INFORMS Journal on Computing

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Efficient Computation of Overlapping Variance Estimators for Simulation

Christos Alexopoulos, Nilay Tanık Argon, David Goldsman, Natalie M. Steiger, Gamze Tokol, James R. Wilson,

To cite this article:
Christos Alexopoulos, Nilay Tanık Argon, David Goldsman, Natalie M. Steiger, Gamze Tokol, James R. Wilson, (2007) Efficient Computation of Overlapping Variance Estimators for Simulation. INFORMS Journal on Computing 19(3):314-327. https://doi.org/10.1287/ijoc.1060.0198

INFORMS Journal on Computing
Vol. 19, No. 3, Summer 2007, pp. 314–327
ISSN 1091-9856 | EISSN 1526-5528 | 07 | 1903 | 0314

DOI 10.1287/ijoc.1060.0198

Efficient Computation of Overlapping Variance Estimators for Simulation

Christos Alexopoulos
H. Milton Stewart School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332, christos@isye.gatech.edu

Nilay Tanık Argon
Department of Statistics and Operations Research, University of North Carolina, Chapel Hill, North Carolina 27599, nilay@unc.edu

David Goldsman
H. Milton Stewart School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332, sman@gatech.edu

Natalie M. Steiger
Maine Business School, University of Maine, Orono, Maine 04469, nsteiger@maine.edu

Gamze Tokol
Decision Analytics, Atlanta, Georgia 30306, gamze@mindspring.com

James R. Wilson
Edward P. Fitts Department of Industrial and Systems Engineering, North Carolina State University, Raleigh, North Carolina 27695, jwilson@ncsu.edu

For a steady-state simulation output process, we formulate efficient algorithms to compute certain estimators of the process variance parameter (i.e., the sum of covariances at all lags), where the estimators are derived in principle from overlapping batches separately and then averaged over all such batches. The algorithms require order-of-sample-size work to evaluate overlapping versions of the area and Cramér–von Mises estimators arising in the method of standardized time series. Recently, Alexopoulos et al. showed that, compared with estimators based on nonoverlapping batches, the estimators based on overlapping batches achieve reduced variance while maintaining similar bias asymptotically as the batch size increases. We provide illustrative analytical and Monte Carlo results for $M/M/1$ queue waiting times and for a first-order autoregressive process. We also present evidence that the asymptotic distribution of each overlapping variance estimator can be closely approximated using an appropriately rescaled chi-squared random variable with matching mean and variance.

Key words: steady-state simulation; simulation output analysis methods; method of batch means; method of standardized time series; area variance estimator; Cramér–von Mises variance estimator; nonoverlapping variance estimator; overlapping variance estimator

History: Accepted by Susan M. Sanchez, Area Editor for Simulation; received January 2006; revised April 2006, June 2006; accepted June 2006. Published online in *Articles in Advance* July 25, 2007.

1. Introduction

The simplest goal in a steady-state simulation experiment is often to estimate the unknown mean μ of a resulting output process, $\{Y_i\colon\ i = 1, 2, \ldots\}$, which is taken to be stationary. Based on a simulation-generated time series of length n, the obvious estimator for μ is the sample mean, $\overline{Y}_n = n^{-1}\sum_{i=1}^{n} Y_i$. In addition, the careful experimenter ought to provide an estimate for the variance of the sample mean, $\mathrm{Var}(\overline{Y}_n) = \mathrm{E}[(\overline{Y}_n - \mu)^2]$. To characterize the asymptotic behavior of $\overline{Y}_n$ as the run length (sample size) n tends to infinity, the experimenter could estimate either (i) the quantities $\{\sigma_n^2 \equiv n\mathrm{Var}(\overline{Y}_n)\colon\ n = 1, 2, \ldots\}$; or (ii) the *variance parameter*, $\sigma^2 \equiv \lim_{n\to\infty} \sigma_n^2$, provided that this limit exists. Among the many techniques in the literature for analyzing steady-state simulation output and for estimating σ^2 are the following: nonoverlapping batch means (NBM) (Schmeiser 1982); overlapping batch means (OBM) (Meketon and Schmeiser 1984); and standardized time series (STS) (Schruben 1983). The STS technique gives rise to such estimators as the STS area estimator (Goldsman et al.

1990) and the STS Cramér–von Mises (CvM) estimator (Goldsman et al. 1999).

A common strategy used with the above estimators involves *batching* the observations. For instance, the NBM, STS area, and STS CvM estimators split the set of observations into adjacent disjoint (nonoverlapping) batches, where each batch is a subseries (i.e., a group of consecutive observations) that is presumably large enough to provide a representative "snapshot" of the overall dependency structure of the entire simulation-generated time series $\{Y_i\colon i = 1, \ldots, n\}$. In the case of NBM, we then assume that the resulting sample (batch) means computed from each batch are approximately independent and identically distributed (i.i.d.) normal random variables; and finally we appropriately rescale the sample variance of the batch means so as to obtain an estimator of σ^2. In the case of the STS estimators, we compute a separate STS estimator of σ^2 from each batch; we assume the resulting STS estimators are i.i.d.; and then we average those estimators to obtain an overall batched estimator of σ^2.

On the other hand, the OBM method uses *overlapping* batches, so that the corresponding batch means are *not* independent (though they are identically distributed and asymptotically normal). As in the computation of the NBM variance estimator, the OBM estimator of σ^2 is an appropriately rescaled version of the sample variance of the overlapping batch means; and it can be shown that as the batch size $m \to \infty$ (so that the simulation run length $n \to \infty$), the OBM estimator asymptotically has (nearly) the same bias as, but smaller variance than, the corresponding NBM estimator for σ^2. Alexopoulos et al. (2006b) formulated overlapping versions of the STS area and CvM estimators; and they found that as the batch size $m \to \infty$, the STS overlapping estimators asymptotically had (nearly) the same bias as, but substantially smaller variance than, their counterparts computed from nonoverlapping batches. In this paper, we present numerically efficient algorithms for computing the new variance estimators, and we evaluate the statistical performance of these estimators in several analytical and Monte Carlo examples.

The rest of this paper is organized as follows. First, in Section 2 we provide a self-contained synopsis of relevant background material. In Section 3 we formulate numerically efficient algorithms for computing the proposed overlapping variance estimators, and we show these algorithms to possess the unexpected advantage of requiring execution times that are linear in the sample size. To complement the theoretical results presented in Alexopoulos et al. (2006b) on the large-sample moment and distributional properties of our new variance estimators, in Section 4 we present both theoretically exact and empirically estimated small-sample results for some illustrative examples with finite but progressively increasing batch sizes. Moreover, Section 4 provides both theoretical and empirical evidence that for a sufficiently large batch size, the distribution of each overlapping variance estimator can be closely approximated by the distribution of a rescaled chi-squared random variable whose scaling factor and degrees of freedom are set to match the mean and variance of the target distribution. To round out the development in Section 4, we formulate approximate confidence intervals for μ and σ^2 based on our approximate distributions for the overlapping variance estimators. Finally, in Section 5 we summarize the main contributions of this work and make recommendations for future research.

Alexopoulos et al. (2004) present a preliminary, abridged version of some of the Monte Carlo results that are more fully described in this paper. In the Online Supplement to this paper on the journal's website, we provide complete derivations of the analytic results for the examples presented in this paper. For efficient computation of the proposed point and confidence-interval estimators of μ and σ^2 from user-supplied data sets, a public-domain Windows-based software package is available online via Alexopoulos et al. (2006a).

2. Background

In this section, we define the relevant variance estimators and review bias and variance results for those estimators. Throughout we consider a stationary stochastic process $\{Y_i\colon i = 1, \ldots, n\}$ that is generated by a simulation in steady-state operation and that satisfies a functional central limit theorem assumption as described in Alexopoulos et al. (2006b).

2.1. Nonoverlapping Batched Estimators

In this section, we work with b contiguous, nonoverlapping batches of observations, each of length m, from the simulation-generated time series $\{Y_i\colon i = 1, \ldots, n\}$ (where we assume that $n = bm$). Thus, the observations $\{Y_{(i-1)m+k}\colon k = 1, \ldots, m\}$ constitute batch i for $i = 1, \ldots, b$. The *standardized time series* computed from batch i is defined as $T_{i,m}(t) \equiv \lfloor mt \rfloor (\overline{Y}_{i,m} - \overline{Y}_{i,\lfloor mt \rfloor})/(\sigma\sqrt{m})$ for $t \in [0, 1]$ and $i = 1, \ldots, b$, where $\overline{Y}_{i,j} \equiv (1/j)\sum_{k=1}^{j} Y_{(i-1)m+k}$ for $i = 1, \ldots, b$ and $j = 1, \ldots, m$ and $\lfloor \cdot \rfloor$ is the floor function. The averages $\{\overline{Y}_{i,m}\colon i = 1, \ldots, b\}$ are the (nonoverlapping) *batch means*.

2.1.1. Weighted Area Estimator. We define the weighted area estimator computed exclusively from batch i as follows:

$$A_i(f; m) \equiv \left[(1/m)\sum_{k=1}^{m} f(k/m)\sigma T_{i,m}(k/m)\right]^2 \quad \text{for } i = 1, \ldots, b,$$

where $f(\cdot)$ is a continuous weighting function on the interval $[0,1]$ and $f(\cdot)$ is normalized so that

$$\int_0^1 \int_0^1 f(s)f(t)(\min(s,t)-st)\,ds\,dt = 1.$$

The *batched area* estimator for σ^2 is

$$\mathcal{A}(f;b,m) \equiv \frac{1}{b}\sum_{i=1}^{b} A_i(f;m) \xrightarrow[m\to\infty]{\mathcal{D}} \sigma^2\chi_b^2/b,$$

where $\xrightarrow[m\to\infty]{\mathcal{D}}$ denotes convergence in distribution as $m\to\infty$ and χ_ν^2 denotes a χ^2 random variable with ν degrees of freedom. Popular choices of weighting functions are

$$\left.\begin{aligned} f_0(t) &= \sqrt{12} \quad \text{(Schruben 1983)},\\ f_2(t) &= \sqrt{840}(3t^2-3t+1/2)\\ &\qquad \text{(Goldsman et al. 1990)}, \quad \text{and}\\ f_{\cos,j}(t) &= \sqrt{8}\pi j\cos(2\pi jt) \quad \text{with } j=1,2,\ldots\\ &\qquad \text{(Foley and Goldsman 1999)}, \end{aligned}\right\} \quad \text{for } t\in[0,1]. \quad (1)$$

Foley and Goldsman (1999) show that we can average the first k area estimators arising from the weighting functions $\{f_{\cos,j}(\cdot)\colon j=1,2,\ldots\}$ to obtain estimators for σ^2 with more degrees of freedom,

$$\bar{\mathcal{A}}_k(f_{\cos};b,m) \equiv \frac{1}{k}\sum_{j=1}^{k}\mathcal{A}(f_{\cos,j};b,m) \xrightarrow[m\to\infty]{\mathcal{D}} \sigma^2\chi_{bk}^2/(bk)$$

$$\text{for } k=1,2,\ldots.$$

2.1.2. Weighted Cramér–von Mises Estimator. The weighted CvM estimator computed exclusively from batch i is defined as follows: $C_i(g;m) \equiv (1/m)\sum_{k=1}^{m} g(k/m)\sigma^2 T_{i,m}^2(k/m)$ for $i=1,\ldots,b$, where $g(\cdot)$ is a normalized weighting function on $[0,1]$ for which $\int_0^1 g(t)t(1-t)\,dt = 1$ and whose second derivative $(d^2/dt^2)g(t)$ is continuous on $[0,1]$. The *batched CvM* estimator for σ^2 is $\mathcal{C}(g;b,m) \equiv (1/b)\sum_{i=1}^{b} C_i(g;m)$; see Goldsman et al. (1999). Popular choices of weighting functions are

$$\left.\begin{aligned} g_0(t) &= 6,\\ g_2^\star(t) &= -24+150t-150t^2, \quad \text{and}\\ g_4^\star(t) &= -\frac{1{,}310}{21}+\frac{19{,}270}{21}t-\frac{25{,}230}{7}t^2\\ &\quad +\frac{16{,}120}{3}t^3-\frac{8{,}060}{3}t^4 \end{aligned}\right\} \quad \text{for } t\in[0,1].$$

2.1.3. NBM Estimator. The NBM estimator for σ^2 is

$$\mathcal{N}(b,m) \equiv \frac{m}{b-1}\sum_{i=1}^{b}(\overline{Y}_{i,m}-\overline{Y}_n)^2 \xrightarrow[m\to\infty]{\mathcal{D}} \sigma^2\chi_{b-1}^2/(b-1);$$

see Glynn and Whitt (1991); Schmeiser (1982); and Steiger and Wilson (2001).

2.2. Overlapping Estimators

In this section, we divide the sample $\{Y_i\colon i=1,\ldots,n\}$ into $n-m+1$ overlapping batches, each of size m so that the observations $\{Y_{i+k}\colon k=0,\ldots,m-1\}$ constitute batch i for $i=1,\ldots,n-m+1$. We will continue to let $b\equiv n/m$ denote the ratio of the sample size to the batch size as we did for the nonoverlapping variance estimators; however, in the context of overlapping batches, b can no longer be interpreted as "the number of batches." The standardized time series computed from overlapping batch i is $T_{i,m}^{\mathrm{O}}(t) \equiv \lfloor mt\rfloor(\overline{Y}_{i,m}^{\mathrm{O}} - \overline{Y}_{i,\lfloor mt\rfloor}^{\mathrm{O}})/(\sigma\sqrt{m})$ for $t\in[0,1]$ and $i=1,\ldots,n-m+1$, where $\overline{Y}_{i,j}^{\mathrm{O}} \equiv (1/j)\sum_{k=0}^{j-1} Y_{i+k}$ for $i=1,\ldots,n-m+1$ and $j=1,\ldots,m$.

2.2.1. Overlapping Area Estimator. We define the overlapping area estimator computed from overlapping batch i by $A_i^{\mathrm{O}}(f;m) \equiv [(1/m)\sum_{k=1}^{m} f(k/m)\cdot\sigma T_{i,m}^{\mathrm{O}}(k/m)]^2$ for $i=1,\ldots,n-m+1$. The *overlapping area* estimator for σ^2 is $\mathcal{A}^{\mathrm{O}}(f;b,m) \equiv (1/(n-m+1))\cdot\sum_{i=1}^{n-m+1} A_i^{\mathrm{O}}(f;m)$. Following a development similar to Section 2.1.1, we can average the first k overlapping area estimators arising from the weighting functions $\{f_{\cos,j}(\cdot)\colon j=1,2,\ldots\}$ to obtain the following class of estimators (with lower variance): $\bar{\mathcal{A}}_k^{\mathrm{O}}(f_{\cos};b,m) \equiv (1/k)\sum_{j=1}^{k}\mathcal{A}^{\mathrm{O}}(f_{\cos,j};b,m)$ for $k=1,2,\ldots$.

2.2.2. Overlapping CvM Estimator. We define the overlapping CvM estimator computed from overlapping batch i by $C_i^{\mathrm{O}}(g;m) \equiv (1/m)\sum_{k=1}^{m} g(k/m)\cdot[\sigma T_{i,m}^{\mathrm{O}}(k/m)]^2$ for $i=1,\ldots,n-m+1$. The *overlapping CvM* estimator for σ^2 is $\mathcal{C}^{\mathrm{O}}(g;b,m) \equiv (1/(n-m+1))\cdot\sum_{i=1}^{n-m+1} C_i^{\mathrm{O}}(g;m)$.

2.2.3. OBM Estimator. The OBM estimator for σ^2 is $\mathcal{O}(b,m) \equiv (nm/((n-m+1)(n-m)))\sum_{i=1}^{n-m+1}(\overline{Y}_{i,m}^{\mathrm{O}} - \overline{Y}_n)^2$; see Meketon and Schmeiser (1984).

2.3. Summary of Properties

Table 1 gives a succinct summary of the bias and variance properties of the estimators reviewed in this section. Following Song and Schmeiser (1995), we define $\gamma \equiv -2\sum_{k=1}^{\infty} kR_k$, provided $\sum_{k=1}^{\infty} k|R_k| < \infty$, where $R_k \equiv \mathrm{Cov}(Y_1, Y_{1+k})$ for $k=0,\pm1,\ldots$. We also use the notation $p(n) = o(q(n))$ to mean that $p(n)/q(n)\to 0$ as $n\to\infty$.

As discussed in Alexopoulos et al. (2006b) and illustrated in Table 1, all the estimators described herein are asymptotically unbiased for σ^2 as the batch size $m\to\infty$. The weighting functions $f_2(\cdot)$ and $f_{\cos,j}(\cdot)$ yield overlapping area estimators that are first-order unbiased, i.e., they have bias of the form $o(1/m)$ as $m\to\infty$. In a similar fashion the weighting functions $g_2^\star(\cdot)$ and $g_4^\star(\cdot)$ yield overlapping CvM estimators that are also first-order unbiased.

Table 1 also shows that the (asymptotic) variances of the nonoverlapping area estimators do not depend

Table 1 Asymptotic Bias and Variance of Variance Estimators as $m \to \infty$

Nonoverlapping estimators			Overlapping estimators			
Form of estimator	(m/γ)Bias	(b/σ^4)Var	Form of estimator	(m/γ)Bias	Var/σ^4	$\approx (b/\sigma^4)$Var
$\mathcal{A}(f_0; b, m)$	3	2	$\mathcal{A}^0(f_0; b, m)$	3	$\frac{24b-31}{35(b-1)^2}$	0.686
$\mathcal{A}(f_2; b, m)$	$o(1)$	2	$\mathcal{A}^0(f_2; b, m)$	$o(1)$	$\frac{3{,}514b-4{,}359}{4{,}290(b-1)^2}$	0.819
$\bar{\mathcal{A}}_1(f_{\cos}; b, m)$	$o(1)$	2	$\bar{\mathcal{A}}_1^0(f_{\cos}; b, m)$	$o(1)$	$\frac{(16\pi^2+30)b-(20\pi^2+33)}{24\pi^2(b-1)^2}$	0.793
$\bar{\mathcal{A}}_2(f_{\cos}; b, m)$	$o(1)$	1	$\bar{\mathcal{A}}_2^0(f_{\cos}; b, m)$	$o(1)$	$\frac{(384\pi^2+1{,}090)b-(480\pi^2+1{,}455)}{1{,}152\pi^2(b-1)^2}$	0.429
$\mathcal{C}(g_0; b, m)$	5	0.8	$\mathcal{C}^0(g_0; b, m)$	5	$\frac{88b-115}{210(b-1)^2}$	0.419
$\mathcal{C}(g_2^*; b, m)$	$o(1)$	1.729	$\mathcal{C}^0(g_2^*; b, m)$	$o(1)$	$\frac{10{,}768b-13{,}605}{13{,}860(b-1)^2}$	0.777
$\mathcal{C}(g_4^*; b, m)$	$o(1)$	1.042	$\mathcal{C}^0(g_4^*; b, m)$	$o(1)$	—	0.477
$\mathcal{N}(b, m)$	$\frac{b+1}{b}$	$\frac{2b}{b-1}$	$\mathcal{O}(b, m)$	$\frac{b^2+1}{b(b-1)}$	$\frac{4b^3-11b^2+4b+6}{3(b-1)^4}$	1.333

on the weighting function, while the corresponding quantities for both types of CvM estimators do exhibit such dependence. In any case, we see that overlapping estimators always have smaller variances than their nonoverlapping counterparts; moreover, all the STS overlapping estimators under study here have smaller variances than the OBM estimator, which in turn has smaller variance than the benchmark NBM estimator.

3. Computational Complexity of Overlapping Variance Estimators

Low-variance estimators such as the overlapping estimators discussed in Section 2 will not do us any good unless they can be calculated in an efficient manner. As a starting point, Foley and Goldsman (1999) and Goldsman et al. (1999) show that, for fixed batch size m, each of the estimators $A_i^{\mathrm{O}}(f; m)$ and $C_i^{\mathrm{O}}(g; m)$ resulting from a single batch can be computed in $O(m)$ time and $O(m)$ space. Hence, each of the overlapping estimators $\mathcal{A}^{\mathrm{O}}(f; b, m)$ and $\mathcal{C}^{\mathrm{O}}(g; b, m)$ can be computed in $O(mn)$ time and $O(m)$ space. Fortunately, we can do better. In this section, we show how to compute all the overlapping area and CvM estimators from Table 1 in just $O(n)$ time. In the Online Supplement we provide formal algorithmic statements of our procedures for computing the proposed overlapping variance estimators in $O(n)$ time.

3.1. Overlapping Area Estimators

Our goal is to "build up" the overlapping area estimator in $O(n)$ time. The game plan is to define some useful expressions, show that these quantities can be initialized in $O(m)$ time, and then construct the estimators step-by-step.

3.1.1. Definitions and Initialization. For integer-valued variables p, i, and m (where $p = 0, 1, \ldots;\ i = 0, \ldots, n-m$; and $l = 1, \ldots, m$), we let

$$\left.\begin{aligned}
S_{i,l}^{(p)} &\equiv \sum_{k=1}^{l} k^p Y_{i+k}, \quad \text{with } S_{i,l} \equiv S_{i,l}^{(0)} = \sum_{j=i+1}^{i+l} Y_j,\\
V_i(f) &\equiv \sum_{k=1}^{m} f\left(\frac{k}{m}\right)\left(\frac{k}{m} S_{i,m} - S_{i,k}\right),\\
W_i(f) &\equiv \sum_{k=1}^{m} f\left(\frac{k}{m}\right) Y_{i+k},\\
Q_i &\equiv \sum_{k=1}^{m} \cos(\alpha_j(i+k)) Y_{i+k},\\
&\qquad \text{with } \alpha_j \equiv 2\pi j/m \text{ for } j = 1, 2, \ldots,\\
Q_i' &\equiv \sum_{k=1}^{m} \sin(\alpha_j(i+k)) Y_{i+k}, \text{ and}\\
a_p(f) &\equiv \sum_{k=1}^{m} k^p f\left(\frac{k}{m}\right).
\end{aligned}\right\} \quad (2)$$

The definitions of Q_i and Q_i' are based on the assumption that the user has selected a weighting function of the form $f_{\cos,j}(\cdot)$ as in (1) for a specific value of j. Because j is fixed throughout the following discussion, no confusion should result from our use of the simplified notation Q_i, Q_i' in which the dependence of these quantities on j is suppressed.

First we initialize the quantities in (2). For fixed p, we can calculate the entire set of auxiliary quanti-

ties $\{S_{0,l}^{(p)}: l=1,\ldots,m\}$ in $O(m)$ time. Further, assuming that each evaluation of the function $f(\cdot)$ takes $O(1)$ work, we can obtain $V_0(f)$, $W_0(f)$, and $a_p(f)$ in $O(m)$ time. Similarly, Q_0 and Q_0' require $O(m)$ work. Thus, all "initialization" activities can be performed in $O(m)$ time.

3.1.2. Construction of Overlapping Area Estimator. For various weighting functions $f(\cdot)$, we show how to calculate efficiently $W_i(f)$, then $V_i(f)$, and finally $\mathscr{A}^{\mathrm{O}}(f;b,m)$.

Remark 1. To begin, note that for $i=1,\ldots,n-m$, we have

$$S_{i,l}^{(p)} = \sum_{j=1}^{l}(j-1)^p Y_{i-1+j} + l^p Y_{i+l} - Y_i I_{\{p=0\}}$$
$$= \sum_{x=0}^{p}\binom{p}{x}(-1)^{p-x} S_{i-1,l}^{(x)} + l^p Y_{i+l} - Y_i I_{\{p=0\}}, \quad (3)$$

where $0^0 \equiv 1$ and the indicator function $I_{\{z\}}=1$ if z is true, and 0 otherwise. Hence, $S_{i,l}^{(p)}$ can be calculated in $O(p)$ time from the intermediate quantities $\{S_{i-1,l}^{(x)}: x=0,\ldots,p\}$ and the original data.

Suppose that the weighting function can be expressed as the polynomial $f_r(t) \equiv \sum_{p=0}^{r} c_p t^p$ for $t \in [0,1]$. Then

$$W_i(f_r) = \sum_{k=1}^{m}\sum_{p=0}^{r} c_p\left(\frac{k}{m}\right)^p Y_{i+k} = \sum_{p=0}^{r}\frac{c_p}{m^p}\sum_{k=1}^{m} k^p Y_{i+k}$$
$$= \sum_{p=0}^{r}\frac{c_p}{m^p} S_{i,m}^{(p)},$$

which, by (3), can be calculated from available information in $O(r^2)$ time. Since r is typically small, say $r \le 3$, this is an $O(1)$ calculation. ◁

Remark 2. Consider the weighting functions $f_{\cos,j}(t)$. For $i=0,\ldots,n-m$, we have

$$W_i(f_{\cos,j}) = \sqrt{8}\pi j(\cos(\alpha_j i)Q_i + \sin(\alpha_j i)Q_i').$$

This shows that $W_i(f_{\cos,j})$ is computable in $O(1)$ time from the available data since both Q_i and Q_i' can be obtained from Q_{i-1} and Q_{i-1}', respectively, in $O(1)$ time from the iterative relations

$$Q_i = Q_{i-1} + \cos(\alpha_j(i+m))Y_{i+m} - \cos(\alpha_j i)Y_i \quad \text{and}$$
$$Q_i' = Q_{i-1}' + \sin(\alpha_j(i+m))Y_{i+m} - \sin(\alpha_j i)Y_i. \quad ◁$$

Now consider the weighting functions $f_r(\cdot)$ and $f_{\cos,j}(\cdot)$. For $i=1,\ldots,n-m$, we have

$$V_i(f) = \sum_{k=1}^{m} f\left(\frac{k}{m}\right)\left[\frac{k}{m}(S_{i-1,m} + Y_{i+m} - Y_i) - (S_{i-1,k} + Y_{i+k} - Y_i)\right]$$
$$= V_{i-1}(f) + \frac{1}{m}(Y_{i+m} - Y_i)a_1(f) - W_i(f) + Y_i a_0(f).$$

This, together with Remarks 1 and 2, shows that we can go from $V_{i-1}(f)$ to $V_i(f)$ in $O(1)$ time; thus the entire set of auxiliary quantities $\{V_i(f): i=1,\ldots,n-m\}$ can be calculated with a total of $O(n)$ work.

Finally, since for $i=1,\ldots,n-m+1$, we have

$$m^3 A_i^{\mathrm{O}}(f;m) = \left[\sum_{k=1}^{m} f\left(\frac{k}{m}\right)\left(\frac{k}{m}S_{i-1,m} - S_{i-1,k}\right)\right]^2$$
$$= V_{i-1}^2(f),$$

it follows that $A_i^{\mathrm{O}}(f;m)$ can be calculated from available information in $O(1)$ time; and so $\mathscr{A}^{\mathrm{O}}(f;b,m) = \sum_{i=1}^{n-m+1} A_i^{\mathrm{O}}(f;m)/(n-m+1)$ needs a total of $O(n)$ work.

3.2. Overlapping CvM Estimators

Our method of attack will be somewhat different when it comes to the overlapping CvM variance estimator $\mathscr{C}^{\mathrm{O}}(g;b,m)$, for there turns out to be a particularly nasty term that we must eventually handle. In particular, we will use a "divide-and-conquer" strategy to determine the required computational effort. To get things going, we assume for convenience that $n \ge 2m$, and we expand the estimator as follows:

$$(n-m+1)m^2\mathscr{C}^{\mathrm{O}}(g;b,m)$$
$$= \sum_{i=0}^{n-m}\left[\frac{a_2(g)S_{i,m}^2}{m^2} - 2S_{i,m}\sum_{k=1}^{m} g\left(\frac{k}{m}\right)\frac{k}{m}S_{i,k} + \sum_{k=1}^{m} g\left(\frac{k}{m}\right)S_{i,k}^2\right]. \quad (4)$$

We show that $O(n)$ time is required to compute each of the three terms on the right-hand side of (4).

3.2.1. First Term of Equation (4). We henceforth assume, as for all our CvM weighting functions in Table 1, that $g(t) \equiv \sum_{p=0}^{r} c_p t^p$ (for $t \in [0,1]$) is an rth-order polynomial. Thus by the discussion after (2), we know that for any fixed p (where $p=0,1,\ldots$), the computation of $a_p(g)$ requires a one-time $O(m)$ expenditure prior to the start of experimentation. Further, since $S_{i,m} = S_{i-1,m} + Y_{i+m} - Y_i$ for $i=1,2,\ldots$, it is obvious that the calculation of the entire set of auxiliary quantities $\{S_{i,m}: i=1,\ldots,n-m\}$ requires a total of $O(n)$ work; so the sum $\sum_{i=0}^{n-m} S_{i,m}^2$, and hence, the first term in (4), can be evaluated in $O(n)$ time.

3.2.2. Second Term of Equation (4). We have already shown that $S_{i,m}$ poses no problem. Thus, consider, for $i=1,\ldots,n-m$,

$$\sum_{k=1}^{m} g\left(\frac{k}{m}\right)\frac{k}{m}S_{i,k} = \sum_{k=1}^{m} g\left(\frac{k}{m}\right)\frac{k}{m}S_{i-1,k} + W_i(g_{r+1}) - \frac{Y_i a_1(g)}{m}, \quad (5)$$

where $g_{r+1}(t) \equiv tg(t)$ is an $(r+1)$st-order polynomial in $t \in [0,1]$. Concerning initialization, we find

that there is a one-time $O(m)$ charge for computing $\sum_{k=1}^{m} g(k/m)(k/m)S_{0,k}$, where we take advantage of the iterative relation $S_{0,k} = S_{0,k-1} + Y_k$ for $k = 2, 3, \ldots$. Further, by the discussion after (2), we know that $W_0(g_{r+1})$ and $a_1(g)$ both require one-time $O(m)$ calculations. Thus, all initialization activities take $O(m)$ time. In addition, Remark 1 implies that evaluating the entire set of auxiliary quantities $\{W_i(g_{r+1})\colon i = 1, \ldots, n-m\}$ needs a total of $O(n)$ work. Now that we have all three terms from the right-hand side of (5) in hand, we can obtain $\sum_{k=1}^{m} g(k/m)(k/m)S_{i,k}$ in $O(1)$ effort. This in turn shows that the second term of (4) can be evaluated in $O(n)$ time.

3.2.3. Third Term of Equation (4). What remains to show is that the last term of (4) can be computed in $O(n)$ time. To this end, let $\tilde{g}_k \equiv g(k/m)$ for $k = 1, \ldots, m$, and consider

$$\begin{aligned}\sum_{i=0}^{n-m}\sum_{k=1}^{m}\tilde{g}_k S_{i,k}^2 &= \sum_{l=1}^{n-m} Y_l \sum_{k=1}^{m}\tilde{g}_k \sum_{i=\max(0,\,l-k)}^{\min(l-1,\,n-m)} S_{i,k} \\ &\quad + \sum_{l=n-m+1}^{n} Y_l \sum_{k=l-n+m}^{m}\tilde{g}_k \sum_{i=\max(0,\,l-k)}^{\min(l-1,\,n-m)} S_{i,k} \\ &= \sum_{l=1}^{n} Y_l M(l), \end{aligned} \tag{6}$$

where

$$M(l) \equiv \sum_{k=\max(1,\,l-n+m)}^{m}\tilde{g}_k \sum_{i=\max(0,\,l-k)}^{\min(l-1,\,n-m)} S_{i,k} \quad \text{for } l = 1, \ldots, n.$$

3.2.3.1. An Iterative Expression for $M(l)$. As in the discussion after (5), we see that $M(1) = \sum_{k=1}^{m}\tilde{g}_k S_{0,k}$ requires a one-time $O(m)$ expenditure. We next express $M(l)$ in terms of $M(l-1)$ for $l = 2, \ldots, n$; and then we show that $M(l)$ can be obtained in $O(1)$ time from $M(l-1)$ and available information.

For $l = 2, \ldots, m$, we have

$$\begin{aligned} M(l) &= \sum_{k=1}^{m}\tilde{g}_k \sum_{i=\max(0,\,l-k)}^{l-1} S_{i,k} \\ &= \sum_{k=1}^{l}\tilde{g}_k \sum_{i=l-k}^{l-1} S_{i,k} + \sum_{k=l+1}^{m}\tilde{g}_k \sum_{i=0}^{l-1} S_{i,k} \\ &= M(l-1) + \sum_{k=1}^{m}\tilde{g}_k S_{l-1,k} - \sum_{k=1}^{l-1}\tilde{g}_k S_{l-1-k,k}. \end{aligned}$$

Similarly, for $l = m+1, \ldots, n-m+1$, we have

$$\begin{aligned} M(l) &= \sum_{k=1}^{m}\tilde{g}_k \sum_{i=l-k}^{l-1} S_{i,k} \\ &= M(l-1) + \sum_{k=1}^{m}\tilde{g}_k S_{l-1,k} - \sum_{k=1}^{m}\tilde{g}_k S_{l-1-k,k}; \end{aligned}$$

and for $l = n-m+2, \ldots, n$, we have

$$\begin{aligned} M(l) &= \sum_{k=l-n+m}^{m}\tilde{g}_k \sum_{i=l-k}^{n-m} S_{i,k} \\ &= \sum_{k=l-n+m}^{m}\tilde{g}_k \sum_{i=l-1-k}^{n-m} S_{i,k} - \sum_{k=l-n+m}^{m}\tilde{g}_k S_{l-1-k,k} \\ &= M(l-1) - \sum_{k=l-n+m-1}^{m}\tilde{g}_k S_{l-1-k,k}. \end{aligned}$$

Hence, for $l = 2, \ldots, n$, we have

$$M(l) = M(l-1) + I_{\{2\le l\le n-m+1\}}\sum_{k=1}^{m}\tilde{g}_k S_{l-1,k} - U_l, \tag{7}$$

where

$$U_l \equiv \sum_{k=\max(1,\,l-n+m-1)}^{\min(l-1,\,m)}\tilde{g}_k S_{l-1-k,k}.$$

3.2.3.2. Obtaining $M(l)$ in $O(1)$ Time. With $M(l-1)$ in hand, we show how to calculate the last two terms of (7) in $O(1)$ time. First, note that for $l = 2, \ldots, n-m+1$,

$$\sum_{k=1}^{m}\tilde{g}_k S_{l-1,k} = \sum_{k=1}^{m}\tilde{g}_k S_{l-2,k} + W_{l-1}(g) - a_0(g)Y_{l-1},$$

all of whose constituents require $O(1)$ computation.

To complete the proof, we must show how to go from U_l to U_{l+1} in $O(1)$ time for each $l \in \{2, \ldots, n-1\}$. Before doing so, we consider the following quantity of interest for positive integers $\alpha \le \beta$,

$$\begin{aligned} Z_{l,\alpha,\beta}^{(p)} &\equiv \sum_{j=\alpha}^{\beta} j^p Y_{l-j} \\ &= \sum_{x=0}^{p}\binom{p}{x} Z_{l-1,\alpha,\beta}^{(x)} + \alpha^p Y_{l-\alpha} - (\beta+1)^p Y_{l-1-\beta}. \end{aligned} \tag{8}$$

At last, we are ready to compute U_l for all $l \in \{2, \ldots, n\}$.

Updating U_l to U_{l+1} for $l = 2, \ldots, m$. We initialize $U_2 = \tilde{g}_1 S_{0,1} = \tilde{g}_1 Y_1$. Then

$$U_{l+1} = U_l + Y_l\sum_{k=1}^{l-1}\tilde{g}_k - \sum_{p=0}^{r}\frac{c_p}{m^p} Z_{l,1,l-1}^{(p)} + \tilde{g}_l S_{0,l}.$$

The terms $Y_l\sum_{k=1}^{l-1}\tilde{g}_k$ and $\tilde{g}_l S_{0,l}$ are easy to compute in $O(1)$ time, simply by keeping running counters for the sums. Furthermore, by (8) with $Y_0 \equiv 0$, we have

$$\begin{aligned} Z_{l,1,l-1}^{(p)} &= \sum_{x=0}^{p}\binom{p}{x} Z_{l-1,1,l-1}^{(x)} + Y_{l-1} - l^p Y_0 \\ &= \sum_{x=0}^{p}\binom{p}{x} Z_{l-1,1,l-2}^{(x)} + Y_{l-1}, \end{aligned}$$

for $p = 0, \ldots, r$ and $l = 2, \ldots, m$. Initialization of $Z^{(p)}_{2,1,1}$ for $p = 0, \ldots, r$ takes a total of $O(1)$ work since r is fixed. Therefore, calculating $Z^{(p)}_{l,1,l-1}$ from the set of intermediate quantities $\{Z^{(x)}_{l-1,1,l-2}\colon x = 0, \ldots, p\}$ and Y_{l-1} takes $O(1)$ effort; and hence, we can go from U_l to U_{l+1} in $O(1)$ time for each of $l \in \{2, \ldots, m\}$.

Updating U_l to U_{l+1} for $l = m+1, \ldots, n-m+1$. By the preceding development, U_{m+1} has already been initialized. Then

$$U_{l+1} = U_l + Y_l a_0(g) - \sum_{p=0}^{r} \frac{c_p}{m^p} Z^{(p)}_{l,1,m}.$$

By (8), we have

$$Z^{(p)}_{l,1,m} = \sum_{x=0}^{p} \binom{p}{x} Z^{(x)}_{l-1,1,m} + Y_{l-1} - (m+1)^p Y_{l-(m+1)},$$

for $p = 0, \ldots, r$ and $l = m+1, \ldots, n-m+1$. As before, initialization of $Z^{(p)}_{m+1,1,m}$ for $p = 0, \ldots, r$ takes a total of $O(1)$ work; and then calculating $Z^{(p)}_{l,1,m}$ from the set of intermediate quantities $\{Z^{(x)}_{l-1,1,m}\colon x = 0, \ldots, p\}$ and previous data takes $O(1)$ effort. Thus, we can obtain U_{l+1} from U_l in $O(1)$ time for each $l \in \{m+1, \ldots, n-m+1\}$.

Updating U_l to U_{l+1} for $l = n-m+2, \ldots, n-1$. By the preceding development, U_{n-m+2} has already been initialized. Then

$$U_{l+1} = U_l + Y_l \sum_{k=l-1-(n-m)}^{m} \tilde{g}_k - \sum_{p=0}^{r} \frac{c_p}{m^p} Z^{(p)}_{l,l-1-(n-m),m} - \tilde{g}_{l-1-(n-m)} S_{n-m+1,l-1-(n-m)}.$$

By (8), we have

$$Z^{(p)}_{l,l-1-(n-m),m} = \sum_{x=0}^{p} \binom{p}{x} Z^{(x)}_{l-1,l-2-(n-m),m} - (m+1)^p Y_{l-(m+1)}$$

for $p = 0, \ldots, r$ and $l = n-m+2, \ldots, n$. Yet again, initialization of $\{Z^{(p)}_{n-m+2,1,m}\colon p = 0, \ldots, r\}$ takes $O(1)$ work; and then calculating $Z^{(p)}_{l,l-1-(n-m),m}$ from $\{Z^{(x)}_{l,l-2-(n-m),m}\colon x = 0, \ldots, p\}$ and previous data takes $O(1)$ effort. This implies that obtaining U_{l+1} from U_l for each $l \in \{n-m+2, \ldots, n-1\}$ takes $O(1)$ work.

3.2.3.3. CvM Computation Wrap-Up. Thus, we have shown that going from U_l to U_{l+1} takes $O(1)$ work for each of $l = 2, 3, \ldots, n-1$, so that calculation of $M(l)$ from previous information via (7) takes $O(1)$ effort. Then, the sum over all l in (6) can be calculated in $O(n)$ time, and hence, all of (4) can be computed in $O(n)$ time.

4. Examples

This section illustrates the performance of the various estimators on a number of stochastic processes. In Section 4.1 we describe a Monte Carlo investigation involving first-order autoregressive and $M/M/1$ queue-waiting-time processes, where we provide estimated expected values and variances for the competing variance estimators. In the Online Supplement, we give exact results for a simple i.i.d. normal process, along with representative runtime comparisons for an $M/M/1$ process. The purpose of Section 4.2 is to study the distributions of the variance estimators, which we found could be closely approximated by the distributions of certain rescaled chi-squared variates. In Section 4.3, we derive approximate confidence interval estimators for μ and σ^2.

4.1. Variance Estimation

In this section, we estimate the expected values and variances of the variance estimators when applied to the following two stationary processes.

1. **First-order autoregressive (AR(1)) process.** A stationary (Gaussian) AR(1) process is defined by $Y_i = \phi Y_{i-1} + \varepsilon_i$ for $i = 1, 2, \ldots$, where $-1 < \phi < 1$; Y_0 is a Nor(0, 1) random variable; and the ε_i's are i.i.d. Nor$(0, 1-\phi^2)$ random variables that are independent of Y_0. The AR(1) process has covariance function $R_k = \phi^{|k|}$ for all $k = 0, \pm 1, \pm 2, \ldots$, so that $\sigma^2 = (1+\phi)/(1-\phi)$ and $\gamma = -2\phi/(1-\phi)^2$; see, e.g., Sargent et al. (1992).

2. $M/M/1$ **queue-waiting-time process.** We also consider the stationary queue-waiting-time process for an $M/M/1$ queue with arrival rate ω and traffic intensity $\rho < 1$; i.e., a queueing system experiencing Poisson arrivals and first-in-first-out i.i.d. exponential service times at a single server. For this process, we have $\sigma^2 = \rho^3(2 + 5\rho - 4\rho^2 + \rho^3)/[\omega^2(1-\rho)^4]$; see, for example, Steiger and Wilson (2001).

Remark 3. Suppose that on one run of a simulation, the conventional NBM variance estimator $\mathcal{N}(b, m)$ requires total computing time (work) with expected value $\mathbb{C}_0$ to deliver a statistic with bias $\mathbb{B}_0$, variance $\mathbb{V}_0$, and mean-squared error $\mathbb{M}_0 = \mathbb{B}_0^2 + \mathbb{V}_0$; and let $\mathbb{C}_1$, $\mathbb{B}_1$, $\mathbb{V}_1$, and $\mathbb{M}_1$ respectively denote the corresponding performance measures for a given overlapping variance estimator whose efficiency is to be evaluated. Taking $\mathcal{N}(b, m)$ as the baseline variance estimator, we can evaluate the relative efficiency of the given overlapping variance estimator in terms of the ratios $\mathbb{Q} = \mathbb{C}_1\mathbb{V}_1/(\mathbb{C}_0\mathbb{V}_0)$ and $\mathbb{Q}' = \mathbb{C}_1\mathbb{M}_1/(\mathbb{C}_0\mathbb{M}_0)$. In the Online Supplement, we present experimental results for our test processes showing that although each overlapping variance estimator required slightly more computing time than $\mathcal{N}(b, m)$ required, this cost difference was offset by significant reductions in variance and mean-squared error; and thus relative to $\mathcal{N}(b, m)$, each of the proposed overlapping variance estimators achieved net efficiency improvements (that is, $\mathbb{Q} < 1$ and $\mathbb{Q}' < 1$) in our test processes. In almost all practical applications of simulation, the basic computing cost to execute the simulation model and to

Table 2 **Estimated Expected Values of Variance Estimators for an AR(1) Process with $\phi = 0.9$, $\sigma^2 = 19$, and $b = 20$**

	Nonoverlapping variance estimators					
m	$\mathcal{A}(f_0; b, m)$	$\mathcal{A}(f_2; b, m)$	$\mathcal{A}(f_{\cos,1}; b, m)$	$\mathcal{C}(g_0; b, m)$	$\mathcal{C}(g_2^\star; b, m)$	$\mathcal{N}(b, m)$
100	13.80	14.87	14.68	11.85	14.83	17.10
300	17.18	18.29	18.30	16.20	18.25	18.41
500	17.89	18.73	18.74	17.28	18.70	18.65
1,000	18.47	18.91	18.91	18.14	18.91	18.82
	Overlapping variance estimators					
m	$\mathcal{A}^0(f_0; b, m)$	$\mathcal{A}^0(f_2; b, m)$	$\mathcal{A}^0(f_{\cos,1}; b, m)$	$\mathcal{C}^0(g_0; b, m)$	$\mathcal{C}^0(g_2^\star; b, m)$	$\mathcal{O}(b, m)$
100	13.80	14.92	14.74	11.86	14.87	17.09
300	17.22	18.34	18.33	16.23	18.29	18.38
500	17.94	18.77	18.78	17.30	18.74	18.67
1,000	18.51	18.98	18.98	18.15	18.96	18.85

generate the output process significantly exceeds the additional cost to apply *any* reasonable variance estimator to the resulting time series; hence there is good reason to expect that in routine use, all the proposed overlapping variance estimators will achieve substantial efficiency improvements relative to the NBM variance estimator. Thus we focus the remaining discussion on the mean and variance of the estimators under consideration. ◁

In Tables 2 and 3, we provide the estimated expected values and variances of several nonoverlapping and overlapping variance estimators for the AR(1) process with $\phi = 0.9$, corresponding to a highly positive autocorrelation structure and variance parameter $\sigma^2 = 19$. Tables 4 and 5 present the estimated expected values and variances of the nonoverlapping and overlapping variance estimators for the $M/M/1$ queue-waiting-time process with arrival rate 0.8 and service rate 1.0 (so that $\rho = 0.8$), corresponding to a highly positive autocorrelation structure and variance parameter $\sigma^2 = 1{,}976$. All entries in Tables 2–5 are based on 10,000 independent replications using $b = 20$ batches. Based on the entries of Table 3, the standard errors of the point estimates in Table 2 have an upper bound of about 0.06. Similarly, given the entries of Table 5, the standard errors of the estimates in Table 4 have an upper bound of about 16. In Tables 3 and 5, the rows labeled "$m \to \infty$" provide the asymptotic variances of the variance estimators obtained analytically in Alexopoulos et al. (2006b) and summarized in Table 1. We used common random numbers across all variance estimation methods based on the combined generator given in Figure 1 of L'Ecuyer (1999) for random number generation.

We summarize our conclusions from Tables 2 and 4 regarding the expected values of the variance estimators under consideration as follows:

- The expected values of all variance estimators converge to σ^2 as m increases, in accordance with our theoretical results.
- The overlapping versions of the various area, CvM, and batch means estimators have expected

Table 3 **Estimated Variances of Variance Estimators for an AR(1) Process with $\phi = 0.9$ and $b = 20$**

	Nonoverlapping variance estimators					
m	$\mathcal{A}(f_0; b, m)$	$\mathcal{A}(f_2; b, m)$	$\mathcal{A}(f_{\cos,1}; b, m)$	$\mathcal{C}(g_0; b, m)$	$\mathcal{C}(g_2^\star; b, m)$	$\mathcal{N}(b, m)$
100	19.72	22.21	21.60	8.39	19.08	30.62
300	30.00	33.55	33.39	12.50	27.99	35.84
500	31.85	35.44	35.58	13.24	29.65	37.04
1,000	33.84	35.52	35.60	13.61	29.96	37.10
$\to \infty$	36.10	36.10	36.10	14.44	31.20	38.00
	Overlapping variance estimators					
m	$\mathcal{A}^0(f_0; b, m)$	$\mathcal{A}^0(f_2; b, m)$	$\mathcal{A}^0(f_{\cos,1}; b, m)$	$\mathcal{C}^0(g_0; b, m)$	$\mathcal{C}^0(g_2^\star; b, m)$	$\mathcal{O}(b, m)$
100	9.25	9.91	9.24	5.12	9.91	24.19
300	12.33	14.43	13.90	7.40	13.93	25.05
500	12.43	14.96	14.46	7.57	14.10	25.88
1,000	12.89	15.34	14.85	7.79	14.53	25.60
$\to \infty$	12.83	15.37	14.89	7.83	14.56	25.56

Table 4 **Estimated Expected Values of Variance Estimators for an $M/M/1$ Waiting-Time Process with $\rho = 0.8$, $\sigma^2 = 1{,}976$, and $b = 20$**

	Nonoverlapping variance estimators					
m	$\mathscr{A}(f_0; b, m)$	$\mathscr{A}(f_2; b, m)$	$\mathscr{A}(f_{\cos,1}; b, m)$	$\mathscr{C}(g_0; b, m)$	$\mathscr{C}(g_2^\star; b, m)$	$\mathscr{N}(b, m)$
512	1,430	1,521	1,497	1,221	1,511	1,761
1,024	1,660	1,792	1,788	1,521	1,779	1,860
2,048	1,839	1,960	1,961	1,755	1,946	1,950
4,096	1,939	2,001	2,000	1,883	1,994	1,964
8,192	1,948	1,969	1,966	1,916	1,969	1,964
16,384	1,965	1,969	1,967	1,948	1,971	1,977
32,768	1,975	1,974	1,972	1,963	1,975	1,972
	Overlapping variance estimators					
m	$\mathscr{A}^0(f_0; b, m)$	$\mathscr{A}^0(f_2; b, m)$	$\mathscr{A}^0(f_{\cos,1}; b, m)$	$\mathscr{C}^0(g_0; b, m)$	$\mathscr{C}^0(g_2^\star; b, m)$	$\mathscr{O}(b, m)$
512	1,420	1,518	1,496	1,227	1,522	1,764
1,024	1,668	1,805	1,798	1,532	1,785	1,851
2,048	1,850	1,950	1,949	1,765	1,946	1,952
4,096	1,930	1,998	1,998	1,883	1,991	1,970
8,192	1,935	1,968	1,968	1,917	1,967	1,968
16,384	1,958	1,976	1,976	1,946	1,975	1,978
32,768	1,971	1,978	1,978	1,966	1,972	1,982

values that are about the same as those of their corresponding nonoverlapping versions.

• For both the nonoverlapping and overlapping versions of the area estimator, the weighting functions $f_2(\cdot)$ and $f_{\cos,1}(\cdot)$ yield similar expected values of the variance estimator for a given value of m, and they provide expected values that converge to σ^2 faster than those from the weighting function $f_0(\cdot)$ as m increases. This is a consequence of the relatively high biases of $\mathscr{A}(f_0; b, m)$ and $\mathscr{A}^{\mathrm{O}}(f_0; b, m)$; see Table 1.

• For both the nonoverlapping and overlapping versions of the CvM estimator, the weighting function $g_2^\star(\cdot)$ provides expected values that converge to σ^2 faster than those from the weighting function $g_0(\cdot)$ as m grows. Note also that among all variance estimators considered, $\mathscr{C}(g_0; b, m)$ and $\mathscr{C}^{\mathrm{O}}(g_0; b, m)$ are

Table 5 **Estimated Variances of Variance Estimators for an $M/M/1$ Waiting-Time Process with $\rho = 0.8$ and $b = 20$**

	Nonoverlapping variance estimators ($\times 10^6$)					
m	$\mathscr{A}(f_0; b, m)$	$\mathscr{A}(f_2; b, m)$	$\mathscr{A}(f_{\cos,1}; b, m)$	$\mathscr{C}(g_0; b, m)$	$\mathscr{C}(g_2^\star; b, m)$	$\mathscr{N}(b, m)$
512	1.472	1.378	1.280	0.694	1.310	2.651
1,024	1.638	1.841	1.808	0.849	1.540	1.866
2,048	1.473	1.873	1.946	0.843	1.441	1.491
4,096	1.206	1.180	1.178	0.658	0.939	0.936
8,192	0.810	0.757	0.762	0.423	0.625	0.660
16,384	0.606	0.568	0.570	0.299	0.475	0.554
32,768	0.489	0.478	0.482	0.223	0.400	0.477
$\to \infty$	0.390	0.390	0.390	0.156	0.337	0.411
	Overlapping variance estimators ($\times 10^6$)					
m	$\mathscr{A}^0(f_0; b, m)$	$\mathscr{A}^0(f_2; b, m)$	$\mathscr{A}^0(f_{\cos,1}; b, m)$	$\mathscr{C}^0(g_0; b, m)$	$\mathscr{C}^0(g_2^\star; b, m)$	$\mathscr{O}(b, m)$
512	0.967	0.977	0.898	0.561	1.030	2.644
1,024	1.023	1.260	1.222	0.709	1.195	1.762
2,048	0.954	1.107	1.095	0.736	1.122	1.367
4,096	0.642	0.793	0.788	0.529	0.727	0.820
8,192	0.380	0.429	0.424	0.314	0.418	0.532
16,384	0.266	0.299	0.294	0.211	0.290	0.415
32,768	0.205	0.235	0.229	0.153	0.229	0.342
$\to \infty$	0.138	0.166	0.161	0.084	0.157	0.276

the most biased, an observation that is consistent with Table 1.

• For small values of m, both the nonoverlapping and overlapping versions of the batch-means estimator yield expected values that are superior to those of the area and CvM estimators. However, in terms of expected values, the first-order unbiased versions of the STS variance estimators quickly "catch up" to the batch-means estimators as m gets larger.

From Tables 3 and 5, we obtain the following conclusions regarding the variances of the variance estimators under consideration:

• The estimated variances for the AR(1) process converge to the theoretical asymptotic values quite quickly as m increases for all variance estimators. Although the estimated variances also appear to converge eventually to the theoretical asymptotic values for the $M/M/1$ process as m grows, the convergence is much slower than that for the AR(1) process. Moreover, the variance estimates for the $M/M/1$ process first increase and then decrease with m. This behavior, first observed in Sargent et al. (1992), is also seen in Table 4 not only for the nonoverlapping estimators $\mathscr{A}(f_2; b, m)$, $\mathscr{A}(f_{\cos,1}; b, m)$, and $\mathscr{C}(g_2^\star; b, m)$ but also for the overlapping variance estimators $\mathscr{A}^{\mathrm{O}}(f_2; b, m)$, $\mathscr{A}^{\mathrm{O}}(f_{\cos,1}; b, m)$, and $\mathscr{C}^{\mathrm{O}}(g_2^\star; b, m)$.

• Among all nonoverlapping variance estimators considered, $\mathscr{C}(g_0; b, m)$ and $\mathscr{N}(b, m)$ have the smallest and largest variances, respectively. Similarly, among all overlapping variance estimators considered, $\mathscr{C}^{\mathrm{O}}(g_0; b, m)$ and $\mathscr{O}(b, m)$ have the smallest and largest variances, respectively.

• The overlapping versions of all variance estimators under consideration provide smaller variances than the respective nonoverlapping versions. In the given test problems, the variance reductions ranged from 20% to 60%.

In summary, this empirical study validates the theoretical asymptotic results obtained in Alexopoulos et al. (2006b) and summarized in Table 1 of this paper. In particular, our numerical experiments support the conclusion that the overlapping versions of the various area, CvM, and batch-means variance estimators have similar expected values but smaller variances than their nonoverlapping counterparts.

4.2. Density Estimation

Beyond the point estimates of the mean and variance of the nonoverlapping and overlapping variance estimators discussed in Section 4.1, we estimated the limiting distributions of all these statistics. To this end, we generated 1,000,000 independent sample paths of a stationary AR(1) process with $\phi = 0.9$, where each sample path contained $n = 20{,}000$ observations. We computed all variance estimators using a batch size of $m = 1{,}000$, which we believe is sufficiently large in this case to ensure approximate convergence of these statistics to their limiting distributions. Figure 1 displays the resulting empirical distributions.

In each of the six subplots of Figure 1, we superimposed the empirical probability density functions (p.d.f.'s) of the corresponding nonoverlapping and overlapping versions of each variance estimator. We represented each empirical p.d.f. by a frequency polygon as detailed in Section 3.3 of Hald (1952). The empirical p.d.f.'s reflect the theoretical moment properties summarized in Table 1. For example, each subplot indicates that the estimator based on overlapping batches has smaller variance than the corresponding estimator based on nonoverlapping batches.

Examination of Figure 1 naturally suggests some basic questions:

• What are the theoretical distributions of $\mathscr{A}^{\mathrm{O}}(f; b, m)$ and $\mathscr{C}^{\mathrm{O}}(g; b, m)$ for given values of m, provided that m is sufficiently large?

• Can we formulate approximations to the theoretical distributions of $\mathscr{A}^{\mathrm{O}}(f; b, m)$ and $\mathscr{C}^{\mathrm{O}}(g; b, m)$ for sufficiently large m that are useful in practical simulation output analysis—for example in constructing confidence intervals on μ and σ^2?

• Can we formulate accurate approximations to the limiting distributions of $\mathscr{A}^{\mathrm{O}}(f; b, m)$ and $\mathscr{C}^{\mathrm{O}}(g; b, m)$ as $m \to \infty$?

Our previous experience with the NBM algorithm ASAP3 (Steiger et al. 2005) strongly suggested the following approach to answering these questions.

If $\{\lambda_j\colon\ j = 1, \ldots, L\}$ denote the positive eigenvalues of the coefficient matrix associated with the quadratic form $\mathscr{A}^{\mathrm{O}}(f; b, m)$ and if m is sufficiently large to ensure adequate convergence to the key limiting result (12) of Alexopoulos et al. (2006b), then it follows from Theorem 2.1 of Box (1954) that

$$\mathscr{A}^{\mathrm{O}}(f; b, m) \dot{\sim} \sum_{j=1}^{L} \lambda_j \chi^2_{1,j}, \tag{9}$$

where the symbol $\dot{\sim}$ means "is approximately distributed as" and $\{\chi^2_{1,j}\colon\ j = 1, \ldots, L\}$ are i.i.d. χ^2 variates each with one degree of freedom. The technique of Satterthwaite (1941) can be used to approximate the distribution of any complex variance estimator that is a weighted sum of independent chi-squared random variables. Since (9) is such a sum, we can apply Satterthwaite's method to obtain the distribution approximation

$$\mathscr{A}^{\mathrm{O}}(f; b, m) \dot{\sim} \mathrm{E}[\mathscr{A}^{\mathrm{O}}(f; b, m)]\chi^2_{\nu_{\mathrm{eff}}}/\nu_{\mathrm{eff}},$$
$$\text{where } \nu_{\mathrm{eff}} = \left[\!\!\left[\frac{2\mathrm{E}^2[\mathscr{A}^{\mathrm{O}}(f; b, m)]}{\mathrm{Var}[\mathscr{A}^{\mathrm{O}}(f; b, m)]} \right]\!\!\right], \tag{10}$$

and $[\![z]\!]$ denotes rounding of z towards the nearest integer. The quantity ν_{eff} is called the "effective" degrees of freedom in (10).

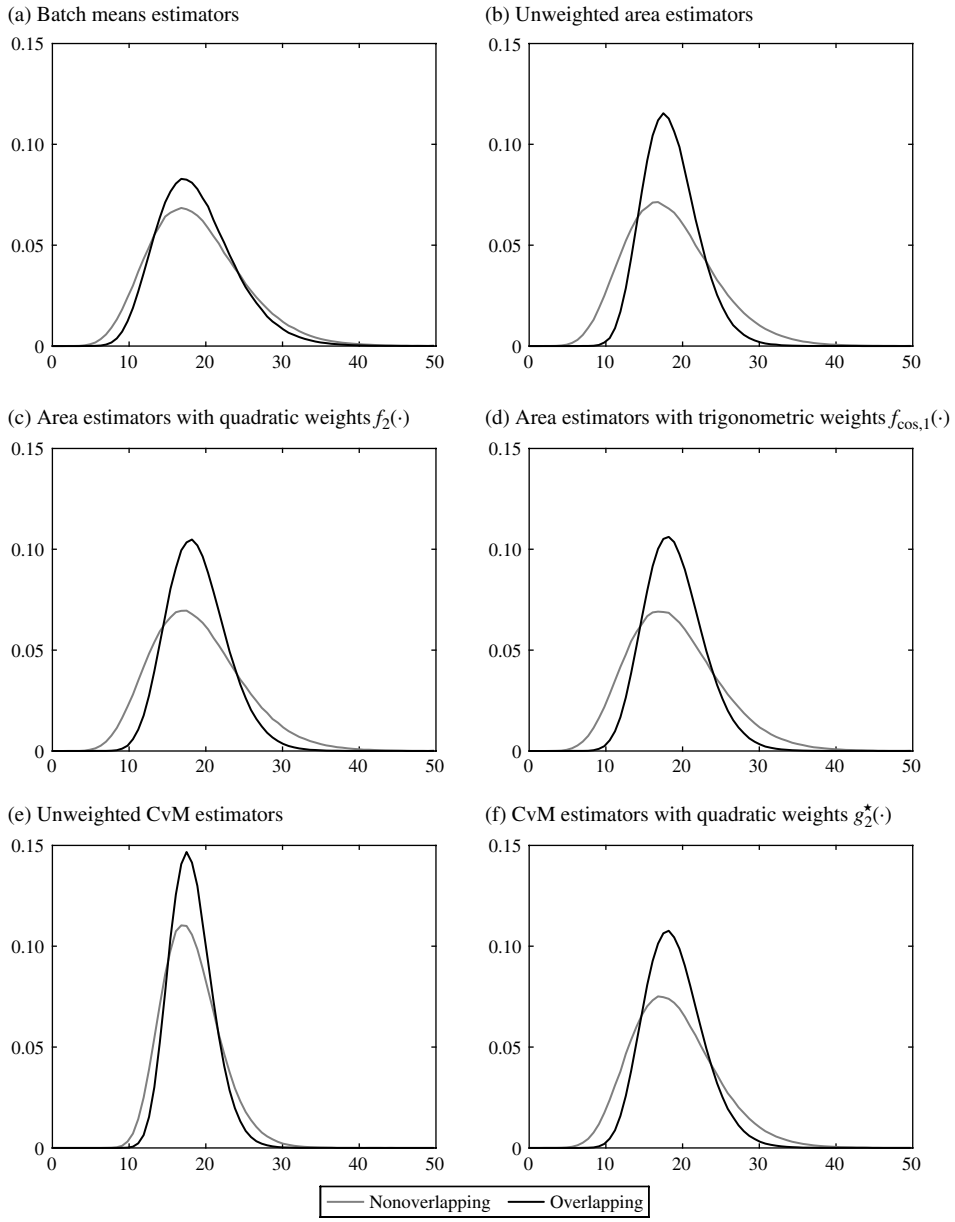

Figure 1 **Estimated p.d.f.'s for Several Variance Estimators Computed from 1,000,000 Replications of a Stationary AR(1) Process with $\phi = 0.9$, Batch Size $m = 1{,}000$, and $b = 20$**

Numerical evidence of the accuracy of approximations such as (10) can be found in Satterthwaite (1941) and Welch (1956). A noteworthy conclusion of these papers is that the approximation is most accurate when the complex variance estimator under discussion involves only independent chi-squared summands with *positive* weights—and this is precisely the situation described by (9). We can also obtain similar results for $\mathscr{C}^{\mathrm{O}}(g; b, m)$ when m is sufficiently large. Below we test the validity of approximations like (10) by examining the empirical p.d.f.'s of three STS estimators based on an AR(1) process with $\phi = 0.9$.

In the case of the overlapping area estimator with a constant weighting function $f_0(\cdot)$ as detailed in Table 1, we see that

$$\mathrm{E}[\mathscr{A}^{\mathrm{O}}(f_0; b, m)] \approx \sigma^2 + 3\gamma/m \quad \text{and}$$
$$\nu_{0,\,\mathrm{eff}} = \left[\!\!\left[\frac{(\sigma^2 + 3\gamma/m)^2 70(b-1)^2}{(24b-31)\sigma^4} \right]\!\!\right], \tag{11}$$

so that we have the approximation

$$\mathscr{A}^{\mathrm{O}}(f_0; b, m) \dot{\sim} (\sigma^2 + 3\gamma/m)\chi^2_{\nu_{0,\,\mathrm{eff}}}/\nu_{0,\,\mathrm{eff}}. \tag{12}$$

Figure 2(a) displays the empirical p.d.f. of $\mathscr{A}^{\mathrm{O}}(f_0; 20, 1{,}000)$ and the fitted p.d.f. based on (11) and (12) with $b = 20$ and $m = 1{,}000$ so that $\nu_{0,\,\mathrm{eff}} = 53$. In this case the distribution of an appropriately scaled chi-squared

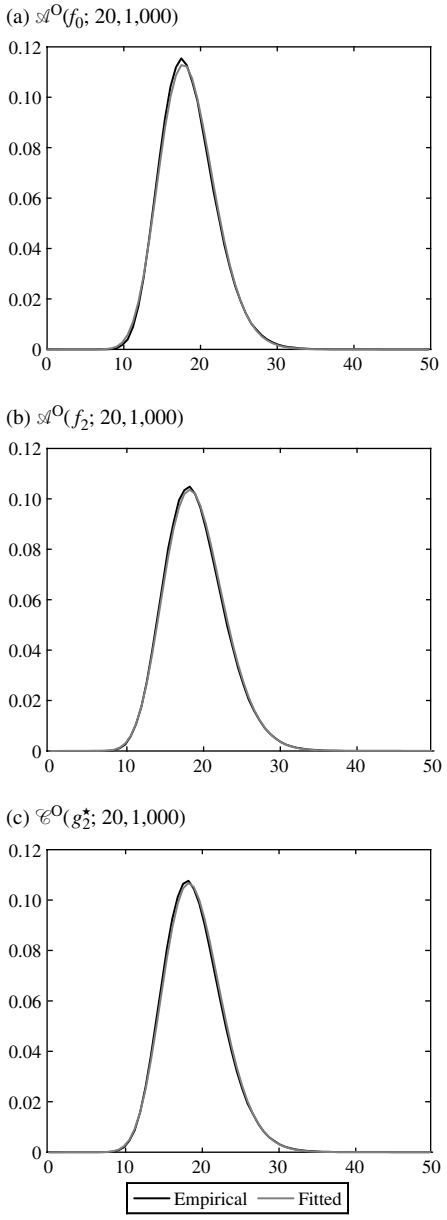

Figure 2 **Empirical and Fitted p.d.f.'s for Three STS Estimators Computed from an AR(1) Process with $\phi = 0.9$**

random variable appeared to provide an excellent approximation to the distribution of $\mathcal{A}^{O}(f_0; 20, 1{,}000)$.

In the case of the overlapping area estimator with the first-order unbiased quadratic weighting function $f_2(\cdot)$ as detailed in Table 1, we see that

$$\mathrm{E}[\mathcal{A}^{O}(f_2; b, m)] \approx \sigma^2 \quad \text{and} \quad \nu_{2,\mathrm{eff}} = \left[\!\!\left[\frac{8{,}580(b-1)^2}{3{,}514b - 4{,}359} \right]\!\!\right], \tag{13}$$

so that we have

$$\mathcal{A}^{O}(f_2; b, m) \dot{\sim} \sigma^2 \chi^2_{\nu_{2,\mathrm{eff}}}/\nu_{2,\mathrm{eff}}. \tag{14}$$

Figure 2(b) displays the empirical p.d.f. of $\mathcal{A}^{O}(f_2; 20, 1{,}000)$ and the fitted p.d.f. based on (13) and (14) with $b = 20$ and $m = 1{,}000$ so that $\nu_{2,\mathrm{eff}} = 47$. Again we believe that we obtained an excellent approximation to the target distribution of an overlapping variance estimator.

As a final example, we consider the overlapping CvM estimator with the weighting function $g_2^{\star}(\cdot)$ that is first-order unbiased with minimum variance over all quadratic weights as detailed in Table 1. We have

$$\mathrm{E}[\mathcal{C}^{O}(g_2^{\star}; b, m)] \approx \sigma^2 \quad \text{and} \quad \nu^{\star}_{2,\mathrm{eff}} = \left[\!\!\left[\frac{27{,}720(b-1)^2}{10{,}768b - 13{,}605} \right]\!\!\right] \tag{15}$$

so that we have the approximation

$$\mathcal{C}^{O}(g_2^{\star}; b, m) \dot{\sim} \sigma^2 \chi^2_{\nu^{\star}_{2,\mathrm{eff}}}/\nu^{\star}_{2,\mathrm{eff}}. \tag{16}$$

Figure 2(c) displays the empirical p.d.f. of $\mathcal{C}^{O}(g_2^{\star}; 20, 1{,}000)$ and the fitted p.d.f. based on (15) and (16) with $b = 20$ and $m = 1{,}000$ so that $\nu^{\star}_{2,\mathrm{eff}} = 50$. As in the two preceding cases, we concluded that we obtained an excellent approximation to the target distribution, which in this case characterizes the behavior of the CvM variance estimator $\mathcal{C}^{O}(g_2^{\star}; 20, 1{,}000)$.

Remark 4. We have

$$\mathcal{A}^{O}(f; b, m) \underset{m\to\infty}{\overset{\mathcal{D}}{\to}} \mathcal{A}^{O}(f; b)$$

as in (18) of Alexopoulos et al. (2006b); and m is sufficiently large to ensure

$$\mathcal{A}^{O}(f; b, m) \dot{\sim} \mathcal{A}^{O}(f; b) \quad \text{and} \quad \mathrm{E}\{[\mathcal{A}^{O}(f; b, m)]^p\} \approx \mathrm{E}\{[\mathcal{A}^{O}(f; b)]^p\} \quad \text{for } p = 1, 2, \tag{17}$$

then from (10) and (17) we also obtain an approximation to the asymptotic distribution of the overlapping area estimator with weighting function $f(\cdot)$,

$$\mathcal{A}^{O}(f; b) \dot{\sim} \sigma^2 \chi^2_{\nu_{\mathrm{eff}}}/\nu_{\mathrm{eff}}, \quad \text{where} \quad \nu_{\mathrm{eff}} = \left[\!\!\left[\frac{2\sigma^4}{\mathrm{Var}[\mathcal{A}^{O}(f; b)]} \right]\!\!\right]. \tag{18}$$

An analogous statement to (18) can be made with regard to the CvM variance estimators. The empirical results displayed in Figure 2 strongly suggest that results such as (18) can provide excellent approximations to the theoretical limiting distributions of overlapping area and CvM variance estimators. ◁

4.3. Confidence Intervals

The distributions of all the overlapping area and CvM variance estimators that we examined were remarkably well approximated by the distributions of scaled χ^2 variates with the appropriate degrees of freedom—at least for the case of an AR(1) process with $b = 20$ and $m = 1{,}000$ as described in Section 4.2; and

these results immediately suggested trying to construct approximate confidence intervals (CIs) for the parameters μ and σ^2. For example, if $\mathcal{A}^{\mathrm{O}}(f; b, m)$ is a first-order unbiased estimator for σ^2, then it follows immediately from (10) that for $\alpha \in (0, 1)$ and a sufficiently large batch size m, an approximate $100(1-\alpha)\%$ two-sided CI for σ^2 is given by

$$\frac{\nu_{\mathrm{eff}}\mathcal{A}^{\mathrm{O}}(f; b, m)}{\chi^2_{1-\alpha/2,\,\nu_{\mathrm{eff}}}} \le \sigma^2 \le \frac{\nu_{\mathrm{eff}}\mathcal{A}^{\mathrm{O}}(f; b, m)}{\chi^2_{\alpha/2,\,\nu_{\mathrm{eff}}}}, \qquad (19)$$

where $\chi^2_{\beta,\,\nu}$ denotes the β-quantile of the χ^2_ν distribution.

Example 1. Consider the overlapping area variance estimator using the quadratic weighting function $f_2(\cdot)$, for which (13) and (14) yield $\mathcal{A}^{\mathrm{O}}(f_2; 20, m) \dot{\sim} \sigma^2\chi^2_{47}/47$ for sufficiently large m. We used the 1,000,000 independent realizations of $\mathcal{A}^{\mathrm{O}}(f_2; 20, 1{,}000)$ computed from the AR(1) process with $\phi = 0.9$ as described in Section 4.2 to estimate the coverage probability of the two-sided 90% CI (19) for σ^2, which in this case had the specific form

$$0.7343\mathcal{A}^{\mathrm{O}}(f_2; 20, 1{,}000) \le \sigma^2 \le 1.4566\mathcal{A}^{\mathrm{O}}(f_2; 20, 1{,}000), \qquad (20)$$

where we recall that the true value of σ^2 is 19. The empirical coverage probability of the CI defined by (20) was 0.9013. Even though $\mathcal{A}^{\mathrm{O}}(f_0; 20, 1{,}000)$ is a biased estimator of σ^2 as seen in (11), we obtained an empirical coverage of 0.9045 for the "naive" two-sided nominal 90% CI (19) for σ^2 based on $\mathcal{A}^{\mathrm{O}}(f_0; 20, 1{,}000)$ with $\nu_{\mathrm{eff}} = 53$. Similarly we obtained an empirical coverage of 0.9006 for the nominal 90% confidence interval that is analogous to (19) based on $\mathcal{C}^{\mathrm{O}}(g_2^\star; 20, 1{,}000)$ with $\nu_{\mathrm{eff}} = 50$. From a practical standpoint, the empirical coverage probabilities achieved by the CIs based on each of the three variance estimators did not differ substantially from the nominal coverage probability. ◁

We can also obtain an approximate CI for the mean μ, provided the batch size m is sufficiently large (so that the sample size n is also sufficiently large). First of all, (12) from Alexopoulos et al. (2006b) implies that for large n, we have (i) $\overline{Y}_n \dot{\sim} \mathrm{Nor}(\mu, \sigma^2/n)$; and (ii) $\overline{Y}_n$ is approximately independent of the standardized time series formed from $\{Y_i\colon i = 1, \ldots, n\}$. It follows that for m sufficiently large, the pivot

$$\frac{\overline{Y}_n - \mu}{\sqrt{\mathcal{A}^{\mathrm{O}}(f; b, m)/n}} \dot{\sim} t_{\nu_{\mathrm{eff}}}, \qquad (21)$$

where ν_{eff} is given by (10) and t_ν denotes a random variable having Student's t-distribution with ν degrees of freedom. Then standard manipulations of (21) yield an approximate $100(1-\alpha)\%$ two-sided CI for μ,

$$\overline{Y}_n - t_{1-\alpha/2,\,\nu_{\mathrm{eff}}}\sqrt{\mathcal{A}^{\mathrm{O}}(f; b, m)/n} \le \mu \le \overline{Y}_n + t_{1-\alpha/2,\,\nu_{\mathrm{eff}}}\sqrt{\mathcal{A}^{\mathrm{O}}(f; b, m)/n}, \qquad (22)$$

where $t_{\beta,\,\nu}$ denotes the β-quantile of a Student's t random variable with ν degrees of freedom.

Example 1 (Continued). Consider the overlapping area variance estimator using the quadratic weighting function $f_2(\cdot)$, for which (13) and (14) yield the following specific version of the approximation (21), $(\overline{Y}_n - \mu)/\sqrt{\mathcal{A}^{\mathrm{O}}(f_2; 20, 1{,}000)/n} \dot{\sim} t_{47}$. We used the 1,000,000 independent realizations of $\overline{Y}_n$ and $\mathcal{A}^{\mathrm{O}}(f_2; 20, 1{,}000)$ computed from the AR(1) process with $\phi = 0.9$ and $n = 20m = 20{,}000$ as described in Section 4.2 to estimate the coverage probability of the two-sided 90% CI for μ, i.e.,

$$\overline{Y}_n - 0.01187\sqrt{\mathcal{A}^{\mathrm{O}}(f_2; 20, 1{,}000)} \le \mu \le \overline{Y}_n + 0.01187\sqrt{\mathcal{A}^{\mathrm{O}}(f_2; 20, 1{,}000)}, \qquad (23)$$

where we recall that the true value of μ is 0. The empirical coverage probability of the CI based on (23) was 0.8992. Similarly, we obtained an empirical coverage of 0.8949 for the two-sided nominal 90% CI (22) for μ based on $\mathcal{A}^{\mathrm{O}}(f_0; 20, 1{,}000)$; and we obtained an empirical coverage of 0.8990 for the CI that is analogous to (22) based on $\mathcal{C}^{\mathrm{O}}(g_2^\star; 20, 1{,}000)$. From a practical standpoint, the empirical coverage probabilities achieved by the CIs based on each of the three variance estimators did not differ substantially from the nominal coverage probability of 0.9. ◁

5. Summary and Conclusions

The general goal of this paper has been to study estimators for the variance parameter arising from a stationary stochastic process, such as the output processes generated by steady-state simulations. Our intention has been to implement and evaluate the new overlapping variance estimators that were formulated in the theory-oriented paper by Alexopoulos et al. (2006b). For these purposes, we presented efficient algorithms to calculate our estimators; and then we presented a battery of Monte Carlo and exact analyses showing that the new estimators perform as advertised. In particular, we found that the new overlapping estimators almost always had the same bias as, but lower variance than, their nonoverlapping counterparts. In many cases, the overlapping variance estimators had variances (and therefore mean squared errors) that were significantly smaller than those of their nonoverlapping counterparts. And in any case, the new estimators seemed to outperform the con-

ventional variance estimators based on the NBM and OBM methods.

A by-product of our Monte Carlo analysis was our finding that the new estimators could often be well-approximated by properly rescaled χ^2 distributions (with the appropriate degrees of freedom). This finding allows for the construction of confidence intervals for the variance parameter as well as for the underlying mean of the stationary stochastic process.

We saw that the confidence intervals based on the appropriately rescaled χ^2 and Student's t distributions gave nearly nominal coverage, at least when they were applied with large enough batch sizes. There is still some room for further study in this area, as our example Monte Carlo work was only for specific stochastic processes with specific choices of batch size, number of observations, and weighting function. In addition, we feel that the variance estimators studied herein can be adapted for use in a sequential procedure such as ASAP3 (Steiger et al. 2005).

Acknowledgments

The authors thank Byeong-Yun Chang for assistance with some of the computations and the Area Editor for several constructive comments and suggestions. This work was partially supported by National Science Foundation Grant DMI-0400260.

References

Alexopoulos, C., N. T. Argon, D. Goldsman, G. Tokol. 2004. Overlapping variance estimators for simulations. R. G. Ingalls, M. D. Rossetti, J. S. Smith, B. A. Peters, eds. *Proc. 2004 Winter Simulation Conf.*, Institute of Electrical and Electronics Engineers, Piscataway, NJ, 737–745.

Alexopoulos, C., N. T. Argon, D. Goldsman, N. M. Steiger, G. Tokol, J. R. Wilson. 2006a. Overlapping variance estimators for simulation: Software and user's manual [online]. Technical Report, School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA. Accessed April 17, 2006, ftp://ftp.eos.ncsu.edu/pub/eos/pub/jwilson/ovesie-soft.exe.

Alexopoulos, C., N. T. Argon, D. Goldsman, G. Tokol, J. R. Wilson. 2006b. Overlapping variance estimators for simulation. *Oper. Res.* Forthcoming.

Box, G. E. P. 1954. Some theorems on quadratic forms applied in the study of analysis of variance problems, I. Effect of inequality of variance in the one-way classification. *Ann. Math. Statist.* **25** 290–302.

Foley, R. D., D. Goldsman. 1999. Confidence intervals using orthonormally weighted standardized times series. *ACM Trans. Model. Comput. Simulation* **9** 297–325.

Glynn, P. W., W. Whitt. 1991. Estimating the asymptotic variance with batch means. *Oper. Res. Lett.* **10** 431–435.

Goldsman, D., K. Kang, A. F. Seila. 1999. Cramér–von Mises variance estimators for simulations. *Oper. Res.* **47** 299–309.

Goldsman, D., M. Meketon, L. Schruben. 1990. Properties of standardized time series weighted area variance estimators. *Management Sci.* **36** 602–612.

Hald, A. 1952. *Statistical Theory with Engineering Applications.* John Wiley & Sons, New York.

L'Ecuyer, P. 1999. Good parameters and implementations for combined multiple recursive random number generators. *Oper. Res.* **47** 159–164.

Meketon, M. S., B. Schmeiser. 1984. Overlapping batch means: Something for nothing? S. Sheppard, U. Pooch, D. Pegden, eds. *Proc. 1984 Winter Simulation Conf.*, Institute of Electrical and Electronics Engineers, Piscataway, NJ, 227–230.

Sargent, R. G., K. Kang, D. Goldsman. 1992. An investigation of finite-sample behavior of confidence interval estimators. *Oper. Res.* **40** 898–913.

Satterthwaite, F. E. 1941. Synthesis of variance. *Psychometrika* **6** 309–316.

Schmeiser, B. 1982. Batch size effects in the analysis of simulation output. *Oper. Res.* **30** 556–568.

Schruben, L. 1983. Confidence interval estimation using standardized time series. *Oper. Res.* **31** 1090–1108.

Song, W. T., B. W. Schmeiser. 1995. Optimal mean-squared-error batch sizes. *Management Sci.* **41** 110–123.

Steiger, N. M., J. R. Wilson. 2001. Convergence properties of the batch-means method for simulation output analysis. *INFORMS J. Comput.* **13** 277–293.

Steiger, N. M., E. K. Lada, J. R. Wilson, J. A. Joines, C. Alexopoulos, D. Goldsman. 2005. ASAP3: A batch means procedure for steady-state simulation output analysis. *ACM Trans. Model. Comput. Simulation* **15** 39–73.

Welch, B. L. 1956. On linear combinations of several variances. *J. Amer. Statist. Assoc.* **51** 132–148.