A Comparison of Block and Semi-Parametric Bootstrap Methods for Variance Estimation in Spatial Statistics

N. Iranpanah[a], M. Mohammadzadeh[*,b], C. C. Taylor[c]

[a] *Department of Statistics, University of Isfahan, 8174673441, Isfahan, Iran.*
[b] *Department of Statistics, Tarbiat Modares University, P.O. Box 14115-175 Tehran, Iran.*
[c] *Department of Statistics, University of Leeds, Leeds LS2 9JT, UK.*

Abstract

Efron (1979) introduced the bootstrap method for independent data but it can not be easily applied to spatial data because of their dependency. For spatial data that are correlated in terms of their locations in the underlying space the moving block bootstrap method is usually used to estimate the precision measures of the estimators. The precision of the moving block bootstrap estimators is related to the block size which is difficult to select. In the moving block bootstrap method also the variance estimator is underestimated. In this paper, first the semi-parametric bootstrap is used to estimate the precision measures of estimators in spatial data analysis. In the semi-parametric bootstrap method, we use the estimation of spatial correlation structure. Then, we compare the semi-parametric bootstrap with a moving block bootstrap for variance estimation of estimators in a simulation study.

*Corresponding author. Tel. and fax: +98 21 82883483;
Email address: mohsen_m@modares.ac.ir (M. Mohammadzadeh)

Preprint submitted to Computational Statistics and Data Analysis *October 31, 2012*

Finally, we use the semi-parametric bootstrap to analyze the coal-ash data.
Key words: Moving block bootstrap; Semi-parametric bootstrap; Plug-in kriging; Monte-Carlo simulation; Coal-ash data.

---

1. Introduction

In environmental studies the data are usually spatially dependent. Determination of the spatial correlation structure of the data and prediction are two important problems in statistical analysis of spatial data. To do so a valid parametric variogram model is often fitted to the empirical variogram of the data. Since there is no closed form for the variogram parameter estimates, they are usually computed numerically. In addition, when data behave as a realization of a non-Gaussian random field, the bootstrap method can be used for statistical inference of spatial data.

The bootstrap technique (Efron, 1979; Efron and Tibshirani, 1993) is a very general method to measure the accuracy of estimators, in particular for parameter estimation from independent identically distributed (iid) variables. For spatially dependent data, the block bootstrap method can be used without requiring stringent structural assumptions. This is an important aspect of the bootstrap in the dependent case, as the problem of model misspecification is more prevalent under dependence and traditional statistical methods are often very sensitive to deviations from model assumptions. A prime example of this issue appears in the seminal paper by Singh (1981), who in addition to providing the first theoretical confirmation of the superiority of the Efron's bootstrap, also pointed out its inadequacy for dependent data. Different variants of spatial subsampling and spatial block bootstrap methods have been proposed in the literature; see Hall (1985), Possolo (1991), Liu and Singh (1992), Politis and Romano (1993, 1994), Sherman and Carlstein (1994), Sherman (1996), Politis, Paparoditis and Romano (1998, 1999), Politis, Romano and Wolf (1999), Bühlman and Künsch (1999), Nordman and Lahiri (2003) and references therein. Here we shall follow the moving block bootstrap (MBB) methods suggested by Lahiri (2003).

On the other hand, the semi-parametric bootstrap (SPB) method has been used by Freedman and Peters (1984) for linear models and Bose (1988) for autoregressive models in time series. In this paper, first, we apply SPB method for estimation of the sampling distribution of estimators in spatial data analysis. Then, the SPB and MBB methods are compared for variance estimation of estimators in a Monte-Carlo simulation study. Finally, the SPB method is used to estimate the bias, variance and distribution of plug-in kriging and variogram parameter estimation for the analysis of the coal-ash data.

In Section 2, spatial statistics, kriging and plug-in kriging are briefly reviewed. The MBB method is given in Section 3. We use the SPB algorithm for analysis of spatial data in Section 4. Section 5 consists of a Monte-Carlo simulation study for comparison of the SPB and MBB methods for variance estimation of estimators. These estimators are; sample mean, GLS plug-in estimator of mean, plug-in kriging and variogram parameters estimator; nugget effect, partial sill and range. In Section 6, we apply the SPB method for estimation of bias, variance and distribution of plug-in kriging and parameter variogram estimators for coal-ash data. In the last section, we will end with discussion and results.

2. Spatial Statistics and Kriging

Usually a random field $\{Z(s) : s \in D\}$ is used for modeling spatial data, where the index set D is a subset of Euclidean space $\mathbb{R}^d$, $d \geq 1$. Suppose $\mathcal{Z} = (Z(s_1), \ldots, Z(s_N))^T$ denotes N realizations of a second-order stationary random field $Z(\cdot)$ with constant unknown mean $\mu = E[Z(s)]$ and covariogram $\sigma(h) = \mathrm{Cov}[Z(s), Z(s+h)]; s, s+h \in D$. The covariogram $\sigma(h)$ is a positive definite function. At a given location $s_0 \in D$ the best linear unbiased predictor for $Z(s_0)$, the ordinary kriging predictor and its variance are given by (Cressie, 1993)

$$\hat{Z}(s_0) = \lambda^T \mathcal{Z}, \quad \sigma_k^2(s_0) = \sigma(0) - \lambda^T \sigma + m, \tag{1}$$

where

$$\lambda^T = (\sigma + \mathbf{1}m)^T \Sigma^{-1}, \quad m = (1 - \mathbf{1}^T \Sigma^{-1} \sigma)(\mathbf{1}^T \Sigma^{-1} \mathbf{1})^{-1}. \tag{2}$$

Here, $\mathbf{1} = (1, \ldots, 1)^T$, $\sigma = (\sigma(s_0 - s_1), \ldots, \sigma(s_0 - s_N))^T$ and Σ is an $N \times N$ matrix whose $(i, j)^{th}$ element is $\sigma(s_i - s_j)$.

In reality, the covariogram is unknown and should be estimated based on the observations. An empirical estimator of covariogram is given by

$$\hat{\sigma}(h) = {N_h}^{-1} \sum_{N(h)} [(Z(s) - \bar{Z})(Z(s+h) - \bar{Z})],$$

where $\bar{Z} = N^{-1} \sum_{i=1}^{N} Z(s_i)$ is the sample mean, $N(h) = \{(s_i, s_j) : s_i - s_j = h; i, j = 1, \cdots, N\}$ and N_h is the number of elements of $N(h)$. The covariogram estimator $\hat{\sigma}(h)$ cannot be used directly for kriging predictor equations, because it is not necessarily positive definite. The idea is to fit a valid parametric covariogram model $\sigma(h; \theta)$ that is closest to the empirical covariogram $\hat{\sigma}(h)$. Various parametric covariogram models such as exponential, spherical, Gaussian, linear are presented in Journel and Huijbregts (1978). For example, the exponential covariogram is given by

$$\sigma(h;\theta) = \begin{cases} c_0 + c_1 & ||h|| = 0 \\ c_1 \exp(\frac{-||h||}{a}) & ||h|| \neq 0, \end{cases} \tag{3}$$

where $\theta = (c_0, c_1, a)^T$ are the nugget effect, partial sill and range, respectively. The maximum likelihood (ML), restricted maximum likelihood (REML), ordinary least squares (OLS) and generalized least squares (GLS) methods can be applied to estimate θ. In these methods, $\hat{\theta}$ is computed numerically with the use of iterative algorithms since there is no closed form. For example, Mardia and Marshall (1984) described the maximum likelihood method for fitting the linear model when the residuals are correlated and when the covariance among the residuals is determined by a parametric model containing unknown parameters. Kent and Mardia (1996) introduced the spectral and circulant approximations to the likelihood for stationary Gaussian random fields. Also, Kent and Mohammadzadeh (1999) obtained a spectral approximation to the likelihood for an intrinsic random field. We will estimate $\text{Var}(\hat{\theta})$ by SPB method.

The plug-in kriging predictor and the plug-in kriging predictor variance are determined by using $\hat{\theta}$ instead of θ in the covariogram $\hat{\sigma}(s_i, s_j) = \sigma(s_i, s_j; \hat{\theta})$

as

$$\hat{\hat{Z}}(s_0) = \hat{Z}(s_0; \hat{\theta}); \quad \hat{\sigma}_k^2(s_0) = \sigma_k^2(s_0; \hat{\theta}). \tag{4}$$

The plug-in kriging predictor is a non-linear function of $\mathcal{Z}$ because $\hat{\theta}$ is a non linear estimator of θ. As a result, properties of the plug-in kriging predictor and the plug-in kriging predictor variance — such as unbiasedness and variance — are unknown. Mardia, Southworth and Taylor (1999) discussed the bias in maximum likelihood estimators. Under the assumption that $Z(\cdot)$ is Gaussian, Zimmerman and Cressie (1992) show that

$$E[\sigma_k^2(s_0; \hat{\theta})] \leq \sigma_k^2(s_0) \leq E[\hat{Z}(s_0; \hat{\theta}) - Z(s_0)]^2,$$

where $\hat{\theta}$ is ML estimator of θ. We can estimate the variance of the plug-in kriging predictor $\sigma^2(s_0) = \text{Var}[\hat{\hat{Z}}(s_0)]$ using the SPB method.

3. Moving Block Bootstrap

Suppose that the sampling region D_n is obtained by inflating the prototype set D_0 by the scaling constant λ_n as

$$D_n = \lambda_n D_0, \tag{5}$$

where $\{\lambda_n\}_{n \geq 1}$ is a positive sequence of scaling factors such that $\lambda_n \to \infty$ as $n \to \infty$ and D_0 is a Borel subset of $(-1/2, 1/2]^d$ containing an open neighborhood of the origin. Suppose that $\{Z(s) : s \in \mathbb{Z}^d\}$ is a stationary random field that is observed at finitely many locations $\mathcal{S}_n = \{s_1, \ldots, s_{N_n}\}$ given by the part of the integer grid $\mathbb{Z}^d$ that lies inside D_n, i.e., the data are $\mathcal{Z} = \{Z(s) : s \in \mathcal{S}_n\}$ for $\mathcal{S}_n = D_n \cap \mathbb{Z}^d$. Let $N \equiv N_n$ denote the sample size or the number of sites in D_n such that N and the volume of the sampling region D_n satisfies the relation $N = \text{Vol}(D_0)\lambda_n^d$, where $\text{Vol}(D_0)$ denotes the volume of D_0.

Let $\{\beta_n\}_{n\geq 1}$ be a sequence of positive integers such that $\beta_n^{-1} + \beta_n/\lambda_n = o(1)$ as $n \to \infty$. Here, β_n gives the scaling factor for the blocks in the spatial block bootstrap method. As a first step, the sampling region D_n is partitioned using blocks of volume β_n^d. Let $\mathcal{K}_n = \{k \in \mathbb{Z}^d : \beta_n(k + \mathcal{U}) \subset D_n\}$ denote the index set of all separate complete blocks $\beta_n(k + \mathcal{U})$ lying inside D_n such that $N = K\beta_n^d$, where $\mathcal{U} = (0, 1]^d$ denotes the unit cube in $\mathbb{R}^d$ and $K \equiv K_n$ denotes the size of $\mathcal{K}_n$. We define a bootstrap version of $\mathcal{Z}_n(D_n)$ by putting together bootstrap replicates of the process $Z(\cdot)$ on each block of D_n given by

$$D_n(k) \equiv D_n \cap [\beta_n(k + \mathcal{U})], \quad k \in \mathcal{K}_n. \tag{6}$$

Let $\mathcal{I}_n = \{i \in \mathbb{Z}^d : i + \beta_n\mathcal{U} \subset D_n\}$ denote the index set of all blocks of volume β_n^d in D_n, with starting points $i \in \mathbb{Z}^d$. Then, $\mathcal{B}_n = \{i + \beta_n\mathcal{U} :$

$i \in \mathcal{I}_n\}$ gives a collection of cubic blocks that are overlapping and contained in D_n. For the MBB method, for each $k \in \mathcal{K}_n$, one block is resampled at random from the collection $\mathcal{B}_n$ independently of the other resampled blocks, giving a version $\mathcal{Z}_n^*(D_n(k))$ of $\mathcal{Z}_n(D_n(k))$ using the observations from the resampled blocks. The bootstrap version $\mathcal{Z}_n^*(D_n)$ of $\mathcal{Z}_n(D_n)$ is now given by concatenating the resampled blocks of observations $\{\mathcal{Z}_n^*(D_n(k)) : k \in \mathcal{K}_n\}$.

Now the bootstrap version of a random variable $T_n = t_n(\mathcal{Z}_n(D_n); \theta)$ is given by $T_n^* = t_n(\mathcal{Z}_n^*(D_n); \hat{\theta}_n)$. For example, the bootstrap versions of $T_n = \sqrt{N}(\bar{Z}_n - \mu)$, where $\bar{Z}_n = N^{-1} \sum_{i=1}^N Z(s_i)$ and $\mu = E[Z(0)]$ is given by $T_n^* = \sqrt{N}(\bar{Z}_n^* - \hat{\mu}_n)$, where $\bar{Z}_n^* = N^{-1} \sum_{i=1}^N Z^*(s_i)$, $\hat{\mu}_n = E_*(\bar{Z}_n^*)$, and E_* denotes the conditional expectation given $\mathcal{Z}$.

Lahiri (2003) shows that the MBB method can be used to derive a consistent estimator of the variance of the sample mean, and more generally, of statistics that are smooth functions of the sample mean. Suppose that $\hat{\theta}_n = H(\bar{Z}_n)$ be an estimator of a parameter of interest $\theta = H(\mu)$, where H is a smooth function. Then, the bootstrap version of $\hat{\theta}_n$ is given by $\theta_n^* = H(\bar{Z}_n^*)$, and the bootstrap estimator of $\sigma_n^2 = N\mathrm{Var}(\hat{\theta}_n)$ is given by $\hat{\sigma}_n^2 \equiv \hat{\sigma}_n^2(\beta_n) = N\mathrm{Var}_*(\theta_n^*)$. He shows that under a weak dependence condition for the random field $\{Z(s) : s \in \mathbb{Z}^d\}$, like a strong mixing condition, then $\hat{\sigma}_n^2 \longrightarrow_p \sigma_\infty^2$ as

$n \longrightarrow \infty$, where $\sigma_\infty^2 \equiv \lim_{n\longrightarrow\infty} N\text{Var}(\hat{\theta}_n) = \frac{1}{\text{Vol}(D_0)} \sum_{i\in\mathbb{Z}^d} EW(0)W(i)$, with $W(i) = \sum_{|\alpha|=1} D^\alpha H(\mu)(Z(i)-\mu)^\alpha$, H is continuously differentiable and the partial derivatives $D^\alpha H(\cdot)$, $|\alpha| = 1$, satisfy Holder's condition. Nordman and Lahiri (2003) and Lahiri (2003) determined the optimal block size by computing $\text{Bias}[\hat{\sigma}_n^2(\beta_n)] = \beta_n^{-2}\gamma_2^2 + o(\beta_n^{-1})$ and $\text{Var}[\hat{\sigma}_n^2(\beta_n)] = N^{-1}\beta_n^d\gamma_1^2 + (1+o(1))$ and minimizing $\text{MSE}[\hat{\sigma}_n^2(\beta_n)] = N^{-1}\beta_n^d\gamma_1^2 + \beta_n^{-2}\gamma_2^2 + o(N^{-1}\beta_n^d + \beta_n^{-2})$ to obtain

$$\beta_n^{opt} = N^{\frac{d}{d+2}}[2\gamma_2^2/d\gamma_1^2]^{\frac{1}{d+2}}(1+o(1)), \tag{7}$$

where $\gamma_1^2 = (\frac{2}{3})^d . \frac{2\sigma_\infty^4}{(\text{Vol}(D_0))^3}$ and $\gamma_2 = -\frac{1}{\text{Vol}(D_0)} \sum_{i\in\mathbb{Z}^d} |i|\sigma_W(i)$ with $\sigma_W(i) = \text{Cov}(W(0), W(i)), i \in \mathbb{Z}^d$ and $|i| = i_1 + \cdots + i_d$ for $i = (i_1, \ldots, i_d) \in \mathbb{Z}^d$. The $\text{Bias}[\hat{\sigma}_n^2(\beta_n)]$ shows that the MBB estimator $\hat{\sigma}_n^2(\beta_n)$ is an underestimator of σ_n^2. Lahiri, Furukawa and Lee (2007) suggested a nonparametric plug-in rule for estimating optimal block sizes in various block bootstrap estimation problems. The optimal block size determination is difficult and sometimes impossible. On the other hand, when using the MBB method the variance estimator $\hat{\sigma}_n^2(\beta_n)$ is underestimated. Therefore, we use the SPB method for spatial data analysis.

4. Semi-Parametric Bootstrap

Suppose $\mathcal{Z} = (Z(s_1), \cdots, Z(s_N))^T$ are observations of a random field $\{Z(s) : s \in D \subset \mathbb{R}^d\}$ with decomposition $Z(s) = \mu(s) + \delta(s)$, where $\mu(\cdot) = E[Z(\cdot)]$ and the error term $\delta(\cdot)$ is a zero-mean stationary random field having $N \times N$ positive-definite covariance matrix $\Sigma \equiv (\sigma(s_i - s_j))$. The Cholesky decomposition allows Σ to be decomposed as the matrix product $\Sigma = LL^T$, where L is a lower triangular $N \times N$ matrix. Let $\epsilon \equiv (\epsilon(s_1), \ldots, \epsilon(s_N))^T = L^{-1}(\mathcal{Z} - \mu)$, be a vector of uncorrelated random variables with zero mean and unit variance from an unknown cumulative distribution $F(\varepsilon)$, where the mean $\mu = (\mu(s_1), \ldots, \mu(s_N))^T$. In the SPB method, we need an empirical distribution $F_N(\varepsilon)$ to estimate $F(\varepsilon)$. The SPB algorithm is described by the following steps:

Step 1. *Estimation and removal of mean structure.*
The trend or mean structure $\mu(\cdot)$ is estimated by the median polish algorithm (Cressie, 1993) or generalized additive models (Hastie, and Tibshirani, 1990) and is removed to obtain $R(s_i) = Z(s_i) - \hat{\mu}(s_i);\quad i = 1, \ldots, N$.

Step 2. *Estimation and removal of correlation structure.*
Estimate the spatial dependence structure of residual $R(s_i)$ by the covariance matrix $\hat{\Sigma}$. Note that, $\hat{\Sigma}$ is an $N \times N$ symmetric positive definite matrix whose

$(i,j)^{th}$ element is an estimate of the covariogram $\hat{\sigma}(s_i - s_j) = \sigma(s_i - s_j; \hat{\theta})$. Then $\hat{\epsilon} \equiv (\hat{\epsilon}(s_1), \ldots, \hat{\epsilon}(s_N))^T = \hat{L}^{-1}\mathcal{R}$ is a vector of uncorrelated residuals, where, $\hat{L}$ is a lower triangular $N \times N$ matrix from Cholesky decomposition $\hat{\Sigma} = \hat{L}\hat{L}^T$ and $\mathcal{R} \equiv (R(s_1), \ldots, R(s_N))^T$ is the vector of residuals.

Step 3. *Computation of empirical distribution $F_N(\varepsilon)$.* Suppose that $\tilde{\epsilon} \equiv (\tilde{\epsilon}(s_1), \ldots, \tilde{\epsilon}(s_N))^T$ is a vector of standardized values $\hat{\epsilon}$, where $\tilde{\epsilon}(s_i) = (\hat{\epsilon}(s_i) - \bar{\hat{\epsilon}})/s_{\hat{\epsilon}}$ and $\bar{\hat{\epsilon}}$, $s_{\hat{\epsilon}}$ denote the sample mean and standard deviation of the residuals, repectively. The empirical distribution function formed from standardized uncorrelated residuals $\{\tilde{\epsilon}(s_1), \ldots, \tilde{\epsilon}(s_N)\}$ is $F_N(\varepsilon) = N^{-1}\sum_{i=1}^{N} I(\tilde{\epsilon}(s_i) \leq \varepsilon)$, where $I(\tilde{\epsilon}(\cdot) \leq \varepsilon)$ is the indicator function equal to 1 when $\tilde{\epsilon}(\cdot) \leq \varepsilon$ and equal to 0 otherwise.

Step 4. *Resampling and Bootstrap sample.* Efron's (1979) bootstrap algorithm is used for the vector of standardized uncorrelated residuals $\tilde{\epsilon}$. We generate N iid bootstrap random variables $\epsilon^*(s_1), \ldots, \epsilon^*(s_N)$ having common distribution $F_N(\varepsilon)$. In other words, $\epsilon^* \equiv (\epsilon^*(s_1), \ldots, \epsilon^*(s_N))^T$ is a simple random sample with replacement from the standardized uncorrelated residuals $\{\tilde{\epsilon}(s_1), \ldots, \tilde{\epsilon}(s_N)\}$. The bootstrap sample $\mathcal{Z}^* \equiv (Z^*(s_1), \cdots, Z^*(s_N))^T$ can be determined using an inverse transform $\mathcal{Z}^* = \hat{\mu} + \hat{L}\epsilon^*$, where $\hat{\mu} = (\hat{\mu}(s_1), \ldots, \hat{\mu}(s_N))^T$ estimates the mean structure.

Step 5. *Bootstrap version of T.*

If $\hat{T} = t(\mathcal{Z}; \hat{\mu}, \hat{\theta})$ is a plug-in estimator of $T = t(\mathcal{Z}; \mu, \theta)$, where $\hat{\theta}$ is the plug-in estimator of θ, then, the SPB version of $\hat{T}$ is given by $T^* = t(\mathcal{Z}^*; \hat{\mu}, \hat{\theta})$.

Step 6. *Bootstrap estimators.*

The bootstrap estimators of the bias, variance and distribution of T are given by

$$\begin{aligned}
\text{Bias}_*(T^*) &= E_*(T^*) - \hat{T}, \\
\text{Var}_*(T^*) &= E_*[(T^*) - E_*(T^*)]^2, \\
G_*(t) &= P_*(T^* \leq t),
\end{aligned}$$

where E_*, Var_* and P_* denote the bootstrap conditional expectation, variance and probability given $\mathcal{Z}$.

Step 7. *Monte-Carlo approximation.*

When the above bootstrap estimators have no closed form, the precision measures of T^* may be evaluated by Monte-Carlo simulation as follows. We repeat Steps 4 and 5, B (e.g., $B = 1000$) times to obtain bootstrap replicates $T_1^*, \ldots, T_B^*$. Then the Monte-Carlo approximations of the bootstrap estimators in step 6 are given by

$$\widehat{\text{Bias}}_*(T^*) = \frac{1}{B} \sum_{b=1}^{B} T_b^* - \hat{T}, \tag{8}$$

$$\widehat{\mathrm{Var}}_*(T^*) = \frac{1}{B}\sum_{b=1}^{B}(T_b^* - \frac{1}{B}\sum_{b=1}^{B}T_b^*)^2, \quad (9)$$

$$\widehat{G}_*(t) = \frac{1}{B}\sum_{b=1}^{B} I(T_b^* \leq t). \quad (10)$$

5. Simulation Study

In this section, we conduct a simulation study to compare the MBB and SPB estimator of $\sigma^2 = \mathrm{Var}(T)$, where T is a statistic of interest. We consider four examples for T: the sample mean; GLS plug-in mean estimator; plug-in kriging; and covariogram parameters estimator. Let $\{Z(s) : s \in Z^2\}$ be a zero mean second-order stationary Gaussian process with the exponential covariogram (3) using parameter values $\theta_1 = (1,1,1)^T$ (weak dependence) and $\theta_2 = (0,2,2)^T$ (strong dependence). We generate realizations of the Gaussian random field $Z(\cdot)$ over three rectangular regions $D = n \times n$; $n = 6, 12, 24$ as spatial sample $\mathcal{Z} = (Z(s_1), \ldots, Z(s_N))^T$ where $N = n^2$.

To apply the MBB method, we identify the above rectangular regions D as $[-3,3) \times [-3,3), [-6,6) \times [-6,6)$ and $[-12,12) \times [-12,12)$, the scaling constants $\lambda = 6, 12, 24$ respectively and the prototype set $D_0 = [-\frac{1}{2}, \frac{1}{2}) \times [-\frac{1}{2}, \frac{1}{2})$. For example, for the sample size $N = \lambda^2 = 144$ and $\beta = 2$, there are $K = |\mathcal{K}| = 36$ subregions in the partition (6), given by $D(k) = [2k_1, 2k_1 + 2) \times [2k_2, 2k_2 + 2)$; $k \in \mathcal{K} = \{(k_1, k_2)^T \in Z^2, -3 \leq k_1, k_2 < 3\}$. To define the MBB version of the random field $Z(\cdot)$ over D we randomly resample 36 times, with replacement from the collection of all observed moving blocks

$$\mathcal{B}(i) = [i_1, i_1 + 2) \times [i_2, i_2 + 2); \quad i \in \mathcal{I} = \{(i_1, i_2)^T \in Z^2, -6 \leq i_1, i_2 < 4\}.$$

The MBB sample $\mathcal{Z}^* = \mathcal{Z}^*(D) = (Z^*(s_1), \ldots, Z^*(s_N))^T$ is given by concatenating the K-many resampled blocks to size β of observations $\{\mathcal{Z}^*(D(k)) : k \in \mathcal{K}\}$.

To define the SPB version of the random field $Z(\cdot)$ over D, we apply steps 2–4 in SPB method. First, the covariance matrix Σ is estimated using the plug-in estimator of the covariogram $\hat{\sigma}(h;\theta) = \sigma(h;\hat{\theta})$, where $\hat{\theta} = (\hat{c}_0, \hat{c}_1, \hat{a})^T$ is an estimator of θ (e.g. ML estimator). Let $\hat{L}$ be the Cholesky decomposition of $\hat{\Sigma}$, then $\hat{\epsilon} = \hat{L}^{-1}\mathcal{Z}$ is a vector of uncorrelated values. Hence, the bootstrap vector $\epsilon^* = (\epsilon^*(s_1), \ldots, \epsilon^*(s_N))^T$ is generated as a simple random sample with replacement from $\{\tilde{\epsilon}(s_1), \ldots, \tilde{\epsilon}(s_N)\}$, where $\tilde{\epsilon}(\cdot)$ denotes standardized uncorrelated values of $\hat{\epsilon}(\cdot)$. Finally, the SPB sample $\mathcal{Z}^* = (Z^*(s_1), \ldots, Z^*(s_N)))^T$ is given by the inverse transform $\mathcal{Z}^* = \hat{L}\epsilon^*$.

Suppose that $T = t(\mathcal{Z})$ is the statistic of interest, then the MBB and SPB versions of T are given by $T^* = t(\mathcal{Z}^*)$. The MBB and SPB estimators $\hat{\sigma}^2 = \mathrm{Var}_*(T^*)$ of $\sigma^2 = \mathrm{Var}(T)$ are approximated based on $B = 1000$ bootstrap replicates (9). For each region D and covariance structure, we compute the variance estimator $\hat{\sigma}^2$ and approximate the normalized bias, variance and mean squared error(MSE)

$$\begin{aligned} \text{NBias}(\hat{\sigma}^2) &= E(\hat{\sigma}^2/\sigma^2) - 1, \\ \text{NVar}(\hat{\sigma}^2) &= \text{Var}(\hat{\sigma}^2/\sigma^2), \\ \text{NMSE}(\hat{\sigma}^2) &= E[(\hat{\sigma}^2/\sigma^2) - 1]^2, \end{aligned}$$

by its empirical version based on 10000 simulations. In MBB method, the variance estimator is determined as $\hat{\sigma}^2 = \hat{\sigma}^2(\beta^{opt})$, where the optimal block size β^{opt} is based on minimal NMSE over various block sizes β.

Example 1. The Sample mean

In this example, we compare the MBB and SPB estimators $\hat{\sigma}_1^2 = N\text{Var}_*(\bar{Z}^*)$ of $\sigma_1^2 = N\text{Var}(\bar{Z}) = N^{-1}\mathbf{1}^T\Sigma\mathbf{1}$, where the sample mean $\bar{Z} = N^{-1}\sum_{i=1}^N Z(s_i)$ is the OLS estimator of mean μ and $\bar{Z}^*$ is a bootstrap sample mean. We consider version T_1^* of the sample mean $T_1 = \sqrt{N}\bar{Z}$ based on a bootstrap sample $\mathcal{Z}^*$ by $T_1^* = \sqrt{N}\bar{Z}^*$. The MBB and SPB estimators $\hat{\sigma}_1^2 = N\text{Var}_*(\bar{Z}^*)$ are approximated based on $B = 1000$ bootstrap replicates (9). The covariogram models that we considered are exponential, spherical and unknown.

Table 1 shows approximates of the NBias, NVar and NMSE for MBB estimators $\hat{\sigma}_1^2$ for various block sizes β based on the exponential covariogram model. The asterisk (*) denotes the minimal value of the NMSE. From Table 1, the optimum block size β^{opt} can be determined based on minimal value of the NMSE. For example, for θ_1 and $n = 6, 12, 24$ the optimum block size is $\beta^{opt} = 2, 3, 6$ and for θ_2 and $n = 6, 12, 24$, $\beta^{opt} = 3, 4, 8$. We have used the optimum block sizes β^{opt} for MBB method in Table 2. To conserve space, we will not further mention the determination of β^{opt} as in Table 1.

Tables 2-4 show true values of σ_1^2, estimates of the NBias, NVar and NMSE for MBB (based on β^{opt}) and SPB estimators $\hat{\sigma}_1^2$ based on exponential covariogram, spherical covariogram with parameter values $\theta_2 = (0, 2, 2)^T$ and $\theta_3 = (0, 2, 4)^T$ and unknown covariogram.

Example 2. The GLS plug-in mean estimate

Let $\hat{\mu} = \mathbf{1}^T\Sigma^{-1}\mathcal{Z}/\mathbf{1}^T\Sigma^{-1}\mathbf{1}$ be the GLS estimator of mean μ with variance $1/\mathbf{1}^T\Sigma^{-1}\mathbf{1}$. We compare MBB and SPB estimators of $\sigma_2^2 = N\mathrm{Var}(\hat{\hat{\mu}})$, where $\hat{\hat{\mu}} = \mathbf{1}^T\hat{\Sigma}^{-1}\mathcal{Z}/\mathbf{1}^T\hat{\Sigma}^{-1}\mathbf{1}$ is GLS plug-in estimator of μ. We define a version T_2^* of the GLS plug-in mean $T_2 = \sqrt{N}\hat{\hat{\mu}}$ based on a bootstrap sample $\mathcal{Z}^*$ by $T_2^* = \sqrt{N}\mu^*$, where $\mu^* = \mathbf{1}^T\hat{\Sigma}^{-1}\mathcal{Z}^*/\mathbf{1}^T\hat{\Sigma}^{-1}\mathbf{1}$.

Example 3. Plug-in kriging

To compare MBB and SPB variance estimators of $\sigma_3^2 = \mathrm{Var}[\hat{\hat{Z}}(s_0)]$, we define the T_3^* version of plug-in ordinary kriging predictor $T_3 = \hat{\hat{Z}}(s_0) = \hat{\lambda}^T\mathcal{Z}$, based

Table 1: Approximates of the NBias, NVar and NMSE for MBB estimators $\hat{\sigma}_1^2 = \hat{\sigma}_1^2(\beta)$ based on exponential covariogram. The asterisk (*) denotes the minimal value of MSE.

		$\theta_1 = (1, 1, 1)^T$			$\theta_2 = (0, 2, 2)^T$		
n	β	NBias	NVar	NMSE	NBias	NVar	NMSE
6	2	-0.569	0.039	0.362*	-0.853	0.008	0.736
	3	-0.624	0.057	0.446	-0.844	0.013	0.725*
	2	-0.561	0.011	0.326	-0.864	0.002	0.750
12	3	-0.475	0.033	0.258*	-0.786	0.009	0.626
	4	-0.452	0.063	0.267	-0.732	0.021	0.557*
	6	-0.563	0.080	0.397	-0.751	0.033	0.597
	2	-0.575	0.003	0.333	-0.874	0.001	0.764
	3	-0.463	0.009	0.233	-0.790	0.003	0.626
24	4	-0.369	0.018	0.174	-0.710	0.008	0.512
	6	-0.320	0.053	0.155*	-0.595	0.029	0.383
	8	-0.328	0.087	0.195	-0.541	0.058	0.351*
	12	-0.507	0.102	0.359	-0.648	0.064	0.484

Table 2: True values of σ_1^2 and approximates of the NBias, NVar and NMSE for MBB and SPB estimators $\hat{\sigma}_1^2$ based on exponential covariogram.

		$\theta_1 = (1,1,1)^T$					$\theta_2 = (0,2,2)^T$				
Method	n	σ_1^2	β^{opt}	NBias	NVar	NMSE	σ_1^2	β^{opt}	NBias	NVar	NMSE
MBB	6	5.279	2	-0.572	0.039	0.366	19.994	3	-0.846	0.014	0.729
SPB				-0.254	0.295	0.359			-0.327	0.367	0.474
MBB	12	6.311	3	-0.471	0.033	0.254	32.074	4	-0.740	0.021	0.569
SPB				-0.059	0.239	0.242			-0.067	0.343	0.347
MBB	24	6.890	6	-0.310	0.054	0.150	40.598	8	-0.558	0.057	0.369
SPB				0.012	0.142	0.143			0.039	0.193	0.195

Table 3: True values of σ_1^2 and approximates of the NBias, NVar and NMSE for MBB and SPB estimators $\hat{\sigma}_1^2$ based on spherical covariogram.

		$\theta_2 = (0,2,2)^T$					$\theta_3 = (0,2,4)^T$				
Method	n	σ_1^2	β^{opt}	NBias	NVar	NMSE	σ_1^2	β^{opt}	NBias	NVar	NMSE
MBB	6	4.728	2	-0.398	0.078	0.236	14.069	3	-0.703	0.051	0.546
SPB				-0.042	0.231	0.232			-0.302	0.275	0.366
MBB	12	5.072	3	-0.285	0.053	0.134	17.046	4	-0.493	0.063	0.306
SPB				-0.046	0.048	0.048			-0.122	0.120	0.135
MBB	24	5.249	4	-0.188	0.029	0.064	18.638	6	-0.313	0.057	0.155
SPB				-0.026	0.011	0.012			-0.048	0.020	0.022

Table 4: True values of σ_1^2 and approximates of the NBias, NVar and NMSE for MBB and SPB estimators $\hat{\sigma}_1^2$ based on unknown covariogram.

		weak dependence					strong dependence				
Method	n	σ_1^2	β^{opt}	NBias	NVar	NMSE	σ_1^2	β^{opt}	NBias	NVar	NMSE
MBB	6	2.593	2	-0.125	0.124	0.140	35.637	3	-0.927	0.004	0.863
SPB				-0.026	0.101	0.102			-0.620	0.353	0.737
MBB	12	3.896	3	-0.032	0.031	0.032	78.315	4	-0.880	0.006	0.781
SPB				-0.011	0.013	0.013			-0.482	0.465	0.697
MBB	24	4.681	4	-0.006	0.009	0.009	126.930	8	-0.754	0.024	0.592
SPB				-0.003	0.004	0.004			-0.422	0.349	0.527

Table 5: True values of σ_2^2 and approximates of the NBias, NVar and NMSE for MBB and SPB estimators $\hat{\sigma}_2^2$.

		$\theta_1 = (1,1,1)^T$					$\theta_2 = (0,2,2)^T$				
Method	n	σ_2^2	β^{opt}	NBias	NVar	NMSE	σ_2^2	β^{opt}	NBias	NVar	NMSE
MBB	6	5.700	2	-0.574	0.044	0.374	16.355	2	-0.749	0.031	0.592
SPB				-0.341	0.201	0.317			-0.274	0.406	0.481
MBB	12	6.242	3	-0.434	0.046	0.235	27.771	4	-0.643	0.045	0.458
SPB				-0.108	0.202	0.214			-0.116	0.286	0.299
MBB	24	6.504	4	-0.329	0.025	0.133	36.802	6	-0.521	0.043	0.315
SPB				-0.039	0.123	0.124			0.006	0.166	0.166

Table 6: True values of σ_3^2 and approximates of the NBias, NVar and NMSE for MBB and SPB estimators $\hat{\sigma}_3^2$.

			$\theta_1 = (1,1,1)^T$					$\theta_2 = (0,2,2)^T$				
Method	n	s_0	σ_3^2	β^{opt}	NBias	NVar	NMSE	σ_3^2	β^{opt}	NBias	NVar	NMSE
MBB	6	(3.5,3.5)	0.496	2	-0.386	0.404	0.553	1.530	2	-0.510	0.133	0.393
SPB					-0.297	0.414	0.503			-0.372	0.168	0.306
MBB	12	(6.5,6.5)	0.415	3	-0.212	0.252	0.297	1.436	4	-0.215	0.114	0.160
SPB					-0.128	0.265	0.282			-0.111	0.087	0.099
MBB	24	(12.5,12.5)	0.381	8	-0.036	0.132	0.133	1.385	8	-0.068	0.059	0.063
SPB					-0.018	0.115	0.115			0.001	0.036	0.036

on a bootstrap sample $\mathcal{Z}^*$ by $T_3^* = Z^*(s_0) = \hat{\lambda}^T \mathcal{Z}^*$.

The MBB and SPB estimators $\hat{\sigma}_2^2 = N\text{Var}_*(\mu^*)$ and $\hat{\sigma}_3^2 = \text{Var}_*[Z^*(s_0)]$ are approximated based on $B = 1000$ bootstrap replicates (9). Tables 5 and 6 show true values of σ_2^2 and σ_3^2, estimates of the NBias, NVar and NMSE for MBB (based on β^{opt}) and SPB estimators $\hat{\sigma}_2^2$ and $\hat{\sigma}_3^2$ based on exponential covariogram for each region D and covariogram parameters θ_1 and θ_2.

Example 4. Covariogram parameters estimator

Let $\hat{\theta} = (T_4, T_5, T_6) = (\hat{c}_0, \hat{c}_1, \hat{a})$ be the MLEs of the covariogram parameters $\theta = (c_0, c_1, a)$. Note that the estimator of $\hat{\theta}$ is computed numerically based

Table 7: True values of σ_4^2 and approximates of the NBias, NVar and NMSE for MBB and SPB estimators $\hat{\sigma}_4^2$.

		$\theta_1 = (1,1,1)^T$					$\theta_2 = (0,2,2)^T$				
Method	n	σ_2^2	β^{opt}	NBias	NVar	NMSE	σ_2^2	β^{opt}	NBias	NVar	NMSE
MBB	6	0.639	2	-0.547	0.240	0.539	0.026	3	-0.037	0.141	0.142
SPB				-0.114	0.237	0.250			-0.072	0.129	0.134
MBB	12	0.378	4	-0.091	0.312	0.321	0.011	4	-0.055	0.100	0.103
SPB				-0.083	0.220	0.227			0.073	0.092	0.097
MBB	24	0.198	6	-0.102	0.291	0.301	0.003	8	-0.148	0.010	0.032
SPB				0.069	0.193	0.198			0.040	0.003	0.005

on the spatial sample $\mathcal{Z}$ as $T_i = t_i(\mathcal{Z})$; $i = 4, 5, 6$ and has no closed form, so $\sigma_i^2 = \text{Var}(T_i)$ is unknown. We define a version $T_i^* = t_i(\mathcal{Z}^*)$ of the estimator T_i based on bootstrap samples $\mathcal{Z}^*$. The MBB and SPB estimators $\hat{\sigma}_i^2 = \text{Var}_*(T_i^*)$ are approximated based on $B = 1000$ bootstrap replicates (9). Tables 7–9 show true values of σ_i^2, estimates of the NBias, NVar and NMSE for MBB (based on β^{opt}) and SPB estimators $\hat{\sigma}_i^2$ based on exponential covariogram for each region D and covariogram parameters θ_1 and θ_2.

Results

Tables 1–9 show that the MBB variance estimations $\hat{\sigma}^2$ are underestimated.

Table 8: True values of σ_5^2 and approximates of the NBias, NVar and NMSE for MBB and SPB estimators $\hat{\sigma}_5^2$.

		$\theta_1 = (1,1,1)^T$					$\theta_2 = (0,2,2)^T$				
Method	n	σ_2^2	β^{opt}	NBias	NVar	NMSE	σ_2^2	β^{opt}	NBias	NVar	NMSE
MBB	6	0.863	2	-0.655	0.233	0.662	0.686	2	-0.363	0.764	0.896
SPB				-0.120	0.258	0.272			-0.297	0.689	0.777
MBB	12	0.409	3	-0.118	0.288	0.302	0.246	4	-0.309	0.702	0.797
SPB				-0.084	0.181	0.188			-0.273	0.507	0.581
MBB	24	0.203	4	-0.145	0.2775	0.298	0.078	6	-0.294	0.624	0.710
SPB				-0.074	0.139	0.144			0.220	0.358	0.406

Table 9: True values of σ_6^2 and approximates of the NBias, NVar and NMSE for MBB and SPB estimators $\hat{\sigma}_6^2$.

		$\theta_1 = (1,1,1)^T$					$\theta_2 = (0,2,2)^T$				
Method	n	σ_2^2	β^{opt}	NBias	NVar	NMSE	σ_2^2	β^{opt}	NBias	NVar	NMSE
MBB	6	0.471	2	-0.714	0.459	0.969	1.477	3	-0.377	0.761	0.903
SPB				-0.616	0.447	0.826			-0.247	0.594	0.655
MBB	12	0.258	4	-0.552	0.312	0.616	0.592	6	-0.302	0.702	0.793
SPB				-0.434	0.195	0.383			-0.206	0.488	0.530
MBB	24	0.162	8	-0.400	0.278	0.438	0.151	8	-0.260	0.639	0.707
SPB				-0.260	0.145	0.213			0.117	0.384	0.398

Tables 2–9 show that the MBB and SPB variance estimations $\hat{\sigma}^2$ are asymptotically unbiased and consistent. Tables 2–9 also indicate that the SPB estimators are preferable to the MBB versions, especially for stronger dependence structure and larger sample sizes. In Tables 5–9, true values of $\sigma_i^2 = \text{Var}(T_i);\ i = 2, \cdots, 6$ have no closed form and they can be approximated based on Monte-Carlo simulation by 10000 times replicates.

6. Analysis of Coal-Ash Data

In this section, we apply the SPB method to analyze the coal-ash data (Cressie, 1993) from Greene County, Pennsylvania. These data are collected with sample size $N = 206$ at locations $\{Z(x, y) : x = 1, \ldots, 16; y = 1, \ldots, 23\}$ with west coordinates greater than 64 000 ft; spatially this defines an approximately square grid, with 2500 ft spacing (Cressie, 1993; Fig. 2.2). Our goal is estimation of bias, variance and distribution of plug-in kriging predictor and variogram parameters estimator by SPB method.

The SPB algorithm is used to estimate and remove the correlation structure. To estimate the correlation structure of the residuals, first, the spherical semi-variogram

$$\gamma(h;\theta) = \begin{cases} 0 & ||h|| = 0 \\ c_0 + c_1(\frac{3}{2}\frac{||h||}{a} - \frac{1}{2}(\frac{||h||}{a})^3) & 0 < ||h|| \le a \\ c_0 + c_1 & ||h|| \ge a \end{cases} \tag{11}$$

is fitted to the empirical semi-variogram estimation of coal-ash data with $\hat{\theta} = (\hat{c}_0, \hat{c}_1, \hat{a}) = (0.817, 0.815, 15.787)$. Figure 1(a) shows the fitted spherical

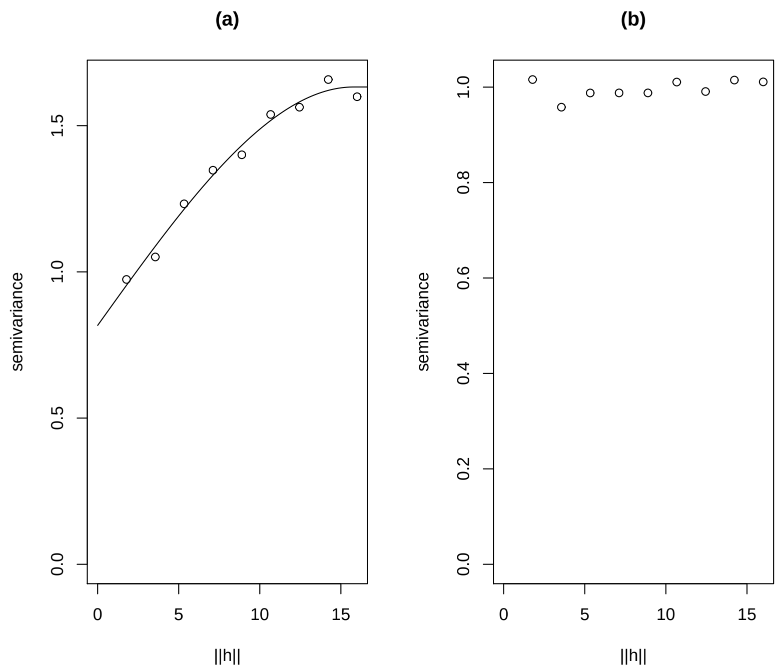

Figure 1: (a) Spherical semi-variogram model $\hat{\gamma}(h;\theta)$ fitted to the empirical semi-variogram $\hat{\gamma}(h)$ before removal correlation structure. (b) Empirical semi-variogram $\hat{\gamma}(h)$ for standardized residuals after removal correlation structure.

semi-variogram. The covariance matrix can be estimated as $\hat{\Sigma} = \sigma(h;\hat{\theta}) = \sigma(0;\hat{\theta}) - \gamma(h;\hat{\theta})$. Then, the uncorrelated residuals $\hat{\epsilon} = \hat{L}^{-1}R$ are used to compute the standardized uncorrelated residuals $\tilde{\epsilon}(s_i) = (\hat{\epsilon}(s_i) - \bar{\hat{\epsilon}})/s_{\hat{\epsilon}};\ i = 1,\ldots,N$. Figure 1(b) shows the fit of a linear semi-variogram to the empirical semi-variogram estimate of the standardized residuals. The linear semi-variogram model in Figure 1(b) shows that the standardized residuals $(\tilde{\epsilon}(s_1),\ldots,\tilde{\epsilon}(s_N))$ are uncorelated. Finally, the bootstrap samples are determined by $\mathcal{Z}^* = \hat{\mu} + \hat{L}\epsilon^*$, where the bootstrap vector ϵ^* is generated by simple random sampling with replacement from the standardized uncorrelated residuals vector $\tilde{\epsilon}$.

Now suppose that the plug-in ordinary kriging $T_1 = \hat{\hat{Z}}(s_0)$ and variogram parameter estimators $\hat{\theta} = (T_2, T_3, T_4) = (\hat{c}_0, \hat{c}_1, \hat{a})$ are the estimators of interest, where $T_i = t_i(\mathcal{Z})$. For example, if $s_0 = (5,6)$ is a new location then, $\hat{\hat{Z}}(s_0) = \hat{\lambda}^T\mathcal{Z} = 10.696$ and also $\hat{\theta} = (\hat{c}_0, \hat{c}_1, \hat{a}) = (0.817, 0.815, 15.787)$. The SPB version T_i^* of T_i is $T_i^* = t_i(\mathcal{Z}^*)$, where $\mathcal{Z}^*$ is the SPB sample. We estimate the precision measures $\text{Bias}(T_i)$ and $\text{Var}(T_i)$ and distribution $G_{T_i}(t)$ by SPB method and B bootstrap replicates $T_{i,1}^*,\ldots,T_{i,B}^*;\ i = 1,2,3,4$ in relations (8)–(10). Table 10 shows estimates of SPB bias and variance for plug-in kriging and estimates of variogram parameters based on $B = 1000$

Table 10: Estimates of SPB bias and variance for plug-in kriging and variogram parameters for coal-ash data.

T_i^*	Bias_*	Var_*
$Z^*(s_0)$	–0.901	0.706
c_0^*	0.002	0.017
c_1^*	0.066	0.037
a^*	–5.829	21.602

bootstrap replicates. Figure 2 shows the histogram of plug-in kriging and variogram parameters estimator based on $B = 1000$ bootstrap replicates.

7. Discussion and Results

Spatial data analysis is based on the estimate of correlation structure, for example, kriging predictor. The estimation of correlation structure is based on parametric covariogram models. Unfortunately, the estimates of covariogram parameters have no closed form and so are computed numerically. If we can estimate the correlation structure as well, then we will use knowledge of the covariogram model which describes the dependence structure in the SPB method. For spatial data the MBB method is usually used to estimate

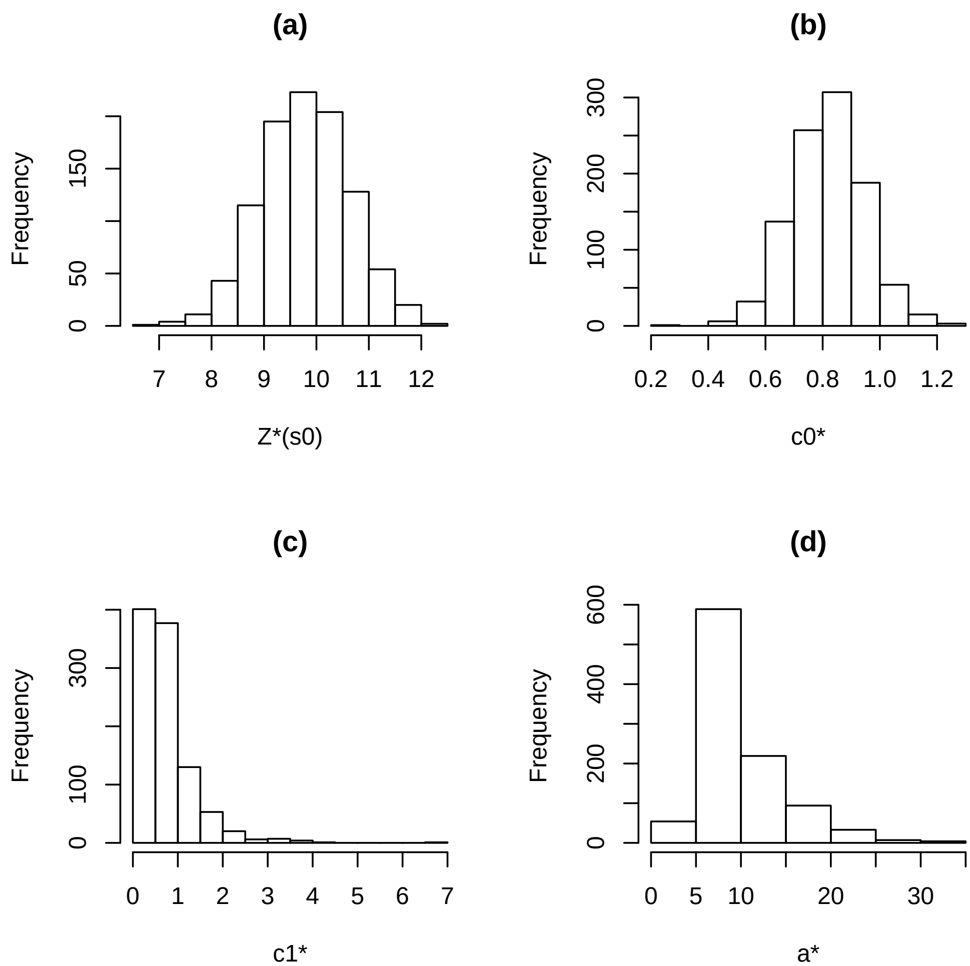

Figure 2: Histogram of (a) plug-in kriging and variogram parameters estimator: (b) nugget effect, (c) partial sill and (d) range for coal-ash data.

the precision measures of the estimators. However, as already pointed out, the MBB method has limitations and weaknesses. We now summarize some advantages of the SPB method as compared with the MBB method:

The precision of the MBB estimators is related to the optimal block size β_n^{opt} in (7) which depends on unknown parameters which are difficult to estimate. In our simulations it is clear that the optimal block size differs for various estimators or precision measures. Note also that the optimal block size determination is impossible for estimators that have no closed form (e.g. covariogram parameters estimator). For some data sets we may not be able to find the block size that satisfies $N = K\beta_n^d$. In other words, there is not always complete blocking and then $N_1 = K\beta_n^d < N$ is the total number of data-values in the resampled complete blocks. As a result, $N - N_1$ observations are ignored.

Establishing the consistency of MBB estimators and estimation of block size requires that the random field satisfies strong-mixing conditions. In the MBB method, our simulations indicate that the variance estimators $\hat{\sigma}^2$ are underestimated. Moreover, our simulations show that the MBB and SPB variance estimations $\hat{\sigma}^2$ are asymptotically unbiased and consistent. In this study, the SPB estimators are more accurate than the MBB estimator, for variance estimation of estimators in spatial data analysis, especially for stronger dependence structure and larger sample sizes. In the SPB method, we use the estimation of spatial correlation structure, therefore the SPB method will perform better than the MBB method. We are studying on comparison of estimation of distribution, spatial prediction interval and confidence interval by SPB and MBB methods.

Acknowledgments

The authors would like to thank the referees for their valuable comments and suggestions. Also partial support from Ordered and Spatial Data Center of Excellence of Ferdowsi University of Mashhad is acknowledged.

References

[1] Bose, A. (1988), Edgeworth correction by bootstrap in autoregression. *Annals of Statistics* **16**, 1709–1722.

[2] Bühlmann, P., Künsch, H.R. (1999), Comments on "Prediction of spatial cumulative distribution functions using subsampling". *Journal of the American Statistical Association* **94**, 97–99.

[3] Cressie, N. (1993), *Statistics for Spatial Data.* 2nd edn, Wiley, New York.

[4] Efron, B. (1979), Bootstrap method; another look at the jackknife. *The Annals of Statistics* **7**, 1–26.

[5] Efron, B., Tibshirani, R. (1993), *An Introduction to the Bootstrap*, Chapman and Hall, London.

[6] Freedman, D.A., Peters, S.C. (1984), Bootstrapping a regression equation: Some empirical results. *Journal of the American Statistical Association* **79**, 97–106.

[7] Hall, P. (1985), Resampling a coverage pattern. *Stochastic Processes and their Application* **20**, 231–246.

[8] Hastie, T.J., Tibshirani, R.J. (1990), *Generalized Additive Models*, Chapman and Hall, London.

[9] Journel, A.G., Huijbregts, C.J. (1978), *Mining Geostatistics*. Academic Press, London.

[10] Kent, J.T., Mardia, K.V. (1996), Spectral and circulant approximations to the likelihood for stationary Gaussian random fields. *Journal of Statistical Planning and Inference* **50**, 379–394.

[11] Kent, J.T., Mohammadzadeh, M. (1999), Spectral approximation to the

likelihood for an intrinsic random field. *Journal of Multivariate Analysis* **70**, 136–155.

[12] Lahiri, S.N. (2003), *Resampling Methods for Dependent data*, Springer-Verlag, New York.

[13] Lahiri, S.N., Furukawa, K., Lee, Y.-D. (2007), A nonparametric plug-in rule for selecting optimal block lengths for block bootstrap methods. *Statistical Methodology* **4**, 292–321.

[14] Liu, R.Y., Singh, K. (1992), Moving blocks jackknife and bootstrap capture weak dependence, *in* R. Lepage and L. Billard, eds, *Exploring the limits of the bootstrap*, Wiley, New York, pp. 225–248.

[15] Mardia, K.V., Marshal, R.J. (1984), Maximum likelihood estimation of models for residuals covariance in spatial regression. *Biometrika* **71**, 135–146.

[16] Mardia, K.V., Southworth, H.R., Taylor, C.C. (1999), On bias in maximum likelihood estimators. *Journal of Statistical Planning and Inference* **76**, 31–39.

[17] Nordman, D., Lahiri, S.N. (2003), On optimal spatial subsample block size for variance estimation. *The Annals of Statistics* **32**, 1981–2027.

[18] Politis, D.N., Paparoditis, E., Romano, J.P. (1998), Large sample inference for irregularly spaced dependent observations based on subsampling. *Sankhya, Series A, Indian Journal of Statistics* **60**, 274–292.

[19] Politis, D.N., Paparoditis, E., Romano, J.P. (1999), 'Resampling marked point processes', in S. Ghosh, ed., *'Multivariate Analysis, Design of Experiments, and Survey Sampling: A Tribute to J. N. Srivastava'* Marcel Dekker, New York, pp. 163–185.

[20] Politis, D.N., Romano, J.P. (1993), Nonparametric resampling for homogeneous strong mixing random fields. *Journal of Multivariate Analysis* **47**, 301–328.

[21] Politis, D.N., Romano, J.P. (1994), Large confidence regions based on subsamples under minimal assumptions. *The Annals of Statistics* **22**, 2031–2050.

[22] Politis, D.N., Romano, J.P., Wolf, M. (1999), *Subsampling*, Springer-Verlag, New York.

[23] Possolo, A. (1991), Subsampling a random field, *in* A. Possolo, ed., *'Spatial Statistics and Imaging'*, IMS Lecture Notes Monograph Series, 20, Institute of Mathematical Statistics, Hayward, CA, pp. 286–294.

[24] Sherman, M. (1996), Variance estimation for statistics computed from spatial lattice data. *Journal of the Royal Statistical Society, Series B* **58**, 509–523.

[25] Sherman, M., Carlstein, E. (1994), Nonparametric estimation of the moments of the general statistics computed from spatial data. *Journal of the American Statistical Association* **89**, 496–500.

[26] Singh, K. (1981), On the asymptotic accuracy of the Efron's bootstrap. *The Annals of Statistics* **9**, 1187-1195.

[27] Zimmerman, D.L., Cressie, N. (1992), Mean squared prediction error in the spatial linear model with estimated covariance parameters. *Annals of the Institute of Statistical Mathematics* **44**, 27–43.