Steady-State Simulation Analysis Using ASAP3*

James R. Wilson
Edward P. Fitts Department of
Industrial and Systems Engineering
North Carolina State University
www.ise.ncsu.edu/jwilson

Joint work with
Natalie Steiger (Univ. of Maine); Emily Lada (SAS Institute);
Jeff Joines (NC State Univ.); and
Christos Alexopoulos and David Goldsman (Georgia Tech)

January 18, 2008

*Partially supported by U.S. NSF under grant number DMI-9900164.
OVERVIEW

I. Introduction
   A. Finite-Horizon vs. Steady-State Simulation Analysis
   B. Method of Nonoverlapping Batch Means (NBM)

II. Overview of ASAP3
   A. Steps in the Operation of ASAP3
   B. Flow Chart of ASAP3

III. Experimental Performance Evaluation
   A. Comparison: ASAP, ASAP3, Law & Carson (LC) Procedure
   B. Test Problems: Queueing Networks, Markov Chains
   C. Performance Measures: CI Coverage & Half-length; Sample Size
   D. Analysis of Results

IV. Conclusions and Recommendations
I. Introduction

A. Finite-Horizon vs. Steady-State Simulation Analysis

A \textit{finite-horizon simulation} is terminated at a specific time or by the occurrence of a specific condition. Examples include the operation of the following discrete-event stochastic systems—

- an inventory system over its five-year economic life;
- a call center over a single day or week;
- a manufacturing cell over the time interval required to process a given set of jobs; and
- the transient (start-up) phase of a power-generation system whose long-run reliability is also of independent interest.
A **steady-state simulation** operates (at least conceptually) into the indefinite future, and interest centers on its long-run behavior.

A performance measure is called a **steady-state parameter** if it is a characteristic of the equilibrium distribution of a simulation-generated output process.

Examples include the operation of the following discrete-event stochastic systems—

- a telecommunications system in which the objective is estimating the long-run average delay in transmitting a packet between selected points; and
- a retail distribution system in which the objective is estimating the long-run service level (that is, the percentage of demands that are satisfied from stock on hand).
Issues in the Analysis of Finite-Horizon and Steady-State Simulations

• By performing independent replications (runs) of a finite-horizon simulation, we can apply classical statistical techniques designed for the analysis of random samples.

• The most widely used statistical techniques require observations to be independent and identically distributed (i.i.d.) and normal.

• Commonly a simulation response accumulated over an individual run of a finite-horizon simulation and delivered at the end of the run is approximately normal because of a “central-limit” effect; and the responses from all such runs are i.i.d. by design.
• Three fundamental problems arise in steady-state simulation analysis.
  – Within a single prolonged run supposedly in steady-state operation, the *start-up problem* is caused by transients in the initial sequence of responses due to the system’s starting conditions. Beyond the warm-up period, these transients are negligible so that: (i) the sequence of successive responses approximately has the steady-state distribution; and (ii) the output parameter of interest is sufficiently close to its steady-state value.
  – The *correlation problem* is caused by pronounced stochastic dependencies among successive responses generated in a single run.
  – The *nonnormality problem* is caused by pronounced departures from normality in the successive responses generated within a single run.

• The focus of this talk is on ASAP3, a procedure for analysis of steady-state simulations that effectively handles the start-up, correlation, and nonnormality problems *automatically*—that is, without intervention by the user.
B. Method of Nonoverlapping Batch Means (NBM)

Given the simulation output \( \{X_i : i = 1, \ldots, n\} \), we seek a \( 100(1 - \alpha)\% \) confidence interval (CI) for the associated steady-state mean \( \mu_X \).

Partition the data into \( k \) adjacent nonoverlapping \textit{batches} each of size \( m \) (assume \( n = km \)) as follows:

\[
X_1, \ldots, X_m, X_{m+1}, \ldots, X_{2m}, \ldots, X_{(k-1)m+1}, \ldots, X_{km}
\]

Compute the \( j \)th batch mean

\[
Y_j(m) = \frac{1}{m} \sum_{i=m(j-1)+1}^{jm} X_i \quad \text{for } j = 1, \ldots, k,
\]

together with the grand mean of the batch means,

\[
\overline{Y} = \overline{Y}(m, k) = \frac{1}{k} \sum_{j=1}^{k} Y_j(m) = \overline{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i,
\]

(1)
and the sample variance of the batch means,

\[ S_{m,k}^2 = \frac{1}{k-1} \sum_{j=1}^{k} \left[ Y_j(m) - \bar{Y}(m,k) \right]^2, \]  

(2)

to yield the usual \(100(1 - \alpha)\)% CI for \(\mu_x\),

\[ \bar{Y}(m,k) \pm t_{1-\alpha/2,k-1} \frac{S_{m,k}}{\sqrt{k}}. \]  

(3)

where \( t_{1-\alpha/2,k-1} \) is the \(1 - \alpha\) quantile of the Student-\(t\) distribution with \(k - 1\) degrees of freedom.

- As \( m \to \infty \) with \( k \) fixed (so that \( n \to \infty \)), the NBM CI (3) has coverage probability \(1 - \alpha\), and thus is asymptotically valid under mild conditions on the output process \( \{X_i\} \).
II. Overview of ASAP3

ASAP3 is a batch means procedure that attempts to deliver an approximately unbiased point estimator of $\mu_X$ as well as an approximately valid CI estimator of $\mu_X$ with user-specified precision and coverage probability; see


ASAP3 requires the following user-supplied inputs:

- the output process \( \{X_i : i = 1, 2, \ldots, n\} \) from which the steady-state mean \( \mu_X \) is to be estimated;

- the desired CI coverage probability \( 1 - \alpha \), where \( 0 < \alpha < 1 \); and

- an upper bound \( H^* \) on the final CI half-length, expressed
  - in absolute terms as the maximum acceptable half-length; or
  - in relative terms as the maximum acceptable fraction \( r^* \) of the magnitude of the CI midpoint \( \overline{Y} \).
ASAP3 delivers the following outputs:

- a nominal $100(1 - \alpha)\%$ CI for $\mu_X$ having the form
  \[ \bar{Y} \pm H, \text{ where } H \leq H^*, \]
  provided no additional simulation is required; or

- a new, larger sample size $n$ to be supplied to the procedure.
ASAP3 builds a correlation-adjusted CI as follows:

- Determine a batch size $m$ sufficiently large to ensure that beyond the first 4 batch means $\{Y_j(m) : j = 1, \ldots, 4\}$, the following conditions hold: (i) any start-up effects are negligible; and (ii) the truncated batch means $\{Y_j(m) : j = 5, \ldots, k\}$ are approximately normal and covariance stationary with mean $\mu_X$ but are not necessarily independent.

- Fit a first-order autoregressive (AR(1)) time-series model to the truncated batch means,

$$ Y_j(m) - \mu_X = \varphi_{Y(m)} [Y_{j-1}(m) - \mu_X] + a_j(m) \text{ for } j = 5, 6, \ldots, $$

(4)

where $\{a_j(m) : j = 5, 6, \ldots\} \overset{\text{i.i.d.}}{\sim} N[0, \sigma^2_a(m)]$.

- Exploit the AR(1) model (4) of the batch means so as to estimate a batch size $m$ sufficiently large to ensure that

$$ \text{Corr}[Y_j(m), Y_{j+1}(m)] = \varphi_{Y(m)} \leq 0.8 \text{ for } j = 5, 6, \ldots $$
• Take an inverse Cornish-Fisher expansion for the usual $t$-ratio

$$t = \frac{\bar{Y}(m, k') - \mu_x}{\sqrt{\frac{S_{m,k'}^2}{k'}}}$$

(5)

using the parameter estimators obtained by fitting the AR(1) model (4) to the truncated set of $k' = k - 4$ batch means so as to compute the following:

- $\hat{\kappa}_2$ and $\hat{\kappa}_4$, estimators of the 2nd and 4th cumulants, respectively, of the $t$-ratio (5);
- $\widehat{\text{Var}}[Y(m)]$, an estimator of the variance of the batch means; and

• Finally compute the correlation-adjusted 100$(1 - \alpha)$% CI for $\mu_x$,

$$\bar{Y}(m, k') \pm \left( \frac{1}{2} + \frac{1}{2} \hat{\kappa}_2 - \frac{1}{8} \hat{\kappa}_4 + \frac{1}{24} \hat{\kappa}_4 z_{1-\alpha/2}^2 \right) \sqrt{\frac{\widehat{\text{Var}}[Y(m)]}{k'}}$$

(6)

Correlation adjustment

where $z_{1-\alpha/2}$ is the $1 - \alpha/2$ quantile of the $N(0, 1)$ distribution.
A. Steps in the Operation of ASAP3

[0] Divide the initial sample into \( k = 256 \) batches with user-defined initial batch size \( m \) (where \( m = 16 \) by default), skip the first 4 batches, and compute batch means for the remaining \( k' = k - 4 = 252 \) batches.

[1] Begin each new iteration of ASAP3 by

- collecting additional data (if any) required by the previous iteration;
- and
- computing the full set of batch means \( \{Y_j(m) : j = 1, \ldots, k\} \) using the current batch size \( m \).
From the current set of $k' = k - 4$ batch means $\{Y_j(m) : j = 5, \ldots, k\}$ that were truncated by skipping the first 4 batches to eliminate start-up effects, select every other group of 4 consecutive batch means to test for multivariate normality according to the following layout:

$$Y_1(m), Y_2(m), Y_3(m), Y_4(m), \quad Y_5(m), Y_6(m), Y_7(m), Y_8(m),$$

ignored spacer (start-up period)

$$Y_9(m), Y_{10}(m), Y_{11}(m), Y_{12}(m), \quad Y_{13}(m), Y_{14}(m), Y_{15}(m), Y_{16}(m),$$

1st $(4 \times 1)$ vector $y_1$

$$Y_9(m), Y_{10}(m), Y_{11}(m), Y_{12}(m), \quad Y_{13}(m), Y_{14}(m), Y_{15}(m), Y_{16}(m),$$

2nd $(4 \times 1)$ vector $y_2$

$$\vdots$$

$$Y_{249}(m), Y_{250}(m), Y_{251}(m), Y_{252}(m), \quad Y_{253}(m), Y_{254}(m), Y_{255}(m), Y_{256}(m).$$

32nd $(4 \times 1)$ vector $y_{32}$
• If the resulting 4-dimensional spaced batch-means vectors \( \{y_j : j = 1, \ldots, 32\} \) fail the Shapiro-Wilk multivariate normality test, then increase the batch size \( m \) according to

\[
m \leftarrow \left\lfloor \sqrt{2m} \right\rfloor
\]

and go to step [1] to begin a new iteration.

• If the vectors \( \{y_j : j = 1, \ldots, 32\} \) pass the Shapiro-Wilk multivariate normality test, then go to step [3].
Fit an AR(1) model,

\[ Y_j(m) - \mu_X = \varphi_{Y(m)} [Y_{j-1}(m) - \mu_X] + a_j(m) \quad \text{for} \quad j = 5, 6, \ldots, \]

to the current set of truncated batch means, where:

- the autoregressive parameter \( \varphi_{Y(m)} \) is the corresponding lag-one correlation

\[ \varphi_{Y(m)} = \text{Corr}[Y_j(m), Y_{j+1}(m)] \quad \text{for} \quad j = 5, 6, \ldots; \]

and

- the residuals \( \{a_{j(m)} : j = 5, 6, \ldots\} \) i.i.d. \( \sim N[0, \sigma_{a(m)}^2] \).
Test the null hypothesis

\[ \varphi_{Y(m)} \leq 0.8 \]

versus the alternative hypothesis

\[ \varphi_{Y(m)} > 0.8. \]

• If the null hypothesis is rejected, then
  – increase the batch size \( m \) by a factor projected to reduce \( \varphi_{Y(m)} \) to 0.80 on the next iteration; and
  – go to step [1] to begin a new iteration.

• If the null hypothesis is accepted, then go to step [4].
[4] Build a correlation-adjusted CI based on an inverse Cornish-Fisher expansion for the usual $t$-ratio

$$
t = \left[ \frac{\overline{Y}(m, k') - \mu_x}{\sqrt{S_{m,k'/k'}}} \right]
$$

using estimates $\hat{\varphi}_{Y(m)}$ and $\hat{\sigma}_{a(m)}^2$ for the AR(1) model (4) fitted to the current set of $k'$ truncated batch means so as to compute:

- $\hat{\kappa}_2$ and $\hat{\kappa}_4$, estimators of the 2nd and 4th cumulants, respectively, of the usual $t$-ratio; and
- $\widehat{\text{Var}}[Y(m)]$, an estimator of the variance of the batch means.

From these quantities, compute the $100(1 - \alpha)\%$ CI for $\mu_x$,

$$
\overline{Y}(m, k') \pm \left[ \left( \frac{1}{2} + \frac{1}{2} \hat{\kappa}_2 - \frac{1}{8} \hat{\kappa}_4 + \frac{1}{24} \hat{\kappa}_4 \hat{\kappa}_2^2 \right) \right] z_{1-\alpha/2} \sqrt{\frac{\widehat{\text{Var}}[Y(m)]}{k'}}.
$$
If the half-length $H$ of the current CI satisfies the precision requirement

$$H \leq H^*, \quad (7)$$

then deliver that CI and stop.

If (7) is not satisfied, then do the following:

- Estimate additional batches needed to satisfy (7),
  $$k'' = \max\left\{ \left\lceil \left( \frac{H}{H^*} \right)^2 k' \right\rceil - k', \, 1 \right\}.$$

- If $k' + k'' \leq 1,504$, then update the batch count
  $$k' \leftarrow k' + k'', \quad k \leftarrow k' + 4$$

  leave the batch size $m$ unchanged, and go to step [1] to begin a new iteration.
• If $k' + k'' > 1,504$, then update the batch size $m$ by the multiplier $\theta$ that is the root of the equation

$$\theta \left( 1 - \hat{\varphi}_+ \right)^2 = \left( \frac{H}{H^*} \right)^2 \left( 1 - \hat{\varphi}_+ \right)^2,$$

where

$$\hat{\varphi}_+ \equiv \max \{ 0, \hat{\varphi}_{Y(m)} \}$$

so that we take

$$m \leftarrow \left\lceil \text{mid} \left( \sqrt{2}, \theta, 4 \right) m \right\rceil;$$

leave the batch count $k$ unchanged; and go to [1] to begin a new iteration.
B. Flow Chart of ASAP3

- Start
- Collect observations; compute batch means statistics
- Stationary multivariate normality test passed?
  - No: Update test level $\delta$
  - Yes: Compute new batch size
- Retain old batch count; compute new batch size
- AR(1) parameter $\phi \leq 0.8$?
  - No: Compute effective degrees of freedom and inverse Cornish-Fisher expansion for $t$-ratio
  - Yes: Construct correlation-adjusted CI
- CI meets precision requirements?
  - Yes: Deliver CI; Stop
  - No: Use new batch count and retain old batch size
- New batch count $> 1,504$?
  - No: Compute new batch size
  - Yes: New batch count $> 1,504$?
Algorithmic Statement of ASAP3

[0] Set iteration index $i \leftarrow 1$, $m_1 \leftarrow$ user-specified initial batch size (default = 16),
initial batch count $k_1 \leftarrow 256$, initial sample size $n_1 \leftarrow k_1 m_1$ with $n_0 \leftarrow 0$,
truncated initial batch count $k'_1 \leftarrow k_1 - 4$,
$1 - \alpha \leftarrow$ user-specified CI coverage probability (default = 0.90),
size of test for autoregressive parameter $\alpha_{\text{arp}} \leftarrow 0.01$,
initial size of test for stationary multivariate normality $\delta_1 \leftarrow 0.1$ with parameter
$\omega \leftarrow 0.18421$ controlling the test size in step [2] on subsequent iterations,
and indicator that normality test was passed $\text{MVTestPassed} \leftarrow \text{‘no’}$;

if a relative precision requirement is given, then set $\text{RelPrec} \leftarrow \text{‘yes’}$ and
$r^* \leftarrow$ the user-specified fraction of the magnitude of the CI midpoint that defines
the maximum acceptable CI half-length;

if an absolute precision requirement is given, then set $\text{RelPrec} \leftarrow \text{‘no’}$ and
$H^* \leftarrow$ the user-specified maximum acceptable CI half-length;

if no precision level is specified then set $\text{RelPrec} \leftarrow \text{‘no’}$, $r^* \leftarrow 0$, and $H^* \leftarrow 0$. 
Algorithmic Statement of ASAP3 (Continued)

[1] Start (or restart) the simulation to generate the data \( \{X_j : j = n_i - 1 + 1, \ldots, n_i\} \) required for the current iteration \( i \); 

Compute the \( k_i \) batch means \( \{Y_j(m_i) : j = 1, \ldots, k_i\} \); and after skipping the initial spacer \( \{Y_1(m_i), Y_2(m_i), Y_3(m_i), Y_4(m_i)\} \), compute the truncated grand mean,

\[
\bar{Y}(m_i, k'_i) \leftarrow \frac{1}{k'_i m_i} \sum_{\ell=4m_i+1}^{n_i} X_\ell = \frac{1}{k'_i} \sum_{j=5}^{k_i} Y_j(m_i);
\]

if MVTestPassed='yes', then goto [3].
Algorithmic Statement of ASAP3 (Continued)

[2] From the truncated batch means \( \{ Y_j(m_i) : j = 5, \ldots, k_i \} \), select every other group of four successive batch means to build the \( 4 \times 1 \) vectors

\[
\{ y_\ell = \begin{bmatrix} Y_{5+(\ell-1)8}(m_i), Y_{6+(\ell-1)8}(m_i), Y_{7+(\ell-1)8}(m_i), Y_{8+(\ell-1)8}(m_i) \end{bmatrix}^T : \ell = 1, \ldots, 32 \};
\]

To test the hypothesis

\( \mathcal{H}_{\text{mvn}} : \{ y_\ell : \ell = 1, \ldots, 32 \} \) are i.i.d. four-dimensional normal random vectors,

evaluate \( \delta_i = \delta_1 \exp[-\omega(i-1)^2] \), the significance level for the test, and \( W_i^* \), the multivariate Shapiro-Wilk statistic computed from the \( \{ y_\ell \} \) according to equations (10)–(12) of Steiger et al. (2005);

if \( W_i^* < w_{\delta_i}^* \), the \( \delta_i \) quantile of the distribution of \( W_i^* \) under the null hypothesis \( \mathcal{H}_{\text{mvn}} \), so that \( \mathcal{H}_{\text{mvn}} \) is rejected at significance level \( \delta_i \), then

set \( i \leftarrow i + 1, \ k_i \leftarrow 256, \ k'_i \leftarrow k_i - 4, \ m_i \leftarrow \left\lfloor \sqrt{2}m_{i-1} \right\rfloor, \) and \( n_i \leftarrow k_i m_i; \)

goto [1];

else

set \( \text{MVTestPassed} \leftarrow \text{‘yes’}; \)

goto [3].
Algorithmic Statement of ASAP3 (Continued)

[3] Fit an AR(1) model (4) to the truncated batch means \( \{Y_j(m_i) : j = 5, \ldots, k_i\} \) so as to obtain the estimator \( \hat{\phi} \) of the autoregressive parameter \( \varphi \);

Test the hypothesis \( \mathcal{H}_{\text{arp}} : \varphi \leq 0.8 \) at the level of significance \( \alpha_{\text{arp}} \) by checking for the condition

\[
\hat{\varphi} \leq \sin\left(0.927 - z_{1-\alpha_{\text{arp}}} / \sqrt{k_i'}\right);
\]

if \( \mathcal{H}_{\text{arp}} \) is rejected at significance level \( \alpha_{\text{arp}} \), then

set \( \theta \leftarrow \text{mid}\left\{\sqrt{2}, \ln\left[\sin\left(0.927 - z_{1-\alpha_{\text{arp}}} / \sqrt{k_i'}\right)\right] / \ln(\hat{\varphi}), 4\right\} \),

\( i \leftarrow i + 1, \ k_i \leftarrow k_{i-1}, \ k_i' \leftarrow k_i - 4, \ m_i \leftarrow \lceil \theta m_{i-1} \rceil, \) and \( n_i \leftarrow k_i m_i; \) \quad \text{goto [1];}

else

\quad \text{goto [4].}
Algorithmic Statement of ASAP3 (Continued)

[4] Using the estimators \( \hat{\phi} \) and \( \hat{\sigma}^2 \) for the AR(1) model (4), compute \( \hat{\text{Var}}[Y(m_i)] \) and \( \hat{\text{Var}}[\bar{Y}(m_i, k'_i)] \) from equations (15)–(16) of Steiger et al. (2005);

For the NOBM \( t \)-ratio (5), compute the estimated effective degrees of freedom \( \hat{\nu}_{\text{eff}} \) from equation (33) of Steiger et al. (2005);

Compute \( \hat{\kappa}_2 \) and \( \hat{\kappa}_4 \), the estimators, respectively, of the second and fourth cumulants of the \( t \)-ratio (5), by inserting \( \hat{\text{Var}}[Y(m_i)], \hat{\text{Var}}[\bar{Y}(m_i, k'_i)] \), and \( \hat{\nu}_{\text{eff}} \) into the computing expressions for \( \kappa_2 \) and \( \kappa_4 \) given in equations (31)–(32) of Steiger et al. (2005);

Calculate the half-length of the correlation-adjusted CI,

\[
H \leftarrow \left[ \left( \frac{1}{2} + \frac{1}{2}\hat{\kappa}_2 - \frac{1}{8}\hat{\kappa}_4 + \frac{1}{24}\hat{\kappa}_4 \hat{\sigma}^2 \right) \right] \frac{z_{1-\alpha/2}}{\sqrt{\text{Var}[Y(m_i)]/k'}}
\]

Construct the correlation-adjusted CI,

\[
\bar{Y}(m_i, k'_i) \pm H.
\]
Algorithmic Statement of ASAP3 (Continued)

[5] if RelPrec='yes' then set $H^* \leftarrow r^*|\overline{Y}(m_i, k_i')|$

if $(H \leq H^*)$ or $(r^* = 0$ and $H^* = 0)$, then
    deliver $\overline{Y}(m_i, k_i') \pm H$ and stop;
else
    Estimate additional batches needed to satisfy the precision requirement (7),
    $k'' = \max\left\{\left\lceil \left(\frac{H}{H^*}\right)^2 k_i' \right\rceil - k_i', 1\right\}$;

If $k_i + k'' \leq 1,504$, then
    set $i \leftarrow i + 1$, $k_i \leftarrow k_{i-1} + k''$, $k_i' \leftarrow k_i - 4$, $m_i \leftarrow m_{i-1}$,
    and $n_i \leftarrow m_i k_i$; goto [1];
else
    Find the root $\theta$ of the equation $\theta \left(1 - \hat{\varphi}_+\right)^2 = \left(\frac{H}{H^*}\right)^2 \left(1 - \hat{\varphi}_+\right)^2$,
    set $\theta \leftarrow \text{mid}(\sqrt{2}, \theta, 4)$, $i \leftarrow i + 1$, $k_i \leftarrow k_{i-1}$, $k_i' \leftarrow k_i - 4$,
    $m_i \leftarrow \lceil \theta m_{i-1} \rceil$, and $n_i \leftarrow m_i k_i$; goto [1].
III. Experimental Performance Evaluation

A. Comparison: ASAP, ASAP3, and the Law & Carson (LC) Procedure


Overall Objective of Law & Carson (LC) Procedure

Deliver \( k = 40 \) batches of a batch size \( m'' \) sufficiently large to ensure the lag-one correlation between the batch means satisfies

\[
\text{Corr}[Y_j(m''), Y_{j+1}(m'')] \leq 0.05
\]

while the associated \( 100(1 - \alpha)\% \) confidence interval for \( \mu_X \),

\[
\bar{Y}(m'', k) \pm H_{\alpha,k} \quad \text{with} \quad H_{\alpha,k} \equiv t_{1-\alpha/2,k-1} \frac{S_{m'',k}}{\sqrt{k}},
\]

satisfies the user-specified relative precision requirement

\[
\left| H_{\alpha,k} / \bar{Y}(m'', k) \right| \leq r^*.
\]
### Algorithmic Statement of Law and Carson (LC) Procedure

**Step 0.** Set the positive integers \( \ell \leftarrow 10, k \leftarrow 40, n_0 \leftarrow 600, \) and \( n_1 \leftarrow 800, \) where \( k' \equiv \ell k = 400 \) and \( k'' \equiv k'/2 = 200. \) Set the stopping value \( c \leftarrow 0.40 \) and the relative precision \( r^* > 0; \) set \( i \leftarrow 1; \) and collect \( n_1 \) observations.

**Step 1.**

a. Divide the overall data set of \( n_i \) observations into \( k' \) batches of size \( m = n_i / k'. \) Compute estimated lag-1 correlation \( \tilde{\rho}_1 (k', m) \) of the batch means \( \{Y_j (m) : j = 1, \ldots, k'\}. \) If \( \tilde{\rho}_1 (k', m) \geq c, \) then go to **Step 2.** If \( \tilde{\rho}_1 (k', m) \leq 0, \) then go to **Step 1c.** Otherwise, go to **Step 1b.**

b. Divide the overall data set of size \( n_i \) into \( k'' \equiv k'/2 \) batches of size \( m' \equiv 2m. \) Compute the estimated lag-1 correlation \( \tilde{\rho}_1 (k'', m') \) of the batch means \( \{Y_j (m') : j = 1, \ldots, k''\}. \) If \( \tilde{\rho}_1 (k'', m') < \tilde{\rho}_1 (k', m), \) then go to **Step 1c.** Otherwise, go to **Step 2.**

c. Divide the overall data set of \( n_i \) observations into \( k \) batches of size \( m'' \equiv \ell m. \) Compute \( \overline{Y} (m'', k) \) and \( S^2_{m'', k} \) as in (1) and (2). Compute the half-length \( H_{\alpha, k} = t_{1-\alpha/2, k-1}S_{m'', k}/\sqrt{k} \) of the CI (3). If \( \left| H_{\alpha, k} / \overline{Y} (m'', k) \right| \leq r^*, \) then deliver (3) and stop. Otherwise, go to **Step 2.**

**Step 2.** Update the iteration counter \( i \leftarrow i + 1 \) and the total sample size \( n_i \leftarrow 2n_i - 2. \) Collect the additional \( n_i - n_{i-1} \) observations and go to **Step 1a.**
B. Test Processes

- Response Times in Central Server Model 3 of Law and Carson (1979)
- Waiting Times in $M/M/1$ LIFO Queue with Utilization 0.8
- Waiting Times in $M/H_2/1$ Queue with Utilization 0.8
- Waiting Times in $M/M/1$ Queue with Utilization 0.8
- Waiting Time in $M/M/1/M/1$ Queue with Utilization 0.8 at Each Station
- Reward in Two-State Markov Chain with $P_{00} = P_{11} = 0.99$
C. Performance Measures

- Confidence Interval (CI) Properties
  - Empirical Coverage Probability
  - Average Relative Precision
  - Average Half-Length
  - Variance of Half-Length
- Average Required Sample Size

D. Analysis of Results

Complete results are available via

www.ise.ncsu.edu/jwilson/lada06iiet.pdf
Table 1: Performance of NBM Procedures for the $M/M/1/LIFO$ Queue Waiting Time Process with Utilization $\tau = 0.8$ Based on Independent Replications of Nominal 90% CIs

<table>
<thead>
<tr>
<th>Precision Requirement</th>
<th>Nominal 90% CIs</th>
<th>ASAP</th>
<th>ASAP3</th>
<th>LC</th>
</tr>
</thead>
<tbody>
<tr>
<td>NO PRECISION</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td># replications</td>
<td>100</td>
<td>400</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>avg. sample size</td>
<td>5,025</td>
<td>53,958</td>
<td>3,120</td>
<td></td>
</tr>
<tr>
<td>coverage</td>
<td>72.0%</td>
<td>87.0%</td>
<td>64.0%</td>
<td></td>
</tr>
<tr>
<td>avg. rel. precision</td>
<td>0.210</td>
<td>0.082</td>
<td>0.236</td>
<td></td>
</tr>
<tr>
<td>avg. CI half length</td>
<td>0.652</td>
<td>0.261</td>
<td></td>
<td></td>
</tr>
<tr>
<td>var. CI half length</td>
<td>0.074</td>
<td>0.106</td>
<td></td>
<td></td>
</tr>
<tr>
<td>±15% PRECISION</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td># replications</td>
<td>100</td>
<td>400</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>avg. sample size</td>
<td>14,317</td>
<td>54,017</td>
<td>13,944</td>
<td></td>
</tr>
<tr>
<td>coverage</td>
<td>77.0%</td>
<td>86.8%</td>
<td>76.0%</td>
<td></td>
</tr>
<tr>
<td>avg. rel. precision</td>
<td>0.119</td>
<td>0.081</td>
<td>0.131</td>
<td></td>
</tr>
<tr>
<td>avg. CI half length</td>
<td>0.372</td>
<td>0.260</td>
<td></td>
<td></td>
</tr>
<tr>
<td>var. CI half length</td>
<td>0.004</td>
<td>0.004</td>
<td></td>
<td></td>
</tr>
<tr>
<td>±7.5% PRECISION</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td># replications</td>
<td>100</td>
<td>400</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>avg. sample size</td>
<td>57,539</td>
<td>68,325</td>
<td>74,624</td>
<td></td>
</tr>
<tr>
<td>coverage</td>
<td>82.0%</td>
<td>87.5%</td>
<td>84.0%</td>
<td></td>
</tr>
<tr>
<td>avg. rel. precision</td>
<td>0.062</td>
<td>0.069</td>
<td>0.064</td>
<td></td>
</tr>
<tr>
<td>avg. CI half length</td>
<td>0.196</td>
<td>0.219</td>
<td></td>
<td></td>
</tr>
<tr>
<td>var. CI half length</td>
<td>6.0E–4</td>
<td>5.1E–4</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 2: Performance of NBM Procedures for the $M/H_2/1$ Queue Waiting Time Process with Utilization $\tau = 0.8$ Based on Independent Replications of Nominal 90% CIs

<table>
<thead>
<tr>
<th>Precision Requirement</th>
<th>Nominal 90% CIs</th>
<th>ASAP</th>
<th>ASAP3</th>
<th>LC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td># replications</td>
<td>100</td>
<td>400</td>
<td>100</td>
</tr>
<tr>
<td>NO PRECISION</td>
<td>avg. sample size</td>
<td>16,716</td>
<td>42,022</td>
<td>86,144</td>
</tr>
<tr>
<td></td>
<td>coverage</td>
<td>76.0%</td>
<td>87.8%</td>
<td>88.0%</td>
</tr>
<tr>
<td></td>
<td>avg. rel. precision</td>
<td>0.539</td>
<td>0.2026</td>
<td>0.106</td>
</tr>
<tr>
<td></td>
<td>avg. CI half-length</td>
<td>4.760</td>
<td>1.614</td>
<td></td>
</tr>
<tr>
<td></td>
<td>var. CI half-length</td>
<td>43.730</td>
<td>0.5962</td>
<td></td>
</tr>
<tr>
<td>±15% PRECISION</td>
<td># replications</td>
<td>100</td>
<td>400</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>avg. sample size</td>
<td>148,820</td>
<td>76,214</td>
<td>86,144</td>
</tr>
<tr>
<td></td>
<td>coverage</td>
<td>88.0%</td>
<td>88.0%</td>
<td>88.0%</td>
</tr>
<tr>
<td></td>
<td>avg. rel. precision</td>
<td>0.102</td>
<td>0.1308</td>
<td>0.106</td>
</tr>
<tr>
<td></td>
<td>avg. CI half-length</td>
<td>0.802</td>
<td>1.0329</td>
<td></td>
</tr>
<tr>
<td></td>
<td>var. CI half-length</td>
<td>0.055</td>
<td>0.0273</td>
<td></td>
</tr>
<tr>
<td>±7.5% PRECISION</td>
<td># replications</td>
<td>100</td>
<td>400</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>avg. sample size</td>
<td>405,854</td>
<td>228,482</td>
<td>229,632</td>
</tr>
<tr>
<td></td>
<td>coverage</td>
<td>93.0%</td>
<td>90.0%</td>
<td>90.0%</td>
</tr>
<tr>
<td></td>
<td>avg. rel. precision</td>
<td>0.053</td>
<td>0.07054</td>
<td>0.067</td>
</tr>
<tr>
<td></td>
<td>avg. CI half-length</td>
<td>0.421</td>
<td>0.5623</td>
<td></td>
</tr>
<tr>
<td></td>
<td>var. CI half-length</td>
<td>0.021</td>
<td>2.0E–3</td>
<td></td>
</tr>
</tbody>
</table>
Table 3: Performance of NBM Procedures for the $M/M/1/M/1$ Queue Waiting Time Process with Utilization $\tau = 0.8$ at Each Station Based on Independent Replications of Nominal 90% CIs

<table>
<thead>
<tr>
<th>Precision Requirement</th>
<th>Nominal 90% CIs</th>
<th>ASAP</th>
<th>ASAP3</th>
<th>LC</th>
</tr>
</thead>
<tbody>
<tr>
<td>NO PRECISION</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td># replications</td>
<td>100</td>
<td>400</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>avg. sample size</td>
<td>3,152</td>
<td>19,133</td>
<td>3,120</td>
<td></td>
</tr>
<tr>
<td>coverage</td>
<td>85.0%</td>
<td>91.0%</td>
<td>64.0%</td>
<td></td>
</tr>
<tr>
<td>avg. rel. precision</td>
<td>0.454</td>
<td>0.154</td>
<td>0.236</td>
<td></td>
</tr>
<tr>
<td>avg. CI half-length</td>
<td>3.250</td>
<td>0.983</td>
<td></td>
<td></td>
</tr>
<tr>
<td>var. CI half-length</td>
<td>14.06</td>
<td>0.165</td>
<td></td>
<td></td>
</tr>
<tr>
<td>±15% PRECISION</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td># replications</td>
<td>100</td>
<td>400</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>avg. sample size</td>
<td>46,610</td>
<td>25,522</td>
<td>13,944</td>
<td></td>
</tr>
<tr>
<td>coverage</td>
<td>93.0%</td>
<td>91.5%</td>
<td>76.0%</td>
<td></td>
</tr>
<tr>
<td>avg. rel. precision</td>
<td>0.103</td>
<td>0.119</td>
<td>0.131</td>
<td></td>
</tr>
<tr>
<td>avg. CI half-length</td>
<td>0.649</td>
<td>0.755</td>
<td></td>
<td></td>
</tr>
<tr>
<td>var. CI half-length</td>
<td>0.030</td>
<td>0.037</td>
<td></td>
<td></td>
</tr>
<tr>
<td>±7.5% PRECISION</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td># replications</td>
<td>100</td>
<td>400</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>avg. sample size</td>
<td>117,339</td>
<td>58,844</td>
<td>49,920</td>
<td></td>
</tr>
<tr>
<td>coverage</td>
<td>90.0%</td>
<td>91.3%</td>
<td>87.0%</td>
<td></td>
</tr>
<tr>
<td>avg. rel. precision</td>
<td>0.050</td>
<td>0.069</td>
<td>0.065</td>
<td></td>
</tr>
<tr>
<td>avg. CI half-length</td>
<td>0.318</td>
<td>0.441</td>
<td></td>
<td></td>
</tr>
<tr>
<td>var. CI half-length</td>
<td>0.008</td>
<td>1.7E–3</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Efficiency Analysis

In the “ideal” case that \( \{ X_i : i = 1, 2, \ldots \} \) is stationary and Gaussian with known steady-state variance parameter (SSVP)

\[
\gamma_X \equiv \lim_{n \to \infty} n \text{Var}(\bar{X}_n) = \sum_{\ell = -\infty}^{\infty} \text{Cov}(X_i, X_{i+\ell}),
\]

the nominal 100\((1 - \alpha)\)% CI, \( \bar{X}_n \pm z_{1-\alpha/2} \sqrt{\gamma_X/n} \), is asymptotically valid:

\[
\lim_{n \to \infty} \Pr \left\{ \mu_X \in \bar{X}_n \pm z_{1-\alpha/2} \sqrt{\gamma_X/n} \right\} = 1 - \alpha.
\]

Thus in the ideal case, an efficient procedure yielding an asymptotically valid 100\((1 - \alpha)\)% CI for \( \mu_X \) with relative precision \( r^* \) will require

\[
n^* = \frac{z_{1-\alpha/2}^2 \gamma_X}{(r^* \mu_X)^2}
\]
observations.
Table 4: Comparison of ASAP3’s Average Sample Size \( \bar{n} \) with the Efficient Sample Size \( n^* \) Required by an Asymptotically Valid 90\% CI for \( \mu_X \) with Relative Precision \( r^* \)

<table>
<thead>
<tr>
<th>Output Process</th>
<th>( r^* )</th>
<th>( n^* )</th>
<th>( \bar{n} )</th>
<th>( \bar{n}/n^* )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( M/M/1 ) Queue Waiting Times</td>
<td>15%</td>
<td>53,306</td>
<td>103,742</td>
<td>1.946</td>
</tr>
<tr>
<td>( \tau = 0.9, \mu_X = 9, \gamma_X = 35,901 ) (Steiger et al., 2005)</td>
<td>7.5%</td>
<td>213,222</td>
<td>287,568</td>
<td>1.349</td>
</tr>
<tr>
<td></td>
<td>3.75%</td>
<td>852,886</td>
<td>969,011</td>
<td>1.136</td>
</tr>
<tr>
<td>( M/M/1 ) Queue Waiting Times</td>
<td>15%</td>
<td>14,853</td>
<td>43,796</td>
<td>2.949</td>
</tr>
<tr>
<td>( \tau = 0.8, \mu_X = 3.2, \gamma_X = 1,264.64 )</td>
<td>7.5%</td>
<td>59,412</td>
<td>72,060</td>
<td>1.213</td>
</tr>
<tr>
<td></td>
<td>3.75%</td>
<td>237,650</td>
<td>256,186</td>
<td>1.078</td>
</tr>
<tr>
<td>( M/H_2/1 ) Queue Waiting Times</td>
<td>15%</td>
<td>45,486</td>
<td>76,214</td>
<td>1.676</td>
</tr>
<tr>
<td>( \tau = 0.8, \mu_X = 8, \gamma_X = 24,204.8 )</td>
<td>7.5%</td>
<td>181,942</td>
<td>228,482</td>
<td>1.256</td>
</tr>
<tr>
<td></td>
<td>3.75%</td>
<td>727,765</td>
<td>798,234</td>
<td>1.097</td>
</tr>
<tr>
<td>Cost Function of Highly Correlated 2-state DTMC</td>
<td>15%</td>
<td>1,323</td>
<td>9,553</td>
<td>7.221</td>
</tr>
<tr>
<td>( \mu_X = 7.5, \gamma_X = 618.75 )</td>
<td>7.50%</td>
<td>5,292</td>
<td>12,283</td>
<td>2.321</td>
</tr>
<tr>
<td></td>
<td>3.75%</td>
<td>21,168</td>
<td>46,469</td>
<td>2.195</td>
</tr>
<tr>
<td></td>
<td>1.875%</td>
<td>84,669</td>
<td>105,190</td>
<td>1.242</td>
</tr>
</tbody>
</table>
IV. Conclusions and Recommendations

A. Main Findings

- For steady-state simulation analysis, completely automated autoregressive–batch means procedures such as ASAP3 can be highly effective.

- Such procedures are based on fitting a first-order autoregressive process to the relevant sequence of batch means, where the batch size is sufficiently large to satisfy relatively loose constraints on the lag-one correlation between the resulting batch means—namely, \( \pm 0.8 \) for ASAP3.
ASAP3 was designed as an improved batch means procedure that:

- generally avoids undercoverage problems often observed with ASAP and LC as well as other popular procedures such as ABATCH and LBATCH;
- generally avoids overcoverage problems sometimes observed with ASAP; and
- completely eliminates excessive variability of CI half-length and final sample size sometimes observed with ASAP, especially with no precision requirement.
• ASAP3 has been tested on an extensive suite of test problems.

• ASAP3 outperforms ABATCH, ASAP, LBATCH, and LC in all test problems used so far.

• ASAP3 delivers well-behaved CIs for processes that are strongly correlated, markedly nonnormal, or highly contaminated by initialization bias (that is, warm-up effects)—provided ASAP3 is used with a meaningful precision requirement.
B. Recommendations for Future Work

- Examine the theoretical asymptotic validity and efficiency of the CIs delivered by ASAP3 as the precision specification $H^*$ or $r^*$ tends to zero.

- Test ASAP3 on queueing network models of production and telecommunications systems having realistic levels of complexity, congestion, and workstation utilization.