# A DISTRIBUTION-FREE TABULAR CUSUM CHART FOR
# CORRELATED DATA WITH AUTOMATED VARIANCE ESTIMATION

Joongsup (Jay) Lee
Christos Alexopoulos
David Goldsman
Seong-Hee Kim
Kwok-Leung Tsui

Georgia Institute of Techology
Atlanta, GA 30318, U.S.A.

James R. Wilson

Edward P. Fitts Department of
Industrial and Systems Engineering
North Carolina State University
Raleigh, NC 27695, U.S.A.

## ABSTRACT

We formulate and evaluate distribution-free statistical process control (SPC) charts for monitoring an autocorrelated process when a training data set is used to estimate the marginal mean and variance of the process as well as its variance parameter (i.e., the sum of covariances at all lags). We adapt variance-estimation techniques from the simulation literature for automated use in DFTC-VE, a distribution-free tabular CUSUM chart for rapidly detecting shifts in the mean of an autocorrelated process. Extensive experimentation shows that our variance-estimation techniques do not seriously degrade the performance of DFTC-VE compared with its performance using exact knowledge of the variance parameter; moreover, the performance of DFTC-VE compares favorably with that of other competing distribution-free SPC charts.

## 1 INTRODUCTION

Statistical process control (SPC) charts are often used to monitor key performance measures of production processes, such as the process mean, and to detect any irregularities represented by gradual or sudden shifts in those quantities. In this article, we let $\text{ARL}_0$ denote the in-control average run length—that is, the expected number of observations taken from the monitored process when it is in control (and thus yields the desired value of the selected performance characteristic) before a false out-of-control alarm is raised. Similarly, let $\text{ARL}_1$ denote the average run length associated with a specific out-of-control condition—that is, the expected number of observations taken from the monitored process before a true out-of-control alarm is raised when the selected performance characteristic of the process deviates from the in-control value by a specific amount. Among several SPC charts that yield a user-selected value of $\text{ARL}_0$, we prefer the chart that yields the smallest values

of $\text{ARL}_1$ for a range of relevant out-of-control conditions. However, it is rarely the case that the exact distribution of the monitored process is known to the user of an SPC chart, and there is always the risk of simply assuming a wrong distribution. Beyond the problem of the monitored process having an unknown distribution (which is sometimes markedly nonnormal), in many SPC applications it is simply incorrect to assume that successive observations of the monitored process are independent. Therefore we seek to develop a distribution-free SPC chart for correlated data. Moreover, when developing distribution-free SPC charts, we must use one or more parameters of the monitored process to determine the control limits that yield the desired $\text{ARL}_0$. We formulate DFTC-VE, a distribution-free tabular CUSUM chart in which the variance parameter and the chart's control limits are computed from a training data set automatically—that is, without the need for any intervention by the user.

We assume that the Phase I process $\{X_i : i = 1, \ldots, n\}$ is covariance stationary with $\mu \equiv \text{E}[X_i]$ and $\sigma^2 \equiv \text{E}\big[(X_i - \mu)^2\big]$ respectively denoting the marginal mean and variance of the process. The usual sample mean and variance of the training data set,

$$\widehat{\mu} = \overline{X}(n) = n^{-1} \sum_{i=1}^{n} X_i \qquad (1)$$

and

$$\widehat{\sigma}^2 = S_n^2 = (n-1)^{-1} \sum_{i=1}^{n} (X_i - \widehat{\mu})^2, \qquad (2)$$

are used to estimate $\mu$ and $\sigma^2$, respectively.

By a certain abuse of notation that should cause no confusion, we will always use $\mu$ and $\sigma^2$ to denote the marginal mean and variance of the data set (either in Phase I or Phase II) that is currently under discussion. Thus in both Phase I and Phase II, we write the in-control condition as

$\mu = \mu_0$, where is $\mu_0$ is the desired value of the process mean. To obtain concrete results, we assume that $\widehat{\sigma}^2$ converges to $\sigma^2$ fast enough as the size $n$ of the training data set grows so that $\widehat{\sigma}^2$ is a sufficiently accurate estimator of $\sigma^2$; moreover, we assume that a shift in the process mean in Phase II is not accompanied by a change in the value of the marginal variance.

For the DFTC-VE chart, we will also need an estimator of the variance parameter

$$\Omega^2 \equiv \lim_{n \to \infty} n \mathrm{Var}\big[\overline{X}(n)\big] = \sum_{\ell = -\infty}^{\infty} \mathrm{Cov}(X_i, X_{i+\ell}) \qquad (3)$$

based on the training data set. As for the marginal variance, we assume that a shift in the process mean in Phase II is not accompanied by a change in the value of the variance parameter (3).

## 2 VARIANCE ESTIMATORS

### 2.1 Standardized Time Series Overlapping Area Estimator

Alexopoulos *et al.* (2007a) show that the so-called standardized time series (STS) overlapping area estimator for $\Omega^2$ outperforms a number of other variance estimators in terms of bias and variance; therefore we chose to incorporate this variance estimator into a version of DFTC-VE. First, we form $n - m + 1$ overlapping batches from the training data, each of size $m$, so that the $i$th batch consists of the observations $\{X_{i+j} : j = 0, \ldots, m-1\}$ for $i = 1, \ldots, n-m+1$. Although the sample-size-to-batch-size ratio $b \equiv n/m$ is fixed, we let the batch size $m$ increase so that the overall sample size $n = bm$ also increases; and all this data must be taken from the training data set to compute the STS estimator of $\Omega^2$. The sample mean of the $i$th overlapping batch with (intermediate) batch size $j$ is denoted by

$$\overline{X}_i^O(j) \equiv \frac{1}{j} \sum_{\ell=0}^{j-1} X_{i+\ell} \text{ for } i = 1, \ldots, n-m+1 \text{ and } j = 1, \ldots, m.$$

Finally, the overlapping STS area estimator for $\Omega^2$ is defined by

$$A^O(f; b, m) \equiv \frac{1}{n-m+1} \sum_{i=1}^{n-m+1} \big[Z_i^O(f; m)\big]^2, \qquad (4)$$

where for a prespecified, continuously differentiable weight function $f(\cdot)$ defined on $[0, 1]$, we let for $i = 1, \ldots, n-m+1$

$$Z_i^O(f; m) \equiv \frac{1}{m^{3/2}} \sum_{j=1}^{m} f\Big(\frac{j}{m}\Big) j \big[\overline{X}_i^O(m) - \overline{X}_i^O(j)\big] \qquad (5)$$

Other weight functions and even other STS estimators for $\Omega^2$ are available for use; our selection here has been based on the comparatively good analytical and empirical performance of the overlapping area estimator (Alexopoulos *et al.* 2007a). The remaining unresolved problem is the selection of the batch size $m$, which affects the bias and variance of the STS estimator (4) as well as the convergence of (4) to its limiting distribution—which is approximately a scaled chi-squared distribution with mean $\Omega^2$ as explained in Alexopoulos *et al.* (2007b).

For use with DFTC-VE, we propose an automated batch-size determination algorithm that uses the same sequential procedure as in Lada and Wilson (2006, 2007) and Lada *et al.* (2008); but instead of using nonoverlapping batch means as the basic data items to be tested for independence and normality, we use the weighted "STS-area" statistics similar to (5) that are computed from nonoverlapping batches in the training data set. For $i = 1, \ldots, b$, the $i$th nonoverlapping batch of size $m$ in the training data set is $\{X_{(i-1)m+j} : j = 1, \ldots, m\}$; and we seek a sufficiently large value of $m$ so that the corresponding weighted "STS-area" statistics $\{Z_i(f; m) : i = 1, \ldots, b\}$ computed from (6) and (7) below approximately constitute a random sample from a normal distribution. Such a batch size $m$ is large enough to ensure approximate convergence of the final overlapping area estimator $A^O(f; b, m)$ to its appropriate limiting distribution; see Alexopoulos *et al.* (2007a, 2007b) for the explicit asymptotic distribution and the scaled chi-squared approximation to the asymptotic distribution. A formal statement of our batch-size determination algorithm for the STS overlapping area estimator is given below.

**Remark 1** *Using the final batch size m delivered by the algorithm given above, we see that the quantity $Z_i^O(f; m)$ in (5) is the signed area under the standardized time series defined on the ith overlapping batch $\{X_{i+j} : j = 0, \ldots, m-1\}$ of size m taken from the entire training data set so that we take $i = 1, \ldots, n-m+1$.*

**Remark 2** *To ensure sufficiently small bias as well as adequate convergence of the STS estimator to its limiting distribution, the final batch size delivered by the algorithm above is usually taken to be three times the batch size required to pass the independence test in step [3] and the normality test in step [5]. This inflation of the final batch size also improves the adequacy of the Brownian-motion approximation to the behavior of the CUSUM statistics used by the J&B and the DFTC charts. If the size of the training data set, n, is not large enough to ensure that the $\{Z_i^O(f; m) : i = 1, \ldots, n'\}$ pass both the independence test in step [3] and the normality in step [5], then we take the final batch size $m \leftarrow \lfloor n/20 \rfloor$, which ensures that $b = n/m \geq 20$. This design for the batch-size algorithm ensures that no matter how the algorithm terminates, the final STS variance*

## Batch-Size Algorithm for the STS Overlapping Area Estimator of $\Omega^2$

**[1]**      Take an initial sample of size $n' \leftarrow 4{,}096$ from the training data set of size $n \leftarrow 10{,}000$; and divide the initial sample into $b \leftarrow 256$ adjacent nonoverlapping batches of size $m \leftarrow 16$. Set the randomness test size $\alpha_{\text{ran}} \leftarrow 0.20$. Set the initial normality test size $\alpha_{\text{nor}}(1) \leftarrow 0.05$, the normality test parameter $\tau \leftarrow 0.184206$, and the normality test iteration counter $k \leftarrow 1$.

**[2]**      Compute the following statistics from the $b$ nonoverlapping batches of the current batch size $m$:

$$\overline{X}_i(j) \equiv \frac{1}{j} \sum_{\ell=1}^{j} X_{m(i-1)+\ell} \quad \text{for} \quad i = 1, \ldots, b \quad \text{and} \quad j = 1, \ldots, m, \tag{6}$$

$$Z_i(f;m) \equiv \frac{1}{m^{3/2}} \sum_{j=1}^{m} f\left(\frac{j}{m}\right) j \left[\overline{X}_i(m) - \overline{X}_i(j)\right] \quad \text{for} \quad i = 1, \ldots, b. \tag{7}$$

**[3]**      Apply the von Neumann test for independence (von Neumann 1941) to the $\{Z_i(f;m) : i = 1, \ldots, b\}$ using the significance level $\alpha_{\text{ran}}$. If the randomness test is passed, then go to step **[5]**; otherwise, go to step **[4]**.

**[4]**      Increase the batch size $m$ and update the overall sample size $n'$ according to $m \leftarrow \lfloor\sqrt{2}m\rfloor$ and $n' \leftarrow bm$, respectively.

     **(a)**      If $n' > n$, then return the final batch size $m \leftarrow \lfloor n/20 \rfloor$ and stop.

     **(b)**      Otherwise, obtain the required additional observations from the training data set and go to step **[2]**.

**[5]**      Apply the Shapiro-Wilk normality test (Shapiro and Wilk 1965) to the current set of statistics (7) using the significance level $\alpha_{\text{nor}}(k) \leftarrow \alpha_{\text{nor}}(1) \exp[-\tau(k-1)^2]$. If the normality test is passed with the current batch size $m$, then return the final batch size $m \leftarrow 3m$ and stop; otherwise go to step **[6]**.

**[6]**      Increase the normality test iteration counter $k$, batch size $m$, and overall sample size $n'$ according to $k \leftarrow k+1$, $m \leftarrow \lfloor\sqrt{2}m\rfloor$, and $n' \leftarrow bm$, respectively.

     **(a)**      If $n' > n$, then return the final batch size $m = \lfloor n/20 \rfloor$ and stop.

     **(b)**      Otherwise, obtain the required additional observations from the training data set; recompute the nonoverlapping batch statistics (6) and (7) and go to step **[5]**.

estimator $A^{\text{O}}(f;b,m)$ in (4) has approximately a scaled chi-squared distribution with at least 48 degrees of freedom. *(Alexopoulos et al. 2007b).*

### 2.2 "Quick-and-Dirty" Autoregressive Variance Estimator (QDARVE)

We also use the "Quick-and-Dirty" Autoregressive Variance Estimator (QDARVE), which is based on results of Steiger *et al.* (2005), Lada and Wilson (2007), and Lada *et al.* (2008), and which is simpler to obtain and easier to automate than the STS variance estimators. If we can find a sufficiently large batch size $m$ such that the nonoverlapping batch means $\{\overline{X}_i(m) : i = 1, \ldots, b\}$ computed from the training data set are adequately modeled by an AR(1) process,

$$\overline{X}_i(m) = \mu + \varphi_{\overline{X}(m)}\left[\overline{X}_{i-1}(m) - \mu\right] + \varepsilon_i(m) \tag{8}$$

for $i = 1, 2, \ldots$, then the variance parameter, $\Omega^2_{\overline{X}(m)} \equiv \sum_{l=-\infty}^{\infty} \text{Cov}\left[\overline{X}_i(m), \overline{X}_{i+l}(m)\right]$, of the AR(1) process in (8)

is given by

$$\Omega^2_{\overline{X}(m)} = \Omega^2/m = \text{Var}\left[\overline{X}(m)\right]\left[\frac{1 + \varphi_{\overline{X}(m)}}{1 - \varphi_{\overline{X}(m)}}\right]. \tag{9}$$

As detailed below, the key idea underlying QDARVE is to do the following: (i) determine a batch size $m$ sufficiently large to ensure the approximate validity of the AR(1) model (8) of the nonoverlapping batch means $\{\overline{X}_i(m)\}$ computed from the training data set; (ii) compute approximately unbiased estimators of $\text{Var}\left[\overline{X}(m)\right]$ and $\varphi_{\overline{X}(m)}$ from the training data set; and (iii) insert these estimators into (9) to obtain a simplified, intuitively appealing estimator of the variance parameter $\Omega^2$ for the original (unbatched) process.

In developing and testing the DFTC chart, Kim *et al.* (2007) find that in practice the adequacy of the Brownian motion approximation to the behavior of the CUSUM on which DFTC is based requires a lag-one correlation of at most 0.5 in the time series of basic data items used to compute that CUSUM. To compensate for the additional uncertainty

## QDARVE Algorithm for Estimating $\Omega^2$

**[1]** Choose the initial batch size $m$, where $m = 1$ by default. Choose the initial batch count $b \geq b_{\min}$, where $b_{\min} = 1024$ by default. Take the initial sample of size $n' = mb$ from the training data set of size $n$, where $n = 10{,}000$ by default. Also set the size of the test for acceptable lag-one correlation in the batch means and the upper bound $\zeta$ on the lag-one correlation in (16) as follows: $\alpha_{\mathrm{cor}} \leftarrow 0.01$ and $\zeta \leftarrow 0.4$, respectively.

**[2]** Calculate the sample statistics for $\{\overline{X}_i(m) : i = 1, \ldots, b\}$, the current set of $b$ nonoverlapping batch means based on batch size $m$,

$$
\left.\begin{array}{c}
\displaystyle \overline{X}(m,b) \equiv b^{-1} \sum_{i=1}^{b} \overline{X}_i(m), \qquad S^2(m,b) \equiv (b-1)^{-1} \sum_{i=1}^{b} \left[\overline{X}_i(m) - \overline{X}(m,b)\right]^2, \\[2ex]
\displaystyle \widehat{\rho}_{\overline{X}(m)}(m,b) \equiv (b-1)^{-1} \sum_{i=1}^{b-1} \left[\overline{X}_i(m) - \overline{X}(m,b)\right]\left[\overline{X}_{i+1}(m) - \overline{X}(m,b)\right] \Big/ S^2(m,b), \\[2ex]
\displaystyle \widehat{\varphi}_{\overline{X}(m)} \equiv 2\widehat{\rho}_{\overline{X}(m)}(m,b) - \left[\widehat{\rho}^{(1)}_{\overline{X}(m)}(m,b/2) + \widehat{\rho}^{(2)}_{\overline{X}(m)}(m,b/2)\right] \Big/ 2,
\end{array}\right\} \tag{10}
$$

where $\widehat{\rho}_{\overline{X}(m)}(m,b)$ denotes the standard lag-one sample correlation between the batch means based on $b$ nonoverlapping batches each of size $m$; and we let $\widehat{\rho}^{(1)}_{\overline{X}(m)}(m,b/2)$ and $\widehat{\rho}^{(2)}_{\overline{X}(m)}(m,b/2)$ respectively denote similar estimators of the lag-one correlation between the batch means based on the first and last $b/2$ nonoverlapping batches each of size $m$.

**[3]** If $\widehat{\varphi}_{\overline{X}(m)} \leq \sin\left[\sin^{-1}(\zeta) - z_{1-\alpha_{\mathrm{cor}}}/\sqrt{b}\right]$ (where $z_{1-\alpha_{\mathrm{cor}}} = \Phi^{-1}(1 - \alpha_{\mathrm{cor}}) = 2.33$), then go to step **[4]**; else estimate the required batch size as follows:

$$
\psi \leftarrow \left\lceil \ln\left\{\sin\left[\sin^{-1}(\zeta) - z_{1-\alpha_{\mathrm{cor}}}/\sqrt{b}\right]\right\} \Big/ \ln\left[\widehat{\varphi}_{\overline{X}(m)}\right] \right\rceil \quad \text{and} \quad m \leftarrow \lceil \mathrm{mid}\{1.1, \psi, 2\}m \rceil,
$$

where $\mathrm{mid}\{u,v,w\}$ denotes the median of the real numbers $u, v, w$; retrieve additional data from the training data set and update $n'$ if necessary to ensure that $b \leftarrow \lfloor n'/m \rfloor$ satisfies $b \geq b_{\min}$; and go to step **[2]**.

**[4]** Compute the estimator of the variance parameter for the batch means,

$$
\widehat{\Omega}^2_{\overline{X}(m)} = S^2(m,b) \left[\frac{b-1}{b - \widehat{C}(m,b)}\right] \left[\frac{1 + \widehat{\varphi}_{\overline{X}(m)}}{1 - \widehat{\varphi}_{\overline{X}(m)}}\right], \tag{11}
$$

where

$$
\widehat{C}(m,b) = \frac{1 + \widehat{\varphi}_{\overline{X}(m)}}{1 - \widehat{\varphi}_{\overline{X}(m)}} - \frac{2\widehat{\varphi}_{\overline{X}(m)}\left[1 - \widehat{\varphi}^b_{\overline{X}(m)}\right]}{b\left[1 - \widehat{\varphi}_{\overline{X}(m)}\right]^2}; \tag{12}
$$

and deliver the final estimator of the variance parameter for the original process,

$$
\widehat{\Omega}^2 = m\widehat{\Omega}^2_{\overline{X}(m)}. \tag{13}
$$

introduced by estimation of the variance parameter, in the formulation of QDARVE we impose the more-stringent upper bound

$$
\varphi_{\overline{X}(m)} = \mathrm{Corr}\left[\overline{X}_i(m), \overline{X}_{i+1}(m)\right] \leq 0.4 \tag{16}
$$

on the basic data items used to compute the CUSUM on which DFTC-VE is based—that is, we require a sufficiently large batch size $m$ so that the nonoverlapping batch means computed from the training data set satisfy (16) with probability close to one. For additional considerations justifying (16), see p. 77 of Bagshaw and Johnson (1975). Steiger *et al.* (2005), Lada and Wilson (2007), and Lada *et al.* (2008) find that if the batch size $m$ is sufficiently large to ensure that $\varphi_{\overline{X}(m)} \leq 0.8$, then their variance estimators similar to QDARVE are sufficiently stable in practice to yield highly

## DFTC-VE: Distribution-Free Tabular CUSUM Chart with Variance Estimation

1. Compute $\widehat{\mu}$, $\widehat{\sigma}^2$, and $\widehat{\Omega}^2$ from the Phase I data set.
2. Set $K = k\widehat{\sigma}$; choose the desired value of the two-sided $ARL_0$; and calculate $H$, the solution to the equation

$$\frac{\widehat{\Omega}^2}{2K^2}\left\{\exp\left[\frac{2K(H+1.166\widehat{\Omega})}{\widehat{\Omega}^2}\right] - 1 - \frac{2K(H+1.166\widehat{\Omega})}{\widehat{\Omega}^2}\right\} = 2ARL_0. \tag{14}$$

3. For $j = 1, 2, \ldots$, use the $j$th observation $Y_j$ from the Phase II data set to update the CUSUMs

$$S^{\pm}(j) \equiv \left\{\begin{array}{ll} 0, & \text{if } j = 0, \\ \max\left\{0, \ S^{\pm}(j-1) \pm (Y_j - \mu_0) - K\right\}, & \text{if } j = 1, 2, \ldots. \end{array}\right\}, \tag{15}$$

and raise an out-of-control alarm if $S^+(j) \geq H$ or $S^-(j) \geq H$.

reliable and accurate confidence-interval estimators of the steady-state mean. On the basis of all these considerations, we concluded that if we take the batch size large enough to satisfy (16) with probability close to one, then we should be able to use QDARVE effectively in distribution-free SPC charts requiring an estimator of the variance parameter.

When the lag-one correlation between individual (unbatched, raw) observations is greater than 0.4, Kim *et al.* (2007) recommend the use of batching in Phase II as well as in Phase I to ensure that the lag-one correlation in the data to be monitored does not exceed 0.4; and thus in Phase II, we use the batch size $m$ determined in Phase I so as to satisfy (16) with probability close to one. A formal statement of the algorithm for computing QDARVE is given below.

**Remark 3** *Taking $b_{min} = 1024$ and $\alpha_{cor} = 0.01$ is based on the computational experience of Steiger et al. (2005) with ASAP3 and Lada and Wilson (2007) and Lada et al. (2008) with SBatch.*

**Remark 4** *The formula for $\widehat{\varphi}_{\overline{X}(m)}$ in the third line of (10) is the jackknifed estimator of $\varphi_{\overline{X}(m)}$ and thus has bias of the form $O(b^{-2})$, whereas the standard correlation estimator $\widehat{\rho}_{\overline{X}(m)}(m,b)$ has bias of the form $O(b^{-1})$. This property coupled with the requirement $b \geq b_{min} = 1024$ ensures that $\widehat{\varphi}_{\overline{X}(m)}$ is an approximately unbiased estimator of $\varphi_{\overline{X}(m)}$.*

**Remark 5** *The final variance estimator (13) follows from (9), Remark 4, and the property*

$$S^2(m,b)\left[\frac{b-1}{b-\widehat{C}(m,b)}\right] \text{ is an (approximately) unbiased}$$

*estimator of $Var\left[\overline{X}(m)\right]$.* (17)

*See Lee et al. (2008) for a derivation of (17).*

## 3 DFTC-VE: DISTRIBUTION-FREE TABULAR CUSUM CHART WITH VARIANCE ESTIMATION

Using the estimators (1), (2), (4), and (13), we modify the DFTC chart of Kim *et al.* (2007) to incorporate automated parameter estimation based on a training data set.

**Remark 6** *For the choice of the reference parameter value, Kim et al. (2007) recommend $k = 0.1$. A search method (such as the bisection algorithm) can be used to solve (14).*

## 4 EXPERIMENTS

In our first set of experiments, we compare the performance of the following distribution-free SPC charts in terms of their average run lengths: (i) the DFTC-VE chart using Phase I and II; and (ii) the DFTC chart using Phase II only based on exact knowledge of the required process parameters as in Kim *et al.* (2007). In our second set of experiments, we study how the DFTC-VE chart performs in comparison with other distribution-free SPC charts—namely, the R&W Shewhart chart and the J&B CUSUM chart—to which we added Phase I estimation of the required process parameters.

**The R&W Shewhart Chart with Estimated Variance:**
From the Phase I data set, determine the batch size $m$ to be used in Phase II such that the nonoverlapping batch means are approximately normal with lag-one autocorrelation at most 0.1. Choose the desired value of $ARL_0$ and find $z_{ON}$

such that

$$\frac{m}{1 - \Phi(z_{\text{ON}}) + \Phi(-z_{\text{ON}})} = \text{ARL}_0.$$

Then raise an out-of-control alarm after observation $i = jm$ in Phase II if the batch mean from the $j$th nonoverlapping batch of size $m$ in Phase II satisfies

$$\left| \overline{Y}_j(m) - \mu_0 \right| \geq z_{\text{ON}} \cdot \sqrt{\widehat{\text{Var}}\left[ \overline{X}(m) \right]},$$

where $\widehat{\text{Var}}\left[ \overline{X}(m) \right]$, the estimated variance of the batch means, is computed from the Phase I data set as in the first line of (10).

**The J&B Chart with Estimated Variance Parameter:** Compute for $j = 1, 2, \ldots,$

$$S^{\pm}(j) = \max\left\{ S^{\pm}(j-1) \pm (Y_j - \mu_0), 0 \right\}$$

with $S^{\pm}(0) = 0$. Choose the desired value of the two-sided $\text{ARL}_0$, and set

$$H = \sqrt{2\text{ARL}_0 V^{(\cdot)}},$$

where $V^{(\cdot)}$ is the estimated variance parameter based on batch size $m$ and computed from the Phase I data set. We use two such estimators: the STS variance estimator $V^A$ given by (4); and the quick-and-dirty autoregressive variance estimator $V^Q$ given by (13). Raise an out-of-control alarm after the $j$th observation in Phase II if $S^+(j) \geq H$ or $S^-(j) \geq H$.

We considered stationary exponential autoregressive processes of order 1 (that is, EAR(1) processes) that have i.i.d. exponential (nonnormal) marginals. We refer readers to Lee *et al.* (2008) for experimental results of other test processes, including stationary AR(1) process and the queue-waiting-time process in a steady-state $M/M/1$ queuing system. For each distribution-free chart, we took the desired value of 10,000 for the two-sided $\text{ARL}_0$; and we performed 4,000 independent replications of each chart for the in-control condition and for a range of out-of-control conditions. On each replication of each SPC chart, the training data set had marginal mean $\mu = \mu_0$, the in-control mean; and the working data set had marginal mean $\mu = \mu_0 + \theta\sigma$, where the shift $\theta$ in mean for each out-of-control condition was expressed in standard deviations of the monitored process. The resulting estimated ARLs are summarized below for each test process.

Throughout this section, we use $m_1$ and $m_2$ to denote the batch sizes used in Phase I and Phase II, respectively. Note that a new set of Phase I data is obtained on each replication; and thus our batch-size determination algorithms

deliver a different batch size on each replication. Over 4,000 independent replications of each SPC chart applied to each test process, we computed $\overline{m}_1$ and $\overline{m}_2$, the average batch sizes used in Phase I and Phase II, respectively. On each replication of the R&W chart applied to a test process, in Phase I we set the batch size $m_1$ just large enough so that the sample estimator of the lag-one autocorrelation of the batch means was at most 0.1 as suggested by Runger and Willemain (1995); and then in Phase II we took $m_2 = m_1$. For the CUSUM-type charts (J&B and DFTC-VE), the batch-size determination algorithm varied depending on which variance estimators were used. On each replication of an SPC chart using the STS variance estimator, in Phase I we used the final batch size $m_1$ delivered by the algorithm in Section 2.1; and in Phase II we used the batch size $m_2 = 1$ so as to compensate for the substantial computational overhead associated with the STS batch-size determination algorithm used in Phase I. Thus the SPC charts with the STS variance estimator always used raw (unbatched) observations in Phase II. On each replication of an SPC chart with QDARVE, in Phase I we used the final batch size $m_1$ delivered by the QDARVE batch-size determination algorithm in Section 2.2 to ensure that (16) was satisfied with high probability; and then in Phase II we took $m_2 = m_1$.

### 4.1 EAR(1) Process

The EAR(1) process is defined by

$$Y_i = \begin{cases} (\mu - 1/\lambda) + \varphi_Y[Y_{i-1} - (\mu - 1/\lambda)], & \text{w.p. } \varphi_Y, \\ (\mu - 1/\lambda) + \varphi_Y[Y_{i-1} - (\mu - 1/\lambda)] + \varepsilon_i, & \text{w.p. } 1 - \varphi_Y, \end{cases} \tag{18}$$

for $i = 1, 2, \ldots$ and where $0 < \varphi_Y < 1$ and the $\{\varepsilon_i\}$ are i.i.d. exponential with mean $1/\lambda = \sigma$ so that the $\{Y_i\}$ have mean $\mu$ and variance $\sigma^2 = 1/\lambda^2$. To ensure the stationarity of the EAR(1) process, we sampled $Y_0$ from an exponential distribution with mean $1/\lambda$.

For the Shewhart-type R&W chart, we employed the batch-size determination scheme as discussed earlier in Section (4). Since the marginal distribution of the EAR(1) process is not normal, one needs batch sizes which yield both small enough lag-one correlation and approximately normal batch means. Consequently, the corresponding values of $\text{ARL}_0$ based on the batch sizes large enough to ensure lag-one autocorrelation of batch means at most 0.1 did not even approximately attain the target level of 10,000; and hence the performance of the R&W chart in terms of the values of $\text{ARL}_1$ could not be validated.

The use of variance estimators in the EAR(1) process did not seriously degrade the performance of the DFTC chart; and because the R&W chart failed to attain the desired value of $\text{ARL}_0$, the DFTC-VE charts were most effective in detecting all sizes of shifts at all levels of the correlation coefficient $\varphi_Y$. As Table 1 shows, the "pull-down" effects were again obtained for the DFTC-VE chart based on the

Table 1: Two-sided ARLs in terms of number of raw observations for an EAR(1) process.

| $\varphi_Y$ | Shift $\theta$ | R&W | J&B | | DFTC-VE | | DFTC Phase II |
|---|---|---|---|---|---|---|---|
| | | | Area | QDARVE | Area | QDARVE | |
| | | $\overline{m}_1 = 4$ | $\overline{m}_1 = 92$ | $\overline{m}_1 = 1$ | $\overline{m}_1 = 90$ | $\overline{m}_1 = 1$ | |
| | | $\overline{m}_2 = 4$ | $\overline{m}_2 = 1$ | $\overline{m}_2 = 1$ | $\overline{m}_2 = 1$ | $\overline{m}_2 = 1$ | $\overline{m}_2 = 1$ |
| 0.25 | 0 | 823 | 8820 | 8864 | 10486 | 9837 | 10557 |
| | 0.25 | 430 | 726 | 729 | 283 | 288 | 279 |
| | 0.5 | 247 | 367 | 366 | 112 | 112 | 112 |
| | 0.75 | 139 | 244 | 245 | 70 | 70 | 70 |
| | 1 | 80 | 183 | 184 | 51 | 50 | 51 |
| | 1.5 | 29 | 122 | 122 | 32 | 32 | 32 |
| | 2 | 11 | 92 | 92 | 24 | 24 | 24 |
| | 2.5 | 5 | 73 | 74 | 19 | 19 | 19 |
| | 3 | 4 | 61 | 61 | 16 | 16 | 16 |
| | 4 | 4 | 46 | 46 | 12 | 12 | 12 |
| | | $\overline{m}_1 = 8$ | $\overline{m}_1 = 98$ | $\overline{m}_1 = 2$ | $\overline{m}_1 = 98$ | $\overline{m}_1 = 1$ | |
| | | $\overline{m}_2 = 8$ | $\overline{m}_2 = 1$ | $\overline{m}_2 = 2$ | $\overline{m}_2 = 1$ | $\overline{m}_2 = 1$ | $\overline{m}_2 = 1$ |
| 0.5 | 0 | 1449 | 8918 | 9816 | 10480 | 13385 | 11497 |
| | 0.25 | 737 | 968 | 1032 | 458 | 516 | 451 |
| | 0.5 | 385 | 489 | 520 | 179 | 223 | 183 |
| | 0.75 | 204 | 326 | 345 | 112 | 142 | 113 |
| | 1 | 104 | 244 | 258 | 80 | 103 | 81 |
| | 1.5 | 36 | 163 | 172 | 51 | 67 | 52 |
| | 2 | 14 | 122 | 129 | 38 | 50 | 38 |
| | 2.5 | 9 | 98 | 103 | 30 | 40 | 30 |
| | 3 | 8 | 81 | 86 | 25 | 33 | 25 |
| | 4 | 8 | 61 | 65 | 19 | 25 | 19 |
| | | $\overline{m}_1 = 19$ | $\overline{m}_1 = 201$ | $\overline{m}_1 = 8$ | $\overline{m}_1 = 201$ | $\overline{m}_1 = 5$ | |
| | | $\overline{m}_2 = 19$ | $\overline{m}_2 = 1$ | $\overline{m}_2 = 8$ | $\overline{m}_2 = 1$ | $\overline{m}_2 = 5$ | $\overline{m}_2 = 1$ |
| 0.7 | 0 | 2826 | 8941 | 9535 | 10973 | 11331 | 12062 |
| | 0.25 | 1214 | 1339 | 1406 | 748 | 955 | 735 |
| | 0.5 | 564 | 668 | 705 | 301 | 434 | 305 |
| | 0.75 | 284 | 444 | 470 | 184 | 282 | 187 |
| | 1 | 141 | 331 | 353 | 132 | 207 | 135 |
| | 1.5 | 45 | 221 | 235 | 84 | 137 | 86 |
| | 2 | 22 | 166 | 177 | 61 | 102 | 63 |
| | 2.5 | 19 | 132 | 143 | 49 | 82 | 50 |
| | 3 | 19 | 110 | 120 | 40 | 68 | 41 |
| | 4 | 19 | 83 | 91 | 30 | 52 | 31 |
| | | $\overline{m}_1 = 58$ | $\overline{m}_1 = 496$ | $\overline{m}_1 = 29$ | $\overline{m}_1 = 496$ | $\overline{m}_1 = 22$ | |
| | | $\overline{m}_2 = 58$ | $\overline{m}_2 = 1$ | $\overline{m}_2 = 29$ | $\overline{m}_2 = 1$ | $\overline{m}_2 = 22$ | $\overline{m}_2 = 1$ |
| 0.9 | 0 | 4685 | 9121 | 10202 | 11897 | 9782 | 12930 |
| | 0.25 | 2267 | 2475 | 2650 | 1941 | 2262 | 1842 |
| | 0.5 | 1094 | 1246 | 1311 | 761 | 1088 | 779 |
| | 0.75 | 560 | 822 | 865 | 468 | 714 | 484 |
| | 1 | 312 | 610 | 653 | 331 | 530 | 349 |
| | 1.5 | 109 | 404 | 437 | 212 | 351 | 224 |
| | 2 | 62 | 304 | 331 | 155 | 264 | 165 |
| | 2.5 | 58 | 241 | 268 | 122 | 213 | 130 |
| | 3 | 58 | 201 | 225 | 101 | 179 | 108 |
| | 4 | 58 | 151 | 172 | 75 | 136 | 80 |

STS variance estimator $V^A$, yielding closer conformance to the desired value of $ARL_0$ than that of the DFTC chart and often yielding smaller values of $ARL_1$. The use of batching in Phase II increased the values of $ARL_1$, especially for high values of $\varphi_Y$; but nevertheless variance estimation did not seem to degrade the performance of the distribution-free CUSUM-type charts compared with the DFTC chart in terms of the values of $ARL_1$.

As pointed out earlier in this section, the R&W chart delivered substantially smaller values of $ARL_0$ than the target value; and this anomalous behavior clearly revealed that the R&W chart cannot be effectively applied to non-normal processes such as the EAR(1) process. Although

not presented here, one could have used different batch sizes in conjunction with the R&W charts; but then the drop in efficiency of the R&W chart to detect even large shifts would be inevitable because the basic observations of this chart are batch means based on large batch sizes. The DFTC-VE chart is again most effective in detecting all sizes of the shifts at all values of $\varphi_Y$.

## 5 CONCLUSION

In this article, we developed and evaluated DFTC-VE, a distribution-free tabular CUSUM chart for autocorrelated data, which incorporates automated variance estimation from a training data set. We examined how the variance estimation affected the performance of DFTC-VE in comparison with that of its predecessor DFTC, which requires exact knowledge of the relevant process parameters. Moreover, we compared the performance of DFTC-VE with that of other distribution-free SPC charts. When used in conjunction with all the distribution-free charts considered in this article, the proposed variance-estimation procedures did not cause dramatic performance degradation in comparison with the corresponding chart performance using exact knowledge of the relevant process parameters. Also, among the different distribution-free charts considered, we observed that the DFTC-VE chart incorporating the STS variance estimator in (4) performed very well; however, the improved performance came at the expense of additional algorithmic and computational complexity. The QDARVE estimator—which is much easier to understand, implement, and apply—also performed well in conjunction with the DFTC chart.

For this study, we assumed that the out-of-control condition only involved a shift in the process mean, with no other changes in the performance characteristics of the monitored process. While the shift detection in the mean of a process is certainly of major interest, various other measures of importance should be included for monitoring study in conjunction with the DFTC chart. This is the subject of ongoing research.

## REFERENCES

Alexopoulos, C., N. T. Argon, D. Goldsman, G. Tokol, and J. R. Wilson. 2007a. Overlapping variance estimators for simulation. *Operations Research* 55 (6):1090–1103.

Alexopoulos, C., N. T. Argon, D. Goldsman, N. M. Steiger, G. Tokol, and J. R. Wilson. 2007b. Efficient computation of overlapping variance estimators for simulation. *INFORMS Journal on Computing* 19 (3):314–327.

Bagshaw, M. and R. A. Johnson. 1975. The effect of serial correlation on the performance of CUSUM Tests II. *Technometrics* 17 (1):73–80.

Fishman, G. S. and L. S. Yarberry. 1997. An implementation of the batch means method. *INFORMS Journal on Computing* 9 (3):296–310.

Jensen, W. A., L. A. Jones-Farmer, C. W. Champ, and W. H. Woodall. 2006. Effects of parameter estimation on control chart properties: A literature review. *Journal of Quality Technology* 38 (4):349–364.

Johnson, R. A. and M. Bagshaw. 1974. The effect of serial correlation on the performance of CUSUM tests. *Technometrics* 16 (1):103–112.

Kim, S. -H., C. Alexopoulos, K. -L. Tsui, and J. R. Wilson. 2007. A distribution-free tabular CUSUM chart for autocorrelated data. *IIE Transactions* 39:317–330.

Kim, S. -H. and J. R. Wilson. 2006. A discussion on "Detection of intrusions in information systems by sequential change-point methods" by Tartakovsky et al. *Statistical Methodology* 3:315–319.

Lada, E. K. and J. R. Wilson. 2006. A wavelet-based spectral procedure for steady-state simulation analysis. *European Journal of Operational Research* 174:1769–1801.

Lada, E. K. and J. R. Wilson. 2007. SBatch: A spaced batch means procedure for simulation analysis, in *Proceedings of the 2007 Winter Simulation Conference* ed. S. G. Henderson, B. Biller, M. -H. Hsieh, J. Shortle, J. D. Tew, and R. R. Barton, 463–471. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers. Available via <www.informs-sim.org/ wsc07papers/055.pdf> [accessed March 30, 2008].

Lada, E. K., N. M. Steiger, and J. R. Wilson. 2008. SBatch: A spaced batch means procedure for steady-state simulation analysis. *Journal of Simulation* forthcoming. Available via <www.ise.ncsu .edu/jwilson/ files/sbjos5.pdf> [accessed March 30, 2008].

Lee, J., C. Alexopoulos, D. Goldsman, S. -H. Kim, K. -L. Tsui, and J. R. Wilson. 2008. A distribution-free tabular CUSUM chart for correlated data with automated variance estimation. *Submitted to IIE Transaction.*

Maragah, H. D. and W. H. Woodall. 1992. The effect of autocorrelation on the retrospective X-chart. *Journal of Statistical Computation and Simulation* 40:29–42.

Montgomery, D. C. 2001. *Introduction to Statistical Quality Control.* 4th ed. New York, New York: John Wiley & Sons.

Rowlands, R. J. and G. B. Wetherill. 1991. Quality control, in *Handbook of Sequential Analysis* ed. B. K. Ghosh and P. K. Sen, 563—580. New York, New York: Marcel Dekker.

Runger, G. C. and T. R. Willemain. 1995. Model-based and model-free control of autocorrelated processes. *Journal of Quality Technology* 27 (4):283–292.

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

Shapiro, S. S. and M. B. Wilk. 1965. An analysis of variance test for normality (Complete samples). *Biometrika* 52:591–611.

Siegmund, D. 1985. *Sequential Analysis: Tests and Confidence Intervals* New York, New York: Springer-Verlag.

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.

Steiger, N. M., E. K. Lada, J. R. Wilson, J. A. Joines, C. Alexopoulos, and D. Goldsman. 2005. ASAP3: A batch means procedure for steady-state simulation analysis. *ACM Transactions on Modeling and Computer Simulation* 15 (1):39–73.

von Neumann, J. 1941. Distribution of the ratio of the mean square successive difference to the variance. *The Annals of Mathematical Statistics* 12:367–395.

## AUTHOR BIOGRAPHIES

**JOONGSUP (JAY) LEE** is a Ph.D. student in the School of Industrial and System Engineering at the Georgia Institute of Technology. He received a B.S. in Mathematical Decision Science from University of North Carolina at Chapel Hill in 2005. His research interests include quality control methodology, stochastic modeling, analysis of simulation, and applied statistics. His e-mail address is <jlee@gatech.edu>.

**CHRISTOS ALEXOPOULOS** is an Associate Professor in the School of Industrial and Systems Engineering at the Georgia Institute of Technology. His research interests are in the areas of simulation, statistics, and optimization of stochastic systems. He is a member of INFORMS and an active participant in the Winter Simulation Conference, having been Proceedings Co-Editor in 1995, Associate Program Chair in 2006, and a member of the Board of Directors since 2007. He is also the Simulation Department Editor of *IIE Transactions*. His e-mail address is <christos@isye.gatech.edu>, and his web page is <www.isye.gatech.edu/~christos>.

**DAVID GOLDSMAN** is a Professor in the School of Industrial and Systems Engineering at the Georgia Institute of Technology. He received his Ph.D. in Operations Research and Industrial Engineering from Cornell University. His research interests are in the areas of simulation, applied statistics, and applications of operations research in medical and humanitarian problems. He is a member of INFORMS and is the IIE Board Representative to the Winter Simulation Conference. His e-mail address is <sman@gatech.edu>.

**SEONG-HEE KIM** is an Associate Professor in the School of Industrial and Systems Engineering at the Georgia Institute of Technology. She received her Ph.D. in Industrial Engineering and Management Sciences from Northwestern University in 2001. Her research interests include optimization via simulation, quality control, and simulation output analysis. She is a member of INFORMS and IIE, and serves on the editorial board of the Simulation Department of *IIE Transactions*. She is also an associate editor in the simulation area of *Operations Research* and the OR/simulation area of *The American Statistician*. Her e-mail and web addresses are <skim@isye.gatech.edu> and <www.isye.gatech.edu/ skim/> respectively.

**KWOK-LEUNG TSUI** is a professor in the School of Industrial and Systems Engineering at the Georgia Institute of Technology. He received his Ph.D. in Statistics from the University of Wisconsin at Madison in 1986. Dr. Tsui was a recipient of the NSF Young Investigator Award in 1992. He was the (elected) President and Vice President of the American Statistical Association Atlanta Chapter in 1992-1993; the Chair of the INFORMS Section in Quality, Statistics, and Reliability in 2000; and the Founding Chair of the INFORMS Section in Data Mining in 2004. He is a fellow of American Statistical Association and US representative in the ISO Technical Committee on Statistical Methods. His current research interests include data mining and surveillance in healthcare and public health, calibration and validation of computer models, bioinformatics, process control and monitoring, and robust design and Taguchi method. His e-mail address is <ktsui@isye.gatech.edu>.

**JAMES R. WILSON** is professor in the Edward P. Fitts Department of Industrial and Systems Engineering at North Carolina State University. He is a member of AAUW, ACM, and ASA; and he is a Fellow of IIE and INFORMS. His e-mail address is <jwilson@ncsu.edu>, and his Web page is <www.ise.ncsu.edu/jwilson>.