THE LINEAR COREGIONALIZATION MODEL AND SIMULTANEOUS DIAGONALIZATION OF THE VARIOGRAM MATRIX FUNCTION

Donald E. Myers
Department of Mathematics
University of Arizona
Tucson, AZ 85721 USA
myers@math.arizona.edu

ABSTRACT

There are at least two general approaches for reducing the analysis of a multivariate data set to the analysis ofa a univariate data set. One corresponds to the analysis of one or more linear combinations of the components in which case the problem is to determine appropriate linear combinations. A second approach is to represent the various components as linear combinations of uncorrelated parts so that each may be analyzed separately. These two approaches are not unrelated and both have been used in geostatistics.

The Linear Coregionalization Model assumes that each variate, or component in the vector random function, is representable as a linear combination of uncorrelated components. These are normally taken to be represented by covariance or variogram models of the same type but with different ranges. The Linear Coregionalization model has the advantage that the positive definiteness condition reduces to a test for positive definiteness of a constant matrix, it has the disadvantage of restricting the choice for the cross-variogram, cross-covariance models and it is most useful in the case of undersampling.

More generally if each component is representable as a linear combination of uncorrelated components this corresponds to a simultaneous diagonalization of the matrix structure function. This formulation has application in image analysis and is related to the near simultaneous diagonalization of multiple matrices.

1.0 INTRODUCTION

In order to apply cokriging in one of its various forms it is necessary to estimate and model the matrix variogram or matrix covariance function. The diagonal entries are variograms, respectively (auto)covariances, and the off-diagonal entries are cross-variograms, respectively cross-

---

in A.G. Fabbri et J.J. Royer (eds), 3rd CODATA Conference on Geomathematics and Geostatistics. Sci. de la Terre, Sér. Inf., Nancy, 1994, 32, 125-139

covariances. In addition to the question of which form of cross-variogram to use, Myers (1992), Cressie (1993), there is the more difficult practical question of how to model these cross-variograms or cross-covariances. One of the reasons that (univariate) geostatistical estimation methods are practical is that there is a ready supply of known valid variograms/covariances. With any choice of these valid variogram models the kriging system is solvable and the kriging variance is positive. More generally it is sufficient to consider a positive linear combination of such known models to ensure that the kriging variance will be positive and that the coefficient matrix in the kriging equations will be invertible. These restrictions are relatively easy to incorporate into a kriging program.

Unfortunately the problem is much more complicated for matrix functions. We do not have a set of known set of valid matrix functions sufficient to generate a wide class of valid models. It will be shown later that even the question of valid linear combinations is more complicated for matrix functions. The problem is also related to the question of what constitutes a valid cross-variogram. That question does not have a simple answer since the cross-variogram must be considered in connection with the associated pair of variograms, Myers(1984, 1988a). Since the easiest way to model a matrix function is to model the individual components we require a method for modeling cross-variograms. Procedures for estimating and modeling the diagonal entries , i.e., variograms or covariances are known even if they are not perfect. The obvious step is to compute a sample cross-variogram, however it is not at all clear what a valid cross-variogram will look like. It is not sufficient to choose a valid variogram/covariance model to model cross-variograms. A number of necessary conditions are known such as the Cauchy-Schwartz inequality but sufficient conditions are more difficult to apply. As a practical solution to this problem, Matheron introduced the linear coregionalization model (LCM) which generalizes the idea of a positive linear combination of valid models. Instead of having to test a positive definiteness condition for a matrix valued function it is sufficient to test several (constant) matrices for positive definiteness. It is still necessary to: (1) determine the number of such matrices, (2) estimate the entries in each of these matrices and (3), estimate and model a corresponding number of variograms. Wackernagel (1985, 1988) proposed using a form of principal components together with a graphical analysis to perform these steps. More recently Goulard and Voltz (1992) have given a variation of Wackernagel's algorthim. Bourgault and Marcotte(1991) have proposed the use of a multivariable variogram. This is the variogram of a particular linear combination of the components. It is useful to re-examine the foundations of the linear coregionalization model and the implications of its use as well as connections with other applications of the matrix structure function such as in image analysis.

2.0 THE GENERAL FORM OF THE LINEAR MODEL

Let $Z(x) = [Z_1(x),...,Z_m(x)]$ be a vector valued random function. Let $Y_1(x), ..., Y_p(x)$ be uncorrelated random functions where p may be greater or less than m . In general the same stationarity assumptions will be imposed on the Y's as on the components of Z. Conversely, the stationarity assumptions imposed on the Y's will imply the same for the components of Z. Suppose now that

$$Z_j(x) = \Sigma Y_k(x)a_{kj} ; j = 1,....., m \qquad (1)$$

In matrix form this becomes

$$[Z_1(x),...,Z_m(x)] = [Y_1(x), ..., Y_p(x)]A \qquad (1')$$

It is easily seen that the spatial structure functions for Z are related to those of the Y's in one of several forms including

$$C_Z(h) = A^T C_Y(h) A \tag{2}$$

and

$$\gamma_Z(h) = A^T \gamma_Y(h) A \tag{3}$$

where $C_Z(h)$, $C_Y(h)$ denote the matrix covariance functions with entries

$$C_{st,Z}(h) = \mathrm{Cov}\{Z_s(x+h), Z_t(x)\} \tag{4}$$

$$C_{uv,Y}(h) = \mathrm{Cov}\{Y_u(x+h), Y_v(x)\}$$

$$= 0, \quad u \neq v \tag{5}$$

and $\gamma_Z(h)$, $\gamma_Y(h)$ the matrix variograms with entries

$$\gamma_{st,Z}(h) = \mathrm{Cov}\{Z_s(x+h) - Z_s(x), Z_t(x+h) - Z_t(x)\} \tag{4'}$$

$$\gamma_{uv,Y}(h) = \mathrm{Cov}\{Y_u(x+h) - Y_u(x), Y_v(x+h) - Y_v(x)\}$$

$$= 0, \; u \neq v \tag{5'}$$

Equations (4') and (5') utilize the standard cross-variogram rather than the pseudo-cross variogram given in Myers (1991). Equations (2) and (3) may be written in non-matrix form and then it is easy to see certain special cases, in particular the linear coregionalization model. Applying equations (4) and (5), or equations (4') and (5'), to the representations given in (1) we obtain

$$C_{st,Z}(h) = \Sigma a_{su} C_{uu,Y}(h) a_{ut} = \Sigma b_{st}^{u} C_{uu,Y}(h) \tag{6}$$

or

$$C_Z(h) = \Sigma B^u C_{u,Y}(h) \tag{7}$$

This is the way the LCM is usually written. Let A_u be the u^{th} column of A then $B^u = A_u^T A_u$. With this construction the B^u's will automatically satisfy the positive definiteness condition which is required for the LCM. It is also easy to see that if a matrix covariance function, or a matrix variogram, is constructed as in (7) and the B^u's are positive definite then the matrix function will satisfy the appropriate positive definiteness condition as described in Myers (1984, 1998a) since the test reduces to showing that a positive linear combination of covariances, or a positive linear combination of variograms, is again a covariance (respectively a variogram). This is, of course, known to be sufficient but it is not a necessary condition. Since the Y's are not directly obervable and hence the covariances/variograms are not directly estimable, various techniques have been proposed for determining the number of Y's, their covariances and the matrices B^u.

3. LINEAR COMBINATIONS

Since positive linear combinations of valid variograms result in valid models it is tempting to consider the same form of construction for matrix functions. However for matrix functions there are at least two different possibilities, one more general than the other. The LCM illustrates both at the same time. It is useful to view that model from both perspectives.

Let $g_1(h),\ldots,g_p(h)$ be conditionally positive definite matrix functions (m x m) as defined in Myers (1984), for generalizations see Myers (1988a, 1992). That is, if $x_1,\ldots,x_n$ are points in Euclidean space and $\Gamma_1,\ldots,\Gamma_n$ are m x m matrices whose sum is the zero matrix then

$$\Sigma\Sigma\Gamma_i^T g_s(x_i-x_j)\Gamma_j > 0; \quad s=1, \ldots, p \tag{8}$$

Definition 1. Let $a_1,\ldots,a_p$ be positive constants (scalars), then

$$a_1 g_1(h) + \ldots + a_p g_p(h) \tag{9}$$

is called a scalar positive linear combination of the $g_1(h),\ldots,g_p(h)$

It is easy to show that such a linear combination is again conditionally positive definite since the trace of a sum is the sum of the traces and matrix multiplication is both left and right distributive. This is the most obvious form of a linear combination of matrix functions.

Defintion 1'. Let $A_1,\ldots,A_p$ be arbitrary (real) m x m matrices then

$$A_1^T g_1(h) A_1 + \ldots + A_p^T g_p(h) A_p \tag{10}$$

is called a matrix positive linear combination of the $g_1(h),\ldots,g_p(h)$.

That such a combination is a valid model is again an easy consequence of the definition of conditional positive definiteness and Definition 1'. For a valid LCM it is sufficient that the matrices B_u be positive definite and hence factorable as $A_u^T A_u$. The LCM can then be viewed as a linear combination under either defintion. For Definition 1 we take $g_k(h)$ to be $A_u^T A_u \gamma(h)$ where $\gamma(h)$ is a valid variogram, in this case the a_j's are simply ones. With respect to Definition 1', the $g_k(h)$'s are scalar matrices, i.e., the product of a valid variogram and the identity matrix. Note that if each of the g_k's is an LCM then under either definition, a linear combination is again an LCM.

Unfortunately it would not be simple nor sufficient to provide a list of valid models in a cokriging program and allow the use to input the coefficient matrices. In part this is because of the variability in the matrix size and in part because of the need to combine different variogram models within a given matrix function. The complexity of the user interface to accomplish this in a cokrging program would be non-trivial.

3. THE MULTIVARIABLE VARIOGRAM

Bourgault and Marcotte (1991) have proposed the use of a multivariable variogram which, in the case of an LCM and an appropriate choice of the matrix M, reduces to the sum of the variograms of the Y's. Suppose that Z satisfies the intrinsic hypothesis. Let M be a symmetric matrix and form

$$G(h) = 0.5E\{[Z(x+h)-Z(x)]M[Z(x+h)-Z(x)]^T\} \tag{11}$$

The multivariable variogram is simply the variogram of a particular linear combination of the components of Z. Let B be an m x 1 vector and set $W(x) = Z(x)B$, moreover $W(x+h) = Z(x+h)B$.

Both W(x) and W(x+h) are scalar valued. Since Z is intrinsic, W will be. Then

$$
\begin{aligned}
\gamma_W(h) &= 0.5E\{[W(x+h)-W(x)]2\} \\
&= 0.5E\{[W(x+h)-W(x)][W(x+h)-W(x)]^T\} \\
&= 0.5E\{[Z(x+h)-Z(x)]BB^T[Z(x+h)-Z(x)]^T\} \\
&= 0.5ETr\{B^T[Z(x+h)-Z(x)]^T[Z(x+h)-Z(x)]B\} \\
&= Tr\{B^T\gamma_Z(h)B\} \qquad (12)
\end{aligned}
$$

which was noted by Bourgault and Marcotte. If $M = BB^T$ then $\gamma_W(h)$ is the multivariable variogram. Since it is essential to require that M be positive definite such a representation is always possible. The trick is to choose B's that produce useful results. One possibility is a vector which has 1's in two places and zeros elsewhere, in this case G(h) is simply the variogram of the sum of two components and as was noted by Myers (1982) it is the sum of the respective variograms and twice the cross-variogram. Likewise if B is a vector with a 1 in one position, -1 in another and zeros elsewhere then G(h) is the variogram of a difference which is also given in terms of the two variograms and the cross-variogram. In order that the multivariable variogram reduce to the sum of the variograms of the components when Z(x) has the representation given by equation (1), and $M = BB^T$ then B should be chosen so that ABB^TA^T is the identity.

These results will all extend equally well to the sample variograms, both scalar and matrix valued. Let $x_1,\ldots,x_n$ be data locations and form the n x m matrix

$Z(x_1)$

$Z(x_2)$

.

.

.

$Z(x_n)$

which we denote simply by Z. We suppose that all components of Z have data values at all the sample locations (the full sampled case). We could interpret Z as n points in m-dimensional space. The rows of ZB are then projections of these points onto the 1-dimensional space determined by B. In many forms of multivariate analysis the question is; for what choice of B do we obtain "interesting" projections? Is there some number of such projections (usually taken to be less than m) which are sufficient to characterize the data set (or nearly characterize it). The sample multivariable variogram is then a measure of the spread of these projections (of the increment points!). In the case of an LCM the spread is maximized and the measure of the spread is the sum of the eigenvalues of the matrix variogram of Z.

4. DIAGONALIZATION

It is easy to see that one interpretation of equations (2) and (3) is that they correspond to (simultaneous) diagonalization of the the matrix functions. That is, consider the set of matrices corresponding to the set of possible values of h, these matrices are simultaneously diagonalized by A. This is a rather strong condition. Suppose given a symmetric matrix valued function S(h) then for each h there exist matrices F(h) and D(h), D(h) a diagonal matrix , such that

$$S(h) = F(h)^T D(h) F(h) \qquad (13)$$

Simultaneous diagonalization with respect to h would imply that the matrices F do not depend on h. For fixed h, the diagonal entries of D(h) are the eigenvalues of S(h). If S(h) is positive definite for *each h* then the F's and D have additional nice properties. Note that a positive definiteness condition on S(h) for each h is not the sufficient condition for S(h) to be a matrix covariance function, respectively a matrix variogram. However if the diagonal entries of D(h) are covariances or variograms and F is any constant matrix then S(h) will be a matrix covariance function, respectively a matrix variogram.

PRINCIPAL COMPONENTS ANALYSIS

Borgman and Framhe (1976) suggested the use of principal components analysis (PCA) as a method for constructing a variant on the LCM. The Bentonite clay data had eleven variables and data locations. By applying PCA the data set is reconstructed with five factors which explain 88% of the variance. The factors are particular 1-dimensional vectors. They are obtained as particular projections and one can compute the sample variograms for each of these projections. After modeling the variograms for each of the factors these were used to reconstruct the variograms of the original variables. Myers and Carr (1984) compared these results with those obtained by the use of sums and differences. Davis and Green (1984) use PCA in a slightly different manner. They experimentally verified that the cross variograms of the factors were null and hence cokriging of the factors reduced to separate kriging. After kriging the factors, the original variables were reconstructed from the PCA.

PCA has often been used in the analysis of multiband images, particularly to "remove" noise. The data for each band is written as a vector, the vectors are adjoined to form a data matrix and PCA is applied. The small eigenvalues are interpreted as the variances of the noise terms and the data is reconstructed using only the factors corresponding to the non-trivial eigenvalues. There are at least two dis-advantages to this usage. First of all, the factors and the eigenvalues obtained by PCA are left unchanged when rows in the data matrix are interchanged. That is, the factors and the eigenvalues are insensitive to the order in which the pixels are arranged when constructing the data matrix, secondly the covariance/correlation matrix used to generate the factors only quantifies intervariable correlation and not spatial correlation.

Switzer and Green (1984) proposed a variation on this use of PCA. Consider a shift of the image one pixel in the horizontal or the vertical directions. At each pixel one then forms the differences between the original image and the shifted image (each band separately). The image of differences is then arranged as a data matrix and PCA is applied to the increments. This is analagous to computing the sample variogram matrix for one lag in both the horizontal and vertical directions. Switzer and Green averaged these two directional sample variograms before diagonalization, claiming that it was sufficient to consider only one lag and only the all directional variogram. Re-analyzing the same data set however it was found that there is an anisotropy and that the lag 2 and lag 3 sample matrix variograms are of interest.

THE COKRIGING EQUIVALENCE

Suppose that $Z(x) = Y(x)A$ where A is invertible. Let $\gamma_Z(h) = A^T \gamma_Y(h) A$ be the matrix variogram for Z(x) expressed in terms of the matrix variogram for Y(x). The cokriging estimator for $Z(x_0)$ is given by

$$Z^*(x_0) = \Sigma\, Z(x_i)\Gamma_i \tag{14}$$

where

$$\Sigma \gamma_Z(x_i - x_j)\, \Gamma_i \;+ \mu \;= \gamma_Z(x_0 - x_j) \tag{15}$$

and

$$\Sigma\Gamma_i = I \tag{16}$$

These are equivalent to

$$Y^*(x_0) = \Sigma Y(x_i)A\Gamma_i A^{-1} = \Sigma Z(x_i)\Gamma_i A^{-1} \tag{17}$$

$$\Sigma\gamma_Y(x_i\text{-}x_j)A\Gamma_i A^{-1} + (A^T)^{-1}\mu\ A^{-1} = \gamma_Y(x_0\text{-}x_j) \tag{18}$$

$$\Sigma A\Gamma_i A^{-1} = I \tag{19}$$

Note that the $A\,\Gamma_1\,A^{-1},\ldots,A\,\Gamma_n\,A^{-1}$ are diagonal matrices. Since $\gamma_Y(h)$ is a diagonal matrix the latter cokriging system is in fact separate kriging of the components of $Y(x_0)$. Moreover the diagonalizability of the $\gamma_Z(h)$ implies the simultaneous diagonalization of the weight matrices for any choice of the points $x_1,\ldots,x_n$.

MATRIX FUNCTION APPROXIMATION

Although in general the variogram matrix function for Z is unknown, the use of an LCM can be considered as an approximation to the "true" function. There are several ways to describe the approximation, e.g., to quantify the error of approximation and several ways to quantify the consequence of the approximation. Suppose that $\gamma_Z(h)$ denotes the "true" but unknown function and $A^T\gamma_Y(h)A$ is the approximating LCM. In general the difference, $[\gamma_Z(h)\text{-}A^T\ \gamma_Y(h)A]$ will depend on h. $Tr\{[\gamma_Z(h)\text{-}A^T\gamma_Y(h)A]^T[\gamma_Z(h)\text{-}\ \gamma_Y(h)A]\}$ is the sum of the squares of the differences, element by element for fixed h. If this is summed over all the data locations to be used, the result is a measure of the error of approximation, equivalentlyone could use the maximum over all data locations to be used. This does not easily translate into a measure of the differences in the cokriging weight matrices nor of the cokriged values of $Z(x_0)$ however. The next section provides an alternative way of characterizing the approximation.

NEAR DIAGONALIZATION

Let $B_1, \ldots\ldots, B_K$ be m x m positive definite matrices. Then A is said to nearly simultaneously diagonalize the B's if the off diagonal entries in each of the $V_k = A^T B_k A$ are collectively small. There are several ways of quantifying how small they are. Without loss of generality we will assume that A is invertible. Flury and Gautshi (1986), Clarkson (1988) have adopted this approach. Clarkson used the sum of the squares of the off-diagonal entries as a measure of near simultaneous diagonalizability. This formulation has the disadvantage that the "new" matrix functions are not diagonal.

Alternatively suppose that there are diagonal matrices $D_1, \ldots, D_K$ such that each $A^T D_k A$ is "close" to B_k. Let $\gamma_Z(h)$ be an unknown variogram matrix but let $\gamma_Z^*(h_i)$, $i = 1, \ldots, N$ be a corresponding sample variogram matrix. If a least squares measure is used then we would minimize

$$\int \{ Tr[_Z(h) - A^T D(h) A]^T [_Z(h) - A^T D(h) A] \} dP(h)$$

where dP(h) is a probability measure on the domain of $\gamma_Z(h)$. The probability measure provides for a weighting. Unfortunately of course $\gamma_Z(h)$ is not known. This suggests using the sample variogram matrices instead. Then we minimize

$$\Sigma Tr[\gamma_Z^*(h_i) - A^T D(h_i) A]^T [\gamma_Z^*(h_i) - A^T D(h_i) A]$$

This is essentially a least squares fitting of the sample variogram matrix to a known model. One choice of a known model is the LCM and the fitting process produces both the diagonal matrix D and the matrix A. Goulard and Voltz (1992) use this formulation and obtain an iterative agorithm for estimating A and D.

SUMMARY AND CONCLUSIONS

The use of an LCM corresponds to the assumption that the variogram matrix function (covariance matrix function) is diagonalizable. It also implies that the cross structure functions are symmetric, i.e., if G(h) is the matrix function then G(h) = G(-h) whereas more generally G would only satisfy $G(-h) = G(h)^T$. Hence the LCM does not incorporate the use of pseudo-cross variograms or non-symmetric cross covariances such are given in Myers (1991), Ver Hoef and Cressie (1993). If the "true" matrix function is approximated by an LCM then in general the kriging variance will be larger and the estimated values will differ.

The multivariable variogram introduced by Bourgault and Marcotte (1992) corresponds to the variogram of a linear combination of the components of Z. In the case of a LCM there is a particular choice of the linear combination that is useful. The method introduced by Myers (1982) which uses the sum and difference corresponds to other linear combinations. The application of principal components analysis to the original data set is analagous to looking for a linear combination for which the variogram is most informative. Note that such linear combinations do not in general provide sufficient information for modeling non-symmetric cross variograms or cross covariances.

The least squares approximation of the sample variogram matrix function by an LCM introduced by Goulard and Voltz (1992) is analagous to near simultaneous diagonalization of the variogram matrix function.

While the LCM provides for an easy check on the positive definiteness condition for the variogram matrix function, the LCM should in general be considered only as an approximation to the "true" model, an approximation that is severely restricted by the implicit diagonalizability condition and the implicit symmetry condition.

REFERENCES

Bourgault,G. and D. Marcotte,(1991) Multivariable Variogram and its application to the Linear

Model of Coregionalization. Math. Geology 23, 899-928

Bourgault, G. and D. Marcotte, (1993) Spatial Filtering under the Linear Coregionalization Model. in Geostatistics Troia '92, A. Soares (ed), Kluwer Academic Publishers, Dordrecht, 237-248

Carr.J., C. Glass, H. Yang and D.E. Myers,(1987) Application of Spatial Statistics to Analyzing Multiple Remote Sensing Data sets. in Geotechnical Applications of remote sensing and remote data transmission,ASTM STP 967, A.I. Johnson and C.B. Peterson (eds), 138-150

Carr,J. and D.E.,(1984) Myers, Application of the Theory of Regionalized Variables to the Spatial Analysis of LANDSAT data. in the Proceedings of the Ninth Pecora Symposium on Remote Sensing, 2-4 Oct. ,IEEE Computer Society Press

Carr,J., D.E. Myers and C. Glass,(1985) Co-kriging:A Computer program. Computers and Geosciences, 11, 111-127

Clarkson, D.B., (1988) A Least Squares Version of Algorithm 211, the F-G diagonalization algorithm. Applied Statistics 37, 317-321

Flury, B. and W. Gautschi, (1986) An Algorithm for Simultaneous Orthogonal Transformation of Several Positive Definite Symmetric Matrices. SIAM J. Sci. Stat. Comp. 7, 169-184

Goulard,M. and M. Voltz,(1992) Linear Coregionalization:Tools for Estimation and Choice of Cross-Variogram Matrix. Math. Geology 24, 269-286

Myers,D.E.,(1982) Matrix Formulation of Cokriging. Math. Geology, 14, no.3, 249-257

Myers,D.E.,(1983) Estimation of Linear Combinations and Cokriging. Math. Geology, 15, no.5, 633-637

Myers,D.E.,(1984) Cokriging:New Developments. in Geostatistics for Natural Resource Characterization, G.Verly et al, eds., D. Reidel Pub. Co., Dordrecht, 295-305

Myers,D.E.,(1985) Co-kriging:Methods and Alternatives. in The Role of Data in Scientific Progress, P. Glaeser,ed., Elsevier Scientific Pub., 425-428

Myers,D.E.,(1988a) Multivariate Geostatistics for Environmental Monitoring. Sciences de la Terre, 27, 411-427

Myers,D.E.,(1988b) Interpolation with Positive Definite Functions. Sciences de la Terre, 28, 251-265

Myers,D.E.,(1988c) Some Aspects of Multivariate Analysis. in Quantitative Analysis of Mineral and Energy Resources, C.F. Chung et al (eds), D. Reidel Publishing Co., Dordrecht, 669-687

Myers,D.E.,(1989a) Borden Field Data and Multivariate Geostatistics. in Hydraulic Engineering, M.A. Ports (ed), Amer. Soc. Civil Eng. 795-800

Myers,D.E.,(1991a) Pseudo-Cross Variograms, Positive Definiteness and Cokriging. Math. Geology 23, 805-816

Myers,D.E.,(1991b) Multivariate, Multidimensional Smoothing. in Spatial Statistics and Imaging: Proceedings of an AMS-IMS-SIAM Joint Summer Research Conference, June 18-24, 1988, Bowden College, Maine. Lecture Notes-Mongraph Series, Institute of Mathematical Statistics, Hayward, CA 275-285

Myers,D.E.,(1992) Kriging, Cokriging, Radial Basis Functions and the Role of Positive Definiteness. Computers Math. Applic. 24, 139-148

Myers,D.E. and J. Carr,(1985) Cokriging and Principal Component Analysis: Bentonite Data revisited. Sciences de la Terre, 21, 65-77

Rouhani,S. and D.E. Myers,(1990) Problems in Space-Time Kriging of Hydrogeological data. Math. Geology, 22, 611-623

Switzer, P. and Green, (1984) Min/Max Autocorrelation Factors for Multivariate Spatial Imagery. Technical Report, Dept. Statistics, Stanford University, 14p.

Ver Hoef, J.M. and N. Cressie, (1993) Multivariable Spatial Prediction. Math. Geology 25, 219-240

Wackernagel,H.,(1985) L'inference d'un modele lineaire en geostatisticque multivariable. These, Ecole Normale Superieure des Mines de Paris.

Wackernagel, H. (1988) Geostatistical techniques for interpreting multivariate spatial information. in Quantitative Analysis of Mineral and Energy Resources, C.F. Chung et al (eds), D. Reidel Publishing Co., Dordrecht, 393-409

Zhang,R., D.E. Myers and A.W. Warrick,(1992) Estimation of the Spatial Distribution of Soil Chemicals using Pseudo-Cross Variograms. Soil Science Society of America Journal 56, 1444-1452