A Parametric Spatial Bootstrap

Liansheng Tang[a]; William R. Schucany*[b]; Wayne A. Woodward[b]; and Richard F. Gunst[b]

[a] *Department of Statistics, George Mason University, Fairfax, VA 22030, USA*

[b] *Department of Statistical Science, Southern Methodist University, P.O. Box 750332, Dallas, TX 75275, USA*

* Corresponding author. Department of Statistical Science, Southern Methodist University, P.O. Box 750332, Dallas, TX 75275, USA

E-mail address: schucany@mail.smu.edu.

Abstract: The classic bootstrap uses valid resamples whenever the observations are independent and identically distributed. Data from a spatial region usually have a correlated structure. If the bootstrap is naively applied to spatially dependent data, these correlations are lost. This paper proposes a new parametric spatial bootstrap. The proposed method is motivated by an analysis of brain images in Spence et al. (2007). The new method combines spatial modeling and the parametric bootstrap to produce valid resamples of spatially correlated, normally distributed data. The coverage of confidence intervals from the proposed method is estimated and compared to an existing spatial bootstrap method proposed by Solow (1985). The two procedures are simulated for several two-dimensional images. The new approach guarantees positive definiteness of covariance matrices and has much better coverage than the existing spatial bootstrap approach.
Keywords: Cholesky; dependent data; Gaussian; resampling; semivariogram

1 Introduction

In spatial data any pair of observations may be correlated due to their positions in space. Such examples can be found in many areas, including brain imaging, geology, and environmental monitoring. In these cases data are collected from various spatial locations and statistical models are often used to represent the dependence between measurements at different locations. The spatial domain may be one, two, or three dimensional. Spatial images can be analyzed regardless of whether the spatial domain is a continuous region or a regular grid. For ease of presentation, we only consider regular grids in this article.

If spatial data have very weak correlations, then a simple basic bootstrap method can yield approximately valid resamples of the observed data ("bootstrap" refers to "resampling with replacement"; see Davison and Hinkley, 1997 for more details). Bootstrap methods provide a simple way to calculate confidence intervals. Bootstrap methods have also been shown in finite samples to provide better coverage than an analytic solution (DiCiccio and Efron,1996). However, in strongly correlated spatial data, naively ignoring the spatial correlation inherent in the data and directly resampling the observations will inevitably destroy the correlation structure among the original observations. In addition, while a point estimator $\hat{T}$ may not be difficult to derive for a parameter of interest from observed spatial data, a confidence interval usually involves complicated calculations of the sampling distribution for $\hat{T}$. Consequently, fidelity to the correlation structure is critically important.

To improve the bootstrap method for dependent data, several different block bootstrap methods (Carlstein, 1986; Kunsch, 1989; Liu and Singh,1992; Politis and Romano, 1992, 1994) have been proposed to nonparametrically replicate the dependence structure of the observations in the resamples. The block bootstrap preserves the dependence structure within the blocks, but distorts the correlation between blocks. Hall et al. (1995) pointed out that the bias and the variance of a block bootstrap estimator are heavily influenced by the block length.

A recent paper by Spence et al. (2007) presented a comprehensive analysis of single-photon emission computed tomography (SPECT) brain images. Based on the authors' findings, the intervoxel correlation structure in SPECT images can be estimated using Gaussian semivariograms. Motivated by the ability to model spatial correlation patterns within deep brain regions, we introduce a parametric spatial bootstrap (PSB) method. Our method adapted a parametric approach

in Sjöstedt-de Luna and Young (2003) and a spatial bootstrap method in Solow (1985) to obtain valid resamples from a Gaussian process or a lognormal process. Our method accounts for the spatial correlation in the data by estimating the correlation structure and then imposing it in the resamples. The proposed method offers a correct way to obtain confidence limits based on the statistic $\hat{T}$. The advantage of our PSB method over Sjostedt-de Luna and Young's (2003) method is that PSB offers more flexibility in spatial modeling. Based on spatial residuals, our method is not only able to obtain interval estimates for a constant mean as in Sjostedt-de Luna and Young (2003), but also is able to obtain confidence intervals on nonlinear trend surfaces of spatial data.

This paper is organized as follows. Section 2 reviews basic terms and issues in spatial statistics. Section 3 introduces PSB algorithm. Section 4 gives the algorithms for the spatial processes used in our simulation study. Different spatial data simulation techniques are discussed with respect to their computational efficiency. The simulation results of the proposed method and an existing method are compared with respect to the coverage of their confidence intervals on 2-D data with constant means. Section 5 contains some conclusions and discussion of this parametric version of the spatial bootstrap. Technical issues on computational speeds are relegated to the appendix.

2 Spatial Statistics

2.1 Semivariogram

Suppose a location $\mathbf{s} = (s_1, s_2, ..., s_n)'$ is a vector in n-dimensional Euclidian space $\Re^n$, and $Z(\mathbf{s})$ is a random magnitude at the location $\mathbf{s}$. A random field is

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

where $\mathbf{s}$ varies over the index set D, which is a subset of $\Re^n$. The random field $Z(\mathbf{s})$ is decomposed into a deterministic mean function $\mu(\mathbf{s})$ and a correlated error process $\delta(\mathbf{s})$ as

$$Z(\mathbf{s}) = \mu(\mathbf{s}) + \delta(\mathbf{s}), \quad \mathbf{s} \in D,$$

(see Cressie, 1993, Section 2.4). The error process $\delta(\cdot)$ is assumed to be a zero-mean intrinsically stationary spatial process. The variogram $2\gamma(d)$ is defined to be

$$2\gamma(d) = var\{Z(\mathbf{s}) - Z(\mathbf{s}')\}, \quad \text{where} \quad \|\mathbf{s} - \mathbf{s}'\| = d \quad \text{and} \quad \mathbf{s}, \mathbf{s}' \in \mathbf{D}.$$

There are a large number of conditionally negative definite semivariogram models in the literature on spatial modeling; e.g., Cressie (1993, Section 2.3.1). Two quite popular isotropic semivariogram models are considered in this paper. The first is the Gaussian model

$$\gamma^G(d;\underline{\theta}) = \begin{cases} 0 & \text{if } d = 0 \\ \theta_1 + \theta_2\{1 - \exp(-(d/\theta_3)^2)\} & \text{if } d > 0, \end{cases}$$

where $\underline{\theta} = (\theta_1, \theta_2, \theta_3)$, and $\theta_1, \theta_2, \theta_3 \geq 0$. The shape of the Gaussian model is displayed in Figure 1 using the nugget parameter $\theta_1 = 0$, the sill (because $\theta_1 = 0$) $\theta_2 = 1$ and the range parameter $\theta_3 = 3$. The vertical dashed line shows the effective range, $\sqrt{3}\theta_3$, and the horizontal dashed line at unit variance is the sill θ_2, where γ^G approaches its limit. The zero nugget is shown as the horizontal dotted line. It is clear from the plot that this semivariogram strictly increases with distance and becomes flat when the distance d gets close to the range.

The other semivariogram model investigated here is the exponential semivariogram model

$$\gamma^E(d;\underline{\theta}) = \begin{cases} 0 & \text{if } d = 0 \\ \theta_1 + \theta_2\{1 - \exp(-d/\theta_3)\} & \text{if } d > 0, \end{cases}$$

where $\underline{\theta} = (\theta_1, \theta_2, \theta_3)$, and $\theta_1, \theta_2, \theta_3 \geq 0$.

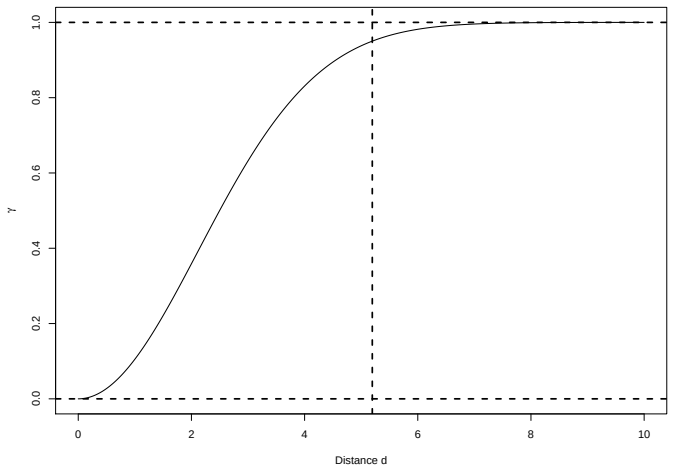

Figure 1. Gaussian semivariogram model with zero nugget, unit sill, and a range of 3.

Initially, it seemed that a nonparametric approach would be preferable in our applications. However, realistic spatial data were simulated and the classic semivariogram estimates obtained.

Shapiro-Botha (1991) semivariogram fits were obtained using R programs from Statlib nonparametric estimation algorithms. The corresponding covariance matrix estimates were calculated and then Cholesky decompositions were attempted on these covariance matrices. It turned out that some of the calculated matrices were not positive-definite so that the algorithm failed to give the Cholesky decomposition.

Even though the selection of a conditionally negative-definite parametric semivariogram family can be subjective, once the parameters of the semivariogram model are obtained, the corresponding fitted semivariogram model provides valid covariance matrix estimates. This property, coupled with results (e.g., Stein 1988) that show the critical feature is not the model selected but good estimation for small lags, led to the decision to use parametric semivariograms for further study.

2.2 Estimation of semivariogram parameters

For the example of a constant mean $\mu(\mathbf{s}) = \mu$, the method-of-moments estimator for the variogram was introduced in Mathéron (1962). It is referred to as the classical variogram estimator by Cressie (1993, Section 2.4) and is given by

$$2\hat{\gamma}(d) = \frac{1}{|N(d)|} \sum_{N(d)} (Z(\mathbf{s}_i) - Z(\mathbf{s}_j))^2,$$

where $N(d) = \{(\mathbf{s}_i, \mathbf{s}_j) : \|\mathbf{s}_i - \mathbf{s}_j\| = d \quad \text{for} \quad i, j = 1, ..., K\}$ includes all distinct pairs of locations, $\mathbf{s}_i, \mathbf{s}_j \in D \subset \Re^n$ that are d units apart and $|N(d)|$ is the number of these pairs.

Several parametric methods for fitting a semivariogram model to estimate semivariogram values $\hat{\gamma}(d)$ are summarized in Cressie (1993, Section 2.4.3). Denote the semivariogram estimator as a vector $\hat{\gamma}$:

$$\hat{\boldsymbol{\gamma}} = (\hat{\gamma}(d_1), \hat{\gamma}(d_2), \cdots, \hat{\gamma}(d_K))'$$

at K lags. For a given semivariogram model $\boldsymbol{\gamma}(\underline{\theta})$, the ordinary least squares (OLS) method obtains the value $\underline{\theta}$ that minimizes

$$(\hat{\boldsymbol{\gamma}} - \boldsymbol{\gamma}(\underline{\theta}))'(\hat{\boldsymbol{\gamma}} - \boldsymbol{\gamma}(\underline{\theta})).$$

If the correlation among estimator values $\hat{\gamma}(d)$ at different lags is taken into account, the generalized least squares (GLS) criterion minimizes $(\hat{\boldsymbol{\gamma}} - \boldsymbol{\gamma}(\underline{\theta}))' V_G^{-1}(\underline{\theta})(\hat{\boldsymbol{\gamma}} - \boldsymbol{\gamma}(\underline{\theta}))$, where $V_G(\theta) = \text{var}(\hat{\boldsymbol{\gamma}})$. However, calculation of all elements in $V_G(\hat{\theta})$ is not a trivial task. Thus, a simpler approach, weighted

least squares (WLS), is widely adopted and only requires the estimation and inversion of a $K \times K$ diagonal matrix $V_W(\theta)$, where

$$V_W(\theta) = \text{diag}\{\text{var}(\hat{\gamma}(d_1)), \text{var}(\hat{\gamma}(d_2)), ..., \text{var}(\hat{\gamma}(d_K))\}.$$

Although GLS can be superior, Cressie (1985) and Grondona and Cressie (1992) show that there is often negligible loss in efficiency using WLS instead of GLS. Thus, WLS is generally the estimator of choice due to its efficiency and ease of calculation.

The weights in WLS are given by the reciprocals of the variances of the semivariogram estimator $\hat{\boldsymbol{\gamma}}$. The diagonal terms of V_W can be approximated by $\text{var}(\hat{\gamma}(d_i)) \simeq 2(\hat{\gamma}(d_i; \underline{\theta}))^2/|N(d_i)|$. Thus, WLS obtains the estimate of $\underline{\theta}$ such that

$$\widehat{\underline{\theta}} = \underset{\underline{\theta}}{\text{argmin}} \left\{ \sum_{i=1}^{K} w(d_i)\{\hat{\gamma}(d_i) - \gamma(d_i; \underline{\theta})\}^2 \right\},$$

where

$$w(d_i) = 1/\text{var}(\hat{\gamma}(d_i)) = |N(d_i)|/2(\hat{\gamma}(d_i; \underline{\theta}))^2 \quad \text{for} \quad i = 1, ..., K.$$

The above equations illustrate that WLS gives more weight to squared deviations that have more pairs of observations, which are at the smaller lags d_i. In particular, the latter condition means that the deviations are weighted more if their lags are closer to the origin, improving the fit of the semivariogram near the origin. Stein (1988) argues the critical importance of estimating semivariogram models well at small lags.

An initial value of the vector $\underline{\theta}$ can be obtained from the sample semivariogram values $\hat{\gamma}(d_i)$ using OLS to obtain the initial estimate $\widehat{\underline{\theta}}^{(0)}$. Then $\widehat{\underline{\theta}}^{(0)}$ is substituted into the matrix V_W to calculate $V_W(\widehat{\underline{\theta}}^{(0)})$, which is used in the $w(d_i)$ to produce $\widehat{\underline{\theta}}^{(1)}$ and the resulting $V_W(\widehat{\underline{\theta}}^{(1)})$. The above procedures are iterated until the parameter estimates converge within a specified tolerance. The final estimate of the parameters, $\widehat{\underline{\theta}}$, is then substituted in the semivariogram model $\gamma(d; \underline{\theta})$, yielding a valid semivariogram estimator $\hat{\gamma}(d; \widehat{\underline{\theta}})$. In practice, there are several nonlinear least squares algorithms in R including optim and nls routines, which include WLS as a special case.

2.3 Estimation of the Covariance Matrix

The covariogram $C(d)$ is often given by $C(d) = cov\{Z(\mathbf{s}), Z(\mathbf{s} + d)\}$. There is a close relationship between the semivariogram $\gamma(d)$ and the covariogram $C(d)$. If $\delta(\cdot)$ is second-order stationary and

$\mu(\mathbf{s}) = \mu$, then

$$\gamma(\|\mathbf{s}_1 - \mathbf{s}_2\|) \;=\; C(0) - C(\|\mathbf{s}_1 - \mathbf{s}_2\|). \tag{1}$$

Furthermore if $\gamma(\cdot)$ and $C(\cdot)$ are functions only of Euclidean distance d, then the above equation implies that $\gamma(d) = C(0) - C(d)$.

The parametric semivariogram method gives a valid estimator $\hat{\gamma}(d; \underline{\hat{\theta}})$. Equation (1) suggests a relationship between the estimator, $\hat{C}(d)$ and $\hat{\gamma}(d; \underline{\hat{\theta}})$, namely,

$$\hat{C}(d) \simeq \hat{C}(0) - \hat{\gamma}(d; \underline{\hat{\theta}}),$$

where $\hat{C}(0) = \hat{\sigma}_\delta^2 = \text{sill}$. An estimate, $\hat{\Sigma}$, of the $n \times n$ covariance matrix, $\Sigma = [cov\{Z(\mathbf{s}_i), Z(\mathbf{s}_j)\}]$, can then be obtained from

$$\hat{\Sigma}(i,j) = \hat{C}(d_{ij}), \tag{2}$$

where $d_{ij} = \|\mathbf{s}_i - \mathbf{s}_j\|$ for $i, j = 1, ..., n$.

The method-of-moments covariogram estimator, $\hat{C}(\cdot)$, yields an estimator $\hat{\Sigma}$ of the covariance matrix, but the resulting covariance matrix is not necessarily positive definite. Therefore it cannot be directly used for the parametric spatial bootstrap method described in Section 3. Fits to the Gaussian or exponential semivariogram models do produce estimated covariance matrices that are positive definite.

3 Parametric Spatial Bootstrap

Direct application of the nonparametric bootstrap method fails to provide valid resamples whenever there is correlation in either time series or spatial data. When this bootstrap is applied to correlated data, it randomizes either the residuals or the observations and destroys the correlation pattern inherent in the joint distribution.

Solow (1985) proposes a spatial bootstrap (SB) method to obtain spatial resamples from the original data. For realizations from the spatial models in Section 2.1, the estimates for the deterministic components $\hat{\boldsymbol{\mu}}$ can be obtained using various methods (e.g., Cressie 1993, Section 3.1). Next he calculates the predicted spatial error process as

$$\hat{\boldsymbol{\delta}} \;=\; \{Z(\boldsymbol{s}_1) - \hat{\mu}(\boldsymbol{s}_1), ..., Z(\boldsymbol{s}_n) - \hat{\mu}(\boldsymbol{s}_n)\}.$$

The covariance matrix is estimated from $\hat{\boldsymbol{\delta}}$ and the SB method is based on the Cholesky decomposition given by $\hat{\Sigma}_{SB} = \hat{L}_{SB}\hat{L}_{SB}^T$, where $\hat{L}$ is a lower triangular $n \times n$ matrix. Appendix 1 compares computation times for these Cholesky and eigenvalue decompositions. The SB method then uses the Cholesky decomposition matrix inverse, $\hat{L}_{SB}^{-1}$, to decorrelate the spatial error sequence

$$(\hat{\varepsilon}_1, \hat{\varepsilon}_2, ..., \hat{\varepsilon}_n) \equiv \hat{\boldsymbol{\varepsilon}} \equiv \hat{L}_{SB}^{-1}\hat{\boldsymbol{\delta}} \tag{3}$$

and then centers the $\hat{\boldsymbol{\varepsilon}}$ to obtain

$$\widetilde{\varepsilon}_i = \hat{\varepsilon}_i - \frac{1}{n}\sum_{j=1}^{n}\hat{\varepsilon}_j$$

for $i = 1, ..., n$. The decorrelated and centered residuals $\widetilde{\varepsilon}_1, \widetilde{\varepsilon}_2, ..., \widetilde{\varepsilon}_n$ are bootstrapped to provide the resampled residuals $\boldsymbol{\varepsilon}_{SB}^* = (\varepsilon_1^*, ..., \varepsilon_n^*)$. The SB resample is obtained by transforming to recorrelate the bootstrapped residuals

$$\boldsymbol{Z}_{SB}^* = \hat{\boldsymbol{\mu}} + \hat{L}_{SB}\boldsymbol{\varepsilon}_{SB}^*. \tag{4}$$

However, the SB method may not be valid because the predicted spatial error process $\hat{\boldsymbol{\delta}}$ may not retain the same correlation structure as the original spatial data. Confidence intervals based on (4) tend to suffer from significant undercoverage because the "decorrelated" residuals are not sufficiently uncorrelated in practice. In this paper we introduce a valid parametric spatial bootstrap (PSB) approach beginning with uncorrelated errors:

$$\boldsymbol{\varepsilon}_{PSB}^* = (\varepsilon_1^*, \varepsilon_2^*, ..., \varepsilon_n^*), \quad \text{where} \quad \varepsilon_j^* \sim N(0,1) \quad \text{for} \quad j = 1, ..., n.$$

We estimate the semivariogram parameter $\widehat{\underline{\theta}}_{PSB}$ and resulting covariance matrix $\hat{\Sigma}_{PSB}$ from the original spatial data $Z(\mathbf{s})$. Next, the spatial resamples are transformed by

$$\boldsymbol{Z}_{PSB}^* = \hat{\boldsymbol{\mu}} + \hat{L}_{PSB}\boldsymbol{\varepsilon}_{PSB}^*, \tag{5}$$

where $\hat{L}_{PSB}$ is the Cholesky decomposition matrix of $\hat{\Sigma}_{PSB}$. Finally, we calculate the statistic of interest, $\hat{T}^*$, from $\boldsymbol{Z}_{SB}^*$ or $\boldsymbol{Z}_{PSB}^*$. The above procedures are repeated B times to estimate the sampling distribution of $\hat{T}$.

Our PSB method does not obtain the errors $\boldsymbol{\varepsilon}^*$ by decorrelating the spatial error process as in the SB algorithm. Instead, the errors are independently generated from a standard normal distribution and transformed by (5). More importantly, because $\hat{\boldsymbol{\delta}}$ may not retain the original spatial structure, we estimate the covariance matrix directly from the original data. This faithfully mimics the spatial

model introduced in Section 4.1. The theoretical foundation is given in Sjöstedt-de Luna and Young (2003).

4 SIMULATIONS

4.1 Simulation algorithm

Cressie (1993, Section 3.6) summarizes several spatial data simulation procedures that are based on the Cholesky decomposition (Cressie and Laslett, 1987) or eigenvalue decomposition. The Cholesky decomposition procedure in Section 3 allows the covariance matrix Σ to be factored into the product of two matrices

$$\Sigma = LL^T.$$

The eigenvalue decomposition calculates the square root of Σ based on

$$\Sigma = Q \text{ diag}\{\lambda_1, ..., \lambda_n\}Q^T,$$

where $\lambda_1, ..., \lambda_n$ are the eigenvalues of Σ and the matrix Q consists of the corresponding eigenvectors of Σ. A matrix $\Sigma^{1/2}$ is obtained by

$$\Sigma^{1/2} = Q \text{ diag}\{\lambda_1^{1/2}, ..., \lambda_n^{1/2}\}Q^T.$$

The Cholesky decomposition is preferred not only because it is widely available from most computer packages with matrix operations, but also because it is more computationally efficient than the eigenvalue decomposition. We report a comparison in the Appendix justifying our use of the Cholesky decomposition.

The algorithm for simulating spatial data in this study is:

1. Choose a constant mean, μ, for the random field (μ=1, 2 or 4).
2. Choose one of the two covariance matrices, Σ^P(power) and Σ^G(Gaussian) (details in Section 4.2).
3. Obtain the Cholesky matrix L_P such that $\Sigma^P = L_P L_P^T$ and similarly for L_G.
4. Generate independent random variables $\boldsymbol{\varepsilon}$ from the standard normal distribution,

$$\boldsymbol{\varepsilon} = (\varepsilon_1, \varepsilon_2, ..., \varepsilon_n), \quad \text{where} \quad \varepsilon_j \sim N(0,1) \quad \text{for} \quad j = 1, ..., n.$$

5. Simulate the spatial data $\boldsymbol{Z}$ using the relationship

$$\boldsymbol{Z} = \mu + L_P \boldsymbol{\varepsilon}.$$

4.2 Data simulation

Two-dimensional spatial data $\boldsymbol{Z}$ are considered here. For simplicity $\boldsymbol{Z}$ has a constant mean μ over the entire field. In the Solow's SB method, the spatial errors $\hat{\boldsymbol{\delta}}$ are estimated by subtracting the sample mean from the realization of the random field,

$$\hat{\boldsymbol{\delta}} = \{Z(\boldsymbol{s}_1) - \frac{1}{n}\sum_{j=1}^{n} Z(\boldsymbol{s}_j), ..., Z(\boldsymbol{s}_n) - \frac{1}{n}\sum_{j=1}^{n} Z(\boldsymbol{s}_j)\}.$$

Two stationary covariance structures are studied here. The first, Σ^P, has the simple power form in which each location is correlated with all other locations according to

$$C(d) = \sigma^2 \rho^d,$$

where d is the Euclidean distance between the locations and σ^2 is the variance of the spatial process. The corresponding power semivariogram is included in the simulation to investigate how the semivariogram models, Gaussian and exponential, perform when the model is incorrectly specified.

Without loss of generality, $\sigma^2 = 1$ in these simulations. The power covariance matrix may be written as $\Sigma^P(i,j) = \rho^{d_{ij}}$, where $d_{ij} = \|\mathbf{s}_i - \mathbf{s}_j\|$ for $i, j = 1, ..., n$. For $\rho = 0.7$ the covariance matrix is illustrated in a grayscale map, motivated by Stein (1988), in Figure 2(a) for an 8×8 spatial image by ordering the pixels from top to bottom and from left to right, and calculating the correlation between each pair of pixels. The bands in the 64×64 correlation image in Figure 2(a) indicate that the correlation is strong among local pixels and weaker with increased distance between pixels. The correlation pattern in this map has box-like shapes, because the ordered pixels from a two-dimensional image do not have strictly decreasing correlation with location in the correlation matrix. For example, the 1st and the 9th ordered pixels from an 8×8 two-dimensional image do not have the correlation associated with pixels eight units apart. To the contrary, they are neighbors in the image and have the strongest correlation of only one unit apart. The theoretical power semivariogram is given by $\gamma(d) = 1 - \rho^d$.

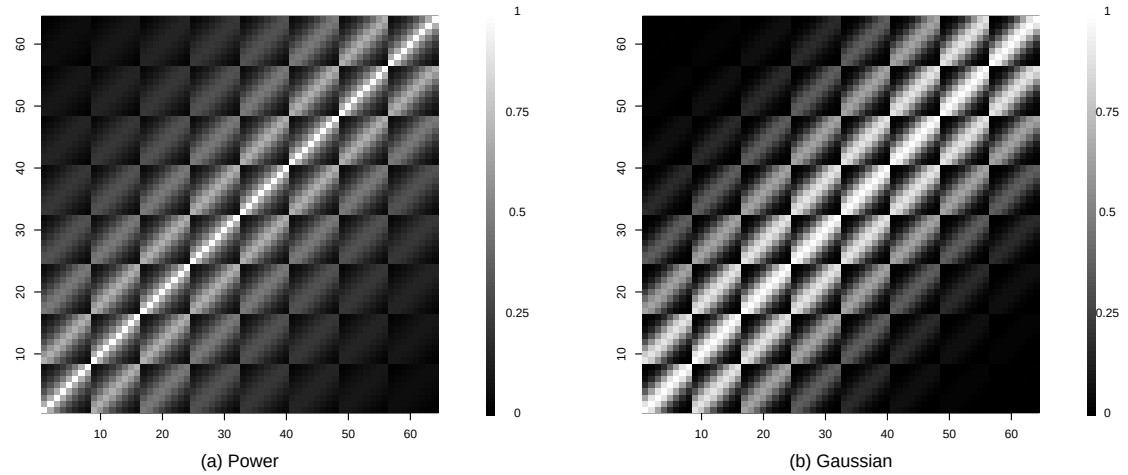

Figure 2. Correlation maps for Power and Gaussian covariance structures. The grayscale bar on the right is the magnitude of the correlations for each pair of pixels. (a) A Power covariance matrix : $C(d) = \rho^d$ with $\rho = 0.7$; (b) A Gaussian covariance matrix.

The second covariance structure for data simulation is based on the Gaussian semivariogram model whose covariance matrix is

$$\begin{aligned}\Sigma^G(i,j) &= C(0) - \gamma^G(d_{ij};\underline{\theta}) \\ &= \sigma^2 - (\theta_1 + \theta_2\{1 - \exp(-(d_{ij}/\theta_3)^2)\}),\end{aligned}$$

where $d_{ij} = \|\mathbf{s}_i - \mathbf{s}_j\|$ for $i, j = 1, ..., n$. In the simulation study the nugget parameter is set to $\theta_1 = 0$ for continuity of the Gaussian semivariogram at the origin, the sill $\sigma^2 = \theta_2 = 1$, and the range parameter $\theta_3 = 3$, which means that the locations about 5 units apart are approximately uncorrelated. This range has been shown to be relevant within deep brain regions in SPECT studies (Spence et al., 2007). The covariance matrix for this Gaussian semivariogram is shown in a grayscale map in Figure 2(b) for an 8×8 spatial image. The correlation here decreases with distance somewhat slower than the one in Figure 2(a) for the power covariance matrix.

Figures 3(a) and 3(b) illustrate simulated 8×8 2D spatial processes and their perspective plots for both the power semivariogram with $\rho = 0.7$ and the Gaussian semivariogram. The top two panels are the maps of the 2D images. From the perspective plots, it can be seen that the power semivariogram of the spatial process in Figure 3(c) has much more abrupt changes than the smoother Gaussian process in Figure 3(d).

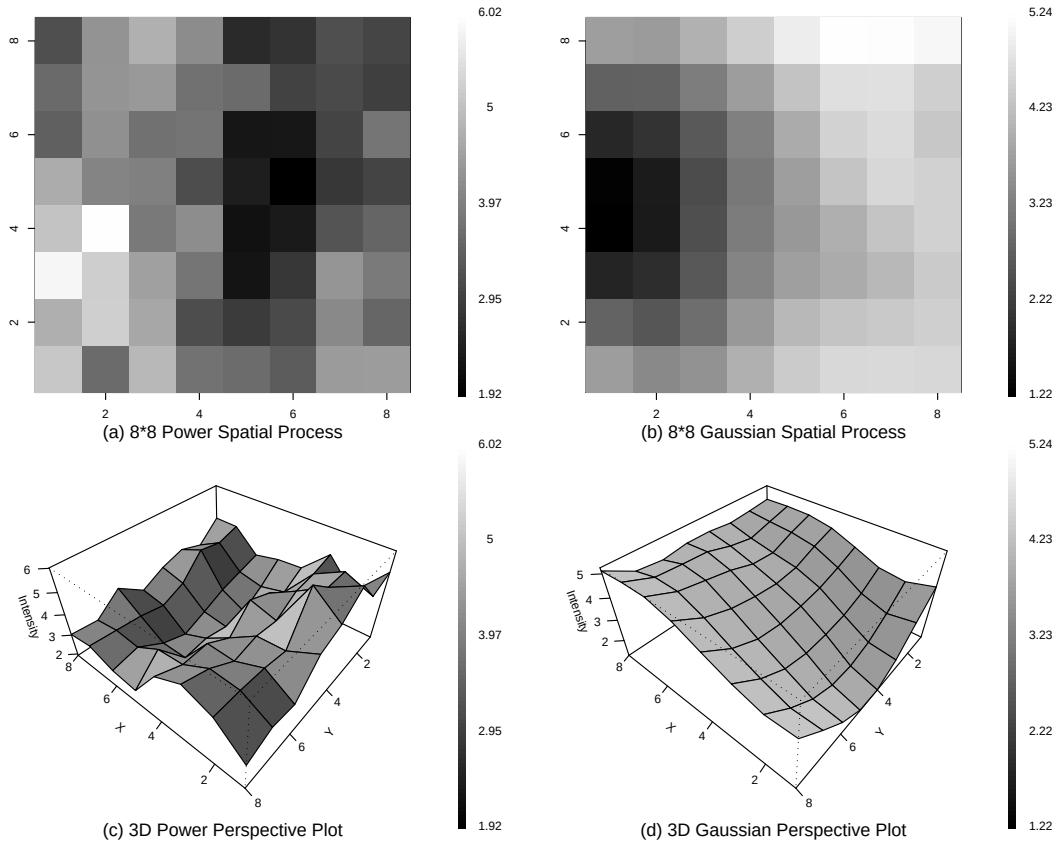

Figure 3. The simulated 2D spatial processes and their corresponding perspective plots. The constant mean for both processes is 4 and the image size is 8×8. Panels (a) and (c) give a power process image in a grayscale map and in a 3D setting, respectively; Panels (b) and (d) show a Gaussian process in a grayscale map and in a 3D setting, respectively.

Table 1: Estimated confidence interval coverages in % using parametric spatial bootstrap (PSB) and Solow's spatial bootstrap (SB). The Power (Pow) and Gaussian (Gau) columns under bootstrap types indicate the semivariogram models generated. The rows labeled Gaussian (Gau) and Exponential (Exp) for each image size indicate the semivariogram (SV) type fit to the data. Simulations are from the Power semivariogram model with $\rho = 0.7$ and a Gaussian model with $\theta_1 = 0, \theta_2 = 1$ and $\theta_3 = 3$. The nominal confidence levels are 90% with standard errors of 0.95%.

		$\mu = 0$		$\mu = 1$		$\mu = 4$	
				PSB			
Size	SV	Pow	Gau	Pow	Gau	Pow	Gau
4×4	Gau	92	74	91	70	92	73
	Exp	88	89	88	85	91	84
8×8	Gau	92	89	91	91	92	89
	Exp	91	89	90	80	88	90
				SB			
4×4	Gau	82	75	82	73	83	76
	Exp	85	79	85	76	82	78
8×8	Gau	80	75	75	74	78	74
	Exp	77	79	74	74	77	76

For each simulated two-dimensional data set the PSB and SB methods are repeated 499 times to obtain approximate 90% confidence limits for the mean μ. Both Gaussian and exponential semivariogram models are fit using WLS estimators. The 2-factor simulation design accounts for three different constant means, $\mu = 0, 1, 2$, and two image sizes, 4×4 and 8×8. All computations are performed in the R language.

4.3 Simulation results

Table 1 gives simulation results at the nominal confidence level of 90% ($SE \approx 0.95\%$) for $\mu = 0$, 1 and 4, respectively. The SB method has coverage of about 80%. The reason for the undercoverage using SB is that the estimated covariance matrix does not decorrelate the spatial error process. We

carried out a small experiment to investigate the correlation of decorrelated error processes. We generated 1000 8×8 Gaussian processes with mean 1. Each simulated process was decorrelated using its estimated Gaussian semivariogram and the procedure in (3). We then used the routine "Moran.I" in the R package "ape" to estimate the Moran's I measure (Moran, 1950), and found that 92.3% of the error processes have significant spatial autocorrelations using $\alpha = 0.05$.

We also examined one realization of the "decorrelated" error. For realization, an 8×8 Gaussian process was generated using the Gaussian semivariogram $1 - \exp(-(d/3)^2)$ with a mean 4. The Gaussian semivariogram model was fit to the semivariogram values and then the covariance matrix was calculated. The spatial error process was decorrelated using the estimated inverse covariance matrix. Figure 4 shows one such approximately decorrelated error process for a typical realization. It is clear that some correlation remains. It follows that the resamples obtained by bootstrapping these "decorrelated" error processes lead to the undercoverage of the resulting confidence intervals.

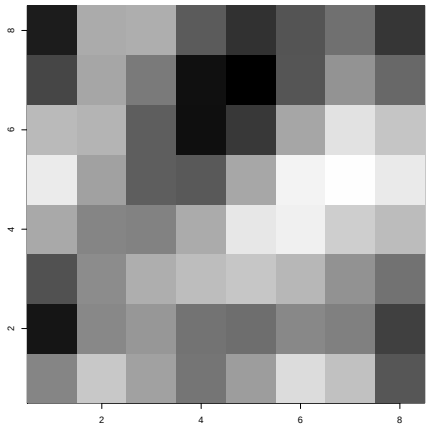

Figure 4. The "decorrelated" error. There are clear patterns of spatial correlation remaining (Estimated Moran's I is 0.0671 with p-value ≈ 0).

The PSB performs better than the spatial bootstrap under all conditions simulated except for 4×4 images with the Gaussian semivariogram. For smaller images the PSB using the exponential semivariogram provides excellent coverage for the means, even when the fitted semivariogram is not the same form as the theoretical one used to generate the data. Although the PSB using the Gaussian semivariogram has undercoverage for 4×4 images, its performance improves as the image

size increases to 8×8.

Figure 5(a) shows the Gaussian and exponential model fits for an 8×8 spatial images when the true semivariogram is of the power form. Figure 5(b) shows the semivariogram fits when the true underlying semivariogram is Gaussian. Note that the Gaussian fit approximates the empirical values well for both semivariogram models but the exponential fit is poor for small distances. Figure 5(a) and Figure 5(b) also illustrate some departures of the empirical semivariogram estimates from the true underlying models.

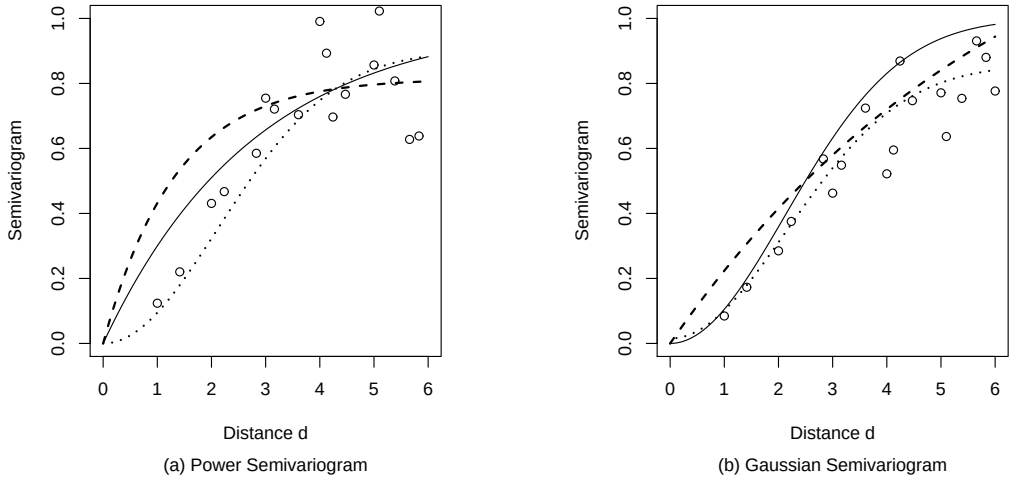

Figure 5. Semivariogram estimates for an 8×8 spatial image. The solid lines are the true semivariograms. The dashed line shows the results of exponential model fitting and the dotted one is the result of Gaussian model fitting. (a) A Power semivariogram $1 - 0.7^d$; (b) A Gaussian semivariogram $1 - \exp(-(d/3)^2)$.

In other simulation studies which were not presented, we applied our method to simulated lognormal processes. Following Cressie (1993, Section 3.2.2), we first transformed lognormal processes to Gaussian scale, and then fit semivariogram models with the transformed data. The coverage results are similar to those in Table 1.

5 Conclusions and Discussion

One of the main advantages of the PSB method over block bootstrap methods used in spatially correlated data is that it does not require one to determine block sizes and partition spatial data

into blocks. The new method also guarantees positive definite covariance matrices and offers better coverage than the spatial bootstrap method proposed in Solow (1985).

The authors used the R routine "optim" to fit the aforementioned semivariograms. Confidence intervals derived from our method give correct coverages of true means for normal errors. The performance of the proposed method relies on the semivariogram fit. When the simulated errors are non-Gaussian, the proposed method may not perform well if the data cannot be properly transformed to the normal as the converages are not close to the nominal level. For example, the authors applied the proposed method with "optim" on simulated spatial data with various types of non-Gaussian errors, including $Uniform(0,1)$, $t(10)$, $Beta(2,2)$, and $Cauchy(0,1)$ errors. We found that simulated coverages of PSB confidence intervals seem to be much lower than the nominal level, except for t-distributed errors which have close resemblance to Gaussian errors. Table 2 lists the coverages for these distributions. Thus, before relying on our parametric bootstrap method, it is recommended that one check the normality of spatial errors. One way to diagnose this is to first decorrelate spatial errors using the Cholesky decompositon matrix $\hat{L}$ estimated in Section 3, and then construct a normal quantile quantile plot to check for nonnormality of the residuals. Other methods for checking the normality of residuals can be found in Barry (1996).

Table 2: Estimated confidence interval coverages using PSB for Uniform (Unif) and t-distributed (t) errors. The nominal confidence levels are 90% with standard errors of 0.95%.

			$\mu = 0$	
Size	Unif	t	Beta	Cauchy
4× 4	57	76	66	81
8×8	48	85	50	51
			$\mu = 1$	
4× 4	61	75	65	79
8×8	47	87	50	53
			$\mu = 4$	
4× 4	58	75	60	78
8×8	49	85	59	58

APPENDIX 1: Computation speeds of two decompositions

Table 3 gives a comparison of the two decompositions on different sizes of Gaussian covariance matrices. The calculations are performed using the R package under the Microsoft Windows XP operation system on a laptop computer with Pentium M 1.6 GHz processor and 512 MB SDRAM memory. The Cholesky decomposition runs about 10 times faster than the eigenvalue decomposition on the higher dimensions. The covariance matrices for spatial data usually have large dimensions. If the size of a two dimensional image is $n \times n$, the corresponding covariance matrix is $n^2 \times n^2$. As n gets large, the covariance matrix quickly grows in size. A computationally efficient algorithm is desired to handle such a large number of matrix operations, and hence we use the Cholesky decomposition in this paper.

	64×64	256×256	2304×2304	4096×4096
Cholesky	0.01	0.2	68.08	438.58
Eigenvalue	0.03	0.72	626.17	4917.47

Table 3: Computation time in seconds for the Cholesky and eigenvalue decompositions. The first row lists the covariance matrix sizes. The decomposition methods are given on subsequent rows.

ACKNOWLEDGEMENTS

This work was part of the first author's PhD dissertation at Southern Methodist University. It was supported by the U.S. Army Medical Research and Materiel Command numbers DAMD17-97-2-7025 and DAMD17-01-1-0741 through a consortium agreement with the University of Texas Southwestern Medical Center at Dallas, U.S. Public Health Service grant M01-RR00633, and by a grant from the Perot Foundation, Dallas, Texas. The content of this paper does not necessarily reflect the position or the policy of the U.S. government, and no official endorsement should be inferred.

REFERENCES

Barry, R.P. (1996), "A Diagnostic to Assess the Fit of a Variogram Model to Spatial Data." *Journal of Statistical Software*, **1**, 1-11.

Carlstein, E. (1986). "The Use of Subseries Methods for Estimating the Variance of a General Statistic from a Stationary Time Series." *Annals of Statistics*, **14**, 1171-1179.

Cherry, S., Banfield, J., and Quimby, W. (1996). "An Evaluation of a Nonparametric Method of Estimating Semivariograms of Isotropic Spatial Processes." *Journal of Applied Statistics*, **23**, 435-449.

Cressie, N. (1985). "Fitting Variogram Models by Weighted Least Squares." *Mathematical Geology*, **17**, 563-585.

Cressie, N. (1993). *Statistics for Spatial Data* (revised ed.), New York: John Wiley and Sons.

Cressie, N. and Laslett, GM (1987). "G. M. Random Set Theory and Problems of Modeling." *SIAM Review*, **29**, 557-574.

Cressie, N. and Grondona, M. O. (1992). "A Comparison of Variogram Estimation with Covariogram Estimation." *The Art of Statistical Science*, K.V. Mardia (ed.), Chinchester: Wiley, 191-208.

Davison, A.C. and Hinkley, D.V. (1997). *Bootstrap Methods and Their Application.* Cambridge University Press.

Genton, M. and Gorsich, D. (2002). "Nonparametric Variogram and Covariogram Estimation with Fourier-Bessel Matrices." *Computational Statistics & Data Analysis*, **41**, 47-57.

DiCiccio, T. and Efron, B. (1996). "Bootstrap Confidence Intervals." *Statistical Science*, **11**, 189-212.

Grondona, M. O. and Cressie, N. (1995). "Residuals Based Estimators of the Covariogram." *Statistics*, **26**, 209-218.

Hall, P. (1985). "Resampling A Coverage Pattern." *Stochastic Processes and their Applications*, **20**, 231-246.

Hall, P., Horowitz, J. L., and Jing, B. (1995). "On Blocking Rules for the Block Bootstrap with Dependent Data." *Biometrika,* **82**, 561-574.

Kunsch, H. R. (1989). "The Jackknife and the Bootstrap for General Stationary Observations."

Annals of Statistics, **17**, 1217-1261.

Liu, R. Y. and Singh, K. (1992). "Moving Blocks Jackknife and Bootstrap Capture Weak Dependence." *In Exploring the Limits of Bootstrap* (R. Lepage and L. Billard, eds.) 225-248. John Wiley, New York.

Mathéron, G. (1962). " Traite de Geostatistique Appliquee, Tome I." *Memoires du Bureaude Recherches Geologiques et Minieres,* No. 14 Editions Technip, Paris.

Mathéron, G. (1971). *The Theory of regionalized variables and its applications.* Les Cahiers du Centre de Morphologie Math'ematique, Fasc. 5, Centre de G'eostatistique, Fontainebleau.

Moran, P.A.P. (1950), "Notes on Continuous Stochastic Phenomena," Biometrika, 37, 1733.

Politis, D. and Romano, J. (1992). "A Circular Block Resampling Procedure for Stationary Data." In *Exploring the Limits of Bootstrap* (R. Lepage and L. Billard, eds.) 263-270. Wiley, New York.

Politis, D. and Romano, J. (1994). "Large Sample Confidence Regions Based on Subsamples under Minimal Assumptions." *Annals of Statistics*, **22**, 2031-2050.

Shapiro, A. and Botha, J. (1991). "Variogram Fitting with a General Class of Conditionally Nonnegative Definite Functions." *Computational Statistics & Data Analysis*, **11**, 87-96.

Sjöstedt-de Luna, S. and Young, A. (2003). "The Bootstrap and Kriging Prediction Intervals." *Scandinavian Journal of Statistics*, **30**, 175-192.

Solow, A. (1985). "Bootstrapping Correlated Data." *Journal of the International Association for Mathematical Geology*, **17**, 769-775.

Spence, J., Carmack, P., Gunst, R., Schucany, W., Woodward, W. and Haley, R. (2007). "Accounting for Spatial Dependence in the Analysis of SPECT Brain Imaging Data." *Journal of the American Statistical Association*, **102**, 464-473.

Stein, M. (1988). "Asymptotically Efficient Prediction of a Random Field with a Misspecified Covariance Function." *Annals of Statistics*, **16**, 55-63.