 DOI: 10.1111/j.1541-0420.2011.01563.x

Filtered Kriging for Spatial Data with Heterogeneous Measurement Error Variances

William F. Christensen*

Department of Statistics, Brigham Young University, Provo, Utah 84602, U.S.A.
**email:* william@stat.byu.edu

Summary. When predicting values for the measurement-error-free component of an observed spatial process, it is generally assumed that the process has a common measurement error variance. However, it is often the case that each measurement in a spatial data set has a known, site-specific measurement error variance, rendering the observed process nonstationary. We present a simple approach for estimating the semivariogram of the unobservable measurement-error-free process using a bias adjustment of the classical semivariogram formula. We then develop a new kriging predictor that filters the measurement errors. For scenarios where each site's measurement error variance is a function of the process of interest, we recommend an approach that also uses a variance-stabilizing transformation. The properties of the heterogeneous variance measurement-error-filtered kriging (HFK) predictor and variance-stabilized HFK predictor, and the improvement of these approaches over standard measurement-error-filtered kriging are demonstrated using simulation. The approach is illustrated with climate model output from the Hudson Strait area in northern Canada. In the illustration, locations with high or low measurement error variances are appropriately down- or upweighted in the prediction of the underlying process, yielding a realistically smooth picture of the phenomenon of interest.

Key words: Heterogeneous variance measurement-error-filtered kriging; HFK; NARCCAP; Spatial prediction; Spatial smoothing.

1. Introduction

Spatial prediction based on properties of an observed spatial process is commonly known as kriging (Matheron, 1963). The early history of the development of kriging and geostatistics generally is discussed in Cressie (1990).

In its simplest form, the ordinary kriging (OK) predictor (see Cressie, 1993, Section 3.2) is an exact interpolator in that the spatial prediction for an observed location is simply the observed value for that location. However, there is often interest in predicting a measurement-error-free or filtered version of the observed process. Such an approach has been developed by Cressie (1993; Section 3.2.1) and further discussed by Schabenberger and Gotway (2005; Section 5.4.3) and Waller and Gotway (2004; Section 8.3.2). Measurement-error-free or measurement-error-filtered kriging (FK) requires that the measurement error variance is known and common across spatial locations. The FK predictor will "smooth" the observed measurements as it predicts the measurement-error-free component of the data.

It is often the case, however, that each measurement in a spatial data set has a measurement error variance that is unique but known. For example, this may occur whenever the spatial data are gathered using different measurement instruments with different measurement properties or when the spatial data set is itself a collection of location-specific parameter estimates with varying standard errors. Kleijnen and Van Beers (2005) use simulation to evaluate the properties of kriging under the heterogeneous variance assumption. They argue that OK is robust but conclude that optimal prediction under the heterogeneous variance assumption is still an open question.

As we later discuss in greater detail, the predominant challenge here is the nonstationarity of such an observed process. The semivariogram/covariance functions for the data are not estimable in the usual sense as in Cressie (1993, Section 2.4) because although the measurement-error-free version of the process may be stationary, the observed process is not. Hence, the observed data are not amenable to direct geostatistical modeling.

In this article, we present a simple approach for estimating the semivariogram/covariance function for the unobservable, measurement-error-free process and develop a new kriging predictor that filters the heterogeneous measurement errors. In Section 2, we consider spatial prediction in the presence of measurement error, reviewing the OK and FK predictors and introducing the heterogeneous variance measurement-error-filtered kriging (HFK) predictor. In Section 3, we propose a modification of HFK based on the variance-stabilizing transformation and trans-Gaussian kriging. This variance-stabilized HFK (or VHFK) approach is necessary whenever the observation process is correlated with the measurement error variance process. A simulation study is presented in Section 4 that illustrates the scenarios for which HFK and VHFK are superior to both OK and a simple adaptation of FK. In Section 5, we present a simple example using climate model output for mean summer temperatures in the Hudson Strait

in northern Canada. The final section contains discussion and conclusions.

2. Spatial Prediction in the Presence of Measurement Error

In this discussion, we consider an observable spatial process $Z(\mathbf{s})$, which is an error-contaminated version of the unobservable process $T(\mathbf{s})$. Specifically, our model of interest for the observation at $\mathbf{s}_i$ is:

$$Z(\mathbf{s}_i) = T(\mathbf{s}_i) + \epsilon(\mathbf{s}_i), \tag{1}$$

where $T(\mathbf{s}_i)$ is an intrinsically stationary spatial process with semivariogram $\gamma_T(\mathbf{s}_i - \mathbf{s}_j) = \frac{1}{2}\text{var}\{T(\mathbf{s}_i) - T(\mathbf{s}_j)\}$ and $\epsilon(\mathbf{s}_i)$ is a site-specific zero-mean error term with (known) variance of $\sigma^2(\mathbf{s}_i)$, and $\text{cov}\{\epsilon(\mathbf{s}_i), \epsilon(\mathbf{s}_{i'})\} = 0$ for $\mathbf{s}_i \neq \mathbf{s}_{i'}$. Finally, $T(\mathbf{s}_i)$ is uncorrelated with $\epsilon(\mathbf{s}_i)$. We have observations of $Z(\mathbf{s})$ for all locations in a sample $\mathbf{D} = \{\mathbf{s}_1, \ldots, \mathbf{s}_n\}$, where $\mathbf{D} \subset \mathbb{R}^d$. In essence, the intrinsic stationarity implies that both the mean and the nature of the spatial covariance remain constant throughout the random field of interest, with no systematic trends that would prevent the researcher from pooling information across regions of the random field to better understand spatial structure.

Spatial prediction (kriging) is traditionally considered to be a method for spatial interpolation. However, when measurement errors are known to contaminate the measured process $Z(\mathbf{s})$, there is often interest in predicting the smoother unobservable process $T(\mathbf{s})$ at both observed and unobserved locations. In this sense, kriging can be used as a spatial smoother. In this section, we consider traditional and newly proposed methods for spatial prediction.

2.1 *Ordinary Kriging*

When it can be assumed that $\epsilon(\mathbf{s}_i) = 0$ for all $\mathbf{s}_i$ (i.e., $Z(\mathbf{s}_i) = T(\mathbf{s}_i)$), the best linear unbiased predictor (BLUP) for $T(\mathbf{s}_0)$ (or $Z(\mathbf{s}_0)$) at location $\mathbf{s}_0$ given data $\mathbf{Z} = (Z(\mathbf{s}_1), \ldots, Z(\mathbf{s}_n))'$ is the OK predictor $\hat{Z}_{OK}(\mathbf{s}_0) = \boldsymbol{\lambda}'_{OK}\mathbf{Z}$, where the kriging coefficients ($\boldsymbol{\lambda}_{OK}$) are found using:

$$\begin{bmatrix} \boldsymbol{\lambda}_{OK} \\ m_{OK} \end{bmatrix} = \begin{bmatrix} \hat{\boldsymbol{\Gamma}}\mathbf{Z} & \mathbf{1} \\ \mathbf{1}' & 0 \end{bmatrix}^{-1} \begin{bmatrix} \hat{\boldsymbol{\gamma}}_{\mathbf{Z},Z(\mathbf{s}_0)} \\ 1 \end{bmatrix}, \tag{2}$$

where $\mathbf{1}$ is an n-vector of ones, the (i, j) element of the $n \times n$ matrix $\hat{\boldsymbol{\Gamma}}_{\mathbf{Z}}$ is $\hat{\gamma}_Z(\mathbf{s}_i - \mathbf{s}_j)$, and $\hat{\gamma}_Z(\mathbf{h})$ is the estimated semivariogram for $Z(\mathbf{s})$ (Matheron, 1963; Cressie, 1990, 1993, Section 3.2). The ith element of the vector $\hat{\boldsymbol{\gamma}}_{\mathbf{Z},Z(\mathbf{s}_0)}$ is technically the cross-semivariogram of $Z(\mathbf{s}_i)$ with $Z(\mathbf{s}_0)$ and is calculated using $\hat{\gamma}_Z(\mathbf{s}_i - \mathbf{s}_0)$ because of the stationarity of $Z(\mathbf{s})$ when $\epsilon(\mathbf{s}) = 0$. Note that the OK predictor is an exact interpolator so that for any $\mathbf{s}_i \in \mathbf{D}$, $\hat{Z}_{OK}(\mathbf{s}_i) = Z(\mathbf{s}_i)$. The quantity m_{OK} is a Lagrange multiplier that ensures that the kriging weights sum to one. The estimated kriging (prediction) variance for $\hat{Z}_{OK}(\mathbf{s}_0)$ is

$$\hat{\tau}^2_{OK}(\mathbf{s}_0) = \boldsymbol{\lambda}'_{OK}\,\hat{\boldsymbol{\gamma}}_{\mathbf{Z},Z(\mathbf{s}_0)} + m_{OK},$$

which is equal to 0 when $\mathbf{s}_0 \in \mathbf{D}$.

2.2 *Measurement-Error-Filtered Kriging*

When $\sigma^2(\mathbf{s})$ in (1) is a known measurement error variance σ^2 for all $\mathbf{s}$, then it is usually preferred to predict the value of $T(\mathbf{s})$, the uncontaminated component of $Z(\mathbf{s})$ (Cressie, 1993, Section 3.2.1; Waller and Gotway, 2004, Section 8.3.2; Schabenberger and Gotway, 2005, Section 5.4.3). Because the measurement error variance is common across sites, $Z(\mathbf{s})$ retains the intrinsic stationarity property of $T(\mathbf{s})$ and prediction of $T(\mathbf{s})$ from $\mathbf{Z}$ is still relatively straightforward, with

$$\gamma_Z(\mathbf{h}) = \begin{cases} \gamma_T(\mathbf{h}) + \sigma^2, & \mathbf{h} \neq \mathbf{0}, \\ 0, & \mathbf{h} = \mathbf{0}. \end{cases}$$

We note here that in general, the nugget effect in the semivariogram for $Z(\mathbf{s})$ is *not* equal to the measurement error variance and should not be treated as such. In general, the nugget effect is equal to the sum of the measurement error variance and the microscale variation (see Cressie, 1993, Section 3.1). Treating the nugget of $Z(\mathbf{s})$ as an estimate of σ^2 will systematically yield predictions for $T(\mathbf{s})$ that are oversmoothed (Cressie, 1993, Section 3.2.1).

The FK predictor for $T(\mathbf{s}_0)$ given $\mathbf{Z}$ is a BLUP that can be written as $\hat{T}_{FK}(\mathbf{s}_0) = \boldsymbol{\lambda}'_{FK}\mathbf{Z}$, where the kriging coefficients ($\boldsymbol{\lambda}_{FK}$) can be found using:

$$\begin{bmatrix} \boldsymbol{\lambda}_{FK} \\ \dot{m}_{FK} \end{bmatrix} = \begin{bmatrix} \hat{\boldsymbol{\Gamma}}_{\mathbf{Z}} & \mathbf{1} \\ \mathbf{1}' & 0 \end{bmatrix}^{-1} \begin{bmatrix} \dot{\boldsymbol{\gamma}}_{\mathbf{Z},T(\mathbf{s}_0)} \\ 1 \end{bmatrix}, \tag{3}$$

where $\hat{\boldsymbol{\Gamma}}_{\mathbf{Z}}$ is defined in Section 2.1 and the vector $\dot{\gamma}_{\mathbf{Z},T(\mathbf{s}_0)}$ is the cross-semivariogram of $\mathbf{Z}$ with $T(\mathbf{s}_0)$. The ith element of the $\dot{\boldsymbol{\gamma}}_{\mathbf{Z},T(\mathbf{s}_0)}$ can be calculated directly from the semivariogram for $Z(\mathbf{s})$ using

$$\begin{aligned} \dot{\gamma}_{Z(\mathbf{s}_i),T(\mathbf{s}_0)}(\mathbf{s}_i - \mathbf{s}_0) &= \frac{1}{2}\widehat{\text{var}}\{Z(\mathbf{s}_i) - T(\mathbf{s}_0)\} \\ &= \frac{1}{2}\widehat{\text{var}}\{T(\mathbf{s}_i) - T(\mathbf{s}_0)\} + \frac{1}{2}\text{var}\{\epsilon(\mathbf{s}_i)\} \\ &= \hat{\gamma}_T(\mathbf{s}_i - \mathbf{s}_0) + \frac{1}{2}\sigma^2 \\ &= \begin{cases} \hat{\gamma}_Z(\mathbf{s}_i - \mathbf{s}_0) - \frac{1}{2}\sigma^2, & \mathbf{s}_i \neq \mathbf{s}_0, \\ \frac{1}{2}\sigma^2, & \mathbf{s}_i = \mathbf{s}_0. \end{cases} \end{aligned}$$

The estimated kriging variance for $\hat{T}_{FK}(\mathbf{s}_0)$ is

$$\hat{\tau}^2_{FK}(\mathbf{s}_0) = \boldsymbol{\lambda}'_{FK}\,\dot{\boldsymbol{\gamma}}_{\mathbf{Z},T(\mathbf{s}_0)} + \dot{m}_{FK}. \tag{4}$$

However, note that the traditional FK formulas (Cressie, 1993, Section 3.2.1; Waller and Gotway, 2004, Section 8.3.2; Schabenberger and Gotway, 2005, Section 5.4.3) are

$$\begin{bmatrix} \boldsymbol{\lambda}_{FK} \\ m_{FK} \end{bmatrix} = \begin{bmatrix} \hat{\boldsymbol{\Gamma}}_{\mathbf{Z}} & \mathbf{1} \\ \mathbf{1}' & 0 \end{bmatrix}^{-1} \begin{bmatrix} \hat{\boldsymbol{\gamma}}_{\mathbf{Z},T(\mathbf{s}_0)} \\ 1 \end{bmatrix}, \tag{5}$$

where the ith element of $\hat{\boldsymbol{\gamma}}_{\mathbf{Z},T(\mathbf{s}_0)}$ is

$$\hat{\gamma}_{Z(\mathbf{s}_i),T(\mathbf{s}_0)}(\mathbf{s}_i - \mathbf{s}_0) = \begin{cases} \hat{\gamma}_Z(\mathbf{s}_i - \mathbf{s}_0), & \mathbf{s}_i \neq \mathbf{s}_0, \\ \sigma^2, & \mathbf{s}_i = \mathbf{s}_0. \end{cases}$$

Notwithstanding the superficial differences between (3) and (5), the standard BLUP-based predictor (3) is equivalent to the conventional FK formula (5). This is particularly important to note because we will use the former to develop a general kriging predictor in Section 2.3 for use with heterogeneous measurement error variances. The equivalence of (3) and (5) is demonstrated in Web Appendix A.

When predicting at previously measured locations in the presence of a known, common measurement error variance at each location, FK yields a smoother prediction and smaller mean squared prediction error (MSPE) than OK (because OK is an exact interpolator). When predicting at previously unmeasured locations, FK yields equivalent predictions to OK, but with smaller prediction variance.

2.3 *Heterogeneous Measurement-Error-Filtered Kriging*

When the known measurement error variance $\sigma^2(\mathbf{s}_i)$ depends upon the location $\mathbf{s}_i$, the problem of predicting $T(\mathbf{s}_0)$ is more complicated. Whereas $T(\mathbf{s})$ is an intrinsically stationary process, $Z(\mathbf{s})$ is not. Thus the semivariogram $\gamma_Z(\cdot)$ is not estimable in the usual sense (as in Cressie, 1993, Section 2.4) because the assumption of stationarity is not met. (In Section 4, we consider the further complication that the site-specific measurement error variances are themselves only estimates of the true values $\sigma^2(\mathbf{s}_i)$. For the present development of the new predictor, we assume that $\sigma^2(\mathbf{s}_i)$ is known.)

Notwithstanding the nonstationarity of $Z(\mathbf{s})$, for the approach we consider here, we use the standard empirical semivariogram formula applied to $Z(\mathbf{s})$ as a tool for estimating the semivariogram for $T(\mathbf{s})$. First, note that under the heterogeneous variance model (1),

$$\begin{aligned} E\left\{\frac{1}{2}(Z(\mathbf{s}_i)-Z(\mathbf{s}_j))^2\right\} &= \frac{1}{2}E\big\{[T(\mathbf{s}_i)-T(\mathbf{s}_j) \\ &\qquad +\epsilon(\mathbf{s}_i)-\epsilon(\mathbf{s}_j)]^2\big\} \\ &= \gamma_T(\mathbf{s}_i-\mathbf{s}_j) \\ &\qquad +\frac{\sigma^2(\mathbf{s}_i)+\sigma^2(\mathbf{s}_j)}{2}, \end{aligned}$$

where $\gamma_T(\mathbf{s}_i-\mathbf{s}_j)$ is the semivariogram for the unobserved process $T(\mathbf{s})$. So,

$$\begin{aligned} E\{\hat{\gamma}_Z(\mathbf{h})\} &= E\left\{\frac{1}{2|N(\mathbf{h})|}\sum_{N(\mathbf{h})}(Z(\mathbf{s}_i)-Z(\mathbf{s}_j))^2\right\} \\ &= \begin{cases} \gamma_T(\mathbf{h})+\dfrac{1}{2|N(\mathbf{h})|}\displaystyle\sum_{N(\mathbf{h})}(\sigma^2(\mathbf{s}_i)+\sigma^2(\mathbf{s}_j)), & \mathbf{h}\neq\mathbf{0}, \\ 0, & \mathbf{h}=\mathbf{0}, \end{cases} \end{aligned} \tag{6}$$

where $N(\mathbf{h})$ is the set of points $(i,\ j)$ such that $\mathbf{s}_i-\mathbf{s}_j$ is within some tolerance region $\mathcal{T}(\mathbf{h})\subset\mathbb{R}^d$ centered at $\mathbf{h}$, and $|N(\mathbf{h})|$ is the number of elements in $N(\mathbf{h})$. The bias term

$$\frac{1}{2|N(\mathbf{h})|}\sum_{N(\mathbf{h})}(\sigma^2(\mathbf{s}_i)+\sigma^2(\mathbf{s}_j))$$

in (6) can be estimated directly for any value $\mathbf{h}$. However, this method-of-moments approach would require that $|N(\mathbf{h})|$ is large enough to yield a stable estimate of the bias for all considered tolerance regions $N(\mathbf{h})$. Fortunately, the bias can be estimated with $\frac{1}{n}\sum_{\ell=1}^{n}\sigma^2(\mathbf{s}_\ell)$ for all $\mathbf{h}$ when the following condition holds:

$$\mathrm{E}\left\{\frac{1}{2}(\sigma^2(\mathbf{s}_i)+\sigma^2(\mathbf{s}_j))\,|\,\mathbf{s}_i-\mathbf{s}_j\in\mathcal{T}(\mathbf{h})\right\}=\frac{1}{n}\sum_{\ell=1}^{n}\sigma^2(\mathbf{s}_\ell)\ \text{for all}\ \mathbf{h}. \tag{7}$$

This condition states that the expected value of $\frac{1}{2}(\sigma^2(\mathbf{s}_i)+\sigma^2(\mathbf{s}_j))$ is simply the mean of the $\sigma^2(\mathbf{s})$ process, regardless of the length or direction of the vector of separation between the pair of points $(\mathbf{s}_i,\ \mathbf{s}_j)$. The condition will be met whenever the large (and small) $\sigma^2(\mathbf{s}_i)$ values are spread throughout the region of interest in $\mathbb{R}^d$. However, the condition is quite unrestrictive and would even hold for cases where high error variance observations are found only in several randomly located clusters in an approximately uniformly spread set of points $\mathbf{D}$ in $\mathbb{R}^d$. For an isotropic process, this condition can be checked by plotting $\|\mathbf{s}_i-\mathbf{s}_j\|$ against $\frac{1}{2}(\sigma^2(\mathbf{s}_i)+\sigma^2(\mathbf{s}_j))$ for all pairs of points in $\mathbf{D}$, looking for evidence of a change in the conditional mean of $\frac{1}{2}(\sigma^2(\mathbf{s}_i)+\sigma^2(\mathbf{s}_j))$ given $\|\mathbf{s}_i-\mathbf{s}_j\|$. (We note that small changes in the conditional mean for pairs of points with $\|\mathbf{s}_i-\mathbf{s}_j\|$ greater than the effective range of $Z(\mathbf{s})$ will have little effect on this bias term approximation.) Under condition (7), the expectation in (6) simplifies to

$$E\{\hat{\gamma}_Z(\mathbf{h})\}\approx\begin{cases}\gamma_T(\mathbf{h})+\dfrac{1}{n}\displaystyle\sum_{\ell=1}^{n}\sigma^2(\mathbf{s}_\ell), & \mathbf{h}\neq\mathbf{0}, \\ 0, & \mathbf{h}=\mathbf{0}.\end{cases} \tag{8}$$

Because condition (7) is so broadly applicable and will yield low-variance estimates of the bias term in (6), we employ it throughout our discussion, simulations, and data-analysis application.

Equation (8) implies that $\gamma_T(\mathbf{h})$ can be estimated using a bias-adjusted version of $\hat{\gamma}_Z(\mathbf{h})$, the empirical semivariogram for $Z(\mathbf{s})$. Further, the bias adjustment of $\hat{\gamma}_Z(\mathbf{h})$ is the same quantity for all values of $\mathbf{h}\neq\mathbf{0}$. Specifically,

$$\hat{\gamma}_T^*(\mathbf{h})\approx\begin{cases}\hat{\gamma}_Z(\mathbf{h})-\dfrac{1}{n}\displaystyle\sum_{\ell=1}^{n}\sigma^2(\mathbf{s}_\ell), & \mathbf{h}\neq\mathbf{0}, \\ 0, & \mathbf{h}=\mathbf{0},\end{cases} \tag{9}$$

where the asterisk indicates the use of a bias-correction term in formulating the estimate. Note that when $\sigma^2(\mathbf{s})=\sigma^2$ and FK (as in Section 2.2) is used, $\hat{\gamma}_Z(\mathbf{h})$ is similarly biased for $\gamma_T(\mathbf{h})$, but the stationarity implied by $\sigma^2(\mathbf{s})=\sigma^2$ allows for a simple adjustment of the kriging equations via (5).

The development of the HFK predictor is the same regardless of whether the bias in $\hat{\gamma}_Z(\mathbf{h})$ is adjusted separately for each choice of $\mathbf{h}$ or if the bias adjustment is the same for all $\mathbf{h}$ (using (9)). The value of the condition (7) and the implied bias-adjusted estimated variogram for $T(\mathbf{s})$ in (9) is that the bias adjustment in (9) is simple and allows for more observations to estimate it, yielding a more stable prediction of $\gamma_T(\mathbf{s})$ than an estimate that requires a separate bias adjustment for each chosen value of $\mathbf{h}$.

In practice, $\hat{\gamma}_T^*(\mathbf{h})$ will be constructed from the fitted semivariogram for $Z(\mathbf{s})$. Suppose, for example, that the semivariogram model fitted to $\hat{\gamma}_Z(\mathbf{h})$ is $\hat{\gamma}_Z(\mathbf{h};\theta)$, where $\boldsymbol{\theta}$ contains parameters defining the parametric model (e.g., exponential or spherical). If $\hat{\gamma}_Z(\mathbf{h};c_Z,r_Z,v_Z)$ has a nugget of c_Z, range of r_Z, and partial sill of v_Z, then an estimated semivariogram model for $T(\mathbf{s})$ is

$$\hat{\gamma}_T^*(\mathbf{h};c_T,r_T,v_T)=\hat{\gamma}_Z\left(\mathbf{h};c_Z-\frac{1}{n}\sum_{\ell=1}^{n}\sigma^2(\mathbf{s}_\ell),r_Z,v_Z\right).$$

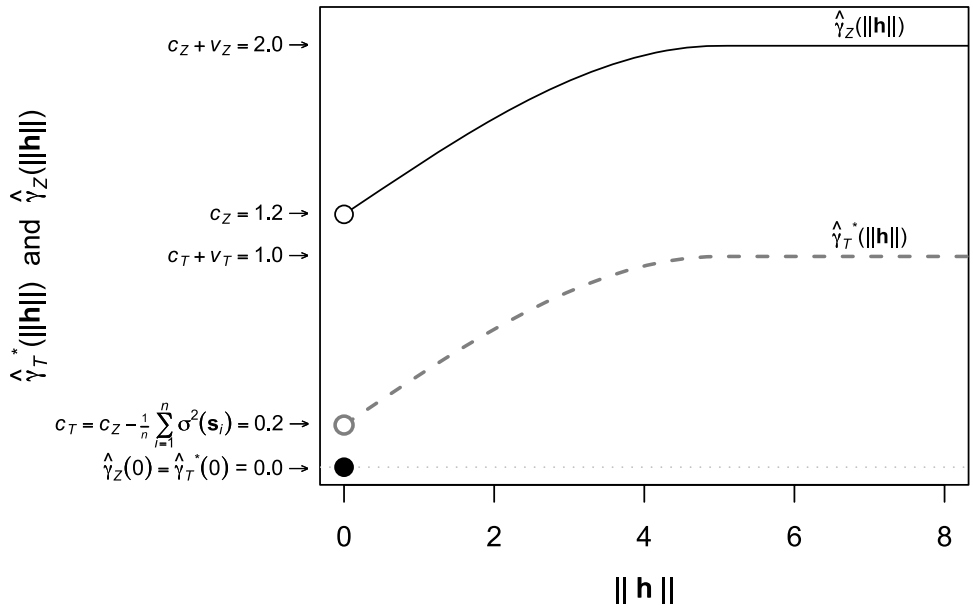

Figure 1. Comparison of $\hat{\gamma}_Z(\mathbf{h})$ and $\hat{\gamma}_T^*(\mathbf{h})$ when the nugget of $\gamma_Z(\mathbf{h})$ is $c_Z = 1.2$, the partial sill is $v_Z = 0.8$, the range is $r_Z = 5$, and the mean measurement error variance for the observations in $\mathbf{D}$ is $\frac{1}{n}\sum_{i=1}^{n}\sigma^2(\mathbf{s}_i) = 1.0$. Thus, $c_T = 0.2$, $v_T = v_Z = 0.8$, and $r_T = r_Z = 5$.

Note that, in general, the range and partial sill for $T(\mathbf{s})$ are unaffected by the bias correction, so $r_T = r_Z$ and $v_T = v_Z$. Example semivariograms for $Z(\mathbf{s})$ and $T(\mathbf{s})$ are illustrated in Figure 1. Note also that under model (1), $c_Z - \frac{1}{n}\sum_{\ell=1}^{n}\sigma^2(\mathbf{s}_\ell) \geqslant 0$, but the estimated nugget may occasionally be less than the average measurement error variance due to sampling error or semivariogram estimation error. When $c_Z - \frac{1}{n}\sum_{\ell=1}^{n}\sigma^2(\mathbf{s}_\ell) < 0$, the nugget for the semivariogram of $T(\mathbf{s})$ (denoted c_T) can be set to zero and the partial sill (v_T) set to equal $v_Z + c_Z - \frac{1}{n}\sum_{\ell=1}^{n}\sigma^2(\mathbf{s}_\ell)$ so that the sill (sum of the nugget plus partial sill) is still reduced by $\frac{1}{n}\sum_{\ell=1}^{n}\sigma^2(\mathbf{s}_\ell)$. Note that the observed nugget for the estimated semivariogram of the $Z(\mathbf{s})$ process is the sum of the microscale variation associated with $T(\mathbf{s})$ (see Cressie, 1993, Section 3.1) and the mean measurement error variance associated with $Z(\mathbf{s})$.

The predictors in the OK and FK settings are examples of BLUPs, which are generally constructed using the covariances among all pairs of observations $(Z(\mathbf{s}_i), Z(\mathbf{s}_j))$ and the covariances between each observation and the process of interest at $\mathbf{s}_0$ (either $Z(\mathbf{s}_0)$ or $T(\mathbf{s}_0)$). In (5), the semivariogram-based formula for the FK weights consists of an inverse of a matrix multiplied by a column vector. Note that in the FK setting, the (i, j) element of $\hat{\mathbf{\Gamma}}_\mathbf{Z}$ is equal to

$$\frac{1}{2}\text{var}\{Z(\mathbf{s}_i) - Z(\mathbf{s}_j)\} = \gamma_T(\mathbf{s}_i - \mathbf{s}_j) + I_{(i\neq j)}\sigma^2,$$

which is directly estimable using $\hat{\gamma}_Z(\mathbf{s}_i - \mathbf{s}_j)$ because of the stationarity of $Z(\mathbf{s})$. In contrast, when following the standard approach for obtaining a BLUP, the HFK predictor will *not* be a simple function of $\hat{\gamma}_Z(\mathbf{s}_i - \mathbf{s}_j)$ because of the nonstationarity of $Z(\mathbf{s})$. Specifically, the HFK weights for the predictor $\hat{T}_{HFK}(\mathbf{s}_0) = \boldsymbol{\lambda}'_{HFK}\mathbf{Z}$ are found using:

$$\begin{bmatrix} \boldsymbol{\lambda}_{HFK} \\ m_{HFK} \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{\Gamma}}^*_\mathbf{Z} & \mathbf{1} \\ \mathbf{1}' & 0 \end{bmatrix}^{-1} \begin{bmatrix} \hat{\gamma}^*_{\mathbf{Z},T(\mathbf{s}_0)} \\ 1 \end{bmatrix}, \tag{10}$$

where the (i, j) element of $\hat{\mathbf{\Gamma}}^*_Z$ is

$$\begin{aligned} &\frac{1}{2}\widehat{\text{var}}\{Z(\mathbf{s}_i) - Z(\mathbf{s}_j)\} \\ &\quad = \begin{cases} \frac{1}{2}\widehat{\text{var}}\{T(\mathbf{s}_i) - T(\mathbf{s}_j)\} + \frac{1}{2}\text{var}\{\epsilon(\mathbf{s}_i) - \epsilon(\mathbf{s}_j)\}, & \mathbf{s}_i \neq \mathbf{s}_j \\ 0, & \mathbf{s}_i = \mathbf{s}_j \end{cases} \\ &\quad = \hat{\gamma}^*_T(\mathbf{s}_i - \mathbf{s}_j) + I_{(i\neq j)}\frac{\sigma^2(\mathbf{s}_i) + \sigma^2(\mathbf{s}_j)}{2}, \end{aligned}$$

the ith element of $\hat{\gamma}^*_{\mathbf{Z},T(\mathbf{s}_0)}$ is

$$\begin{aligned} \frac{1}{2}\widehat{\text{var}}\{Z(\mathbf{s}_i) - T(\mathbf{s}_0)\} &= \frac{1}{2}\widehat{\text{var}}\{T(\mathbf{s}_i) - T(\mathbf{s}_0)\} + \frac{1}{2}\text{var}\{\epsilon(\mathbf{s}_i)\} \\ &= \hat{\gamma}^*_T(\mathbf{s}_i - \mathbf{s}_0) + \frac{\sigma^2(\mathbf{s}_i)}{2}, \end{aligned}$$

and the indicator function I_A equals 1 when condition A is true, and equals 0 otherwise. A comparison of HFK with FK and a demonstration that HFK reduces to FK in the case of $\sigma^2(\mathbf{s}) = \sigma^2$ is given in Web Appendix B.

The kriging variance for $\hat{T}_{HFK}(\mathbf{s}_0)$ is

$$\tau^2_{HFK}(\mathbf{s}_0) = \boldsymbol{\lambda}'_{HFK}\,\hat{\gamma}^*_{\mathbf{Z},T(\mathbf{s}_0)} + m_{HFK}. \tag{11}$$

When $\sigma^2(\mathbf{s}_i) = \sigma^2$ for all $\mathbf{s}_i$, we can use an argument similar to that used in Web Appendix A to show that (4), (A.2) in Web Appendix A, and (11) are equivalent.

Web Appendix C presents the HFK predictors in terms of covariograms instead of semivariograms.

3. Variance-Stabilized HFK When $\sigma^2(\mathbf{s}_i)$ Is Correlated with $T(\mathbf{s}_i)$

In many scenarios, the location-specific measurement error variance $\sigma^2(\mathbf{s}_i)$ is itself a function of $T(\mathbf{s}_i)$. This is problematic because the development of the estimate of $\gamma_T(\mathbf{h})$ in (9) and the kriging equations in (10) require that $T(\mathbf{s})$ and $\epsilon(\mathbf{s})$ be uncorrelated. If $\sigma^2(\mathbf{s}_i)$ is a function of $T(\mathbf{s})$, then in general, $\text{cor}\{T(\mathbf{s}),\, \epsilon(\mathbf{s})\} \neq 0$ and the approach for constructing the HFK predictor will be flawed. Specifically, a relationship between $\sigma^2(\mathbf{s})$ and $T(\mathbf{s})$ can lead to bias in $\hat{T}_{HFK}(\mathbf{s}_0)$ because HFK downweights the influence of the high-variance observations in the linear combination $\hat{T}_{HFK}(\mathbf{s}_0) = \boldsymbol{\lambda}'_{HFK}\mathbf{Z}$. For example, there are many common scenarios in environmental monitoring where the unobservable quantity of interest $T(\mathbf{s})$ is positively correlated with $\sigma^2(\mathbf{s})$. In the area of receptor modeling of air-quality data, ambient measures of particulate matter abundances are generally associated with available measurement error variances that tend to increase with the measured abundance (Lingwall, Christensen, and Reese, 2008). The HFK predictor for such data will exhibit downward bias, particularly among the regions of the observation space where $T(\mathbf{s})$ and $\sigma^2(\mathbf{s})$ are high.

When $\sigma^2(\mathbf{s})$ is correlated with $T(\mathbf{s})$, we consider the use of variance-stabilizing transformations proposed by Box and Cox (1964) and developed by many other authors including Hocking (2003; Section 3.3), Box, Hunter, and Hunter (1978; Section 7.8), and Kuehl (1994, Section 4.6). Suppose that the model relating $\sigma^2(\mathbf{s})$ to $Z(\mathbf{s})$ is

$$\sigma^2(\mathbf{s}_i) = \alpha^2(Z(\mathbf{s}_i))^{2\beta} \tag{12}$$

so that

$$\log\sigma(\mathbf{s}_i) = \log\alpha + \beta\log Z(\mathbf{s}_i).$$

The coefficients $\hat{\alpha}$ and $\hat{\beta}$ can be obtained via regression. Then, following Hocking (2003) and Kuehl (1994), define the transformed measurement $\tilde{Z}(\mathbf{s}_i)$ using

$$\tilde{Z}(\mathbf{s}_i) = f(Z(\mathbf{s}_i)) = \frac{1}{\hat{\alpha}(1-\hat{\beta})}(Z(\mathbf{s}_i))^{1-\hat{\beta}} \tag{13}$$

and define

$$\begin{aligned}\tilde{\sigma}^2(\mathbf{s}_i) &= \text{var}\{\tilde{Z}(\mathbf{s}_i)\} = \sigma^2(\mathbf{s}_i) \times [f'(Z(\mathbf{s}_i))]^2 \\ &= \sigma^2(\mathbf{s}_i) \times \left[\frac{1}{\hat{\alpha}}(Z(\mathbf{s}_i))^{-\hat{\beta}}\right]^2. \end{aligned} \tag{14}$$

Note that if model (12) is valid, the correlation between $\tilde{Z}(\mathbf{s})$ and $\tilde{\sigma}^2(\mathbf{s})$ will be roughly zero.

Thus, the decomposition of signal and measurement error defined in our original data model (1) is best applied to the transformed data $\tilde{Z}(\mathbf{s})$, $\mathbf{s} \in \mathbf{D}$, where components $\tilde{T}(\mathbf{s})$ and $\tilde{\epsilon}(\mathbf{s})$ are approximately uncorrelated. The estimator $\hat{\tilde{T}}(\mathbf{s}_0)$ is found by applying HFK to $\tilde{Z}(\mathbf{s}_i)$ and $\tilde{\sigma}^2(\mathbf{s}_i)$ in (13) and (14). Define $\tilde{m}_{HFK}$ and $\tilde{\tau}^2(\mathbf{s}_0)$ to be the Lagrange multiplier and estimated kriging variance obtained from (10) and (11), respectively.

To transform back to the original scale, we define

$$g(\tilde{Z}(\mathbf{s})) = f^{-1}(\tilde{Z}(\mathbf{s})) = \left(\hat{\alpha}(1-\hat{\beta})\tilde{Z}(\mathbf{s})\right)^{\frac{1}{1-\hat{\beta}}},$$

$$g'(\tilde{Z}(\mathbf{s})) = \hat{\alpha}\left(\hat{\alpha}(1-\hat{\beta})\tilde{Z}(\mathbf{s})\right)^{\frac{\hat{\beta}}{1-\hat{\beta}}},$$

and

$$g''(\tilde{Z}(\mathbf{s})) = \hat{\alpha}^2\hat{\beta}\left(\hat{\alpha}(1-\hat{\beta})\tilde{Z}(\mathbf{s})\right)^{\frac{2\hat{\beta}-1}{1-\hat{\beta}}}.$$

Let $\hat{C}_Z(\mathbf{h}; c_Z, r_Z, v_Z)$ be the covariogram model fitted to $\hat{C}_Z(\mathbf{h})$, with a nugget of c_Z, range of r_Z, and partial sill of v_Z. Then an estimated covariogram model for $T(\mathbf{s})$ is

$$\hat{C}^*_T(\mathbf{h}; c_T, r_T, v_T) = \hat{C}_Z\left(\mathbf{h}; c_Z - \frac{1}{n}\sum_{\ell=1}^{n}\sigma^2(\mathbf{s}_\ell), r_Z, v_Z\right),$$

and $\hat{\mathbf{C}}^*_Z$ is an $n \times n$ matrix with $(i,\, j)$ element equal to $\hat{C}^*_T(\mathbf{s}_i - \mathbf{s}_j) + I_{(i=j)}\sigma^2(\mathbf{s}_i)$. (See Web Appendix C for additional discussion of HFK using covariograms.) Then, using trans-Gaussian kriging (Cressie, 1993, Section 3.2.2), we define the VHFK predictor as

$$\hat{T}_{VHFK}(\mathbf{s}_0) = g\left(\hat{\tilde{T}}(\mathbf{s}_0)\right) + g''(\hat{\mu}_{\tilde{T}}) \times \left(\tilde{\tau}^2(\mathbf{s}_0)/2 - \tilde{m}_{HFK}\right), \tag{15}$$

where $\hat{\mu}_{\tilde{T}} = \mathbf{1}'\tilde{\mathbf{C}}^{*-1}_Z\tilde{\mathbf{Z}}/(\mathbf{1}'\tilde{\mathbf{C}}^{*-1}_Z\mathbf{1})$. The second term in (15) is an adjustment that reduces the bias of the predictor. Schabenberger and Gotway (2005; Section 5.6.2) discuss scenarios in which the adjustment may not be needed, drawing upon the discussion of Zimmerman and Cressie (1992). As an alternative to (15), we define the simpler variance-stabilized predictor

$$\ddot{T}_{VHFK}(\mathbf{s}_0) = g\left(\hat{\tilde{T}}(\mathbf{s}_0)\right). \tag{16}$$

Applying the results of Cressie (1993, Section 3.2.2) to the VHFK predictor, we obtain the estimated kriging variance for (15) and (16)

$$\hat{\tau}^2_{VHFK}(\mathbf{s}_0) = (g'(\hat{\mu}_{\tilde{T}}))^2 \times \tilde{\tau}^2(\mathbf{s}_0), \tag{17}$$

where $g'(\cdot), \hat{\mu}_{\tilde{T}}$, and $\tilde{\tau}^2(\mathbf{s}_0)$ are defined as in (15).

4. Simulation Studies

We use a computer simulation experiment to evaluate the performance of HFK in the scenarios where (1) $\sigma^2(\mathbf{s}_i)$ is uncorrelated with $T(\mathbf{s}_i)$ (using the HFK predictor), and (2) $\sigma^2(\mathbf{s}_i)$ is correlated with $T(\mathbf{s}_i)$ (using the variance-stabilized HFK or VHFK predictor). In all scenarios, we will compare HFK or VHFK (discussed in Sections 2.3 and 3) with both OK and FK (discussed in Sections 2.1 and 2.2). Because FK is not directly applicable to the heterogeneous measurement error variance scenario, when calculating FK predictions we use the mean of the assumed measurement error variances as the "common" measurement error variance. (Note that this mean measurement error variance is equivalent to $\frac{1}{n}\sum_{\ell=1}^{n}\sigma^2(\mathbf{s}_\ell)$, which is used as the bias adjustment in (9).) Although this yields only an approximation of the "common" measurement error variance, it allows us to compare our new predictors against another method that accounts for the presence of measurement error. For the case of $\sigma^2(\mathbf{s}_i)$ correlated with $T(\mathbf{s}_i)$, we use the variance-stabilizing adjustment of HFK (VHFK) discussed in Section 3. When generating spatial processes and when fitting semivariograms, we use isotropic models, but the isotropy assumption is not necessary for any of the discussed predictors. When using HFK or VHFK, we use the simple bias-adjusted semivariogram estimator $\hat{\gamma}^*_T(\mathbf{h})$ given in (9) because condition (7) is satisfied in the simulations.

We note here that although FK can be used in the heterogeneous measurement error variance scenario by using the mean of the measurement error process $\sigma^2(\mathbf{s})$ as an estimate of the common measurement error variance, there are clear hazards in taking such an approach. First, as we show below, the HFK/VHFK approach has much smaller MSPE in cases where the location-to-location variation in measurement error variance is large relative to the average measurement error variance. Second, the use of FK in the heterogeneous measurement error variance scenario will yield prediction variances that are incorrect—too small in the locations with above-average measurement error variances and too large in the locations with below-average measurement error variances.

4.1 $\sigma^2(\mathbf{s}_i)$ *Is Uncorrelated with* $T(\mathbf{s}_i)$

In this simulation, we generate data from (1) with $\mathbf{D}$ comprising 400 observations on a 20×20 grid. The signal $T(\mathbf{s})$ is generated from a normal distribution with the spatial correlation governed by an isotropic spherical semivariogram with nugget of 0.2, partial sill of 0.8, and range of 5 units. Various values for the relative size of the nugget were explored, with little effect on outcomes. The values of $\sigma^2(\mathbf{s})$ were independent and identically distributed draws from a lognormal distribution with mean of μ_{σ^2} and coefficient of variation (CV) of κ. Thus, κ controls the variability of the measurement error variances across the region of interest, with $\kappa = 0.5$ denoting that the standard deviation of the values of $\sigma^2(\mathbf{s})$ is 50% of the value of the mean measurement error variance μ_{σ^2}. Figure 1 illustrates examples of the semivariograms for $Z(\mathbf{s})$ and $T(\mathbf{s})$ that are used in this simulation, with $\mu_{\sigma^2} = 1$ and $\kappa = 1$.

We expect that as κ increases, the advantage of $\hat{T}_{HFK}(\mathbf{s})$ over $\hat{T}_{FK}(\mathbf{s})$ should generally increase. The process of generating from a lognormal with specified mean and variance is described in detail in Lingwall and Christensen (2007). We consider drawing the value of $\epsilon(\mathbf{s})$ from a normal distribution with mean of zero and variance of $\sigma^2(\mathbf{s})$, and from a right-skewed distribution with mean of zero and variance of $\sigma^2(\mathbf{s})$. The right-skewed distribution is a "shifted lognormal" distribution obtained by drawing values from a lognormal distribution with a mean of 1 and variance of $\sigma^2(\mathbf{s})$, and then subtracting 1 from each value. The skewed distribution is evaluated to identify whether or not a small number of locations with very large measurement errors affect the relative performance of the estimators.

Instead of using the true measurement error variance process $\sigma^2(\mathbf{s})$ when predicting $T(\mathbf{s})$, we incorporate the additional real-world complication of errors in the "known" measurement error variances. That is, we create a set of assumed measurement error variances $\breve{\sigma}^2(\mathbf{s})$, $\mathbf{s} \in \mathbf{D}$, where $\breve{\sigma}^2(\mathbf{s}_i)$ is a draw from a lognormal distribution with mean of $\sigma^2(\mathbf{s}_i)$ and CV of ϕ. Thus the quality of our assumed measurement error variances as surrogates for the true values of $\sigma^2(\mathbf{s})$ is governed by ϕ, where $\phi = 0.5$ implies that the assumed measurement error variances tend to be off by roughly 50% of the true value. We expect that as ϕ increases, the advantage of $\hat{T}_{HFK}(\mathbf{s})$ over $\hat{T}_{FK}(\mathbf{s})$ should decrease, because site-specific information about measurement error variance will be of less value.

When obtaining estimates of the semivariograms for calculating the predictions for $T(\mathbf{s})$, we assume the correct (isotropic spherical) parametric form, but then estimate the semivariogram parameters using the weighted least squares approach of Cressie (1985) as applied to the simulated data set. For each of the 100 scenarios considered, 200 data sets were simulated and predictions of $T(\mathbf{s})$ were obtained using each method. Because $\hat{T}_{HFK}(\mathbf{s})$ and $\hat{T}_{FK}(\mathbf{s})$ were each vastly superior to $\hat{T}_{OK}(\mathbf{s})$ when the average measurement error (μ_{σ^2}) is nonzero, in this simulation we focus only on $\hat{T}_{HFK}(\mathbf{s})$ and $\hat{T}_{FK}(\mathbf{s})$, where the HFK predictor uses $Z(\mathbf{s}_i)$ and $\breve{\sigma}^2(\mathbf{s}_i)$, and the FK predictor uses $Z(\mathbf{s}_i)$ and $\frac{1}{n}\sum_{\ell=1}^{n}\breve{\sigma}^2(\mathbf{s}_\ell)$ as the estimate of the "common" measurement error variance.

Table 1

Simulation results when $\sigma^2(\mathbf{s}_i)$ is uncorrelated with $T(\mathbf{s}_i)$, and $\epsilon(\mathbf{s}_i)$ is drawn from a distribution with variance $\sigma^2(\mathbf{s}_i)$. Table values are the MSPE values for $\hat{T}_{HFK}(\mathbf{s})$ using (10) relative to the MSPE for $\hat{T}_{FK}(\mathbf{s})$ using (5). Simulation parameters include: μ_{σ^2} (the average size of measurement variances), κ (governing the variability of measurement error variances across sites), and ϕ (governing the variability of $\breve{\sigma}^2(\mathbf{s})$ as an estimate of the true measurement error variance $\sigma^2(\mathbf{s})$).

		κ				
ϕ	μ_{σ^2}	0.1	0.25	0.5	1	1.5
0.1	0.1	1.00	0.99	0.97	0.90	0.84
0.1	0.25	1.00	0.99	0.96	0.86	0.78
0.1	0.5	1.00	0.99	0.95	0.83	0.74
0.1	1	1.00	0.98	0.94	0.83	0.71
0.1	1.5	1.00	0.99	0.92	0.81	0.76
0.5	0.1	1.02	1.02	1.00	0.93	0.87
0.5	0.25	1.04	1.03	0.99	0.87	0.81
0.5	0.5	1.05	1.04	0.99	0.86	0.77
0.5	1	1.06	1.02	0.99	0.87	0.76
0.5	1.5	1.06	1.05	0.99	0.87	0.78

Because the results obtained from the simulations with $\epsilon(\mathbf{s})$ drawn from a normal distribution and from a shifted lognormal distribution are nearly identical, we combine the results from the two simulations in one discussion. Table 1 gives the MSPE of the HFK prediction of $T(\mathbf{s})$ (for all $\mathbf{s} \in \mathbf{D}$) divided by the MSPE of the FK prediction. When the available measurement error variances $\breve{\sigma}^2(\mathbf{s}_i)$ are reasonably accurate ($\phi = 0.1$) and there is some variability in the measurement error variances across the spatial field ($\kappa \geqslant 0.25$), HFK predicts $T(\mathbf{s})$ with a MSPE that is smaller than the MSPE for FK. We see that when the measurement error variances are uncorrelated with the process of interest, using HFK instead of FK reduces the MSPE by as much as 32%. For the values considered in this simulation, as the measurement error's mean (μ_{σ^2}) and CV (ϕ) increases, the HFK predictor generally becomes more superior. Note that because the sill of $T(\mathbf{s})$ was equal to 1, one can use the values of μ_{σ^2} in the simulations to identify ranges of values for the ratio

$$\frac{\mu_{\sigma^2}}{\text{var}\{T(\mathbf{s})\}},$$

where one would expect to find HFK superior to FK. When $\sigma^2(\mathbf{s}_i)$ is uncorrelated with $T(\mathbf{s}_i)$ and the available measurement error variances $\breve{\sigma}^2(\mathbf{s}_i)$ are reasonably accurate ($\phi = 0.1$),

Table 2

Simulation results when $\sigma^2(\mathbf{s}_i)$ is correlated with $T(\mathbf{s}_i)$. Table values are the MSPE values for $\ddot{T}_{VHFK}(\mathbf{s})$ using (16) relative to the MSPE for $\hat{T}_{FK}(\mathbf{s})$ using (5). Simulation parameters include: μ_{σ^2} (the average size of measurement variances), κ (governing the variability of measurement error variances across sites), and ϕ (governing the variability of $\breve{\sigma}^2(\mathbf{s})$ as an estimate of the true measurement error variance $\sigma^2(\mathbf{s})$).

		κ				
ϕ	μ_{σ^2}	0.1	0.25	0.5	1	1.5
0.1	0.1	1.00	1.00	1.00	0.99	0.97
0.1	0.25	0.99	1.00	0.99	0.97	0.94
0.1	0.5	0.99	0.99	0.98	0.94	0.90
0.1	1	0.99	0.98	0.96	0.92	0.88
0.1	1.5	1.00	0.99	0.96	0.91	0.89
0.5	0.1	1.02	1.02	1.02	1.01	0.99
0.5	0.25	1.03	1.03	1.03	1.00	0.96
0.5	0.5	1.04	1.04	1.03	0.98	0.94
0.5	1	1.06	1.04	1.01	0.95	0.90
0.5	1.5	1.05	1.03	0.98	0.90	0.90

we can recommend HFK over FK without reservation. If the available measurement error variances are only approximate ($\phi = 0.5$), we can only recommend HFK over FK when the CV of the $\sigma^2(\mathbf{s})$ process is at least 0.5.

Additional small-scale simulations were run using spatial samples $\mathbf{D}$ comprising as few as 25 observations (on a 5×5 grid) with little or no changes in the MSPE values for HFK relative to FK. We also note that although our simulations are based on gridded data, the methods discussed here are not affected when analyzing nongridded data.

4.2 $\sigma^2(\mathbf{s}_i)$ *Is Correlated with* $T(\mathbf{s}_i)$

Simulations for the scenario when $\sigma^2(\mathbf{s}_i)$ is correlated with $T(\mathbf{s}_i)$ were carried out in a manner similar to Section 4.1 with the exception of the choice of the value for $\sigma^2(\mathbf{s}_i)$. In this section, we summarize the simulation results and compare the VHFK predictor (16) with the FK predictor (5). The motivation and methodology for these simulations are discussed in detail in Web Appendix D.

Table 2 gives results for this simulation in a format identical to that of Table 1. Comparing to Table 1, note that when $\sigma^2(\mathbf{s}_i)$ is correlated with $T(\mathbf{s}_i)$, the improvement of VHFK over FK is less pronounced, but VHFK is still superior in most cases. When the available measurement error variances are only approximate ($\phi = 0.5$), the set of scenarios where VHFK improves on FK is limited to the cases where κ (the CV of $\sigma^2(\mathbf{s})$) is greater than 0.5. The process for constructing the VHFK predictor (described in Section 3) is complicated by the additional model (12) that must be fitted. Hence, the small differences between values in the $\sigma^2(\mathbf{s})$ process when $\kappa \leqslant 0.5$ does not warrant the adoption of the VHFK predictor with its increased potential for estimation error. However, VHFK is superior to FK in most cases. An advantage of VHFK is that one can reasonably predict when it will be superior to FK. The VHFK predictor is the preferred choice whenever the $\sigma^2(\mathbf{s})$ process is known with reasonable precision and/or the $\sigma^2(\mathbf{s})$ process has substantial location-to-location variability (i.e, the CV of $\sigma^2(\mathbf{s})$ is greater than 0.5).

Finally, we note that in their simulation study, Kleijnen and Van Beers (2005) concluded that OK is "not very sensitive to variance heterogeneity." However, their conclusions were based on OK's accuracy relative to several competing methods that use repeated measurements of $Z(\mathbf{s})$ at each location. Our study employs a newly developed BLUP approach, which assumes that the measurement error variance at each location is known. In contrast to Kleijnen and Van Beers (2005), we conclude that notable improvement can be realized using HFK/VHFK, with MSPE values in our simulations reduced by as much as 64% compared with OK and as much as 32% compared with FK.

5. Smoothing of Climate Parameter Estimates

To illustrate HFK/VHFK, we consider a sample of output from an analysis of North American Region Climate Change Assessment Program (NARCCAP) data (`http://www.narccap.ucar.edu/`). In this example, we focus on mean summer temperatures (1980–2000) from the Hudson Strait in northern Canada, which connects the Hudson Bay on the west with the Labrador Sea on the east (see panel (a) of Figure 2). Climate change in the Hudson Strait (Houser and Gough, 2003) and throughout the Arctic region (Moritz, Bitz, and Steig, 2002) is of increasing interest to climate science researchers, but Arctic climate can be complex and difficult to model, particularly near water/land boundaries. Thus, although the NARCCAP models tend to yield similar pictures of North American climate averaged over decades, they are not in as strong agreement for mean summer temperatures in the Hudson Strait.

Christensen and Sain (2011) consider factor analysis models for multimodel ensemble data. They provide descriptions of climate processes in common to the ensemble and measures of commonality (dependence on common factors) for each climate model. The data $Z(\mathbf{s})$ for the illustration given here is the commonality process associated with one of the NARCCAP models—the Pennsylvania State University/National Center for Atmospheric Research (PSU/NCAR) mesoscale model (or "MM5I"). See Mearns et al. (2009) for a description of NARCCAP. This process is parameterized such that $Z(\mathbf{s})$ values greater (less) than 1 indicate that MM5I has a relationship with the common climate factor that is stronger (weaker) than average when compared with the other models. Associated with each estimated commonality $Z(\mathbf{s})$ is a standard error $\sigma(\mathbf{s})$, which we will treat as a measurement error standard deviation. Our purpose is to use model (1) to obtain an estimate of the error-filtered model-commonality process $T(\mathbf{s})$.

Because there is a reasonably strong positive association between $Z(\mathbf{s})$ and $\sigma(\mathbf{s})$ (with a linear relationship between log $Z(\mathbf{s})$ and log $\sigma(\mathbf{s})$ and a correlation of 0.64), the standard HFK predictor will be biased (see Section 3), and the VHFK approach is most appropriate. The use of VHFK seems particularly promising because the data scenario implied by the $Z(\mathbf{s})$ and $\sigma^2(\mathbf{s})$ processes falls within the collection of cases where VHFK was recommended in Section 4.2. The mean of the measurement error variances for this region is 0.31 and

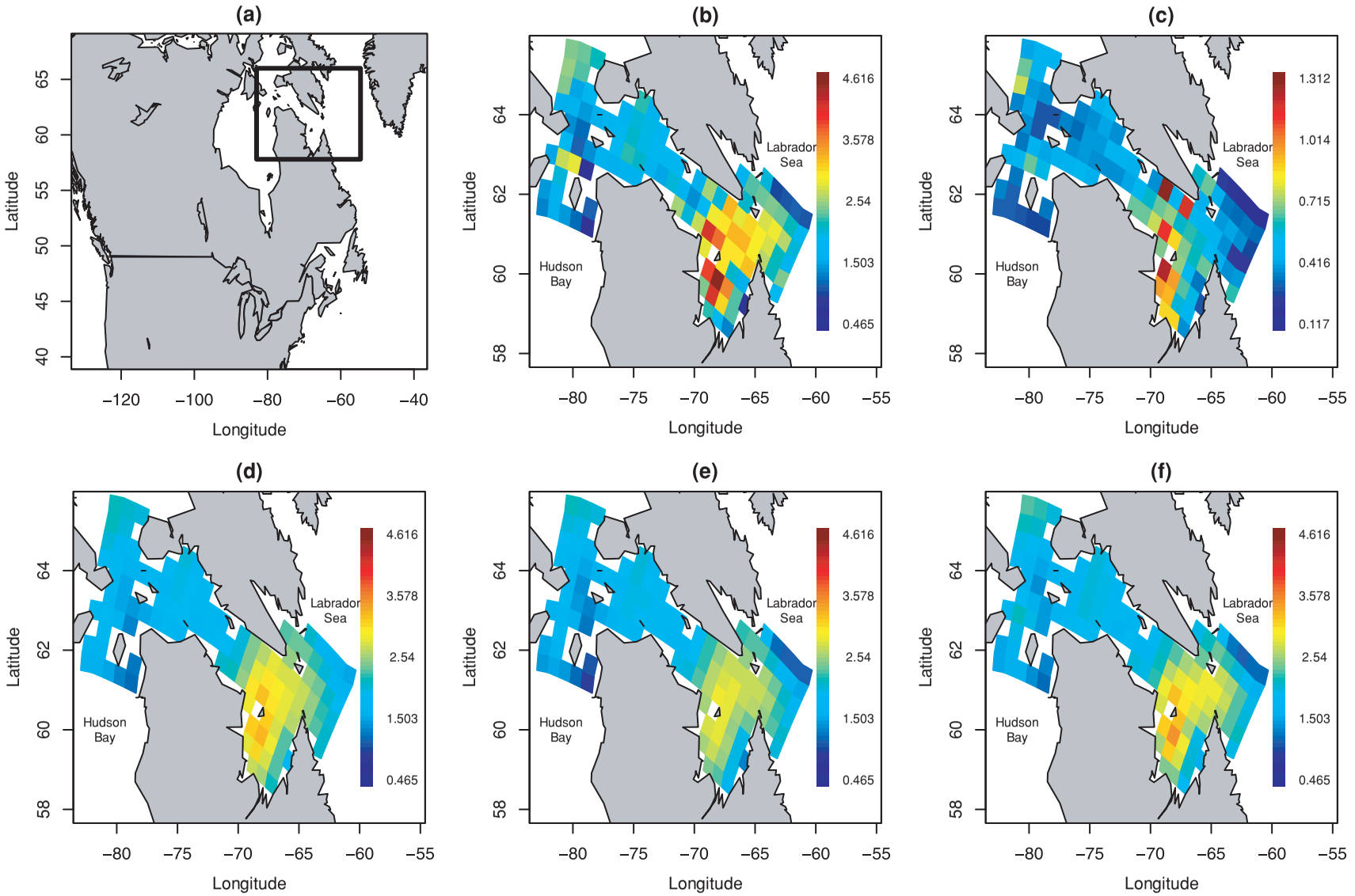

Figure 2. Prediction of the model-commonality process for the MM5I model, where values of $Z(\mathbf{s})$ (and $T(\mathbf{s})$) greater than 1 indicate that the MM5I model is strongly associated with the climate model ensemble's latent common factor. (a) Location of the Hudson Strait region. (b) The observed data $Z(\mathbf{s})$ denoting site-specific estimates of model commonality for the MM5I. (c) Site-specific standard errors ($\sigma(\mathbf{s})$) associated with the estimates in $Z(\mathbf{s})$. (d) FK prediction $\hat{T}_{FK}(\mathbf{s})$ from (5). (e) HFK prediction $\hat{T}_{HFK}(\mathbf{s})$ from (10). (f) VHFK prediction $\ddot{T}_{VHFK}(\mathbf{s})$ from (16). This figure appears color in the electronic version of this article.

the sample variance of $Z(\mathbf{s})$ is 0.56. Then, the logic behind (8) and Figure 1 implies that the variance of $T(\mathbf{s})$ should be approximately 0.25 (0.56 − 0.31). Thus, the ratio of μ_{σ^2} to var {T(s)} is roughly 1.25. The CV of the given $\sigma^2(\mathbf{s})$ process (formed using the squared standard error of the estimated commonality process at each location) can be calculated directly and is approximately 1 for these data. Thus, these data scenarios would be similar to a scenario in the simulations in Section 4.2 with $\mu_{\sigma^2} = 1.25$ and $\kappa = 1$. We see from Table 2 that this is clearly a case where VHFK should be superior. Because the condition (7) is verified (see Web Appendix E), we use the bias adjustment in (9).

Based on the simulation studies in Section 4, we expect FK to be the best alternative to VHFK. In comparing VHFK with FK, we are particularly interested in locations where $\ddot{T}_{VHFK}(\mathbf{s})$ from (16) is significantly different from $\hat{T}_{FK}(\mathbf{s})$ from (5). Statistical significance is assessed with the test statistic

$$t_{\text{diff}}(\mathbf{s}_0) = \frac{\ddot{T}_{VHFK}(\mathbf{s}_0) - \hat{T}_{FK}(\mathbf{s}_0)}{\sqrt{\hat{\tau}^2_{\text{diff}}(\mathbf{s}_0)}} \dot{\sim} N(0,1), \qquad (18)$$

which is described in Web Appendix F.

Figure 2 compares the raw data $Z(\mathbf{s})$ with the FK, HFK, and the VHFK predictions. Note that because of the correlation between $Z(\mathbf{s})$ and $\sigma(\mathbf{s})$, we see that—compared to VHFK—HFK dramatically oversmooths the surface, particularly in the locations with high values of $Z(\mathbf{s})$ and $\sigma(\mathbf{s})$. Figure 3 plots the difference between the variance-stabilized VHFK prediction and the FK prediction. Locations outlined with a black box are those for which the VHFK prediction is significantly different from the FK prediction at level $\alpha = 0.10$ (i.e., $|t_{\text{diff}}(\mathbf{s})| > 1.645$). Locations significant at the $\alpha = 0.05$ level are cross-hatched. Note that the areas where the predictors deviate most from each other can be identified by examining the $Z(\mathbf{s})$ and $\sigma(\mathbf{s})$ processes plotted in panels (b) and (c) of Figure 2. The locations marked with an "A" in Figure 3 have both a very low value of $Z(\mathbf{s})$ *and* a very low standard error. The FK predictor smooths these locations upward toward the mean more than the VHFK predictor because FK does not account for these measurements' high reliability. The area surrounding location "B" in Figure 3 is characterized by a mix of high and low Z(s) values, with the high values associated with large standard errors. The FK predictor implicitly assumes equal reliability for these measurements and effectively averages the high and low values for $Z(\mathbf{s})$ in this area. In contrast, the VHFK predictor gives more weight to the low-standard-error locations, yielding a lower predicted value for this area. The location marked with "C" in Figure 3 has a higher $Z(\mathbf{s})$ value than its neighboring locations, but is accompanied by a relatively low standard error, ensuring that the prediction at this location is not overly smoothed downward by the VHFK predictor. The kriging variances associated with the FK and VHFK predictions are relatively similar, with VHFK generally having smaller prediction variances in this example. Regardless of their relative sizes, we know that the heterogeneous measurement error variances render the estimated FK kriging variances invalid.

The practical importance of VHFK in this setting is that it accentuates the regions where model-commonality is low or high, giving climate modelers specific insight about geographical locations where climate models may be inadequately incorporating important geographical or climatological variables (e.g., sea ice or water/land boundaries). Without proper spatial prediction in the heterogeneous variance setting, there is a tendency to either understate or overstate the impact of local parameter estimation errors in the assessment of the climate model ensemble's performance. In general, the VHFK predictor provides an improved picture of the measurement-error-filtered process by exploiting the given measurement error variance information and yielding a more precise picture of which locations are anomalous or interesting.

6. Discussion and Conclusion

In this article, we present a new approach for kriging in the presence of heterogeneous (site-specific) measurement error variances. Although the predictor itself is based on a BLUP of the same structure as other kriging predictors, the real challenge in forming the BLUP is the nonstationarity of the process, which substantially complicates the estimation of the semivariogram/covariance function required to construct the predictor. We present a novel but fairly simple method for estimating the semivariogram of the measurement-error-free spatial process, which is based on a bias adjustment of the classical semivariogram formula, and we construct kriging equations that properly use this semivariogram along with the known measurement error variances. The HFK predictor in (10) reduces to the standard FK predictor in (5) when $\sigma^2(\mathbf{s})$ is the same value for all $\mathbf{s}$.

For use with processes where each site's measurement error variance is a function of the process of interest, we recommend an approach which uses a variance-stabilizing transformation. VHFK uses the notion of trans-Gaussian kriging to form a predictor.

Using simulation, we compare the HFK and VHFK predictors with the standard FK predictor (where the mean measurement error variance is used for the "common" measurement error variance for FK). When the measurement error variances are uncorrelated with the process of interest, using HFK instead of FK reduces the MSPE by as much as 32%. The only case where HFK would be inadvisable is when the measurement error variances are themselves measured with substantial error *and* the measurement error variance process $\sigma^2(\mathbf{s})$ has a CV $\leqslant 0.25$. That is, HFK is preferable when measurement error variances can be identified with reasonable accuracy and/or measurement error variances are noticeably heterogeneous across the region of interest. When site-specific measurement error variances are a function of the site-specific measured values, VHFK is also an improvement over FK except when the measurement error variances are measured with substantial error *and* the measurement error variance process $\sigma^2(\mathbf{s})$ has a CV $\leqslant 0.5$.

In our example using climate model output, we find that VHFK provides an improved predictor of the error-filtered model-commonality process of interest. Locations with high or low values of $\sigma^2(\mathbf{s})$ are appropriately down- or upweighted in the prediction of the process of interest, yielding a realistically smooth picture of the specific locations in which the MM5I regional climate model is in greatest agreement with the ensemble's common underlying climate factor.

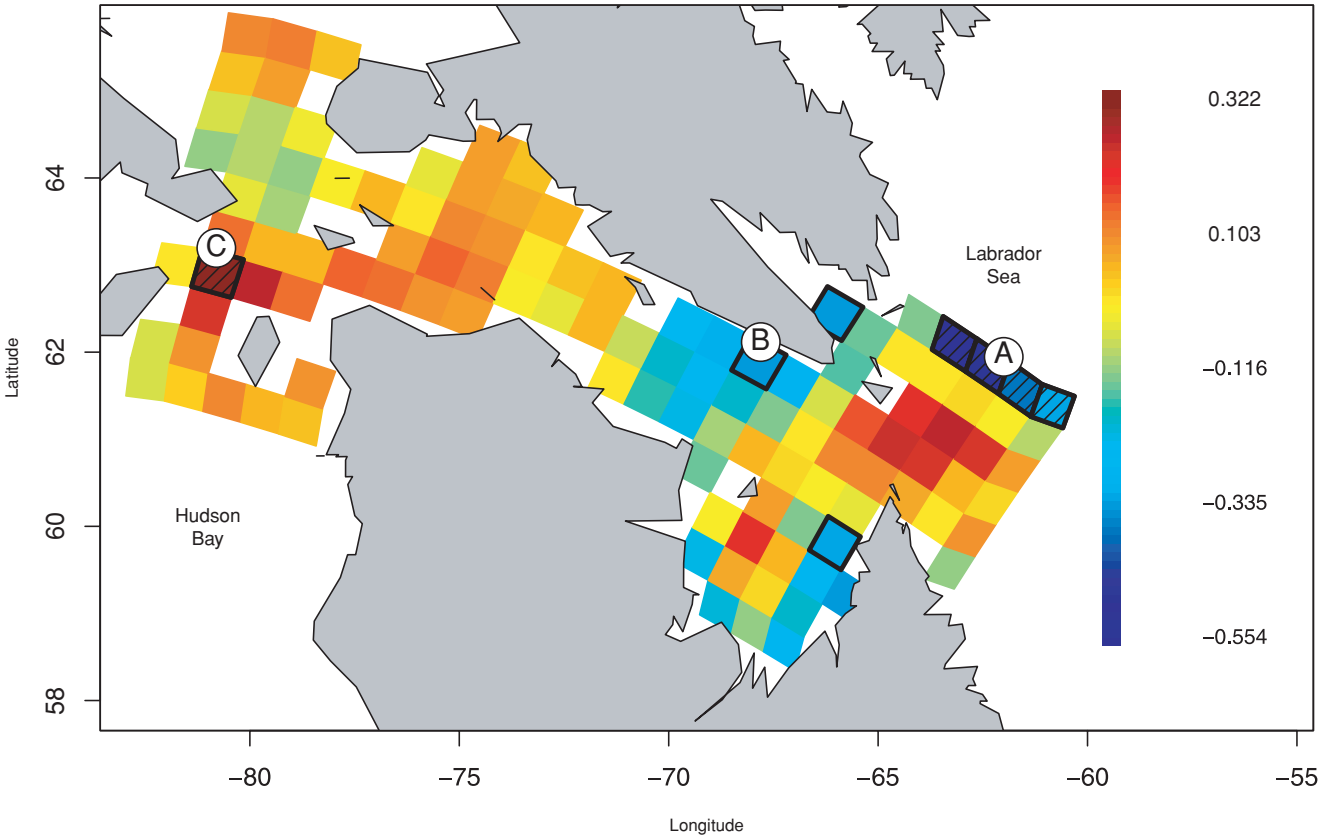

Figure 3. Differences $\ddot{T}_{VHFK}(\mathbf{s}) - \hat{T}_{FK}(\mathbf{s})$. Locations with significant differences between VHFK and FK are denoted with black boxes ($\alpha = 0.10$) and cross-hatching ($\alpha = 0.05$). This figure appears color in the electronic version of this article.

7. Supplementary Materials

Web Appendices referenced in Sections 2.2, 2.3, 3, 4.2, and 5 are available under the Paper Information link at the *Biometrics* website `http://www.biometrics.tibs.org/`.

Acknowledgements

I thank the NARCCAP for providing the data used in this article. NARCCAP is funded by the National Science Foundation, the U.S. Department of Energy, the National Oceanic and Atmospheric Administration, and the U.S. Environmental Protection Agency Office of Research and Development. I particularly thank Stephan Sain and Linda Mearns for their assistance with the NARCCAP data. I also thank the editors and reviewers for helpful comments and insights.

References

Box, G. E. P. and Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society, Series B* **26,** 211–243.

Box, G. E. P., Hunter, W. G., and Hunter, J. S. (1978). *Statistics for Experimenters*. New York: John Wiley & Sons.

Christensen, W. F. and Sain, S. R. (2011). Spatial latent variable modeling for integrating output from multiple climate models. To appear in *Mathematical Geosciences*, doi:10.1007/s11004-011-9321-1.

Cressie, N. (1985). Fitting variogram models by weighted least squares. *Mathematical Geology* **17,** 563–586.

Cressie, N. (1990). The origins of kriging. *Mathematical Geology* **22,** 239–252.

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

Hocking, R. R. (2003). *Methods and Applications of Linear Models: Regression and the Analysis of Variance*. Hoboken, New Jersey: John Wiley & Sons.

Houser, C. and Gough, W. A. (2003). Variations in sea ice in the Hudson Strait: 1971–1999. *Polar Geography* **27,** 1–14.

Kleijnen, J. P. C. and Van Beers, W. C. M. (2005). Robustness of kriging when interpolating in random simulation with heterogeneous variances: Some experiments. *European Journal of Operational Research* **165,** 826–834.

Kuehl, R. O. (1994). *Statistical Principles of Research Design and Analysis*. Belmont, California: Wadsworth.

Lingwall, J. W. and Christensen, W. F. (2007). Pollution source apportionment using a priori information and Positive Matrix Factorization. *Chemometrics and Intelligent Laboratory Systems* **87,** 281–294.

Lingwall, J. W., Christensen, W. F., and Reese, C. S. (2008). Dirichlet based Bayesian multivariate receptor modeling. *Environmetrics* **19,** 618–629.

Matheron, G. (1963). Principles of geostatistics. *Economic Geology* **58,** 1246–1266.

Mearns, L. O., Gutowski, W., Jones, R., Leung, R., McGinnis, S., Nunes, A., and Qian, Y. (2009). A regional climate change

assessment program for North America. *Eos Transactions of the American Geophysical Union* **90,** doi:10.1029/2009EO360002.

Moritz, R. E., Bitz, C. M., and Steig, E. J. (2002). Dynamics of recent climate change in the Arctic. *Science* **297,** 1497–1502.

Schabenberger, O. and Gotway, C. A. (2005). *Statistical Methods for Spatial Data Analysis*. Boca Raton, Florida: Chapman & Hall/CRC Press.

Waller, L. A. and Gotway, C. A. (2004). *Applied Spatial Statistics for Public Health Data*. Hoboken, New Jersey: John Wiley & Sons.

Zimmerman, D. L. and Cressie, N. A. (1992). Mean squared prediction error in the spatial linear model with estimated covariance parameters. *Annals of the Institute of Statistical Mathematics* **31,** 1–15.

Received June 2009. *Revised August* 2010.
Accepted October 2010.