Subsampling Variance Estimation for Statistics Computed from Nonstationary Spatial Lattice Data

Magnus Ekström

Research Report
Centre of Biostochastics

Swedish University of Agricultural Sciences

Report 2005:7
ISSN 1651-8543

Subsampling Variance Estimation for Statistics Computed from Nonstationary Spatial Lattice Data

Magnus Ekström[1]

Centre of Biostochastics
Swedish University of Agricultural Sciences
SE-901 83 Umeå, Sweden

Abstract

Most of the proposed subsampling and resampling methods in the literature assume stationary data. In many empirical applications, however, the hypothesis of stationarity can easily be rejected. In this paper we demonstrate that moment and variance estimators based on the subsampling methodology can be employed also for different types of nonstationarity data. Consistency of estimators are demonstrated under mild moment and mixing conditions. Rates of convergence are provided, giving guidance into the appropriate choice of subshape size. Results from a small simulation study on finite sample properties are also reported.

Keywords: Subsampling, resampling, block bootstrap, nonstationary random field, spatial lattice data, mixing.

AMS 2000 subject classifications. Primary 62G09; secondary 62G05, 60G60.

[1]E-mail address to the correspondence author: magnus.ekstrom@sekon.slu.se

1 Introduction

Resampling and subsampling methods have been suggested in the literature to nonparametrically estimate different characteristics, such as the variance and the distribution, of statistics computed from, e.g., time series and spatial lattice data. The advantage of such methods is that statistical inference can be done without knowledge of the underlying dependence mechanism and marginal distributions that generated the data. Furthermore, for the user no explicit theoretical derivation is necessary but instead intensive computing. Most of the proposed methods in the literature assume stationary data. However, when modeling real-life data (in, e.g., agricultural experiments and applications of remote-sensing imagery), the hypothesis of stationarity often must be rejected. In the current paper we will mainly focus on subsampling variance estimation of statistics computed from *nonstationary* data.

For stationary time series data, the subsampling variance estimator of a statistic g use “replicates” of g computed on subseries of consecutive observations. Thus, within each subseries the dependence structure of the original observations is preserved and if the common length of the subseries increases to infinity with the sample size, asymptotically valid inference can ensue. Pioneering work in this direction has been performed by Carlstein (1986); see Künsch (1989) for related results. When a statistic g is computed on some spatially indexed data observed in some region $A \subset \mathbb{R}^2$, then the subsampling variance estimator of g use “replicates” of g computed on subshapes of A. Such extensions to stationary spatial lattice data have been provided by Possolo (1991), Politis and Romano (1993), Sherman and Carlstein (1994), Sherman (1996), Politis et al. (1999), and Lahiri (1993), among others. In Fukuchi (1999) and Politis et al. (1999), the results of Carlstein (1986) are extended to nonstationary time series data. See also Belyaev (1996).

In the current paper, the general moment and variance estimators for spatial lattice data, as proposed by Sherman (1996), will be regarded in a nonstationary context, and consistency will be shown under assumptions similar to those for time series in Fukuchi (1999, Theorem 1) and Politis et al. (1999, Lemma 4.6.1). Such results show that the subsampling methodology introduced for dependent but stationary observations can still be employed, even if the assumption of stationarity is violated. For example, heteroscedasticity is a problem that often arises. We demonstrate that the subsampling variance estimator allow for considerable heteroscedasticity.

At the same time it is clear that it is not difficult to find conditions of nonstationarity under which the subsampling variance estimator fails. In some of these cases, however, the estimation method can be modified in order to obtain valid inference, as illustrated in the example below.

EXAMPLE 1. Assume that $\{X_{\boldsymbol{i}} : \boldsymbol{i} \in \mathbb{Z}^2\}$ is a random field and that we observe the $X_{\boldsymbol{i}}$'s in some region $A \subset \mathbb{R}^2$, i.e., we observe $\{X_{\boldsymbol{i}} : \boldsymbol{i} \in A \cap \mathbb{Z}^2\}$. If the statistic of interest is the sample mean $\bar{X}_A$ over A, then the subsampling estimator of variance use the sample means computed on (overlapping) subshapes of A as "replicates", i.e., the estimator is a normalized sample variance of the subshape means. Although this estimator can handle considerable heteroscedasticity, it is sensitive to variation in the expected values $\{E[X_{\boldsymbol{i}}]\}$. Unless the subshape means have (essentially) the same expected value, the subsampling estimator of variance will fail, since it cannot distinguish the variation in the expected values from the random variation. This happens, for example, when the expected values $\{E[X_{\boldsymbol{i}}]\}$ decomposes additively into directional components (i.e., $E[X_{\boldsymbol{i}}] = \mu + c_{i_1} + r_{i_2}$ for all $\boldsymbol{i} = (i_1, i_2)$). In geostatistics it is common to detrend data of this type (Cressie, 1993). That is, the directional components are estimated (e.g., by mean or median polishing) and subtracted from the observations. The analysis is then performed on the detrended observations (as if they are a sample from a stationary random field). In Ekström (2002) it is shown that the subsampling/resampling method applied on detrended observations (through mean polishing) gives a consistent estimator of the variance of $\bar{X}_A$. Unfortunately, the estimator based on detrended observations is heavily biased for small samples. In Ekström and Sjöstedt-de Luna (2004a) a more promising method is introduced by modifying the subsampling estimator of variance so that it can handle directional components (as well as smoothly varying expected values) without the need of estimating them. This modified estimator is based on "crosswise differences" of subshape means rather than single subshape means.

In the current paper we establish, under weak moment and mixing conditions, that the modified subsampling estimator of the variance of sample means can handle heteroscedastic data that also possess a spatial trend (e.g., a periodic trend, a smooth trend, directional components, or any combination of these). Our conditions on the dependency structure and the spatial trend are weaker than in the corresponding results by Ekström and Sjöstedt-de Luna (2004a). The question on how to choose the distances between the subshapes that defines the crosswise differences appeared as an open problem in Ekström and Sjöstedt-de Luna (2004a). In this paper we introduce a method for finding appropriate distances between these subshapes.

The remaining of the paper is organized as follows. Section 2 defines our notation and the general subsampling estimator of variance. In Section 3, consistency and rates of convergence for the general moment and variance estimators are established under weak moment and mixing conditions. Section 4 is focused on the special case when the statistic is a sample mean, and in this section we consider the (original) subsampling variance estimator as well as Ekström and Sjöstedt-de Luna's modified estimator. In Section 5 we introduce a method for choosing the distances between the subshapes that defines the crosswise differences in the modified variance estimator. A small simulation study is given in Section 6 and the proofs of the theorems are given in the Appendix.

2 Preliminaries

Assume that the boundary of a set $A_{\mathbf{1}} \subseteq (0,1] \times (0,1]$ is a simple closed curve of finite length, and for two integer values n_1, n_2, that $A = A_{\boldsymbol{n}} \subseteq (0, n_1] \times (0, n_2]$, $\boldsymbol{n} = (n_1, n_2)$, is obtained from $A_{\mathbf{1}}$ in the following sense: $A = A_{\boldsymbol{n}} = \{\boldsymbol{x} \in \mathbb{R}^2 : (x_1/n_1, x_2/n_2) \in A_{\mathbf{1}}\}$. We now consider the region A, with our data $\{X_{\boldsymbol{i}} \in \mathbb{R}\}$ being observed at the indices in $A \cap \mathbb{Z}^2$. A subshape of A can be constructed as $A_{\boldsymbol{k}} = \{\boldsymbol{x} \in \mathbb{R}^2 : (x_1/k_1, x_2/k_2) \in A_{\mathbf{1}}\}$, formed in the subrectangle $B_{\mathbf{0}} = (0, k_1] \times (0, k_2]$, where $k_j < n_j$, $j = 1, 2$. We can construct other (overlapping) subshapes in a similar way by identifying them in the subrectangles $B_{\boldsymbol{j}} = (j_1, j_1 + k_1] \times (j_2, j_2 + k_2]$, $j_1, j_2 = 0, 1, \ldots$. However, we only use the $t_{\boldsymbol{n}}$ subshapes $A_{\boldsymbol{k},t}$, $t = 1, \ldots, t_{\boldsymbol{n}}$, which are completely contained in A. Note that the $A_{\boldsymbol{k},t}$'s do not need to be subshapes in the strict sense. Actually, $A_{\boldsymbol{k},t}$ is a "true subshape" only when $k_1/k_2 = n_1/n_2$. The number of indices in $A_{\boldsymbol{n}} \cap \mathbb{Z}^2$ and $A_{\boldsymbol{k}} \cap \mathbb{Z}^2$ will be denoted by S and s, respectively. Let $K = k_1 k_2$ and $N = n_1 n_2$. If $A_{\mathbf{1}} = (0,1] \times (0,1]$, then $A = A_{\boldsymbol{n}} = (0, n_1] \times (0, n_2]$, $s = K$, $S = N$, and $t_{\boldsymbol{n}} = (n_1 - k_1 + 1)(n_2 - k_2 + 1)$.

Suppose, for the moment, that $\{X_{\boldsymbol{i}} \in \mathbb{R}\}$ is a stationary random field. Suppose further that a statistic $g(A)$ is computed from the $X_{\boldsymbol{i}}$'s in region A and that we want an estimate of the variance $\gamma_{\boldsymbol{n}} = \mathrm{var}[\sqrt{S} g(A)]$. The motivation of the subsampling method is that if the subshapes $\{A_{\boldsymbol{k},t}\}$ are large enough, enough of the original dependence will be preserved in the subshapes that statistics $g(A_{\boldsymbol{k},t})$ will have approximately the same distribution (if properly normalized) as values $g(A)$ calculated from replicates of the original data.

Further, the $\{g(A_{\boldsymbol{k},t})\}$ are identically distributed random variables (r.v.'s), and if the dependency is weak enough we can expect that

$$\tilde{\gamma}_{\boldsymbol{n}} = \frac{s}{t_{\boldsymbol{n}}} \sum_{t=1}^{t_{\boldsymbol{n}}} (g(A_{\boldsymbol{k},t}) - \bar{g}')^2, \tag{1}$$

where $\bar{g}' = \sum_{t=1}^{t_{\boldsymbol{n}}} g(A_{\boldsymbol{k},t})/t_{\boldsymbol{n}}$, is a valid estimator of $\mathrm{var}[\sqrt{s}g(A_{\boldsymbol{k},t})]$. Now, if $\gamma_{\boldsymbol{n}}$ tends to some limit γ, as the region A expands more or less uniformly, and if the subshapes are large enough, then we have $\mathrm{var}[\sqrt{S}g(A)] \approx \gamma$ and $\mathrm{var}[\sqrt{s}g(A_{\boldsymbol{k},t})] \approx \gamma$. Therefore we can regard $\tilde{\gamma}_{\boldsymbol{n}}$ as an estimator of $\gamma_{\boldsymbol{n}} = \mathrm{var}[\sqrt{S}g(A)]$ as well. In the next section it will be shown that $\tilde{\gamma}_{\boldsymbol{n}}$ is a consistent estimator of $\gamma_{\boldsymbol{n}}$, even if the assumptions of stationarity and the existence of a limiting variance γ are violated.

The weak dependency in the random field $\{X_{\boldsymbol{i}} \in \mathbb{R}\}$ will be quantified through a model-free mixing coefficient. A mixing condition says essentially that observations separated by large distances are approximately independent. In, e.g., Doukhan (1994), Guyon (1995), and Lin and Lu (1996) a collection of strong mixing coefficients is defined by

$$\begin{aligned} \alpha_l(m) &= \sup\{|P(B_1 \cap B_2) - P(B_1)P(B_2)| : \\ &\quad B_j \in \mathcal{F}(U_j), |U_j| \leq l, j = 1, 2, \rho(U_1, U_2) \geq m\}, \end{aligned}$$

where $\mathcal{F}(U)$ is the σ-algebra generated by $\{X_{\boldsymbol{i}}, \boldsymbol{i} \in U\}$ and $\rho(U_1, U_2) = \inf\{\rho(\boldsymbol{i}, \boldsymbol{j}) : \boldsymbol{i} \in U_1, \boldsymbol{j} \in U_2\}$, $\rho(\boldsymbol{i}, \boldsymbol{j}) = \max\{|i_1 - j_1|, |i_2 - j_2|\}$. If we put $\alpha(m) = \alpha_\infty(m)$ we obtain Rosenblatt's (1956) usual strong mixing coefficients and it is apparent that $\alpha_l(m) \leq \alpha(m)$. If $\alpha(m) \to 0$ as $m \to \infty$ then the random field is said to be *strong mixing.*

Henceforth we use the following notation: $a \vee b$ denotes the maximum of a and b; $a_n \sim b_n$ means that for two sequences $\{a_n\}_{n\geq 1}$ and $\{b_n\}_{n\geq 1}$ there exist two constants, $0 < c_1 < c_2 < \infty$, such that $c_1 < a_n/b_n < c_2$, $n = 1, 2, ...$; $\mathbf{a} = (a_1, a_2) \to \infty$ means that *both* a_1 and a_2 tend to ∞; ${}^\eta X = XI_{\{|X|>\eta\}}$; $||X||_p = (E[|X|^p])^{1/p}$. A sequence $\xi_1, \xi_2, ...$ of r.v.'s will be said to converge to the random variable ξ in L_p norm, $0 < p < \infty$, if $||\xi_n - \xi||_p \to 0$ as $n \to \infty$. Suppose $f(n)$ and $g(n)$ are two functions. We write $f(n) = O(g(n))$ if and only if there exists a constant c such that $|f(n)| \leq c|g(n)|$ for all sufficiently large values of n. It will be observed that our formulae involving $O(\cdot)$ will not usually be reversible. Thus '$O(n) = O(n^2)$' (that is, 'if $f(n) = O(n)$ then $f(n) = O(n^2)$') is true, but '$O(n^2) = O(n)$' is false. The letter c will denote a constant which may have different values from equation to equation. U.I. stands for uniformly integrable (or uniform integrability).

3 Consistency and convergence rates of general subsampling estimators

In this section we extend results of Sherman (1996) to *nonstationary* spatial lattice data, i.e., it is shown, under weak moment and mixing conditions, that the general moment estimator $\bar{g}'$ and the variance estimator $\tilde{\gamma}_{\boldsymbol{n}}$ are consistent estimators of $E[g(A)]$ and $\gamma_{\boldsymbol{n}}$, respectively. Convergence rates in mean square are also given. These extensions are useful, since in many real-life situations, the hypothesis of stationarity can easily be rejected. Examples of such situations include forestry applications of satellite data (Ekström and Sjöstedt-de Luna, 2004a) and landscape ecology, where certain indices are used for describing the spatial structures of landscapes (Ekström and Sjöstedt-de Luna, 2004b).

Henceforth we assume that $n_1 \sim n_2$, $k_1 \sim k_2$, $\boldsymbol{k} \to \infty$, and $k_j/n_j \to 0$, $j = 1, 2$, as $\boldsymbol{n} \to \infty$.

Theorem 1 *Assume that*

(a) $\{(g(A_{\boldsymbol{k},t}))^2 : t = 1, ..., t_{\boldsymbol{n}}, n_1, n_2 = 1, 2, ...\}$ *is U.I.,*

(b) $E[\bar{g}'] - E[g(A)] \to 0$ *as* $\boldsymbol{n} \to \infty$,

(c) $\alpha_s(k_1 \vee k_2) \to 0$ *as* $\boldsymbol{n} \to \infty$.

Then $\bar{g}' - E[g(A)]$ *converge to 0 in* L_2 *norm as* $\boldsymbol{n} \to \infty$.

Theorem 2 *Assume that*

(a) $\{(\sqrt{s}\,(g(A_{\boldsymbol{k},t}) - E[g(A_{\boldsymbol{k},t})]))^4 : t = 1, ..., t_{\boldsymbol{n}}, n_1, n_2 = 1, 2, ...\}$ *is U.I.,*

(b) $t_{\boldsymbol{n}}^{-1} \sum_{t=1}^{t_{\boldsymbol{n}}} \mathrm{var}[\sqrt{s}\, g(A_{\boldsymbol{k},t})] - \gamma_{\boldsymbol{n}} \to 0$ *as* $\boldsymbol{n} \to \infty$,

(c) $t_{\boldsymbol{n}}^{-1} \sum_{t=1}^{t_{\boldsymbol{n}}} (E[\sqrt{s}\, g(A_{\boldsymbol{k},t})] - E[\sqrt{s}\, \bar{g}'])^2 \to 0$ *as* $\boldsymbol{n} \to \infty$,

(d) $\alpha_s(k_1 \vee k_2) \to 0$ *as* $\boldsymbol{n} \to \infty$.

Then $\tilde{\gamma}_{\boldsymbol{n}} - \gamma_{\boldsymbol{n}}$ *converge to 0 in* L_2 *norm as* $\boldsymbol{n} \to \infty$.

The assumptions in Theorems 1 and 2 resemble the assumptions in Fukuchi (1999, Theorem 1a) and Politis et al. (1999, Lemma 4.6.1), respectively, although an assumption like (c) in Theorem 2 appears to be missing in Politis et al. (1999). Further, we do not require the existence of the limits of $E[g(A)]$

and $\gamma_{\boldsymbol{n}}$ as $\boldsymbol{n} \to \infty$, as in the corresponding results for time series by Fukuchi and Politis et al.

If $\{X_{\boldsymbol{i}}\}$ is stationary and the limits of $E[g(A)]$ and $\gamma_{\boldsymbol{n}}$ exist as $\boldsymbol{n} \to \infty$, then assumption (b) in Theorem 1 and assumptions (b) and (c) in Theorem 2 are trivially fulfilled. Note also that (b) in Theorem 2 is a weak condition on the variances. For example, if $g(A)$ is the sample mean and if, for some $\delta, \beta_\delta > 0$, $||X_{\boldsymbol{i}} - E[X_{\boldsymbol{i}}]||_{2+\delta} \le \beta_\delta$ for all $\boldsymbol{i}, \boldsymbol{n}$ and $\sum_{m=0}^{\infty}(m+1)\alpha_1(m)^{\delta/(2+\delta)} < \infty$, then condition (b) is fulfilled (see Lemma A.4 in the Appendix). Thus, the subsampling variance estimator allow for considerable heteroscedasticity.

The main practical issue in applying subsampling and block resampling methods is the choice of the subshape/block size. This issue is shared by all "blocking" methods, such as Künsh's (1989) moving blocks bootstrap, Carlstein's (1986) variance estimator, or the estimator of variance in Ekström and Sjöstedt-de Luna (2004a). The asymptotic conditions for consistency are typically fulfilled for a broad range of choices of the rate of subshape/block size. Although any choice of rate satisfying these conditions will provide the desired consistency, the conditions do not give much guidance on how to choose the subshape/block size in the case of a finite sample. In subsampling and block resampling methods, the optimal asymptotic rate for the subshapes/blocks depends critically on the context. When S denotes the total number of observations and s the desired subshape/block size, the optimal asymptotic formula for block and subshape size is typically $s \sim S^\beta$, where the value of β is known, determined by context (see, e.g., Hall et al. (1995)). Based on this formula, Hall et al. (1995) suggested an empirical rule for estimating the optimal block size in a stationary time series context. A similar rule is applied in a spatial setting in Ekström and Sjöstedt-de Luna (2004a) and it is shown that it works also in cases of nonstationarity. For obtaining an optimal asymptotic formula for the subshape size for the variance estimator $\tilde{\gamma}_{\boldsymbol{n}}$, we will investigate the rate of convergence for $\tilde{\gamma}_{\boldsymbol{n}}$. This will be done in Corollary 2 below. First we establish, in Corollary 1, a rate of convergence for the general moment estimator $\bar{g}'$.

Corollary 1 *Assume that*

(a) $\{(g(A_{\boldsymbol{k},t}))^4 : t = 1, ..., t_{\boldsymbol{n}}, n_1, n_2 = 1, 2, ...\}$ *is U.I.,*

(b) $|E[\bar{g}'] - E[g(A)]| = O(\kappa_{\boldsymbol{n}})$ *for some nonnegative function* $\kappa_{\boldsymbol{n}}$*,*

(c) $s^2\alpha_s(k_1 \vee k_2) = O(\lambda_{\boldsymbol{n}})$ *for some nonnegative function* $\lambda_{\boldsymbol{n}}$*.*

Then $E[(\bar{g}' - E[g(A)])^2] = O(K/N + 1/K + \kappa_{\boldsymbol{n}}^2 + \lambda_{\boldsymbol{n}})$.

Corollary 2 *Assume that*

(a) $\{(\sqrt{s}\,(g(A_{\boldsymbol{k},t}) - E[g(A_{\boldsymbol{k},t})]))^8 : t = 1, ..., t_{\boldsymbol{n}}, n_1, n_2 = 1, 2, ...\}$ *is U.I.,*

(b) $|t_{\boldsymbol{n}}^{-1} \sum_{t=1}^{t_{\boldsymbol{n}}} \mathrm{var}[\sqrt{s}\, g(A_{\boldsymbol{k},t})] - \gamma_{\boldsymbol{n}}| = O(\varrho_{\boldsymbol{n}})$ *for some nonnegative function* $\varrho_{\boldsymbol{n}}$*,*

(c) $t_{\boldsymbol{n}}^{-1} \sum_{t=1}^{t_{\boldsymbol{n}}} (E[\sqrt{s}\, g(A_{\boldsymbol{k},t})] - E[\sqrt{s}\, \bar{g}'])^2 = O(\zeta_{\boldsymbol{n}})$ *for some nonnegative function* $\zeta_{\boldsymbol{n}}$*,*

(d) $s^2 \alpha_s(k_1 \vee k_2) = O(\lambda_{\boldsymbol{n}})$ *for some nonnegative function* $\lambda_{\boldsymbol{n}}$*.*

Then $E[(\tilde{\gamma}_{\boldsymbol{n}} - \gamma_{\boldsymbol{n}})^2] = O(K/N + 1/K + \varrho_{\boldsymbol{n}}^2 + \zeta_{\boldsymbol{n}} + \lambda_{\boldsymbol{n}})$*.*

If the random field $\{X_{\boldsymbol{i}}\}$ is stationary, then assumption (c) in Corollary 2 is trivially fulfilled (i.e., we may put $\zeta_{\boldsymbol{n}} \equiv 0$). If $\varrho_{\boldsymbol{n}}^2 + \zeta_{\boldsymbol{n}} + \lambda_{\boldsymbol{n}} = O(K/N + 1/K)$, then the rate of convergence for $\tilde{\gamma}_{\boldsymbol{n}}$ is minimized when k_i is of the order of $n_i^{1/2}$, $i = 1, 2$, i.e., when $s \sim S^{1/2}$. Note that this makes it possible to apply an empirical rule for estimating the optimal subshape size, by extending the ideas in Hall et al. (1995) to spatial lattice data. In Ekström and Sjöstedt-de Luna (2004a), a rule of this type was considered for a (modified) subsampling estimator of variance for nonstationary spatial lattice data.

4 The case of the sample mean

Let $\bar{X}_A$ be the sample mean of the $X_{\boldsymbol{i}}$'s in region A, $\bar{X}_t$ the sample mean of the $X_{\boldsymbol{i}}$-values in $A_{\boldsymbol{k},t}$, and $\bar{X}' = \sum_{t=1}^{t_{\boldsymbol{n}}} \bar{X}_t / t_{\boldsymbol{n}}$. Then we can write the estimator (1) of $\gamma_{\boldsymbol{n}} = \mathrm{var}[\sqrt{S}\bar{X}_A]$ as

$$\tilde{\gamma}_{\boldsymbol{n}} = \frac{s}{t_{\boldsymbol{n}}} \sum_{t=1}^{t_{\boldsymbol{n}}} (\bar{X}_t - \bar{X}')^2.$$

In Theorem 3 and Corollary 3 below, the results of Theorem 2 and Corollary 2 are reformulated and the U.I. conditions are replaced by moment and mixing conditions.

The following assumption is used in the next two theorems.

Assumption A1. The random field $\{X_{\boldsymbol{i}}\}$ satisfies the following conditions, for some $\delta > 0$,

(i) $||X_{\boldsymbol{i}} - E[X_{\boldsymbol{i}}]||_{4+\delta} \leq \beta_\delta$, for some $\beta_\delta > 0$ and all $\boldsymbol{i}, \boldsymbol{n}$,

(ii) $st_{\boldsymbol{n}}^{-1} \sum_{t=1}^{t_n} (E[\bar{X}_t] - E[\bar{X}'])^2 \to 0$ as $\boldsymbol{n} \to \infty$,

(iii) $\alpha_l(m) \leq cm^{-\tau_1} l^{\tau_2}$, for some $\tau_1 > 10(6+\delta)/\delta$, $0 \leq \tau_2 < \tau_1/2$, and all $l, m \geq 1$.

Theorem 3 *If A1 holds, then* $\tilde{\gamma}_{\boldsymbol{n}} - \gamma_{\boldsymbol{n}} \to 0$ *in* L_2 *norm as* $\boldsymbol{n} \to \infty$.

Note, in A1 there is no assumption on the variances of the $X_{\boldsymbol{i}}$'s, other than a uniform bound from above. The assumption on the mixing coefficients in A1 resembles the assumptions of Lahiri (2003, Theorem 12.1).

Recall that k_1 and k_2 determine the subshape size and note that the L_2 consistency holds for a broad range of choices of k_1 and k_2 in Theorem 3. In a particular application the values of k_1 and k_2 must be specified. From the following corollary it follows that k_i of the order of $n_i^{1/2}$, $i = 1, 2$, is a good choice when, for example, $E[X_{\boldsymbol{i}}] \equiv 0$. The corollary gives a rate of convergence for $\tilde{\gamma}_{\boldsymbol{n}}$ and is established under the following assumption.

Assumption A2. The random field $\{X_{\boldsymbol{i}}\}$ satisfies the following conditions, for some $\delta > 0$,

(i) $||X_{\boldsymbol{i}} - E[X_{\boldsymbol{i}}]||_{4+\delta} \leq \beta_\delta$, for some $\beta_\delta > 0$ and all $\boldsymbol{i}, \boldsymbol{n}$,

(ii) $st_{\boldsymbol{n}}^{-1} \sum_{t=1}^{t_n} (E[\bar{X}_t] - E[\bar{X}'])^2 = O(\psi_{\boldsymbol{n}})$ for some nonnegative function $\psi_{\boldsymbol{n}}$,

(iii) $\alpha_l(m) \leq cm^{-\tau_1} l^{\tau_2}$, for some $\tau_1 > 10(6+\delta)/\delta$, $0 \leq \tau_2 \leq \tau_1/2 - (2+\delta)/\delta$, and all $l, m \geq 1$.

Corollary 3 *If A2 holds, then* $E[(\tilde{\gamma}_{\boldsymbol{n}} - \gamma_{\boldsymbol{n}})^2] = O(K/N + 1/K + \psi_{\boldsymbol{n}}^2)$.

The subsampling estimator $\tilde{\gamma}_{\boldsymbol{n}}$ can handle cases when the data have "asymptotically equal" expected values, as is demonstrated in the following example.

Example 2. If $E[X_{\boldsymbol{i}}] = \mu + \delta_{\boldsymbol{i}}$, where $|\delta_{\boldsymbol{i}}| =\leq c\nu_{\boldsymbol{n}}$, for all $\boldsymbol{i}, \boldsymbol{n}$, then we may take $\psi_{\boldsymbol{n}} = K\nu_{\boldsymbol{n}}^2$ in assumption A2(ii).

If, for example, the expected values $\{E[X_{\boldsymbol{i}}]\}$ vary smoothly over region A, then the estimator $\tilde{\gamma}_{\boldsymbol{n}}$ will not be consistent due to variation coming from $\{E[X_{\boldsymbol{i}}]\}$. To reduce this variation, Ekström and Sjöstedt-de Luna (2004a) propose a modified subsampling method based on "crosswise subshape differences". For the creation of a crosswise subshape difference, identify the subshapes in the four subrectangles $\mathcal{B}_{\boldsymbol{j}}, \mathcal{B}_{j_1+d_1+k_1,j_2}, \mathcal{B}_{\boldsymbol{j}+\boldsymbol{d}+\boldsymbol{k}}, \mathcal{B}_{j_1,j_2+d_2+k_2}$, and name the subshapes $A_{\boldsymbol{k},t_1}$, $A_{\boldsymbol{k},t_2}$, $A_{\boldsymbol{k},t_3}$, and $A_{\boldsymbol{k},t_4}$, respectively. If all four subshapes are completely contained in A, then we define the crosswise subshape difference Z_t from the corresponding mean values, i.e., $Z_t = (\bar{X}_{t_1} - \bar{X}_{t_2} + \bar{X}_{t_3} - \bar{X}_{t_4})/2$. Assume there are $t'_{\boldsymbol{n}}$ such crosswise differences, Z_t, $t = 1, ..., t'_{\boldsymbol{n}}$, and let $\bar{Z} = \sum_{t=1}^{t'_{\boldsymbol{n}}} Z_t/t'_{\boldsymbol{n}}$. Then the modified subsampling estimator of $\gamma_{\boldsymbol{n}}$ is defined as

$$\hat{\gamma}_{\boldsymbol{n}} = \frac{s}{t'_{\boldsymbol{n}}} \sum_{t=1}^{t'_{\boldsymbol{n}}} (Z_t - \bar{Z})^2.$$

In the next theorem, concerning the modified subsampling estimator of variance, it is assumed that the random field $\{X_{\boldsymbol{i}}\}$ can be decomposed as $X_{\boldsymbol{i}} = \mu_{\boldsymbol{i}} + Y_{\boldsymbol{i}}$, where $\{Y_{\boldsymbol{i}}\}$ satisfies A1 and $\{\mu_{\boldsymbol{i}}\}$ is some additive deterministic "trend". With $\{Z_t\}$ denoting the crosswise subshape differences created from the $X_{\boldsymbol{i}}$-values, we can write $Z_t = Z_{\mu,t} + Z_{Y,t}$, where $\{Z_{\mu,t}\}$ and $\{Z_{Y,t}\}$ are the crosswise subshape differences created from $\{\mu_{\boldsymbol{i}}\}$ and $\{Y_{\boldsymbol{i}}\}$, respectively. Let $\bar{Z}_\mu = \sum_{t=1}^{t'_{\boldsymbol{n}}} Z_{\mu,t}/t'_{\boldsymbol{n}}$.

Theorem 4 *Assume that $X_{\boldsymbol{i}} = \mu_{\boldsymbol{i}} + Y_{\boldsymbol{i}}$, where $\{Y_{\boldsymbol{i}}\}$ satisfies A1 and*

$$\frac{s}{t'_{\boldsymbol{n}}} \sum_{t=1}^{t'_{\boldsymbol{n}}} (Z_{\mu,t} - \bar{Z}_\mu)^2 \to 0 \text{ as } \boldsymbol{n} \to \infty. \tag{2}$$

Further, assume that $d_i = O(k_i)$, $i = 1, 2$. Then $\hat{\gamma}_{\boldsymbol{n}} - \gamma_{\boldsymbol{n}} \to 0$ in L_2 norm as $\boldsymbol{n} \to \infty$.

Corollary 4 *Assume that $X_{\boldsymbol{i}} = \mu_{\boldsymbol{i}} + Y_{\boldsymbol{i}}$, where $\{Y_{\boldsymbol{i}}\}$ satisfies A2 and*

$$\frac{s}{t'_{\boldsymbol{n}}} \sum_{t=1}^{t'_{\boldsymbol{n}}} (Z_{\mu,t} - \bar{Z}_\mu)^2 = O(\phi_{\boldsymbol{n}}), \tag{3}$$

for some nonnegative function $\phi_{\boldsymbol{n}}$. Further, assume that $d_i = O(k_i)$, $i = 1, 2$. Then $E[(\hat{\gamma}_{\boldsymbol{n}} - \gamma_{\boldsymbol{n}})^2] = O(K/N + 1/K + \psi_{\boldsymbol{n}}^2 + \phi_{\boldsymbol{n}}^2)$.

It should be noted that the above assumptions on the expected values $\{E[X_{\boldsymbol{i}}]\}$ are more general than in Ekström and Sjöstedt-de Luna (2004b). Further, Ekström and Sjöstedt-de Luna (2004b) assume that $\{X_{\boldsymbol{i}}\}$ is an m-dependent random field, i.e., that $\alpha_l(r) = 0$ for all $r > m$ and $l \geq 1$, which is special case of assumptions A1(iii) and A2(iii). If $E[X_{\boldsymbol{i}}] \equiv \mu$, then the assumptions A2(ii), (2), and (3), with $\psi_{\boldsymbol{n}} \equiv 0$ and $\phi_{\boldsymbol{n}} \equiv 0$, are trivially fulfilled and the obtained convergence rate in mean square in Corollaries 3 and 4 becomes $O(K/N + 1/K)$. This is the same order of convergence as for stationary random fields (e.g. Sherman 1996) where $\tilde{\gamma}_{\boldsymbol{n}}$ is used.

We now present a number of further examples of "trends" $\{\mu_{\boldsymbol{i}}\}$ that satisfy assumptions (2) and (3).

Example 3. If $\mu_{\boldsymbol{i}}$ can be decomposed additively into directional components such that $\mu_{\boldsymbol{i}} = \mu + c_{i_1} + r_{i_2}$ for all $\boldsymbol{i}$, then (2) and (3), with $\phi_{\boldsymbol{n}} \equiv 0$, are satisfied.

Example 4. If there is some periodicity in $\{\mu_{\boldsymbol{i}}\}$ such that

$$\mu_{\boldsymbol{i}} = \mu_{i_1+p_1,i_2} \text{ (or } \mu_{\boldsymbol{i}} = \mu_{i_1,i_2+p_2}) \text{ for all } \boldsymbol{i}, \tag{4}$$

then (2) and (3), with $\phi_{\boldsymbol{n}} \equiv 0$, are satisfied whenever $d_1 + k_1 = mp_1$ (or $d_2 + k_2 = mp_2$) for some integer m.

Example 5. If

$$\mu_{\boldsymbol{i}} - \mu_{i_1+p_1,i_2} + \mu_{i_1+p_1,i_2+p_2} - \mu_{i_1,i_2+p_2} = 0 \text{ for all } \boldsymbol{i}, \tag{5}$$

then (2) and (3), with $\phi_{\boldsymbol{n}} \equiv 0$, are satisfied whenever $d_1 + k_1 = mp_1$ and $d_2 + k_2 = mp_2$ for some integer m. Note that if both equalities in (4) are satisfied, then so is (5), but (5) does not necessarily imply any of the equalities in (4).

Example 6. If $d_i = O(k_i)$, $i = 1, 2$, and the $\mu_{\boldsymbol{i}}$-values are smoothly varying in the sense that they satisfy a Lipschitz condition of order β, i.e., if $\mu_{\boldsymbol{i}} = f(i_1/n_1, i_2/n_2)$ for all $\boldsymbol{i}, \boldsymbol{n}$, where, for some $0 < \beta \leq 1$, $|f(\boldsymbol{x}) - f(\boldsymbol{y})| \leq c||\boldsymbol{x} - \boldsymbol{y}||^\beta$ over the set $A_{\boldsymbol{1}}$, then (3) is valid with $\phi_{\boldsymbol{n}} = K(K/N)^\beta$. Assumption (2) holds if $K(K/N)^\beta \to 0$ as $\boldsymbol{n} \to \infty$. Here $||\cdot||$ is the Euclidian norm.

Example 7. If $d_i = O(k_i)$, $i = 1, 2$, and the $\mu_{\boldsymbol{i}}$-values are smoothly varying in the sense that $\mu_{\boldsymbol{i}} = f(i_1/n_1, i_2/n_2)$ for all $\boldsymbol{i}, \boldsymbol{n}$, where, for some $0 < \beta \leq 1$,

$$|f(x_1, x_2) - f(x_1 + h_1, x_2) + f(x_1 + h_1, x_2 + h_2) - f(x_1, x_2 + h_2)| \leq c|h_1 h_2|^\beta \tag{6}$$

over the set $A_{\boldsymbol{1}}$, then (3) is valid with $\phi_{\boldsymbol{n}} = K(K/N)^{2\beta}$. Assumption (2) holds if $K(K/N)^{2\beta} \to 0$ as $\boldsymbol{n} \to \infty$. Note that if f is such that its first-order partial derivatives exist and are continuous on $A_{\mathbf{1}}$, and if the mixed partial derivative f_{12} (or f_{21}) is continuous and bounded on $A_{\mathbf{1}}$, then (6) is satisfied with $\beta = 1$.

EXAMPLE 8. The different conditions on $\{\mu_{\boldsymbol{i}}\}$ in Examples 3-7 can be combined. For instance, if $d_i = O(k_i)$, $i = 1, 2$, and $\mu_{\boldsymbol{i}} = \mu_{3,\boldsymbol{i}} + \mu_{7,\boldsymbol{i}}$, where $\mu_{3,\boldsymbol{i}}$ and $\mu_{7,\boldsymbol{i}}$ satisfy the conditions in Examples 3 and 7, respectively, then (3) is valid with $\phi_{\boldsymbol{n}} = K(K/N)^{2\beta}$ and assumption (2) holds if $K(K/N)^{2\beta} \to 0$ as $\boldsymbol{n} \to \infty$.

Assume that assumption A2(ii) holds with $\psi_{\boldsymbol{n}} = N^{-1/4}$ when $k_i = O(n_i^{1/2})$, $i = 1, 2$, and that β in Examples 7 and 8 is larger or equal to $3/4$. The optimal convergence rate in mean square in Examples 7 and 8 is then $O(N^{-1/2})$, which we get if we choose $k_i = O(n_i^{1/2})$, $i = 1, 2$. This is the same optimal order of convergence as for stationary random fields (e.g. Sherman 1996) where $\tilde{\gamma}_{\boldsymbol{n}}$ is used. In Example 6 with $\beta = 1$, on the other hand, the corresponding optimal rate is only $O(N^{-2/5})$, achieved if we choose $k_i = O(n_i^{2/5})$, $i = 1, 2$.

5 Choice of $\boldsymbol{d}$

Crosswise differences are created to reduce the variation coming from the expected values. If the expected values vary smoothly, then the reduction tends to be more successful when d_1 and d_2 are small. On the other hand, if d_1 and d_2 are small, then the subshapes defining a crosswise difference may be strongly dependent, which is an argument for not choosing d_1 and d_2 too small. Thus, it is not obvious how to choose $\boldsymbol{d} = (d_1, d_2)$ in practice. However, if we can find small values of d_1 and d_2 such that $\bar{X}_{t_1}$, $\bar{X}_{t_2}$, $\bar{X}_{t_3}$, and $\bar{X}_{t_4}$ are independent or at most weakly dependent, then this choice of d_1 and d_2 is typically a good choice. But how to tell whether $\bar{X}_{t_1}$, $\bar{X}_{t_2}$, $\bar{X}_{t_3}$, and $\bar{X}_{t_4}$ are at most weakly independent or not for a particular choice of $\boldsymbol{d}$? In the case of stationary data, a study of an estimated correlogram (Cressie, 1993) may yield the answer, but when having nonstationary data with, e.g., smoothly varying expected values, it is often difficult to find reasonable estimators of the correlation structure of $\{X_{\boldsymbol{i}}\}$. Instead we will use a quantity closely related to $\hat{\gamma}_{\boldsymbol{n}}$ with $k_1 \equiv 1$ and $k_2 \equiv 1$ for finding appropriate values of d_1 and d_2.

Let $X_{\boldsymbol{i},\boldsymbol{d},1} = X_{\boldsymbol{i}}$, $X_{\boldsymbol{i},\boldsymbol{d},2} = X_{(i_1+d_1+1,i_2)}$, $X_{\boldsymbol{i},\boldsymbol{d},3} = X_{\boldsymbol{i}+\boldsymbol{d}+\boldsymbol{1}}$, $X_{\boldsymbol{i},\boldsymbol{d},4} = X_{(i_1,i_2+d_2+1)}$, and

$$T_{\boldsymbol{n}}^{\bullet} = \{\boldsymbol{i} : \boldsymbol{i}, (i_1 + d_1 + 1, i_2), \boldsymbol{i} + \boldsymbol{d} + \boldsymbol{1}, (i_1, i_2 + d_2 + 1) \in A_{\boldsymbol{n}} \cap \mathbb{Z}^2\}.$$

Define crosswise differences $Z_{\boldsymbol{i}}^{\bullet} = X_{\boldsymbol{i},\boldsymbol{d},1} - X_{\boldsymbol{i},\boldsymbol{d},2} + X_{\boldsymbol{i},\boldsymbol{d},3} - X_{\boldsymbol{i},\boldsymbol{d},4}$, $\boldsymbol{i} \in T_{\boldsymbol{n}}^{\bullet}$, and let $\bar{Z}^{\bullet}$ be the sample mean of $\{Z_{\boldsymbol{i}}^{\bullet} : \boldsymbol{i} \in T_{\boldsymbol{n}}^{\bullet}\}$. Define

$$\gamma_{\boldsymbol{n}}^{\bullet} = \gamma_{\boldsymbol{n}}^{\bullet}(\boldsymbol{d}) = \frac{1}{t_{\boldsymbol{n}}^{\bullet}} \sum_{\boldsymbol{i} \in T_{\boldsymbol{n}}^{\bullet}} (Z_{\boldsymbol{i}}^{\bullet} - \bar{Z}^{\bullet})^2,$$

where $t_{\boldsymbol{n}}^{\bullet} = \#T_{\boldsymbol{n}}^{\bullet}$, i.e., $t_{\boldsymbol{n}}^{\bullet}$ is the number of elements in $T_{\boldsymbol{n}}^{\bullet}$. If the region A is rectangular, $A = A_{\boldsymbol{n}} = (0, n_1] \times (0, n_2]$, and $k_1 \equiv 1$ and $k_2 \equiv 1$, then note that $\gamma_{\boldsymbol{n}}^{\bullet} = \hat{\gamma}_{\boldsymbol{n}}$.

If, for example, $X_{\boldsymbol{i}} = \mu_{\boldsymbol{i}} + Y_{\boldsymbol{i}}$, with $E[Y_{\boldsymbol{i}}] = 0$ and $\mu_{\boldsymbol{i}} = \mu_{3,\boldsymbol{i}} + \mu_{7,\boldsymbol{i}}$ for all $\boldsymbol{i}$, and where $\mu_{3,\boldsymbol{i}}$ and $\mu_{7,\boldsymbol{i}}$ satisfy the conditions in Examples 3 and 7, respectively, then it can be shown that

$$E[\gamma_{\boldsymbol{n}}^{\bullet}(\boldsymbol{d})] = \frac{1}{4t_{\boldsymbol{n}}^{\bullet}} \sum_{\boldsymbol{i} \in T_{\boldsymbol{n}}^{\bullet}} m_{\boldsymbol{i}}(\boldsymbol{d}) + O\Big(\frac{D^{\beta}}{N^{\beta}}\Big),$$

where $D = d_1 d_2$ and

$$m_{\boldsymbol{i}}(\boldsymbol{d}) = \sum_{u=1}^{4} \mathrm{var}[X_{\boldsymbol{i},\boldsymbol{d},u}] + 2 \sum_{u=1}^{3} \sum_{v=u+1}^{4} \mathrm{cov}[X_{\boldsymbol{i},\boldsymbol{d},u}, X_{\boldsymbol{i},\boldsymbol{d},v}](-1)^{u+v}. \tag{7}$$

Thus, the contribution from the varying expected values $\{\mu_{\boldsymbol{i}}\}$ to $E[\gamma_{\boldsymbol{n}}^{\bullet}(\boldsymbol{d})]$ is of order $(D/N)^{\beta}$.

In many practical situations, it is reasonable to assume that the covariances in (7) are nonnegative and monotonically tending to zero as d_1 and d_2 increases. In this case, the double sum on the right-hand side of (7) is nonpositive and $m_{\boldsymbol{i}}(\boldsymbol{d})$ is a nondecreasing function of d_1 and d_2. This implies that $\gamma_{\boldsymbol{n}}^{\bullet}(\boldsymbol{d})$ typically has a positive increase until a specific point $\boldsymbol{d}^{\bullet} = (d_1^{\bullet}, d_2^{\bullet})$, after which the function flattens out, i.e., if $d_1 \geq d_1^{\bullet}$ and $d_2 \geq d_2^{\bullet}$, then the r.v.'s $X_{\boldsymbol{i},\boldsymbol{d},1}$, $X_{\boldsymbol{i},\boldsymbol{d},2}$, $X_{\boldsymbol{i},\boldsymbol{d},3}$, and $X_{\boldsymbol{i},\boldsymbol{d},4}$ can be considered as independent or, at most, weakly dependent. This implies that if $d_1 = d_1^{\bullet}$ and $d_2 = d_2^{\bullet}$, then $\bar{X}_{t_1}$, $\bar{X}_{t_2}$, $\bar{X}_{t_3}$, and $\bar{X}_{t_4}$ are independent or at most weakly dependent. The value of $\boldsymbol{d}^{\bullet}$ can be found from a visual inspection of the function $\gamma_{\boldsymbol{n}}^{\bullet}(\boldsymbol{d})$.

6 Simulation study

We investigated the performance of the estimators $\tilde{\gamma}_{\boldsymbol{n}}$ and $\hat{\gamma}_{\boldsymbol{n}}$ of variances of sample means for samples of size 50×50 observations from four different Gaussian random field models.

Let $\{Y_{\boldsymbol{i}}\}$ be an (intrinsically) stationary Gaussian random field, with an exponential covariance function,

$$\mathrm{cov}[Y_{\boldsymbol{i}}, Y_{\boldsymbol{j}}] = 0.245 \exp(||\boldsymbol{i} - \boldsymbol{j}||/0.723), \tag{8}$$

and with $E[Y_{\boldsymbol{i}}] = 0$ for all $\boldsymbol{i}$. The covariance function (8) is taken from Lee and Lahiri (2002), who fitted a variogram model to agricultural field data.

The four Gaussian random field models are defined as follows:

1. $X_{\boldsymbol{i}} = Y_{\boldsymbol{i}}$,
2. $X_{\boldsymbol{i}} = Y_{\boldsymbol{i}} + \cos(i_1/n_1) + \cos(i_2/n_2)$,
3. $X_{\boldsymbol{i}} = Y_{\boldsymbol{i}} a_{\boldsymbol{i}}^{1/2} b + a_{\boldsymbol{i}}$,
4. $X_{\boldsymbol{i}} = Y_{\boldsymbol{i}} a_{\boldsymbol{i}}^{1/2} b$,

where $a_{\boldsymbol{i}} = \cos(i_1/n_1)\cos(i_2/n_2)$ and $b = 1.194...$, $i_1 = 1, ..., n_1$ and $i_2 = 1, ..., n_2$. Thus, Model 2 is a two-way linear model, and in Model 3, the variances, $\mathrm{var}[X_{\boldsymbol{i}}]$, are nonconstant and proportional to the expected values, $E[X_{\boldsymbol{i}}]$. Model 4 is defined as Model 3, but without the spatial trend component. The parameter b is chosen so that the true variance $\gamma_{\boldsymbol{n}} = \gamma_{\boldsymbol{50}} \approx 0.852$ in all four models. Realizations of the random field $\{Y_{\boldsymbol{i}}\}$ were generated using the *rfsim* function in S+SpatialStats, an add-on module to the S-Plus version 6.0 statistical software package.

Empirical rules for choosing the subshape size are computationally demanding for spatial lattice data, and for getting the simulations done within a reasonable amount of time we decided to use subshapes of size 7×7 throughout this study. With this subshape size, k_i is of the order of $n_i^{1/2}$, $i = 1, 2$, which is desired for obtaining the optimal rate of convergence of $\tilde{\gamma}_{\boldsymbol{n}}$ and $\hat{\gamma}_{\boldsymbol{n}}$.

Although the values $\boldsymbol{d}^\bullet$ may be found from visual inspections of plots of realizations of $\gamma_{\boldsymbol{n}}^\bullet(\boldsymbol{d})$, this would be impracticable in a simulation study. Therefore we use an empirical rule, where we define $d^\bullet = d_1^\bullet = d_2^\bullet$ as the smallest $d = d_1 = d_2$ for which $\gamma_{\boldsymbol{n}}^\bullet(d+1, d+1) - \gamma_{\boldsymbol{n}}^\bullet(d, d) < t$, and where t is some specified threshold value. The rule is simple, but seems to work well. In Figure 1, three realizations of $\gamma_{\boldsymbol{n}}^\bullet(d, d)$ are plotted against d. If the threshold

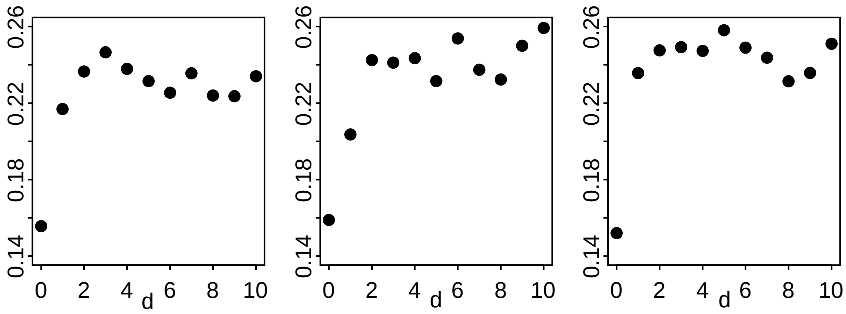

Figure 1: Three realizations of $\gamma_{\boldsymbol{n}}^{\bullet}(d_1, d_2)$, with $d = d_1 = d_2$ and $\boldsymbol{n} = (50, 50)$, from Model 3.

equals 0.01, we get $d^{\bullet}$ equal to 2, 2, and 1, respectively, and if the threshold is 0.02, we obtain the same results as for the threshold 0.01. These values of $d^{\bullet}$ agree well with what would have been obtained from a visual inspection.

The results of the simulations are summarized in Tables 1 and 2, where each entry is based on 1,000 replicates of $\{Y_{\boldsymbol{i}}\}$. In Models 1 and 4, where the $X_{\boldsymbol{i}}$'s have zero expectation, the variance estimator $\tilde{\gamma}_{\boldsymbol{n}}$ performs slightly better than the modified estimator $\hat{\gamma}_{\boldsymbol{n}}$ in terms of root mean square error (RMSE), as expected. Observe that both $\tilde{\gamma}_{\boldsymbol{n}}$ and $\hat{\gamma}_{\boldsymbol{n}}$ do well in the heteroscedastic Model 4. In Models 2 and 3, where the expected values $E[X_{\boldsymbol{i}}]$ vary, the estimator $\hat{\gamma}_{\boldsymbol{n}}$ is clearly the winner, and $\tilde{\gamma}_{\boldsymbol{n}}$ computed on the original data $\{X_{\boldsymbol{i}}\}$ fails to provide valid estimates in these two cases. If $\tilde{\gamma}_{\boldsymbol{n}}$ is computed on detrended observations (obtained through mean polishing) in Model 2, then the resulting RMSE is still more than 50% larger than for $\hat{\gamma}_{\boldsymbol{n}}$. The estimator $\tilde{\gamma}_{\boldsymbol{n}}$ has a large negative bias in this case, and a plausible reason for this could be that the mean polishing not only removes the trend but also some of the random variation. The modified estimator $\hat{\gamma}_{\boldsymbol{n}}$, on the other hand, gives similar values of RMSE for all four models and the two choices of threshold value. Thus, if we do not have good knowledge of whether the data have constant expected values, then a safe policy is to choose $\hat{\gamma}_{\boldsymbol{n}}$ instead of $\tilde{\gamma}_{\boldsymbol{n}}$.

Table 1

Estimated biases and RMSEs of $\tilde{\gamma}_{\boldsymbol{n}}$

Model	Data	$\widehat{\text{bias}}[\tilde{\gamma}_{\boldsymbol{n}}]$	$\widehat{\text{RMSE}}[\tilde{\gamma}_{\boldsymbol{n}}]$
1	Original	-0.169	0.205
2	Original	1.330	1.375
1 and 2	Detrended	-0.356	0.366
3	Original	0.917	0.967
4	Original	-0.160	0.199

Table 2

Estimated biases and RMSEs of $\hat{\gamma}_{\boldsymbol{n}}$, together with corresponding mean values and sample variances of $d^\bullet$

Model	Data	Threshold	$\widehat{\text{bias}}[\hat{\gamma}_{\boldsymbol{n}}]$	$\widehat{\text{RMSE}}[\hat{\gamma}_{\boldsymbol{n}}]$	$\bar{d}^\bullet$	$\widehat{\text{var}}(d^\bullet)$
1 and 2	Original	0.01	-0.158	0.237	2.07	0.46
1 and 2	Original	0.02	-0.167	0.234	1.47	0.28
3	Original	0.01	-0.137	0.231	2.12	0.51
3	Original	0.02	-0.147	0.227	1.51	0.30
4	Original	0.01	-0.138	0.231	2.13	0.51
4	Original	0.02	-0.147	0.227	1.51	0.30

Appendix

Define $\mathcal{S}_{\boldsymbol{n}} = A_{\boldsymbol{n}} \cap \mathbb{Z}^2$, $\mathcal{S}_{\boldsymbol{k},t} = A_{\boldsymbol{k},t} \cap \mathbb{Z}^2$, and $\mathcal{V}_{\boldsymbol{n}} = \{\boldsymbol{i} : \boldsymbol{i} + \boldsymbol{h} \in \mathcal{S}_{\boldsymbol{n}}$ for all $\boldsymbol{h}$ such that $|h_1| \leq k_1$ and $|h_2| \leq k_2\}$. Then $\mathcal{V}^c_{\boldsymbol{n}} = \mathcal{S}_{\boldsymbol{n}} \setminus \mathcal{V}_{\boldsymbol{n}}$ defines a strip around the outside of $\mathcal{V}_{\boldsymbol{n}}$. Let $V = \#\mathcal{V}_{\boldsymbol{n}}$.

The abbreviation E&S is used throughout the Appendix for the article by Ekström and Sjöstedt-de Luna (2004a). To simplify the notation, we write $g_A = g(A)$ and $g_t = g(A_{\boldsymbol{k},t})$, $t = 1, ..., t_{\boldsymbol{n}}$.

Lemma 1 (E&S) *There exist a constant $n > 0$ such that for all $n_1, n_2 > n$, $N \geq S \geq t_{\boldsymbol{n}} \geq V \geq cN$ and $K \geq s \geq cK$. Further, $V/S \to 1$, $t'_{\boldsymbol{n}}/t_{\boldsymbol{n}} \to 1$, and $S - V = O((KN)^{1/2})$ as $\boldsymbol{n} \to \infty$.*

It follows from Lemma 1 that S, $t_{\boldsymbol{n}}$, and V are all of $O(N)$ and that $s = O(K)$. Henceforth we will not refer to Lemma 1 when applying these simple consequences of it.

Lemma 2 *Let Y_j be measurable with respect to the σ-field $\mathcal{F}(U_j)$, where $|U_j| \leq l$, $j = 1, 2$, and $\rho(U_1, U_2) \geq m$.*

(a) If $||Y_j||_{2+\delta} < \infty$, $j = 1, 2$, $\delta > 0$, then

$$|cov[Y_1, Y_2]| \leq c||Y_1||_{2+\delta}||Y_2||_{2+\delta}\alpha_l(m)^{\delta/(2+\delta)}$$

(b) If $|Y_j| \leq v$ a.s., $j = 1, 2$, then $|cov[Y_1, Y_2]| \leq cv^2\alpha_l(m)$.

(c) If $||Y_j||_2 \leq v$, $j = 1, 2$, then for any $\eta > 0$,

$$|cov[Y_1, Y_2]| \leq c(\eta^2\alpha_l(m) + v(||^\eta Y_1||_2 + ||^\eta Y_2||_2))$$

.

Proof. For (a) and (b), see Lin and Lu (1996, Lemma 6.1.1). Inequality (c) follows from (b) and the argument in Carlstein (1986, Lemma 1). □

Proof of Theorem 1. The proof follows closely that of Sherman (1996, Theorem 1). Since $E[(\bar{g}' - E[g_A])^2] = \mathrm{var}[\bar{g}'] + (E[\bar{g}'] - E[g_A])^2$, it suffices to show that $\mathrm{var}[\bar{g}'] \to 0$ as $\boldsymbol{n} \to \infty$. Let $k = k_1 \vee k_2$. We have

$$\mathrm{var}[\bar{g}'] = \frac{1}{t^2_{\boldsymbol{n}}} \sum_{t_a, t_b = 1}^{t_{\boldsymbol{n}}} \mathrm{cov}[g_{t_a}, g_{t_b}] = \frac{1}{t^2_{\boldsymbol{n}}} {\sum}' \mathrm{cov}[g_{t_a}, g_{t_b}] + \frac{1}{t^2_{\boldsymbol{n}}} {\sum}'' \mathrm{cov}[g_{t_a}, g_{t_b}], \quad (9)$$

where $\sum'$ denotes the sum over t_a, t_b such that $\rho(\mathcal{S}_{\boldsymbol{k},t_a}, \mathcal{S}_{\boldsymbol{k},t_b}) \leq k$ and $\sum''$ denotes the sum over t_a, t_b such that $\rho(\mathcal{S}_{\boldsymbol{k},t_a}, \mathcal{S}_{\boldsymbol{k},t_b}) > k$. There are at most $16k^2 t_{\boldsymbol{n}}$ summands in $\sum'$ and it follows from the U.I. condition that there exists a constant v such that $||g_t||_2 \leq v$ and $|\text{cov}[g_{t_a}, g_{t_b}]| \leq ||g_{t_a}||_2 ||g_{t_b}||_2 \leq v^2$ uniformly in $t, t_a, t_b, \boldsymbol{n}$. Thus, $t_{\boldsymbol{n}}^{-2} \sum' \text{cov}[g_{t_a}, g_{t_b}] \leq 16v^2k^2t_{\boldsymbol{n}}^{-1} \to 0$ as $\boldsymbol{n} \to \infty$. Finally, by Lemma 2(c) and the U.I. condition,

$$\overline{\lim_{\boldsymbol{n}\to\infty}}\, t_{\boldsymbol{n}}^{-2} \sum\nolimits'' \text{cov}[g_{t_a}, g_{t_b}] \leq \lim_{\eta\to\infty} \overline{\lim_{\boldsymbol{n}\to\infty}}\, c\eta^2 \alpha_s(k) + 2cv \lim_{\eta\to\infty} \sup_{t,\boldsymbol{n}} ||^\eta g_t||_2 = 0.$$

□

Proof of Theorem 2. The proof follows closely that of Sherman (1996, Theorem 2). Namely, we can write the estimator of variance as

$$\begin{aligned} \tilde{\gamma}_{\boldsymbol{n}} &= st_{\boldsymbol{n}}^{-1} \sum (g_t - E[g_t])^2 - s(\bar{g}' - E[\bar{g}'])^2 + st_{\boldsymbol{n}}^{-1} \sum (E[g_t] - E[\bar{g}'])^2 \\ &\quad + 2st_{\boldsymbol{n}}^{-1} \sum (g_t - E[g_t])(E[g_t] - E[\bar{g}']) \\ &\triangleq T_{\boldsymbol{n}1} + T_{\boldsymbol{n}2} + T_{\boldsymbol{n}3} + T_{\boldsymbol{n}4}. \end{aligned} \tag{10}$$

By applying Theorem 1 with $h_A = S(g_A - E[g_A])^2$ we see that $T_{\boldsymbol{n}1} - \gamma_{\boldsymbol{n}} \overset{L_2}{\to} 0$. Moreover, by applying Theorem 1 with $h_A = \sqrt{S}(g_A - E[g_A])$ we see that $\sqrt{s}(\bar{g}' - E[\bar{g}']) \overset{L_2}{\to} 0$, and, by the arguments in the proof of Theorem 2 in Sherman (1996), we actually have $\sqrt{s}(\bar{g}' - E[\bar{g}']) \overset{L_4}{\to} 0$, i.e., $T_{\boldsymbol{n}2} \overset{L_2}{\to} 0$. We have, by assumption, $T_{\boldsymbol{n}3} \to 0$, and by the Cauchy-Schwarz inequality, $E[T_{\boldsymbol{n}4}^2] \leq 4E[T_{\boldsymbol{n}1}]T_{\boldsymbol{n}3}$. By the U.I. condition there exists a constant c such that $\text{var}[\sqrt{s}g_t] \leq c$ uniformly in $t, \boldsymbol{n}$, implying that $E[T_{\boldsymbol{n}1}] \leq c$ for all $\boldsymbol{n}$. Hence, $E[T_{\boldsymbol{n}4}^2] \leq 4cT_{\boldsymbol{n}3} \to 0$ and we have shown that $\tilde{\gamma}_{\boldsymbol{n}} - \gamma_{\boldsymbol{n}} = T_{\boldsymbol{n}1} + T_{\boldsymbol{n}2} + T_{\boldsymbol{n}3} + T_{\boldsymbol{n}4} - \gamma_{\boldsymbol{n}} \overset{L_2}{\to} 0$ as $\boldsymbol{n} \to \infty$. □

Proof of Corollary 1. From the proof of Theorem 1,

$$\text{var}[\bar{g}'] \leq 16v^2k^2t_{\boldsymbol{n}}^{-1} + c\eta^2\alpha_s(k) + 2cv\,\text{sup}_t ||^\eta g_t||_2.$$

If we put $\eta = s$, then

$$\begin{aligned} \text{var}[\bar{g}'] &\leq 16v^2k^2t_{\boldsymbol{n}}^{-1} + cs^2\alpha_s(k) + 2cv\,\text{sup}_t ||^s g_t||_2 \\ &\leq 16v^2k^2t_{\boldsymbol{n}}^{-1} + cs^2\alpha_s(k) + 2cvs^{-1}\,\text{sup}_t ||^s g_t||_4^2 \\ &= O(K/N + 1/K + \lambda_{\boldsymbol{n}}). \end{aligned}$$

The desired result follows from the identity $E[(\bar{g}'-E[g_A])^2] = \text{var}[\bar{g}']+(E[\bar{g}']-E[g_A])^2$ and the assumption that $|E[\bar{g}'] - E[g_A]| = O(\kappa_{\boldsymbol{n}})$. □

Proof of Corollary 2. Consider (10). By applying Corollary 1 with $h_A = S(g_A - E[g_A])^2$ we see that $E[(T_{\boldsymbol{n}1} - \gamma_{\boldsymbol{n}})^2] = O(K/N + 1/K + \varrho^2_{\boldsymbol{n}} + \lambda_{\boldsymbol{n}})$. Moreover, by applying Corollary 1 with $h_A = \sqrt{S}(g_A - E[g_A])$ we see that $E[T^2_{\boldsymbol{n}2}] = O(K/N + 1/K + \lambda_{\boldsymbol{n}})$. By assumption, $T^2_{\boldsymbol{n}3} = O(\zeta^2_{\boldsymbol{n}})$, and by the proof of Theorem 2, $E[T^2_{\boldsymbol{n}4}] \leq 4cT_{\boldsymbol{n}3} = O(\zeta_{\boldsymbol{n}})$. This completes the proof of Corollary 2. □

Define $a_{\boldsymbol{h}} = \#\{\boldsymbol{i} : \boldsymbol{i}, \boldsymbol{i}+\boldsymbol{h} \in \mathcal{S}_{\boldsymbol{k}}\}$, i.e., for each $\boldsymbol{h}$, $a_{\boldsymbol{h}}$ is the number of pixels $\boldsymbol{i}$ such that both $\boldsymbol{i}$ and $\boldsymbol{i}+\boldsymbol{h}$ belong to $\mathcal{S}_{\boldsymbol{k}}$. Further, let $w_{\boldsymbol{h}} = a_{\boldsymbol{h}}/a_{\boldsymbol{0}} = a_{\boldsymbol{h}}/s$, and note that $0 \leq w_{\boldsymbol{h}} \leq 1$ for all $\boldsymbol{h}$.

Lemma 3 $|1 - w_{\boldsymbol{h}}| \leq cK^{-1/2}(|h_1| \vee |h_2|)$ *uniformly in* $\boldsymbol{h}$.

Proof. The result is implicitly contained in the proof of Lemma 3 in E&S. □

Lemma 4 *Assume for some* $\delta, \beta_\delta > 0$ *that* $||X_{\boldsymbol{i}} - E[X_{\boldsymbol{i}}]||_{2+\delta} \leq \beta_\delta$ *for all* $\boldsymbol{i}, \boldsymbol{n}$ *and* $\sum_{m=0}^{\infty}(m+1)\alpha_1(m)^{\delta/(2+\delta)} < \infty$. *Then* $t_{\boldsymbol{n}}^{-1}\sum_{t=1}^{t_n} \mathit{var}[\sqrt{s}\bar{X}_t] - \gamma_{\boldsymbol{n}} \to 0$ *as* $\boldsymbol{n} \to \infty$.

Proof. Let $\gamma^*_{\boldsymbol{n}} = t_{\boldsymbol{n}}^{-1}\sum_{t=1}^{t_n} \text{var}[\sqrt{s}\bar{X}_t]$, $c_{\boldsymbol{ih}} = \text{cov}[X_{\boldsymbol{i}}, X_{\boldsymbol{i}+\boldsymbol{h}}]$, $I_{\boldsymbol{ih}} = I_{\{\boldsymbol{i},\boldsymbol{i}+\boldsymbol{h}\in\mathcal{S}_n\}}$, and $w_{\boldsymbol{ih}} = a_{\boldsymbol{ih}}/s$, where $a_{\boldsymbol{ih}} = \#\{t : \boldsymbol{i}, \boldsymbol{i}+\boldsymbol{h} \in \mathcal{S}_{\boldsymbol{k},t}\}$, that is, the number of subshapes that includes both pixels $\boldsymbol{i}$ and $\boldsymbol{i}+\boldsymbol{h}$. Then $\gamma_{\boldsymbol{n}} = S^{-1}\sum_{\boldsymbol{h}}\sum_{\boldsymbol{i}} c_{\boldsymbol{ih}}I_{\boldsymbol{ih}}$, $\gamma^*_{\boldsymbol{n}} = t_{\boldsymbol{n}}^{-1}\sum_{\boldsymbol{h}}\sum_{\boldsymbol{i}} c_{\boldsymbol{ih}}w_{\boldsymbol{ih}}I_{\boldsymbol{ih}}$, and

$$\begin{aligned}
\gamma^*_{\boldsymbol{n}} - \gamma_{\boldsymbol{n}} &= (t_{\boldsymbol{n}}^{-1} - S^{-1})\sum_{\boldsymbol{h}}\sum_{\boldsymbol{i}} c_{\boldsymbol{ih}}I_{\boldsymbol{ih}} + t_{\boldsymbol{n}}^{-1}\sum_{\boldsymbol{h}}\sum_{\boldsymbol{i}\in\mathcal{V}_n} c_{\boldsymbol{ih}}(w_{\boldsymbol{ih}} - 1)I_{\boldsymbol{ih}} \\
&\quad + t_{\boldsymbol{n}}^{-1}\sum_{\boldsymbol{h}}\sum_{\boldsymbol{i}\in\mathcal{V}^c_{\boldsymbol{n}}} c_{\boldsymbol{ih}}(w_{\boldsymbol{ih}} - 1)I_{\boldsymbol{ih}} \\
&\triangleq U_{\boldsymbol{n}1} + U_{\boldsymbol{n}2} + U_{\boldsymbol{n}3}.
\end{aligned}$$

By Lemma 2(a), $|c_{\boldsymbol{ih}}| = |\text{cov}[X_{\boldsymbol{i}} - E[X_{\boldsymbol{i}}], X_{\boldsymbol{i}+\boldsymbol{h}} - E[X_{\boldsymbol{i}+\boldsymbol{h}}]]| \leq c\beta^2_\delta\alpha_1(h)^{\delta'}$, where $h = |h_1| \vee |h_2|$ and $\delta' = \delta/(2+\delta)$. Thus

$$|\sum_{\boldsymbol{h}}\sum_{\boldsymbol{i}} c_{\boldsymbol{ih}} I_{\boldsymbol{ih}}| \le c\beta_\delta^2 \sum_{\boldsymbol{h}}\sum_{\boldsymbol{i}} \alpha_1(h)^{\delta'} I_{\boldsymbol{ih}} \le c\beta_\delta^2 S \sum_{h_1,h_2=-\infty}^{\infty} \alpha_1(h)^{\delta'}$$
$$\le\ 8c\beta_\delta^2 S \sum_{h_1=0}^{\infty}\sum_{h_2=0}^{h_1} \alpha_1(h_1)^{\delta'} \le 8c\beta_\delta^2 S \sum_{h_1=0}^{\infty} (h_1+1)\alpha_1(h_1)^{\delta'} = O(S). \quad (11)$$

By Lemmas 1 and 2 in E&S, $t_{\boldsymbol{n}}^{-1} - S^{-1} = S^{-1}O((K/N)^{1/2})$, and together with (11) this implies that

$$U_{\boldsymbol{n}1} = O(K^{1/2}/N^{1/2}). \quad (12)$$

Consider $U_{\boldsymbol{n}2}$. Let $k = k_1 \vee k_2$. Note that $w_{\boldsymbol{ih}} = w_{\boldsymbol{h}}$ if $\boldsymbol{i} \in \mathcal{V}_{\boldsymbol{n}}$ and that $V/t_{\boldsymbol{n}} \le 1$ by Lemma 1. Then, by Lemma 2(a) and Lemma 3,

$$|U_{\boldsymbol{n}2}| \le c_1\beta_\delta^2 t_{\boldsymbol{n}}^{-1} \sum_{|h_j|\le k_j}\sum_{\boldsymbol{i}\in\mathcal{V}_{\boldsymbol{n}}} \alpha_1(h_1)^{\delta'}|w_{\boldsymbol{h}}-1| \le c_1\beta_\delta^2 \sum_{|h_j|\le k_j} \alpha_1(h_1)^{\delta'}|w_{\boldsymbol{h}}-1|$$
$$\le 8c_1c_2\beta_\delta^2 K^{-1/2} \sum_{h_1=0}^{k}\sum_{h_2=0}^{h_1} h_1\alpha_1(h_1)^{\delta'} \le 8c_1c_2\beta_\delta^2 K^{-1/2} \sum_{h=0}^{k} h(h+1)\alpha_1(h)^{\delta'}$$
$$\le 8c_1c_2\beta_\delta^2 K^{-1/2}[k^{1/2}\sum_{0\le h\le k^{1/2}} (h+1)\alpha_1(h)^{\delta'} + k\sum_{k^{1/2}<h\le k}(h+1)\alpha_1(h)^{\delta'}], \quad (13)$$

for some constants c_1 and c_2. Since $\sum_{m=0}^{\infty}(m+1)\alpha_1(m)^{\delta'}$ is finite, the right hand side above tends to 0 as $\boldsymbol{n}\to\infty$, implying that $U_{\boldsymbol{n}2}\to 0$ as $\boldsymbol{n}\to\infty$.

Similarly,

$$|U_{\boldsymbol{n}3}| = O((S-V)/t_{\boldsymbol{n}}) = O(K^{1/2}/N^{1/2}). \quad (14)$$

where the last equality follows from Lemma 1. □

Lemma 5 *Assume, for some* $\delta, \beta_\delta > 0$, $r\in\mathbb{N}$, *and* $\max\{2(r-1),2\}\le\tau\le 2r$, *that* $||X_{\boldsymbol{i}} - E[X_{\boldsymbol{i}}]||_{\tau+\delta} \le \beta_\delta$ *for all* $\boldsymbol{i},\boldsymbol{n}$ *and that* $\sum_{m=1}^{\infty} m^{4r-3}\alpha_{2r-1}(m)^{\delta/(2r+\delta)}$ *is finite. Then*

$$||\sum_{\boldsymbol{i}\in\mathcal{S}_{\boldsymbol{n}}}(X_{\boldsymbol{i}} - E[X_{\boldsymbol{i}}])||_\tau \le c\sqrt{S},$$

where c *is a constant that depends on* r,τ,β_δ, *and the strong mixing coefficients, but not on the set* $\mathcal{S}_{\boldsymbol{n}}$.

Proof. Follows from Fazekas et al. (2000, Theorem 1) and Jensen's inequality. □

Lemma 6 *Assume, for some* $\delta, \beta_\delta > 0$*, that* $||X_{\boldsymbol{i}} - E[X_{\boldsymbol{i}}]||_{4+\delta} \leq \beta_\delta$ *for all* $\boldsymbol{i}, \boldsymbol{n}$ *and that* $\sum_{m=1}^{\infty} m^9 \alpha_5(m)^{\delta/(6+\delta)} < \infty$*. Then* $\{(\sqrt{s}(\bar{X}_t - E[\bar{X}_t]))^4 : t = 1, ..., t_{\boldsymbol{n}}, n_1, n_2 = 1, 2, ...\}$ *is U.I.*

Proof. By Shiryayev (1984, Lemma 3, p. 188) it suffices to show that $\sup_{t,\boldsymbol{n}} E[(\sqrt{s}(\bar{X}_t - E[\bar{X}_t]))^{4+\delta}]$ is finite. The finiteness follows as a direct consequence of Lemma 5. □

Proof of Theorem 3. Follows from Theorem 2 and Lemmas 4 and 6. □

Proof of Theorem 4. We have

$$\begin{aligned} \hat{\gamma}_{\boldsymbol{n}} &= \frac{s}{t'_{\boldsymbol{n}}} \sum_{t=1}^{t'_{\boldsymbol{n}}} (Z_t - E[Z_t])^2 - s(\bar{Z} - E[\bar{Z}])^2 + \frac{s}{t'_{\boldsymbol{n}}} \sum_{t=1}^{t'_{\boldsymbol{n}}} (E[Z_t] - E[\bar{Z}])^2 \\ &\quad + \frac{2s}{t'_{\boldsymbol{n}}} \sum_{t=1}^{t'_{\boldsymbol{n}}} (Z_t - E[Z_t])(E[Z_t] - E[\bar{Z}]) \\ &\triangleq R_{\boldsymbol{n}1} + R_{\boldsymbol{n}2} + R_{\boldsymbol{n}3} + R_{\boldsymbol{n}4}. \end{aligned} \tag{15}$$

Since $R_{\boldsymbol{n}1}$ and $R_{\boldsymbol{n}2}$ are based on centered versions of Z_t and $\bar{Z}$, respectively, we assume, without loss of generality, that $E[X_{\boldsymbol{i}}] \equiv 0$ when considering $R_{\boldsymbol{n}1}$ and $R_{\boldsymbol{n}2}$.

We have

$$\begin{aligned} R_{\boldsymbol{n}1} &= \frac{s}{4t'_{\boldsymbol{n}}} \sum_{t=1}^{t'_{\boldsymbol{n}}} \sum_{u=1}^{4} \bar{X}_{t_u}^2 + \frac{s}{2t'_{\boldsymbol{n}}} \sum_{t=1}^{t'_{\boldsymbol{n}}} \sum_{u=2}^{4} \sum_{v=1}^{u-1} (-1)^{u+v} \bar{X}_{t_u} \bar{X}_{t_v} \\ &\triangleq R_{\boldsymbol{n}1a} + R_{\boldsymbol{n}1b}. \end{aligned} \tag{16}$$

From Lemma 1, Lemma 5, and Theorem 3,

$$E[(R_{\boldsymbol{n}1a} - \gamma_{\boldsymbol{n}})^2] \to 0 \text{ as } \boldsymbol{n} \to \infty. \tag{17}$$

The proof of (17) is essentially the same as the corresponding part of the proof of Theorem 2 in E&S (only a slight change of notation is needed and the references to Lemma 4 and Theorem 1 in E&S should be changed to Lemma 5 and Theorem 3 in the current paper).

Let $Q_{\boldsymbol{n}uv} = s(t'_{\boldsymbol{n}})^{-1}\sum_{t=1}^{t'_{\boldsymbol{n}}} \bar{X}_{t_u}\bar{X}_{t_v}$. Then

$$R_{\boldsymbol{n}1b} = \sum_{u=2}^{4}\sum_{v=1}^{u-1}(-1)^{u+v}Q_{\boldsymbol{n}uv}/2$$

and by Lemma 2,

$$|E[Q_{\boldsymbol{n}uv}]| = \frac{1}{st'_{\boldsymbol{n}}}|\sum_{t=1}^{t'_{\boldsymbol{n}}}\sum_{\boldsymbol{i}\in\mathcal{S}_{\boldsymbol{k},t_u}}\sum_{\boldsymbol{j}\in\mathcal{S}_{\boldsymbol{k},t_v}}\text{cov}[X_{\boldsymbol{i}}, X_{\boldsymbol{j}}]| \leq \frac{c\beta_\delta^2}{s}\sum_{\boldsymbol{i}\in\mathcal{S}_{\boldsymbol{k},t_u}}\sum_{\boldsymbol{j}\in\mathcal{S}_{\boldsymbol{k},t_v}}\alpha_1(l_1 \vee l_2)^{\delta'},$$

where $l_1 = |i_1 - j_1|$, $l_2 = |i_2 - j_2|$, and $\delta' = \delta/(2+\delta)$. Consider the case $u = 2$ and $v = 1$. By a counting argument,

$$\begin{aligned}
|E[Q_{\boldsymbol{n}21}]| &\leq \frac{c\beta_\delta^2}{s}\sum_{\boldsymbol{i}\in\mathcal{S}_{\boldsymbol{k},t_u}}\sum_{\boldsymbol{j}\in\mathcal{S}_{\boldsymbol{k},t_v}}\alpha_1(l_1 \vee l_2)^{\delta'} \\
&= \frac{c\beta_\delta^2}{s}\sum_{l_1=d_1+1}^{d_1+k_1-1}\sum_{l_2=0}^{k_2}(1+I_{\{l_2\neq 0\}})(k_2-l_2)(l_1-d_1)\alpha_1(l_1\vee l_2)^{\delta'} \\
&\quad + \frac{c\beta_\delta^2}{s}\sum_{l_1=d_1+k_1}^{d_1+2k_1}\sum_{l_2=0}^{k_2}(1+I_{\{l_2\neq 0\}})(k_2-l_2)(2k_1+d_1-l_1)\alpha_1(l_1\vee l_2)^{\delta'} \\
&\leq \frac{2c\beta_\delta^2 k_2}{s}(\sum_{l_1=d_1+1}^{d_1+k_1-1}\sum_{l_2=0}^{k_2\wedge(l_1-1)} l_1\alpha_1(l_1)^{\delta'} + \sum_{l_1=d_1+1}^{d_1+k_1-1}\sum_{l_2=k_2\wedge l_1}^{k_2} l_1\alpha_1(l_2)^{\delta'}) \\
&\quad + \frac{2c\beta_\delta^2 k_2(2k_1+d_1)}{s}(\sum_{l_1=d_1+k_1}^{d_1+2k_1}\sum_{l_2=0}^{k_2\wedge(l_1-1)}\alpha_1(l_1)^{\delta'} + \sum_{l_1=d_1+k_1}^{d_1+2k_1}\sum_{l_2=k_2\wedge l_1}^{k_2}\alpha_1(l_2)^{\delta'}) \\
&= O(k_1^{-1}),
\end{aligned}$$

where the last equality follows from A1(iii). In general, we have, by using similar arguments as above, $E[Q_{\boldsymbol{n}uv}] = O(k_1^{-1} + k_2^{-1})$ for all $u \neq v$, implying that

$$E[R_{\boldsymbol{n}1b}] = O(k_1^{-1} + k_2^{-1}). \tag{18}$$

Further,

$$\text{var}[Q_{\boldsymbol{n}uv}] = \frac{s^2}{(t'_{\boldsymbol{n}})^2}\sum{}'\text{cov}[\bar{X}_{t_u}\bar{X}_{t_v}, \bar{X}_{l_u}\bar{X}_{l_v}] + \frac{s^2}{(t'_{\boldsymbol{n}})^2}\sum{}''\text{cov}[\bar{X}_{t_u}\bar{X}_{t_v}, \bar{X}_{l_u}\bar{X}_{l_v}],$$

where $\sum'$ denotes the sum over t, l such that $\rho(\mathcal{S}_{\boldsymbol{k},t_u} \cup \mathcal{S}_{\boldsymbol{k},t_v}, \mathcal{S}_{\boldsymbol{k},l_u} \cup \mathcal{S}_{\boldsymbol{k},l_v}) \leq k_*$, $\sum''$ denotes the sum over t, l such that $\rho(\mathcal{S}_{\boldsymbol{k},t_u} \cup \mathcal{S}_{\boldsymbol{k},t_v}, \mathcal{S}_{\boldsymbol{k},l_u} \cup \mathcal{S}_{\boldsymbol{k},l_v}) > k_*$, and $k_* = (2k_1 + d_1) \vee (2k_2 + d_2)$. There are at most $16k_*^2 t'_{\boldsymbol{n}}$ summands in $\sum'$ and by applying Lemma 5 we see that

$$
\begin{aligned}
&|\mathrm{cov}[\bar{X}_{t_u}\bar{X}_{t_v}, \bar{X}_{l_u}\bar{X}_{l_v}]| \leq (\mathrm{var}[\bar{X}_{t_u}\bar{X}_{t_v}]\mathrm{var}[\bar{X}_{l_u}\bar{X}_{l_v}])^{1/2} \\
&\quad \leq (E[\bar{X}_{t_u}^2\bar{X}_{t_v}^2]E[\bar{X}_{l_u}^2\bar{X}_{l_v}^2])^{1/2} \leq (E[\bar{X}_{t_u}^4]E[\bar{X}_{t_v}^4]E[\bar{X}_{l_u}^4]E[\bar{X}_{l_v}^4])^{1/4} \leq cs^{-2}.
\end{aligned}
$$

Thus, $s^2(t'_{\boldsymbol{n}})^{-2}|\sum' \mathrm{cov}[\bar{X}_{t_u}\bar{X}_{t_v}, \bar{X}_{l_u}\bar{X}_{l_v}]| \leq 16ck_*^2(t'_{\boldsymbol{n}})^{-1} = O(K/N)$. Let $K_* = (2k_1 + d_1)(2k_2 + d_2)$. Then, by Lemmas 2(a) and 5,

$$
\begin{aligned}
&|\mathrm{cov}[\bar{X}_{t_u}\bar{X}_{t_v}, \bar{X}_{l_u}\bar{X}_{l_v}]| \leq c(E[|\bar{X}_{t_u}\bar{X}_{t_v}|^{2+\delta}]E[\bar{X}_{l_u}\bar{X}_{l_v}|^{2+\delta}])^{1/(2+\delta)}\alpha_{K_*}(k_*)^{\delta'} \\
&\quad \leq c_1(E[|\bar{X}_{t_u}|^{4+2\delta}]E[|\bar{X}_{t_v}|^{4+2\delta}]E[|\bar{X}_{l_u}|^{4+2\delta}]E[|\bar{X}_{l_v}|^{4+2\delta}])^{1/(4+2\delta)}\alpha_{K_*}(k_*)^{\delta'} \\
&\quad \leq c_1c_2s^{-2}\alpha_{K_*}(k_*)^{\delta'},
\end{aligned}
$$

for some constants c_1 and c_2. Thus, $s^2(t'_{\boldsymbol{n}})^{-2}|\sum'' \mathrm{cov}[\bar{X}_{t_u}\bar{X}_{t_v}, \bar{X}_{l_u}\bar{X}_{l_v}]| \leq c\alpha_{K_*}(k_*)^{\delta'} = O(K^{(\tau_2 - \tau_1/2)\delta'})$, and we have shown that

$$
\mathrm{var}[R_{\boldsymbol{n}1b}] = O(K^{(\tau_2-\tau_1/2)\delta'} + K/N). \tag{19}
$$

By Lemma 5,

$$
E[R_{\boldsymbol{n}2}^2] = s^2E[(\frac{1}{t'_{\boldsymbol{n}}}\sum_{\boldsymbol{i}\in\mathcal{S}_{\boldsymbol{n}}}\beta_{\boldsymbol{i}}X_{\boldsymbol{i}})^4] \leq \frac{cs^2S^2}{(t'_{\boldsymbol{n}})^4} = O(\frac{K^2}{N^2}), \tag{20}
$$

for some "weights" $\{\beta_{\boldsymbol{i}}\}$ satisfying $|\beta_{\boldsymbol{i}}| \leq 1$ for all $\boldsymbol{i}$.

Let $\bar{Z}_y = \sum_{t=1}^{t'_{\boldsymbol{n}}} Z_{Y,t}/t'_{\boldsymbol{n}}$. Then,

$$
\begin{aligned}
R_{\boldsymbol{n}3} &\leq \frac{2s}{t'_{\boldsymbol{n}}}\sum_{t=1}^{t'_{\boldsymbol{n}}}(E[Z_{Y,t}] - E[\bar{Z}_Y])^2 + \frac{2s}{t'_{\boldsymbol{n}}}\sum_{t=1}^{t'_{\boldsymbol{n}}}(Z_{\mu,t} - \bar{Z}_\mu)^2 \\
&\triangleq R_{\boldsymbol{n}3a} + R_{\boldsymbol{n}3b}.
\end{aligned} \tag{21}
$$

Let $\bar{Y}'_u = \sum_{t=1}^{t'_{\boldsymbol{n}}} \bar{Y}_{t_u}/t'_{\boldsymbol{n}}$, $u = 1, 2, 3, 4$. By Jensen's inequality,

$$
\begin{aligned}
R_{\boldsymbol{n}3a} &= \frac{s}{2t'_{\boldsymbol{n}}}\sum_{t=1}^{t'_{\boldsymbol{n}}}[\sum_{u=1}^{4}(-1)^u(E[\bar{Y}_{t_u}] - E[\bar{Y}'_u])]^2 \leq \frac{2s}{t'_{\boldsymbol{n}}}\sum_{u=1}^{4}\sum_{t=1}^{t'_{\boldsymbol{n}}}(E[\bar{Y}_{t_u}] - E[\bar{Y}'_u])^2 \\
&\leq \frac{4s}{t'_{\boldsymbol{n}}}\sum_{u=1}^{4}\sum_{t=1}^{t'_{\boldsymbol{n}}}(E[\bar{Y}_{t_u}] - E[\bar{Y}'])^2 + 4s\sum_{u=1}^{4}(E[\bar{Y}'_u] - E[\bar{Y}'])^2 \\
&\triangleq R'_{\boldsymbol{n}3a} + R''_{\boldsymbol{n}3a}.
\end{aligned} \tag{22}
$$

Again, by Jensen's inequality,

$$R''_{\boldsymbol{n}3a} = 4s\sum_{u=1}^{4}[(t'_{\boldsymbol{n}})^{-1}\sum_{t=1}^{t'_{\boldsymbol{n}}}(E[\bar{Y}_{t_u}] - E[\bar{Y}'])]^2 \le R'_{\boldsymbol{n}3a}, \tag{23}$$

and $R'_{\boldsymbol{n}3a} \to 0$ by assumption A1(ii) and since $t'_{\boldsymbol{n}}/t_{\boldsymbol{n}} \to 1$ (Lemma 1). Thus, $R_{\boldsymbol{n}3a} \to 0$ and by assumption (2), $R_{\boldsymbol{n}3b} \to 0$, which together with (21) imply

$$R_{\boldsymbol{n}3} \to 0 \text{ as } \boldsymbol{n} \to \infty. \tag{24}$$

Finally, by the Cauchy-Schwarz inequality, (24), and by noting that $E[R_{\boldsymbol{n}1}] = E[R_{\boldsymbol{n}1a}] + E[R_{\boldsymbol{n}1b}] = O(1)$, we see that

$$E[R^2_{\boldsymbol{n}4}] \le 4E[R_{\boldsymbol{n}1}]R_{\boldsymbol{n}3} \to 0 \text{ as } \boldsymbol{n} \to \infty. \tag{25}$$

From (15-20) and (24-25), $\tilde{\gamma}_{\boldsymbol{n}} - \gamma_{\boldsymbol{n}} = R_{\boldsymbol{n}1} + R_{\boldsymbol{n}2} + R_{\boldsymbol{n}3} + R_{\boldsymbol{n}4} - \gamma_{\boldsymbol{n}} \to 0$ in L_2 norm as $\boldsymbol{n} \to \infty$, as was to be proved. □

Proof of Corollary 3. We write

$$\begin{aligned}\tilde{\gamma}_{\boldsymbol{n}} &= st^{-1}_{\boldsymbol{n}}\sum(\bar{X}_t - E[\bar{X}_t])^2 - s(\bar{X}' - E[\bar{X}'])^2 + st^{-1}_{\boldsymbol{n}}\sum(E[\bar{X}_t] - E[\bar{X}'])^2 \\ &\quad +2st^{-1}_{\boldsymbol{n}}\sum(\bar{X}_t - E[\bar{X}_t])(E[\bar{X}_t] - E[\bar{X}']) \\ &\triangleq T_{\boldsymbol{n}1} + T_{\boldsymbol{n}2} + T_{\boldsymbol{n}3} + T_{\boldsymbol{n}4}.\end{aligned} \tag{26}$$

Let $\delta' = \delta/(2+\delta)$. Since $\sum_{h=0}^{\infty} h(h+1)\alpha_1(h)^{\delta'}$ is finite by assumption, we see from (12-14) that

$$(E[T_{\boldsymbol{n}1}] - \gamma_{\boldsymbol{n}})^2 \le 3(U^2_{\boldsymbol{n}1} + U^2_{\boldsymbol{n}2} + U^2_{\boldsymbol{n}3}) = O(K/N + 1/K). \tag{27}$$

Further, if $g_t = s(\bar{X}_t - E[\bar{X}_t])^2$, then (9) holds and $\text{var}[T_{\boldsymbol{n}1}] = \text{var}[\bar{g}']$. Let $\sum'$ and $\sum''$ be defined as in the proof of Theorem 1. By Lemma 5, $|\text{cov}[g_{t_a}, g_{t_b}]| \le (E[g^2_{t_a}]E[g^2_{t_b}])^{1/2} \le c$, for some $c > 0$, and by the arguments in the proof of Theorem 1, $t^{-2}_{\boldsymbol{n}}\sum'\text{cov}[g_{t_a}, g_{t_b}] = O(K/N)$. By Lemma 2(a) and A2(iii), $t^{-2}_{\boldsymbol{n}}\sum''\text{cov}[g_{t_a}, g_{t_b}] \le c\alpha_s(k_1 \vee k_2)^{\delta'} = O(1/K)$. Thus

$$\text{var}[T_{\boldsymbol{n}1}] = O(K/N + 1/K). \tag{28}$$

By Lemma 5,

$$E[T^2_{\boldsymbol{n}2}] = s^2 t^{-4}_{\boldsymbol{n}} E[(\sum_{\boldsymbol{i}\in\mathcal{S}_{\boldsymbol{n}}} \beta_{\boldsymbol{i}}(X_{\boldsymbol{i}} - E[X_{\boldsymbol{i}}])^4] = O(K^2/N^2), \tag{29}$$

where $|\beta_{\boldsymbol{i}}| \leq 1$ for all $\boldsymbol{i}$, and by assumption A2(ii),

$$T_{\boldsymbol{n}3}^2 = O(\psi_{\boldsymbol{n}}^2). \tag{30}$$

Moreover,

$$\begin{aligned} E[T_{\boldsymbol{n}4}^2] &= \frac{4s^2}{t_{\boldsymbol{n}}^2}{\sum}' \text{cov}[\bar{X}_{t_a}, \bar{X}_{t_b}](E[\bar{X}_{t_a}] - E[\bar{X}'])(E[\bar{X}_{t_b}] - E[\bar{X}']) \\ &\quad + \frac{4s^2}{t_{\boldsymbol{n}}^2}{\sum}'' \text{cov}[\bar{X}_{t_a}, \bar{X}_{t_b}](E[\bar{X}_{t_a}] - E[\bar{X}'])(E[\bar{X}_{t_b}] - E[\bar{X}']) \\ &\triangleq V_{\boldsymbol{n}1} + V_{\boldsymbol{n}2} \end{aligned} \tag{31}$$

Let $\rho_{ab} = \rho(\mathcal{S}_{\boldsymbol{k},t_a}, \mathcal{S}_{\boldsymbol{k},t_b})$ and $k = k_1 \vee k_2$. By Lemma 5, $|\text{cov}[\bar{X}_{t_a}, \bar{X}_{t_b}]| \leq (\text{var}[\bar{X}_{t_a}]\text{var}[\bar{X}_{t_b}])^{1/2} \leq c/s$, and, by the Cauchy-Schwarz inequality,

$$\begin{aligned} |V_{\boldsymbol{n}1}| &\leq \frac{4cs}{t_{\boldsymbol{n}}^2} \sum_{t_a,t_b=1}^{t_{\boldsymbol{n}}} |(E[\bar{X}_{t_a}] - E[\bar{X}'])(E[\bar{X}_{t_b}] - E[\bar{X}'])| I_{\{\rho_{ab} \leq k\}} \\ &\leq 4ct_{\boldsymbol{n}}^{-1} T_{\boldsymbol{n}3} ({\sum}' 1)^{1/2} = O((K/N)^{1/2}\psi_{\boldsymbol{n}}) = O(K/N + \psi_{\boldsymbol{n}}^2), \end{aligned} \tag{32}$$

since there are at most $16k^2 t_{\boldsymbol{n}}$ summands in $\sum'$. By Lemma 2(a), $|\text{cov}[\bar{X}_{t_a}, \bar{X}_{t_b}]| \leq cs^{-1}\alpha_s(\rho_{ab})^{\delta/(2+\delta)}$, and, by the Cauchy-Schwarz inequality, A2(ii), and A2(iii),

$$\begin{aligned} |V_{\boldsymbol{n}2}| &\leq \frac{4cs\alpha_s(k)^{\delta/(2+\delta)}}{t_{\boldsymbol{n}}^2} \sum_{t_a,t_b=1}^{t_{\boldsymbol{n}}} |(E[\bar{X}_{t_a}] - E[\bar{X}'])(E[\bar{X}_{t_b}] - E[\bar{X}'])| \\ &\leq 4c\alpha_s(k)^{\delta/(2+\delta)} T_{\boldsymbol{n}3} = O(\psi_{\boldsymbol{n}}/K) = O(1/K). \end{aligned} \tag{33}$$

The desired result follows from (26-33). □

Proof of Corollary 4. Consider (15-16) and (21). From Lemma 1, Lemma 5, and Corollary 3,

$$E[(R_{\boldsymbol{n}1a} - \gamma_{\boldsymbol{n}})^2] = O(K/N + 1/K). \tag{34}$$

The proof of (34) is essentially the same as the corresponding part of the proof of Theorem 2 i E&S (only a slight change of notation is needed and the references to Lemma 4 and Theorem 1 in E&S should be changed to Lemma 5 and Corollary 3 in the current paper). By (18-19), A2(iii), and (20),

$$E[R_{\boldsymbol{n}1b}^2] = O(K/N + 1/K) \quad \text{and} \quad E[R_{\boldsymbol{n}2}^2] = O(K^2/N^2). \tag{35}$$

By (22-23), A2(ii), and assumption (3),

$$R^2_{\boldsymbol{n}3a} = O(\psi^2_{\boldsymbol{n}}) \quad \text{and} \quad R^2_{\boldsymbol{n}3b} = O(\phi^2_{\boldsymbol{n}}). \tag{36}$$

Finally, consider $R_{\boldsymbol{n}4}$. We have,

$$\begin{aligned} E[R^2_{\boldsymbol{n}4}] &= \frac{4s^2}{(t'_{\boldsymbol{n}})^2} E[\{\sum_{t=1}^{t'_{\boldsymbol{n}}} (Z_{Y,t} - E[Z_{Y,t}])(E[Z_{Y,t}] - E[\bar{Z}_Y] + Z_{\mu,t} - \bar{Z}_\mu)\}^2] \\ &\leq \frac{2s^2}{(t'_{\boldsymbol{n}})^2} E[\{\sum_{t=1}^{t'_{\boldsymbol{n}}} (Z_{Y,t} - E[Z_{Y,t}])(E[Z_{Y,t}] - E[\bar{Z}_Y])\}^2] \\ &\quad + \frac{2s^2}{(t'_{\boldsymbol{n}})^2} E[\{\sum_{t=1}^{t'_{\boldsymbol{n}}} (Z_{Y,t} - E[Z_{Y,t}])(Z_{\mu,t} - \bar{Z}_\mu)\}^2] \\ &\triangleq W_{\boldsymbol{n}1} + W_{\boldsymbol{n}2}. \end{aligned} \tag{37}$$

Let $\delta' = \delta/(2+\delta)$ and let k_*, K_*, and $\sum'$ be defined as in the proof of Theorem 4. Then, by reasoning as in (31-33),

$$\begin{aligned} W_{\boldsymbol{n}1} &= \frac{2s^2}{(t'_{\boldsymbol{n}})^2} \sum_{t_a,t_b=1}^{t'_{\boldsymbol{n}}} (E[Z_{Y,t_a}] - E[\bar{Z}_Y])(E[Z_{Y,t_b}] - E[\bar{Z}_Y])\text{cov}[Z_{Y,t_a}, Z_{Y,t_b}] \\ &\leq \Big(\frac{2cs}{(t'_{\boldsymbol{n}})^2}\sum{}' + \frac{2cs\alpha_{K_*}(k_*)^{\delta'}}{(t'_{\boldsymbol{n}})^2}\sum{}''\Big)|(E[Z_{Y,t_a}] - E[\bar{Z}_Y])(E[Z_{Y,t_b}] - E[\bar{Z}_Y])| \\ &\leq 2cR_{\boldsymbol{n}3a}((t'_{\boldsymbol{n}})^{-1}(\sum{}' 1)^{1/2} + \alpha_{K_*}(k_*)^{\delta'}) = O(K/N + 1/K + \psi^2_{\boldsymbol{n}}). \end{aligned} \tag{38}$$

Likewise,

$$W_{\boldsymbol{n}2} \leq 2cR_{\boldsymbol{n}3b}((t'_{\boldsymbol{n}})^{-1}(\sum{}' 1)^{1/2} + \alpha_{K_*}(k_*)^{\delta'}) = O(K/N + 1/K + \phi^2_{\boldsymbol{n}}), \tag{39}$$

The desired result follows from (15-16), (21), and (34-39). □

References

Belyaev, Yu.K. (1996). Central limit resampling theorems for m-dependent heterogeneous random variables. Research Report 1996-5, Dept. Mathematical Statistics, Umeå University, Sweden.

Carlstein, E. (1986). The use of subseries values for estimating the variance of a general statistic from a stationary sequence. *Ann. Statist.* **14** 1171-1179.

Cressie, N. (1993). *Statistics for Spatial Data*, 2nd ed. Wiley, New York.

Doukhan, P. (1994). *Mixing: Properties and Examples.* Springer, Berlin.

Ekström, M. (2002). Non-Parametric Estimation of the Variance of Sample Means Based on Non-Stationary Spatial Data. *Comm. Statist. Theory Methods* **31** 1743-1775.

Ekström, M. and Sjöstedt-de Luna, S. (2004a). Subsampling methods to estimate the variance of sample means based on nonstationary spatial data with varying expected values. *J. Amer. Statist. Assoc.* **99** 82-95.

Ekström, M. and Sjöstedt-de Luna, S. (2004b). Resampling methods to estimate the distribution and the variance of functions of sample means based on nonstationary spatial data. In *Programme, Abstracts and Directory of Participants*, 6th World Congress of the Bernoulli Society for Mathematical Statistics and Probability and 67th Annual Meeting of the Institute of Mathematical Statistics, Barcelona, pp. 94-95.

Fazekas, I., Kukush, A.G. and Tomacs, T. (2000). On the Rosenthal inequality for mixing fields. *Ukrainian Math. J.* **52** 305-318.

Fukuchi, J. (1999). Subsampling and model selection in time series analysis. *Biometrika* **86** 591-604.

Guyon, X. (1995). *Random Fields on a Network: Modeling, Statistics, and Applications.* Springer, Berlin.

Hall, P., Horowitz, J.L. and Jing, B.-Y. (1995). On blocking rules for the bootstrap with dependent data. *Biometrika* **82** 561-574.

Künsch, H. (1989). The jackknife and bootstrap for general stationary observations. *Ann. Statist.* **17** 1217-1241.

Lahiri, S.N. (2003). *Resampling Methods for Dependent data.* Springer, New York.

Lee, Y.D. and Lahiri, S.N. (2002). Least squares variogram fitting by spatial subsampling. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **64** 837-854.

Lin, Z.Y. and Lu, C.R. (1996). *Limit Theory for Mixing Dependent Random Variables.* Kluwer, Beijing.

Politis, D.N. and Romano, J.P. (1993). Nonparametric resampling for homogeneous strong mixing random fields. *J. Multivariate Anal.* **47** 301-328.

Politis, D.N., Romano, J.P. and Wolf, M. (1999). *Subsampling.* Springer, New York.

Possolo, A. (1991). Subsampling a Random Field. In *Spatial Statistics and Imaging*, IMS Lecture Notes Monograph Series, Vol. 20, ed. A. Possolo, Hayward, CA: Institute of Mathematical Statistics, pp. 286-294.

Sherman, M. (1996). Variance Estimation for Statistics Computed From Spatial Lattice Data. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **58** 509-523.

Sherman, M. and Carlstein, E. (1994). Nonparametric Estimation of the Moments of a General Statistic Computed From Spatial Data. *J. Amer. Statist. Assoc.* **89** 496-500.

Shiryaev, A.N. (1984). *Probability.* Springer, New York.