Table 1. (Refer to Journel and Huijbregts, 1978, for definitions of these models.)

Table 1. Semivariogram Model Parameters for the Wesepe Data Set

Texture	Model	c_0 (%^{2})	c_1 (%^{2})	$a_{\max}$ (m)	$a_{\min}$ (m)	ϕ (°)
Clay	Exponential	1.91	9.84	261	109	105
Silt	Exponential	1.13	26.9	188	103	105
Sand	Exponential	12.4	46.3	338	220	105
Alr(clay)	Spherical	8.15	26.2	116	116	—
Alr(silt)	Spherical	0.00	8.83	174	174	—

Note. Parameter a is given for the direction of maximum ($a_{\max}$) and minimum ($a_{\min}$) continuity. Angle ϕ gives the direction of maximum continuity in clockwise direction from the North.

For both compositional kriging and the alr-transform, Aitchison's distance between the predicted $\hat{\mathbf{z}}(x_i)$ and observed $\mathbf{z}(x_i)$ was computed for all validation points x_i. It is defined as (Martin-Fernandez, Barcelo-Vidal, and Pawlowsky-Glahn, 1998; Martin-Fernandez, Olea-Meneses, and Pawlowsky-Glahn, submitted)

$$D_{\mathrm{A}}(\mathbf{z}(x_i), \hat{\mathbf{z}}(x_i)) = \sqrt{\sum_{k=1}^{p} \left[\ln\left(\frac{z_k(x_i)}{\left[\prod_{j=1}^{p} z_j(x_i)\right]^{1/p}} \right) - \ln\left(\frac{\hat{z}_k(x_i)}{\left[\prod_{j=1}^{p} \hat{z}_j(x_i)\right]^{1/p}} \right) \right]^2} \tag{3}$$

and meets all criteria Aitchison (1992) formulated for measures of compositional difference. At two observation points, $\mathbf{z}(x_i)$ contained a zero component. These points were removed from the data set because Eq. (3) is only defined for positive components. A scatter plot of D_{A} for the alr-transform versus D_{A} for compositional kriging is given in Figure 3. A one-tailed paired difference t test showed that the null hypothesis of no difference between the average D_{A} for compositional kriging and the alr-transform should be rejected ($p \approx 0$). It can be concluded that for the Wesepe case study, predictions obtained with compositional kriging were more accurate than those obtained with the alr-transform.

CASE STUDY II: WALKER LAKE

The second case study concerns a composition derived from a digital elevation model (DEM) of the Walker Lake area in Nevada. Two data sets were constructed, an exhaustive data set for validation purposes, and a much smaller sample data set. The exhaustive data were located at the nodal points of a regular grid of 300 rows by 260 columns which was superimposed on the DEM. Isaaks and Srivastava (1989) describe how at each nodal point i two artificial variables $U(x_i)$ and $V(x_i)$ were derived as a function of 25 elevation points in the neighborhood

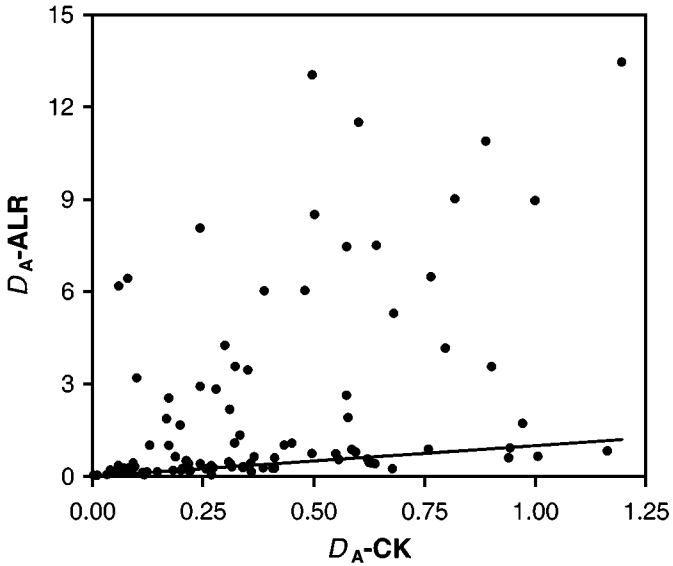

Figure 3. Aitchison's distance for the alr-transform (D_A-ALR) and compositional kriging (D_A-CK) for the Wesepe case study. The solid line is the 1:1 line.

of x_i. To obtain a regionalized composition, we added a third variable $W(x_i)$, given by

$$W(x_i) = M - U(x_i) - V(x_i)$$

where M is the maximum sum of $U(x_i)$ and $V(x_i)$ on the grid. A regionalized composition $\mathbf{z}(x_i)$ is then constructed by

$$\mathbf{z}(x_i) = [U'(x_i), V'(x_i), W'(x_i)]^{\mathrm{T}} = \frac{c}{M}[U(x_i), V(x_i), W(x_i)]^{\mathrm{T}}$$

where $c = 100\%$.

The sample data set was obtained by sequential sampling of the exhaustive grid, which has an assumed grid spacing of 1 distance unit. In the first phase, 195 samples were selected on a regular grid with a grid spacing of 20 distance units that was superimposed on the exhaustive grid. During the remaining phases, sampling was restricted to the neighborhood of target points, i.e. points sampled during previous phases where $V(x_i) > 600$ units. In the second phase, 8 additional samples were taken corresponding to a grid spacing of 10 around all target points of the first phase. The first two phases are supposed to reveal the major patterns in the study area. The aim of the third phase was to delineate these patterns more

Table 2. Semivariogram Model Parameters for the Walker Lake Data Set

	Model	c_0 (%^{2})	c_1 (%^{2})	a_{max}	a_{min}	ϕ (°)
U	Spherical	39.0	6.24	18.4	5.74	166
V	Spherical	2.94	6.33	52.7	29.9	166
W	Spherical	49.8	30.6	39.5	30.2	166
Alr(U)	Spherical	2.77	15.5	17.3	9.0	166
Alr(V)	Spherical	6.07	12.3	22.1	9.9	166

Note. Parameter a is given for the direction of maximum (a_{max}) and minimum (a_{min}) continuity. Angle ϕ gives the direction of maximum continuity in clockwise direction from the North.

clearly by sampling the exhaustive grid 5 distance units to the east and 5 distance units to the west of target points. The second and third phase yielded 147 and 114 sample points, respectively, resulting in a total of 456 sample points.

Anisotropic spherical models were fitted to the sample semivariograms. The model parameters are listed in Table 2. Compositional kriging and the alr-transform were applied to predict compositions at the unsampled nodes of the exhaustive grid. Again, the basis method could not be applied because no information was available on the regionalized size $s(x_i)$.

Aitchison's distance was computed between the predicted and true compositions. Summary statistics are given in Table 3. Because Aitchison's distance is only defined for positive components, compositions with zero components were not considered. Table 3 shows that the alr-transform performed slightly better than compositional kriging. There was no need for statistical testing, because the entire population (except for a relatively small sample) was used for validation.

CONCLUSIONS

Compositional kriging is an unbiased predictor that minimizes the prediction error variance and that complies fully with the nonnegativity and constant sum constraints of compositional data. In the Wesepe case study compositional kriging resulted in significantly more accurate results than the alr-transform whereas in

Table 3. Summary Statistics of Aitchison's Distance for the Walker Lake Validation Study

Method	N	Mean	p-25	Median	p-75
CK	70140	1.59	0.61	1.17	2.20
Alr	71617	1.47	0.55	1.09	1.94

Note. N: number of validation points, p-25 and p-75 are the lower and upper quartiles respectively, CK: compositional kriging, Alr: additive logratio-transform.

the Walker Lake study it gave comparable results. Compositional kriging has the additional advantages that it can handle compositions with zero components and that it does not need a regionalized basis like the basis method.

ACKNOWLEDGMENTS

This research was supported by the Technology Foundation STW (applied science division of NWO and the technology program of the Ministry of Economic Affairs of the Netherlands) and the research programme 358 (GIS & Remote Sensing) of the Dutch Ministry of Agriculture. We thank Marc Bierkens, Vera Pawlowsky-Glahn, Gert Jan Weltje, and Raphael Viscarra Rossel for their useful comments on an earlier draft of this paper.

REFERENCES

Aitchison, J., 1986, The statistical analysis of compositional data: Chapman and Hall, London, 416 p.

Aitchison, J., 1992, On criteria for measures of compositional difference: Math. Geol., v. 24, no. 4, p. 365–379.

Barnes, R. J., and You, K., 1992, Adding bounds to kriging: Math. Geol., v. 24, no. 2, p. 171–176.

Bezdek, J. C., Ehrlich, R., and Full, W., 1984, FCM: The fuzzy c-means clustering algorithm: Comput. Geosci., v. 10, nos. 2/3, p. 191–203.

Bierkens, M. F. P., and Burrough, P. A., 1993a, The indicator approach to categorical soil data. I Theory: J. Soil Sci., v. 44, p. 361–368.

Bierkens, M. F. P., and Burrough, P. A., 1993b, The indicator approach to categorical soil data. II Application to mapping and land use suitability analysis: J. Soil Sci., v. 44, p. 369–381.

Burrough, P. A., Van Gaans, P. F. M., and Hootsmans, R., 1997, Continuous classification in soil survey: Spatial correlation, confusion and boundaries: Geoderma, v. 77, no. 2–4, p. 115–135.

Davis, J. C., 1986, Statistics and data analysis in geology, 2nd edn: Wiley, New York, 646 p.

De Gruijter, J. J., and McBratney, A. B., 1988, A modified fuzzy k-means method for predictive classification, *in* Bock, H., ed., Classification and related methods of data analysis: Elsevier, Amsterdam, p. 97–104.

De Gruijter, J. J., Walvoort, D. J. J., and Van Gaans, P. F. M., 1997, Continuous soil maps—a fuzzy set approach to bridge the gap between aggregation levels of process and distribution models: Geoderma, v. 77, no. 2–4, p. 169–195.

Deutsch, C. V., 1996, Correcting for negative weights in ordinary kriging: Comput. Geosci., v. 22, no. 7, p. 765–773.

Isaaks, E. H., and Srivastava, R. M., 1989, An introduction to applied geostatistics: Oxford University Press, New York, 561 p.

Journel, A. G., 1983, Nonparametric estimation of spatial distributions: Math. Geol., v. 15, no. 3, p. 445–468.

Journel, A. G., and Huijbregts, Ch. J., 1978, Mining geostatistics: Academic Press, London, 600 p.

Martin-Fernandez, J. A., Barcelo-Vidal, C., and Pawlowsky-Glahn, V., 1998, Measures of difference for compositional data and hierarchical clustering methods, *in* Buccianti, A., Nardi, G., and Potenza, R., eds., Proceedings of IAMG'98, p. 526–539.

Martin-Fernandez, J. A., Olea-Meneses, R. A., and Pawlowsky-Glahn, V., submitted, Criteria to compare estimation methods of regionalized compositions: De Frede Publishing, Naples, Submitted.

Myers, D. E., 1982, Matrix formulation of co-kriging: Math. Geol., v. 14, no. 3, p. 249–257.

Odeh, I. O. A., McBratney, A. B., and Chittleborough, D. J., 1992, Fuzzy-c-means and kriging for mapping soil as a continuous system: Soil Sci. Soc. Am. J., v. 56, p. 1848–1854.

Olea, R. A., Pawlowsky, V., and Davis, J. C., 1993, Volumetric calculations in an oil field: The basis method: Comput. Geosci., v. 19, no. 10, p. 1517–1527.

Pawlowsky, V., 1989, Cokriging of regionalized compositions: Math. Geol., v. 21, no. 5, p. 513–521.

Pawlowsky-Glahn, V., and Barcelo-Vidal, C., 1999, Confidence regions in ternary diagrams (theoretical foundations and practical applications), *in* Ristedt, H., ed., Terra Nostra: Schriften der Alfred-Wegener-Stiftung 99/1, p. 37–47.

Pawlowsky, V., and Burger, H., 1992, Spatial structure analysis of regionalized compositions: Math. Geol., v. 24, no. 6, p. 675–691.

Pawlowsky, V., Olea, R. A., and Davis, J. C., 1994, Additive logratio estimation of regionalized compositional data: An application to calculation of oil reserves, *in* Dimitrakopoulos, R., ed., Geostatistics for the next century: Kluwer Academic, The Netherlands, p. 371–382.

Pawlowsky, V., Olea, R. A., and Davis, J. C., 1995, Estimation of regionalized compositions: A comparison of three methods: Math. Geol., v. 27, no. 1, p. 105–127.

Szidarovszky, F., Baafi, E. Y., and Kim, C., 1987, Kriging without negative weights: Math. Geol., v. 19, no. 6, p. 549–559.

Wismer, D. A., and Chattergy, R., 1978, Introduction to nonlinear optimization. A problem solving approach: Elsevier North-Holland, Amsterdam, The Netherlands, 395 p.