JOURNAL OF TIME SERIES ANALYSIS
*J. Time. Ser. Anal.* **36**: 315–326 (2015)
Published online 28 November 2014 in Wiley Online Library
(wileyonlinelibrary.com) DOI: 10.1111/jtsa.12109

SPECIAL ISSUE

# THE DEPENDENT RANDOM WEIGHTING

SRIJAN SENGUPTA[a,†] XIAOFENG SHAO[a,‡] AND YINGCHUAN WANG[b]

[a] *Department of Statistics, University of Illinois at Urbana-Champaign, Champaign, IL, USA*
[b] *Department of Statistics, University of Michigan, Ann Arbor, MI, USA*

We propose a new resampling method, the dependent random weighting, for both time series and random fields. The method is a generalization of the traditional random weighting in that the weights are made to be temporally or spatially dependent and are adaptive to the configuration of the data. Unlike the block-based bootstrap or subsampling methods, the dependent random weighting can be used for irregularly spaced time series and spatial data without any implementational difficulty. Consistency of the distribution approximation is shown for both equally and unequally spaced time series. Simulation studies illustrate the finite sample performance of the dependent random weighting in comparison with the existing counterparts for both one-dimensional and two-dimensional irregularly spaced data.

*Received 22 January 2014; Revised 26 July 2014; Accepted 22 October 2014*

Keywords: Block bootstrap; irregularly spaced; spatial data; time series.

## 1. INTRODUCTION

Resampling methodology for dependent data such as time series and spatial data have undergone rapid developments since Künsch (1989) and Liu and Singh (1992) introduced the moving block bootstrap independently. The block-based bootstrap and subsampling methods [Politis and Romano (1994)] have been proved to be very useful non-parametric resampling techniques in the inference of regularly spaced time series and spatial data. The block-based resampling/subsampling methodology, although still applicable and theoretically justified to irregularly spaced time series and spatial data, is practically inconvenient to use. Here, we mention Hall (1985), Politis and Romano (1993), Sherman and Carlstein (1994), Sherman (1996), Garcia-Soidan and Hall (1997), Politis *et al.* (1998), Lahiri (1999), Lahiri *et al.* (1999), Politis and Sherman (2001), and Nordman and Lahiri (2004) among others for important work along this line. For time series data, the irregularity can occur if there are missing values for an equally spaced time series or the time points at which the observation are taken are generated from a one-dimensional point process. In the spatial setting, the irregularly spaced data, which can be in the form of lattice data with an irregular shape of the sampling region or non-lattice data with spatial locations generated from a spatial point process, are quite common. For irregularly spaced data, the main difficulty associated with the block-based resampling/subsampling approach is that the partition of sampling region into complete and incomplete blocks requires careful programming efforts and it depends on temporal/spatial configuration to a large extent. This makes the use of block-based methods less automatic, so it would be desirable to develop alternative methods whose implementation does not depend on the irregular temporal/spatial configuration. Recently, Shao (2010) proposed the dependent wild bootstrap (DWB, hereafter) for stationary and weakly dependent time series, which has no implementational difficulty when applied to irregularly spaced time series. However, the applicability of the DWB is limited to the smooth function model, and it cannot be used to approximate the sampling distribu-

* Correspondence to: Srijan Sengupta, Department of Statistics, University of Illinois at Urbana-Champaign, Champaign, IL, USA. E-mail: ssengpt2@illinois.edu
‡Xiaofeng Shao's research was partially supported by NSF under grant no. DMS08-04937 and DMS11-04545.

Copyright © 2014 Wiley Publishing Ltd

tion and variance of some other quantities, such as sample median. In this article, we propose a new resampling method, called the dependent random weighting (DRW, hereafter), which has wider applicability than the DWB and possesses considerable implementational advantage than the block-based bootstrap and subsampling methods for irregularly spaced dependent data. The random weighting method [Zheng (1987)] has been well studied for i.i.d. data and for linear models; see Shao and Tu (1995), Chapter 10, for a detailed introduction. Instead of generating resamples from the data, the random weighting method assigns a random weight to each observation. Random weighting can be regarded as an extension of the Bayesian bootstrap [Rubin (1981)] and a smoothing of Efron's bootstrap. Often, the weights can be written as

$$w_i = \frac{Z_i}{\sum_{i=1}^n Z_i},\ i = 1, \cdots, n, \tag{1}$$

where $Z_i$ are non-negative i.i.d. random variables. So far, it seems that the methodological and theoretical developments are confined to the independent data setting. For dependent data, such as time series and spatial data, the original random weighting method, which typically allows the weights to be exchangeable, does not work in general. To capture the dependence in the data, we extend the traditional random weighting to the time series/spatial setting by allowing the $Z_i$ involved in the random weighting method to be dependent, so it is capable of mimicking the dependence in the original series. Section 2 describes the DRW and demonstrates the distribution consistency of the DRW estimator for regular and irregular spaced time series. Section 3 reports results from simulation studies for irregular time series data (one-dimensional) and spatial data (two-dimensional). Section 4 concludes, and technical details are gathered in technical appendix.

## 2. DRW FOR TIME SERIES

We shall first provide a description of the DRW in the time series context. Suppose that we have a stationary $p$-dimensional time series $(X_t)_{t\in\mathbb{Z}}$ and the parameter of interest is $\theta = T(F)$, where $T$ is a given functional and $F$ is the marginal distribution function of $X_t \in \mathbb{R}^p$. Examples include the mean, marginal variance and quantiles of $X_t$. The estimator of $\theta$ is $\hat{\theta}_n = T(F_n)$, where $F_n$ is the empirical distribution function based on the observations $\{X_{t_j}\}_{j=1}^n$ and $\{t_j\}_{j=1}^n$ are the time points at which the data are observed. In the equally spaced case, $t_j = j$. The randomly weighted empirical distribution function $F_n^*$ is defined as $F_n^*(x) = \sum_{i=1}^n w(t_i)\mathbf{1}(X_{t_i} \le x)$, where $\{w(t_i)\}_{i=1}^n$ are the random weights. We assume that the weights take the form of (1), in particular

$$w(t_i) = \frac{Z(t_i)}{\sum_{i=1}^n Z(t_i)},$$

where $\{Z(t_i)\}$ are a realization from a non-negative continuous-time process $Z(t), t \in \mathbb{R}$.

**Assumption 1.** The random variables $\{Z(t_j)\}_{j=1}^n$ are independent of the data and are a realization of a stationary process with $cov(Z(t_j), Z(t_{j'})) = a\{(t_j - t_{j'})/l\}$, where $a : \mathbb{R} \to [0, 1]$ is continuous, symmetric and has compact support on $[-1, 1]$. Further, assume that $\{Z(t)\}$ is $l$-dependent.

Several commonly used windows (kernels) in spectral analysis, such as Bartlett, Parzen and Tukey-Hanning windows, satisfy Assumption 1 on $a(\cdot)$. The bandwidth parameter $l$ plays a similar role as that in the DWB or the block size in the moving block bootstrap.

Let $\hat{\theta}_n^* = T(F_n^*)$. Then we can approximate the sampling distribution or variance of $\sqrt{n}(\hat{\theta}_n - \theta)$ by the conditional distribution or conditional variance of $\sqrt{n}(\hat{\theta}_n^* - \hat{\theta}_n)S_Z$ given the data, where $S_Z = \mathbb{E}(Z(1))/\sqrt{\mathrm{var}\,(Z(1))}$ is a scaling factor. It is worth noting that the scaling constant $S_Z$ also comes up in the original random weighting method and it is in fact possible to select the distribution of $(Z(t))$ so that $S_Z = 1$, as demonstrated in the following example.

 Copyright © 2014 Wiley Publishing Ltd *J. Time. Ser. Anal.* **36**: 315–326 (2015)
DOI: 10.1111/jtsa.12109

**Example 1.** In the equally spaced case, let $Z_t = (Y_t + c)^2$, where $\{Y_t\}_{t=1}^n \sim N(0, \Sigma)$, where $\Sigma$ is an $n \times n$ matrix with $(i, j)$th entry defined as $\Sigma(i, j) = W((i-j)/l_n)$, where $W$ is a symmetric kernel function. Assuming that $W(0) = 1$, then $\mathbb{E}(Z_1) = \mathbb{E}\left(Y_1^2\right) + c^2 = c^2 + 1$ and $\operatorname{var}(Z_1) = \mathbb{E}(Y_1 + c)^4 - (\mathbb{E}(Z_1))^2 = \mathbb{E}\left(Y_1^4\right) + 6c^2\mathbb{E}\left(Y_1^2\right) + c^4 - (c^2 + 1)^2 = 4c^2 + 2$. Setting $S_Z = 1$, we obtain $4c^2 + 2 = \left(c^2 + 1\right)^2$, which yields $c^2 = (1+\sqrt{2})$. Note that in this case, $\operatorname{cov}(Z_t, Z_{t'}) = 2W^2((t-t')/l_n) + 4c^2W((t-t')/l_n)$. The same argument applies to the unequally spaced case (Section 3).

### 2.1. Equally Spaced Time Series

We shall first study the asymptotic properties of the DRW estimator when the time series is evenly spaced; that is, $t_j = j$. Following Shao (2010), we focus on the framework of the smooth function model, which contains a large class of quantities of interest in time series analysis. Let $\theta = H(\mu)$ where $\mu = \mathbb{E}(X_t)$ and $H : \mathbb{R}^p \to \mathbb{R}$ is a smooth function. Given observations $(X_t)_{t=1}^n$, the estimator is $\hat{\theta}_n = H(\hat{\mu}_n)$, where $\hat{\mu}_n = \bar{\mathbf{X}}_n = n^{-1}\sum_{t=1}^n X_t$. The DRW counterpart of $\hat{\theta}_n$ is $\hat{\theta}^*_{n,DRW} = H\left(\hat{\mu}^*_{n,DRW}\right)$, where $\hat{\mu}^*_{n,DRW} = \sum_{t=1}^n w_t X_t$. Let $\sigma_n^2 = n \operatorname{var}(\hat{\theta}_n)$ and $\boldsymbol{\nabla}(\mathbf{x}) = \{\partial H(\mathbf{x})/\partial x_1, \partial H(\mathbf{x})/\partial x_2, \ldots, \partial H(\mathbf{x})/\partial x_p\}'$ be the vector of first-order partial derivatives of $H$ at $\mathbf{x}$. Denote by $\boldsymbol{\nabla} = \boldsymbol{\nabla}(\boldsymbol{\mu})$ and $\boldsymbol{\Sigma}_\infty = \sum_{k=-\infty}^{\infty} \operatorname{cov}(\mathbf{X}_0, \mathbf{X}_k)$. Under some suitable conditions, we have $\sqrt{n}(\hat{\theta}_n - \theta_0) \to_D N\left(0, \tau_\infty^2\right)$, where $\tau_\infty^2 = \boldsymbol{\nabla}'\boldsymbol{\Sigma}_\infty\boldsymbol{\nabla} > 0$.

Denote by $\alpha(k)$ strong mixing coefficients of the process $\mathbf{X}_t$ and by $X_{t,i}$ the $i$th component of $\mathbf{X}_t$. Assumptions 2 and 3 are needed to state the consistency of the DWB in distribution approximation.

**Assumption 2.** Assume that there exists a $\delta \geq 2$ such that $\sum_{j=1}^\infty \alpha(j)^{\delta/(2+\delta)} < \infty$ and $\mathbb{E}\|\mathbf{X}_1\|^{2+\delta} < \infty$. Also suppose that $\boldsymbol{\Sigma}_\infty$ is non-singular.

**Assumption 3.** For any $(i_1, i_2, i_3, i_4) \in \{1, 2, \cdots, p\}^4$, we have

$$\sum_{t_1,t_2,t_3=-\infty}^{\infty} |\operatorname{cum}(X_{0,i_1}, X_{t_1,i_2}, X_{t_2,i_3}, X_{t_3,i_4})| < \infty.$$

See Section 3 in Shao (2010) for the discussion of the previous assumptions.

**Theorem 1.** Assume that the function $H$ is differentiable in a neighbourhood of $\boldsymbol{\mu}$, that is, $N_H = \{\mathbf{x} \in \mathbb{R}^p : \|\mathbf{x} - \boldsymbol{\mu}\| \leq \epsilon\}$ for some $\epsilon > 0$, $\sum_{|\boldsymbol{\alpha}|=1} |D^{\boldsymbol{\alpha}} H(\boldsymbol{\mu})| \neq 0$, and the first partial derivatives of $H$ satisfy a Lipschitz condition of order $s > 0$ on $N_H$. Suppose that Assumptions 1, 2, 3 and $l^{-1} + l/n^{\delta/(2+2\delta)} = o(1)$ hold. Further, assume that $Z_t \in \mathcal{L}^{2+\delta}$ for $\delta \geq 2$ (i.e. $E[Z_t^{2+\delta}] < \infty$). Then

$$\sup_{x\in\mathbb{R}} \left|P\left[\sqrt{n}\{H\left(\bar{\mathbf{X}}_n\right) - H(\boldsymbol{\mu})\} \leq x\right] - P^*\left[\sqrt{n}\{H\left(\bar{\mathbf{X}}^*_{n,DRW}\right) - H\left(\bar{\mathbf{X}}_n\right)\}S_Z \leq x\right]\right| = o_p(1).$$

**Remark 1.** The smooth function model framework covers several important parameters and their estimators; for example, autocovariances, autocorrelations and autoregressive coefficients. The median and other quantiles, however, do not fall in the class of smooth function models. For more details on the general class of estimators covered under the smooth function model, we refer the interested reader to Chapter 4 of Lahiri (2003b), specifically the examples of that chapter.

In general, for approximately linear statistic $T(F_n)$, we can expand $T(F_n)$ around $T(F)$ as $T(F_n) = T(F) + n^{-1}\sum_{t=1}^n IF(X_t; F) + R_n$, where $IF(X_t; F)$ is the influence function and $R_n$ is the remainder term. Similarly, we have $T\left(F_n^*\right) = T(F) + n^{-1}\sum_{t=1}^n w_t IF(X_t; F) + R_n^*$. To show the consistency of $\sqrt{n}\left(T\left(F_n^*\right) - T(F_n)\right)$ as an estimator of $\sqrt{n}(T(F_n) - T(F))$ in terms of distribution approximation, a typical strategy is to find appropriate

 Copyright © 2014 Wiley Publishing Ltd 
DOI: 10.1111/jtsa.12109

regularity conditions on $T$ and the weak dependence of $X_t$ to guarantee the asymptotic negligibility of $\sqrt{n}R_n$ and $\sqrt{n}R_n^*$ (conditional on the data); the latter of which may require non-trivial details.

**Remark 2.** We encounter some technical difficulty in establishing the consistency for the DRW variance estimator. In particular, it is difficult to obtain a sharp rate for $E\{(1+U_n)^{-1}-1\}^2$ ($U_n$ defined in the proof of Theorem 1; see Technical Appendix), which seems necessary to show the variance consistency. The random variable is of the form $\frac{X^2}{(1+X)^2}$ that is hard to control when $X$ is close to $-1$. In our approach, we tried using the power series expansion for $\frac{x}{1+x}$, but that results in a series involving higher moments of the underlying variable, and hence, controlling such a quantity would require putting bounds on these higher moments. Such restrictions on higher-order moments seem to require stronger assumptions than the standard set of assumptions usually found in the literature. However, from our simulation results in Section 3, it appears that variance consistency holds for the model used in our simulations.

**Remark 3.** The DRW is closely related to the DWB, which, in a sense, is also a random weighting method [see Section 3 of Shao (2010)]. However, the weights of the DWB can be negative, and the corresponding bootstrapped measure is not a valid probability measure, which limits its applicability. By contrast, the DRW corresponds to a proper probability measure conditional on the data, so it has wider applicability than the DWB. In particular, it can be used to approximate the sampling distribution of sample median and empirical processes for which the DWB is not applicable. The DRW can also be regarded as an extension of the extended tapered block bootstrap [Shao (2009)], where the tapering is applied to the bootstrapped empirical measure corresponding to the moving block bootstrap. However, the extended tapered block bootstrap is still block-based, and it encounters implementational difficulty when applied to irregularly spaced data.

### 2.2. Irregularly Spaced Time Series

To allow for irregularly spaced time points, we shall use the theoretical framework in Section 5 of Shao (2010) to study the asymptotic properties of the DRW estimator. In particular, we assume a stochastic sampling design, which was used by Lahiri (2003), Lahiri and Mukherjee (2004) and Lahiri and Zhu (2006) in the study of spatial block bootstrap for irregularly spaced spatial data. Assume that $t_j = \lambda_n v_j$, $j = 1, \ldots, n$, where $v_j$ takes values in $\mathbb{R}_0$ ($\mathbb{R}_0$ is a Borel subset of $(-1/2, 1/2]$, which is the prototype sampling region) and $\{v_j\}_{j=1}^n$ are a realization of the i.i.d. random variables $V_1, \ldots, V_n$. This formulation allows a non-uniform design across the region, and the expected number of points in two regions of the same size could be different. Assume that there is an underlying one-dimensional continuous-time stationary process $\{X(t)\}$. Given the observations $\{X(t_j)\}_{j=1}^n$, our interest is in the inference of the mean. Let $\gamma(v_1) = \operatorname{cov}(X(0), X(v_1))$, $C_4(v_1, v_2, v_3) = \operatorname{cum}\{X(0), X(v_1), X(v_2), X(v_3)\}$ denote the autocovariance and the fourth-order cumulant for $v_1, v_2, v_3 \in \mathbb{R}$. Let $\mu = \mathbb{E}(X(t))$ and $\bar{X}_n = n^{-1}\sum_{j=1}^n X(t_j)$. To estimate the distribution and the variance of $\sqrt{n}\left(\bar{X}_n - \mu\right)$, we note that the DRW counterpart of $\sqrt{n}\left(\bar{X}_n - \mu\right)$ is $\sqrt{n}\left(\bar{X}^*_{n,DRW} - \bar{X}_n\right) S_Z$, where $\bar{X}^*_{n,DRW} = \sum_{j=1}^n w(t_j) X(t_j)$. Without loss of generality, we assume that $\{V_n\}_{n\geq 1}$, $\{X(t), t \in \mathbb{R}\}$ and the bootstrap variables $\{Z(t_j), t \in \mathbb{R}\}$ are all defined on a common probability space $(\Omega, \mathcal{F}, P)$. Let $P_V$ denote the joint probability distribution of the sequence of i.i.d. random variables $V_1, V_2, \ldots, V_n$ with density $\eta(v)$, $v \in R_0$. We shall use $\mathbb{E}_V$ ($\operatorname{var}_V$) to denote the expectation (variance) with respect to the joint distribution $P_V$; use $\mathbb{E}_{X|V}$ ($\operatorname{var}_{X|V}$) to denote the conditional expectation (variance) with respect to $P_X$ (i.e. the joint probability distribution for $\{X(t), t \in \mathbb{R}\}$) given $\{V_n\}_{n\geq 1}$. Following Shao (2010), we assume Assumptions 4 and 5 on the sampling region $R_0$ and sampling density $\eta(\cdot)$.

**Assumption 4.** Define $R_0$ to be a Borel subset of $(-1/2, 1/2]$ containing an open neighbourhood of the origin such that, for any sequence of positive real numbers $a_n \to 0$, the number of cubes of the scaled lattice $a_n\mathbb{Z}$ that intersect $R_0$ and $R_0^c$ is $O(1)$ as $n \to \infty$.

 Copyright © 2014 Wiley Publishing Ltd 
DOI: 10.1111/jtsa.12109

**Assumption 5.** The probability distribution function $\eta(x)$ is continuous, everywhere positive with support $\bar{R}_0$ and $\int_{s\in R_0}\eta(s)ds = 1$.

Denote by $\iota = \int_{s\in R_0}\eta^2(s)ds$. Lahiri (2003) showed that depending on the magnitude of $\kappa := \lim_{n\to\infty} n/\lambda_n$, this formulation accommodates both pure-increasing-domain asymptotics (i.e. $\kappa < \infty$) and mixed-increasing-domain asymptotics (i.e. $\kappa = +\infty$). Let $\xi_n = \text{var}\left(\bar{X}_n\right)$. Lemma 5.2 of Lahiri (2003) implies that under appropriate conditions, we have that (i) if $\kappa \in (0,\infty)$, then $n\xi_n \to \gamma(0) + \kappa\iota\int_{\mathbb{R}}\gamma(s)ds$, a.s. $(P_V)$ and (ii) if $\kappa = \infty$, then $\lambda_n\xi_n \to \iota\int_{\mathbb{R}}\gamma(s)ds$, a.s. $(P_V)$. Here, a.s. $(P_V)$ means that the result holds with probability one under $P_V$; that is, for almost all realizations of the sequence $\{V_n\}_{n\geq 1}$. In the work of Lahiri (2003), the distribution of $\bar{X}_n$ is regarded as conditional distribution given $\{V_n\}_{n\geq 1}$, and $\xi_n$ is regarded as a function of the randomly sampled locations. Whereas in our treatment, we view $\xi_n$ as an unknown quantity, where the randomness due to $\{V_n\}_{n\geq 1}$ has been removed by the expectation. Theorem 2 states the distribution consistency of the DRW estimator.

**Theorem 2.** Suppose that Assumptions 1, 4 and 5 hold. Assume that $l_n/\sqrt{n} + l_n/\lambda_n = o(1)$,

$$\int_{\mathbb{R}} |\gamma(v)|dv < \infty, \tag{2}$$

$$\text{and} \int_{\mathbb{R}^3} |C_4(v_1, v_2, v_3)|dv_1dv_2dv_3 < \infty. \tag{3}$$

Further, assume that $Z(t) \in \mathcal{L}^4$. We have that (i) if $\kappa \in (0,\infty)$, then

$$\sup_{x\in\mathbb{R}} |P[\sqrt{n}\left(\bar{X}_n - \mu\right) \leq x] - P^*[\sqrt{n}\left(\bar{X}^*_{n,DRW} - \bar{X}_n\right) S_Z \leq x]| = o_p(1),$$

and that (ii) if $\kappa = \infty$, then

$$\sup_{x\in\mathbb{R}} |P[\sqrt{\lambda_n}\left(\bar{X}_n - \mu\right) \leq x] - P^*[\sqrt{\lambda_n}\left(\bar{X}^*_{n,DRW} - \bar{X}_n\right) S_Z \leq x]| = o_p(1).$$

**Remark 4.** Lahiri and Zhu (2006) showed that a naive application of block bootstrap, called data-site-shifted block bootstrap (DSSBB), is not suitable for irregularly spaced data when the spatial sampling density is non-uniform. The DRW does not suffer from the same problem. To quote from their paper,

> The failure of the DSSBB method seems to be an artifact of the interaction between the nonuniform design density and of the additional randomness in the data-site-shifted blocks induced by the random locations of the sampling sites.

In the case of DRW, resampling takes place by assigning random weights to the data points without shifting their locations. These random weights are independent of the data and spatially correlated to reflect the dependence in the data. Therefore, our resampling scheme is free from the interaction alluded to by Lahiri and Zhu (2006) in their explanation of why the DSSBB fails.

## 3. SIMULATION RESULTS

Here, we investigate the finite sample performance of DRW and its competitors for irregular time series and spatial data under the framework of stochastic sampling design. Let $R_0 = (-1/2, 1/2]$, sample size $n = 200$ and $\lambda_n = 18$ or 36. The time points $\{t_j\}_{j=1}^n$ are generated by taking i.i.d. draws from truncated N(0,1) density


Copyright © 2014 Wiley Publishing Ltd

DOI: 10.1111/jtsa.12109

Table I. Top panel: the normalized MSEs for the bootstrap variance estimators of $nvar\,[median\,(x_1, \cdots, x_n)]$ using (a) the grid-based block bootstrap and (b) the dependent random weighting; the box for each row indicates the smallest normalized MSE among $l_n = 1, \ldots, 10$

| | | | $l$ | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\lambda_n$ | $\rho$ | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| 18 | 0.5 | (a) | 0.57 | 0.48 | $\boxed{0.47}$ | 0.50 | 0.52 | 0.54 | 0.57 | 0.62 | 0.73 | 0.72 |
| | | (b) | 0.59 | 0.48 | $\boxed{0.45}$ | 0.46 | 0.47 | 0.49 | 0.50 | 0.51 | 0.52 | 0.52 |
| | 1 | (a) | 0.38 | $\boxed{0.34}$ | 0.38 | 0.43 | 0.49 | 0.50 | 0.52 | 0.57 | 0.83 | 0.75 |
| | | (b) | 0.39 | $\boxed{0.32}$ | 0.34 | 0.37 | 0.39 | 0.41 | 0.43 | 0.44 | 0.45 | 0.45 |
| | 2 | (a) | $\boxed{0.27}$ | 0.29 | 0.37 | 0.40 | 0.45 | 0.44 | 0.43 | 0.49 | 0.67 | 0.62 |
| | | (b) | $\boxed{0.27}$ | 0.28 | 0.31 | 0.35 | 0.37 | 0.39 | 0.41 | 0.41 | 0.42 | 0.43 |
| 36 | 0.5 | (a) | 0.51 | 0.38 | $\boxed{0.35}$ | 0.35 | 0.37 | 0.41 | 0.43 | 0.46 | 0.51 | 0.53 |
| | | (b) | 0.53 | 0.39 | 0.34 | $\boxed{0.33}$ | 0.34 | 0.35 | 0.36 | 0.37 | 0.38 | 0.40 |
| | 1 | (a) | 0.32 | $\boxed{0.25}$ | 0.27 | 0.29 | 0.33 | 0.37 | 0.39 | 0.40 | 0.45 | 0.46 |
| | | (b) | 0.33 | $\boxed{0.25}$ | 0.25 | 0.27 | 0.29 | 0.30 | 0.32 | 0.34 | 0.34 | 0.35 |
| | 2 | (a) | $\boxed{0.18}$ | 0.19 | 0.22 | 0.25 | 0.28 | 0.33 | 0.33 | 0.35 | 0.40 | 0.41 |
| | | (b) | 0.18 | $\boxed{0.17}$ | 0.20 | 0.22 | 0.24 | 0.27 | 0.27 | 0.29 | 0.30 | 0.31 |
| 18 | 0.5 | (a) | 64 | 71 | $\boxed{72}$ | 71 | 69 | 67 | 63 | 59 | 65 | 59 |
| | | (b) | 63 | 70 | $\boxed{72}$ | 72 | 71 | 71 | 70 | 68 | 68 | 67 |
| | 1 | (a) | 76 | 80 | $\boxed{81}$ | 79 | 76 | 75 | 72 | 66 | 70 | 63 |
| | | (b) | 75 | 79 | $\boxed{81}$ | 81 | 79 | 79 | 77 | 75 | 74 | 74 |
| | 2 | (a) | 85 | $\boxed{86}$ | 84 | 82 | 80 | 78 | 76 | 70 | 73 | 66 |
| | | (b) | 85 | $\boxed{86}$ | 86 | 85 | 83 | 82 | 81 | 80 | 79 | 78 |
| 36 | 0.5 | (a) | 68 | 77 | 79 | 80 | $\boxed{81}$ | 79 | 79 | 76 | 76 | 74 |
| | | (b) | 67 | 76 | 79 | 80 | $\boxed{82}$ | 82 | 82 | 81 | 80 | 80 |
| | 1 | (a) | 81 | 85 | $\boxed{88}$ | 88 | 87 | 86 | 85 | 84 | 80 | 80 |
| | | (b) | 80 | 85 | 86 | $\boxed{87}$ | 87 | 87 | 86 | 86 | 85 | 84 |
| | 2 | (a) | 88 | $\boxed{90}$ | 89 | 89 | 88 | 85 | 85 | 83 | 80 | 80 |
| | | (b) | 88 | $\boxed{90}$ | 90 | 89 | 89 | 88 | 87 | 87 | 86 | 86 |

Bottom panel: the empirical coverage (in percent) for the bootstrap-based confidence intervals of the median using (a) and (b). The box for each row indicates the best coverage among $l_n = 1, \ldots, 10$ (nominal level is 95%).

function over $R_0$ and multiplying by the scaling constant $\lambda_n$. Given the sampled time points, we then generate the observations $\{X(t_j)\}_{j=1}^n$ from a one-dimensional zero-mean Gaussian process with exponential covariance function $\gamma(z) = \exp(-\rho|z|)$, $z \in \mathbb{R}$, where $\rho = 0.5$, 1 and 2. The random weights are generated by following Example 1 and letting $S_Z = 1$ and $W$ to be the Bartlett kernel.

Table I shows the normalized mean squared error and the empirical coverages in percentage for the bootstrap approximation of the variance and distribution of the sample median based on the grid-based block bootstrap [Lahiri and Zhu (2006)] and DRW. For the variance estimator, let the true variance be $\sigma_n$, and let $\sigma_n^{(j)}$ be its estimate based on 1000 bootstrap samples for the $j$th replicate, where $j = 1, 2, \ldots, 1000$ because 1000 Monte Carlo replications are used. Then the normalized MSE is calculated as $\frac{1}{1000}\sum_{j=1}^{1000}(\frac{n\sigma_n^{(j)}}{n\sigma_n} - 1)^2$. We calculate MSE for the DRW variance estimator, even though we have not demonstrated its asymptotic consistency. It can be seen that in terms of smallest normalized MSE or best empirical coverage (boxed values), DRW typically performs at par with grid-based block bootstrap (GBBB) and sometimes marginally outperforms GBBB. For larger block sizes, DRW typically outperforms GBBB by a substantial margin. Larger $\rho$ corresponds to smaller MSE and superior coverage, which is expected because larger $\rho$ implies weaker dependence. Moreover, the MSE decreases, and coverage probability improves as $\lambda_n$ decreases. For the mean case (Table II), DRW, DWB and GBBB perform similarly in terms of best result in each row (boxed values). For larger block sizes, DRW typically outperforms GBBB

 Copyright © 2014 Wiley Publishing Ltd 
DOI: 10.1111/jtsa.12109

Table II. Top panel: the normalized MSEs for the bootstrap variance estimators of $nvar(\bar{x}_n)$ using (a) the dependent wild bootstrap, (b) the dependent random weighting and (c) the grid-based block bootstrap; the box for each row indicates the smallest normalized MSE among $l = 1, \ldots, 10$

| | | | $l$ | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\lambda_n$ | $\rho$ | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| 18 | 0.5 | (a) | 0.67 | 0.54 | 0.48 | $\boxed{0.47}$ | 0.47 | 0.48 | 0.50 | 0.52 | 0.54 | 0.55 |
| | | (b) | 0.69 | 0.56 | 0.51 | 0.49 | $\boxed{0.48}$ | 0.49 | 0.49 | 0.50 | 0.51 | 0.52 |
| | | (c) | 0.67 | 0.55 | $\boxed{0.48}$ | 0.50 | 0.50 | 0.53 | 0.57 | 0.62 | 0.67 | 0.73 |
| | 1 | (a) | 0.45 | 0.33 | $\boxed{0.32}$ | 0.34 | 0.37 | 0.41 | 0.44 | 0.46 | 0.48 | 0.50 |
| | | (b) | 0.48 | 0.36 | $\boxed{0.33}$ | 0.34 | 0.35 | 0.38 | 0.40 | 0.41 | 0.43 | 0.44 |
| | | (c) | 0.45 | 0.34 | $\boxed{0.33}$ | 0.36 | 0.40 | 0.42 | 0.47 | 0.55 | 0.62 | 0.67 |
| | 2 | (a) | 0.25 | $\boxed{0.22}$ | 0.26 | 0.30 | 0.35 | 0.38 | 0.40 | 0.43 | 0.45 | 0.47 |
| | | (b) | 0.27 | $\boxed{0.22}$ | 0.24 | 0.27 | 0.30 | 0.33 | 0.35 | 0.36 | 0.38 | 0.39 |
| | | (c) | 0.25 | $\boxed{0.22}$ | 0.27 | 0.31 | 0.36 | 0.36 | 0.39 | 0.47 | 0.56 | 0.63 |
| 36 | 0.5 | (a) | 0.62 | 0.47 | 0.39 | 0.35 | $\boxed{0.34}$ | 0.34 | 0.34 | 0.35 | 0.37 | 0.38 |
| | | (b) | 0.64 | 0.49 | 0.41 | 0.37 | 0.35 | $\boxed{0.34}$ | 0.35 | 0.35 | 0.36 | 0.37 |
| | | (c) | 0.62 | 0.46 | 0.38 | 0.36 | $\boxed{0.35}$ | 0.35 | 0.36 | 0.38 | 0.41 | 0.42 |
| | 1 | (a) | 0.39 | 0.26 | $\boxed{0.22}$ | 0.22 | 0.24 | 0.25 | 0.27 | 0.30 | 0.32 | 0.34 |
| | | (b) | 0.42 | 0.28 | 0.24 | $\boxed{0.23}$ | 0.23 | 0.25 | 0.26 | 0.27 | 0.29 | 0.30 |
| | | (c) | 0.39 | 0.26 | $\boxed{0.22}$ | 0.23 | 0.24 | 0.28 | 0.29 | 0.32 | 0.37 | 0.37 |
| | 2 | (a) | 0.19 | $\boxed{0.14}$ | 0.15 | 0.18 | 0.20 | 0.23 | 0.25 | 0.27 | 0.29 | 0.31 |
| | | (b) | 0.21 | $\boxed{0.15}$ | 0.15 | 0.17 | 0.18 | 0.20 | 0.22 | 0.23 | 0.25 | 0.26 |
| | | (c) | 0.19 | $\boxed{0.14}$ | 0.16 | 0.18 | 0.21 | 0.25 | 0.26 | 0.28 | 0.34 | 0.35 |
| 18 | 0.5 | (a) | 58 | 67 | 69 | $\boxed{70}$ | 70 | 68 | 68 | 67 | 67 | 64 |
| | | (b) | 56 | 65 | 67 | $\boxed{69}$ | 67 | 67 | 66 | 65 | 65 | 64 |
| | | (c) | 59 | 67 | $\boxed{70}$ | 69 | 68 | 67 | 63 | 59 | 54 | 48 |
| | 1 | (a) | 71 | 78 | $\boxed{79}$ | 79 | 77 | 75 | 73 | 72 | 71 | 69 |
| | | (b) | 71 | 77 | $\boxed{78}$ | 77 | 75 | 74 | 73 | 73 | 71 | 71 |
| | | (c) | 71 | $\boxed{79}$ | 79 | 77 | 75 | 74 | 71 | 64 | 58 | 50 |
| | 2 | (a) | 82 | $\boxed{85}$ | 84 | 82 | 80 | 79 | 78 | 76 | 74 | 72 |
| | | (b) | 81 | $\boxed{84}$ | 83 | 82 | 81 | 80 | 79 | 79 | 78 | 76 |
| | | (c) | 81 | $\boxed{85}$ | 82 | 81 | 78 | 77 | 75 | 69 | 61 | 55 |
| 36 | 0.5 | (a) | 63 | 72 | 76 | 78 | 79 | $\boxed{80}$ | 79 | 79 | 79 | 77 |
| | | (b) | 62 | 71 | 74 | 76 | 78 | $\boxed{78}$ | 78 | 77 | 77 | 77 |
| | | (c) | 65 | 72 | 77 | 78 | 78 | $\boxed{79}$ | 78 | 76 | 74 | 74 |
| | 1 | (a) | 77 | 84 | 85 | $\boxed{86}$ | 86 | 86 | 85 | 84 | 82 | 82 |
| | | (b) | 76 | 82 | 84 | $\boxed{86}$ | 86 | 85 | 85 | 85 | 84 | 84 |
| | | (c) | 77 | 83 | $\boxed{85}$ | 85 | 85 | 84 | 84 | 83 | 80 | 80 |
| | 2 | (a) | 85 | $\boxed{88}$ | 88 | 87 | 86 | 86 | 84 | 84 | 84 | 83 |
| | | (b) | 85 | 87 | $\boxed{88}$ | 87 | 87 | 87 | 86 | 85 | 85 | 85 |
| | | (c) | 85 | $\boxed{88}$ | 88 | 87 | 86 | 84 | 84 | 82 | 80 | 79 |

Bottom panel: the empirical coverage (in percent) for the bootstrap-based confidence intervals of the mean using (a), (b) and (c). The box for each row indicates the best coverage among $l_n = 1, \ldots, 10$ (nominal level is 95%).

by a substantial margin as before and marginally outperforms DWB. Note that the implementation of grid-based block bootstrap is rather complicated, whereas the DRW can be easily programmed and is also computationally less expensive.

Conceptually, the extension of the DRW to irregularly spaced spatial data is straightforward, but technically, it seems non-trivial and quite challenging. Here, we shall provide a description and some finite sample results for DRW applied to spatial data. Given $n$ spatial locations $\{s_i\}_{i=1}^n$, the observations are assumed to be $\{X(s_i)\}_{i=1}^n$. We shall assume that the observations are from a stationary random field in $\mathbb{R}^2$ (for the sake of simplicity) and the locations can be in a lattice with fixed spacing or irregularly spaced. Let $\theta = T(F)$ be the parameter of


Copyright © 2014 Wiley Publishing Ltd

DOI: 10.1111/jtsa.12109

interest, where $F$ is the marginal cumulative distribution function (cdf) of $X(s)$ and $T$ is a given functional. This framework includes spatial mean and quantiles. Let $F_n$ be the empirical cdf based on $\{X(s_i)\}_{i=1}^n$. Then the sampling distribution or variance of $\sqrt{n}\,(T(F_n) - T(F))$ can be approximated by the random weighting counterpart $\sqrt{n}\left(T\left(F_n^*\right) - T\left(F_n\right)\right)$, with the random weighted empirical cdf defined by

$$F_n^*(x) = \sum_{j=1}^{n} w(s_j)\,\mathbf{1}(X_j \le x), \text{ where } w(s_j) = \frac{Z(s_j)}{\sum_{j=1}^{n} Z(s_j)}.$$

Here, $\{Z(s_j)\}_{j=1}^n$ are non-negative random variables that are independent of the data and spatially correlated. In particular, we can mimic Example 1 and let $Z(s_j) = (Y(s_j) + c)^2$, where $\{Y(s_j)\}_{j=1}^n \sim N(0, \Sigma)$ and $\Sigma$ is a $n \times n$ matrix with $(i, j)$th entry defined as $\Sigma(i, j) = W((\|s_i - s_j\|)/l_n)$, where $W$ is a kernel function and $\|s\| = \sqrt{s_{(1)}^2 + s_{(2)}^2}$ for any $s = (s_{(1)}, s_{(2)}) \in \mathbb{R}^2$. Again, $c$ can be chosen such that the scaling constant $S_Z = 1$. For spatial data, both subsampling and block-based bootstrap implicitly have some requirements on the sampling design (e.g. they may not work well when the sampling design is very heterogeneous), and their implementation is quite involved in the irregularly spaced case. By contrast, the DWB does not involve block sampling but rather generate random and spatially correlated weights to the data. The irregular configuration does not really cause any difficulty in its implementation.

Following the discussion earlier, we also performed simulations for the two-dimensional case, where $\rho$ is fixed at 1 and two sample sizes, $n = 200$ and 400, are used. The normalized MSE for the variance estimator and coverage rates for bootstrap-based intervals for sample mean and sample median are shown in Tables III and IV respectively. Similar to the one-dimensional case, we observe that in terms of best result in each row (boxed

Table III. Top panel: the normalized MSEs for the bootstrap variance estimators of $nvar(median(X_1, \cdots, X_n))$ using (a) the grid-based block bootstrap and (b) the dependent random weighting; the box for each row indicates the smallest normalized MSE among $l = 1, \ldots, 10$

| | | | $l$ | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\lambda_n$ | $n$ | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| 18 | 200 | (a) | 0.27 | 0.20 | $\boxed{0.18}$ | 0.19 | 0.21 | 0.25 | 0.26 | 0.28 | 0.34 | 0.36 |
| | | (b) | 0.28 | 0.21 | $\boxed{0.17}$ | 0.17 | 0.17 | 0.18 | 0.19 | 0.20 | 0.21 | 0.22 |
| | 400 | (a) | 0.39 | 0.25 | 0.19 | $\boxed{0.18}$ | 0.20 | 0.25 | 0.27 | 0.31 | 0.39 | 0.41 |
| | | (b) | 0.41 | 0.27 | 0.20 | 0.18 | $\boxed{0.17}$ | 0.17 | 0.19 | 0.20 | 0.21 | 0.22 |
| 36 | 200 | (a) | $\boxed{0.13}$ | 0.13 | 0.14 | 0.15 | 0.16 | 0.19 | 0.20 | 0.21 | 0.24 | 0.25 |
| | | (b) | $\boxed{0.13}$ | 0.13 | 0.13 | 0.13 | 0.14 | 0.14 | 0.15 | 0.15 | 0.16 | 0.16 |
| | 400 | (a) | 0.17 | 0.13 | $\boxed{0.12}$ | 0.12 | 0.12 | 0.13 | 0.15 | 0.15 | 0.18 | 0.18 |
| | | (b) | 0.17 | 0.13 | 0.12 | $\boxed{0.11}$ | 0.11 | 0.11 | 0.12 | 0.12 | 0.13 | 0.14 |
| 18 | 200 | (a) | 81 | 84 | $\boxed{86}$ | 85 | 83 | 81 | 78 | 76 | 70 | 65 |
| | | (b) | 80 | 83 | 85 | $\boxed{86}$ | 85 | 86 | 85 | 85 | 85 | 84 |
| | 400 | (a) | 72 | 82 | $\boxed{85}$ | 84 | 83 | 81 | 78 | 74 | 68 | 66 |
| | | (b) | 72 | 79 | 83 | $\boxed{84}$ | 84 | 84 | 83 | 84 | 83 | 83 |
| 36 | 200 | (a) | 87 | $\boxed{89}$ | 89 | 89 | 88 | 88 | 87 | 87 | 85 | 85 |
| | | (b) | 88 | 88 | 88 | $\boxed{89}$ | 89 | 89 | 89 | 89 | 89 | 89 |
| | 400 | (a) | 83 | 86 | $\boxed{87}$ | 87 | 87 | 87 | 87 | 85 | 85 | 84 |
| | | (b) | 83 | 85 | 87 | 87 | 88 | $\boxed{89}$ | 89 | 89 | 88 | 88 |

Bottom panel: the empirical coverage (in percent) for the bootstrap-based confidence intervals of the median using (a) and (b). The box for each row indicates the best coverage among $l_n = 1, \cdots, 10$ (nominal level is 95%). Two-dimensional case: $n = 200, 400$, $\lambda_n = 18, 36$ and $\rho = 1$ is fixed.

 Copyright © 2014 Wiley Publishing Ltd 
DOI: 10.1111/jtsa.12109

Table IV. Top panel: the normalized MSEs for the bootstrap variance estimators of $nvar(\bar{x}_n)$ using (a) the dependent wild bootstrap, (b) the dependent random weighting and (c) the grid-based block bootstrap; the box for each row indicates the smallest normalized MSE among $l = 1, \ldots, 10$

| | | | $l$ | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\lambda_n$ | $n$ | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| 18 | 200 | (a) | 0.36 | 0.24 | 0.17 | $\boxed{0.15}$ | 0.15 | 0.16 | 0.18 | 0.19 | 0.21 | 0.23 |
| | | (b) | 0.37 | 0.25 | 0.19 | $\boxed{0.16}$ | 0.16 | 0.16 | 0.17 | 0.17 | 0.18 | 0.19 |
| | | (c) | 0.36 | 0.23 | 0.17 | $\boxed{0.16}$ | 0.17 | 0.21 | 0.23 | 0.26 | 0.35 | 0.38 |
| | 400 | (a) | 0.49 | 0.31 | 0.22 | 0.18 | $\boxed{0.17}$ | 0.18 | 0.19 | 0.21 | 0.23 | 0.25 |
| | | (b) | 0.51 | 0.33 | 0.24 | 0.20 | $\boxed{0.18}$ | 0.18 | 0.19 | 0.20 | 0.21 | 0.22 |
| | | (c) | 0.49 | 0.30 | 0.20 | $\boxed{0.19}$ | 0.19 | 0.23 | 0.25 | 0.30 | 0.38 | 0.40 |
| 36 | 200 | (a) | 0.12 | 0.080 | 0.061 | 0.055 | $\boxed{0.054}$ | 0.057 | 0.064 | 0.071 | 0.080 | 0.087 |
| | | (b) | 0.12 | 0.085 | 0.065 | 0.057 | $\boxed{0.055}$ | 0.056 | 0.060 | 0.064 | 0.071 | 0.077 |
| | | (c) | 0.11 | 0.079 | 0.061 | $\boxed{0.055}$ | 0.058 | 0.064 | 0.073 | 0.084 | 0.11 | 0.11 |
| | 400 | (a) | 0.22 | 0.14 | 0.094 | 0.073 | $\boxed{0.064}$ | 0.064 | 0.067 | 0.074 | 0.083 | 0.095 |
| | | (b) | 0.23 | 0.15 | 0.10 | 0.080 | 0.069 | $\boxed{0.066}$ | 0.067 | 0.071 | 0.077 | 0.084 |
| | | (c) | 0.22 | 0.14 | 0.092 | 0.072 | $\boxed{0.068}$ | 0.070 | 0.079 | 0.086 | 0.11 | 0.12 |
| 18 | 200 | (a) | 79 | 84 | 87 | 88 | 88 | 88 | $\boxed{89}$ | 88 | 87 | 86 |
| | | (b) | 78 | 84 | 86 | $\boxed{88}$ | 87 | 88 | 88 | 88 | 88 | 87 |
| | | (c) | 78 | 85 | $\boxed{88}$ | 87 | 86 | 83 | 80 | 77 | 71 | 67 |
| | 400 | (a) | 69 | 80 | 85 | 87 | $\boxed{88}$ | 87 | 87 | 86 | 86 | 85 |
| | | (b) | 69 | 79 | 84 | 85 | $\boxed{87}$ | 87 | 87 | 87 | 86 | 86 |
| | | (c) | 70 | 82 | $\boxed{85}$ | 85 | 84 | 81 | 78 | 74 | 70 | 66 |
| 36 | 200 | (a) | 89 | 91 | 91 | 92 | $\boxed{93}$ | 92 | 92 | 92 | 92 | 92 |
| | | (b) | 89 | 90 | $\boxed{92}$ | 92 | 92 | 92 | 92 | 92 | 92 | 92 |
| | | (c) | 89 | 91 | 91 | $\boxed{92}$ | 92 | 90 | 90 | 89 | 88 | 87 |
| | 400 | (a) | 83 | 87 | 89 | 90 | 90 | $\boxed{91}$ | 91 | 91 | 91 | 91 |
| | | (b) | 83 | 87 | 89 | 89 | $\boxed{91}$ | 90 | 91 | 91 | 91 | 91 |
| | | (c) | 83 | 87 | 89 | $\boxed{90}$ | 89 | 89 | 88 | 87 | 86 | 85 |

Bottom panel: the empirical coverage (in percent) for the bootstrap-based confidence intervals of the mean using (a), (b) and (c). The box for each row indicates the best coverage among $l_n = 1, \cdots, 10$ (nominal level is 95%). Two-dimensional case: $n = 200, 400$, $\lambda_n = 18, 36$ and $\rho = 1$ is fixed.

values), DRW and GBBB (and DWB for the mean case) perform similarly, while for larger block sizes, the DRW typically outperforms GBBB by a substantial margin and DWB by a small margin.

## 4. CONCLUSION

In this article, we proposed a new resampling method, the DRW, for time series and briefly mention its extension to spatial data. The main attraction of this new method lies in its adaptiveness to the irregularity of temporal or spatial configurations as its implementation in the irregularly spaced case is the same as regularly spaced case, unlike the block-based bootstrap or subsampling methods. Under suitable conditions, we proved its consistency in distribution approximation for both equally and unequally spaced time series. It is expected that additional theoretical results, such as consistency of bootstrapping empirical processes in both equally and unequally spaced time series (see Bühlmann (1994), Naik-Nimbalkar and Rajarshi (1994) and Peligrad (1998) among others for consistency of block-based bootstrap) and consistency in distribution approximation in the spatial case, can hold under certain regularity conditions. However, this may require a very technical analysis, and we leave this for future work. Another topic that is worthy of investigation is the optimal choice of bandwidth parameter $l_n$ for a given kernel function. Also, it is of interest to see if one can borrow the recently popular fixed-$b$ asymptotics [Kiefer and Vogelsang (2005)] and calibrate the bootstrap-based inference; see Shao and Politis (2013), for a recent attempt along this direction.

 Copyright © 2014 Wiley Publishing Ltd 
DOI: 10.1111/jtsa.12109

## REFERENCES


Bhattacharya RN, Rao RR. 1986. *Normal Approximation and Asymptotic Expansions. Krieger*. Melbourne: FL.

Bühlmann P. 1994. Blockwise bootstrapped empirical process for stationary sequences. *Annals of Statistics* **22**: 995–1012.

Garcia-Soidan PH, Hall P. 1997. On sample reuse methods for spatial data. *Biometrics* **53**: 273–281.

Hall P. 1985. Resampling a coverage pattern. *Stochastic Processes and their Applications* **20**: 231–246.

Kiefer NM, Vogelsang TJ. 2005. A new asymptotic theory for heteroskedasticity-autocorrelation robust tests. *Econometric Theory* **21**: 1130–1164.

Künsch HR. 1989. The jackknife and the bootstrap for general stationary observations. *Annals of Statistics* **17**(3): 1217–1241.

Lahiri SN. 1999. Asymptotic distribution of the empirical spatial cumulative distribution function predictor and prediction bands based on a subsampling method. *Probability Theory and Related Fields* **114**: 55–84.

Lahiri SN. 2003. Central limit theorems for weighted sums of a spatial process under a class of stochastic and fixed designs. *Sankhyā* **65**: 356–388.

Lahiri SN. 2003b. *Resampling Methods for Dependent Data*. New York: Springer.

Lahiri SN, Mukherjee K. 2004. Asymptotic distributions of M-estimators in a spatial regression model under some fixed and stochastic spatial sampling designs. *Annals of the Institute of Statistical Mathematics* **56**: 225–250.

Lahiri SN, Zhu J. 2006. Resampling methods for spatial regression models under a class of stochastic designs. *Annals of Statistics* **34**: 1774–1813.

Lahiri SN, Kaiser MS, Cressie N, Hsu N-J. 1999. Prediction of spatial cumulative distribution functions using subsampling. *Journal of the American Statistical Association* **94**: 86–110.

Liu RY, Singh K. 1992. *Moving blocks jackknife and bootstrap capture weak dependence*. In *Exploring the Limits of Bootstrap*, LePage R, Billard L (eds.), pp. 225–248.

Naik-Nimbalkar UV, Rajarshi MB. 1994. Validity of blockwise bootstrap for empirical processes with stationary observations. *Annals of Statistics* **22**: 980–994.

Nordman DJ, Lahiri SN. 2004. On optimal spatial subsample size for variance estimation. *Annals of Statistics* **32**: 1981–2027.

Peligrad M. 1998. On the blockwise bootstrap for empirical processes for stationary sequences. *Annals of Probability* **26**: 877–901.

Politis DN, Paparoditis E, Romano JP. 1998. Large sample inference for irregularly spaced dependent observations based on subsampling. *Sankhyā* **60**: 274–292.

Politis DN, Romano JP. 1993. Nonparametric resampling for homogeneous strong mixing random fields. *Journal of Multivariate Analysis* **47**: 301–328.

Politis DN, Romano JP. 1994. Large sample confidence regions based on subsamples under minimal assumptions. *Annals of Statistics* **22**: 2031–2050.

Politis DN, Sherman M. 2001. Moment estimation for statistics from marked point processes. *Journal of the Royal Statistical Society: Series B* **63**: 261–275.

Rubin DB. 1981. The Bayesian bootstrap. *Annals of Statistics* **9**: 130–134.

Shao J, Tu D. 1995. *The Jackknife and Bootstrap*. New York: Springer-Verlag.

Shao X. 2009. Extended tapered block bootstrap. *Statistica Sinica* **20**: 807–821.

Shao X. 2010. The dependent wild bootstrap. *Journal of the American Statistical Association* **105**: 218–235.

Shao X, Politis DN. 2013. Fixed-b subsampling and block bootstrap: improved confidence sets based on p-value calibration. *Journal of the Royal Statistical Society: Series B* **75**: 161–184.

Sherman M, Carlstein E. 1994. Nonparametric estimation of the moments of a general statistic computed from spatial data. *Journal of the American Statistical Association* **89**: 496–500.

Sherman M. 1996. Variance estimation for statistics computed from spatial lattice data. *Journal of the Royal Statistical Society: Series B* **58**: 509–523.

Zheng. 1987. Random weighting methods. *Acta Mathematicae Applicatae Sinica* **10**: 247–253.


## TECHNICAL APPENDIX

*Proof of Theorem 1*

Let $\Phi(\mathbf{x}; \boldsymbol{\Sigma}_\infty)$ be the distribution function of $N(\mathbf{0}, \boldsymbol{\Sigma}_\infty)$ on $\mathbb{R}^p$. We first show that

$$\sup_{\mathbf{x}\in\mathbb{R}^p} \left| P\left\{\sqrt{n}\left(\bar{\mathbf{X}}_n - \boldsymbol{\mu}\right) \le \mathbf{x}\right\} - P^*\left\{\sqrt{n}\left(\bar{\mathbf{X}}^*_{n,DRW} - \bar{\mathbf{X}}_n\right) S_Z \le \mathbf{x}\right\}\right| = o_p(1). \tag{A1}$$

 Copyright © 2014 Wiley Publishing Ltd 
DOI: 10.1111/jtsa.12109

Since $\sqrt{n}\left(\bar{\mathbf{X}}_n - \boldsymbol{\mu}\right) \to_D N\left(\mathbf{0}, \boldsymbol{\Sigma}_\infty\right)$ under Assumption 2, it follows from a multivariate version of Polyā's theorem (Bhattacharya and Rao, 1986) that $\sup_{\mathbf{x}\in\mathbb{R}^p} \left|P\{\sqrt{n}\left(\bar{\mathbf{X}}_n - \boldsymbol{\mu}\right) \le \mathbf{x}\} - \Phi\left(\mathbf{x}; \boldsymbol{\Sigma}_\infty\right)\right| = o(1)$. Then (A1) follows if we can show that

$$\sup_{\mathbf{x}\in\mathbb{R}^p} \left|P^*\left\{\sqrt{n}(\bar{\mathbf{X}}^*_{n,DRW} - \bar{\mathbf{X}}_n)S_Z \le \mathbf{x}\right\} - \Phi\left(\mathbf{x}; \boldsymbol{\Sigma}_\infty\right)\right| = o_p(1). \tag{A2}$$

To this end, we shall first establish the relation between the DRW estimator and the DWB estimator introduced by Shao (2010). For DWB,

$$T^*_{n,DWB} = \sqrt{n}\left(\bar{X}^*_{n,DWB} - \bar{X}_n\right) = \frac{1}{\sqrt{n}}\sum_{t=1}^{n}\left(X_t - \bar{X}_n\right)\delta_t,$$

where $\delta_t$ is independent of $\mathcal{X}_n$, $E\left(\delta_t\right) = 0$, $Var\left(\delta_t\right) = 1$ and $Cov\left(\delta_t, \delta_{t'}\right) = a\left(\frac{t-t'}{l}\right)$. For DRW,

$$\begin{aligned}
T^*_{n,DRW} &= \sqrt{n}\left(\bar{X}^*_{n,DRW} - \bar{X}_n\right)S_Z = \sqrt{n}\sum_{t=1}^{n}\frac{Z_t}{\sum_{t=1}^{n} Z_t}\left(X_t - \bar{X}_n\right)S_Z, \\
&= \frac{1}{\sqrt{n}}\left(\frac{1}{n}\sum_{t=1}^{n} Z_t\right)^{-1}\sum_{t=1}^{n}\left[\left(Z_t - E\left(Z_1\right)\right)\left(X_t - \bar{X}_n\right)\right]\frac{E\left(Z_1\right)}{\sqrt{Var\left(Z_1\right)}}, \\
&= \left(\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\left(X_t - \bar{X}_n\right)\frac{Z_t - E\left(Z_1\right)}{\sqrt{Var\left(Z_1\right)}}\right)\left(\frac{\frac{1}{n}\sum_{t=1}^{n} Z_t}{E\left(Z_1\right)}\right)^{-1}, \\
&= \left(\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\left(X_t - \bar{X}_n\right)\delta_t\right)\left(\frac{\frac{1}{n}\sum_{t=1}^{n} Z_t}{E(Z_1)}\right)^{-1} = T^*_{n,DWB}\left(1 + U_n\right)^{-1},
\end{aligned}$$

where $\delta_t = \frac{Z_t - E(Z_1)}{\sqrt{Var(Z_1)}}$ and $U_n = \frac{1}{nS_Z}\sum_{t=1}^{n}\delta_t$. Note that

$$E^*\left[\left(\sum_{t=1}^{n}\delta_t\right)^2\right] = \sum_{t_1,t_2=1}^{n} E\left[\delta_{t_1}\delta_{t_2}\right] \le 2\sum_{1\le t_1\le t_2\le n}\left|E\left[\delta_{t_1}\delta_{t_2}\right]\right| = O(nl),$$

in view of the fact that under $l$-dependence, $E\left[\delta_{t_1}\delta_{t_2}\right] \ne 0$ only when $|t_1 - t_2| \le l$. Thus,

$$E^*[U_n^2] = \frac{1}{n^2 S_Z^2}E^*\left[\left(\sum_{t=1}^{n}\delta_t\right)^2\right] = O\left(\frac{l}{n}\right) \to 0 \text{ as n} \to \infty,$$

which implies $U_n \to^P 0$. Further, note that for $|x| < \frac{1}{2}$, $\left|\frac{x}{1+x}\right| < 2|x|$ and hence for any $0 < \epsilon < \frac{1}{2}$, $1 \leftarrow P\left[|U_n| \le \frac{\epsilon}{2}\right] \le P\left[\left|\frac{1}{1+U_n} - 1\right| \le \epsilon\right]$, that is,

$$(1 + U_n)^{-1} \to^P 1.$$

Hence, (A2) holds by conditional Slutsky's theorem (see Lemma 4.1 of Lahiri (2003b)) and the fact that

$$T^*_{n,DWB} = \frac{1}{\sqrt{n}}\sum_{t=1}^{n}\left(X_t - \bar{X}_n\right)\delta_t \to_D N\left(0, \boldsymbol{\Sigma}_\infty\right), \tag{A3}$$


Copyright © 2014 Wiley Publishing Ltd

DOI: 10.1111/jtsa.12109

in probability conditional on the data, the latter of which has been shown in the work of Shao (2010). Finally, our conclusion follows from the argument in the proof of theorem 4.1 of Lahiri (2003b). We omit the details here. □

*Proof of Theorem 2*
We prove case (i) only, as case (ii) can be dealt with in a similar fashion. For case (i), following the argument in the proof of Theorem 1, we can write

$$T^*_{n,DRW} = \sqrt{n}\left(\bar{X}^*_{n,DRW} - \bar{X}_n\right) S_Z = \left(\frac{1}{\sqrt{n}}\sum_{j=1}^{n}\left(X\left(t_j\right) - \bar{X}_n\right)\delta\left(t_j\right)\right)(1+U_n)^{-1} = \frac{T^*_{n,DWB}}{1+U_n},$$

where $\delta_{t_j} = \frac{Z_{t_j} - E(Z_1)}{\sqrt{Var(Z_1)}}$ and $U_n = \frac{1}{nS_Z}\sum_{j=1}^{n}\delta(t_j)$. We want to show that $E^*\left[U_n^2\right] = E[U_n^2] \to 0$. Note that

$$E\left[\left(\sum_{j=1}^{n}\delta\left(t_j\right)\right)^2\right] = E_V\left[E\left[\left(\sum_{j=1}^{n}\delta\left(t_j\right)\right)^2 |V\right]\right] = E_V\left[\sum_{j,j'=1}^{n} a\left(\frac{t_j - t_{j'}}{l_n}\right)\right].$$

For $j = j'$, clearly $E_V\left[a\left(\frac{t_j - t_{j'}}{l_n}\right)\right] = 1$. For $j \neq j'$,

$$E_V\left[a\left(\frac{t_j - t_{j'}}{l_n}\right)\right] = \int_{R_0^2} a\left(\frac{\lambda_n(x-y)}{l_n}\right)\eta(x)\eta(y)dxdy.$$

Let $R_1 = \{x - y : x \in R_0, y \in R_0\}$, and for $z \in R_1$, let $R(z) = (R_0 + z) \cap R_0$. Then, the previous integral equals

$$\int_{R_1}\int_{x\in R(z)} a\left(\frac{\lambda_n z}{l_n}\right)\eta(x)\eta(x-z)dxdz = \frac{l_n}{\lambda_n}\int_{\frac{\lambda_n}{l_n}R_1}\int_{x\in R\left(\frac{tl_n}{\lambda_n}\right)} a(t)\eta(x)\eta\left(x - \frac{tl_n}{\lambda_n}\right)dxdt = \frac{l_n}{\lambda_n}I_n.$$

Since $a(\cdot)$ has compact support on $[-1, 1]$ and for $|t| \leq 1$, $\frac{tl_n}{\lambda_n} = o(1)$, so it follows from the continuity of $\eta(\cdot)$ that

$$\limsup |I_n| \leq \limsup \int_{|t|\leq 1}\int_{x\in R\left(\frac{tl_n}{\lambda_n}\right)} \eta(x)\eta\left(x - \frac{tl_n}{\lambda_n}\right)dxdt = \iota < \infty,$$

and hence,

$$E(U_n^2) = E_V\left[\frac{1}{n^2}\sum_{j,j'=1}^{n} a\left(\frac{t_j - t_{j'}}{l_n}\right)\right] = \frac{1}{n^2} \times O\left(n + n^2\frac{l_n}{\lambda_n}\right) = o(1),$$

which implies that $\frac{1}{U_n+1} \to^P 1$ similar to the regular time series case. The conclusion follows from the consistency of the DWB (see Theorem 5.2 of Shao (2010)) and an application of conditional Slutsky's theorem (Lemma 4.1. of Lahiri (2003b)). The proof for case (ii) follows in a similar fashion, and we skip the details. □

 Copyright © 2014 Wiley Publishing Ltd 
DOI: 10.1111/jtsa.12109