Mathematical Geology, Vol. 38, No. 8, November 2006 (© 2007)
DOI: 10.1007/s11004-006-9055-7

On the Use of Non-Euclidean Distance Measures in Geostatistics[1]

Frank C. Curriero[2]

In many scientific disciplines, straight line, Euclidean distances may not accurately describe proximity relationships among spatial data. However, non-Euclidean distance measures must be used with caution in geostatistical applications. A simple example is provided to demonstrate there are no guarantees that existing covariance and variogram functions remain valid (i.e. positive definite or conditionally negative definite) when used with a non-Euclidean distance measure. There are certain distance measures that when used with existing covariance and variogram functions remain valid, an issue that is explored. The concept of isometric embedding is introduced and linked to the concepts of positive and conditionally negative definiteness to demonstrate classes of valid norm dependent isotropic covariance and variogram functions, results many of which have yet to appear in the mainstream geostatistical literature or application. These classes of functions extend the well known classes by adding a parameter to define the distance norm. In practice, this distance parameter can be set a priori to represent, for example, the Euclidean distance, or kept as a parameter to allow the data to choose the metric. A simulated application of the latter is provided for demonstration. Simulation results are also presented comparing kriged predictions based on Euclidean distance to those based on using a water metric.

KEY WORDS: conditionally negative definite, euclidean distance, isometric embedding, positive definite, spatial dependence.

INTRODUCTION

Characterizing spatial dependence of random processes via the covariance or variogram function is cornerstone to many geostatistical related applications. Because these functions represent a second moment structure they must be of specific type, positive definite for covariance functions and conditionally negative definite for variograms. Available to practitioners are parametric families of known valid covariance and variogram functions. Under the pragmatic assumptions of stationarity and isotropy, covariance functions and variograms are functions of

[1]Received 5 April 2006; accepted 6 June 2006; Published online: 28 February 2007.
[2]Department of Environmental Health Sciences and Department of Biostatistics, The Johns Hopkins University, Bloomberg School of Public Health, 615 N. Wolfe Street, Baltimore, MD 21205 U.S.A.; e-mail: fcurrier@jhsph.edu

the straight line, Euclidean inter-point distance. There is a large body of literature pertaining to the validity and mathematical characterization of covariance and variogram functions (Christakos, 1984; Schlather, 1999). A topic less covered is the concept of using different (non-Euclidean) measures of inter-point distance to characterize isotropic spatial dependence.

Reasons to consider a non-Euclidean distance could include physical properties of how the process under study disperses or has come to exist in space, such as in the use of geodetic distances on the earth's surface (Cressie, Gotway, and Grondona, 1990, Banerjee, 2005). Sampling non-convex spatial domains such as irregular waterways suggests a water distance measure honoring boundaries and flow patterns (Cressie and Majure, 1997a,b, Little, Edwards, and Porter, 1997; Rathbun, 1998; Kern and Higdon, 2000; Loland and Host, 2003; Krivoruchko and Gribov, 2004; Ver Hoef, Peterson, and Theobald, 2006). Distances based on travel times is another possible consideration (Krivoruchko and Gribov, 2004). In other applications (Dominici, Samet, and Zeger, 2000) focus is on regression coefficients and covariance or variograms functions are commonly used to characterize residual spatial variation, which may be quite complicated, for example due to contagious agents and/or a combination of missing covariates, as well as being dependent on the spatial design of sampled locations. In practice our goal is to characterize spatial dependence as best as possible and consideration to possible non-Euclidean distances to describe proximity relationships among spatial data may prove beneficial.

The purpose of this paper is to demonstrate some of the technical details involved in using a non-Euclidean inter-point distance to characterize isotropic spatial dependence. A simple motivating example is provided in Section 2 to caution against the naive use of a non-Euclidean distance measure with existing covariance and variogram functions. In Section 3, the mathematical concepts of distance metrics and isometric embedding are introduced. These concepts are then integrated with the concepts of conditional negative definiteness and positive definiteness to create classes of valid covariance and variogram functions that can be used with certain non-Euclidean distance measures (Section 4). These functions extend the well known classes by adding a parameter to define the distance measure used. In practice, this distance parameter can be set a priori to represent, for example, the Euclidean distance, or kept as a parameter to allow the data to choose the distance metric. Simulated applications of the latter are provided for demonstration in Section 5 as well as results comparing kriged predictions based on Euclidean distance to those based on using a water metric.

MOTIVATING EXAMPLE

As a motivating example consider a simple four point regular grid configuration in $\Re^2$ with unit spacing, points represented by (x_i, y_i), $i = 1, \ldots, 4$, and

consider using the "city block," distance measure defined by $d_{ij} = |x_i - x_j| + |y_i - y_j|$, as an alternative to the straight line or Euclidean distance measure. This yields the following matrix of inter-point city block distances,

$$\begin{pmatrix} 0 & 1 & 1 & 2 \\ 1 & 0 & 2 & 1 \\ 1 & 2 & 0 & 1 \\ 2 & 1 & 1 & 0 \end{pmatrix},$$

which when used with the Gaussian covariance function, $20\exp(-d_{ij}^2/4)$, nugget, sill, and range parameters arbitrarily set at (0,20,4) respectively, results in the following variance-covariance matrix,

$$\begin{pmatrix} 20.00 & 15.58 & 15.58 & 7.36 \\ 15.58 & 20.00 & 7.36 & 15.58 \\ 15.58 & 7.36 & 20.00 & 15.58 \\ 7.36 & 15.58 & 15.58 & 20.00 \end{pmatrix}.$$

The characteristic roots (eigenvalues) of a positive definite matrix must be positive, and conversely, if one root is negative the matrix cannot be positive definite (Graybill, 1983). The characteristic roots of this matrix are (58.52, 12.64, 12.64, −3.80), implying the Gaussian covariance function is no longer positive definite when used city block distances. Using the same matrix of city block distances and parameter settings, the same conclusion can be drawn from other known covariance functions such as the spherical, rational quadratic, and various forms from the Matern class. On the contrary, the exponential covariance function, $\tau^2 + \sigma^2\exp(-d_{ij}/\phi)$ with positive parameters (τ^2, σ^2, ϕ) remains positive definite in dimensions ≥ 1 when used with the city block distance measure. This fact is straight forward to show since the exponential covariance function with the city block distance measure in $\Re^N$ reduces to the product of one dimensional exponential covariance functions based on the Euclidean distance measure in $\Re^1$ and hence positive definite, a separable covariance function as noted by Cressie (1991, p. 68).

The message from this example is clear, there are no guarantees that the common set of positive definite functions used in geostatistical related applications to represent covariances will remain positive definite (and hence valid) when used with distance measures other than the Euclidean distance. This message also pertains to the pool of known valid isotropic variogram functions (see subsequent text). It is therefore essential that applications involving a non-Euclidean distance measure provide proof that the proposed family of covariance or variogram

functions remain valid when used with the alternative distance measure. Attention to such detail has been less than consistent in the literature.

The water distance used in Cressie and Majure (1997a,b) is actually calculated as though the process was an irregular one dimensional transect by assuming the winding streams have negligible width for their application. In some instances, distances calculated along such a structure can be shown to be equivalent to Euclidean distances along a corresponding "stretched out" one dimensional transect (isometric embedding). However, this representation is lost if the original winding stream structure branches off as it appears to do in their application. Ver Hoef, Peterson, and Theobald (2006) do consider such a stream network (again approximated by assuming zero width) and develop a moving average based covariance function that is shown to be valid with resulting stream distances. Their work also incorporates flow patterns upstream and downstream. The water distance used in Little, Edwards, and Porter (1997) and Rathbun (1998), is computed accounting for water body width, however the validity of its use with known variogram functions is suspect. The water distance used in Kern and Higdon (2000) hinges on satisfying conditions of a metric, which is demonstrated above as not being sufficient. Gneiting (1999) discusses results that justify the great-arc distance used in Cressie, Gotway, and Grondana (1990). The non-Euclidean distance used in Dominici, Samet, and Zeger (2000) is binary, locations within a common geographic region are given a distance one and infinity otherwise. The consequence being that spatial correlation is constant within geographic regions and zero between regions. Such binary distances can always be represented as Euclidean distances between points in some higher dimension, and thus are valid to use provided the correlation function is valid in the higher dimension. This example and the stretched out stream scenario describes the concept of isometric imbedding, the formal definition of which is provided next. Development of valid functions to characterize spatial dependence with non-Euclidean distance measures would be very useful in a variety of geostatistical application areas.

DEFINITIONS AND NOTATION

Let the spatial process be represented by the random field

$$\{Z(\mathbf{s}) : \mathbf{s} \in D \subset \Re^N\},$$

where $\mathbf{s} \in \Re^N$ is a generic spatial location varying continuously over a region D. Characterizing the second moment structure of such processes plays a key role in statistical inference and is usually carried out with the covariance or variogram function, which represents the $Cov(Z(\mathbf{s}_i), Z(\mathbf{s}_j))$ and the $Var(Z(\mathbf{s}_i) - Z(\mathbf{s}_j))$, respectively, $\forall\ \mathbf{s}_i, \mathbf{s}_j \in D$. It is well known that these functions must be of a

specific type, positive definite for covariance functions and conditionally negative definite for variograms. Probably less well known is the connection between these definitions and the concept of isometric embedding (Wells and Williams, 1970). Some general definitions regarding distance measures are provided before reviewing these connections.

Let $\mathbf{S}$ represent an arbitrary collection of objects, such as spatial locations $\mathbf{s} \in \Re^N$, and define the real valued function $d(\cdot, \cdot)$ to represent a distance function operating on $\mathbf{S} \times \mathbf{S}$ such that $d : \mathbf{S} \times \mathbf{S} \to [0, \infty)$. The distance function d is said to satisfy the conditions of a *metric* if:

$$
\begin{aligned}
&d(\mathbf{s}_i, \mathbf{s}_j) \geq 0 \text{ and } d(\mathbf{s}_i, \mathbf{s}_j) = 0 \text{ iff } \mathbf{s}_i = \mathbf{s}_j, \\
&d(\mathbf{s}_i, \mathbf{s}_j) = d(\mathbf{s}_j, \mathbf{s}_i), \text{ and} \\
&d(\mathbf{s}_i, \mathbf{s}_j) \leq d(\mathbf{s}_i, \mathbf{s}_k) + d(\mathbf{s}_k, \mathbf{s}_j) \text{ (Triangle inequality)}
\end{aligned}
$$

for all $\mathbf{s}_i, \mathbf{s}_j, \mathbf{s}_k \in \mathbf{S}$.

For example, consider the Euclidean distance measure, known to be a metric, in $\Re^2$ between points $\mathbf{s}_1 = (x_1, y_1)$ and $\mathbf{s}_2 = (x_2, y_2)$,

$$d(\mathbf{s}_1, \mathbf{s}_2) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}.$$

Clearly the distance between $\mathbf{s}_1$ and $\mathbf{s}_2$, $d(\mathbf{s}_1, \mathbf{s}_2)$, is always positive and zero only when the two locations coincide (condition 1 of a metric). Computing the distance between locations $\mathbf{s}_1$ and $\mathbf{s}_2$ is the same as computing the distance between locations $\mathbf{s}_2$ and $\mathbf{s}_1$ (condition 2 of a metric). The shortest distance between two points is a straight line, so the Euclidean distance between locations $\mathbf{s}_1$ and $\mathbf{s}_2$ would satisfy the triangle inequality (condition 3 of a metric). The city block distance from the previous section is another example of a metric distance. The concept of isometric embedding is now defined.

DEFINITION 1. Let $d_{ij} = d(\mathbf{s}_i, \mathbf{s}_j)$ represent distance between points $\mathbf{s}_i$ and $\mathbf{s}_j$ of some metric space represented by $(\mathbf{S}, d)$. The metric space $(\mathbf{S}, d)$ is said to be *isometrically embedded* in a Euclidean space of dimension N^* if there exists points $\mathbf{s}_i^*$ and $\mathbf{s}_j^*$ and a function ϕ such that

$$d_{ij} = d(\mathbf{s}_i, \mathbf{s}_j) = \|\mathbf{s}_i^* - \mathbf{s}_j^*\|,$$

for all $\mathbf{s}_i, \mathbf{s}_j \in \mathbf{S}$ and where $\phi(\mathbf{s}) = \mathbf{s}^*$.

Isometric embedding in a Euclidean space (hereafter referred to as embedding), thus defines the situation when a metric distance function is equivalent to a Euclidean distance metric. The ramifications for the topic at hand is readily apparent. If a non-Euclidean distance function (meaning non-Euclidean in the dimension the process is observed) is embeddable, then the distance function

used with existing covariance and variogram functions will retain the positive and conditionally negative definite properties provided these functions are valid in the embedding dimension. The winding stream scenario mentioned previously provides an example. Suppose distances between geographic coordinates $\mathbf{s} \in \Re^2$ along the stream are calculated assuming the stream has no width, like traveling through the center of the stream. If the stream doesn't branch off then these distances are equivalent to a set of Euclidean distances calculated from a new set of locations $\mathbf{s}^* \in \Re^1$, located along the stretched out stream now in one dimension. Stream distances under these conditions are thus embeddable in a one dimensional Euclidean space. Contrary to this simple example, the embedding dimension N^* is often assumed to be much larger than the observed dimension N.

Although it is necessary that a distance function d satisfy the conditions of a metric for embedding (since Euclidean distance is a metric), it is clearly not sufficient as was previously demonstrated. The following theorem, due originally to Schoenberg (1937), see also Young and Householder (1938), provides a necessary and sufficient condition for the embedding of a finite metric space.

THEOREM 1. (Schoenberg, 1937). The finite metric space $(\mathbf{S}, d)$, where $\mathbf{S} = \{\mathbf{s}_0, \mathbf{s}_1, \ldots, \mathbf{s}_n\}$ $n > 2$, is embeddable in $\Re^n$ if and only if

$$(1/2)\sum_{i=1}^{n}\sum_{j=1}^{n}\left\{d(\mathbf{s}_0, \mathbf{s}_i)^2 + d(\mathbf{s}_0, \mathbf{s}_j)^2 - d(\mathbf{s}_i, \mathbf{s}_j)^2\right\}\xi_i\xi_j \geq 0 \tag{1}$$

for all choices of real numbers $\xi_0, \xi_1, \ldots, \xi_n$.

As pointed out in (Wells and Williams, 1970), the quadratic form condition (1) is equivalent to

$$\sum_{i=0}^{n}\sum_{j=0}^{n} d(\mathbf{s}_i, \mathbf{s}_j)^2\xi_i\xi_j \leq 0, \tag{2}$$

for all choices of real numbers $\xi_0, \xi_1, \ldots, \xi_n$ such that $\sum_0^n \xi_i = 0$. This is precisely the conditionally negative definite property used to characterize variograms (Christakos, 1984; Cressie, 1991). Therefore, a distance function d is embeddable if and only if d^2 is conditionally negative definite. Put another way, for a given distance function d, $d^{1/2}$ is embeddable, and hence preserves the positive and conditionally negative definite properties of covariance and variogram functions that are valid in all dimensions, if and only if d is conditionally negative definite. This explains another less well known fact that the square root of the variogram (or any conditionally negative definite function) is equivalent to a Euclidean metric.

The embedding and conditionally negative definite property are linked to positive definiteness by the following result (e.g. Wells and Williams, 1970),

$$\exp(-ad(\cdot)) \text{ is positive definite } \forall a > 0 \text{iff } d(\cdot) \text{ is conditionally negative definite} \tag{3}$$

Note, the multiplication and addition by positively restricted parameters (τ^2, σ^2), for example $\tau^2 + \sigma^2 \exp(-ad(\cdot))$ do not change the result.

In practice spatial processes are usually assumed stationary. Letting $\mathbf{h} = \mathbf{s}_i - \mathbf{s}_j, \forall\ \mathbf{s}_i, \mathbf{s}_j \in D$, second-order stationarity is defined for $Z(\mathbf{s})$ by a constant mean and covariance a function of $\mathbf{h}$, denoted by the covariance function $C(\mathbf{h})$. Intrinsic stationarity is defined as a constant mean and variance of the increments $Z(\mathbf{s}_i) - Z(\mathbf{s}_j)$ to be a function of $\mathbf{h}$, denoted by the variogram $2\gamma(\mathbf{h})$, $\gamma(\mathbf{h})$ the semivariogram. Isotropy further assumes covariance functions and variograms to only be a function of distance with the notation $\|\mathbf{h}\|$ used to denote Euclidean distance. Geometric anisotropy refers to the linear transformation of coordinates to achieve isotropy, denoted by $\|\mathbf{Ah}\|$, with matrix $\mathbf{A}$ representing in geostatistical terminology the rotation and stretching transformation of coordinates $\mathbf{h}$ (Cressie, 1991).

As reviewed in the literature, the positive definite property fully characterizes the class of valid covariance functions. Hence, the eigenvalue approach used in Section 2 provides a simple way to exclude candidate models. Valid variograms are necessarily conditionally negative definite as in (2) and also must grow more slowly than $\|\mathbf{h}\|^2$ (Matheron, 1973; Christakos, 1984). Since the square root of a conditionally negative definite function must represent a Euclidean metric, the multidimensional scaling technique of Mardia, Kent, and Bibby (1995, Theorem 14.2.1, p. 397) can be used for verification. By recasting the quadratic form condition (1) into a matrix that must be positive semi-definite, this theorem provides a straight forward computational method for determining if a given distance matrix can be represented as a Euclidean metric. This approach was applied to the motivating example in Section 2 to establish that the Gaussian and other referenced corresponding variograms (excluding the exponential) are no longer conditionally negative definite when used with the city block metric.

NORM DEPENDENT COVARIANCE AND VARIOGRAM FUNCTIONS

There are certain covariance and variogram functions that retain their positive definite and conditionally negative definite properties when used with distance measures other than Euclidean. Many of these results have yet to appear in the mainstream geostatistical literature or application. Introduced first is the concept of a vector norm which is closely related to a distance metric.

A *vector norm* is a function $f : \Re^N \to [0, \infty)$ that satisfies the following properties:

$$\begin{array}{lll} f(\mathbf{h}) \geq 0 & \mathbf{h} \in \Re^N & (f(\mathbf{h}) = 0 \text{ iff } \mathbf{h} = 0) \\ f(\mathbf{h} + \mathbf{h}^*) \leq f(\mathbf{h}) + f(\mathbf{h}^*) & \mathbf{h}, \mathbf{h}^* \in \Re^N & \\ f(\alpha \mathbf{h}) = |\alpha| f(\mathbf{h}) & \alpha \in \Re, \ \mathbf{h} \in \Re^N. & \end{array}$$

A vector norm becomes a distance metric by defining $d(\mathbf{s}_i, \mathbf{s}_j) = f(\mathbf{h})$, with $\mathbf{h} = \mathbf{s}_i - \mathbf{s}_j$ as is the customary notation used in geostatistics. The common α-norm distance metrics for $\alpha \geq 1$ are defined as

$$\|\mathbf{h}\|_\alpha = (|h_1|^\alpha + |h_2|^\alpha + \cdots |h_N|^\alpha)^{1/\alpha}, \tag{4}$$

where $\mathbf{h} = (h_1, \ldots, h_N)'$. When $\alpha = 1, 2$, and ∞, for example, we have

$$\begin{array}{ll} \|\mathbf{h}\|_1 = |h_1| + |h_2| + \cdots |h_N| & \text{(Manhattan or City Block)} \\ \|\mathbf{h}\|_2 = \left(h_1^2 + h_2^2 + \cdots h_N^2\right)^{1/2} & \text{(Euclidean)} \\ \|\mathbf{h}\|_\infty = \text{Max}|h_i| & \text{(Supremum).} \end{array}$$

Note, $\|\mathbf{h}\|$ without the subscript is taken to represent the Euclidean distance and for $\alpha < 1$, $\|\mathbf{h}\|_\alpha$ no longer satisfies the conditions of a metric.

The demonstration following hinges on results from Richards (1985) who provides the following sufficient conditions for which certain power transforms of α-norms are conditionally negative definite. For related mathematical developments see also Koldobskii (1992) and Zastavyni (1993, 2000).

Proposition Richards (1985).

(a) *On* $\Re^2$, $\|\mathbf{h}\|_\alpha^\beta$ *is conditionally negative definite if*

(i) $0 < \beta \leq 1, \ 1 \leq \alpha \leq \infty$, or
(ii) $0 < \beta \leq \alpha \leq 2$.

(b) *On* $\Re^N$, $N \geq 3$, $\|\mathbf{h}\|_\alpha^\beta$ *is conditionally negative definite if*

(i) $0 < \beta \leq \alpha \leq 2$, *and if*
(ii) $\alpha > 2$ *it is not conditionally negative definite for* $\beta > 1$.

These results in combination with Schoenberg's Theorem and (3) can now be used to a create class of valid covariance and variogram functions that can be used with non-Euclidean norm dependent measures of distance. Greater flexibility is gained with processes restricted to $\Re^2$, and since most applications involve analyzing data in $\Re^2$ these extensions are stated separately.

To illustrate, the above results in combination with (3) leads to the following class of norm dependent isotropic powered exponential covariance functions. For $\mathbf{h} \in \Re^2$, the functions

$$C(\mathbf{h}) = \tau^2 + \sigma^2 \exp(-\|\mathbf{h}\|_\alpha^\beta/\phi),\ 0 < \beta \le 1,\ 1 \le \alpha \le \infty \quad \text{or} \quad 0 < \beta \le \alpha \le 2$$

and for $\mathbf{h} \in \Re^N,\ N \ge 3$, the functions

$$C(\mathbf{h}) = \tau^2 + \sigma^2 \exp\big(-\|\mathbf{h}\|_\alpha^\beta/\phi\big),\ 0 < \beta \le \alpha \le 2,$$

are positive definite and hence valid covariance functions for $\tau^2, \sigma^2 > 0$. The usual isotropic exponential and Gaussian covariance functions based on the Euclidean distance measure can be obtained by setting (α, β) to (2,1) and (2,2) respectively. Fixing $\alpha = 2$ provides the current definition of the powered exponential covariance function (Stein 1999, p. 32–33). Setting $\alpha = \beta = 1$ demonstrates the city block metric with the exponential covariance function, whereas $\alpha = 1$ and $\beta = 2$ (city block metric with the Gaussian covariance function) is not admissible, as was demonstrated previously with the motivating example. For $\mathbf{h} \in \Re^2$, all norms are admissible provided $0 < \beta \le 1$.

Combining the results from Richards (1985) and Schoenberg's Theorem provides conditions for which $\|\mathbf{h}\|_\alpha^{\beta/2}$ is embeddable and thus can be used with existing isotropic covariance and variogram functions that are valid in all dimensions. This approach is applied to the Matern class of covariance functions isotropic with respect to Euclidean distance (Cressie, 1991), which is now shown for $\mathbf{h} \in \Re^2$, to include the functions

$$C(\mathbf{h}) = \tau^2 + \sigma^2\Big\{\Big(2^{\kappa-1}\Gamma(\kappa)\Big)^{-1}\Big(\|\mathbf{h}\|_\alpha^{\beta/2}/\phi\Big)^\kappa K_\kappa\Big(\|\mathbf{h}\|_\alpha^{\beta/2}/\phi\Big)\Big\},\ 0 < \beta \le 1,\ 1 \le \alpha \le \infty \quad \text{or} \quad 0 < \beta \le \alpha \le 2$$

and for $\mathbf{h} \in \Re^N,\ N \ge 3$, to include the functions

$$C(\mathbf{h}) = \tau^2 + \sigma^2\big\{\big(2^{\kappa-1}\Gamma(\kappa)\big)^{-1}\big(\|\mathbf{h}\|_\alpha^{\beta/2}/\phi\big)^\kappa K_\kappa\big(\|\mathbf{h}\|_\alpha^{\beta/2}/\phi\big)\big\},\ 0 < \beta \le \alpha \le 2,$$

for $\tau^2, \sigma^2 > 0$, where $K_\kappa(\cdot)$ represents the modified Bessel function of the third kind of order κ. Setting $\alpha = \beta = 2$ provides the class of Matern covariance functions isotropic with respect to Euclidean distance and for $\kappa = 0.5, \infty$ in this case

the Matern covariance function reduces to the exponential and Gaussian covariance function respectively. Again, for $\mathbf{h} \in \Re^2$, all norms are admissible provided $0 < \beta \leq 1$. However, unlike for the powered norm dependent exponential covariance function above, the exact form of the Matern covariance function is not retained due the exponent $\beta/2$ which equals 1 only when $\alpha = \beta = 2$.

Forms of other existing covariance functions can be used to demonstrate other classes of norm dependent isotropic covariance functions in a similar fashion. Assuming second-order stationarity, relation $\gamma(\mathbf{h}) = C(\mathbf{0}) - C(\mathbf{h})$ demonstrates corresponding classes of norm dependent isotropic (semi)variogram functions.

Assuming only intrinsic stationarity, the embedding approach can also be used to demonstrate classes of norm dependent conditionally negative definite functions. For example, consider the power variogram function isotropic with respect to Euclidean distance currently defined for $\mathbf{h} \in \Re^N,\ N \geq 1$, to be

$$2\gamma(\mathbf{h}) = \tau^2 + \phi\|\mathbf{h}\|^{\delta},\ 0 < \delta < 2,$$

for $\tau^2, \phi > 0$. Substituting the embeddable norms $\|\mathbf{h}\|_{\alpha}^{\beta/2}$ for the Euclidean norm $\|\mathbf{h}\|$ in above yields the following class of norm dependent isotropic conditionally negative definite functions. To ensure identifiability, the functions are parameterized with a single exponent parameter $\lambda = \beta\delta/2$. For $\mathbf{h} \in \Re^2$, the functions

$$2\gamma(\mathbf{h}) = \tau^2 + \phi\|\mathbf{h}\|_{\alpha}^{\lambda},\ \ 0 \leq \lambda \leq 1, \leq \alpha \leq \infty,$$
$$\text{or}$$
$$0 < \lambda < 2,\ \lambda \leq \alpha \leq 2,$$

and for $\mathbf{h} \in \Re^N,\ N \geq 3$, the functions

$$2\gamma(\mathbf{h}) = \tau^2 + \phi\|\mathbf{h}\|_{\alpha}^{\lambda},\ 0 < \lambda < 2,\ \lambda \leq \alpha \leq 2,$$

for $\tau^2, \phi > 0$, are conditionally negative definite. Setting $\alpha = 2$ provides the family of power variogram functions isotropic with respect to Euclidean distance. For $\mathbf{h} \in \Re^2$, all norms yield a conditionally negative definite function provided $0 \leq \lambda \leq 1$.

Note, for intrinsic stationarity care was taken not to refer to the class of norm dependent conditionally negative definite functions as valid variograms. As stated previously there is a growth condition variograms must satisfy (Matheron, 1973). This condition would need to be evaluated for the more general norm dependent class of conditionally negative definite functions before labeling these functions as valid variograms. Pragmatically speaking though, the greater mathematical flexibility achieved by assuming intrinsic stationarity over second-order stationarity is not often realized in applications.

SIMULATED APPLICATIONS

Two simulation applications are provided demonstrating different aspects concerning the use of a non-Euclidean distance measure to characterize isotropic spatial dependence in geostatistics. The first example considers norm dependent distance measures and investigates whether data can be used to choose the distance norm. The second application considers the winding stream scenario and compares kriged predictions based on Euclidean distance to those using the stream distance. All simulated data were Gaussian with mean zero (assumed unknown) and spatial covariance structures as described. The simulations are kept simple only to highlight these concepts, with more application specific details provided in future work. All computing was performed in R (R Development Core Team, 2005) with necessary modifications applied to functions from the geoR (Ribeiro and Diggle, 2001) contributed package. The simulated stream design was generated using ArcGIS Desktop (Environmental Systems Research Institute, 2004).

Norm Dependent Distances

Data were simulated on a 10×10 regular grid ($n = 100$) with unit spacing. The norm dependent exponential covariance function

$$C(\mathbf{h}) = \tau^2 + \sigma^2 \exp(-\|\mathbf{h}\|_\alpha/\phi) \quad 1 \leq \alpha \leq \infty$$

$\tau^2, \sigma^2, \phi > 0$, obtained by fixing the exponent parameter $\beta = 1$, was used to characterize spatial structure. Covariance parameters τ^2, σ^2, and ϕ were set at 0, 10, and 2 respectively. Four scenarios were considered based on setting the distance norm parameter $\alpha = 1, 2, 3$, and 4 with 1000 data sets simulated for each setting. For each data set, parameters were estimated via restricted maximum likelihood considering (a) the distance norm parameter α to be fixed at 2 representing isotropy with respect to Euclidean distance and (b) allowing the distance norm parameter α to vary $1 \leq \alpha \leq \infty$ representing isotropy with respect to a non-Euclidean norm dependent distance. Ratios of the minimized negative log restricted likelihoods based on using the Euclidean distance (a) to that from allowing the data to choose the distance norm (b) are used to compare the two approaches. Distributional summaries for the distance norm parameter α when it was estimated are also presented. Results summarized for each scenario are listed in Table 1.

For isotropy with respect to the non-Euclidean distance norm $\alpha = 1$ (city block metric), the approach based on allowing the data to estimate the distance norm resulted in minimized negative log restricted likelihoods about 4.2% smaller than the approach based on assuming isotropy with respect to Euclidean distance $\alpha = 2$ fixed. Distributional summaries for the estimated α were targeted to their

true value of $\alpha = 1$, noting this to be on the boundary of the parameter space. For isotropy with respect to Euclidean distance $\alpha = 2$, both approaches produced minimized negative log restricted likelihoods that were relatively equal most of the time. However, more variability is seen in the estimated α parameter. For isotropy with respect to the $\alpha = 3, 4$ norm distances there were surprisingly no apparent difference in the minimized negative log restricted likelihoods casting doubt as to whether data can help distinguish the distance norm as is suggested with results from $\alpha = 1, 2$. The estimated α norms for these scenarios displayed even more variability.

Note for some simulated data sets (0.9%, 5.2%, and 10.5% for those simulated under $\alpha = 2, 3, 4$ respectively) the distance norm parameter α was estimated very high causing calculations in the norm distance measure (4) to be beyond computer accuracy. For the two dimensional spatial design considered here this occurred when α was estimated higher than 300. The α parameter was therefore restricted to be $0 \leq \alpha \leq 300$ when minimizing log likelihoods. The results listed in Table 1 exclude these cases. Their inclusion, in terms of the ratio of the minimized negative log restricted likelihoods, produced results that were qualitatively indistinguishable from those presented. Excluding these cases as can be expected though have a more profound effect on summaries for the estimated α-norm parameter, providing more favorable support for the proposal of data driven distance norm estimation. Even so, these results overall for such a proposal are less than convincing.

It was thought that for this spatial design α values greater than 300 might effectively be interpreted as infinity and distance calculations using (4) for $\alpha >$ 300 be calculated using $\|\mathbf{h}\|_\infty$ (the supremum norm). Estimated large α norm parameters may be similar in spirit to the situations where the range and sill parameters are sometimes estimated (via a likelihood based approach) far outside our expected parameter space but end up producing a variogram shape consistent

Table 1. Distributional Summaries for the Ratios of the Minimized Negative Log Restricted Likelihoods (NLRL) Based on Using the Euclidean Distance to that from Allowing the Data to Choose the Distance Norm

	Ratio NLRL($\alpha = 2$) to NLRL($\hat{\alpha}$)			Estimated norm parameter $\hat{\alpha}$			
Distance norm	Mean	5th % tile	95th % tile	Median	Mean	5th % tile	95th % tile
$\alpha = 1$	1.043	1.013	1.080	1.000	1.042	1.000	1.248
$\alpha = 2$	1.002	1.000	1.009	1.908	2.259	1.191	3.744
$\alpha = 3$	1.005	1.000	1.017	2.977	4.993	1.662	10.912
$\alpha = 4$	1.010	1.000	1.026	4.004	7.809	2.132	21.477

Note. Summaries for the estimated α norm parameter also included the median since these distributions were skewed unlike the distributions of NLRL ratios. Results presented for each α norm setting $\alpha = 1, 2, 3, 4$ were summarized over the 1000 simulated data sets for each scenario.

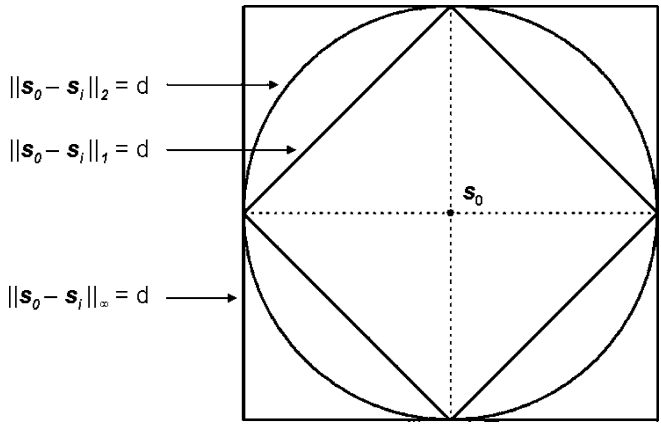

Figure 1. α-norm distance buffers, $\alpha = 1, 2, \infty$.

with the data. Addressing these and possibly other technical issues can be included in future more comprehensive investigations.

It is worth noting some numerical comparisons between distance norms to further explore issues related to their involvement in geostatistics. For example,

$$\|\mathbf{h}\|_{\alpha_1} \geq \|\mathbf{h}\|_{\alpha_2} \quad \text{for } \alpha_1 \geq \alpha_2.$$

A geometric interpretation of which is provided by letting $\mathbf{s}_0$ represent a point of origin and consider other locations $\mathbf{s}_i$ a fixed α-norm distance from $\mathbf{s}_0$, say $\|\mathbf{s}_0 - \mathbf{s}_i\|_\alpha = \mathrm{d}$. The diagram shown in Fig. 1 displays the shapes of the distance buffers around $\mathbf{s}_0$ such that for all locations $\mathbf{s}_i$, $\|\mathbf{s}_0 - \mathbf{s}_i\|_\alpha = \mathrm{d}$, for $\alpha = 1, 2, \infty$. For $\alpha = 2$ Euclidean norm, all points within a distance d of $\mathbf{s}_0$ fall within a circle of radius d (i.e. radial distance). In contrast, all points within an α-norm distance d of $\mathbf{s}_0$, $\alpha = 1, \infty$, correspond to diamond and square shaped buffers respectively. Shapes for distance buffers based on α norms not shown fit respectively within those displayed.

In terms of the traditional graphical approach towards characterizing spatial dependence Fig. 2 displays estimated variograms using the method of moments estimator (Cressie, 1991), adjusted to consider isotropy with respect to $\|\mathbf{h}\|_\alpha$ distances. Using a simulated data set from above for $\alpha = 2$, isotropy with respect to Euclidean distance, shown are estimated variograms based on $\alpha = 1, 2, 3, 4$. Immediate from Fig. 2 is the similarity in estimates, especially for the more important distances near the origin. This is an artifact not only of the sample design but that distance norms themselves not being very different for relatively small distances. Add to this the practice of distance binning, common for real data not sampled on a regular grid, that may further mask any differences when considering

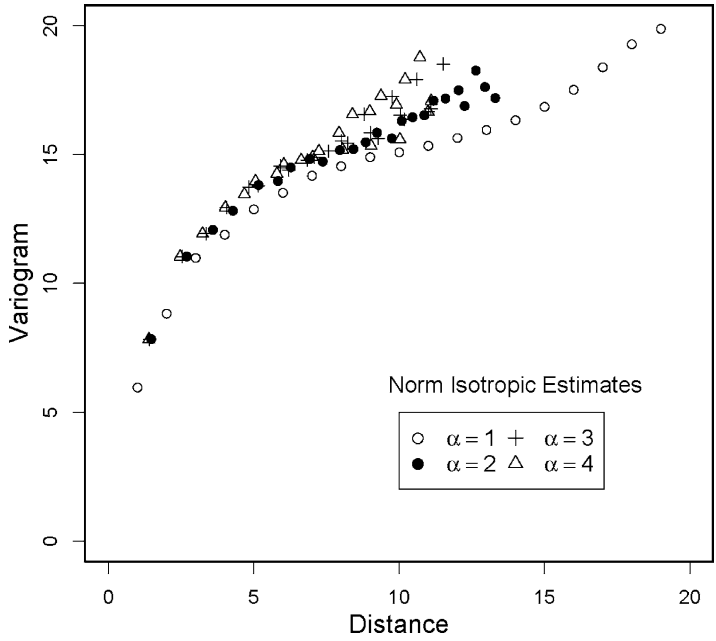

Figure 2. Omnidirectional variogram estimates for a data set simulated with isotropic spatial dependence based on the Euclidean distance measure, $\alpha = 2$ norm. Shown are the method of moments variogram estimates using various α norm distances, $\alpha = 1, 2, 3, 4$.

different norm dependent distance measures to characterize spatial dependence. Interpretations from this visual inspection only applies to the method of moments variogram estimator and similar graphical procedures used to characterize spatial dependence.

Stream Distance versus Euclidean Distance

To provide a demonstration of kriging with different distance measures a winding stream design was generated as follows. Shown highlighted in Fig. 3 (left) is the Potomac River running from Washington, DC through the Potomac River Branch tributary of the Chesapeake Bay. The solid line drawn through the center of this waterway is an example of the winding stream scenario that doesn't branch off and has no width, whether the width is negligible is not an assumption addressed here. The full length of the generated stream shown is 134 miles in terms of stream distance (traveling along the line), compared to the straight line Euclidean distance of 70 miles between the two dimensional geographic coordinates at each

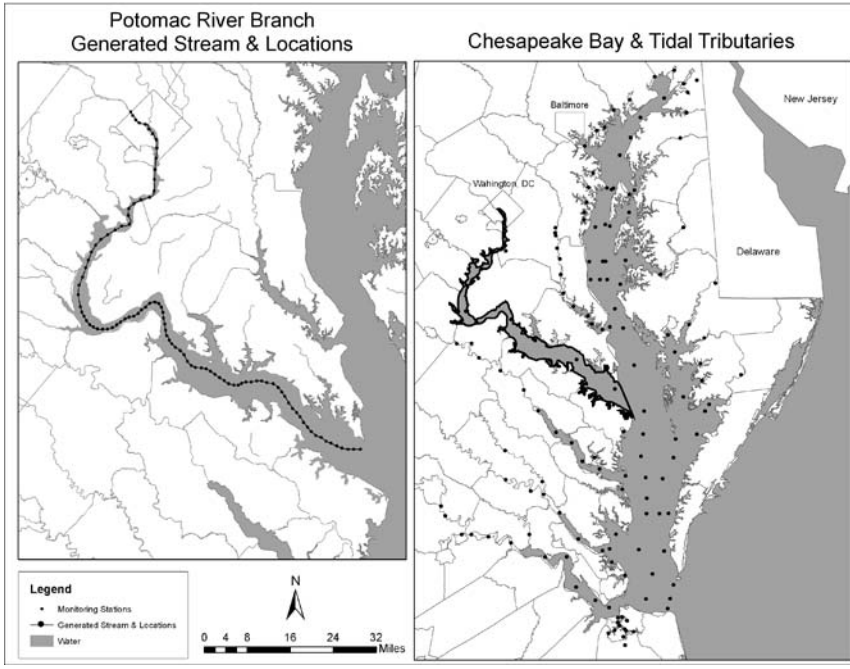

Figure 3. Shown on the right is the Chesapeake Bay region and its tidal tributaries with the fixed monitoring stations used in the collection of water quality parameters. Shown highlighted on the left is the Potomac River running from Washington, DC through the Potomac River Branch Tributary with the generate stream and accompanying 100 sample locations used for the simulated application.

end of the line. The Euclidean distance measure along such a stream can intersect land, a clear motivation for considering the stream distance. Placed along this generated stream are 100 locations approximately 1.3 miles apart (stream miles). This design was chosen because, as previously demonstrated, distances from such a stream configuration represent a special case of isometric embedding since these distances are actually equivalent to Euclidean distances in the transformed (stretched out) one dimensional space and hence valid for use in geostatistics. The technique provided in Mardia, Kent, and Bibby (1995, Theorem 14.2.1, p. 397) was used to generate these one dimensional coordinates which were used below for the analysis based on stream distances.

Simulation results are used to compare kriging based on stream distances and Euclidean distances. Spatial data were simulated at the 100 locations based on stream distance with an exponential variogram, nugget and sill fixed at 0 and 10 respectively and effective range parameter varied from 15, 30 and 90 stream miles, representing shorter to longer range spatial dependence. For each simulated data set a hold out sample of 25 locations were randomly selected. Based on

Table 2. Root Mean Squared Error in Predictions (RMSEP) for Kriging Based on Stream Distance Compared to Those Based on Euclidean Distance

	Ratio RMSEP Euclidean to RMSEP Stream		
Spatial dependence	Mean	5th % tile	95th % tile
Shorter	1.042	0.946	1.232
Medium	1.103	0.936	1.442
Longer	1.141	0.941	1.612

Note. Shown are distributional summaries, mean, 5t and 95t percentiles, for the ratio RMSEP for Euclidean distance to RMSEP for stream distance from the 1000 iterations for each spatial dependence range setting, Shorter (effective range 15 miles), Medium (effective range 45 miles), and Longer (effective range 90 miles).

the remaining 75 locations variogram parameters were estimated (via restricted maximum likelihood) and used to krige at the 25 hold out locations. Variogram estimation and kriging were performed using the stream distance and Euclidean distance (based on the two dimensional geographic coordinates). Using the hold out data at the 25 locations root mean squared error in predictions (RMSEP) were calculated for the two kriging approaches as a measure of prediction accuracy, RMSEP based on stream distance and RMSEP based on Euclidean distance. For each spatial dependence range this process iterated through 1000 simulated data sets. For each iteration the ratio of Euclidean RMSEP to stream RMSEP was generated.

Results summarized in Table 2 show on average kriging using the stream distance provides more accurate predictions to those based on Euclidean distance. For shorter range spatial dependence (15 miles) kriging based on stream distance produced predictions that were 4.2% more accurate than those based on Euclidean distance. This effect increased to 10.3% for medium range spatial dependence (45 miles) and 14.1% for longer range spatial dependence (90 miles). The pattern across spatial dependence ranges is intuitive, since the difference between stream and Euclidean distances in this design are less for locations close and greater for locations further away. Kriging with longer range spatial dependence is influenced more by data further away (compared to shorter range spatial dependence) and hence is impacted more by the appropriate distance metric.

The Chesapeake Bay region and its tidal tributaries shown in Fig. 3 (right) is the spatial setting of a large water-quality assessment that is being conducted by the Chesapeake Bay Program, which is the six state-federal partnership leading efforts to restore the Bay. Water quality parameters such as dissolved oxygen, turbidity, and chlorophyll are used to assess the ecological health of the Bay and

are routinely collected from a fixed network of monitoring stations, also shown in Fig. 3 (right). To fully assess water quality within the Bay and support decisions on attainment of established standards, water quality measures need to be interpolated to areas between those sampled at the fixed monitoring stations (EPA, 2003). For this endeavor geostatistics is an application motivating the use of a water distance measure more sophisticated than a negligible width winding stream network. Developing geostatistical methods that are valid with such a distance measure as well a resource for computing these types of distances are topics currently under investigation.

DISCUSSION

A simple example was used to demonstrate there are no guarantees that the existing pool of isotropic covariance and variogram functions remain valid when used with a distance measure other than Euclidean. It is therefore essential to establish the validity of these functions when an alternative measure of distance is proposed. By linking the concepts of isometric embedding, conditionally negative definiteness, and positive definiteness, an approach for demonstrating classes of norm dependent isotropic covariance and variogram functions was provided. An appealing proposition from this is that in practice data can be used to estimate the distance norm, as was demonstrated with the simulated application in the previous section. However, those results and the discussion following are not convincingly supportive of such an approach with further work necessary to address more technical issues. On the contrary, results from the simulated stream application demonstrated benefits in terms of kriging accuracy when considering a distance measure more appropriately suited for the spatial setting that was different from the usual straight line Euclidean distance.

Define the non-Euclidean distance problem in geostatistical related applications to include issues arising from the proposed use of a non-Euclidean distance (at least non-Euclidean in the dimension the process is observed) to characterize isotropic spatial dependence via a covariance or variogram function. As demonstrated here, existing isotropic functions are likely norm dependent, such as Euclidean distance or the extensions outlined in Section 4. Not considered here are the situations involving a distance measure d that is not necessarily a norm function, for example distances traveled through complex waterways or roads, such as in the Chesapeake Bay water quality assessment application. Establishing the validity of $C(d)$ or $2\gamma(d)$ as functions of isotropic spatial dependence, either for known covariance and variogram functions or for newly developed classes of such functions, may be mathematically challenging. Methods for dealing with such situations has not received much attention. Another point not to be overlooked is the calculation of such distances, which may prove nontrivial compared to the straight forward calculations involved in norm dependent distance measures.

One approach for using a general non-Euclidean distance measure d for geostatistical applications could be based on multidimensional scaling (MDS). Multidimensional scaling (Mardia, Kent, and Bibby, 1995) is a multivariate statistical technique concerned with the problem of constructing a set of points so that the Euclidean distance between these points matches (exact or most often approximate) a set of given distances that are likely not Euclidean. The concept of isometric embedding relates to the situation when such a configuration can be found for an exact match. For geostatistical applications a matrix of non-Euclidean inter-point distances (such as those traveled through complex waterways) would be approximated by the Euclidean distance between a set of points (often in a much higher dimension) generated by multidimensional scaling. The analysis would proceed using the approximate Euclidean distances hence avoiding issues of covariance/variogram function validity. In a sense transforming the application to the new Euclidean space determined by the multidimensional scaling. Sampson and Guttorp (1992) propose a similar approach to a different problem.

For dealing with non-Euclidean distance measures in geostatistics, such an MDS approach was originally proposed in Curriero (1996), more recently applied in Loland and Host (2003), Schabenberger and Gotway (2005), and could serve as an approach in the Chesapeake Bay Program's water quality assessments. A potential drawback of this approach is based on the fact that the multidimensional scaling Euclidean distance approximation does not consider spatial variation directly, that is it only considers approximating inter-point distances and ignores the outcome data. Further, it is sample design dependent, in the sense that adding and/or deleting a location (and hence a series of distances) can change the distance approximation elsewhere. Combining the sampling and prediction locations is a solution to the latter but this increases (often dramatically) the number of distances to be approximated and hence could adversely effect the MDS accuracy. In addition, there are two classes of MDS algorithms, those attempting to approximate the provided set of non-Euclidean distances (metric MDS) and those attempting to just preserve the rank ordering of distances (non-metric MDS). Which of these methods if any is preferred for the non-Euclidean distance problem in geostatistics is an open question.

Its worth mentioning a few valid criticism on using non-Euclidean distance measures to describe proximity relationships among spatial data. First, in the norm dependent case when the data are used to guide the distance norm, one to two extra parameters (α for the norm and β for its power) require estimation in addition to the usual range, sill, and nugget parameters. Issues of identifiability and reliable estimation which have not been addressed here certainly come into play. Although in regards to reliable estimation the same can be said for the two extra rotation and stretching parameters involved with geometric anisotropy. Alternatively, the distance norm parameter α and/or β can be set a priori to represent several possible choices and evaluated. A second issue is the fact that for geostatistical applications

characterizing spatial dependence is most crucial for smaller distances near the origin of the covariance or variogram function. It may be such that non-Euclidean inter-point distances are very close to their Euclidean counterparts at these smaller distances, a fact that is certainly true for distance norms, and became evident when combined with the strength of spatial dependence in the kriging results from the simulated stream application.

ACKNOWLEDGMENTS

This work was partially supported by The Johns Hopkins Bloomberg School of Public Health Faculty Innovation Grant 2006. The author would like to thank Tilmann Gneiting, Department of Statistics, University of Washington, for comments on a previous draft, Kathryn Kulbicki, GIS Database Specialist, Department of Environmental Health Sciences, The Johns Hopkins Bloomberg School of Public, for her efforts constructing the simulated stream network, and the Chesapeake Bay Program for their continued data and technical support.

REFERENCES

Banerjee, S., 2005, On geodetic distance computations in spatial modeling: Biometrics, v. 61, p. 617–625.

Christakos, G., 1984, On the problem of permissible covariance and variogram models: Water Resou. Res., v. 20, p. 251–265.

Cressie, N., 1991, Statistics for spatial data: Wiley, New York, 928 p.

Cressie, N., Gotway, C. A., and Grondona, M. O., 1990, Spatial prediction from networks: Chemometr. Intell. Lab. Systems, v. 7, p. 251-271.

Cressie, N., and Majure, J. J., 1997a, Non-point source pollution of surface waters over a watershed, *in* Barnett, V., and Turkman, K., eds., Statistics for the environment 3: pollution assessment and control, Wiley, New York, p. 210–224.

Cressie, N., and Majure, J. J., 1997b, Spatio-temporal statistical modeling of livestock waste in streams: J. Agric. Biol. Envir. Statist., v. 2, p. 24–47.

Curriero, F. C., 1996, The Use of non-Euclidean distances in geostatistics: PhD Dissertation, Department of Statistics, Kansas State University, 213 p.

Diggle, P. J., Tawn, J. A., and Moyeed, R. A., 1998, Model based geostatistics (with discussion): Appl. Statist., v. 47, p. 299–350.

Dominici, F., Samet, J. M., and Zeger, S. L., 2000, Combining evidence on air pollution and daily mortality from the 20 largest US cities: A hierarchical modelling strategy: J. Roy. Stat. Soc. Series A, v. 163, p. 263–302.

Environmental Protection Agency (EPA), 2003, Ambient water quality criteria for dissolved oxygen, water clarity, and chlorophyll a for the chesapeake Bay and its tidal tributaries: EPA 903-R-03-002.

Gneiting, T., 1999, Correlation functions for atmospheric data analysis: QJ Roy. Meteor. Soc., v. 125, p. 2449–2464.

Graybill, F. A., 1983, Matrices with applications in statistics, 2nd ed.: Wadsworth Belmont, California, 480 p.

Kern, J. C., and Higdon, D. M., 1999, A distance metric to account for edge effects in spatial data analysis, *in* Proceeding of the American Statistical Society, Biometrics Section, Alexandria, VA, p. 49–52.

Koldobskii, A. L., 1992, Schoenberg's problem on positive definite fucntions: St. Petersburg Math. J., v. 3, 563–570.

Krivoruchko K., and Gribov, A., 2004, Geostatistical interpolation and simulation in the presence of barriers, *in* geoENV IV Geostatistics for Environmental Applications, Proceedings, Barcelona, Spain, p. 331–342.

Little, L. S., Edwards, D., and Porter, E. E., 1997, Kriging in estuaries: As the crow flies, or as the fish swims?: J. Exp. Mar. Biol. Ecol., v. 213, p. 1–11.

Loland, A., and Host, G., 2003, Spatial covariance modelling in a complex coastal domain by multidimensional scaling: Environmetrics, v. 14, p. 307–321.

Mardia, K. V., Kent, J. T., and Bibby, J. M., 1995, Multivariate analysis: Academic Press, New York, 521 p.

Matheron, G., 1973, The intrinsic random functions and their applications: Adv. Appl. Probab., v. 5, p. 439–468.

R Development Core Team, 2005, R: A language and environment for statistical computing: R Foundation for Statistical Computing, Vienna, Austria, http://www.R-project.org.

Rathbun, S. L., 1998, Spatial modelling in irregularly shaped regions: kriging estuaries: Environmetrics, v. 9, p. 109–129.

Ribeiro, P. J., Jr., and Diggle, P. J., 2001, geoR: A package for geostatistical analysis, R-NEWS, v. 1, no. 2.

Richards, D. St. P., 1985, Positive definite symmetric functions on finite dimensional spaces II: Stat. Probabil. Lett., v. 3, p. 325–329.

Sampson, P. D., and Guttorp, P., 1992, Nonparametric estimation of nonstationary spatial covariance structure: J. Am. Stat. Assoc., v. 87, p. 108–119.

Schabenberger, O., and Gotway, C. A., 2005, Statistical methods for spatial data analysis: Chapmann and Hall/CRC Press, Florida, 512 p.

Schlather, M., 1999, Introduction to positive definite functions and to unconditional simulation of random fields: Technical Report ST-99-10, Lancaster University, UK.

Schoenberg, I. J., 1937, On certain metric spaces arising from euclidean spaces by a change of metric and their imbedding in Hilbert space: Ann. Math., v. 38, p. 787–793.

Stein, M. L., 1999, Interpolation of spatial data: some theory for kriging: Springer, New York, 247 p.

Ver Hoef, J. M. Peterson, E., and Theobald, D., 2006, Spatial statistical models that use flow and stream distance: Environmental and Ecological Statistics (in press).

Wells, J. H., and Williams, L. R., 1975, Embeddings and extensions in analysis: Springer-Verlag, New York, 108 p.

Young, G., and Householder, A. S., 1938, Discussion of a set of points in terms of their mutual distances: Psychometrika, v. 3, p. 19–22.

Zastavyni, V. P., 1993, Positive-definite functions that depend on a norm: Russian Acad. Sci. Dokl. Math., 46, p. 112–114.

Zastavyni, V. P., 2000, On positive definiteness of some functions: J. Multivariate Anal., v. 73, p. 55–81.