Proceedings of the 2005 Winter Simulation Conference
M. E. Kuhl, N. M. Steiger, F. B. Armstrong, and J. A. Joines, eds.

LINEAR COMBINATIONS OF OVERLAPPING VARIANCE ESTIMATORS FOR SIMULATIONS

Tûba Aktaran-Kalaycı
David Goldsman

School of Industrial and Systems Engineering
Georgia Institute of Technology
Atlanta, GA 30332, U.S.A.

ABSTRACT

We examine properties of linear combinations of overlapping standardized time series area estimators for the variance parameter of a stationary stochastic process. We find that the linear combination estimators have lower bias and variance than their overlapping constituents and nonoverlapping counterparts; in fact, the new estimators also perform particularly well against the benchmark batch means estimator. We illustrate our findings with analytical and Monte Carlo examples.

1 INTRODUCTION

Simulations are used to analyze a variety of complicated stochastic systems. A common goal is to estimate the unknown mean μ of the steady-state output process, $Y_1, Y_2, \ldots, Y_n$. In this case, the sample mean $\bar{Y}_n$ is the estimator of choice for μ; and to provide a measure of the precision (variability) of the sample mean, one might also attempt to estimate the quantity $\sigma_n^2 \equiv n\,\mathrm{Var}(\bar{Y}_n)$, or almost equivalently, the *variance parameter*, $\sigma^2 \equiv \lim_{n\to\infty} \sigma_n^2$. Of course, it is well known that steady-state simulation output is rarely amenable to elementary statistical analysis, for it is almost never independent, identically distributed (i.i.d.), and normal—thus rendering as untrustworthy "standard" estimators such as the sample variance estimator for σ^2.

There are many techniques in the literature concerning the estimation of σ^2. Popular methods include: nonoverlapping batch means (NBM), overlapping batch means (OBM), and standardized time series (STS) (see Law and Kelton 2000 for a quick synopsis). These methodologies all use some form of *batching*, as will be explained in more detail later. Batching tends to increase variance estimator bias (a drawback), but decrease variance (an advantage)—though its effects on mean-squared error require more-careful analysis (cf. Song and Schmeiser 1995). The use of overlapping batches yields variance estimators having the same bias as, but lower variance than, the analogous estimators arising from nonoverlapping batches (Meketon and Schmeiser 1984, Alexopoulos et al. 2005ab).

The present paper discusses a generalization of overlapping variance estimators that can be used in the analysis of steady-state simulations. The article is organized as follows. We present background material in §2, where we introduce a number of benchmark estimators that use traditional batching. §3 reviews properties of certain estimators that employ overlapping batches; and we find that the overlapping estimators almost always outperform their nonoverlapped counterparts in terms of mean-squared error. §4, which contains our main contributions, shows how to apply overlapping batch techniques to the STS area estimator using *different batch sizes* and then to construct a variance-optimal linear combination of these estimators. The resulting estimators have lower bias and variance than their constituent overlapping predecessors. We illustrate these findings on a number of simple analytical and Monte Carlo examples. We end with concluding remarks in §5.

2 BACKGROUND

This section reviews relevant assumptions, definitions, and background results.

2.1 Fundamentals

To get things going, we briefly discuss some basics. First of all, throughout this paper, we shall consider a stationary stochastic process $\{Y_i, i \ge 1\}$, which we assume satisfies a Functional Central Limit Theorem (FCLT).

Assumption FCLT There exist constants μ and positive σ such that as $n \to \infty$, $X_n \Rightarrow \sigma\mathcal{W}$, where $\mathcal{W}$ is a standard Brownian motion process, "$\Rightarrow$" denotes weak convergence as $n \to \infty$, and

$$X_n(t) \equiv \frac{\lfloor nt \rfloor(\bar{Y}_{\lfloor nt \rfloor} - \mu)}{\sqrt{n}} \quad \text{for} \quad t \ge 0,$$

where $\bar{Y}_j \equiv \sum_{k=1}^{j} Y_k/j$, $j = 1, 2, \ldots$, and $\lfloor\cdot\rfloor$ is the greatest integer function.

This assumption applies to a broad class of processes, and will allow us to determine the limiting properties of the various variance estimators under consideration in this paper. Glynn and Iglehart (1990) list several different sets of sufficient conditions for Assumption FCLT to hold—usually in the form of moment and mixing conditions. The constants μ and σ^2 in the assumption can be identified with the process mean and variance parameter, respectively.

In this section, we will work with b contiguous, nonoverlapping batches of observations, each of length m, from the simulation output, $Y_1, Y_2, \ldots, Y_n$ (where we assume that $n = bm$). Obviously, the observations $Y_{(i-1)m+1}, Y_{(i-1)m+2}, \ldots, Y_{im}$ constitute batch i, $i = 1, \ldots, b$.

Schruben (1983) defines the *standardized time series* from batch i as

$$T_{i,m}(t) \equiv \frac{\lfloor mt \rfloor (\bar{Y}_{i,\lfloor mt \rfloor} - \bar{Y}_{i,m})}{\sigma\sqrt{m}},$$

for $0 \le t \le 1$ and $i = 1, \ldots, b$, where

$$\bar{Y}_{i,j} \equiv \frac{1}{j}\sum_{k=1}^{j} Y_{(i-1)m+k},$$

for $j = 1, \ldots, m$ and $i = 1, \ldots, b$. Then we have

Theorem 1 (cf. Alexopoulos et al. 2005b, among others) *Define $Z_i(m) \equiv \sqrt{m}(\bar{Y}_{i,m} - \mu)$, $i = 1, \ldots, b$. Then under Assumption FCLT,*

$$\begin{aligned}(Z_1(m), \ldots, Z_b(m); \sigma T_{1,m}, \ldots, \sigma T_{b,m}) \\ \Rightarrow (\sigma Z_1, \ldots, \sigma Z_b; \sigma\mathcal{B}_{0,1}, \ldots, \sigma\mathcal{B}_{b-1,1}),\end{aligned} \quad (1)$$

where the Z_i's are i.i.d. standard normal random variables, and $\mathcal{B}_{u,v}$ denotes a Brownian bridge process on $[u, u+v]$, for $u \in [0, b-v]$, i.e.,

$$\mathcal{B}_{u,v}(t) = \frac{\mathcal{W}(u+tv) - \mathcal{W}(u) - t[\mathcal{W}(u+v) - \mathcal{W}(u)]}{\sqrt{v}},$$

for $0 \le t \le 1$. It is easy to see that $\mathcal{B}_{0,1}, \ldots, \mathcal{B}_{b-1,1}$ are independent Brownian bridges.

2.2 Batched Area Estimator

This subsection deals with the (nonoverlapping) batched area estimator for σ^2 (Goldsman et al. 1990; Goldsman and Schruben 1990).

We will work with the square of the weighted area under the standardized time series from the ith batch,

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

and its limiting functional

$$A_i(f) \equiv \left[\int_0^1 f(t)\sigma\mathcal{B}_{i-1,1}(t)\,dt\right]^2,$$

for $i = 1, \ldots, b$, where $f(t)$ is continuous on the interval $[0, 1]$ and normalized so that $\text{Var}(\int_0^1 f(t)\mathcal{B}_{0,1}(t)\,dt) = 1$. Under mild conditions (see Alexopoulos et al. 2005b), one can show that $A_i(f; m) \xrightarrow{\mathcal{D}} A_i(f)$, $i = 1, \ldots, b$, where "$\xrightarrow{\mathcal{D}}$" denotes convergence in distribution as $m \to \infty$; and further, $A_1(f), \ldots, A_b(f)$ are i.i.d. $\sigma^2\chi_1^2$. This result motivates construction of the *batched area estimator* for σ^2,

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

Denote the covariance function $R_k \equiv \text{Cov}(Y_1, Y_{1+k})$, $k = 0, \pm1, \pm2, \ldots$, and the associated quantity $\gamma \equiv -2\sum_{k=1}^{\infty} kR_k$ (cf. Song and Schmeiser 1995). In addition, the notation $p(n) = o(q(n))$ means that $p(n)/q(n) \to 0$ as $n \to \infty$. The next theorem gives the expected value and variance of the area estimator.

Theorem 2 (see, e.g., Foley and Goldsman 1999) *Suppose that $\{Y_i, i \ge 1\}$ is a stationary process for which Assumption FCLT holds, $\sum_{k=1}^{\infty} k^2|R_k| < \infty$, and $\sigma^2 > 0$. Further, suppose that $A^2(f; b, m)$ is uniformly integrable (cf. Billingsley 1968). If we define $F^\star \equiv [(F - \bar{F})^2 + \bar{F}^2]/2$, then*

$$\text{E}[\mathcal{A}(f; b, m)] = \sigma^2 + \frac{F^\star\gamma}{m} + o(1/m)$$

and

$$\text{Var}(\mathcal{A}(f; b, m)) \to \text{Var}(\sigma^2\chi_b^2/b) = 2\sigma^4/b$$

as $m \to \infty$, where the quantities $F \equiv \int_0^1 f(t)\,dt$ and $\bar{F} \equiv \int_0^1\int_0^t f(s)\,ds\,dt$. Note that the limiting variance $2\sigma^4/b$ does not depend on the form of the weighting function.

Example 1 Schruben (1983) studied the area estimator with constant weighting function $f_0(t) \equiv \sqrt{12}$, for $0 \le t \le 1$. In this case, Theorem 2 implies that $\text{E}[A(f_0; b, m)] = \sigma^2 + 3\gamma/m + o(1/m)$.

Example 2 If one chooses weights having $F = \bar{F} = 0$, the resulting estimator is *first-order unbiased* for σ^2, i.e., its bias is $o(1/m)$. An example of such a weighting function is the quadratic $f_2(t) \equiv \sqrt{840}(3t^2 - 3t + 1/2)$ (Goldsman et al. 1990; Goldsman and Schruben 1990).

Example 3 Foley and Goldsman (1999) give an "orthonormal" sequence of first-order unbiased weights, $f_{\cos,j}(t) = \sqrt{8}\pi j \cos(2\pi j t)$, $j = 1, 2, \ldots$. It can be shown that the orthonormal estimators' limiting functionals $A_i(f_{\cos,1}), A_i(f_{\cos,2}), \ldots$ are i.i.d. $\sigma^2\chi_1^2$.

2.3 NBM Estimator

The quantities $\bar{Y}_{i,m}$, $i = 1, \ldots, b$, are referred to as the *batch means* of the process $\{Y_i\}$, and are often assumed to be i.i.d. normal random variables, at least for large enough batch size m. This assumption immediately suggests the *NBM estimator* for σ^2,

$$\mathcal{N}(b, m) \equiv \frac{m}{b-1}\sum_{i=1}^{b}(\bar{Y}_{i,m} - \bar{Y}_n)^2 \xrightarrow{\mathcal{D}} \frac{\sigma^2\chi^2_{b-1}}{b-1},$$

as $m \to \infty$ with b fixed (cf. Glynn and Whitt 1991, Schmeiser 1982, and Steiger and Wilson 2001). The NBM estimator is one of the most popular for σ^2, and serves as a benchmark for comparison with other estimators.

Theorem 3 (Chien et al. 1997, Goldsman and Meketon 1986, Song and Schmeiser 1995, among others) *Under mild conditions,*

$$\mathrm{E}[\mathcal{N}(b, m)] = \sigma^2 + \frac{\gamma(b+1)}{bm} + o(1/m).$$

Further, for fixed b,

$$\lim_{m\to\infty}(b-1)\,\mathrm{Var}(\mathcal{N}(b, m)) = 2\sigma^4.$$

So we see from Theorems 2 and 3 that as $m \to \infty$, the STS batched area and NBM estimators are all asymptotically unbiased for σ^2; and the batched area estimator with certain weighting functions can outperform NBM in terms of first-order bias. We also notice that the variances of these estimators are all about equal and inversely proportional to the number of batches (at least for sufficiently large batch size). In the next section, we will show that the use of overlapping batches with respect to any particular estimator preserves its expected value, while reducing its variance.

3 OVERLAPPING ESTIMATORS

Here we review estimators based on *overlapping* batches, first proposed by Meketon and Schmeiser (1984) in the context of OBM estimators. See Goldsman and Meketon (1986), Welch (1987), Song (1988), Damerdji (1991, 1994, 1995), Song and Schmeiser (1993), Pedrosa and Schmeiser (1993, 1994), and Alexopoulos et al. (2005ab) for additional discussions.

3.1 Overlapping Fundamentals

Suppose we have n observations $Y_1, Y_2, \ldots, Y_n$ on hand and that we form $n - m + 1$ overlapping batches, each of size m. In particular, the observations $Y_i, Y_{i+1}, \ldots, Y_{i+m-1}$ constitute batch i, $i = 1, \ldots, n - m + 1$. We will continue to denote $b \equiv n/m$ as before, though b can no longer be interpreted as "the number of batches."

To parallel the discussion in §2.1, the standardized time series from overlapping batch i is

$$T^{\mathrm{O}}_{i,m}(t) \equiv \frac{\lfloor mt\rfloor\left(\bar{Y}^{\mathrm{O}}_{i,\lfloor mt\rfloor} - \bar{Y}^{\mathrm{O}}_{i,m}\right)}{\sigma\sqrt{m}},$$

for $0 \le t \le 1$ and $i = 1, \ldots, n - m + 1$, where

$$\bar{Y}^{\mathrm{O}}_{i,j} \equiv \frac{1}{j}\sum_{k=0}^{j-1} Y_{i+k},$$

for $i = 1, \ldots, n - m + 1$ and $j = 1, \ldots, m$. Under the same mild conditions as before,

$$\sigma T^{\mathrm{O}}_{\lfloor um\rfloor,m} \Rightarrow \sigma\mathcal{B}_{u,1}, \quad 0 \le u \le b-1, \;\; u \text{ fixed}.$$

3.2 Overlapping Area Estimator

The square of the weighted area under the standardized time series from the ith overlapping batch is

$$A^{\mathrm{O}}_i(f; m) \equiv \left[\frac{1}{m}\sum_{k=1}^{m} f(k/m)\sigma T^{\mathrm{O}}_{i,m}(k/m)\right]^2,$$

$i = 1, \ldots, n - m + 1$. The *overlapping area estimator* for σ^2 is

$$\mathcal{A}^{\mathrm{O}}(f; b, m) \equiv \frac{1}{n-m+1}\sum_{i=1}^{n-m+1} A^{\mathrm{O}}_i(f; m).$$

Alexopoulos et al. (2005b) use the continuous mapping theorem to show that as $m \to \infty$,

$$\mathcal{A}^{\mathrm{O}}(f; b, m) \xrightarrow{\mathcal{D}} \mathcal{A}^{\mathrm{O}}(f; b) \equiv \frac{\sigma^2}{b-1}\int_0^{b-1}\left(\int_0^1 f(t)\mathcal{B}_{u,1}(t)\,dt\right)^2 du. \quad (2)$$

It is easy to see that the expected value of the overlapping area estimator equals that of the corresponding batched area estimator. Thus, Theorem 2 gives

Theorem 4 *Under mild conditions similar to those of Theorem 2,*

$$\mathrm{E}[\mathcal{A}^{\mathrm{O}}(f;b,m)] = \sigma^2 + \frac{F^\star \gamma}{m} + o(\frac{1}{m}).$$

Calculation of the variance of the overlapping area estimator can be undertaken using the right-hand side of Equation (2) along with some algebraic elbow grease. Some examples from Alexopoulos et al. (2005b) reveal that the limiting ($m \to \infty$) variance of the overlapping area estimator depends on the choice of weighting function.

Example 4 For the overlapping area estimator with constant weight $f_0(t)$ from Example 1, we have that as $m \to \infty$,

$$\mathrm{Var}(\mathcal{A}^{\mathrm{O}}(f_0;b,m)) \to \mathrm{Var}(\mathcal{A}^{\mathrm{O}}(f_0;b)) = \frac{24b-31}{35(b-1)^2}\sigma^4.$$

This compares favorably to the batched constant-weighted area estimator's asymptotic ($m \to \infty$) variance, $\mathrm{Var}(\mathcal{A}(f_0;b)) = 2\sigma^4/b$ (see Theorem 2).

Example 5 For the overlapping area estimator with first-order unbiased quadratic weighting function $f_2(t)$ from Example 2, we have an asymptotic ($m \to \infty$) variance of

$$\mathrm{Var}(\mathcal{A}^{\mathrm{O}}(f_2;b)) = \frac{3514b-4359}{4290(b-1)^2}\sigma^4.$$

This compares well to the analogous batched quadratically weighted area estimator's asymptotic variance, $\mathrm{Var}(\mathcal{A}(f_2;b)) = 2\sigma^4/b$.

Example 6 For the overlapping area estimators from the family of orthonormal first-order unbiased weights $f_{\cos,j}(t)$, $j = 1, 2, \ldots$, from Example 3, we have

$$\mathrm{Var}(\mathcal{A}^{\mathrm{O}}(f_{\cos,j};b)) \doteq \frac{8\pi^2 j^2 + 15}{12\pi^2 j^2 b}\sigma^4.$$

Yet again, the analogous batched weighted area estimator has asymptotic variance $\mathrm{Var}(\mathcal{A}(f_{\cos,j};b)) = 2\sigma^4/b$.

Remark 1 One can average the orthonormal estimators $\mathcal{A}^{\mathrm{O}}(f_{\cos,j};b,m)$, $j = 1, 2, \ldots$, and use knowledge of the covariances of these estimators to obtain estimators with even smaller variance (cf. Alexopoulos et al. 2005b).

3.3 OBM Estimator

The ith overlapping batch mean is given by $\bar{Y}^{\mathrm{O}}_{i,m}$, $i = 1, \ldots, n-m+1$. The OBM estimator for σ^2, originally studied by Meketon and Schmeiser (1984) (with a slightly different scaling constant), is

$$\mathcal{O}(b,m) \equiv \frac{nm}{(n-m+1)(n-m)} \sum_{i=1}^{n-m+1} (\bar{Y}^{\mathrm{O}}_{i,m} - \bar{Y}_n)^2.$$

Theorem 5 *Under mild conditions, Goldsman and Meketon (1986) and Song and Schmeiser (1995) show that, for large b,*

$$\mathrm{E}[\mathcal{O}(b,m)] = \sigma^2 + \frac{\gamma}{m} + o(1/m).$$

Further, Meketon and Schmeiser (1984), Damerdji (1995), and Alexopoulos et al. (2005b) find that as $m \to \infty$,

$$\mathrm{Var}(\mathcal{O}(b,m)) \to \frac{(4b^3 - 11b^2 + 4b + 6)\sigma^4}{3(b-1)^4} \doteq \frac{4\sigma^4}{3b},$$

with the approximate result holding for large b.

We see from Theorems 4 and 5, along with the accompanying examples, that as $m \to \infty$, the overlapping area and OBM estimators are asymptotically unbiased for σ^2. In fact, the overlapping estimators preserve the bias properties of their nonoverlapping counterparts. Thus, we find that the overlapping area estimator with certain weighting functions is first-order unbiased, whereas OBM is not. An added feature is that the overlapping estimators also defeat their nonoverlapped counterparts in terms of variance, sometimes by quite a bit.

4 LINEAR COMBINATIONS OF OVERLAPPING ESTIMATORS

We show how to apply overlapping batch techniques to the area estimator using *different batch sizes*, with the intent of constructing a variance-optimal linear combination of these estimators.

4.1 Motivation

The idea behind the new estimators is simple.

- Form an overlapping area estimator using batch size m, i.e., $\mathcal{A}^{\mathrm{O}}(f;b,m)$.
- Change the batch size to rm, where $r > 1$, and form $\mathcal{A}^{\mathrm{O}}(f;b/r,rm)$. Generate several such overlapping area estimators with different batch sizes, say $m_1, m_2, \ldots, m_k$.
- Form a linear combination of these k estimators and scale appropriately. Denote this new estimator by $\mathcal{A}^{\mathrm{LO}}(f;\boldsymbol{M})$, where $\boldsymbol{M} \equiv \{m_1, \ldots, m_k\}$.

If $m \le \min\{m_1, \ldots, m_k\}$, then the bias of $\mathcal{A}^{\mathrm{LO}}(f;\boldsymbol{M})$ will be lower than that of $\mathcal{A}^{\mathrm{O}}(f;b,m)$. In addition, if the k constituent overlapping area estimators used to form $\mathcal{A}^{\mathrm{LO}}(f;\boldsymbol{M})$ are all first-order unbiased, then $\mathcal{A}^{\mathrm{LO}}(f;\boldsymbol{M})$ will also be unbiased. Further, the linear combination estimator will have lower variance, even if some of the constituent estimators are correlated (which will likely be the case).

4.2 Expected Value and Variance

As a simple example, we combine two overlapping area estimators, one based on batches of size m, and the other on batches of size $2m$, i.e.,

$$\begin{aligned}&\mathcal{A}^{\mathrm{LO}}(f;\{m,2m\})\\ &\quad= \alpha\mathcal{A}^{\mathrm{O}}(f;b,m)+(1-\alpha)\mathcal{A}^{\mathrm{O}}(f;\tfrac{b}{2},2m).\end{aligned}$$

Then

$$\begin{aligned}&\mathrm{E}[\mathcal{A}^{\mathrm{LO}}(f;\{m,2m\})]\\ &\quad= \alpha\mathrm{E}[\mathcal{A}^{\mathrm{O}}(f;b,m)]+(1-\alpha)\mathrm{E}[\mathcal{A}^{\mathrm{O}}(f;\tfrac{b}{2},2m)]\\ &\quad= \sigma^2+\frac{(1+\alpha)F^\star\gamma}{2m}+o(1/m).\end{aligned}$$

If $\alpha\in[-3,1]$ and we ignore small-order terms, we see that

$$\left|\mathrm{Bias}[\mathcal{A}^{\mathrm{LO}}(f;\{m,2m\})]\right| \;\le\; \left|\mathrm{Bias}[\mathcal{A}^{\mathrm{O}}(f;b,m)]\right|.$$

Further,

$$\begin{aligned}&\mathrm{Var}\left(\mathcal{A}^{\mathrm{LO}}(f;\{m,2m\})\right)\\ &\quad= \alpha^2\,\mathrm{Var}(\mathcal{A}^{\mathrm{O}}(f;b,m))+(1-\alpha)^2\,\mathrm{Var}(\mathcal{A}^{\mathrm{O}}(f;\tfrac{b}{2},2m))\\ &\qquad+2\alpha(1-\alpha)\,\mathrm{Cov}(\mathcal{A}^{\mathrm{O}}(f;b,m),\mathcal{A}^{\mathrm{O}}(f;\tfrac{b}{2},2m))\\ &\quad\doteq \alpha^2\,\mathrm{Var}(\mathcal{A}^{\mathrm{O}}(f;b))+(1-\alpha)^2\,\mathrm{Var}(\mathcal{A}^{\mathrm{O}}(f;\tfrac{b}{2}))\\ &\qquad+2\alpha(1-\alpha)\,\mathrm{Cov}(\mathcal{A}^{\mathrm{O}}(f;b),\mathcal{A}^{\mathrm{O}}(f;\tfrac{b}{2})),\end{aligned}\tag{3}$$

for large m. It would now be efficacious if we could choose α to minimize $\mathrm{Var}(\mathcal{A}^{\mathrm{LO}}(f;\{m,2m\}))$.

4.3 Minimizing the Variance

Motivated by the previous discussion, we will find a general expression for

$$c(f;b,m;r,s)\;\equiv\;\mathrm{Cov}(\mathcal{A}^{\mathrm{O}}(f;\tfrac{b}{r},rm),\mathcal{A}^{\mathrm{O}}(f;\tfrac{b}{s},sm)).$$

First of all, we need the following theorem, which generalizes (2).

Theorem 6 *Fix $r\in[1,b-1]$. As $m\to\infty$,*

$$\mathcal{A}^{\mathrm{O}}(f;\tfrac{b}{r},rm)\;\stackrel{\mathcal{D}}{\to}\;\frac{\sigma^2}{b-r}\int_0^{b-r}\left(\int_0^1 f(t)\mathcal{B}_{u,r}(t)\,dt\right)^2du.$$

The theorem, along with a great deal of algebra, allows us to make the necessary covariance calculations.

Example 7 For large b, $m\to\infty$, and $1\le s\le r$, we have

$$c(f_0;b,m;r,s)\;\doteq\;\frac{6s^2(7r^2-3s^2)\sigma^4}{35r^3b}.$$

Example 8 For large b, $m\to\infty$, and $1\le s\le r$, we have

$$\begin{aligned}&c(f_2;b,m;r,s)\\ &\quad\doteq\;\frac{7(3003r^5-3250r^4s+875r^2s^3-126s^5)s^3\sigma^4}{4290\,r^7\,b}.\end{aligned}$$

We can use these results in conjunction with Equation (3).

Example 9 From Examples 4 and 7, we have

$$\begin{aligned}&\mathrm{Var}\left(\mathcal{A}^{\mathrm{LO}}(f_0;\{m,2m\})\right)\\ &\quad= \alpha^2\,\mathrm{Var}(\mathcal{A}^{\mathrm{O}}(f_0;b,m))+(1-\alpha)^2\,\mathrm{Var}(\mathcal{A}^{\mathrm{O}}(f_0;\tfrac{b}{2},2m))\\ &\qquad+2\alpha(1-\alpha)\,c(f_0;b,m;2,1)\\ &\quad\doteq \alpha^2\frac{24\sigma^4}{35b}+(1-\alpha)^2\frac{24\sigma^4}{35(b/2)}+2\alpha(1-\alpha)\frac{15\sigma^4}{28b},\end{aligned}$$

for large b and $m\to\infty$. This quantity is minimized by taking $\alpha=0.8478$, whence

$$\mathrm{Var}(\mathcal{A}^{\mathrm{LO}}(f_0;\{m,2m\}))\;\doteq\;0.6629\sigma^4/b,$$

which compares to $\mathrm{Var}(\mathcal{A}^{\mathrm{O}}(f_0;b,m))\doteq 0.6857\sigma^4/b$ from Example 4.

4.4 General Linear Combinations

Now consider the general linear combination of overlapping estimators,

$$\mathcal{A}^{\mathrm{LO}}(f;\boldsymbol{M})\;\equiv\;\sum_{i=1}^{k}\alpha_i\mathcal{A}^{\mathrm{O}}(f;\tfrac{b}{r_i},r_im),$$

where $\sum_{i=1}^{k}\alpha_i=1$ and $r_1,r_2,\ldots,r_k\ge 1$. Then

$$\mathrm{E}[\mathcal{A}^{\mathrm{LO}}(f;\boldsymbol{M})]\;=\;\sigma^2+F^\star\gamma\sum_{i=1}^{k}\frac{\alpha_i}{r_im}+o(1/m)$$

and

$$\text{Var}(\mathcal{A}^{\text{LO}}(f;\boldsymbol{M})) \quad (4)$$
$$= \sum_{i=1}^{k} \alpha_i^2 \,\text{Var}(\mathcal{A}^{\text{O}}(f;\tfrac{b}{r_i},r_i m)) + 2\sum_{i=2}^{k}\sum_{j=1}^{i-1}\alpha_i\alpha_j\text{Cov}_{ij},$$

where $\text{Cov}_{ij} \equiv c(f;b,m;r_i,r_j)$.

For large b and $m \to \infty$, one can (approximately) minimize the variance given in Equation (4) subject to the constraint that $\sum_{i=1}^{k}\alpha_i = 1$. Table 1 illustrates the asymptotic bias and minimized variance of our linear combination estimators for a variety of choices of weights f and batch size sets $\boldsymbol{M}$. Note that the $\boldsymbol{M} = \{m\}$ case corresponds to the basic overlapping estimator $\mathcal{A}^{\text{O}}(f;b,m)$.

Table 1: Approximate Performance of Various Variance-Optimal Linear Combinations of Overlapping Estimators

	$\mathcal{A}^{\text{LO}}(f_0;\boldsymbol{M})$		$\mathcal{A}^{\text{LO}}(f_2;\boldsymbol{M})$	
$\boldsymbol{M}$	$\frac{m}{\gamma}$Bias	$\frac{b}{\sigma^4}$Var	$\frac{m}{\gamma}$Bias	$\frac{b}{\sigma^4}$Var
$\{m\}$	3.00	0.686	$o(1)$	0.819
$\{m,2m\}$	2.77	0.663	$o(1)$	0.782
$\{m,2m,3m\}$	2.71	0.638	$o(1)$	0.731
$\{m,\ldots,4m\}$	2.67	0.630	$o(1)$	0.722
$\{m,\ldots,10m\}$	2.61	0.615	$o(1)$	0.695
$\{m,\ldots,20m\}$	2.59	0.610	$o(1)$	0.688

We see that the estimator $\mathcal{A}^{\text{LO}}(f_0;\boldsymbol{M})$ has bias of the form c/m, where the constant c decreases a bit from 3.00 to 2.59 as we add more and more terms to the linear combination; the estimator $\mathcal{A}^{\text{LO}}(f_2;\boldsymbol{M})$ only has $o(1/m)$ bias. Further, in this example, the standardized variance of $\mathcal{A}^{\text{LO}}(f_0;\boldsymbol{M})$ decreases from 0.686 to 0.610 (about 12%) as we add more terms, while that of $\mathcal{A}^{\text{LO}}(f_2;\boldsymbol{M})$ decreases from 0.819 to 0.688 (about 16%). In any case, $\mathcal{A}^{\text{LO}}(f_0;\boldsymbol{M})$ has higher bias but lower variance than $\mathcal{A}^{\text{LO}}(f_2;\boldsymbol{M})$—the familiar bias-variance tradeoff. Finally, for comparison purposes, we note that the NBM and OBM estimators both have an analogous bias constant of about 1, and respective variance constants of 2.000 and 1.333.

A simple Monte Carlo example shows that the estimators perform as advertised.

Example 10 Consider an i.i.d. standard normal sequence $Y_1,Y_2,\ldots,Y_n$, with $n = 20{,}000$. Based on 10,000 replications of this process, we estimated the expected values and variances for a variety of linear combinations of overlapping area variance estimators, $\mathcal{A}^{\text{LO}}(f_0;\boldsymbol{M})$. Representative results are given in Table 2, where $b = 20$ and $m = 1000$. The last column in the table provides the asymptotic ($m \to \infty$) variance of each variance estimator that we have obtained analytically. We see that the empirical results match up with the theory.

Table 2: Empirical Performance of Variance Estimators $\mathcal{A}^{\text{LO}}(f_0;\boldsymbol{M})$ for $b = 20$ and $m = 1000$

$\boldsymbol{M}$	Sample Var	True Var
$\{m,2m\}$	0.0344	0.0346
$\{m,2m,3m\}$	0.0334	0.0334
$\{m,2m,\ldots,10m\}$	0.0324	0.0328

5 SUMMARY AND CONCLUSIONS

In this paper, we introduced linear combinations of overlapping versions of standardized times series area variance estimators for steady-state simulations. We obtained asymptotic expressions for the expected values and variances of these linear combination variance estimators, and we compared them with the corresponding nonoverlapping and overlapping estimators as well as with the nonoverlapping and overlapping batch means estimators. We showed that the linear combination estimators have slightly smaller bias than their nonoverlapping and overlapping counterparts as well as lower variance—sometimes substantially lower. We supported these asymptotic results by a simple empirical example that showed that the linear combination variance estimators perform as predicted by the theory.

Ongoing research includes a battery of analytical and Monte Carlo examples, as well as the development of linear combinations of other varieties of overlapping estimator, e.g., those based on the Cramér–von Mises estimators from Goldsman et al. (1999).

ACKNOWLEDGMENT

Partial support for our research was provided by NSF Grant DMI-0400260.

REFERENCES

Alexopoulos, C., N. T. Argon, D. Goldsman, N. M. Steiger, G. Tokol, and J. R. Wilson. 2005a. Overlapping variance estimators for simulation: II. Implementation and evaluation. Technical report, School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA.

Alexopoulos, C., N. T. Argon, D. Goldsman, G. Tokol, and J. R. Wilson. 2005b. Overlapping variance estimators for simulation: I. Theory. Technical Report, School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA.

Billingsley, P. 1968. *Convergence of Probability Measures*. New York: John Wiley & Sons.

Chien, C., D. Goldsman, and B. Melamed. 1997. Large-sample results for batch means. *Management Science* 43:1288–1295.

Damerdji, H. 1991. Strong consistency and other properties of the spectral variance estimator. *Management Science* 37:1424–1440.

Damerdji, H. 1994. Strong consistency of the variance estimator in steady-state simulation output analysis. *Mathematics of Operations Research* 19:494–512.

Damerdji, H. 1995. Mean-square consistency of the variance estimator in steady-state simulation output analysis. *Operations Research* 43:282–291.

Foley, R. D., and D. Goldsman. 1999. Confidence intervals using orthonormally weighted standardized times series. *ACM Transactions on Modeling and Computer Simulation* 9:297–325.

Glynn, P. W., and D. L. Iglehart. 1990. Simulation output analysis using standardized time series. *Mathematics of Operations Research* 15:1–16.

Glynn, P. W., and W. Whitt. 1991. Estimating the asymptotic variance with batch means. *Operations Research Letters* 10:431–435.

Goldsman, D., K. Kang, and A. F. Seila. 1999. Cramér–von Mises variance estimators for simulations. *Operations Research* 47:299–309.

Goldsman, D., and M. S. Meketon. 1986. A comparison of several variance estimators. Technical Report J-85-12, School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, Georgia.

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

Goldsman, D., and L. W. Schruben. 1990. New confidence interval estimators using standardized time series. *Management Science* 36:393–397.

Law, A. M., and W. D. Kelton. 2000 *Simulation Modeling and Analysis*, 3rd ed. New York: McGraw-Hill.

Meketon, M. S., and B. W. Schmeiser. 1984. Overlapping batch means: Something for nothing? In *Proceedings of the 1984 Winter Simulation Conference*, eds. S. Sheppard, U. W. Pooch, and C. D. Pegden, 227–230. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers.

Pedrosa, A., and B. W. Schmeiser. 1993. Asymptotic and finite-sample correlations between OBM estimators. In *Proceedings of the 1993 Winter Simulation Conference*, eds. G. W. Evans, M. Mollaghssemi, E. C. Russell, and W. E. Biles, 481–488. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers.

Pedrosa, A., and B. W. Schmeiser. 1994. Estimating the variance of the sample mean: Optimal batch-size estimation and 1-2-1 overlapping batch means. Technical Report SMS94-3, School of Industrial Engineering, Purdue University, West Lafayette, Indiana.

Schmeiser, B. W. 1982. Batch size effects in the analysis of simulation output. *Operations Research* 30:556–568.

Schruben, L. W. 1983. Confidence interval estimation using standardized time series. *Operations Research* 31:1090–1108.

Song, W.-M. T. 1988. *Estimators of the variance of the sample mean: Quadratic forms, optimal batch sizes, and linear combinations*. Ph.D. Dissertation, School of Industrial Engineering, Purdue University, West Lafayette, Indiana.

Song, W.-M. T., and B. W. Schmeiser. 1993. Variance of the sample mean: Properties and graphs of quadratic-form estimators. *Operations Research* 41:501–517.

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

Steiger, N. M., and J. R. Wilson. 2001. Convergence properties of the batch means method for simulation output analysis. *INFORMS Journal on Computing* 13(4):277–293.

Welch, P. D. 1987. On the relationship between batch means, overlapping batch means and spectral estimation. In *Proceedings of the 1987 Winter Simulation Conference*, eds. A. Thesen, H. Grant, and W. D. Kelton, 320–323. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers.

AUTHOR BIOGRAPHIES

DAVID GOLDSMAN is a Professor in the School of Industrial and Systems Engineering at the Georgia Institute of Technology. His research interests include simulation output analysis and ranking and selection. He is an active participant in the Winter Simulation Conference, having been Program Chair in 1995, and having served on the WSC Board of Directors since 2002. His e-mail address is <sman@isye.gatech.edu>, and his web page is <http://www.isye.gatech.edu/~sman>.

TÛBA AKTARAN-KALAYCI is a Ph.D candidate in the School of Industrial and Systems Engineering at the Georgia Institute of Technology. Her expected graduation date is May 2006. Her research interests include simulation output analysis and Markov decision processes with applications to queueing and pricing. Her email address is <taktaran@isye.gatech.edu>.