John Gabrosek
and Noel Cressie

The Effect on Attribute Prediction of Location Uncertainty in Spatial Data

A datum is considered spatial if it contains location information. Typically, there is also attribute information, whose distribution depends on its location. Thus, error in location information can lead to error in attribute information, which is reflected ultimately in the inference drawn from the data. We propose a statistical model for incorporating location error into spatial data analysis. We investigate the effect of location error on the spatial lag, the covariance function, and optimal spatial linear prediction (that is, kriging). We show that the form of kriging after adjusting for location error is the same as that of kriging without adjusting for location error. However, location error changes entries in the matrix of explanatory variables, the matrix of covariances between the sample sites, and the vector of covariances between the sample sites and the prediction location. We investigate, through simulation, the effect that varying trend, measurement error, location error, range of spatial dependence, sample size, and prediction location have on kriging after and without adjusting for location error. When the location error is large, kriging after adjusting for location error performs markedly better than kriging without adjusting for location error, in terms of both the prediction bias and the mean squared prediction error.

Data are considered spatial if they contain location information. Typically, there is also attribute information available. The distribution of the attribute varies from location to location. Attribute information consists of the measured response (or responses), which can be either discrete (for example, counts of animal populations) or continuous (for example, soil pH). With the advent of optimal spatial linear prediction (that is, kriging), the analysis of spatially dependent data has progressed rapidly in the past forty years. Cressie (1990) lists three components necessary for the development of ordinary kriging: (i) use of covariances (or variograms) to weight observations,

The authors thank the referees for their very helpful comments. This research was supported in part by the United States Environmental Protection Agency under Assistance Agreement Number CR 822919-01-1 (at Iowa State University) and R 827257-01-0 (at The Ohio State University).

John Gabrosek is assistant professor of statistics at Grand Valley State University. E-mail: gabrosej@gvsu.edu *Noel Cressie, Professor of Statistics and Director, Program in Spatial Statistics and Environmental Sciences, The Ohio State University. E-mail:* ncressie@stat.ohio-state.edu

Geographical Analysis, Vol. 34, No. 3 (July 2002) The Ohio State University
Submitted: 04/17/01. Revised version accepted: 01/04/02

(ii) use of the best linear unbiased estimator (BLUE) for an unknown constant mean μ, and (iii) use of spatial locations in the determination of covariances (or variograms).

Since locations are used to determine covariances, and covariances are used to weight observations, uncertainty in locations leads to a further component of variability in kriging predictions. Almost without fail, texts and articles about kriging assume that locations are known without error. Here, we explicitly account for location uncertainty and quantify its effect on statistical inference. We show how to incorporate location error by adjusting the component parts of the kriging equations and then proceeding as if the spatial process were sampled without location error.

Suppose that we have one (spatially incomplete) observation on the random process $\{Z(\mathbf{s}): \mathbf{s} \in D \subset \Re^d\}$. Without further assumptions about $Z(\cdot)$, data $\{Z(\mathbf{s}_1),\ldots,Z(\mathbf{s}_n)\}$ represent a sample from a single realization and statistical inference is not possible. In subsequent sections of this article, we shall assume that $Z(\cdot)$ has a linear mean and a stationary covariance function:

$$\begin{aligned}
&(i) && \mathrm{var}(Z(\mathbf{s})) < \infty; && \text{for all } \mathbf{s} \in D\,,\\
&(ii) && \mu(\mathbf{s}) = E(Z(\mathbf{s})) = \mathbf{f}(\mathbf{s})'\boldsymbol{\beta}; && \text{for all } \mathbf{s} \in D\,,\\
&(iii) && \mathrm{cov}(Z(\mathbf{s}_1), Z(\mathbf{s}_2)) = C(\mathbf{s}_2 - \mathbf{s}_1); && \text{for all } \mathbf{s}_1, \mathbf{s}_2 \in D\,.
\end{aligned} \tag{1}$$

Condition (i) ensures that the first two moments of the process are defined. Condition (ii) assumes that the mean function is linear in the $(q + 1) \times 1$ vector of unknown parameters $\boldsymbol{\beta}$, where the explanatory variables $\mathbf{f}(\mathbf{s})$ are known functions of $\mathbf{s} \in D$. Condition (iii) states that the covariance between locations $\mathbf{s}_1$ and $\mathbf{s}_2$ depends only on the spatial lag $\mathbf{h} = \mathbf{s}_2 - \mathbf{s}_1$. The lag plays a central role in geostatistics; however, little attention has been paid to the effect of location uncertainty (that is, error in $\mathbf{s}$ and thus in $\mathbf{h}$) on spatial methods such as kriging (see, for example, Clark, 1979, for a brief discussion).

Models for Location Error in Spatial Data

Geographers have attempted to account for location uncertainty. When creating maps, location is of paramount interest. At the other extreme, many statistical analyses take no account of location, modeling data as if it were statistically independent. Geostatistics is between these extremes. Geostatisticians use location to model trend and correlation between attribute values over a geographic region; however, they ignore uncertainty in locations.

The advent and then ubiquity of geographic information systems (GIS) has led to an explosion of information available from spatial databases. The easy storage and quick retrieval possible within a GIS requires concomitant development of spatial statistical methodology. Incorporation of attribute-error analysis is often handled through geostatistics, but there is an urgent need for statistical research and software developments to deal with both location error and attribute error (for example, Griffith, Haining, and Arbia 1999).

Geographers and users of raster-based GIS often model the effects of location error in spatial data by assuming that the attribute value is discrete (often a gray-scale value) and that the spatial domain is a fixed grid of pixels. A commonly used model for the Bayesian restoration of images, attributed to Geman and Geman (1984), has been adapted by researchers working with GIS to investigate how errors in source maps propagate through a GIS to output maps (for example, Goodchild 1989; Haining and Arbia 1993; and Arbia, Griffith, and Haining 1998). Output maps result from overlay operations that combine two or more source maps at potentially different scales of spatial resolution.

Following Arbia, Griffith, and Haining (1998), let T be the *true* source map (that is, ground truth). Suppose that T is an $n \times m$ matrix of gray-scale values. The (i,j)th element of T, the gray-scale value for the (i,j)th pixel, is denoted by T_{ij}; $i = 1,2,\ldots,n, j = 1,2,\ldots,m$. Similarly, Z_{ij} is the gray-scale value for the (i,j)th element of the observable map Z. Let (g,h) be a shift in the location of pixel (i,j), such that its location is $(i+g,j+h)$, where $g,h = 0, \pm 1, \pm 2,\ldots$ Consider the *corruption model*, which is defined as

$$Z_{ij} = \sum_g \sum_h w_{i,j}(g,h) T_{i+g,j+h} + u_{i,j}\,, \tag{2}$$

where

$$\sum_g \sum_h w_{i,j}(g,h) = 1$$

and

$$\{w_{i,j}(g,h) \geq 0 : i = 1,\ldots,n;\ j = 1,\ldots,m;\ g,h = 0,\pm 1,\pm 2,\ldots\}\,. \tag{3}$$

The probability that we measure pixel $(i+g,j+h)$ instead of pixel (i,j) is $w_{i,j}(g,h)$. The attribute error associated with pixel (i,j) is $u_{i,j}$. If we assume that the location error is independent of (i,j), then $w_{i,j}(g,h)$ is independent of (i,j) and (3) can be replaced by the $n \times n$ *blurring matrix* W. In that case, (2) becomes

$$Z = WT + U\,.$$

To proceed with an error-propagation analysis, we specify W and a distribution for U. Arbia , Griffith, and Haining (1998) use Monte Carlo simulation to assess the effect of location error on output maps. They generate realizations $\{T_k : k = 1,\ldots,K\}$ of the true map. With W fixed and a known distribution for U, they obtain maps $\{Z_k : k = 1,\ldots,K\}$ where pixel (i,j) has observed attribute values $Z_{ij1}, Z_{ij2},\ldots, Z_{ijK}$. They assess the effect of error (both attribute and location) by comparing the observed attribute value Z_{ijk} to the true value T_{ijk}; $k = 1,\ldots,K$. Attempts to extend error-propagation analyses to continuous source maps have included attribute error but not location error (Heuvelink 1996).

Users of GIS are becoming aware that determining the suitability of a spatial database for a spatial analysis depends on the effect of location error on the analysis being done. Lewis and Hutchinson (2000) conclude that location error has less effect than the structure of the underlying spatial random process. Here, we quantify the effect of location error on kriging.

Morphometrics, which studies the geometrical form and arrangement of organisms, is also concerned with uncertainty of locations. Morphological data usually include information on landmarks, whose relative positions correspond from organism to organism within a species (Bookstein 1986). For instance, the tip of the nose and the point of the chin are often used as landmarks in Homo sapiens. While morphometrics does not provide an immediate solution to the problem of location error in spatial prediction, it does provide models to deal with location error. For example, Bookstein (1986) describes a null model for analyzing configurations: K landmarks, $L_1,L_2,\ldots,L_K$, are distributed about their centroids, $W_1,W_2,\ldots,W_K$, according to a circular normal model (that is, a bivariate normal distribution with covariance matrix

equal to a multiple of the identity matrix). Independent normally distributed location errors is one of two models used throughout this article.

Regression rarely incorporates either errors in locations or errors in explanatory variables. Measurement-error models, often called *errors-in-variables* models, incorporate uncertainty in explanatory (x) variables (for example, Fuller 1987). Following Fazekas et al. (1999), consider the model

$$z_i = g(\varsigma_i,\beta) + \delta_i$$
$$x_i = \varsigma_i + \varepsilon_i \,, \tag{4}$$

where z_i is an observable (possibly vector-valued) response, $g(\varsigma_i,\beta)$ is the mean function measured at the (unobservable) nonrandom, error-free explanatory variables ς_i, β are mean-function parameters, $\{\delta_i\}$ are zero-mean independent measurement errors, $\mathbf{x}_i$ are (observable) values of the explanatory variables, and $\{\varepsilon_i\}$ are zero-mean independent errors on the explanatory variables. Fazekas and Kukush (2000) prove consistency and asymptotic normality results for the usual least-squares estimator of β. Though ς_i is not observable, they assume that possible values for ς_i are identifiable. The least-squares estimator results are proved for infill asymptotics, which choose more and more points within a finite spatial domain, following Lahiri et al. (1999). Now, if we interpret ς_i as an error-free spatial location and $\mathbf{x}_i$ as its observed location, then the model (4) can be interpreted as one that includes both attribute error and location error.

Sources of Location Error in Spatial Data

Consider a collection of n intended sample sites $S \equiv \{\mathbf{s}_1,\mathbf{s}_2,\ldots,\mathbf{s}_n\}$ chosen a priori. In Figure 1, the open squares represent S. Suppose that an attempt is made to go to the sites in S using a positioning instrument that is not infinitely accurate, and hence the realized sites, $R \equiv \{\mathbf{r}_1,\mathbf{r}_2,\ldots,\mathbf{r}_n\}$, are displaced from those in S. In Figure 1, the solid squares represent R.

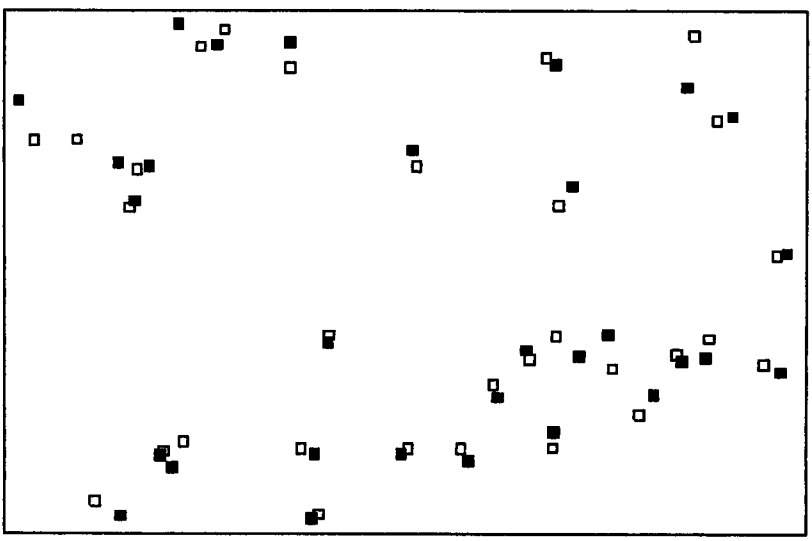

FIG. 1. The Intended Sampling Sites, S, Are Represented by the Open Squares. The Realized Sampling Sites, R, Are Represented by the Solid Squares.

Location error enters through the inability to go to (and sample from) the predetermined sites S. This scenario, which we call the *design model*, embodies the most common source of location error in geostatistics [see, for example, Krumbein and Graybill 1965, ch. 7)], where core samples in mining are taken). In design-model sampling, an appropriate statistical model relating the realized sites to the intended sites is

$$R \mid S \sim (S,\Sigma_p) ; \tag{5}$$

that is, the positioning instrument is assumed to be unbiased and the realized sites R are distributed about the intended sites S according to a probability distribution with mean S and covariance matrix Σ_p. Note that here S is known but R is not.

Although the design model is assumed throughout this article, we describe an alternative model for location error. Location error can occur when the resource to be sampled determines the sample locations. Although the resource may have well-defined locations, the geographic coordinates are not known a priori.

Consider sampling redwood trees in a forest, the locations of which are not known a priori. Each tree's location is recorded using a positioning instrument. As above, the instrument is not infinitely accurate. Let $B \equiv \{\mathbf{b}_1,\mathbf{b}_2,\ldots,\mathbf{b}_n\}$ be the reported locations returned by the positioning instrument. Let $A \equiv \{\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_n\}$ be the actual locations, which are fixed but unknown. In this scenario, which we call the *resource model*, neither A nor B are predetermined. In resource-model sampling, an appropriate statistical model relating the reported tree locations to the actual tree locations is

$$B \mid A \sim (A,\Gamma_p) ; \tag{6}$$

that is, the positioning instrument is assumed to be unbiased and the reported sites are distributed about the actual sites according to a probability distribution with mean A and covariance matrix Γ_p. The location error arises from the inability to report a resource's location infinitely accurately. In contrast to the design model (5), here B is observed but A is unknown.

The design model and the resource model are modifications of two scenarios incorporating error in the response and explanatory variables for the regression of x and y, as described by Berkson (1950). Consider the model

$$\begin{aligned} x &= u + d , \\ y &= v + e ; \end{aligned} \tag{7}$$

where u and v are the "true" values of x and y. If x is an intended location and u is the realized location, then we have design-model sampling. If u is the actual location of a resource and x is the reported location, then we have resource-model sampling. Berkson (1950) calls the two scenarios a controlled observation and an uncontrolled observation.

Location error may be knowingly introduced to simplify a statistical analysis of spatial data. Suppose a datum occurs at a known $\mathbf{r}_i$, but for the purposes of spatial analysis is assumed to have occurred at $\mathbf{s}_i$; $i = 1,\ldots,n$. Griffith, Haining, and Arbia (1999) describe this type of location error in attributing census-tract data to zip-code centroids. Although such errors are not random, we can treat them as if they were random to make efficient predictions. Thus, the design model is useful.

In this article, we incorporate location error by building on established kriging methodology. In the following section, we describe the model, discuss estimation of the location-error variance, investigate distribution of the spatial lag, and examine the

effect that uncertainty in the lag has on the first two moments of the spatial random process. In section 2, we describe kriging results without adjusting for location error. In section 3, we incorporate the location error into our calculation of the optimal kriging weights and present kriging after adjusting for location error. In section 4, the techniques described are investigated using simulation. In section 5, we summarize our results. An appendix contains the proofs of two important results.

1. DESIGN MODEL FOR LOCATION ERROR

1.1 Model Description

In design-model sampling, location error causes an intended site $\mathbf{s}_i$ to be displaced to a realized site $\mathbf{r}_i = \mathbf{s}_i + \mathbf{p}(\mathbf{s}_i)$. Thus, the design model is

$$Z(\mathbf{s}) = Y(\mathbf{r}) + \varepsilon(\mathbf{r}) ,$$
$$\mathbf{r} = \mathbf{s} + \mathbf{p}(\mathbf{s}); \mathbf{s} \in D . \tag{8}$$

The observed attribute value $Z(\mathbf{s})$, at the intended location $\mathbf{s}$, is equal to the error-free attribute value $Y(\mathbf{r})$, at the realized location $\mathbf{r}$, plus measurement error $\varepsilon(\mathbf{r})$ that is independent of $Y(\mathbf{r})$. Throughout, we assume that $\varepsilon(\cdot)$ is a white-noise process with mean zero and variance σ_ε^2. The realized location $\mathbf{r}$ is equal to the intended location $\mathbf{s}$ plus location error $\mathbf{p}(\mathbf{s})$ that is independent of $\mathbf{s}$. The observed attribute values, $\mathbf{Z} \equiv \{Z(\mathbf{s}_1),\ldots,Z(\mathbf{s}_n)\}$, comprise the data. The intended locations, $S \equiv \{\mathbf{s}_1,\ldots,\mathbf{s}_n\}$, are a collection of known parameters.

Typically, the data provide no information on $R|S$ in (5). If a sample is taken with a single positioning instrument, then to estimate R we must have prior knowledge of the positioning instrument's accuracy. One way to gather information on $R|S$ is to locate S with the usual positioning instrument and then to record R with a highly accurate instrument. The location error is then given by $\mathbf{p}(\mathbf{s}_i) \equiv \mathbf{r}_i - \mathbf{s}_i$; $i = 1,\ldots,n$.

Throughout, we assume one of the following two distributions for location error:

A1. *iid* normal location error; that is, $(vecR - vecS)|vecS \sim N(\mathbf{0},\sigma_p^2 I_{dn})$, where $vec\,R \equiv (\mathbf{r}_1',\ldots,\mathbf{r}_n')'$, $\mathbf{r}_i \in \Re^d$; $i = 1,\ldots,n$, $vec\,S$ is defined similarly, I_{dn} is the identity matrix of dimension $dn \times dn$, and d is the dimension of the spatial domain.

A2. *iid* uniform location error; that is, $\left(\{r_{ij}\} - \{s_{ij}\}\right)|\{s_{ij}\} \overset{iid}{\sim} Unif\left(-\sqrt{3}\sigma_p, \sqrt{3}\sigma_p\right)$; $i = 1,\ldots,n; j = 1,\ldots,d$, where d is the dimension of the spatial domain.

In A1 and A2, independent location errors are assumed from site to site and in the d directions for each site. For the design model and normal errors A1, the maximum likelihood estimate (MLE) of the location error variance, σ_p^2, is

$$\hat{\sigma}_p^2 = \frac{\sum_{i=1}^{n} (\mathbf{r}_i - \mathbf{s}_i)'(\mathbf{r}_i - \mathbf{s}_i)}{dn} ,$$

which is unbiased for σ_p^2. For the design model and uniform errors A2, the MLE of σ_p^2 is

$$\hat{\sigma}_p^2 = \frac{1}{3}(\max|r_{ij} - s_{ij}|)^2 ,$$

an unbiased estimator of σ_p, based on the MLE, is $\left[(dn+1)/\left(\sqrt{3}dn\right)\right](\max|r_{ij}-s_{ij}|)$.

1.2 Distribution of the Spatial Lag for the Design Model

The spatial lag between $\mathbf{s}_i$ and $\mathbf{s}_j$ is defined as $\mathbf{h}_{ij} \equiv \mathbf{s}_j - \mathbf{s}_i$. In geostatistical theory, no uncertainty is associated with $\mathbf{h}_{ij}$. However, in the presence of location error, $\mathbf{h}_{ij}$ is different from the actual lag $\mathbf{v}_{ij} \equiv \mathbf{r}_j - \mathbf{r}_i$. For the design model and assuming $\mathbf{p}(\mathbf{s})$ given by (8) are $iid(\mathbf{0}, \sigma_p^2 I_d)$; $\mathbf{s} \in D$, then $E(\mathbf{v}_{ij}) = \mathbf{h}_{ij}$ and $\mathrm{var}(\mathbf{v}_{ij}) = 2\sigma_p^2 I_d$. (Note that normal location errors A1 and uniform location errors A2 are examples of such $\mathbf{p}(\mathbf{s})$.)

For illustration, consider data in $\Re^1$, and assume that $\mathbf{h}_{ij} = s_j - s_i > 0$. With location error, it is possible that $v_{ij} = r_j - r_i < 0$. Since the covariance is invariant to the order of the sites, our interest is in $\xi = |v_{ij}|$. For the design model and normal errors A1, it is straightforward to derive the following:

$$E(\xi) = \frac{2\sigma_p}{\sqrt{\pi}} \exp\left\{\frac{-h^2}{4\sigma_p^2}\right\} + h\left[\Phi\left(\frac{h}{\sigma_p\sqrt{2}}\right) - \Phi\left(\frac{-h}{\sigma_p\sqrt{2}}\right)\right]$$

and

$$\mathrm{var}(\xi) = 2\sigma_p^2 + h^2 - \left(\frac{4\sigma_p^2}{\pi}\exp\left\{\frac{-h^2}{2\sigma_p^2}\right\} + \frac{4\sigma_p h}{\sqrt{\pi}}\exp\left\{\frac{-h^2}{4\sigma_p^2}\right\}\left[\Phi\left(\frac{h}{\sigma_p\sqrt{2}}\right) - \Phi\left(\frac{-h}{\sigma_p\sqrt{2}}\right)\right] + h^2\left[\Phi\left(\frac{h}{\sigma_p\sqrt{2}}\right) - \Phi\left(\frac{-h}{\sigma_p\sqrt{2}}\right)\right]^2\right),$$

where $h = h_{ij}$ for notational convenience.

Curves A and B in Figure 2 show the expectation, $E(\xi)$, and the standard deviation, $\sigma(\xi) \equiv (\mathrm{var}(\xi))^{\frac{1}{2}}$, as a function of h, for normal errors A1 and $\sigma_p^2 = 100$. As h increases, the effect of location error diminishes. In the limit, as h approaches ∞, $E(\xi)$ approaches h, represented by line D, and $\sigma(\xi)$ approaches $\sqrt{2}\sigma_p$, represented by line C. As h increases, the probability that the order of r_i and r_j is a *flip* of the order of s_i and s_j decreases. Therefore, ξ converges (in probability) to $v_{ij} \sim N(h, 2\sigma_p^2)$. For small h, $E(\xi) > h$ and $\sigma(\xi) < \sqrt{2}\sigma_p$. For example, at $h = \sigma_p = 10$, we have $E(\xi) = 1.4h$ and $\sigma(\xi) = 1.02\sigma_p$. Therefore, *iid* normal location error has an appreciable effect on the expected lag only for small h. From a simulation, Atkinson (1997) concludes that location error increases the variogram only at short lags (see also Lewis and Hutchinson 2000).

For the design model and uniform errors A2, it is straightforward to derive

$$E(\xi) = \begin{cases} \dfrac{1}{12\sigma_p^2}\left[-\dfrac{1}{3}h^3 + 2\sqrt{3}\sigma_p h^2 + 8\sqrt{3}\sigma_p^3\right]; & \text{if } h \le 2\sqrt{3}\sigma_p, \\ h & ; \text{ otherwise,} \end{cases}$$

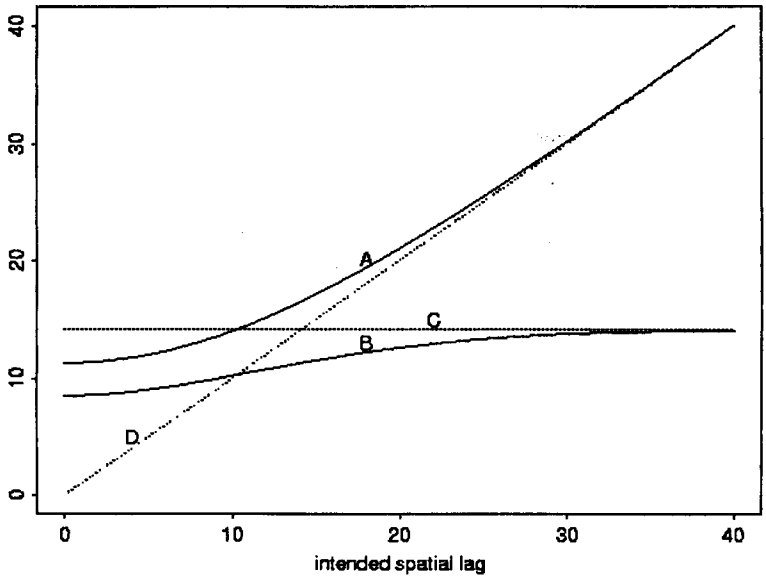

FIG. 2. The Case of *iid* Normal Errors A1 in $\Re^1$: Curve A shows $E(\xi)$, curve B shows $\sigma(\xi)$, line C is $\sqrt{2}\,\sigma_p$, and line D is h.

and

$$\mathrm{var}(\xi) = \begin{cases} \dfrac{2}{3}\sigma_p^2 - \dfrac{1}{1296\sigma_p^4}h^6 + \dfrac{1}{108\sigma_p^3}\sqrt{3}h^5 \\ -\dfrac{1}{12\sigma_p^2}h^4 + \dfrac{1}{27\sigma_p}\sqrt{3}h^3 + \dfrac{1}{3}h^2; & \text{if } h \le 2\sqrt{3}\sigma_p\,, \\ 2\sigma_p^2 & ;\ \text{otherwise.} \end{cases}$$

Curves A and B in Figure 3 show $E(\xi)$ and $\sigma(\xi)$ as a function of h, for uniform errors A2 and $\sigma_p^2 = 100$. As h increases, the effect of location error diminishes until $E(\xi) = h$ and $\sigma(\xi) = \sqrt{2}\,\sigma_p$; $h \ge 2\sqrt{3}\,\sigma_p$, because the order of r_i and r_j must be the same as the order of s_i and s_j for $h \ge 2\sqrt{3}\,\sigma_p$. There is little difference between these curves and those in Figure 2. This suggests that the distribution of the location error is relatively unimportant, at least in the *iid* case.

1.3 Attribute Moments for the Design Model

First, we assume that the hidden process $Y(\cdot)$, introduced in (8), satisfies (1). The expectation of the observable process $Z(\cdot)$ is found by integrating the expectation of $Y(\cdot)$ over the locations realizable under the location error distribution. For the design model, and assuming that the measurement error $\xi(\cdot)$ has a mean of zero, the first moment after adjusting for location error is,

$$E(Z(\mathbf{s})) = \int \mu(\mathbf{s} + \mathbf{u})g(\mathbf{u})d\mathbf{u}\,, \tag{9}$$

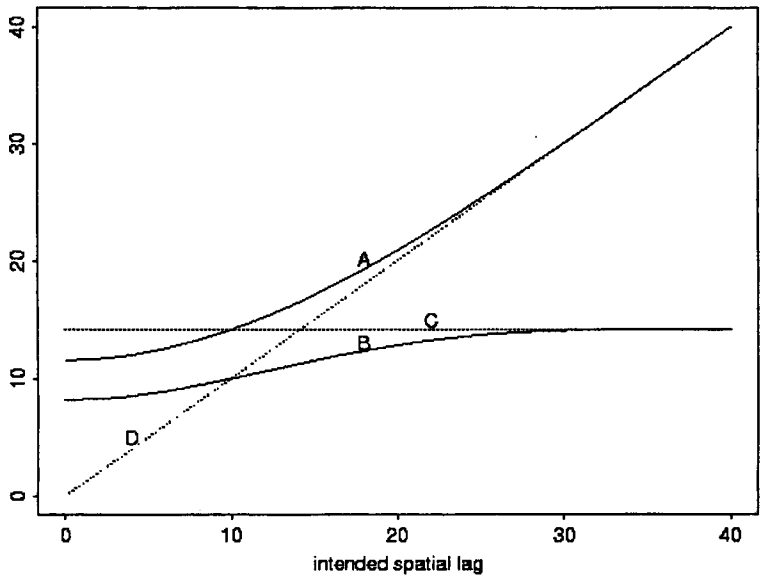

FIG. 3. The Case of *iid* Uniform Errors A2 in $\Re^1$: Curve A shows $E(\xi)$, curve B shows $\sigma(\xi)$, line C is $\sqrt{2}\,\sigma_p$, and line D is h.

where $\mu(\cdot) = E(Y(\cdot)) = \mathbf{f}(\cdot)'\beta$, and the location error $\mathbf{u}$ has probability density function (pdf) g. Note that (9) holds even when $\mu(\cdot)$ is not linear in the parameters β.

Similarly, we adjust the covariance of $Z(\cdot)$ by integrating the covariance of $Y(\cdot)$ over the locations realizable under the location error distribution. For the design model, and assuming that the location errors are *iid* mean zero, the covariance after adjusting for location error is,

$$C_p(\mathbf{h}) \equiv \mathrm{cov}(Z(\mathbf{s}), Z(\mathbf{s}+\mathbf{h})) = \int\int C(\mathbf{h}+\mathbf{v}-\mathbf{u})g(\mathbf{u})g(\mathbf{v})d\mathbf{u}d\mathbf{v}\,, \tag{10}$$

where $C(\cdot)$ is the covariance function given in (1) for $Y(\cdot)$. Note that, because we assume the location error has a density and no discrete mass at $\mathbf{0}$, the variance of $\varepsilon(\cdot)$, σ_ε^2, does not enter into (10).

The covariance after adjusting for location error must be positive semidefinite; that is, $C_p(\mathbf{h})$ must satisfy

$$\sum_{i=1}^{m}\sum_{j=1}^{m} a_i a_j C_p(\mathbf{s}_j - \mathbf{s}_i) \geq 0\,,$$

for any finite collection of sites $\{\mathbf{s}_i\colon i = 1,\ldots,m\}$ and real numbers $\{a_i\colon i = 1,\ldots,m\}$. Since $C(\cdot)$ in (10) is positive semidefinite, $C_p(\cdot)$ is also positive semidefinite.

Unfortunately, the integration in (10) is analytically intractable for many commonly used covariance models, assuming even the simplest location error models. The integrals in (9) and (10) can be evaluated using Monte Carlo methods. We evaluate (9) by

(i) drawing a random sample $\{\mathbf{u}_i\colon i = 1,\ldots,n\}$ from $g(\cdot)$;

(ii) estimating $E(Z(\mathbf{s}))$ as $\frac{1}{n}\sum_{i=1}^{n}\mu(\mathbf{s}+\mathbf{u}_i)$.

Similarly, we can evaluate (10).

Here, we discuss two analytically tractable covariance models to illustrate the effect of location error. Assume $p(s) \overset{iid}{\sim} Unif\left(-\sqrt{3}\sigma_p, \sqrt{3}\sigma_p\right); s \in D \subset \Re^1$ and the covariance is given by the isotropic exponential covariance function,

$$C(h) = \sigma^2 \exp\left\{-\frac{|h|}{\theta}\right\}, \tag{11}$$

where σ^2 is the variance of $Y(\cdot)$, h is the intended lag, and θ is a scale parameter. Small values of θ correspond to only localized spatial dependence, and large values correspond to longer-range spatial dependence. By substituting (11) into (10), we obtain

$$C_p(h) = \begin{cases} \dfrac{\sigma^2}{12\sigma_p^2}\left[\begin{array}{l} 4\sqrt{3}\sigma_p\theta - 2h\theta - 2\theta^2 \exp\left\{-\dfrac{h}{\theta}\right\} \\ + \theta^2 \exp\left\{\dfrac{h-2\sqrt{3}\sigma_p}{\theta}\right\} + \theta^2 \exp\left\{\dfrac{-h-2\sqrt{3}\sigma_p}{\theta}\right\} \end{array}\right], & \text{if } h \le 2\sqrt{3}\sigma_p, \\ \dfrac{\sigma^2}{12\sigma_p^2}\left[\begin{array}{l} -2\theta^2 \exp\left\{-\dfrac{h}{\theta}\right\} + \theta^2 \exp\left\{\dfrac{-h+2\sqrt{3}\sigma_p}{\theta}\right\} \\ + \theta^2 \exp\left\{\dfrac{-h-2\sqrt{3}\sigma_p}{\theta}\right\} \end{array}\right], & \text{if } h > 2\sqrt{3}\sigma_p. \end{cases} \tag{12}$$

Figure 4 illustrates the effect of *iid* uniform location error on the exponential covariance function in $\Re^1$ when the parameters $\sigma^2 = 1000$, $\sigma_p^2 = 100$, and $\theta = 100$ are fixed. For each intended lag $h = 0,1,\ldots,25$, we proceed as follows:

(i) Calculate the exponential covariance (12) after adjusting for location error. Let w denote the resulting function of h.

(ii) Calculate the exponential covariance (11) without adjusting for location error, using the intended lag h. Let x denote the resulting function of h.

In Figure 4, curve A is a plot of w against h and B is a plot of x against h. The effect of location error on the covariance is most pronounced for small lags, where the realized lag can be quite different from the intended lag. At $h = 0$, $x(0) = 1000$, and $w(0) = 893.9$. Thus, failing to adjust gives covariances that can be too large by more than 10 percent. As h increases, the effect of location error diminishes.

To obtain second moments of the effect of location error on covariances, we use Monte Carlo simulation. Table 1 gives the consequences of *iid* uniform location error on the exponential covariance function in $\Re^1$. Case I shows the effect of small location error ($\sigma_p^2 = 10$), case II moderate location error ($\sigma_p^2 = 100$), and case III large location error ($\sigma_p^2 = 1000$). For each, we consider several intended lags ($h = 0, 25, 100$) and several values of the scale parameter ($\theta = 50, 100, 500$). The process variance is fixed at $\sigma^2 = 1000$. (The relative effect of location error is not a function of the process variance.) For each combination of σ_p^2, θ, and h, the Monte Carlo simulation proceeds as follows:

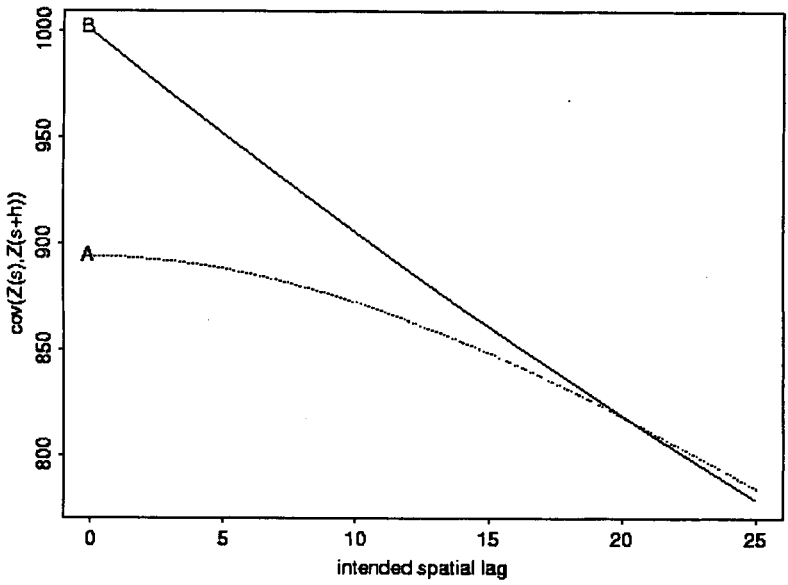

FIG. 4. Curve A Is the Exponential Covariance (12) after Adjusting for Location Error in $\Re^1$. Curve B Is the Exponential Covariance (11) without Adjusting for Location Error in $\Re^1$.

(i) Generate 1000 realizations of independent location errors, u and v, according to the *iid* uniform distribution on the interval $\left(-\sqrt{3}\sigma_p, \sqrt{3}\sigma_p\right)$.

(ii) Calculate the exponential covariance (12) after adjusting for location error. Let w denote the resulting function of h.

(iii) Calculate the exponential covariance (11) without adjusting for location error, using the intended lag h. Let x denote the resulting function of h.

(iv) Calculate the exponential covariance (11) without adjusting for location error, using the realized lag $(h + v - u)$. Let y denote the resulting function of h.

The mean squared errors (MSE) are obtained for the covariance after adjusting for location error (MSE^c) and the covariance without adjusting for location error (MSE^d) using the Monte Carlo estimators,

$$\mathrm{MSE}^c(h) = \frac{1}{1000}\sum_{i=1}^{1000}\left(w(h)^{(i)} - y(h)^{(i)}\right)^2$$

and

$$\mathrm{MSE}^d(h) = \frac{1}{1000}\sum_{i=1}^{1000}\left(x(h)^{(i)} - y(h)^{(i)}\right)^2,$$

where i denotes the ith realization simulated. Table 1 shows that the MSE^c is smaller than the MSE^d for all combinations of σ_p^2, θ, and h, where the difference is noteworthy. From Figure 4 and Table 1, we conclude:

TABLE 1
The Effect of *iid* Uniform Location Error on the Exponential Covariance Function in $\Re^1$

Case		Cov[a]	Cov[b]	MSE[c]	MSE[d]	Percent Change[e]
I	$\theta = 50, h = 0$	930.80	1000.00	2294	7207	−68.17
	$\theta = 50, h = 25$	608.97	606.53	3016	3020	−0.13
	$\theta = 50, h = 100$	135.88	135.34	150	150	0.00
	$\theta = 100, h = 0$	964.46	1000.00	626	1964	−68.13
	$\theta = 100, h = 25$	779.58	778.80	120	120	0.00
	$\theta = 100, h = 100$	368.24	367.88	283	283	0.00
	$\theta = 500, h = 0$	992.74	1000.00	26	78	−66.67
	$\theta = 500, h = 25$	951.27	951.23	71	71	0.00
	$\theta = 500, h = 100$	818.76	818.73	53	53	0.00
II	$\theta = 50, h = 0$	804.09	1000.00	15973	55285	−71.11
	$\theta = 50, h = 25$	626.20	606.53	29530	29752	−0.75
	$\theta = 50, h = 100$	140.84	135.34	1631	1662	−1.87
	$\theta = 100, h = 0$	893.88	1000.00	5178	16937	−69.43
	$\theta = 100, h = 25$	784.13	778.80	11621	11603	0.16
	$\theta = 100, h = 100$	371.57	367.88	2863	2862	0.04
	$\theta = 500, h = 0$	977.30	1000.00	250	749	−66.62
	$\theta = 500, h = 25$	951.11	951.23	664	665	−0.15
	$\theta = 500, h = 100$	819.06	818.73	521	520	0.19
III	$\theta = 50, h = 0$	542.79	1000.00	57562	266352	−78.39
	$\theta = 50, h = 25$	504.35	606.53	62384	73492	−15.11
	$\theta = 50, h = 100$	198.43	135.43	32337	35474	−8.84
	$\theta = 100, h = 0$	716.39	1000.00	29402	104112	−71.76
	$\theta = 100, h = 25$	685.90	778.80	33989	41240	−17.58
	$\theta = 100, h = 100$	405.93	367.88	31515	32706	−3.64
	$\theta = 500, h = 0$	930.80	1000.00	2247	7039	−68.08
	$\theta = 500, h = 25$	921.42	951.23	2961	3889	−23.86
	$\theta = 500, h = 100$	821.96	818.73	5367	5364	0.06

NOTES: [a]Exponential covariance (12) after adjusting for location error. [b] Exponential covariance (11) without adjusting for location error, using the intended lag. [c] The MSE for the covariance after adjusting for location error. [d] The MSE for the covariance without adjusting for location error. [e] Relative change in the expected MSE, found as $100(\text{MSE}^c - \text{MSE}^d)/\text{MSE}^d$.

(i) As σ_p^2 increases so does the absolute effect of location error on the covariance.
(ii) While the relative change in the MSE is insensitive to θ, the magnitude of the MSE is sensitive to θ.
(iii) Both the relative change in and the magnitude of the MSE are affected by h.

In two-dimensional Euclidean space, assume location error $\mathbf{p}(\mathbf{s}) \overset{iid}{\sim} N(\mathbf{0}, \sigma_p^2 I_2)$ $\mathbf{s} \in D \subset \Re^2$ and the covariance is modeled by the isotropic Gaussian covariance function

$$C(\mathbf{h}) = \sigma^2 \exp\left\{-\frac{\|\mathbf{h}\|^2}{\phi^2}\right\}, \tag{13}$$

where σ^2 is the variance of $Y(\cdot)$, $\mathbf{h}$ is the intended lag, and ϕ is a scale parameter. Then,

$$C(\mathbf{h}) = \frac{\sigma^2\phi^2}{4\sigma_p^2 + \phi^2} \exp\left\{-\frac{\|\mathbf{h}\|^2}{\left(4\sigma_p^2 + \phi^2\right)}\right\}. \tag{14}$$

Similar to the previous discussion, Figure 5 and Table 2 show the consequences of *iid* normal location error on the Gaussian covariance function in $\Re^2$. Comparing Tables 1 and 2, we draw the same three conclusions as we did for *iid* uniform error in

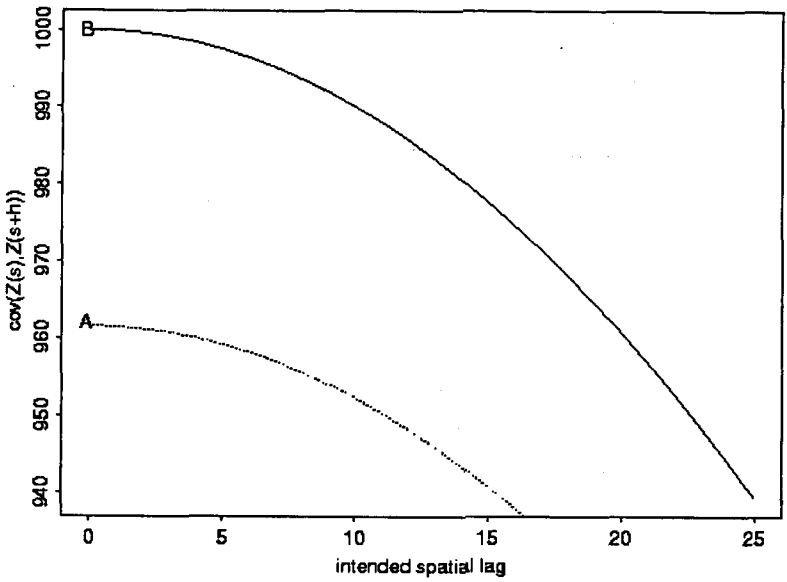

FIG. 5. Curve A is the Gaussian Covariance (14) after Adjusting for Location Error in $\Re^2$. Curve B Is the Gaussian Covariance (13) without Adjusting for Location Error in $\Re^2$.

$\Re^1$. However, the magnitude of the effect of location error depends on the covariance model. For $\theta = \phi = 50$ and $h = 0$, the covariance after adjusting for location error is 930.80 in the uniform-exponential case (Table 1) and 984.25 in the normal-Gaussian case (Table 2). In either case, these values are to be compared to 1,000, the covariance without adjusting for location error.

We must also adjust the covariance between an intended site $\mathbf{s}$ and the location-error-free and measurement-error-free prediction site $\mathbf{s}_o$. For the design model and assuming zero-mean measurement error, the covariance after adjusting for location error between $\mathbf{s}$ and $\mathbf{s}_o$ is

$$\mathrm{cov}(Z(\mathbf{s}), Y(\mathbf{s}_o)) = \int C(\mathbf{s}_o - \mathbf{s} + \mathbf{u}) g(\mathbf{u}) d\mathbf{u}\,, \tag{15}$$

where $C(\cdot)$ is the covariance function for $Y(\cdot)$ and the location error $\mathbf{u}$ has pdf g.

2. KRIGING WITHOUT ADJUSTING FOR LOCATION ERROR

The model used for kriging without adjusting for location error is

$$\begin{aligned} Z(\mathbf{s}) &= Y(\mathbf{s}) + \varepsilon(\mathbf{s})\,, \\ Y(\mathbf{s}) &= \mu(\mathbf{s}) + \delta(\mathbf{s})\,. \end{aligned} \tag{16}$$

Therefore, at location $\mathbf{s}$, the observed attribute value $Z(\mathbf{s})$ is equal to the mean $\mu(\mathbf{s})$ plus spatially dependent error $\delta(\mathbf{s})$ plus measurement error $\varepsilon(\mathbf{s})$. Kriging finds the best linear unbiased predictor at a given location for a known covariance or variogram model. If $\mu(\mathbf{s}) = \mu$, for all $\mathbf{s} \in D$, we can use ordinary kriging to predict the value of

TABLE 2
The Effect of *iid* Gaussian Location Error on the Gaussian Covariance Function in $\Re^2$

Case		Cov[a]	Cov[b]	MSE[c]	MSE[d]	Percent Change[e]
I	$\phi = 50, h = 0$	984.25	1000.00	232	474	−51.05
	$\phi = 50, h = 25$	769.56	778.80	4280	4299	−0.44
	$\phi = 50, h = 100$	19.20	18.32	46	47	−2.13
	$\phi = 100, h = 0$	996.50	1000.00	16	32	−50.00
	$\phi = 100, h = 25$	935.90	939.41	440	452	−2.65
	$\phi = 100, h = 100$	367.88	367.88	1034	1034	0.00
	$\phi = 200, h = 0$	999.00	1000.00	1	2	−50.00
	$\phi = 200, h = 25$	983.53	984.50	32	32	0.00
	$\phi = 200, h = 100$	778.22	778.80	309	309	0.00
II	$\phi = 50, h = 0$	862.07	1000.00	13774	32567	−57.71
	$\phi = 50, h = 25$	694.94	778.80	36474	46516	−21.59
	$\phi = 50, h = 100$	27.41	18.32	952	991	−3.94
	$\phi = 100, h = 0$	961.54	1000.00	1421	2958	−51.96
	$\phi = 100, h = 25$	905.46	939.41	4682	5947	−21.27
	$\phi = 100, h = 100$	367.60	367.88	10164	10165	−0.01
	$\phi = 200, h = 0$	990.10	1000.00	100	199	−49.75
	$\phi = 200, h = 25$	974.90	984.50	378	466	−18.88
	$\phi = 200, h = 100$	773.00	778.80	3001	3025	−0.79
III	$\phi = 50, h = 0$	384.62	1000.00	87494	474053	−81.54
	$\phi = 50, h = 25$	349.36	778.80	92349	279979	−57.71
	$\phi = 50, h = 100$	82.58	18.32	27361	31216	−12.35
	$\phi = 100, h = 0$	714.29	1000.00	51023	139518	−63.43
	$\phi = 100, h = 25$	683.10	939.41	58076	129898	−55.27
	$\phi = 100, h = 100$	349.67	367.88	61375	61923	−0.88
	$\phi = 200, h = 0$	909.09	1000.00	7295	15966	−54.31
	$\phi = 200, h = 25$	896.27	984.50	9311	17662	−45.75
	$\phi = 200, h = 100$	724.28	778.80	25945	29435	−11.86

NOTES: [a] Gaussian covariance (14) after adjusting for location error. [b] Gaussian covariance (13) without adjusting for location error, using the intended lag. [c] The MSE for the covariance after adjusting for location error. [d] The MSE for the covariance without adjusting for location error. [e] Relative change in the expected MSE, found as $100(\mathrm{MSE}^c - \mathrm{MSE}^d)/\mathrm{MSE}^d$.

the $Y(\cdot)$ process at a location $\mathbf{s}_o \in D$. If $\mu(\mathbf{s}) = \mathbf{f}(\mathbf{s})'\beta$, where $\mathbf{f}(\mathbf{s})$ and β are defined as in (1), then we can use universal kriging for prediction. The universal (and ordinary) kriging predictor is $P(\mathbf{Z};\mathbf{s}_o) = \lambda'\mathbf{Z}$, where λ is an optimal (minimizes mean squared prediction error) vector of kriging coefficients and $\mathbf{Z}$ is the observed attribute (data) vector. The simplification to ordinary kriging (where $q = 0$ and $f_0(\mathbf{s}) \equiv 1$) is straightforward.

The assumptions (1) are made more specific as follows:

B1. Let $Y(\cdot)$ be a spatial random process with $E(Y(\mathbf{r})) = \mathbf{f}(\mathbf{r})'\beta$; $\mathbf{r} \in D$, where the $(q+1)\times 1$ vector of known explanatory variables is $\mathbf{f}(\cdot) = \{f_j(\cdot): j = 0,1,\ldots,q\}$ and β is a $(q+1)\times 1$ vector of unknown regression coefficients. We write the $(q+1)\times n$ matrix $F' \equiv (\mathbf{f}(\mathbf{s}_1),\ldots,\mathbf{f}(\mathbf{s}_n))$.

B2. Let $C(\mathbf{r}_2 - \mathbf{r}_1) = \sigma^2 K_\theta(\mathbf{r}_2 - \mathbf{r}_1)$ be the covariance function of $Y(\cdot)$, where σ^2 is the variance of $Y(\cdot)$, $K_\theta(\cdot)$ is a valid parametric class of correlation models, $\theta \in \Theta$ is an $m \times 1$ vector of correlation-model parameters, and Θ is an open set in $\Re^m$.

Then, the kriging coefficients are given by

$$\lambda' = \mathbf{b}'(F'K^{-1}F)^{-1}F'K^{-1} + \mathbf{k}'K^{-1},$$

where $\mathbf{b} = \mathbf{f}(\mathbf{s}_o) - F'K^{-1}\mathbf{k}$, F is the $n\times(q+1)$ matrix of explanatory variables at S, $\mathbf{f}(\mathbf{s}_o)$ is the vector of explanatory variables at $\mathbf{s}_o$, K is the $n \times n$ correlation matrix between

$Z(\cdot)$ at S, and $\mathbf{k}$ is the vector of correlations between $Y(\cdot)$ at $\mathbf{s}_o$ and $Z(\cdot)$ at S. The universal kriging prediction variance is,

$$\sigma_k^2(\mathbf{s}_o) = \sigma^2(1 - \mathbf{k}'K^{-1}\mathbf{k} + \mathbf{b}'(F'K^{-1}F)^{-1}\mathbf{b}) - \sigma_\varepsilon^2 .$$

Proofs of these results can be found in Cressie (1993, pp. 128 and 153–55). Note that the kriging predictor and the prediction variance depend on the correlation matrix, whose entries are functions of the lag.

3. KRIGING AFTER ADJUSTING FOR LOCATION ERROR (KAALE)

Assuming (1), the choice of kriging weights is a function of the lag. Kriging methodology assumes that the lag is known without error. In this section, we develop methodology for kriging after adjusting for location error (KAALE). We build upon standard kriging methodology by adjusting the component parts to include location error.

The KAALE predictor is a linear combination of the attribute data, $\lambda'\mathbf{Z}$, where λ is an $n \times 1$ vector of kriging coefficients. The KAALE results differ from those of standard kriging because expectations and covariances have to be adjusted for location error using (9) and (10), respectively.

Assuming B1 and B2, the matrix of explanatory variables after adjusting for location error F_p is

$$F_p = \begin{pmatrix} \int f_0(\mathbf{s}_1 + \mathbf{u})g(\mathbf{u})d\mathbf{u} \ldots \int f_q(\mathbf{s}_1 + \mathbf{u})g(\mathbf{u})d\mathbf{u} \\ \int f_0(\mathbf{s}_2 + \mathbf{u})g(\mathbf{u})d\mathbf{u} \ldots \int f_q(\mathbf{s}_2 + \mathbf{u})g(\mathbf{u})d\mathbf{u} \\ \vdots \qquad \vdots \qquad \vdots \\ \int f_0(\mathbf{s}_n + \mathbf{u})g(\mathbf{u})d\mathbf{u} \ldots \int f_q(\mathbf{s}_n + \mathbf{u})g(\mathbf{u})d\mathbf{u} \end{pmatrix}, \tag{17}$$

where $\{f_j(\cdot): j = 0,1,\ldots,q\}$ are known functions, typically polynomial trend terms. The optimal vector of kriging coefficients λ, after adjusting for location error is

$$\lambda' = \mathbf{b}_p'(F_p'K_p^{-1}F_p)^{-1}F_p'K_p^{-1} + \mathbf{k}_p'K_p^{-1} ,$$

where $\mathbf{b}_p \equiv \mathbf{f}(\mathbf{s}_o) - F_p'K_p^{-1}\mathbf{k}_p$; F_p is defined in (17); K_p is the correlation matrix given by (10); and $\mathbf{k}_p$ is the correlation vector given by (15). (See Proposition 1 in the Appendix for details.)

The prediction variance $\sigma_k^2(\mathbf{s}_o)$ is

$$\sigma_k^2(\mathbf{s}_o) = \sigma^2\{1 - \mathbf{k}_p'K_p^{-1}\mathbf{k}_p + \mathbf{b}_p'(F_p'K_p^{-1}F_p)^{-1}\mathbf{b}_p\} - \sigma_\varepsilon^2 ,$$

where σ_ε^2 is the measurement-error variance. (See Proposition 2 in the Appendix for details.)

In (10), we give $C_p(\cdot)$, the covariance after adjusting for location error. While $C_p(\cdot)$ will be positive semidefinite if $C(\cdot)$ in (1) is positive semidefinite, the method-of-moments estimator $\hat{C}_p$ of C_p may not be positive semidefinite. Define,

$$\hat{C}_p(\mathbf{h}) = \frac{1}{|N(\mathbf{h})|} \sum_{N(\mathbf{h})} (Z(\mathbf{s}_i) - \bar{Z})(Z(\mathbf{s}_j) - \bar{Z}), \tag{18}$$

where $N(\mathbf{h}) \equiv \{(i,j): \mathbf{s}_j - \mathbf{s}_i = \mathbf{h}\}$, $|N(\mathbf{h})|$ is the number of pairs of sites separated by $\mathbf{h}$, and $\bar{Z} = \frac{1}{n}\sum Z(\mathbf{s}_i)$. In practice, we let $N(\mathbf{h}) \equiv \{(i,j): \mathbf{s}_j - \mathbf{s}_i \in (\mathbf{h} - \Delta\mathbf{h}, \mathbf{h} + \Delta\mathbf{h})\}$, where $\Delta\mathbf{h}$ is a bandwidth. To insure that C_p is positive semidefinite, we fit a valid covariance model, after adjusting for location error, to (18). Estimation of the parameters proceeds similarly to estimation of the parameters of the covariance model without adjusting for location error (see, for example, Cressie 1993, section 2.4).

4. SIMULATION STUDY

In this section, we investigate the effect of location error using a simulation study. We assume an exponential covariance function in $\Re^1$, and use a 2^6 factorial design where the factors are the prediction location, the measurement-error variance, the scale parameter (amount of spatial dependence), the location-error variance, the sample size, and the trend. The spatial domain D is the one-dimensional subset of $\Re^1$ from 100 to 200. The "low" level of the sample size, $n = 20$, corresponds to the intended sites $S \equiv \{100,105,\ldots,195\}$ and the "high" level of the sample size, $n = 40$, to the intended sites $S \equiv \{100,102.5,\ldots,197.5\}$. Data are generated according to the design model (8) and uniform errors A2.

Levels for the six factors are given in Table 3. The low level of the prediction site, $s_o = 155$, corresponds to a location near the middle of the spatial domain and the high level, $s_o = 200$, to a location on the boundary. The low level of the measurement error variance, $\sigma_\varepsilon^2 = 10$, corresponds to small measurement error and the high level, $\sigma_\varepsilon^2 = 10{,}000$, to large measurement error. The low level of the scale parameter, $\theta = 5$, corresponds to localized spatial dependence and the high level, $\theta = 50$, to longer-range spatial dependence. The low level of the location error variance, $\sigma_p^2 = 1$, corresponds to small location uncertainty and the high level, $\sigma_p^2 = 64$, to large location uncertainty. The low level of the trend, $\mu(r) = 1000$, corresponds to a constant mean and the high level, $\mu(r) = 1000 + r - .01r^2$, to a quadratic trend that reaches its maximum value when $r = 100$ and its minimum value when $r = 200$.

For the sixty-four combinations of the experimental factors, we perform a Monte Carlo simulation study as follows:

(i) For each of five hundred realizations, realized sites R are generated by adding *iid* uniform location error on $\left(-\sqrt{3}\sigma_p, \sqrt{3}\sigma_p\right)$ to each of the intended sites S.

TABLE 3
Factors for the Monte Carlo Simulation Study in $\Re^1$

Factor	Low Level	High Level
Prediction site	$s_o = 155$	$s_o = 200$
Measurement-error variance	$\sigma_\varepsilon^2 = 10$	$\sigma_\varepsilon^2 = 10{,}000$
Range parameter	$\theta = 5$	$\theta = 50$
Location-error variance	$\sigma_p^2 = 1$	$\sigma_p^2 = 64$
Sample size	$n = 20$	$n = 40$
Trend	[a]no trend	[b]strong trend

NOTES: [a] No trend corresponds to the constant mean $\mu(r) = 1000$, where r is the realized location. [b] Strong trend corresponds to the mean function $\mu(r) = 1000 + r - .01r^2$, where r is the realized location.

(ii) For each realization, R and the prediction site s_o are used to generate a Gaussian spatial random vector according to the exponential covariance function (11). This gives $(\mathbf{Y}', Y(s_o))'$.
(iii) For each realization, *iid* normal measurement error is added to $\mathbf{Y}$. This gives the observed values $\mathbf{Z}$.
(iv) For each realization, the kriging predictor at s_o, $P(\mathbf{Z}; s_o)$, is found using (i) the intended sites and the exponential covariance (11) without adjusting for location error and (ii) the intended sites and the exponential covariance (12) after adjusting for location error.

The bias and the mean squared prediction error (MSPE) of the predictor $P(\mathbf{Z}; s_o)$ are

$$\text{Bias}(P(\mathbf{Z}; s_o)) \equiv \text{E}(P(\mathbf{Z}; s_o) - Y(s_o)) ,$$

$$\text{MSPE}(P(\mathbf{Z}; s_o)) \equiv \text{E}((P(\mathbf{Z}; s_o) - Y(s_o))^2) .$$

We estimate these quantities from the simulation as follows:

$$\widehat{\text{Bias}}(P(\mathbf{Z}; s_o)) \equiv \frac{1}{500}\sum_{i=1}^{500}\left[P(\mathbf{Z}; s_o)^{(i)} - Y(s_0)^{(i)}\right], \tag{19}$$

$$\widehat{\text{MSPE}}(P(\mathbf{Z}; s_o)) \equiv \frac{1}{500}\sum_{i=1}^{500}\left[P(\mathbf{Z}; s_o)^{(i)} - Y(s_0)^{(i)}\right]^2 , \tag{20}$$

where i denotes the ith realization simulated. We have simplified the analysis, isolating the effect of location error on kriging by assuming that the mean and covariance functions are known. Hence, the study does not enable us to investigate relationships between mean-function estimation, covariance estimation, and kriging.

The Monte Carlo estimators (19) and (20) are calculated for the exponential covariance (11) without adjusting for location error and the exponential covariance (12) after adjusting for location error. Tables 4 and 5 give results of the study. A value of 0 represents the low level of the factor and a value of 1 represents the high level. Table 6 gives an analysis of variance (ANOVA) of the difference in the MSPE between kriging without adjusting for location error and kriging after adjusting for location error (KAALE). The ANOVA partitions the total sum of squares into sums of squares due to various treatments and a residual sum of squares. Associated with each sum of squares is its degrees of freedom; mean squares are obtained by dividing sums of squares by their corresponding degrees of freedom. The corresponding variance ratio (VR) for each combination of factors is calculated as

$$VR = \frac{\text{Mean square treatment}}{\text{Mean square residual}} .$$

A large VR implies differences due to the treatment. We pool all treatments involving four or more factors to obtain the residual sum of squares. In an initial run of the full model (that is, all treatments), those of degree four or more had small mean squares.

From the tables, we conclude:

(i) The measurement-error variance plays an important role. At the high level of σ_ε^2, kriging without adjusting for location error and KAALE both have larger bias and

TABLE 4

Results of the Monte Carlo Simulation Study for the Exponential Covariance Function in $\Re^1$ for $n = 20$

Site	σ_ε^2	θ	σ_μ^2	trend	n	[a]*Bias*	[a]*MSPE*	[b]*Bias*	[b]*MSPE*
0	0	0	0	0	0	−1.52	281	−1.52	283
1	0	0	0	0	0	−1.94	845	−1.93	844
0	1	0	0	0	0	−3.62	9749	−3.63	9882
1	1	0	0	0	0	−1.99	2667	−1.94	2752
0	0	1	0	0	0	−0.49	40	−0.49	40
1	0	1	0	0	0	−0.86	180	−0.85	180
0	1	1	0	0	0	−2.60	9530	−2.60	9564
1	1	1	0	0	0	−0.08	8834	−0.07	8870
0	0	0	1	0	0	−3.21	1334	−2.66	1251
1	0	0	1	0	0	−0.73	823	−0.83	795
0	1	0	1	0	0	−5.31	10538	−3.29	10250
1	1	0	1	0	0	−0.78	2633	−0.37	2360
0	0	1	1	0	0	−0.94	269	−0.57	212
1	0	1	1	0	0	0.13	244	−0.12	201
0	1	1	1	0	0	−3.04	9630	−1.25	7761
1	1	1	1	0	0	0.91	8977	1.28	5113
0	0	0	0	1	0	−3.19	2017	−3.18	2017
1	0	0	0	1	0	−3.42	2556	−3.39	2553
0	1	0	0	1	0	−5.23	11490	−5.22	11491
1	1	0	0	1	0	−3.86	10380	−3.75	10584
0	0	1	0	1	0	−2.21	1864	−2.20	1864
1	0	1	0	1	0	−2.49	1619	−2.40	1619
0	1	1	0	1	0	−4.25	11361	−4.24	11362
1	1	1	0	1	0	−1.14	16346	−1.12	16386
0	0	0	1	1	0	−3.99	3272	−3.37	3282
1	0	0	1	1	0	−2.73	2926	−2.75	2770
0	1	0	1	1	0	−6.04	12408	−5.46	12495
1	1	0	1	1	0	−3.17	10630	−2.70	8771
0	0	1	1	1	0	−1.72	2215	−1.10	2218
1	0	1	1	1	0	−1.21	2408	−1.82	1948
0	1	1	1	1	0	−3.76	11509	−3.21	11559
1	1	1	1	1	0	0.05	17060	−0.17	9879

NOTES: [a]*Bias* and *MSPE* are found using the exponential covariance (11) without adjusting for location error. [b]*Bias* and *MSPE* are found using the exponential covariance (12) after adjusting for location error.

TABLE 5

Results of the Monte Carlo Simulation Study for the Exponential Covariance Function in $\Re^1$ for $n = 40$

Site	σ_ε^2	θ	σ_μ^2	trend	n	[a]*Bias*	[a]*MSPE*	[b]*Bias*	[b]*MSPE*
0	0	0	0	0	1	0.41	387	0.41	394
1	0	0	0	0	1	0.75	630	0.70	629
0	1	0	0	0	1	−3.01	10280	−3.13	10611
1	1	0	0	0	1	−5.34	4995	−5.61	5325
0	0	1	0	0	1	0.14	51	0.14	51
1	0	1	0	0	1	0.14	104	0.14	104
0	1	1	0	0	1	−3.25	9830	−3.22	9717
1	1	1	0	0	1	−8.60	9861	−8.60	9845
0	0	0	1	0	1	−2.03	1417	−1.56	1318
1	0	0	1	0	1	3.05	895	3.96	743
0	1	0	1	0	1	−5.89	11780	−4.97	10819
1	1	0	1	0	1	−3.00	5170	2.51	2629
0	0	1	1	0	1	−1.04	283	−0.75	217
1	0	1	1	0	1	0.96	230	1.81	144
0	1	1	1	0	1	−4.63	10328	−3.76	7367
1	1	1	1	0	1	−8.43	9962	−0.80	3893
0	0	0	0	1	1	−5.41	5282	−5.40	5282
1	0	0	0	1	1	−4.53	4337	−4.58	4338
0	1	0	0	1	1	−10.27	15506	−10.26	15507
1	1	0	0	1	1	−11.73	11640	−11.96	12046
0	0	1	0	1	1	−5.67	5095	−5.66	5095
1	0	1	0	1	1	−4.90	3806	−4.88	3806
0	1	1	0	1	1	−10.52	15185	−10.51	15185
1	1	1	0	1	1	−14.26	16267	−14.21	16123
0	0	0	1	1	1	−10.47	8801	−9.83	8796
1	0	0	1	1	1	−5.00	6579	−4.02	6106
0	1	0	1	1	1	−15.79	19098	−15.15	19115
1	1	0	1	1	1	−12.70	13851	−8.04	9469
0	0	1	1	1	1	−10.06	7766	−9.40	7761
1	0	1	1	1	1	−6.99	6332	−6.17	5760
0	1	1	1	1	1	−15.38	17786	−14.70	17799
1	1	1	1	1	1	−17.11	18751	−11.21	10208

NOTES: [a]*Bias* and *MSPE* are found using the exponential covariance (11) without adjusting for location error. [b]*Bias* and *MSPE* are found using the exponential covariance (12) after adjusting for location error.

TABLE 6

ANOVA of MSPE for Kriging without Adjusting for Location Error minus MSPE for Kriging after Adjusting for Location Error Using the Exponential Covariance in $\Re^1$

Factor	Df	Sum of Squares	Mean Square	Variance Ratio
s_o	1	13995081	13995081	38.775
σ_ε^2	1	21423012	21423012	59.354
θ	1	7688143	7688143	21.301
σ_p^2	1	30567077	30567077	84.689
trend	1	293764	293764	0.814
n	1	1756950	1756950	4.868
s_o*σ_ε^2	1	11027381	11027381	30.552
s_o*θ	1	3034564	3034564	8.408
s_o*σ_p^2	1	15023376	15023376	41.623
s_o*trend	1	4337848	4337848	12.018
s_o*n	1	738311	738311	2.046
σ_ε^2*θ	1	7271112	7271112	20.145
σ_ε^2*σ_p^2	1	24596640	24596640	68.147
σ_ε^2*trend	1	77702	77702	0.215
σ_ε^2*n	1	1361306	1361306	3.772
θ*σ_p^2	1	5561343	5561343	15.408
θ*trend	1	12321	12321	0.034
θ*n	1	156	156	0
σ_p^2*trend	1	213444	213444	0.591
σ_p^2*n	1	1940449	1940449	5.376
trend*n	1	48290	48290	0.134
s_o*σ_ε^2*θ	1	2703558	2703558	7.490
s_o*σ_ε^2*σ_p^2	1	11987175	11987175	33.211
s_o*σ_ε^2*trend	1	2788900	2788900	7.727
s_o*σ_ε^2*n	1	521284	521284	1.444
s_o*θ*σ_p^2	1	2572816	2572816	7.128
s_o*θ*trend	1	801473	801473	2.221
s_o*θ*n	1	37346	37346	0.103
s_o*σ_p^2*trend	1	4832303	4832303	13.388
s_o*σ_p^2*n	1	808651	808651	2.240
s_o*trend*n	1	32761	32761	0.091
σ_ε^2*θ*σ_p^2	1	5218940	5218940	14.460
σ_ε^2*θ*trend	1	57002	57002	0.158
σ_ε^2*θ*n	1	3630	3630	0.010
σ_ε^2*σ_p^2*trend	1	41311	41311	0.114
σ_ε^2*σ_p^2*n	1	1512285	1512285	4.190
σ_ε^2*trend*n	1	81225	81225	0.225
θ*σ_p^2*trend	1	2809	2809	0.008
θ*σ_p^2*n	1	73984	73984	0.205
θ*trend*n	1	58202	58202	0.161
σ_p^2*trend*n	1	76591	76591	0.212
Residuals	22	7940566	360935	

larger MSPE than for the low level of σ_ε^2. For example, from Table 4, when all factors are low (000000) the bias and MSPE are −1.52 and 281 without adjusting for location error and −1.52 and 283 after adjusting for location error. Changing the measurement error to high (010000) yields bias and MSPE of −3.62 and 9749 without adjusting for location error and −3.63 and 9882 after adjusting for location error. This pattern holds regardless of the levels of the other factors.

(ii) The location-error variance plays an important role. In general, when σ_p^2 is low, the bias and MSPE of kriging without adjusting for location error and KAALE are quite similar, regardless of the levels of the other factors. For the example cited above, σ_p^2 is low in both situations; we have already seen that the MSPE for kriging without adjusting for location error and the MSPE for KAALE are similar. When σ_p^2 is high, the absolute bias is 15 percent to 50 percent greater and the MSPE is 10 percent to 30 percent greater for kriging without adjusting for location error. For example, for the treatment 001101, the bias and MSPE are −1.04 and 283 without adjusting for location error and −0.75 and 217 after adjusting for location error.

(iii) The absolute bias and MSPE are markedly greater for high levels of location error when the scale parameter θ is low (little spatial dependence) as compared to when it is high, regardless of whether or not an adjustment is made for location error. For example, consider kriging without adjusting for location error. Here the bias and MSPE for all levels low except for the location error (000100) are −3.21 and 1334; when the scale parameter θ is also high (001100), the bias and MSPE are −0.94 and 269. This agrees with simulation results reported by Griffith, Haining, and Arbia (1999) for error propagation in map overlay operations in a GIS.

(iv) There is not a consistent relationship between the location of the prediction site and either the bias or the MSPE. For example, when all levels are low (000000), the bias and MSPE are −1.52 and 281 without adjusting for location error. Changing the site to high (100000), the bias and MSPE are −1.94 and 845. However, when the measurement error is high and all other factors are low (010000), the bias and MSPE without adjusting for location error are −3.62 and 9749. Changing the location of the prediction site to high (110000), the bias and MSPE are −1.99 and 2667. Similar results occur after adjusting for location error. Thus, placing the site near the boundary does not lead to a predictable increase in bias or MSPE.

(v) Table 6 shows that interactions involving multiple factors have large variance ratios suggesting other relationships that are quite complex.

Results of the simulation study in $\Re^1$ indicate the importance of using the most accurate positioning instrument possible (that is, with small σ_p^2). For small levels of location uncertainty, kriging is not adversely affected by location error; however, as the level of location uncertainty increases, so does its effect on kriging. With large location uncertainty, KAALE performs markedly better than kriging without adjusting for location error.

5. SUMMARY

Location error that occurs because of the inability to match a predetermined collection of sites, or by knowingly introducing a location displacement, can be dealt with using the design model (5). Based on this model, we have developed expressions for the mean function and the covariance function that are adjusted for location error. We have used these results to develop kriging after adjusting for location error (KAALE). The *form* of the kriging-coefficients vector and the prediction variance are unchanged—what changes are the values of the mean function and the correlation matrix. When the location error is substantial, KAALE performs markedly better than kriging without adjusting for location error.

APPENDIX

To find the optimal kriging coefficients (, after adjusting for location error, we use the technique of Lagrange multipliers and write the constrained optimization problem as the minimization of

$$\mathrm{E}\left(\left(Y(\mathbf{s}_o)-\sum_{i=1}^{n}\lambda_i Z(\mathbf{s}_i)\right)^2\right)-2\sum_{j=1}^{q+1}m_{j-1}\left\{\sum_{i=1}^{n}\lambda_i\int f_{j-1}(\mathbf{s}_i+\mathbf{u})g(\mathbf{u})d\mathbf{u}-f_{j-1}(\mathbf{s}_o)\right\},$$

with respect to $\lambda_1,\ldots,\lambda_n,m_0,m_1,\ldots m_q$.

PROPOSITION 1: *Assume the design model for location error as given in (8), and assume B1 and B2. The vector of universal kriging coefficients (is given by*

$$\lambda' = \mathbf{b}_p'(F_p'K_p^{-1}F_p)^{-1}F_p'K_p^{-1} + \mathbf{k}_p'K_p^{-1},$$

and the vector of Lagrange multipliers $\mathbf{m}$ *is given by*

$$\mathbf{m}' = \sigma^2(\mathbf{f}(\mathbf{s}_o) - F_p'K_p^{-1}\mathbf{k}_p)'(F_p'K_p^{-1}F_p)^{-1},$$

where $\mathbf{b}_p = \mathbf{f}(\mathbf{s}_o) - F_p'K_p^{-1}\mathbf{k}_p$; F_p *is the* $n\times(q+1)$ *matrix of explanatory variables as defined in (17);* K_p *is the* $n \times n$ *correlation matrix given by (10); and* $\mathbf{k}_p$ *is the* $n \times 1$ *correlation vector given by (15).*

PROOF: Write $Z(\mathbf{s}) = \mathrm{E}(Z(\mathbf{s})) + (_p(\mathbf{s})$; $\mathbf{s}$ (D, where from (9),

$$\mathrm{E}(Z(\mathbf{s})) = \int \mu(\mathbf{s}+\mathbf{u})g(\mathbf{u})d\mathbf{u}.$$

We have seen in (10) that

$$\mathrm{E}(\delta_p(\mathbf{s})\cdot\delta_p(\mathbf{s}+\mathbf{h})) = C_p(\mathbf{h}) = \int\int C(\mathbf{h}+\mathbf{v}-\mathbf{u})g(\mathbf{u})g(\mathbf{v})d\mathbf{u}d\mathbf{v},$$

and in (15) that

$$\mathrm{E}(\delta_p(\mathbf{s})\cdot\delta(\mathbf{s}_o)) = \int C(\mathbf{s}_o-\mathbf{s}+\mathbf{u})g(\mathbf{u})d\mathbf{u}.$$

Now, because of the unbiasedness condition, $\mathrm{E}(Y(\mathbf{s}_o)) = \mathrm{E}\left(\sum_{i=1}^{n}\lambda_i Z(\mathbf{s}_i)\right)$, the squared prediction error is

$$\begin{aligned}\left(Y(\mathbf{s}_o)-\sum_{i=1}^{n}\lambda_i Z(\mathbf{s}_i)\right)^2 &= \left(\delta(\mathbf{s}_o)-\sum_{i=1}^{n}\lambda_i\delta_p(\mathbf{s}_i)\right)^2\\ &= \delta(\mathbf{s}_o)^2+\sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_i\lambda_j\delta_p(\mathbf{s}_i)\delta_p(\mathbf{s}_j)-2\sum_{i=1}^{n}\lambda_i\delta_p(\mathbf{s}_i)\delta(\mathbf{s}_o).\end{aligned}$$

Therefore, using the method of Lagrange multipliers to account for the unbiasedness condition, we wish to minimize,

$$C(\mathbf{0}) + \sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i \lambda_j C_p(\mathbf{s}_i - \mathbf{s}_j) - 2\sum_{i=1}^{n} \lambda_i \int C(\mathbf{s}_o - \mathbf{s}_i + \mathbf{u}) g(\mathbf{u}) d\mathbf{u}$$
$$- 2\sum_{j=1}^{q+1} m_{j-1} \left\{ \sum_{i=1}^{n} \lambda_i \int f_{j-1}(\mathbf{s}_i + \mathbf{u}) g(\mathbf{u}) d\mathbf{u} - f_{j-1}(\mathbf{s}_o) \right\};$$

see, for example, Cressie (1993, p. 153). The standard universal kriging results (for example, Cressie 1993, p. 153) yield,

$$\lambda' = \mathbf{b}_p'(F_p' K_p^{-1} F_p)^{-1} F_p' K_p^{-1} + \mathbf{k}_p' K_p^{-1},$$
$$\mathbf{m}' = \sigma^2 (\mathbf{f}(\mathbf{s}_o) - F_p' K_p^{-1} \mathbf{k}_p)' (F_p' K_p^{-1} F_p)^{-1}. \qquad \blacksquare$$

PROPOSITION 2: *Assume the design model for location error as given in (8), and assume B1 and B2 from above. The prediction variance is given by*

$$\sigma_k^2(\mathbf{s}_o) = \sigma^2 \left\{ 1 - \mathbf{k}_p' K_p^{-1} \mathbf{k}_p + \mathbf{b}_p' (F_p' K_p^{-1} F_p)^{-1} \mathbf{b}_p \right\} - \sigma_\varepsilon^2,$$

where $\mathbf{b}_p \equiv \mathbf{f}(\mathbf{s}_o) - F_p' K_p^{-1} \mathbf{k}_p$, F_p *is the* $n \times (q+1)$ *matrix of explanatory variables as defined in (17),* K_p *is the* $n \times n$ *correlation matrix given by (10),* $\mathbf{k}_p$ *is the* $n \times 1$ *correlation vector given by (15), and* σ_ε^2 *is the measurement-error variance.*

PROOF: Using the expression for squared prediction error given in the proof of Proposition 1, and the unbiasedness constraint, we have

$$\sigma_k^2(\mathbf{s}_o) = C(\mathbf{0}) - 2\sum_{i=1}^{n} \lambda_i C(\mathbf{s}_o - \mathbf{s}_i + \mathbf{u}) g(\mathbf{u}) d\mathbf{u} + \sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i \lambda_j C_p(\mathbf{s}_i - \mathbf{s}_j).$$

Applying standard universal kriging results (Cressie 1993, p. 154) to λ given in Proposition 1, we obtain

$$\sigma_k^2(\mathbf{s}_o) = \sigma^2 \left\{ 1 - \mathbf{k}_p' K_p^{-1} \mathbf{k}_p + \mathbf{b}_p' (F_p' K_p^{-1} F_p)^{-1} \mathbf{b}_p \right\} - \sigma_\varepsilon^2. \qquad \blacksquare$$

LITERATURE CITED

Arbia, G., D. Griffith, and R. Haining (1998). "Error Propagation Modelling in Raster GIS: Overlay Operations." *International Journal of Geographic Information Science* 12, 145–67.

Atkinson, P. M. (1997). "Simulating Location Error in Field-based Measurements of Reflectance." In *geoEnv I - Geostatistics for Environmental Applications*, edited by A. Soares, J. Gomez-Hernandez, and R. Froidevaux, pp. 297–308. Dordrecht: Kluwer Academic Publishers.

Berkson, J. (1950). "Are There Two Regressions?" *Journal of the American Statistical Association* 45, 164–80.

Bookstein, F. L. (1986). "Size and Shape Spaces for Landmark Data in Two Dimensions." *Statistical Science* 1, 181–242.

Clark, I. (1979). *Practical Geostatistics*. London: Applied Science Publishers.

Cressie, N. (1990). "The Origins of Kriging." *Mathematical Geology* 22, 239–52.

_____ (1993). *Statistics for Spatial Data,* revised edition. John Wiley & Sons, New York.

Fazekas, I., S. Baran, A. Kukush, and J. Lauridsen (1999). "Asymptotic Properties in Space and Time of an Estimator in Nonlinear Functional Errors-in-Variables Models." *Random Operations and Stochastic Equations* 7, 389–412.

Fazekas, I., and A. G. Kukush (2000). "Infill Asymptotics inside Increasing Domains for the Least Squares Estimator in Linear Models." *Statistical Inference for Stochastic Processes* 3, 199–223.

Fuller, W. A. (1987). *Measurement Error Models*. New York: John Wiley & Sons.

Geman, A., and D. Geman (1984). "Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images." *IEEE Transactions on Pattern Analysis and Machine Intelligence* PAMI-6, 721–41.

Goodchild, M. (1989). "Modeling Error in Objects and Fields." In *The Accuracy of Spatial Databases*, edited by M. Goodchild and S. Gopal, pp. 107–13. London: Taylor & Francis.

Griffith, D. A., R. P. Haining, and G. Arbia (1999). "Uncertainty and Error Propagation in Map Analyses Involving Arithmetic Operations: Inventory and Prospects." In *Spatial Accuracy Assessment: Land Information Uncertainty in Natural Resources*, edited by K. Lowell and A. Jaton, pp. 11–25. Chelsea, Mich.: Ann Arbor Press.

Haining, R. and Arbia, G. (1993). "Error Propagation through Map Operations." *Technometrics* 35, 293–305.

Heuvelink, G.B.M. (1996). "Identification of Field Attribute Error under Different Models of Spatial Variation." *International Journal of Geographic Information Systems* 10, 921–35.

Krumbein, W. C., and F. A. Graybill (1965). *An Introduction to Statistical Models in Geology*. New York: McGraw-Hill.

Lahiri, S. N., M. S. Kaiser, N. Cressie, and N. J. Hsu (1999). "Prediction of Spatial Distribution Functions Using Subsampling." *Journal of the American Statistical Association* 94, 86–110.

Lewis, A., and M. F. Hutchinson (2000). "From Data Accuracy to Data Quality: Using Spatial Statistics to Predict the Implications of Spatial Error in Point Data." In *Quantifying Spatial Uncertainty in Natural Resources: Theory and Applications for GIS and Remote Sensing*, edited by H. T. Mowrer and R. G. Congalton, pp. 17–35. Chelsea, Mich.: Ann Arbor Press.