Phaedon C. Kyriakidis

A Geostatistical Framework for Area-to-Point Spatial Interpolation

The spatial prediction of point values from areal data of the same attribute is addressed within the general geostatistical framework of change of support; the term support *refers to the domain informed by each datum or unknown value. It is demonstrated that the proposed geostatistical framework can explicitly and consistently account for the support differences between the available areal data and the sought-after point predictions. In particular, it is proved that appropriate modeling of all area-to-area and area-to-point covariances required by the geostatistical framework yields coherent (mass-preserving or pycnophylactic) predictions. In other words, the areal average (or areal total) of point predictions within any arbitrary area informed by an areal-average (or areal-total) datum is equal to that particular datum. In addition, the proposed geostatistical framework offers the unique advantage of providing a measure of the reliability (standard error) of each point prediction. It is also demonstrated that several existing approaches for area-to-point interpolation can be viewed within this geostatistical framework. More precisely, it is shown that (i) the choropleth map case corresponds to the geostatistical solution under the assumption of spatial independence at the point support level; (ii) several forms of kernel smoothing can be regarded as alternative (albeit sometimes incoherent) implementations of the geostatistical approach; and (iii) Tobler's smooth pycnophylactic interpolation, on a quasi-infinite domain without non-negativity constraints, corresponds to the geostatistical solution when the semivariogram model adopted at the point support level is identified to the free-space Green's functions (linear in 1-D or logarithmic in 2-D) of Poisson's partial differential equation. In lieu of a formal case study, several 1-D examples are given to illustrate pertinent concepts.*

1. INTRODUCTION

Going from one spatial support (domain informed by each measurement or unknown value) to another is of critical importance to numerous scientific disciplines. Coarse spatial resolution predictions of general circulation models, for example, need to be downscaled to the watershed level (or even finer in the case of spatially distrib-

P. C. Kyriakidis is an assistant professor in the Department of Geography, University of California Santa Barbara (phaedon@geog.ucsb.edu).

uted models) for hydrologic impact assessment studies. Similarly, socioeconomic variables reported on census tracts need to be downscaled to smaller regions for detailed modeling. Scaling issues continue to be a critical and vibrant research topic in Geography; a recent review of such issues and some of their geostatistical solutions can be found in Atkinson and Tate (2000).

Area-to-point interpolation is a particular case of change of support, whereby areal data are used to predict point values; these points need not lie on a regular grid or comprise a surface. For a recent comprehensive review of existing approaches for change of support, see Gotway and Young (2002). Alternatively, area-to-point interpolation can be viewed as a special case of areal interpolation (Haining 2003), whereby both source data and target values pertain to the same spatial attribute and are defined respectively over areal units and points. Routine applications of area-to-point interpolation in geography (Lam 1983), however, tend to ignore several critical issues: (i) the explicit account of the different supports of the areal data and sought-after point predictions, for example, the areal data are often incorrectly collapsed into their respective polygon centroids; (ii) the coherence of predictions, for example, the areal average of point predictions within any area comprising an areal average datum, should be equal to that datum (if the latter is assumed error free); and (iii) the uncertainty in the resulting point predictions.

The geostatistical framework for area-to-point prediction presented in this paper is a special case of the original development of Kriging, which was formulated as the spatial prediction of an areal value using available areal data of the same or different variable (Matheron 1971; Journel and Huijbregts 1978; Gotway and Young 2002). Most geostatistical textbooks and applications regarding change of support, however, address only the problem of point-to-area interpolation via the use of block Kriging (Isaaks and Srivastava 1989; Cressie 1993; Wackernagel 1995; Goovaerts 1997). This is partially a consequence of the mining practice, where most of the original geostatistical developments were made. In such applications, the mineral content of blocks of ore is predicted from point support core data, that is, mining applications call for point-to-area interpolation and rarely (if ever) for the reverse procedure of area-to-point interpolation.

In this work, the applicability of the general geostatistical framework for addressing the area-to-point spatial interpolation problem is demonstrated. Only the case of a single variable is treated here; the use of auxiliary predictor variables is not addressed (although it is possible). Several of the little known, but extremely important, characteristics of this framework are show-cased, namely that the resulting point predictions are coherent, and several existing methods for area-to-point spatial interpolation can be formulated in terms of the geostatistical framework. Such methods include choropleth mapping, several forms of kernel smoothing (Brillinger 1990; Brillinger 1994; Bracken and Martin 1989; Martin 1996), and smooth pycnophylactic interpolation (Tobler, 1979) on a quasi-infinite domain without non-negativity constraints. In addition, the geostatistical framework has the unique advantage of yielding a measure of the reliability (uncertainty) associated with each point prediction.

Kelsall and Wakefield (2002) also adopted area-to-point Kriging in a lognormal setting for estimating the underlying (latent) relative risk surface of disease incidences. In that work, however, the coherence of the resulting area-to-point predictions is not demonstrated, and no connections are established with existing methods of area-to-point spatial interpolation. In addition, the resulting predictions are not guaranteed to be coherent due to precisely the nonlinearity of the logarithmic transformation involved in lognormal Kriging. Another approach that explicitly honors the coherence (or mass-balance) constraint is the multiresolution, autoregressive, tree-structured model of Huang, Cressie, and Garbosek (2002). As noted in Gotway and Young (2002), however, the basic algorithm used for prediction in multiscale tree models implicitly dictates the covariance at any particular scale, as well as across scales. In

addition, because two neighboring points may lie in different supports, the correlation function at the point level may show unrealistic "blocky" artifacts.

Viewed within the comprehensive review of change of support issues provided by Gotway and Young (2002), this paper fills in an important gap: the identification of formal links between choropleth mapping, several forms of kernel smoothing, and Tobler's Laplacian smooth pycnophylactic interpolation. In particular, this work clarifies two subtle (but extremely important for geographical research) points made in the review of Gotway and Young (2002). The first one pertains to the assumption underlying choropleth mapping, which entails that the target variable is evenly distributed within any source unit (Gotway and Young 2002, p. 643). In Section 5.1, it is shown that this assumption is a consequence of another (and more important) prior assumption: that of lack of spatial correlation at the point support level. In a geographical context, it is much more difficult to accept the latter assumption, than the assumption of constant target (point) values within a source (areal) unit. The second issue pertains to the partially incorrect statement that both Brillinger's kernel smoothing and Tobler's pycnophylactic interpolation assume that the areal data are spatially uncorrelated (Gotway and Young 2002, p. 643). In Sections 5.2 and 5.3, it is shown that only Brillinger's method, not Tobler's, implicitly invokes such an assumption.

In the next section, the geostatistical framework for area-to-point interpolation is presented, and in Section 3, it is demonstrated that the resulting point predictions are coherent. In Section 4, the smoothness of the area-to-point predictions is linked to the point covariance model adopted for Kriging. In Section 5, it is shown that various existing methods for area-to-point interpolation, namely choropleth mapping, kernel smoothing, and pycnophylactic interpolation with Laplacian smoothing (on a quasi-infinite domain without non-negativity constraints) yield the same point predictions with those obtained using the geostatistical framework if one assumes particular types of point covariance models. In Section 6, the problem of inferring a point covariance model from the available areal data is briefly addressed, and in Section 7 some conclusions are drawn.

2 GEOSTATISTICAL PREDICTION OF POINT VALUES FROM AREAL DATA

Let $z(\mathbf{s})$ denote the true (unknown) point support value at a location with coordinate vector $\mathbf{s}$ within a study domain D. In the geostatistical framework, the set of all point support values $\{z(\mathbf{s}), \mathbf{s} \in D\}$ is viewed as a joint realization (outcome) of a collection of spatially correlated random variables (RVs) $\{Z(\mathbf{s}), \mathbf{s} \in D\}$, or in short a realization of a random function (RF), e.g., Cressie (1993).

In this work, I consider the case of an intrinsically stationary RF model, with a constant but unknown mean: $E\{Z(\mathbf{s})\} = m_Z, \forall \mathbf{s} \in D$. The semivariogram $\gamma_Z(\mathbf{h}) = \frac{1}{2}\text{Var}\{[Z(\mathbf{s}+\mathbf{h}) - Z(\mathbf{s})]\}$ and covariance $C_Z(\mathbf{h}) = E\{[Z(\mathbf{s}+\mathbf{h}) - m_Z][Z(\mathbf{s}) - m_Z]\}$ between any two RVs $Z(\mathbf{s})$ and $Z(\mathbf{s}+\mathbf{h})$ at two locations $\mathbf{s}$ and $\mathbf{s}+\mathbf{h}$ separated by a vector $\mathbf{h}$ depend only on the modulus and possibly orientation of that vector. The semivariogram $\gamma_Z(\mathbf{h})$ is a richer tool for quatifying spatial correlation than the covariance $C_Z(\mathbf{h})$, because it is defined even in situations where the covariance $C_Z(\mathbf{h})$ is not, that is, even if $\text{Var}\{Z(\mathbf{s})\} \to \infty$. For a finite variance $\text{Var}\{Z(\mathbf{s})\} = C_Z(\mathbf{0})$, the covariance and the semivariogram are linked by $C_Z(\mathbf{h}) = C_Z(\mathbf{0}) - \gamma_Z(\mathbf{h})$.

Let $v_k = v\mathbf{s}_k$ denote the support of the k-th areal datum with centroid at location $\mathbf{s}_k$ and otherwise arbitrary shape, size, and orientation. It should be stressed that support centroids are used here solely for indexing purposes; an areal datum is never collapsed into its support centroid. The measure (length in 1-D, area in 2-D, volume in 3-D) of that support is denoted as $|v_k|$. The k-th observed areal datum $\bar{z}(v_k)$ is defined as the average of all point values within support v_k:

$$\bar{z}(v_k) = \frac{1}{|v_k|} \int_{\mathbf{s} \in v_k} z(\mathbf{s}) \, d\mathbf{s} \simeq \frac{1}{P_k} \sum_{i=1}^{P_k} z(\mathbf{s}_i), \, \mathbf{s}_i \in v_k \tag{1}$$

where, in the discrete case, P_k denotes the number of points within support v_k. The above definition constitutes a regularization (averaging) of point z-values by the support v_k.

In the geostatistical framework, the K observed areal data $\{\bar{z}(v_k), k = 1, \dots, K\}$ are also viewed as a joint realization of K RVs $\{\bar{Z}(v_k), k = 1, \dots, K\}$, since any areal datum is defined as the arithmetic average (an equally weighted linear combination) of point support values, which themselves are viewed as outcomes of point RVs. The mean and covariance of the areal data can be readily derived from those of the point support values, provided the latter are known (Gotway and Young 2002; Matheron 1971; Journel and Huijbregts 1978; Cressie 1993; Chilès and Delfiner 1999). In what follows, I assume knowledge of the stationary point covariance $C_Z(\mathbf{h})$, or equivalently of the stationary point semi-variogram $\gamma_Z(\mathbf{h})$, in order to consistently derive all area-to-area and area-to-point covariance terms required by area-to-point Kriging (see hereafter). This latter consistent covariance modeling is precisely what guarantees the coherence of the resulting point predictions. In Section 6, I describe possible approaches for estimating this point covariance model $C_Z(\mathbf{h})$ from areal data, which itself is a nontrivial problem. In Section 4, I illustrate how the choice of a particular point support co-variance model $C_Z(\mathbf{h})$ affects the spatial distribution (or smoothness) of the resulting area-to-point Kriging predictions.

The objective of area-to-point spatial interpolation is to predict any unknown point value $z(\mathbf{s})$ using the K areal data $\{\bar{z}(v_k), k = 1, \dots, K\}$. In this paper, I consider the case of disjointed areal supports, i.e., $v_k \cap v_l = \varnothing$. The prediction locations are arbitrary, that is, they need not comprise a regular grid and they can lie inside or outside any areal support v_k (see Figure 1).

The predicted point value $z^*(\mathbf{s})$ is expressed as a weighted linear combination of the available K areal data comprising the $(K \times 1)$ data vector $\bar{\mathbf{z}} = [\bar{z}(v_k), k = 1, \dots, K]'$:

$$z^*(\mathbf{s}) = \sum_{k=1}^{K} \lambda_\mathbf{s}(v_k)\bar{z}(v_k) = \boldsymbol{\lambda}'_\mathbf{s}\bar{\mathbf{z}} \tag{2}$$
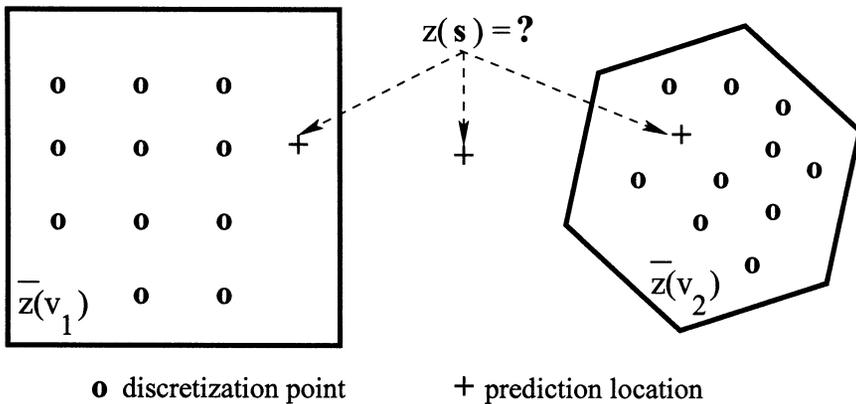

FIG. 1: Configuration example for an area-to-point spatial interpolation problem. Prediction locations are marked with crosses and discretization locations with open circles.

where $\boldsymbol{\lambda_s} = [\lambda_s(v_k), k = 1,\ldots,K]'$ denotes a location-specific $(K \times 1)$ vector of weights, whose k-th entry $\lambda_s(v_k)$ is the weight assigned to the k-th areal datum $\bar{z}(v_k)$ for prediction at location $\mathbf{s}$. Since the stationary point mean m_Z is unknown, one has to introduce a constraint on the weights to ensure unbiasedness (as in the point-to-point prediction case):

$$\sum_{k=1}^{K} \lambda_s(v_k) = 1 \quad \text{or} \quad \mathbf{1}_K' \boldsymbol{\lambda_s} = 1, \ \forall \mathbf{s} \in D \tag{3}$$

where $\mathbf{1}_K = [1,\ldots,1]'$ denotes a $(K \times 1)$ vector of unit entries.

The vector $\boldsymbol{\lambda_s}$ of K weights that yield the minimum prediction error variance among all constrained weighted linear combinations of areal data is obtained per solution of the system of normal equations, also termed area-to-point Ordinary Kriging (OK) system:

$$\begin{cases} \sum_{l=1}^{K} \lambda_s(v_l)\overline{C}_Z(v_k,v_l) - \chi_s = \overline{C}_Z(\mathbf{s},v_k), k = 1,\ldots,K \\ \sum_{l=1}^{K} \lambda_s(v_l) = 1 \end{cases} \tag{4}$$

where χ_s denotes a location-specific Lagrange multiplier that accounts for the unit sum constraint on the weights.

Term $\overline{C}_Z(v_k, v_l)$ of equation (4) denotes the (regularized) covariance between any two areal RVs $\overline{Z}(v_k)$ and $\overline{Z}(v_l)$:

$$\overline{C}_Z(v_k,v_l) = Cov\{\overline{Z}(v_k),\overline{Z}(v_l)\} = \frac{1}{|v_k|}\frac{1}{|v_l|} \int\limits_{\mathbf{s} \in v_k} \int\limits_{\mathbf{s}' \in v_l} Cov\{Z(\mathbf{s}),Z(\mathbf{s}')\}\, d\mathbf{s}'d\mathbf{s}$$

$$\simeq \frac{1}{P_k}\frac{1}{P_l} \sum_{i=1}^{P_k} \sum_{j=1}^{P_l} C_Z(\mathbf{s}_i - \mathbf{s}_j), \mathbf{s}_i \in v_k, \mathbf{s}_j \in v_l \tag{5}$$

where $C_Z(\mathbf{s}_i - \mathbf{s}_j)$ is the point covariance between any two discretization locations $\mathbf{s}_i \in v_k$ and $\mathbf{s}_j \in v_l$, and P_k, P_l denote the respective number of points discretizing the two supports v_k and v_l (see Figure 2). In words, the covariance $\overline{C}_Z(v_k, v_l)$ between any two areal supports v_k and v_l is the average of point covariance values $C_Z(\mathbf{s} - \mathbf{s}')$ corresponding to vectors $\mathbf{s} - \mathbf{s}'$ formed by all possible pairs of points $\mathbf{s} \in v_k$ and $\mathbf{s}' \in v_l$. Note that in this work, when the prime superscript accompanies a coordinate vector, e.g., $\mathbf{s}'$, it denotes another coordinate vector, not transposition. If the areal data were incorrectly collapsed into their respective centroids, then term $\overline{C}_Z(v_k, v_l)$ of equation (4) would be replaced by $C_Z(\mathbf{s}_k - \mathbf{s}_l)$, as in the point-to-point OK case.

Similarly, term $\overline{C}_Z(\mathbf{s},v_k)$ of equation (4) denotes the (regularized) covariance between any point RV $Z(\mathbf{s})$ and any areal RV $\overline{Z}(v_k)$:

$$\overline{C}_Z(\mathbf{s},v_k) = Cov\{Z(\mathbf{s}),\overline{Z}(v_k)\} = \frac{1}{|v_k|} \int\limits_{\mathbf{s}' \in v_k} Cov\{Z(\mathbf{s}),Z(\mathbf{s}')\}\, d\mathbf{s}'$$

$$\simeq \frac{1}{P_k} \sum_{i=1}^{P_k} C_Z(\mathbf{s} - \mathbf{s}_i), \mathbf{s}_i \in v_k \tag{6}$$
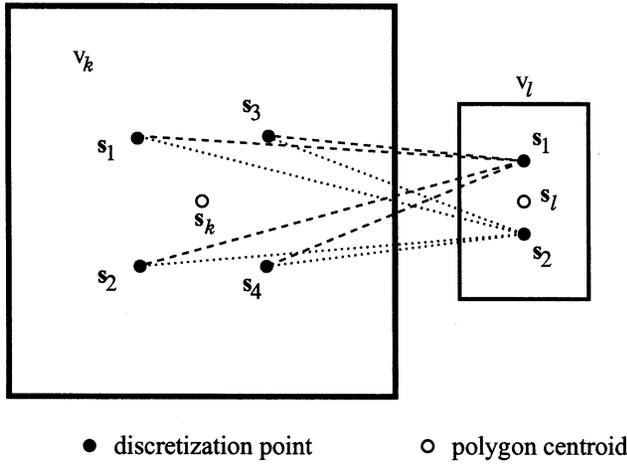
FIG. 2: Example of an area-to-area covariance calculation, based on four discretization points in support v_k, and two discretization points in support v_l; dashed lines represent vectors $\mathbf{s}_i - \mathbf{s}_j$, for which point covariance values $C_Z(\mathbf{s}_i - \mathbf{s}_j)$ are evaluated.

where $C_Z(\mathbf{s} - \mathbf{s}_i)$ is the point covariance between the prediction location $\mathbf{s}$ and any discretization location $\mathbf{s}_i \in v_k$, see Figure 3. In words, the covariance $\overline{C}_Z(\mathbf{s}, v_k)$ between any support v_k and any point $\mathbf{s}$ is the average of point covariance values $C_Z(\mathbf{s} - \mathbf{s}')$ corresponding to vectors $\mathbf{s} - \mathbf{s}'$ formed by all possible pairs of points $\mathbf{s}$ and $\mathbf{s}' \in v_k$. If again the areal data were incorrectly collapsed into their respective centroids, then term $\overline{C}_Z(\mathbf{s}, v_k)$ of equation (4) would be replaced by $C_Z(\mathbf{s} - \mathbf{s}_k)$, as in the point-to-point OK case.

The area-to-point OK system of equation (4) can be written more compactly in matrix form as:

$$\begin{bmatrix} \overline{\mathbf{C}} & \mathbf{1}_K \\ \mathbf{1}'_K & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{\lambda}_\mathbf{s} \\ -\chi_\mathbf{s} \end{bmatrix} = \begin{bmatrix} \overline{\mathbf{c}}_\mathbf{s} \\ 1 \end{bmatrix} \tag{7}$$

where $\overline{\mathbf{C}} = [\overline{C}_Z(v_k, v_l), k = 1,\ldots,K, l = 1,\ldots,K]$ denotes the $(K \times K)$ matrix of area-to-area covariances, and $\overline{\mathbf{c}}_\mathbf{s} = [\overline{C}_Z(\mathbf{s}, v_k), k = 1,\ldots,K]'$ denotes the $(K \times 1)$ location-specific vector of area-to-point covariances.

If and only if: (i) the matrix $\overline{\mathbf{C}}$ is positive definite, a requirement that is ensured if all entries $\overline{C}_Z(v_k, v_l)$ of that matrix are evaluated via equation (5) using a positive definite point covariance function $C_Z(\mathbf{h})$, and (ii) no datum support coincides with any other, that is, if $v_k \neq v_l$, $\forall k, l$, then the above system has a unique solution:

$$\begin{bmatrix} \boldsymbol{\lambda}_\mathbf{s} \\ -\chi_\mathbf{s} \end{bmatrix} = \begin{bmatrix} \overline{\mathbf{C}} & \mathbf{1}_K \\ \mathbf{1}'_K & 0 \end{bmatrix}^{-1} \begin{bmatrix} \overline{\mathbf{c}}_\mathbf{s} \\ 1 \end{bmatrix} \tag{8}$$

Note that when prediction is performed at P locations $\{\mathbf{s}_p, p = 1,\ldots,P\}$, a different system of the form of equation (4) must be solved at each prediction location $\mathbf{s}_p$, because the corresponding area-to-point covariance vector $\overline{\mathbf{c}}_p = \overline{\mathbf{c}}_{\mathbf{s}_p}$ changes from one location $\mathbf{s}_p$ to another.
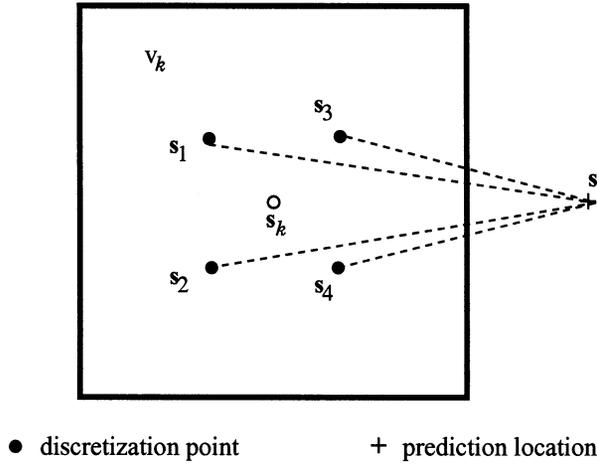
FIG. 3: Example of an area-to-point covariance calculation, based on four discretization points in support v_k; dashed lines represent vectors $\mathbf{s} - \mathbf{s}_i$, for which point covariance values $C_Z(\mathbf{s} - \mathbf{s}_i)$ are evaluated.

The resulting OK minimum prediction error variance $\sigma_E^2(\mathbf{s})$ at location $\mathbf{s}$ is:

$$\sigma_E^2(\mathbf{s}) = C_Z(\mathbf{0}) - [\boldsymbol{\lambda}_\mathbf{s}' - \chi_\mathbf{s}]\begin{bmatrix}\overline{\mathbf{c}}_\mathbf{s} \\ 1\end{bmatrix} = C_Z(\mathbf{0}) - \sum_{k=1}^{K} \lambda_\mathbf{s}(v_k)\overline{C}_Z(\mathbf{s}, v_k) + \chi_\mathbf{s} \qquad (9)$$

It should be noted that when area-to-point Kriging is performed at a support centroid, for example, when $\mathbf{s} = \mathbf{s}_k$, the resulting prediction error variance $\sigma_E^2(\mathbf{s}_k)$ of equation (9) is not the same as that obtained by point-to-point Kriging, because precisely the areal datum $\bar{z}(v_k)$ is not collapsed into its centroid. For the same reason, the area-to-point Kriging prediction $z^*(\mathbf{s}_k)$ of equation (2) at a support centroid $\mathbf{s}_k$ is not the same as that obtained by point-to-point Kriging.

Characteristics of Area-to-Point Kriging

- Area-to-point Kriging is a special case of the more general development of Kriging, which was originally formulated as the spatial prediction of an areal value $\bar{z}(v_\mathbf{s})$ using available areal data of the same or different variable (Matheron 1971; Journel and Huijbregts 1978); see also Gotway and Young (2002). In this case, the left-hand-side matrix of the Kriging system of equation (7) would be still populated with area-to-area covariance values $\overline{C}_Z(v_k, v_l)$, since the available data are defined over areal units. The right-hand-side vector of that system, however, would contain area-to-area covariance values $\overline{C}_Z(v_\mathbf{s}, v_k)$ for any pair of prediction and data supports. In the case of area-to-point Kriging, one needs to replace the elements of the right-hand-side vector of that system by area-to-point covariance values $\overline{C}_Z(\mathbf{s}, v_k)$, since now the unknown represents a point value and the known data areal measurements. All the above modifications, as well as equations (5) and (6), are direct consequences of the bilinearity of the covariance operator (Anderson 1958). On the computational side, accurate approximations of $\overline{C}_Z(v_k, v_l)$ and $\overline{C}_Z(\mathbf{s}, v_k)$ by their discrete counterparts can be obtained using numerical integration rules, such as Gaussian quadrature (Journel and Huijbregts 1978; Carr and Palmer 1993).

- In the case of unequal data supports $v_k \neq v_l$, the area-to-area (data-to-data) covariance matrix $\bar{\mathbf{C}}$ is non-stationary, for example, its diagonal elements are different: $\bar{\mathbf{C}}_{kk} = Var\{\bar{Z}(v_k)\} \neq Var\{\bar{Z}(v_l)\} = \bar{\mathbf{C}}_{ll}$. Another example of nonstationarity includes the case of two pairs of supports $\{v_k, v_l\}$ and $\{v_{k'}, v_{l'}\}$, for which centroid-to-centroid distances $|\mathbf{s}_k - \mathbf{s}_l|$ and $|\mathbf{s}_{k'} - \mathbf{s}_{l'}|$ are equal. Equation (6) entails that $\bar{C}_Z(v_k, v_l) \neq \bar{C}_Z(v_{k'}, v_{l'})$, unless all four supports have the same size, shape, and orientation. Similarly, in the case of unequal data supports, the area-to-point covariance is nonstationary. Consider, for example, the case of two pairs of points and supports $\{\mathbf{s}, v_k\}$ and $\{\mathbf{s}', v_l\}$, for which the point-to-centroid distances $|\mathbf{s} - \mathbf{s}_k|$ and $|\mathbf{s}' - \mathbf{s}_l|$ are equal. Equation (6) entails that $\bar{C}_Z(\mathbf{s}, v_k) \neq \bar{C}_Z(\mathbf{s}', v_l)$, unless supports v_k and v_l have the same size, shape, and orientation.
- The expressions for area-to-point OK prediction, OK system, and OK prediction error variance, equations (2) through (9), are completely general. They can accommodate different point covariance models $C_Z(\mathbf{h})$, for example, exponential or linear decay, including or not a white noise component and nested structures each possibly with its own anisotropy; and arbitrary and possibly partially overlapping areal data supports v_k and v_l. In addition, the area-to-point OK weights, and the resulting prediction error variance account for: (i) the statistical distance between any areal datum $\bar{z}(v_k)$ and the prediction location $\mathbf{s}$, via $\bar{C}_Z(\mathbf{s}, v_k)$; (ii) the relative redundancy (clustering) of any two areal data pairs $\bar{z}(v_k)$ and $\bar{z}(v_l)$, which accounts for the geometry of their configuration (including their size, shape, and orientation), via $\bar{C}_Z(v_k, v_l)$; and (iii) the type of spatial correlation of the point support values, that is, the functional form of the point covariance model $C_Z(\mathbf{h})$. The Lagrange parameter $\chi_\mathbf{s}$ represents a penalty on the prediction error variance due to lack of knowledge of the true point support mean m_Z.
- In analogy to classical point-to-point Kriging, the area-to-point OK predictor $Z^*(\mathbf{s})$ is the Best (in the least squares sense) Linear Unbiased Predictor (BLUP). The area-to-point OK variance $\sigma_E^2(\mathbf{s})$ is homoscedastic, that is, it does not depend on the actual areal data values, but only on the relative geometry (configuration) of their supports via the area-to-point covariances $\bar{C}_Z(\mathbf{s}, v_k)$ and the weights $\lambda_\mathbf{s}(v_k)$. In the multivariate Gaussian case, the area-to-point OK prediction and variance coincide with the conditional mean $E\{Z(\mathbf{s})|\bar{\mathbf{z}}\}$ and variance $Var\{Z(\mathbf{s})|\bar{\mathbf{z}}\}$ of the point RV $Z(\mathbf{s})$ given the K areal data.
- Strictly speaking, equations (2) through (9) constitute a co-Kriging procedure, whereby areal $\bar{z}$-data are used to predict point z-values. Since the two variables are functionally related per equation (1), both the auto-covariance of the areal data, that is, the area-to-area covariance $\bar{C}_Z(v_k, v_l)$, as well as the cross-covariance between the areal data and the point values—the area-to-point covariance $\bar{C}_Z(\mathbf{s}, v_k)$—are completely specified in terms of the point covariance model $C_Z(\mathbf{h})$; see equations (5) and (6).

In the general nonstationary case, whereby the z-mean is a function of monomials of spatial coordinates, or more generally a function of auxiliary variables, one needs to impose additional constraints on the weights (Matheron 1971; Journel and Huijbregts 1978; Cressie 1993; Gotway and Young 2002). In this paper I only focus on the intrinsic stationary case whereby second-order stationarity applies to the increments of the z-process. All the derivations presented in this paper generalize in a straightforward manner to the nonstationary case.

3. COHERENCE OF AREA-TO-POINT KRIGING PREDICTIONS

Coherence is a term used in geostatistics to characterize interpolation procedures that satisfy constraints expressed as linear combinations of data; see, for example,

Chilès and Delfiner (1999). Along the same lines, I adopt the term *coherence* in this paper to describe the fact that the areal-average of the resulting area-to-point OK predictions within a datum support v_k is equal to the corresponding areal datum $\bar{z}(v_k)$. In what follows, I prove the coherence property of the area-to-point OK predictions using the dual form of Ordinary Kriging, in short DOK (Matheron 1971; Davis and Grivet 1984; Cressie 1993; Chilès and Delfiner 1999). The coherence proof given in this paper builds upon, and generalizes, the results of Journel (1999).

Using equation (8), the OK prediction of equation (2) can also be written as:

$$z^*(\mathbf{s}) = \boldsymbol{\lambda}_\mathbf{s}' \bar{\mathbf{z}} = \bar{\mathbf{z}}' \boldsymbol{\lambda}_\mathbf{s} = [\bar{\mathbf{z}}' \; 0] \begin{bmatrix} \boldsymbol{\lambda}_\mathbf{s} \\ -\chi_\mathbf{s} \end{bmatrix} = [\bar{\mathbf{z}}' \; 0] \begin{bmatrix} \overline{\mathbf{C}} & \mathbf{1}_K \\ \mathbf{1}_K' & 0 \end{bmatrix}^{-1} \begin{bmatrix} \bar{\mathbf{c}}_\mathbf{s} \\ 1 \end{bmatrix}$$

from which the DOK prediction is derived as:

$$z^*(\mathbf{s}) = [\boldsymbol{\omega}' \; \psi] \begin{bmatrix} \bar{\mathbf{c}}_\mathbf{s} \\ 1 \end{bmatrix} = \sum_{k=1}^{K} \omega(v_k) \overline{C}_Z(\mathbf{s}, v_k) + \psi \tag{10}$$

where $\boldsymbol{\omega} = [\omega(v_k), \; k = 1,\dots,K]'$ is a $(K \times 1)$ vector of location-independent DOK weights, whose k-th entry $\omega(v_k)$ is the DOK weight assigned to the (area-to-point) co-variance $\overline{C}_Z(\mathbf{s}, v_k)$ between the k-th support v_k and the prediction location $\mathbf{s}$; the term ψ denotes a location-independent Lagrange multiplier.

The DOK system of equations simply emerges as the system that defines the vector of DOK weights $\boldsymbol{\omega}$ and the Lagrange multiplier ψ:

$$\begin{bmatrix} \overline{\mathbf{C}} & \mathbf{1}_K \\ \mathbf{1}_K' & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{\omega} \\ \psi \end{bmatrix} = \begin{bmatrix} \bar{\mathbf{z}} \\ 0 \end{bmatrix} \quad \text{or} \quad \begin{cases} \sum_{l=1}^{K} \omega(v_l) \overline{C}_Z(v_k, v_l) + \psi = \bar{z}(v_k), \; \; k = 1,\dots,K \\ \sum_{l=1}^{K} \omega(v_l) = 0 \end{cases} \tag{11}$$

whose solution yields:

$$\begin{bmatrix} \boldsymbol{\omega} \\ \psi \end{bmatrix} = \begin{bmatrix} \overline{\mathbf{C}} & \mathbf{1}_K \\ \mathbf{1}_K' & 0 \end{bmatrix}^{-1} \begin{bmatrix} \bar{\mathbf{z}} \\ 0 \end{bmatrix} \tag{12}$$

It should be stressed that the DOK system of equation (11) is independent of the prediction location $\mathbf{s}$, that is, the DOK system needs to be solved once for prediction at all points within the entire study region. This entails that the DOK vector of weights $\boldsymbol{\omega}$ is also independent of the prediction location $\mathbf{s}$, hence the use of notation $\boldsymbol{\omega}$ instead of $\boldsymbol{\omega}_\mathbf{s}$, and $\omega(v_k)$ instead of $\omega_\mathbf{s}(v_k)$. The DOK prediction $z^*(\mathbf{s})$ of equation (10) is then computed via a simple functional evaluation of the entries $\overline{C}_Z(\mathbf{s}, v_k)$ of the corresponding vector $\bar{\mathbf{c}}_\mathbf{s}$. When all the K areal data are used for interpolation, or, when performing Kriging with a global neighborhood, computing the set of all DOK predictions at all prediction locations is extremely fast. Essentially, the dual form of Kriging is implicitly adopted by most software for geostatistical interpolation using a global neighborhood, since inversion of the left-hand-side matrix of equation (7) is typically performed once for all prediction locations.

The interpretation of the DOK equations is left for the next section, where the smoothness of the resulting DOK predictions is addressed. It suffices to say here that one can treat the DOK formalism as an alternative but strictly equivalent way of expressing the OK prediction of equation (2). Most importantly, the dual form of Kriging allows a straightforward demonstration of the coherence of DOK predictions.

Consider now the areal average of the point DOK predictions $\frac{1}{|v_k|}\int_{\mathbf{s} \in v_k} z^*(\mathbf{s})\, d\mathbf{s}$ within the support v_k of the k-th areal datum $\bar{z}(v_k)$, and rewrite the DOK point prediction of equation (10) at any location $\mathbf{s}$ as $z^*(\mathbf{s}) = \sum_{l=1}^K \omega(v_l)\overline{C}_Z(\mathbf{s},v_l) + \psi$. It readily follows that:

$$\frac{1}{|v_k|} \int_{\mathbf{s} \in v_k} z^*(\mathbf{s})\, d\mathbf{s} = \frac{1}{|v_k|} \int_{\mathbf{s} \in v_k} \left(\sum_{l=1}^K \omega(v_l)\overline{C}_Z(\mathbf{s}, v_l) + \psi \right) d\mathbf{s}$$

$$= \sum_{l=1}^K \omega(v_l) \frac{1}{|v_k|} \int_{\mathbf{s} \in v_k} \overline{C}_Z(\mathbf{s}, v_l)\, d\mathbf{s} + \psi$$

The integral $\frac{1}{|v_k|}\int_{\mathbf{s} \in v_k} \overline{C}_Z(\mathbf{s},v_l)\, d\mathbf{s}$ of the area-to-point covariance $\overline{C}_Z(\mathbf{s},v_l)$ over the support v_k is none other than the area-to-area covariance $\overline{C}_Z(v_k, v_l)$. Hence, the above equation becomes

$$\frac{1}{|v_k|} \int_{\mathbf{s} \in v_k} z^*(\mathbf{s})\, d\mathbf{s} = \sum_{l=1}^K \omega(v_l)\overline{C}_Z(v_k, v_l) + \psi \ ,$$

which is the left-hand-side term of the top equation of the DOK system, see equation (11).

The right-hand-side of that DOK system is precisely the k-th datum value $\bar{z}(v_k)$, which entails exact reproduction of that k-th areal datum:

$$\frac{1}{|v_k|} \int_{\mathbf{s} \in v_k} z^*(\mathbf{s})\, d\mathbf{s} = \bar{z}(v_k), \quad \forall k \ .$$

This proves the coherence characteristic of area-to-point Kriging when used for interpolation: *the mean (or sum) of point OK predictions within any areal datum support identifies the corresponding areal-average (or areal-total) datum.* In the case of predicting point density values from areal density data, the area-to-point Kriging predictions satisfy the pycnophylactic constraint.

The above proof can be modified (Kyriakidis and Yoo 2004) to account for data defined as other forms of linear combinations of point values, for example, areal data expressed as convolutions of point values (spatial averaging being a special case of convolution) or point data pertaining to spatial derivatives (differentiation being another form of convolution). The Kriging-derived point predictions will satisfy any linear constraint, provided that the latter is included as a datum in the expression of the Kriging prediction and in the Kriging system, and that its covariance with the sought-after predictions is consistently modeled based on the functional definition of that constraint. Since a point datum is a special case of an areal datum whose support collapses to a point, the above derivations also hold for the case where a subset of the K available data is defined on a point support. This entails that one can account for known point values in conjunction with areal data using the geostatistical framework.

Such point data could be obtained from GIS coverages or from ground surveys, for example, zero population density values over water bodies or high altitude regions. In all cases, the consideration of additional point values does not alter the coherence property of the predicted surface with respect to the available areal data.

It should be stressed that in the above derivations there was no explicit specification of the type (e.g., exponential or spherical) or the parameters (e.g., range or relative nugget) of the point covariance model $C_Z(\mathbf{h})$: *the coherence property of area-to-point Kriging is independent of the particular point covariance model adopted for prediction and is always satisfied as long as the area-to-area and area-to-point covariance values are derived consistently from the point covariance model.* In what follows, I illustrate the dependence of the smoothness of the resulting point predictions on the particular covariance model $C_Z(\mathbf{h})$ adopted for interpolation.

4. SMOOTHNESS OF AREA-TO-POINT KRIGING PREDICTIONS

The dual Kriging formalism allows a very informative investigation of the smoothness properties of DOK, and equivalently OK. The case of point-to-point DOK prediction is considered first, in order to illustrate the differences with area-to-point DOK prediction. The former case corresponds to the usual practice of collapsing the areal data into their support centroids. It should be noted here that point-to-point predictions obtained by Kriging are equivalent to those obtained by splines and radial basis functions under certain point covariance models $C_Z(\mathbf{h})$. This equivalence of predictions has been demonstrated long ago and discussed in many references (Matheron 1981; Dubrule 1983; Davis and Grivet 1984; Myers 1987; Wahba 1990; Cressie 1993).

Assume for a moment that the areal data have been collapsed into their respective support centroids, and hence are incorrectly considered as point support data. The DOK point prediction $z^*(\mathbf{s})$ at any location $\mathbf{s}$ from K available data $\mathbf{z} = [z(\mathbf{s}_k), k = 1,...,K]'$ is derived as a special case of equation (10):

$$z^*(\mathbf{s}) = [\boldsymbol{\eta}'\ \xi]\begin{bmatrix}\mathbf{c_s}\\1\end{bmatrix} = \sum_{k=1}^{K}\eta(\mathbf{s}_k)C_Z(\mathbf{s}-\mathbf{s}_k)+\xi \tag{13}$$

where $\boldsymbol{\eta} = [\eta(\mathbf{s}_k), k = 1,...,K]'$ is a $(K \times 1)$ vector of DOK weights, whose k-th entry $\eta(\mathbf{s}_k)$ is the weight assigned to the k-th covariance term $C_Z(\mathbf{s}-\mathbf{s}_k)$, and $\mathbf{c_s}$ denotes the $(K \times 1)$ vector of point-to-point covariance values whose k-th entry $C_Z(\mathbf{s}-\mathbf{s}_k)$ is the point covariance between the prediction location $\mathbf{s}$ and the k-th support centroid $\mathbf{s}_k$. The K weights and the Lagrange multiplier in the point-to-point DOK case are not the same as those in the area-to-point DOK case, hence the different notation.

The vector $\boldsymbol{\eta}$ of K weights is derived per solution of the point-to-point DOK system, a special case of equation (11):

$$\begin{bmatrix}\mathbf{C} & \mathbf{1}_K\\\mathbf{1}'_K & 0\end{bmatrix}\begin{bmatrix}\boldsymbol{\eta}\\\xi\end{bmatrix}=\begin{bmatrix}\mathbf{z}\\0\end{bmatrix}\ \text{or}\ \begin{cases}\sum_{l=1}^{K}\eta(\mathbf{s}_l)C_Z(\mathbf{s}_k-\mathbf{s}_l)+\xi=z(\mathbf{s}_k),\ k=1,...,K\\\sum_{l=1}^{K}\eta(\mathbf{s}_l)=0\end{cases} \tag{14}$$

where $\mathbf{C} = [C_Z(\mathbf{s}_k-\mathbf{s}_l), k = 1,...,K, l = 1,...,K]$ denotes the $(K \times K)$ matrix of point-to-point covariances between any two support centroids $\mathbf{s}_k$ and $\mathbf{s}_l$.

Note that if the prediction location $\mathbf{s}$ coincides with the k-th support centroid $\mathbf{s}_k$, then $z^*(\mathbf{s}_k) = z(\mathbf{s}_k)$, $\forall k$, no matter the point covariance model $C_Z(\mathbf{h})$. This is the well-known data exactitude characteristic of Kriging: the interpolated surface (or profile in 1D) "passes" through the sample data points.

In other words, the DOK point-to-point prediction $z^*(\mathbf{s})$ of equation (13) can be regarded as a weighted linear combination of K interpolation functions, namely the K point-to-point covariances $\{C_Z(\mathbf{s} - \mathbf{s}_k), k = 1,...,K\}$ of vector $\mathbf{c}_s$. The weights are the decorrelated (declustered) K point data values of vector $\mathbf{C}^{-1}\mathbf{z}$, which can also be interpreted as the projection of the data onto the vector space spanned by the eigenvectors of the covariance matrix $\mathbf{C}$. The physical analogy behind the DOK predictions is that of potential interpolation: the potential at any location $\mathbf{s}$ generated by K point electrical charges (sources) at the sample locations (in this case the K decorrelated point data values $\{z(\mathbf{s}_k), k = 1,...,K\}$) is a weighted linear combination of K potential functions (in this case the K point-to-point covariances $\{C_Z(\mathbf{s} - \mathbf{s}_k), k = 1,...,K\}$) evaluated at these K locations. The same interpretation can be adopted for the case of area-to-point prediction, with a point sample datum $z(\mathbf{s}_k)$ being now replaced by an areal-averaged datum $\bar{z}(v_k)$, and the point-to-point covariance function $C_Z(\mathbf{s} - \mathbf{s}_k)$, or kernel, being now replaced by the area-to-point covariance function $\bar{C}_Z(\mathbf{s},v_k)$, or regularized kernel.

If prediction is performed at P locations $\{\mathbf{s}_p, p = 1,...,P\}$ using the vector $\mathbf{z}$ of K point support data, the $(P \times 1)$ vector of point DOK predictions $\mathbf{z}^* = [z^*(\mathbf{s}_p), p = 1,...,P]'$ is written as in Equation (14),

$$\mathbf{z}^* = [\mathbf{Q}\ \mathbf{1}_P]\begin{bmatrix}\boldsymbol{\eta}\\\boldsymbol{\xi}\end{bmatrix} = [\mathbf{Q}\ \mathbf{1}_P]\begin{bmatrix}\mathbf{C} & \mathbf{1}_K\\\mathbf{1}_K' & 0\end{bmatrix}^{-1}\begin{bmatrix}\mathbf{z}\\0\end{bmatrix} \tag{15}$$

where $\mathbf{1}_P$ denotes a $(P \times 1)$ vector of unit entries, and $\mathbf{Q} = [C_Z(\mathbf{s}_p,\mathbf{s}_k), p = 1,...,P, k = 1,...,K]$ is a $(P \times K)$ matrix of point-to-point covariance values:

$$\mathbf{Q} = \begin{bmatrix}\begin{bmatrix}C_Z(\mathbf{s}_1 - \mathbf{s}_1)\\ \cdots \\ C_Z(\mathbf{s}_p - \mathbf{s}_1)\\ \cdots \\ C_Z(\mathbf{s}_P - \mathbf{s}_1)\end{bmatrix} & \cdots & \begin{bmatrix}C_Z(\mathbf{s}_1 - \mathbf{s}_k)\\ \cdots \\ C_Z(\mathbf{s}_p - \mathbf{s}_k)\\ \cdots \\ C_Z(\mathbf{s}_P - \mathbf{s}_k)\end{bmatrix} & \cdots & \begin{bmatrix}C_Z(\mathbf{s}_1 - \mathbf{s}_K)\\ \cdots \\ C_Z(\mathbf{s}_p - \mathbf{s}_K)\\ \cdots \\ C_Z(\mathbf{s}_P - \mathbf{s}_K)\end{bmatrix}\end{bmatrix}$$

where the p-th row of matrix $\mathbf{Q}$ is the vector $\mathbf{c}_p' = [C_Z(\mathbf{s}_p - \mathbf{s}_k), k = 1,...,K]$ of point covariance values between the p-th prediction location $\mathbf{s}_p$ and all K support centroids, while the k-th column of matrix $\mathbf{Q}$ is the vector $[C_Z(\mathbf{s}_p - \mathbf{s}_k), p = 1,...,P]'$ of point covariance values between all P prediction locations and the k-th support centroid $\mathbf{s}_k$.

To illustrate the influence of the covariance model on the shape of the predicted profiles, I consider a transect of $P = 100$ equally spaced prediction locations $\{\mathbf{s}_p, p = 1,...,100\}$, and $K = 2$ point support data $z(\mathbf{s}_{30}) = 20$ and $z(\mathbf{s}_{70}) = 30$ at the 30th and 70th points along that transect. The spacing between any two prediction locations is one distance unit. Four different covariance models are considered in this example: (i) an exponential covariance with practical range 10 distance units: $C_Z(|\mathbf{s} - \mathbf{s}'|) = \exp\left(-3\frac{|\mathbf{s} - \mathbf{s}'|}{10}\right)$; (ii) an exponential covariance with practical range 40

distance units: $C_Z\left(\left|\mathbf{s}-\mathbf{s}'\right|\right)=\exp\left(-3\frac{|\mathbf{s}-\mathbf{s}'|}{40}\right)$; (iii) a Gaussian covariance with prac-

tical range 40 distance units: $C_Z\left(\left|\mathbf{s}-\mathbf{s}'\right|\right)=\exp\left(-3\frac{|\mathbf{s}-\mathbf{s}'|^2}{40^2}\right)$; and (iv) an exponential

covariance with practical range 40 distance units plus a significant (50%) relative

nugget effect: $C_Z\left(\left|\mathbf{s}-\mathbf{s}'\right|\right)=0.5\delta_{\mathbf{ss}'}+0.5\exp\left(-3\frac{|\mathbf{s}-\mathbf{s}'|}{40}\right)$, where $\delta_{\mathbf{ss}'}$ is the Kronecker

δ, defined as $\delta_{\mathbf{ss}'}=1$ if $\mathbf{s}=\mathbf{s}'$, and $\delta_{\mathbf{ss}'}=0$ if $\mathbf{s}\neq\mathbf{s}'$.

The resulting covariance kernels $[C_Z(\mathbf{s}_p-\mathbf{s}_{30}),\ p=1,\dots,P]'$ and $[C_Z(\mathbf{s}_p-\mathbf{s}_{70}),\ p=1,\dots,P]'$ for the four different covariance models, that is, the resulting two columns of matrix $\mathbf{Q}$, are plotted in Figure 4. The difference between the various covariance kernels is easily appreciated. All covariance kernels attain a maximum of one, $C_Z(\mathbf{0})=1$, even in the case of a 0.5 relative nugget effect (Figure 4d) where that maximum is attained with a sharp discontinuity. Note that in the case of the exponential covariance kernel with a practical range of 10 units (Figure 4a), the two sample point data are uncorrelated (for all practical purposes). In other words, the two kernels do not overlap (strictly speaking, they do so, but they have quasi-zero values at the region of overlap).

The corresponding four profiles of DOK predictions computed via equation (15) for the four different covariance models considered above, are shown in Figure 5.
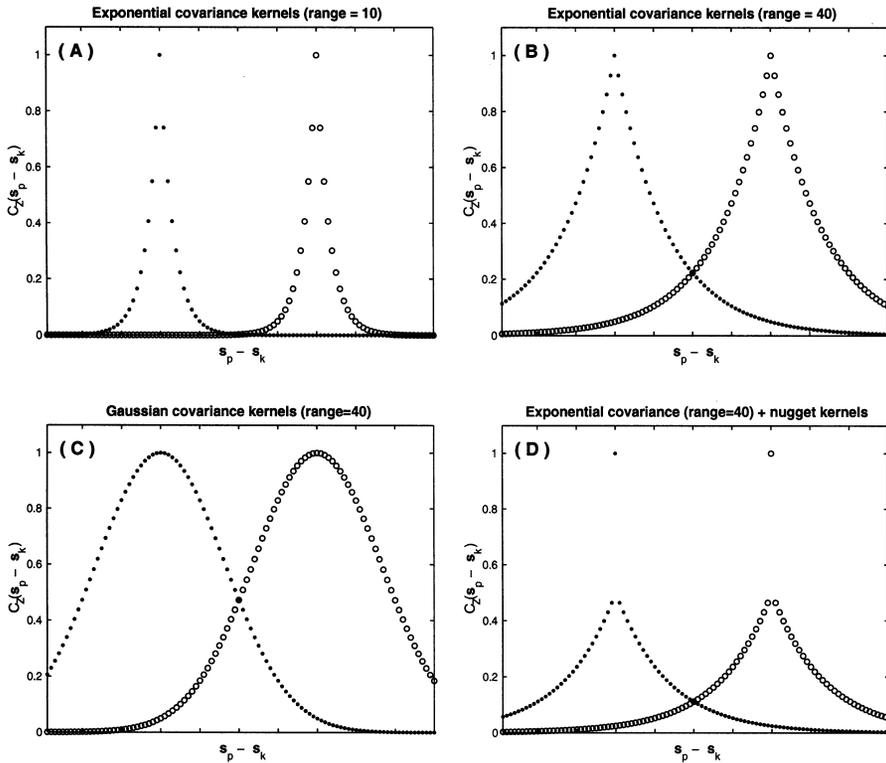

FIG. 4: Point covariance kernels $[C_Z(\mathbf{s}_p-\mathbf{s}_k),\ p=1,\dots,100]'$ corresponding to four different covariance models, evaluated at $K=2$ sample locations $\mathbf{s}_{30}$ (filled circles) and $\mathbf{s}_{70}$ (open circles).
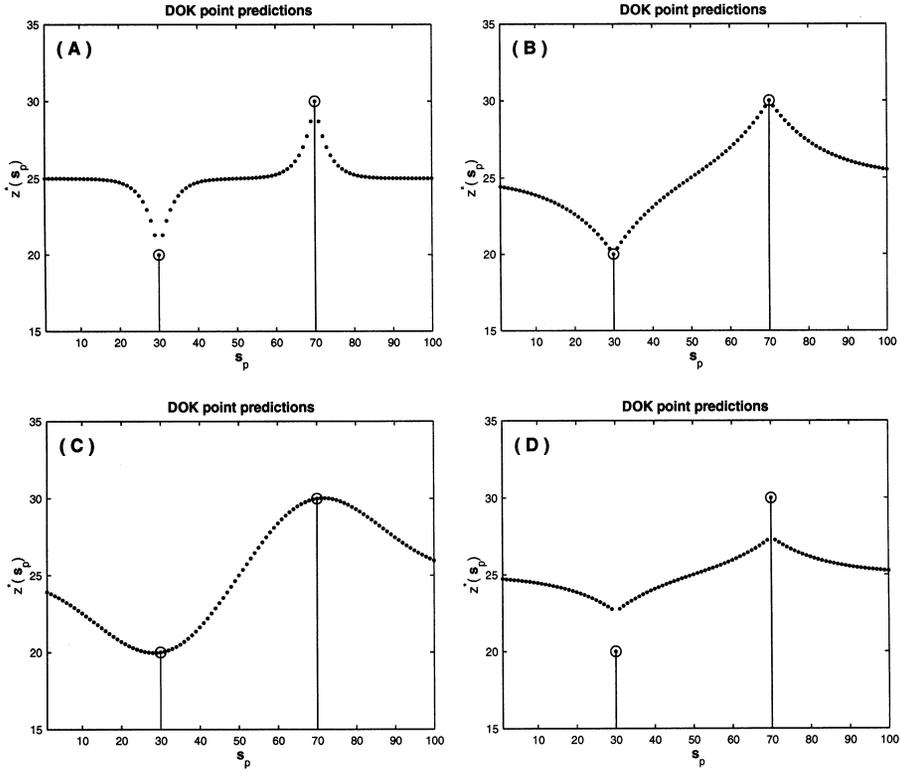
FIG. 5: Four different profiles (dotted lines) of DOK point-to-point predictions at $P = 100$ transect locations resulting from the four different point covariance kernels of figure 4 using the $K = 2$ point support sample data $z(\mathbf{s}_{30}) = 20$ and $z(\mathbf{s}_{70}) = 30$ (open circles).

Clearly, the shape of the covariance model is of paramount importance for the shape of the resulting predicted profiles. All profiles "pass" through the sample data values at the sample locations (i.e., $z^*(\mathbf{s}_k) = z(\mathbf{s}_k)$, $\forall k$). In the presence of a large nugget effect, however, this sample data reproduction is achieved via a sharp discontinuity in the predicted profile (Figure 5d). In the absence of a nugget effect, the resulting profile gets smoother as the range of the covariance kernel increases; compare Figures 5a and 5b. For the same range, the shape of the covariance kernel near the origin dictates the shape of the resulting predicted profile: for the exponential covariance case, the predicted profile at the vicinity of the data locations is more linear (less smooth) than that obtained via a Gaussian covariance; compare Figures 5b and 5c.

If, on the other hand, area-to-point prediction is performed at P locations $\{\mathbf{s}_p, p = 1,\ldots,P\}$ using the vector $\bar{\mathbf{z}}$ of K areal-average data, the $(P \times 1)$ vector of area-to-point DOK predictions $\mathbf{z}^* = [z^*(\mathbf{s}_p), p = 1,\ldots,P]'$ is written as in equation (11)

$$\mathbf{z}^* = [\overline{\mathbf{Q}}\ \mathbf{1}_P] \begin{bmatrix} \boldsymbol{\omega} \\ \boldsymbol{\psi} \end{bmatrix} = [\overline{\mathbf{Q}}\ \mathbf{1}_P] \begin{bmatrix} \overline{\mathbf{C}} & \mathbf{1}_K \\ \mathbf{1}'_K & 0 \end{bmatrix}^{-1} \begin{bmatrix} \bar{\mathbf{z}} \\ 0 \end{bmatrix} \tag{16}$$

where $\overline{\mathbf{Q}} = [\bar{C}_Z(\mathbf{s}_p, v_k), p = 1,\ldots,P, k = 1,\ldots,K]$ is a $(P \times K)$ matrix of area-to-point covariance values:

$$\overline{\mathbf{Q}} = \begin{bmatrix} \begin{bmatrix} \overline{C}_Z(\mathbf{s}_1,v_1) \\ \cdots \\ \overline{C}_Z(\mathbf{s}_p,v_1) \\ \cdots \\ \overline{C}_Z(\mathbf{s}_P,v_1) \end{bmatrix} & \cdots & \begin{bmatrix} \overline{C}_Z(\mathbf{s}_1,v_k) \\ \cdots \\ \overline{C}_Z(\mathbf{s}_p,v_k) \\ \cdots \\ \overline{C}_Z(\mathbf{s}_P,v_k) \end{bmatrix} & \cdots & \begin{bmatrix} \overline{C}_Z(\mathbf{s}_1,v_K) \\ \cdots \\ \overline{C}_Z(\mathbf{s}_p,v_K) \\ \cdots \\ \overline{C}_Z(\mathbf{s}_P,v_K) \end{bmatrix} \end{bmatrix}$$

where the p-th row of matrix $\overline{\mathbf{Q}}$ is the vector $\overline{\mathbf{c}}'_p = [\overline{C}_Z(\mathbf{s}_p,v_k), k = 1,\ldots,K]$ of area-to-point covariance values between the p-th prediction location $\mathbf{s}_p$ and all K supports, while the k-th column of matrix $\overline{\mathbf{Q}}$ is the vector $[\overline{C}_Z(\mathbf{s}_p,v_k), p = 1,\ldots,P]'$ of area-to-point covariance values between all P prediction locations and the k-th support v_k.

To illustrate the influence of the regularized covariance kernels on the shape of the predicted profiles, I consider again $P = 100$ equally spaced prediction locations $\{\mathbf{s}_p, p = 1,\ldots,100\}$ along a 1D transect. Now the available data consist of $K = 2$ areal-average measurements $\bar{z}(v_1) = 20$ and $\bar{z}(v_2) = 30$ of unequal support. The support v_1 is centered at location $\mathbf{s}_{30}$ and is a line segment of length $|v_1| = 21$ units, whereas the support v_2 is centered at location $\mathbf{s}_{70}$ and is a line segment of length $|v_2| = 11$ units. The same four covariance models considered above are also used for this example. The resulting regularized covariance kernels $[\overline{C}_Z(\mathbf{s}_p,v_1), p = 1,\ldots,P]'$ and $[\overline{C}_Z(\mathbf{s}_p,v_2), p = 1,\ldots,P]'$ for the four different covariance models, that is, the resulting two columns of matrix $\overline{\mathbf{Q}}$, are plotted in Figure 6.

The differences between the original covariance kernels (Figure 4) and the regularized ones (Figure 6) are evident. In general, all regularized kernels do not reach a maximum of one, that is, their peaks are lower than those of the kernels in Figure 4. The difference in these peaks is a function of the support size and the covariance model: for a given plot, the kernel on the left is regularized by a larger support ($|v_1| = 21$) than the kernel on the right ($|v_2| = 11$). All regularized kernels appear much smoother than their original counterparts, with the exception of the kernels in Figure 6c for which the differences are minimal: regularization of covariance kernels with parabolic shape at the origin and large range (with respect to the extent of the support) yields regularized kernels that are similar to their original counterparts. This is due to the fact that averaging now includes very similar point covariance values (especially near the origin).

One of the most pronounced changes is that of the regularized kernels of Figure 6a with respect to those of Figure 4a: regularization of covariance kernels with linear shape at the origin and small range (with respect to the extent of the support) yields regularized kernels with nonlinear shape at the origin that are much smoother and less peaked than their original counterparts. Another important difference is that found between Figure 6d and Figure 4d: the nugget effect in the original kernels is reduced in their regularized counterparts. This reduction of the point support nugget effect due to regularization depends on the relative magnitude of that nugget effect and on the size (and shape) of the support. The regularized kernels, however, are still discontinuous; that discontinuity has now "migrated" at separation distances $\mathbf{s} - \mathbf{s}_i$ equal to the size of the support v_k. Note that $\mathbf{s}_i$ denotes a discretization location; see equation (6).

The corresponding four profiles of DOK area-to-point predictions computed in equation (16) for the four different covariance models considered above, are shown in Figure 7. The shape of the regularized covariance kernel is again of paramount importance for the shape of the resulting predicted profiles. All predicted profiles are coherent, because the sample areal data are reproduced by construction, that is, $1/21 \sum_{p=20}^{40} z^*(\mathbf{s}_p) = 20 = \bar{z}(v_1)$, and $1/11 \sum_{p=65}^{75} z^*(\mathbf{s}_p) = 30 = \bar{z}(v_2)$.
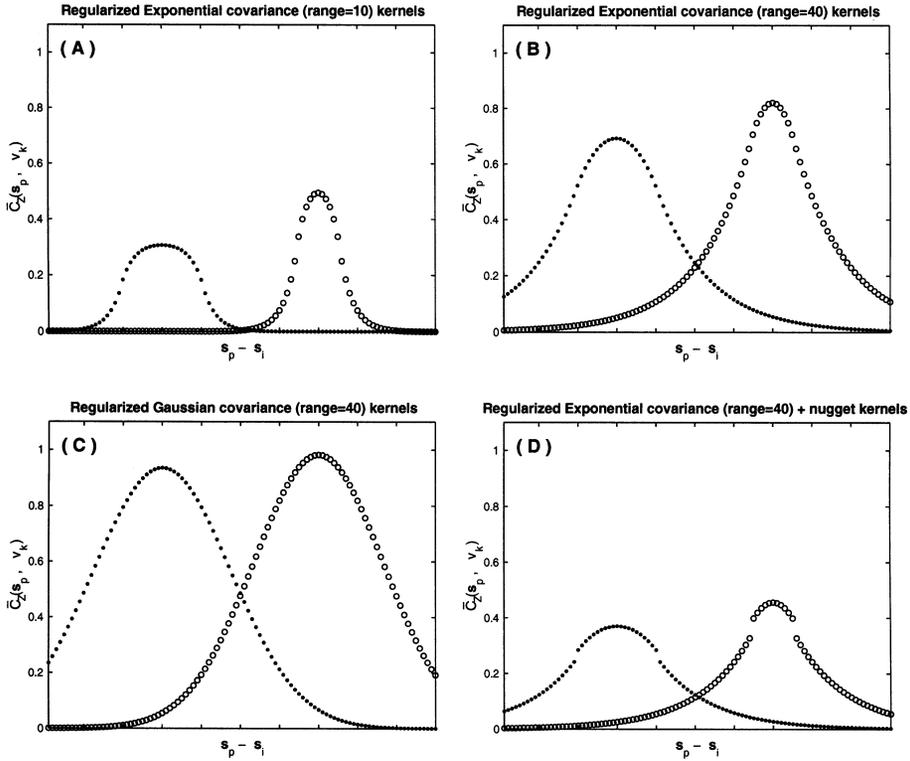
FIG. 6: Regularized covariance kernels $[\bar{C}_Z(\mathbf{s}_p, v_k), p = 1,\ldots,100]'$ corresponding to four different point covariance models. Regularization is performed by $K = 2$ supports: v_1 with length $|v_1| = 21$ centered at location $\mathbf{s}_{30}$ (corresponding regularized kernel marked with filled circles), and v_2 with length $|v_2| = 11$ centered at location $\mathbf{s}_{70}$ (corresponding regularized kernel marked with open circles).

In the presence of a large nugget effect, the reproduction of the areal data is achieved via a discontinuity in the predicted profile at the support boundaries (Figure 7d). Area-to-point interpolation yields point surfaces with sharper discontinuities at support boundaries as the relative nugget of the point covariance model $C_Z(\mathbf{h})$ increases (compare Figures 7b and 7d). At the limit, area-to-point interpolation in the case of a pure nugget effect for $C_Z(\mathbf{h})$ yields the choropleth map; see section 5.1. In the absence of a nugget effect, the resulting profile becomes smoother as the range of the covariance kernel increases; compare Figures 7a and 7b. Note also that, in the case of the Gaussian covariance kernel, the resulting area-to-point profile (Figure 7c) is very similar to its point-to-point prediction counterpart (Figure 5c), with the very important difference that the latter is not coherent with respect to the areal data (it is coherent with respect to the point data).

5. LINKS WITH OTHER AREA-TO-POINT SPATIAL INTERPOLATION METHODS

In this section, I identify key links between the proposed geostatistical framework and several existing approaches for area-to-point interpolation, namely choropleth mapping, different forms of kernel smoothing, and pycnophylactic interpolation with Laplacian smoothing, on a quasi-infinite domain without non-negativity constraints. More precisely, I illustrate that the above methods invoke particular (implicit or ex-
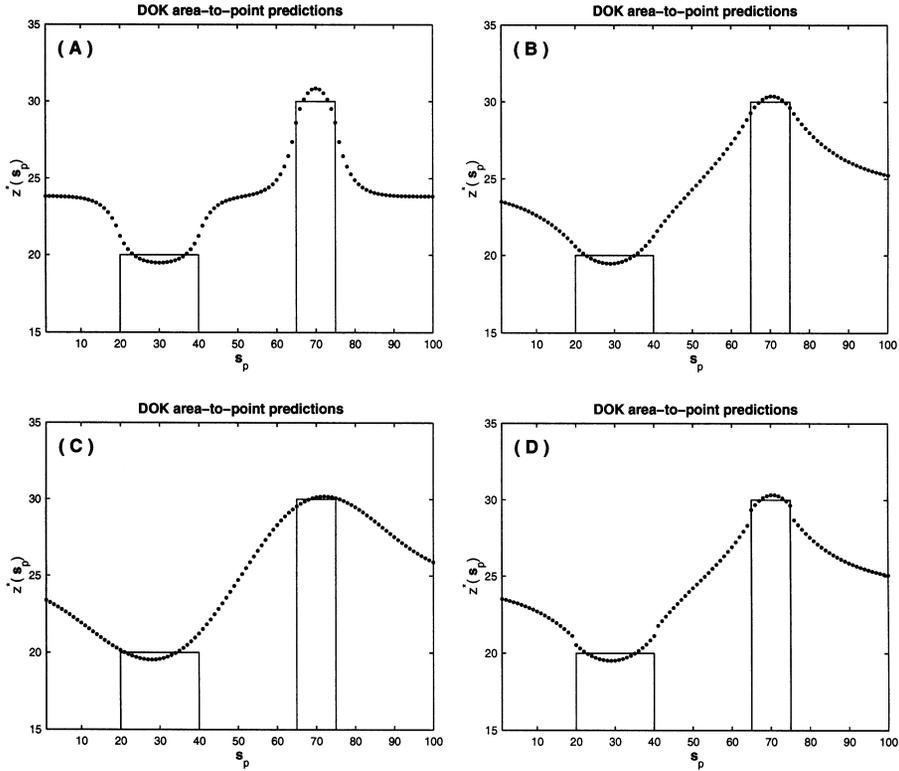
FIG. 7: Four different profiles (dotted lines) of area-to-point DOK predictions at $P = 100$ transect locations resulting from the four different regularized covariance kernels of figure 6 using the $K = 2$ areal-average sample data $\bar{z}(v_1) = 20$ and $\bar{z}(v_2) = 30$; solid lines delineate the supports of the two areal data: v_1 with length $|v_1| = 21$ centered at location $\mathbf{s}_{30}$, and v_2 with length $|v_2| = 11$ centered at location $\mathbf{s}_{70}$.

plicit) assumptions regarding the point covariance model $C_Z(\mathbf{h})$, and thus can be regarded as special cases (or particular implementations) of the proposed general geostatistical framework for area-to-point interpolation.

Before proceeding, it is necessary to clarify some terms that often have different meaning in geography and spatial statistics. Any method for area-to-point prediction that reproduces the areal data at their supports is actually an interpolation method, in analogy to methods of point-to-point prediction that reproduce the point support sample data at their locations. Smoothing (or filtering) pertains to prediction methods that do not satisfy such areal data constraints, in analogy to methods of point-to-point prediction that do not reproduce the point support sample data at their locations. The term smoothing, however, is often used in Geography to indicate: (i) changes in areal values assimilated to point centroids, or (ii) elimination of the abrupt discontinuities incurred by the choropleth map. None of these reasons, however, is linked to the areal data constraint, which is precisely what should dictate the use of the qualifier interpolation versus smoothing.

5.1. The Choropleth Map Case

The choropleth map is the simplest coherent method for area-to-point interpolation, whereby any predicted value $z^*(\mathbf{s})$ within the support v_k of an areal-average datum $\bar{z}(v_k)$ is equal to that datum value. Although choropleth mapping is typically

used as a visualization tool, it can be regarded as a special case of areal interpolation, whereby the source data are defined on areal supports and the target data are point values of the same variable. In many applications (including visualization), and in the absence of other information on the spatial distribution of point values, choropleth maps are used for inference purposes while acknowledging the pitfalls associated with such an inference. In addition, it is not rare to advocate the use of choropleth mapping under the assumption of constant point support attribute values within areal supports. As it is shown in this section, this assumption is a consequence of another assumption: lack of spatial correlation at the point support level.

Assume, for a moment, that the point support values are spatially uncorrelated, or formally:

$$C_Z(\mathbf{s} - \mathbf{s}') = C_Z(\mathbf{0})\delta_{\mathbf{ss}'} = \begin{cases} C_Z(\mathbf{0}) & \text{if } \mathbf{s} = \mathbf{s}' \\ 0 & \text{if } \mathbf{s} \neq \mathbf{s}' \end{cases}, \tag{17}$$

which is the classical definition of the covariance of a white noise process with variance $C_Z(\mathbf{0})$, representing complete lack of spatial correlation (also termed pure nugget effect in the geostatistical jargon).

Recall from equation (6) that the area-to-point covariance $\bar{C}_Z(\mathbf{s}, v_l)$ between a point $\mathbf{s} \in v_k$ and a support v_l is the average point covariance $C_Z(\mathbf{s} - \mathbf{s}')$ between all possible vectors formed by two points $\mathbf{s} \in v_k$ and $\mathbf{s}' \in v_l$. Under the assumption of spatial independence, and for disjointed supports, it follows that $\bar{C}_Z(\mathbf{s}, v_l) = 0, \forall l \neq k$, since $\delta_{\mathbf{ss}'} = 0, \forall \mathbf{s} \in v_k, \mathbf{s}' \in v_l$. In addition, the covariance $\bar{C}_Z(\mathbf{s}, v_k)$ between a point $\mathbf{s} \in v_k$ and the support v_k can be expressed as:

$$\bar{C}_Z(\mathbf{s}, v_k) = \frac{1}{|v_k|} \int_{\mathbf{s}' \in v_k} C_Z(\mathbf{s} - \mathbf{s}') \, d\mathbf{s}' = \frac{1}{|v_k|} \int_{\mathbf{s}' \in v_k} C_Z(\mathbf{0})\delta_{\mathbf{ss}'} \, d\mathbf{s}' = \frac{C_Z(\mathbf{0})}{|v_k|}, \quad \forall \mathbf{s} \in v_k.$$

Recall from equation (5) that the area-to-area covariance $\bar{C}_Z(v_k, v_l)$ between two supports v_k and v_l is the average point covariance $C_Z(\mathbf{s} - \mathbf{s}')$ between all possible vectors formed by two points $\mathbf{s} \in v_k$ and $\mathbf{s}' \in v_l$. Under the assumption of spatial independence and for disjointed supports, it follows that $\bar{C}_Z(v_k, v_l) = 0, \forall k \neq l$, since $\delta_{\mathbf{ss}'} = 0, \forall \mathbf{s} \in v_k, \mathbf{s}' \in v_l$. In addition, the variance $\bar{C}_Z(v_k, v_k)$ of the k-th areal support v_k can be expressed as:

$$\bar{C}_Z(v_k, v_k) = \frac{1}{|v_k|} \frac{1}{|v_k|} \int_{\mathbf{s} \in v_k} \int_{\mathbf{s}' \in v_k} C_Z(\mathbf{s} - \mathbf{s}') \, d\mathbf{s}' d\mathbf{s} = \frac{1}{|v_k|^2} \int_{\mathbf{s} \in v_k} \int_{\mathbf{s}' \in v_k} C_Z(\mathbf{0})\delta_{\mathbf{ss}'} \, d\mathbf{s}' d\mathbf{s}$$

$$= \frac{1}{|v_k|^2} C_Z(\mathbf{0}) |v_k| = \frac{C_Z(\mathbf{0})}{|v_k|}.$$

In summary, for the case of disjointed supports, the assumption of spatial independence for the point support values yields the following area-to-point and area-to-area covariances:

$$\bar{C}_Z(\mathbf{s}, v_k) = \begin{cases} 0 & \forall \mathbf{s} \notin v_k \\ C_Z(\mathbf{0}) / |v_k| & \forall \mathbf{s} \in v_k \end{cases} \quad \text{and} \quad \bar{C}_Z(v_k, v_l) = \begin{cases} 0 & \text{if } k \neq l \\ C_Z(\mathbf{0}) / |v_k| & \text{if } k = l \end{cases}.$$

Using the above expressions for the area-to-area covariance terms, the top relation of the DOK system of equation (11) becomes

$$\omega(v_k)\overline{C}_Z(v_k, v_k) + \sum_{l \neq k}^{K-1} \omega(v_l)\overline{C}_Z(v_k, v_l) + \psi = \overline{z}(v_k),$$

$$\forall k \Rightarrow \omega(v_k)\frac{C_Z(\mathbf{0})}{|v_k|} + \psi = \overline{z}(v_k), \forall k$$

with the bottom relation $\sum_{l=1}^{K} \omega(v_l) = 0$ of the DOK system of equation (11) just being used to solve for ψ.

Using the above expressions for the area-to-point covariance terms, the DOK prediction $z^*(\mathbf{s})$ of equation (10) for any location $\mathbf{s} \in D$ becomes

$$z^*(\mathbf{s}) = \underbrace{\omega(v_k)\overline{C}_Z(\mathbf{s}, v_k)}_{\mathbf{s} \in v_k} + \sum_{l \neq k}^{K-1} \omega(v_l)\overline{C}_Z(\mathbf{s}, v_l) + \psi = \omega(v_k)\frac{C_Z(\mathbf{0})}{|v_k|} + \psi \ .$$

From the last two equations, it is evident that $z^*(\mathbf{s}) = \overline{z}(v_k)$, $\forall \mathbf{s} \in v_k$, that is, the predicted values within any support v_k are all equal to the corresponding areal-average datum $\overline{z}(v_k)$. Note that the above result does not depend on the variance $C_Z(\mathbf{0})$ of the point support values.

It has thus been shown that the choropleth map solution of the area-to-point interpolation problem can be viewed as a particular case of the geostatistical framework, under the assumption of spatial independence regarding the point support values. Evidently, this assumption of spatial independence at the point support level is rather unrealistic. Consequently, choropleth mapping could be extremely misleading and should be used (if at all) with utmost caution.

5.2. The Kernel Smoothing Case and Its Variants

In what follows, I illustrate the connections between the kernel smoothing method of Brillinger (1990, 1994), as well as two closely related approaches suggested by Bracken and Martin (1989) and Martin (1996), and the proposed geostatistical framework. In particular, I demonstrate that Brillinger's local weighting scheme can be viewed as a particular (albeit incoherent) implementation of the proposed geostatistical approach. The original development of Brillinger's method did not consider the areal data constraint and was correctly termed a smoothing method. The motivation for this method was the construction of contour lines from areal data for visual display and exploratory data analysis. Brillinger's method, however, has also been used for spatial prediction and has been contrasted to Tobler's pycnophylactic interpolation; see, for example, Gotway and Young (2002).

In this subsection, $\tilde{z}(v_k)$ denotes the integral (or sum in the discrete case) of all point values $z(\mathbf{s})$ in support v_k:

$$\tilde{z}(v_k) = \int_{\mathbf{s} \in v_k} z(\mathbf{s})\, d\mathbf{s} \simeq \sum_{i=1}^{P_k} z(\mathbf{s}_i), \mathbf{s}_i \in v_k \ .$$

Similarly, $\tilde{C}_Z(\mathbf{s}, v_k)$ denotes the total area-to-point covariance, that is, the sum of point covariance values $C_Z(\mathbf{s} - \mathbf{s}')$ between a prediction location $\mathbf{s}$ and all locations $\mathbf{s}' \in v_k$:

$$\tilde{C}_Z(\mathbf{s}, v_k) = \int\limits_{\mathbf{s}' \in v_k} Cov\{Z(\mathbf{s}), Z(\mathbf{s}')\}\, d\mathbf{s}' \simeq \sum_{i=1}^{P_k} C_Z(\mathbf{s} - \mathbf{s}_i),\, \mathbf{s}_i \in v_k\,. \tag{18}$$

Finally, $\tilde{C}_Z(v_k, v_l)$ denotes the total area-to-area covariance, that is, the sum of the point covariance values $C_Z(\mathbf{s} - \mathbf{s}')$ between any two locations $\mathbf{s} \in v_k$ and $\mathbf{s}' \in v_l$:

$$\tilde{C}_Z(v_k, v_l) = \int\limits_{\mathbf{s} \in v_k} \int\limits_{\mathbf{s}' \in v_l} Cov\{Z(\mathbf{s}), Z(\mathbf{s}')\}\, d\mathbf{s}' d\mathbf{s} \simeq \sum_{i=1}^{P_k} \sum_{j=1}^{P_l} C_Z(\mathbf{s}_i - \mathbf{s}_j),\, \mathbf{s}_i \in v_k, \mathbf{s}_j \in v_l\,. \tag{19}$$

Consider now the case of area-to-point Simple Kriging (SK) with K areal-total data $\{\tilde{z}(v_k), k = 1,\ldots,K\}$, and a known stationary mean $m_Z = 0$ (in this case, the mean $\tilde{m}_Z$ of the areal data is also zero):

$$z^*(\mathbf{s}) = m_Z + \sum_{k=1}^{K} \phi_{\mathbf{s}}(v_k)[\tilde{z}(v_k) - \tilde{m}_Z] = \sum_{k=1}^{K} \phi_{\mathbf{s}}(v_k)\tilde{z}(v_k) \tag{20}$$

where $\phi_{\mathbf{s}}(v_k)$ denotes the weight assigned to the k-th areal datum $\tilde{z}(v_k)$ for prediction at location $\mathbf{s}$.

The corresponding system of normal equations (area-to-point SK system) is written as

$$\sum_{l=1}^{K} \phi_{\mathbf{s}}(v_l)\tilde{C}_Z(v_k, v_l) = \tilde{C}_Z(\mathbf{s}, v_k),\ \ k = 1,\ldots,K\,, \tag{21}$$

which is a variant of the area-to-point OK system of equation (4) with no constraint on the weights.

Brillinger (1990, 1994) proposed a smoothing scheme very similar to that of equations (20) and (21), which accounts explicitly for the support differences between the available areal data and the point predictions, under the assumption of spatial independence at the area level only. In other words, although the areal data are assumed uncorrelated, that is, $\tilde{C}_Z(v_k, v_l) = 0$, $\forall k \neq l$, they are still correlated with the point values, that is, $\tilde{C}_Z(\mathbf{s}, v_k) \neq 0$, even for locations $\mathbf{s} \notin v_k$; it is very difficult, however, to conceptualize a process that exhibits such a pattern of spatial correlation. A possible, but very particular, scenario would correspond to a point covariance with a range that does not exceed the extent of any support. In this case, all the areal covariance terms would be zero, and the only nonzero area-to-point covariance term would be $\tilde{C}_Z(\mathbf{s}, v_k$, only for $\mathbf{s} \in v_k$.

If one adopts the above assumptions, the solution of the above area-to-point SK system yields the following K weights:

$$\phi_{\mathbf{s}}(v_k) = \frac{\tilde{C}_Z(\mathbf{s}, v_k)}{\tilde{C}_Z(v_k, v_k)},\ \ k = 1,\ldots,K$$

and the area-to-point SK prediction of equation (20) becomes

$$z^*(\mathbf{s}) = \sum_{k=1}^{K} \frac{\tilde{C}_Z(\mathbf{s}, v_k)}{\tilde{C}_Z(v_k, v_k)} \tilde{z}(v_k),$$

or, in words, only the areal datum $\tilde{z}(v_k)$ at support v_k within which prediction is performed is taken into account.

The denominator $\tilde{C}_Z(v_k, v_k)$ and numerator $\tilde{C}_Z(\mathbf{s}, v_k)$ of the k-th weight $\phi_\mathbf{s}(v_k)$ are related as

$$\tilde{C}_Z(v_k, v_k) = \int\limits_{\mathbf{s} \in v_k} \tilde{C}_Z(\mathbf{s}, v_k) \, d\mathbf{s} \simeq \sum_{i=1}^{P_k} \tilde{C}_Z(\mathbf{s}_i, v_k), \, \mathbf{s}_i \in v_k \, , \tag{22}$$

or, in words, the total (co)variance of the k-th support v_k is the integral of the total area-to-point covariance $\tilde{C}_Z(\mathbf{s}, v_k)$ within that support.

If the integral $\int_{\mathbf{s} \in v_k} \phi_\mathbf{s}(v_k) \, d\mathbf{s}$ of all weights applied to the k-th areal datum were one, then the resulting set of point predictions would be coherent, that is, the sum of the predicted point values within any support v_k would identify the corresponding areal-total datum $\tilde{z}(v_k)$. Not all prediction locations, however, are located within the support v_k, and consequently equation (22) holds only for a subset of prediction locations $\{\mathbf{s} \in v_k\}$. In other words, the common denominator $\tilde{C}_Z(v_k, v_k)$ of each weight $\phi_\mathbf{s}(v_k)$ is not equal to the sum of all numerator terms, that is, $\tilde{C}_Z(v_k, v_k) \neq \int_{\mathbf{s} \in v_k} \tilde{C}_Z(\mathbf{s}, v_k) \, d\mathbf{s}$, which entails that the sum of the weights applied to the k-th areal datum is not equal to one.

The above variant of kernel smoothing does not yield coherent predictions, due precisely to the assumption of uncorrelated areal data, or, due to $\tilde{C}_Z(v_k, v_k) = 0$, $\forall k \neq l$. This assumption leads to a set of weights that do not satisfy the relation $\int_{\mathbf{s} \in v_k} \phi_\mathbf{s}(v_k) \, d\mathbf{s} = 1$, which can be regarded as another way of ensuring the coherence of point predictions.

It has thus been shown that the kernel smoothing method of Brillinger (1990, 1994) can be viewed as a variant of the geostatistical approach, which does not yield coherent predictions due to the inconsistent modeling of area-to-area and area-to-point covariances (in particular due to the assumption of uncorrelated areal data).

To ensure the coherence of point predictions Bracken and Martin (1989) proposed to standardize the above weights so that they sum to one. Their approach, however, assumes that the areal datum $\tilde{z}(v_k)$ can be collapsed to a point datum $z(\mathbf{s}_k)$ typically located at the centroid of support v_k, and thus has the following shortcomings: (i) it incorrectly associates an areal datum with a location; (ii) the support differences between the data and sought-after predictions is not accounted for; (iii) the areal data are assumed uncorrelated; and (iv) the predicted point values are rendered coherent by a simple standardization of the resulting weights, which is not equivalent to solving the theoretically correct Kriging system. In addition, since their method reproduces the areal data, it should be termed kernel interpolation instead of kernel smoothing.

In what follows, I investigate the proposal of Martin (1996) to truncate the covariance kernel at the boundaries of each support v_k. Note that, contrary to the original approach of Martin (1996) that treats areal data as point data assigned to support centroids (with all associated shortcomings mentioned above), the support differences are hereafter correctly accounted for.

Consider the case of adopting the following model of nonstationary spatial correlation at the point support level:

$$C_Z(\mathbf{s}, \mathbf{s}') = \begin{cases} C_Z(\mathbf{s} - \mathbf{s}') & \text{if } \mathbf{s}, \mathbf{s}' \in v_k \\ 0 & \text{if } \mathbf{s} \in v_k \text{ and } \mathbf{s}' \notin v_k \end{cases}, \tag{23}$$

which is a rather unrealistic assumption because it entails that support boundaries are essentially barriers. In most cases, however, this is not true, since such boundaries are arbitrary and artificial.

The above equation implies the following area-to-point covariance $\tilde{C}_Z(\mathbf{s}, v_k)$:

$$\tilde{C}_Z(\mathbf{s}, v_k) = \begin{cases} \int\limits_{\mathbf{s}' \in v_k} C_Z(\mathbf{s} - \mathbf{s}') \, d\mathbf{s}' & \text{if } \mathbf{s}, \mathbf{s}' \in v_k \\ 0 & \text{if } \mathbf{s} \in v_k \text{ and } \mathbf{s}' \notin v_k \end{cases}. \tag{24}$$

In light of this covariance structure, the area-to-point SK prediction of equation (20) at any location $\mathbf{s} \in v_k$ is written as

$$z^*(\mathbf{s}) = \frac{\tilde{C}_Z(\mathbf{s}, v_k)}{\tilde{C}_Z(v_k, v_k)} \tilde{z}(v_k) + \sum_{l \neq k}^{K-1} \frac{\tilde{C}_Z(\mathbf{s}, v_l)}{\tilde{C}_Z(v_l, v_l)} \tilde{z}(v_l) = \frac{\tilde{C}_Z(\mathbf{s}, v_k)}{\tilde{C}_Z(v_k, v_k)} \tilde{z}(v_k), \quad \mathbf{s} \in v_k$$

or, in words, only the areal datum $\tilde{z}(v_k)$ at support v_k within which prediction is performed is taken into account.

Per equation (22), the denominator $\tilde{C}_Z(v_k, v_k)$ of each weight $\phi_{\mathbf{s}}(v_k)$ is a normalization constant, independent of the prediction location $\mathbf{s}$, which ensures that the integral of all weights applied to the k-th areal datum $\tilde{z}(v_k)$ equals one, or, $\int_{\mathbf{s} \in v_k} \phi_{\mathbf{s}}(v_k) \, d\mathbf{s} = 1$. The truncation procedure thus ensures coherent point predictions.

Recall, however, that (per the dual Kriging formalism) the interpolated surface is a superposition of covariance functions, which entails that the discontinuous covariance structure of equation (24) will induce discontinuities in the predicted point surface at the boundaries of the K supports. The nonstationary covariance structure of equation (24) should therefore be avoided if support boundaries are artificial and interactions across areas are important to model. Since both conditions typically apply in practice, I do not recommend this local truncation of the covariance kernel, which yields coherent predictions at the expense of artifact discontinuities at the support boundaries that call for rather adhoc procedures to correct them a posteriori.

It has thus been shown that: the truncated area-to-point kernel smoothing method, which is the theoretically correct interpretation of the approach proposed by Martin (1996) can be viewed as a particular implementation of the geostatistical approach, which yields coherent but discontinuous (at the support boundaries) predictions.

5.3. The Case of Laplacian Smooth Pycnophylactic Interpolation

Tobler (1979) proposed a method for area-to-point spatial interpolation, whereby the smoothness of point support values is dictated by Laplace's partial differential equation (PDE) and some prescribed boundary conditions (BCs). In his method, Tobler iteratively solves Laplace's PDE based on the prescribed BCs via a finite difference approximation, under the constraint of non-negative point predictions $z^*(\mathbf{s}) \geq 0$ and mass preservation $\int_{\mathbf{s} \in v_k} z^*(\mathbf{s}) \, d\mathbf{s} = \tilde{z}(v_k)$. In what follows, I derive the particular point support semivariogram types that are implicitly associated with Tobler's method. In addition, I demonstrate that Tobler's solution, on a quasi-infinite domain without the non-negativity constraint, can be derived as a particular case of the proposed geostatistical framework.

Assume that there is prior knowledge regarding the variable under study, which dictates that the unknown true point support surface $\{z(\mathbf{s}), \mathbf{s} \in D\}$ satisfies the 2-D Poisson's PDE. This time-independent PDE is widely used in electrostatic (potential) theory, and is expressed as

$$\nabla^2[z(\mathbf{s})] = f(\mathbf{s}) \Rightarrow \frac{\partial^2 z(\mathbf{s})}{\partial x^2} + \frac{\partial^2 z(\mathbf{s})}{\partial y^2} = f(\mathbf{s}), \ \mathbf{s} \in D \tag{25}$$

where ∇^2 denotes the linear differentiation operator, and $f(\mathbf{s})$ denotes a spatially distributed source term. The homogeneous version of the above PDE, that is, with no source term $f(\mathbf{s}) = 0$, $\forall \mathbf{s}$, is none other than Laplace's equation. In this paper, I focus on the free-space solution of Poisson's equation, and consider a quasi-infinite domain D with constant flux at its boundary: $\partial z(\mathbf{s})/\partial \boldsymbol{\tau} = 0$, where $\boldsymbol{\tau}$ denotes a unit vector perpendicular to the boundary of D, with outward direction.

The solution to equation (25) in terms of linear operators is simply:

$$z(\mathbf{s}) = (\nabla^2)^{-1}[f(\mathbf{s})], \ \mathbf{s} \in D$$

where $(\nabla^2)^{-1}$ denotes the inverse of the differential operator, or the integral operator.

Since integration can also be seen as a convolution operation, the above solution can be defined as

$$z(\mathbf{s}) = (\nabla^2)^{-1}[f(\mathbf{s})] = \int_{\mathbf{s}' \in D} G(\mathbf{s}, \mathbf{s}') f(\mathbf{s}') \, d\mathbf{s}', \ \mathbf{s} \in D \tag{26}$$

where the convolution kernel $G(\mathbf{s}, \mathbf{s}')$ is the *Green's function* associated with the differential operator ∇^2 (Greenberg 1971).

The Green's function is defined by the following relation:

$$\nabla^2[G(\mathbf{s}, \mathbf{s}')] = \delta(\mathbf{s} - \mathbf{s}'), \ \mathbf{s}, \mathbf{s}' \in D$$

where $\delta(\mathbf{s} - \mathbf{s}')$ denotes the Dirac delta function that satisfies $\int_{\mathbf{s}' \in D} \delta(\mathbf{s} - \mathbf{s}') \, d\mathbf{s}' = 1$ and $\int_{\mathbf{s}' \in D} \delta(\mathbf{s} - \mathbf{s}') f(\mathbf{s}') \, d\mathbf{s}' = f(\mathbf{s})$.

In the terminology of linear systems theory, the Green's function is the impulse response of the differential operator ∇^2 to a unit source $\delta(\mathbf{s} - \mathbf{s}')$ placed at location $\mathbf{s}$. If the solution to Poisson's equation is interpreted as a potential field, then the Green's function quantifies the influence of a unit charge, or source, $\delta(\mathbf{s} - \mathbf{s}')$ placed at location $\mathbf{s}$ on any other point $\mathbf{s}'$ in the domain. In the general case, the Green's function is nonstationary, hence the notation $G(\mathbf{s}, \mathbf{s}')$ instead of $G(\mathbf{s} - \mathbf{s}')$, because it depends on the prescribed BCs. In other words, the influence of a unit point source $\delta(\mathbf{s} - \mathbf{s}')$ placed at location $\mathbf{s}$ on another location $\mathbf{s}'$ is different when $\mathbf{s}'$ is close to the domain boundary than when it is not. From this view point, equation (26) states that the solution to Poisson's equation is a superposition of elementary solutions given by the product of Green's functions $G(\mathbf{s}, \mathbf{s}')$ with the (nonunit) source term $f(\mathbf{s}')$.

It can be easily shown that $(\nabla^2)^{-1} f(\mathbf{s})$ solves equation (25), since

$$\nabla^2[z(\mathbf{s})] = \nabla^2 \left[\int_{\mathbf{s}' \in D} G(\mathbf{s}, \mathbf{s}') f(\mathbf{s}') \, d\mathbf{s}' \right] = \int_{\mathbf{s}' \in D} \nabla^2[G(\mathbf{s}, \mathbf{s}')] f(\mathbf{s}') \, d\mathbf{s}'$$

$$= \int_{\mathbf{s}' \in D} \delta(\mathbf{s} - \mathbf{s}') f(\mathbf{s}') \, d\mathbf{s}' = f(\mathbf{s}), \quad \mathbf{s} \in D \,.$$

In the case of a quasi-infinite domain, the boundary conditions play no role, and the (free-space) Green's function is stationary, or $G(\mathbf{s},\mathbf{s}') = G(\mathbf{s} - \mathbf{s}')$. In addition, the free-space Green's function associated with the Laplacian operator ∇^2 depends on the dimensionality of the problem at hand (Greenberg 1971). In particular

- in 1-D, $G(\mathbf{s},\mathbf{s}') = |\mathbf{s} - \mathbf{s}'|$: the Green's function is the Euclidean distance between any two 1-D-coordinates;
- in 2-D, $G(\mathbf{s},\mathbf{s}') = \frac{1}{2\pi} \log(|\mathbf{s} - \mathbf{s}'|)$: the Green's function is the logarithm of the Euclidean distance between any two 2-D-coordinate vectors; and
- in 3-D, $G(\mathbf{s},\mathbf{s}') = -\frac{1}{4\pi}|\mathbf{s} - \mathbf{s}'|^{-1}$: the Green's function is the negative inverse Euclidean distance between any two 3-D-coordinate vectors.

It is necessary here to introduce the concept of a generalized covariance function $K_Z(\mathbf{h})$ for an intrinsic RF model of order zero (IRF-0), that is, a RF with stationary increments and a possibly unbounded (without sill) semivariogram $\gamma_Z(\mathbf{h})$. The term *order zero* implies that the trend component is a polynomial of order zero, that is, a constant m_Z; see Chilès and Delfiner (1999). In this case, the generalized covariance is $K_Z(\mathbf{h}) = A - \gamma_Z(\mathbf{h})$, where A is an arbitrary constant such that $A - \gamma_Z(\mathbf{h}) \geq 0$. In other words, the generalized covariance of order zero is equal to $-\gamma_Z(\mathbf{h})$ up to an arbitrary constant. This arbitrary constant A in the definition of a generalized covariance does not affect the solution of the DOK system of equation (14) or the DOK prediction of equation (13), due to the zero sum constraint on the weights.

By replacing the ordinary covariance $C_Z(\mathbf{h})$ with the generalized covariance $K_Z(\mathbf{h})$ in the DOK prediction of equation (13), one can express that prediction in terms of the point semivariogram model $\gamma_Z(\mathbf{h})$:

$$z^*(\mathbf{s}) = \sum_{k=1}^{K} \eta(\mathbf{s}_k)\gamma_Z(\mathbf{s} - \mathbf{s}_k) + \xi, \quad \mathbf{s} \in D$$

where $\gamma_Z(\mathbf{s} - \mathbf{s}_k)$ denotes the semivariogram between the prediction location $\mathbf{s}$ and any sample location $\mathbf{s}_k$. The DOK weights $\{\eta(\mathbf{s}_k), k = 1,\ldots,K\}$ and the Lagrange parameter ξ are obtained by solving the DOK system of equation (14), whereby the data-to-data covariance term $C_Z(\mathbf{s}_k - \mathbf{s}_l)$ is replaced by the semivariogram term $\gamma_Z(\mathbf{s}_k - \mathbf{s}_l)$. The solution of this semivariogram-based system yields the same weights and Lagrange parameter as the DOK system of equation (14).

The set of all point DOK predictions can be regarded as a discrete approximation to the continuous solution of equation (26), with K point sources $\{\eta(\mathbf{s}_k), k = 1,\ldots,K\}$ replacing the continuous source term $f(\mathbf{s}')$, and the stationary point semivariogram term $\gamma_Z(\mathbf{s} - \mathbf{s}_k)$ playing the role of the free-space Green's function (Matheron 1971).

Indeed, if one applies the differential operator ∇^2 to the DOK-predicted surface $\{z^*(\mathbf{s}), \mathbf{s} \in D\}$ of equation (13), the linearity of ∇^2 entails

$$\nabla^2[z^*(\mathbf{s})] = \nabla^2\left[\sum_{k=1}^{K} \eta(\mathbf{s}_k)\gamma_Z(\mathbf{s} - \mathbf{s}_k) + \xi\right] = \sum_{k=1}^{K} \eta(\mathbf{s}_k)\nabla^2[\gamma_Z(\mathbf{s} - \mathbf{s}_k)] + \underbrace{\nabla^2[\xi]}_{=0}, \quad \mathbf{s} \in D \ .$$

The DOK-predicted surface will therefore satisfy Poisson's equation if the point semivariogram function $\gamma_Z(\mathbf{h})$ used for prediction is the free-space Green's function of the differential operator (Kitanidis 1999), that is, if $\gamma_Z(\mathbf{s} - \mathbf{s}_k) = G(\mathbf{s}_k - \mathbf{s}_l)$, and $\gamma_Z(\mathbf{s}_k - \mathbf{s}_l) = G(\mathbf{s} - \mathbf{s}_k)$, in which case

$$\nabla^2[z^*(\mathbf{s})] = \sum_{k=1}^{K} \eta(\mathbf{s}_k)\delta(\mathbf{s} - \mathbf{s}_k), \quad \mathbf{s} \in D \,,$$

which entails that the resulting DOK surface is harmonic at any location $\mathbf{s} \in D$ not coinciding with a sample location $\mathbf{s}_k$, and exhibits K singularities at the K sample locations $\{\mathbf{s}_k, k = 1,\ldots,K\}$ where it attains the corresponding sample data values. Note that in the 2-D and 3-D cases, the scale (variance) terms $\frac{1}{2\pi}$ and $\frac{1}{4\pi}$ appearing in the corresponding free-space Green's functions do not play any role for prediction, because they cancel out in the DOK system.

Indeed, it is well known (e.g., Chilès and Delfiner 1999) that Ordinary Kriging in 1-D with a linear semivariogram yields a perfect screening effect: only the two neighboring samples located on either side of the prediction location receive non-zero weights. In this case, the resulting interpolated profile is piecewise linear, that is, harmonic at all unsampled locations. In addition, it can be shown through the application of the divergence theorem that the above solution satisfies the constant flux condition at the boundary of the domain (Kitanidis 1999).

Finite domains with different shapes and various types of BCs could be also handled analytically by modifying the free-space Green's functions to become nonstationary, via conformal mapping (Greenberg 1971). Such analytical modifications could be used to construct non-stationary and permissible (conditionally negative definite) semivariogram models at the point support level. Alternatively, BCs could be handled as values or known local derivatives within the geostatistical framework (Chilès and Delfiner 1999). In addition, one can formally associate other types of smoothness criteria with Green's functions and covariance kernels (Hilgers 1976; Matheron 1981; Wahba 1990; Smola, Schölkopf, and Müller 1998; Evgeniou, Pontil, and Poggio 2000), but the treatment of these issues is beyond the scope of this paper.

In summary, I have shown that Tobler's smoothness criterion implies the following point support semivariogram models: linear in 1-D, logarithmic in 2-D, and negative inverse distance in 3-D. It should be stressed that these particular semivariogram models are parameter-free (apart from a multiplicative constant that cancels out in the Kriging system). In other words, these semivariogram models need not be fitted to any observations, for example, there is no range parameter that has to be adjusted according to the areal data. This explains the fact that Tobler's approach has been used without paying much attention to inference: in this case, inference is tantamount to the adoption of the method, that is, the assumption that the spatial distribution of point support values is governed by Poisson's PDE. It is also interesting to note that both the linear and logarithmic (or de Wijsian in the geostatistical jargon) semivariogram models imply self-similarity and can be regarded as fractal models (Chilès and Delfiner 1999). In particular, in 1-D, a linear semivariogram corresponds to the process of Brownian motion (or Wiener-Lévy process).

Going back to the case of area-to-point interpolation, the DOK prediction $z^*(\mathbf{s})$ is given by a modified (but equivalent) version of equation (10), written in terms of semivariograms:

$$z^*(\mathbf{s}) = [\boldsymbol{\omega}' \ \psi]\begin{bmatrix} \bar{\boldsymbol{\gamma}}_\mathbf{s} \\ 1 \end{bmatrix} = \sum_{k=1}^{K} \omega(v_k)\bar{\gamma}_Z(\mathbf{s}, v_k) + \psi \tag{27}$$

where $\bar{\boldsymbol{\gamma}}_\mathbf{s} = \bar{\gamma}_Z(\mathbf{s}, v_k), k = 1,\ldots,K]'$ denotes the $(K \times 1)$ location-specific vector of area-to-point regularized semivariogram values between the prediction location $\mathbf{s}$ and any areal support v_k.

The corresponding DOK system of equations is also given by a modified (but equivalent) version of equation (11), written in terms of regularized semivariograms:

$$
\begin{bmatrix} \bar{\Gamma} & \mathbf{1}_K \\ \mathbf{1}'_K & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{\omega} \\ \psi \end{bmatrix} = \begin{bmatrix} \bar{\mathbf{z}} \\ 0 \end{bmatrix} \quad \text{or} \quad \begin{cases} \sum_{l=1}^{K} \omega(v_l) \bar{\gamma}_Z(v_k, v_l) + \psi = \bar{z}(v_k), \ \ k = 1, \dots, K \\ \sum_{l=1}^{K} \omega(v_l) = 0 \end{cases} \tag{28}
$$

where $\bar{\Gamma} = \bar{\gamma}_Z(v_k, v_l), k = 1, \dots, K, l = 1, \dots, K]$ denotes the $(K \times K)$ matrix of area-to-area semivariogram values between any two supports v_k and v_l.

The DOK point surface computed via equations (27) and (28) is coherent, that is, reproduces the areal average data, since the above equations are equivalent to equations (10) and (11), and harmonic only outside the support of each areal datum, that is, the interpolated surface does not satisfy Poisson's PDE at all points within the K areal data supports (as opposed to the point-to-point prediction case where singularities occur only at the K point sample locations). Note, however, that the interpolated surface is the closest possible approximation to a harmonic surface, under the constraint of reproduction of the available areal-average data.

In summary, I have shown that Tobler's pycnophylactic interpolation with Laplacian smoothing, on a quasi-infinite domain without non-negativity constraints, corresponds to a particular solution of the geostatistical framework, whereby the semivariogram model adopted at the point support level is identified to the free-space Green's function of the Laplace differential operator.

It should be noted here that Tobler's original approach calls for a regular grid of prediction locations in order to apply the finite difference approximation, whereas in the geostatistical framework the prediction locations can be arbitrarily specified. Similarly, the recently developed variant of pycnophylactic interpolation using triangular irregular (TIN) surface representation (Rase 2001) alleviates the regular grid restriction of Tobler's method.

On the computational side, in the case of P prediction locations and K areal data, Tobler's approach calls for the inversion of a $(P + K) \times (P + K)$ matrix, whereas the geostatistical framework calls for the inversion of a much smaller $K \times K$ matrix, and can thus handle much larger problems without resorting to iterative system solvers. When the number of areal data is large ($K > 5000$), the geostatistical solution would call for the inversion of a large $K \times K$ matrix, which could be computationally prohibitive.

In such situations, one could perform DOK with moving local neighborhoods, thus restricting the number of nearby areal data considered for prediction from K to $K(\mathbf{s})$, with $K(\mathbf{s}) << K$. This would avoid the inversion of a single large matrix $\bar{\mathbf{C}}$ to compute the DOK weights; see equation (8). The extent of such neighborhoods is typically identified to the range of the covariance model $C_Z(\mathbf{h})$. In DOK with moving local neighborhoods, however, the area-to-area covariance matrix $\bar{\mathbf{C}}$ is location specific, hence one would have to solve P local (but small) DOK systems, one per prediction location $\mathbf{s}$. The coherence of the resulting area-to-point predictions is still guaranteed, if one uses the same $K(\mathbf{s})$ areal data to predict all the unknown values $\{z(\mathbf{s}), \mathbf{s} \in v_k\}$ at all locations $\mathbf{s}$ within any support v_k. The above local DOK shortcut is tantamount to an adaptive kernel smoothing procedure, whereby the covariance kernel remains fixed (including its functional form and parameters), and only the number $K(\mathbf{s})$ of nearby areal data changes from one prediction location $\mathbf{s}$ to another. Note, however, that the equivalence between Kriging, splines, radial basis functions, and pycnophylactic interpolation applies to the global (not the local) prediction scheme.

6. INFERENCE OF A POINT SUPPORT COVARIANCE MODEL

Since it has been demonstrated that the common thread of most area-to-point interpolation methods is the point covariance model $C_Z(\mathbf{h})$, and this model is a free parameter in the proposed geostatistical framework; it behooves us to address the issue of inferring such a point covariance model from the available areal data. As in the point-to-point Kriging case, the particular point covariance model adopted will have a major impact on the standard error of the resulting predictions, not so much on the actual predictions themselves (Cressie 1993; Chilès and Delfiner 1999).

Recall from equation (5) that the regularized covariance is computed from the original point covariance $C_Z(\mathbf{h})$ via an arithmetic averaging (regularization) procedure. Essentially, regularization of a point covariance model $C_Z(\mathbf{h})$ by a support v_k is a special case of double convolution of function $C_Z(\mathbf{h})$ with a weight function whose elements are all equal to $\frac{1}{|v_k|}$ (Journel and Huijbregts 1978; Vanmarcke 1983; Chilès and Delfiner 1999). Consequently, the inference of the point covariance from the regularized one calls for the reverse procedure of deconvolution or deregularization. More formally, this inference problem constitutes an underdetermined inverse problem (Menke 1989; Bertero and Boccacci 1998; Vogel 2002).

In the case of equal data supports, $|v_k|$=constant, $\forall k$, the areal covariance is stationary and can be readily computed from the available areal data. In the general case of unequal data supports, however, there is an important complication: the regularized covariance $\overline{C}_Z(v_k, v_l)$ between any two areal data supports v_k and v_l is nonstationary, and cannot be estimated from a single areal data realization. In other words, one cannot simply assimilate the areal values to their respective support centroids and then compute a sample covariance $\overline{C}_Z(v_k - v_l)$; that covariance calculation presupposes stationarity. If one ignores this important complication, and computes a pseudo-stationary covariance $\overline{C}_Z(v_k - v_l)$ using support centroids, then there are various solutions to the inference problem; the term pseudo-stationary is used here to distinguish between the actual nonstationary areal covariance $\overline{C}_Z(v_k, v_l)$ and the covariance $\overline{C}_Z(v_k - v_l)$ computed from, say, support centroids.

One possible solution is to compute the pseudo-stationary covariance values $\overline{C}_Z(v_k - v_l)$, possibly fit a functional model to them, assume a constant support v_k, and perform deconvolution (Journel and Huijbregts 1978; Mockus 1998; Atkinson and Martin 1999). That constant support can be (subjectively) chosen as the most often repeated one; alternative definitions of an "average" support are also possible. Inference typically proceeds in an iterative fashion, whereby a parametric point covariance function $C_Z(\mathbf{h};\boldsymbol{\theta})$ is first postulated, and all regularized area-to-area covariance values $\overline{C}_Z(v_k, v_l;\boldsymbol{\theta})$ are calculated for all possible combinations of support pairs v_k and v_l. Here $\boldsymbol{\theta}$ denotes a vector of parameter values (e.g., range and sill) specifying the point covariance model. The resulting regularized area-to-area covariance values are then compared to the pre-computed, pseudo-stationary covariance values $\overline{C}_Z(v_k - v_l)$, and a summary measure of their discrepancy is recorded. A new set of parameters for the point covariance model is then proposed, and the entire procedure is repeated until the above discrepancy is smaller than a predefined tolerance value. This procedure can be also extended to account for uncertainty in the parameter vector $\boldsymbol{\theta}$ using a hierarchical Bayesian framework (Gelfand, Zhu, and Carlin 2001).

A more general approach that does not rely on the computation of a pseudo-stationary covariance $\overline{C}_Z(v_k - v_l)$ would call for a prior classification of the shape and size (and possibly orientation) of the available data supports. A genuine stationary covariance could then be computed from each set of supports with similar size and shape, and its deconvolution could be performed. Since several areal support covariances would be calculated, based on the prior shape and size classification, there could be more than one deconvolved covariance value for the same lag $\mathbf{h}$. The resulting point

covariance values should then be combined into a single, final value of the point support covariance $C_Z(\mathbf{h})$ for lag $\mathbf{h}$. Last, a parametric covariance model should be fitted to these combined covariance values to ensure positive definiteness.

Alternatively, if one assumes that the point support values are multivariate Gaussian, then the method of maximum likelihood (ML) could be used to infer a parametric point covariance model. In the multivariate Gaussian case, the likelihood $p(\mathbf{z}|\boldsymbol{\theta})$ of the point support values is written as (Mardia and Marshall 1984; Cressie 1993; Pardo-Igúzquiza 1998):

$$p(\mathbf{z}|\boldsymbol{\theta}) = (2\pi)^{-\frac{K}{2}}|\mathbf{C_\theta}|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(\mathbf{z} - \mathbf{m}_Z)'\mathbf{C_\theta}^{-1}(\mathbf{z} - \mathbf{m}_Z)\right)$$

where $|\mathbf{C_\theta}|$ denotes the determinant of the point support covariance matrix (indexed by the parameter vector $\boldsymbol{\theta}$), and $\mathbf{m}_Z$ denotes a $(K \times 1)$ vector of mean values for the point support data.

The estimated parameter vector $\boldsymbol{\theta}^*$ is the one that maximizes the probability of observing the particular combination $\mathbf{z}$ of K point values. The ML procedure involves iterative minimization of the negative loglikelihood (which is nonlinear in the parameter $\boldsymbol{\theta}$), and thus becomes computationally very expensive for a large number $(K > 1000)$ of areal data.

When only areal data are available, the above equation could be modified as

$$p(\bar{\mathbf{z}}|\boldsymbol{\theta}) = (2\pi)^{-\frac{K}{2}}|\overline{\mathbf{C}}_\theta|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(\bar{\mathbf{z}} - \overline{\mathbf{m}}_Z)'\overline{\mathbf{C}}_\theta^{-1}(\bar{\mathbf{z}} - \overline{\mathbf{m}}_Z)\right)$$

where $|\overline{\mathbf{C}}_\theta|$ denotes the determinant of the areal data-to-data covariance matrix (still indexed by the parameter vector $\boldsymbol{\theta}$ of the point covariance model), and $\overline{\mathbf{m}}_Z$ denotes a $(K \times 1)$ vector of mean values for the areal data. The only difference with the point support case, is that now the likelihood is defined with respect to the areal data, which entails that the point support data vector $\mathbf{z}$ is replaced by $\bar{\mathbf{z}}$, the point mean vector $\mathbf{m}_Z$ is replaced by $\overline{\mathbf{m}}_Z$, and the point-to-point covariance matrix $\mathbf{C_\theta}$ is replaced by $\overline{\mathbf{C}}_\theta$.

The very important advantage of the ML procedure over the previously described approaches is that no comparison is made between a proposed set of area-to-area covariance values $\overline{C}_Z(v_k, v_l; \boldsymbol{\theta})$ and some pseudo-stationary covariance values $\overline{C}_Z(v_k - v_l)$. The ML approach essentially circumvents the problem of calculating the areal data covariance (which is nonstationary), but it relies heavily on the multivariate Gaussian assumption.

7. DISCUSSION

The geostatistical framework for area-to-point prediction presented in this paper constitutes a general unifying framework for interpolating point values from available areal data. It has been shown that this framework explicitly accounts for the different data supports and their nonstationary spatial correlation, and it is coherent, that is, yields point predictions that satisfy any linear areal-average or areal-total datum constraint. In addition, the proposed geostatistical framework allows an explicit investigation of the links between various existing methods for area-to-point interpolation. All such methods make specific implicit or explicit assumptions regarding the covariance $C_Z(\mathbf{h})$ or semivariogram $\gamma_Z(\mathbf{h})$ model of the point support values. In the choro-

pleth map case, that point covariance model is assumed to be a pure nugget effect (complete absence of spatial correlation), whereas in the case of kernel smoothing that covariance model is none other than the kernel function adopted for interpolation. In the case of Tobler's pycnophylactic interpolation with Laplacian smoothing, the point semivariogram model is linked to the smoothness criterion (or the shape of the resulting harmonic surface) implied by Laplace's equation, and it is linear in 1-D and logarithmic in 2-D.

The various existing methods for area-to-point interpolation differ in their degree of flexibility with respect to the point covariance or semivariogram model: the choropleth map and Tobler's pycnophylactic interpolation specify implicitly (fix) that covariance model, thus excluding any possible user control over that model. On the one hand, the kernel smoothing method allows for a user-specified kernel or equivalently for a user-specified covariance model, but its current applications either do not account for the difference in data supports or do not yield coherent predictions. The proposed geostatistical framework allows for a wide spectrum of parametric point covariance models by essentially leaving their functional form and their parameters unspecified, while always ensuring the coherence of predictions.

The most critical requirement for the geostatistical framework is the availability of the point support covariance model. This requirement, however, is not a drawback, but rather an advantage because it explicates the subjective decisions made at the point support level by all existing methods for area-to-point interpolation. The very attempt to infer a point covariance model, even in the rather futile case of unequal areal data supports, is far more data-consistent than the arbitrary adoption of a kernel or a smoothness criterion expressed in terms of partial derivatives. Of course, if there exists physical justification dictating that the variable under study satisfies a particular PDE, then this extremely valuable prior information should be incorporated in any prediction endeavor. Since such information can also be expressed in terms of a point covariance model, the geostatistical framework can easily cope with this situation, and thus offers a flexible and more general alternative to existing area-to-point interpolation methods. The only restriction on the particular point covariance model $C_Z(\mathbf{h})$ adopted for area-to-point interpolation is that of positive definiteness (or conditional negative definiteness in the case of the semivariogram $\gamma_Z(\mathbf{h})$). In other words, any arbitrary weighted linear combination of covariance values should yield a non-negative variance. This condition is linked to the requirement that the covariance be a permissible "distance" measure, and places restrictions on the types of functions that can serve as point covariance models. For example, the tent function (also known as triangular covariance model) suggested by Tobler (1999) is only positive definite in 1-D, and hence cannot be used for interpolation in two or three dimensions (Chilès and Delfiner 1999). In short, one cannot adopt any arbitrary-shaped kernel as covariance function without first ensuring that it is indeed positive definite.

One of the most striking results of this paper is the derivation of the choropleth map solution as a limiting case of the geostatistical framework under the rather unrealistic assumption of spatially uncorrelated point support values. The choropleth map is sometimes advocated as a valuable representation of, say, population density surfaces, because such surfaces have indeed discontinuities that can coincide with support boundaries. It was shown, however, that the discontinuities induced in the predicted point surface via the choropleth map are a result of the overwhelming contribution of the nugget effect, that is, they are surface characteristics stemming from second-order effects. Surface discontinuities, however, are first-order effects, and they should always be treated and generated as such. In other words, attribute values that change abruptly between valid (not artificial) zone boundaries should be modeled as the consequence of a nonstationary mean component, and not as the artifact of the unrealistic assumption of spatially uncorrelated point values. One should first

and foremost strive to assess the fidelity of zone boundaries and then incorporate them as first-order effects in area-to-point interpolation.

Future research endeavors should be directed towards: (i) handling areal data defined as arbitrary (linear or nonlinear) combinations of point values; (ii) developing and testing methods for inferring a point support covariance model from available areal data; (iii) incorporating observed data of the same variable at the point support level, such as boundary conditions or sets of known-valued locations comprising geographical features; (iv) incorporating additional covariates at the areal or point support level; (v) accounting for error-prone areal data that do not warrant exact reproduction; (vi) generating alternative synthetic realizations of point support values using stochastic simulation subject to the above data constraints; and (vii) accounting for non-Gaussian data. For ongoing research in items (i), (iii), and (vi), the reader is referred to Kyriakidis and Yoo (2004).

LITERATURE CITED

Anderson, T. W. (1958). *An Introduction to Multivariate Statistical Analysis*. New York: John Wiley & Sons.

Atkinson, P. M., and D. Martin. (1999). "Investigating the Effect of Support Size on Population Surface Models." *Geographical & Environmental Modelling* 3, 101–19.

Atkinson, P. M., and N. J. Tate. (2000). "Spatial Scale Problems and Geostatistical Solutions: A Review." *Professional Geographer* 52, 607–23.

Bertero, M., and P. Boccacci. (1998). *Introduction to Inverse Problems in Imaging*. Bristol: Institute of Physics Publishing.

Bracken, I., and D. Martin. (1989). "The Generation of Spatial Population Distributions from Census Centroid Data." *Environment and Planning A* 21, 537–43.

Brillinger, D. R. (1990). "Spatial-Temporal Modeling of Spatially Aggregate Birth Data." *Survey Methodology* 16, 255–69.

Brillinger, D. R. (1994). "Examples of Scientific Problems and Data Analyses in Demography, Neurophysiology, and Seismology." *Journal of Computation and Graphical Statistics* 3, 1–22.

Carr, J. R., and J. A. Palmer. (1993). "Revisiting the Accurate Calculation of Block-Sample Covariances Using Gauss Quadrature." *Mathematical Geology* 25, 507–24.

Chilès, J. P., and P. Delfiner. (1999). *Geostatistics: Modeling Spatial Uncertainty*. New York: John Wiley & Sons.

Cressie, N. A. C. (1993). *Statistics for Spatial Data.* Revised edition. New York: John Wiley & Sons.

Davis, M. M. W., and C. Grivet. (1984). "Kriging in a Global Neighborhood." *Mathematical Geology* 16, 249–65.

Dubrule, O. (1983). "Two Methods with Different Objectives: Splines and Kriging." *Mathematical Geology* 15, 245–57.

Evgeniou, T., M. Pontil, and T. Poggio. (2000). "Regularization Networks and Support Vector Machines." *Advances in Computational Mathematics* 13, 1–50.

Gelfand, A. E., L. Zhu, and B. P. Carlin. (2001). "On the Change of Support Problem for Spatio-Temporal Data." *Biostatistics* 2, 31–45.

Goovaerts, P. (1997). *Geostatistics for Natural Resources Evaluation.* New York: Oxford University Press.

Gotway, C. A., and L. J. Young. (2002). "Combining Incompatible Spatial Data." *Journal of the American Statistical Association* 97, 632–48.

Greenberg, M. D. (1971). *Application of Green's Functions in Science and Engineering.* Englewood Cliffs, N.J.: Prentice-Hall.

Haining, R. (2003). *Spatial Data Analysis: Theory and Practice.* Cambridge: Cambridge University Press.

Hilgers, J. W. (1976). "On the Equivalence of Regularization and Certain Reproducing Kernel Hilbert Space Approaches for Solving First Kind Problems." *SIAM Journal of Numerical Analysis* 12, 172–84.

Huang, H.-C., N. Cressie, and J. Garbosek. (2002). "Fast, Resolution-Consistent Spatial Prediction of Global Processes from Satellite Data." *Journal of Computational and Graphical Statistics* 11, 63–88.

Isaaks, E., and R. Srivastava. (1989). *An Introduction to Applied Geostatistics.* New York: Oxford University Press.

Journel, A. G. (1999). "Conditioning Geostatistical Operations to Non-linear Volume Averages." *Mathematical Geology* 31, 931–53.

Journel, A. G., and C. J. Huijbregts. (1978). *Mining Geostatistics*. New York: Academic Press.

Kelsall, J., and J. Wakefield. (2002). "Modeling Spatial Variation in Disease Risk: A Geostatistical Approach." *Journal of the American Statistical Association* 97, 692–701.

Kitanidis, P. K. (1999). "Generalized Covariance Functions Associated with the Laplace Equation and Their Use in Interpolation and Inverse Problems." *Water Resources Research* 35, 1361–67.

Kyriakidis, P. C., and E.-H. Yoo. (2004). "Geostatistical Prediction and Simulation of Point Values from Areal Data." *Geographical Analysis* (in press).

Lam, N. S.-N. (1983). "Spatial Interpolation Methods: A Review." *The American Cartographer* 10, 129–49.

Mardia, K. V., and R. J. Marshall. (1984). "Maximum Likelihood Estimation of Models for Residual Covariance in Spatial Regression." *Biometrika* 71, 135–46.

Martin, D. (1996). "An Assessment of Surface and Zonal Models of Population." *International Journal of Geographical Information Systems* 10, 973–89.

Matheron, G. (1971). *The Theory of Regionalized Variables and its Applications*. Paris: Ecole Nat. Sup. des Mines.

Matheron, G. (1981). "Splines and Kriging: Their Formal Equivalence." In *Down-to-Earth Statistics: Solutions Looking for Geological Problems*, edited by D. Merriam, 77–95. Syracuse, N.Y.: Syracuse University Geology Contributions.

Menke, W. (1989). *Geophysical Data Analysis: Discrete Inverse Theory*. Revised Edition. San Diego: Academic Press.

Mockus, A. (1998). "Estimating Dependencies from Spatial Averages." *Journal of Computational and Graphical Statistics* 7, 501–13.

Myers, D. E. (1987). "Interpolation with Positive Definite Functions." *Science de la Terre* 28, 251–65.

Pardo-Igúzquiza, E. (1998). "Maximum Likelihood Estimation of Spatial Covariance Parameters." *Mathematical Geology* 30, 95–107.

Rase, W.-D. (2001). "Volume-Preserving Interpolation of a Smooth Surface from Polygon-Related Data." *Journal of Geographical Systems* 3, 199–213.

Smola, A. J., B. Schölkopf, and K.-R. Müller. (1998). "The Connection Between Regularization Operators and Support Vector Kernels." *Neural Networks* 11, 637–49.

Tobler, W. (1999). "Linear Pycnophylactic Reallocation—Comment on a Paper by D. Martin." *International Journal of Geographical Information Systems* 13, 85–90.

Tobler, W. R. (1979). "Smooth Pycnophylactic Interpolation for Geographical Regions." *Journal of the American Statistical Association* 74, 519–36.

Vanmarcke, E. (1983). *Random Fields: Analysis and Synthesis.* Cambridge, MA: The M.I.T. Press.

Vogel, C. R. (2002). *Computational Methods for Inverse Problems.* Philadelphia: Society for Industrial and Applied Mathematics.

Wackernagel, H. (1995). *Multivariate Geostatistics.* Berlin: Springer-Verlag.

Wahba, G. (1990). *Spline Models for Observational Data.* Philadelphia: Society for Industrial and Applied Mathematics.