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

EXPLORING EXPONENTIALLY WEIGHTED MOVING AVERAGE CONTROL CHARTS TO DETERMINE THE WARM-UP PERIOD

Manuel D. Rossetti
Zhe Li
Peng Qu

Industrial Engineering Department
4207 Bell Engineering Center
University of Arkansas
Fayetteville, AR 72701, U.S.A.

ABSTRACT

In this paper, we examine the use of exponentially weighted moving average (EWMA) control charts for the detection of initialization bias in steady state simulation experiments. EWMA charts have the interesting property of being more sensitive to shifts in the data as compared to other control charting techniques. We exploit this sensitivity by developing a criteria for searching for the deletion point when the EWMA is applied to the reversed data sequence. This allows us to more easily detect and count the number of times the smoothed sequence remains in control. Our results indicate that the procedure can quickly find and recommend a deletion point. In addition, the properties of the resulting estimators are good if the dataset that is being analyzed does not have an overtly large amount of biased data points. We use experimental test cases to illustrate the properties of the procedure.

1 INTRODUCTION

Consider the output stochastic process $\{Y_i\}$ of the simulation. Let $F_i(y \mid I)$ be the conditional cumulative distribution function of $\{Y_i\}$ where I represents the initial conditions used to start the simulation at time 0. If $F_i(y \mid I) \to F(y)$ when $i \to \infty$, for all initial conditions *I*, then $F(y)$ is called the steady state distribution of the output process. (Law & Kelton, 2002). In steady state simulation, we are often interested in estimating parameters of the steady state distribution, $F(y)$, such as the steady state mean, μ. Because the initial distributions $F_i(y \mid I)$ tend to depend more heavily on the initial conditions, estimators of μ such as the sample average, $\bar{Y} = (1/n)\sum_{i=1}^{n} Y_i$, will tend to be biased. This is the so called initialization bias problem in steady state simulation. Unless we can generate the initial conditions of the simulation according to $F(y)$, which we do not know, we must focus on methods that detect and/or mitigate the presence of initialization bias. One strategy for initialization bias mitigation is to find an index, d, for the output process, $\{Y_i\}$, so that $\{Y_i; i = d+1, \ldots\}$ will have substantially similar distributional properties as the steady state distribution $F(y)$. This is called the simulation warm up problem, where d is called the warm up point, and $\{i = 1, \ldots, d\}$ is called the warm up period for the simulation.

Over the years, many methods and rules have been proposed to detect the warm up period. We refer the interested reader to Wilson and Pritsker (1978), Lada et al. (2003), Litton and Harmonsoky (2002), White et al. (2000), Cash et al. (1992), and Rossetti and Delaney (1995) for an overview of such methods. Robinson (2002) presents a Statistical Process Control (SPC) approach for estimating the length of the warm-up period. In Robinson (2002), the relationships between SPC and steady state simulation output analysis are clarified. The Shewhart Control Chart is used to detect whether the model is running under steady state. A series of procedures for implementing this approach are discussed in detail in the paper. To determine whether the selected warm-up period is sufficiently long to remove initialization bias, Schruben (1982) and Schruben et al. (1983) initialization bias test was applied. In the example application, the length of warm-up period identified by the SPC approach is significantly reduced comparing with the method described by Welch (1983). According to the paper, the SPC approach provides a set of clear rules for determining when the model enters steady state. The disadvantage is that more data are required in terms of replication and run length in order to get normalized and uncorrelated data, which is required by the discussed SPC approach.

In this paper, we examine the potential of using exponentially weighted moving average (EWMA) control charting techniques to help in identifying the warm up period. The EWMA control chart was introduced by Roberts

(1959). The exponentially weighted moving average is defined as follows:

$$Z_i = \lambda Y_i + (1-\lambda) Z_{i-1} \quad (1)$$

where $0 < \lambda < 1$ and the starting value $Z_0 = \mu_0$. EWMA has been extensively studied and used for a long time due to its robustness to non-normality and its ability to detect initial small shifts in the process. In an EWMA control chart, the control limits are calculated as: (See Montgomery 2001 for a detailed discussion)

$$UCL = \mu_0 + L\sigma\sqrt{\frac{\lambda}{(2-\lambda)}\left[1-(1-\lambda)^{2i}\right]} \quad (2)$$

$$\text{Center Line} = \mu_0 \quad (3)$$

$$LCL = \mu_0 - L\sigma\sqrt{\frac{\lambda}{(2-\lambda)}\left[1-(1-\lambda)^{2i}\right]} \quad (4)$$

where σ is the variance of the underlying data and L is a constant used to control the average run length or probability of being out of control. As we can see from (2) and (4), the control limits are functions of observation indices i. Hence, the control limits for initial data are narrow. As time goes on, the control limits will eventually converge towards their process limits. Due to this characteristic, EWMA control charts are more powerful than Shewhart based control charts to detect the initial out-control points. Because of these characteristics, we decided to examine the use of EWMA for initialization bias detection.

Our procedure is of the class of initialization bias detection/mitigation methods that we term, *post-run* methods. Post-run methods are applied to an already collected dataset. Welch's procedure as well as Robinson's SPC procedure are within this class of methods. The other class of initialization bias detection/mitigation methods is termed *within-run* methods. Post-run methods have the disadvantage of having to store all the data necessary for the analysis, but by virtue of that fact can possibly take better advantage of the data. Within-run methods have the advantage of being applied during the simulation run and thus do not necessarily have to store all of the data. Within-run methods have the disadvantage of having a limited view of the data, i.e. they can only react to data as it is collected. The references within Spratt (1998) discuss a variety of both of these classes of methods.

In the following section, we discuss the basic procedure that we will be studying. Then, in the following sections, we present initial results of applying the method to various simulation output processes. Finally, we discuss our plans for future work in this area.

2 THE BASIC IDEA

The basic idea behind our approach is best illustrated from the examination of control charts applied to some sample data. The application of SPC methods and in our case the EWMA depends on the initial estimation of the center line, μ_0 and the process variance, σ^2, two quantities that we are interested in estimating via simulation in the first place. In other words, if we knew μ_0 we would not need to simulate and would not have an initialization bias problem to worry about. A key question in the application of SPC techniques to the initialization bias problem is how to overcome the problem of estimating μ_0 and σ^2.

In the figures that follow, we generated sample data from functions for which we can know the appropriate deletion point. Data from these functions will also be used as part of the testing of our procedure. The appendix to this paper describes the development and characteristics of these sample data generating functions. For the function illustrated in Figure 1, the appropriate deletion point is 380 which we will denote as d^* The data illustrated in Figure 1 has a monotonically decreasing bias function with uncorrelated random noise with a true mean, $\mu_0 = 0.0$, after d^*.

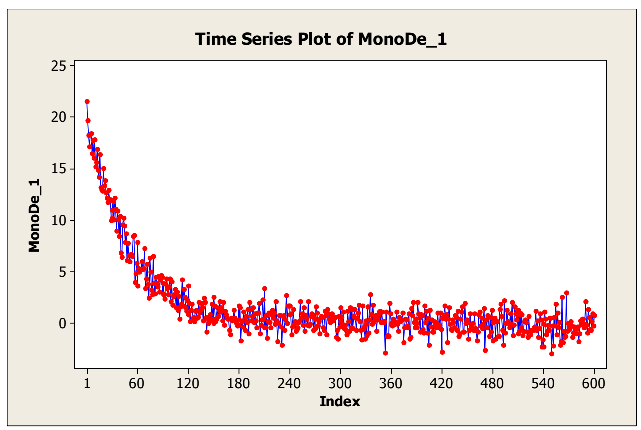

Figure 1: Sample Data from Monotonically Decreasing Bias Function with Random Noise

Figure 2 shows the application of an EWMA chart to the entire data set based on estimating μ_0 and σ^2 from the data past the known deletion point. In other words, we deleted the data prior to d^* and formed estimators for μ_0 and σ^2 from the remaining data. In this case, we simply used the sample average and the sample variance as estimators for μ_0 and σ^2 as per the following equations.

$$\bar{Y}(d) = \frac{\sum_{d+1}^{n} Y_i}{n-d} \quad (5)$$

$$S^2(d) = \frac{\sum_{d+1}^{n} \left(Y_i - \overline{Y}(d)\right)^2}{n-d-1} \tag{6}$$

with $d = d^*$. The values of L and λ directly affect the UCL and LCL and thus the average run length properties of the EWMA control chart. In this example, we set the smoothing constant $\lambda = 0.05$ and the average run length constant $L = 3$. Three sigma limits ($L = 3.$) work reasonably well for our settings of λ and are standard values in practice. In general the settings of these parameters will affect the behavior of our procedures; however, further optimization of these parameters will be saved for future work.

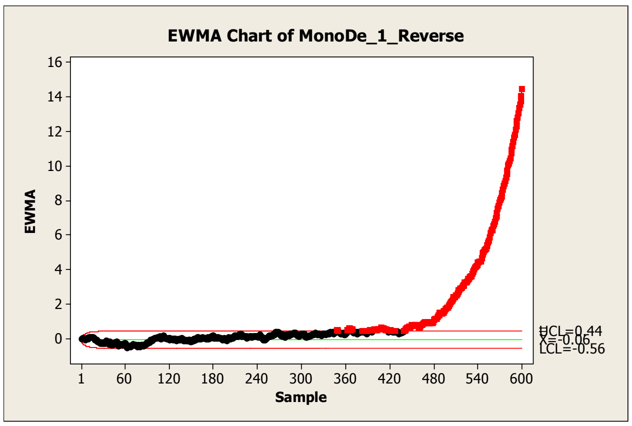

Figure 2 : EWMA Control Chart ($d = 380$)

The control chart given in Figure 2 was developed *backwards* That is, we take the last data point Y_n and consider it as the first data point when applying Equations (2)-(4). Thus, we reversed the sequence of the data. We do this because we assume that the initialization bias will be in the early part of the simulation run and we want to "detect" when we have gone out of control. We might consider the control chart given in Figure 2 as the best we can do since is based on the true deletion point, $d = d^*$. From Figure 2, we can also note that a large proportion of the data falls within the control limits, to the right-side of the deletion point, as we might expect.

Now let us consider the case of not deleting enough of the data. In Figure 3 we illustrate an EWMA control chart applied to the same data set as illustrated in Figure 1 except that we estimated μ_0 and σ^2 based on $d = 10$. As indicated in Figure 3 the centerline is significantly higher because the points we did not delete are "biased". The control chart limits are a little wider because the biased data is further away from the true centerline. This also increases the variability in the data. Remember that since the data has been reversed for this illustration, we have points going out of control at the end of the chart, indicating some bias in the early part of the un-reversed sequence. Even for the case of not deleting enough data, the control chart is telling us something about the presence of bias. In addition, we can clearly see that the proportion of points that fall within the control limits has decreased.

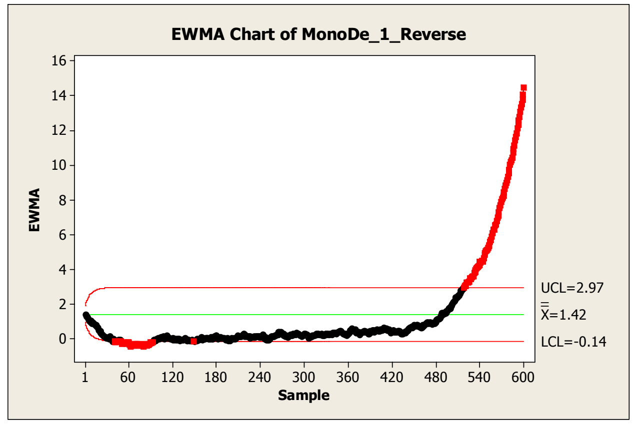

Figure 3 : EWMA Control Chart ($d = 10$)

Now let us consider the case of deleting too much data. In Figure 4, we illustrate an EWMA control chart applied to the same data set as illustrated in Figure 1 except that we estimated μ_0 and σ^2 based on $d = n - 10$. As indicated in Figure 4, the centerline is close to the true centerline because we deleted the biased points. In addition, we can see that the proportion of points that fall within the control limits is high; however, the estimators for μ_0 and σ^2 will not be as good as in the case of Figure 1 because we have only used 10 data points to form the estimates. In fact, the control chart limits are wider than for the case illustrated in Figure 2.

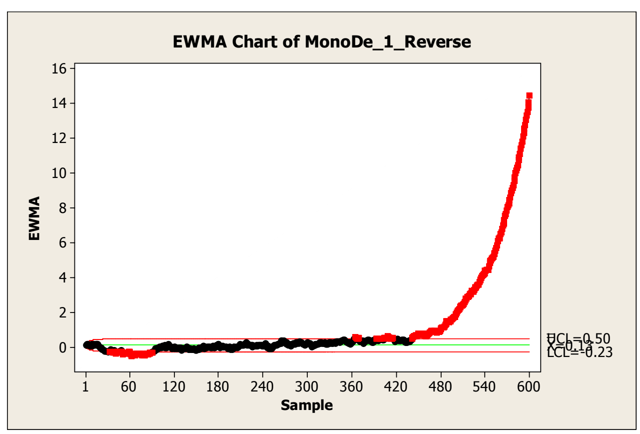

Figure 4 : EWMA Control Chart ($d = 590$)

Consideration of all three figures indicates that there is a trade-off in the application of the control charts. If the control chart is well-formed (centerline is estimated from non-biased data) then we can expect a high proportion of the data to fall within the control limits. If the control chart is not well-formed (centerline is biased from included data), then we can expect a lower proportion of the data to fall within the control limits. This motivates us to look for the smallest deletion point that trades-off the proportion of the data that is in control versus the amount of data deleted.

3 DESCRIPTION OF THE APPROACH

In our approach, we consider the proportion of the data not deleted that remains in control. To be specific, we define $\hat{p}_2(d)$ to be the observed proportion of exponentially weighted data points falling within the control limits to the right side of deletion point, i.e. the observed proportion of the exponentially weighted data remaining after deletion that fall within the control limits. Define an indicator variable to indicate whether a exponentially weighted point is in control or not.

$$I_i(d) = \begin{cases} 1 & \text{if } Z_i \in [LCL(d), UCL(d)] \\ 0 & \text{otherwise} \end{cases} \quad (7)$$

Therefore, $\hat{p}_2(d)$ can be defined as:

$$\hat{p}_2(d) = \frac{\sum_{i=d+1}^{n} I_i(d)}{n-d} \quad (8)$$

Notice that this observed proportion will depend on the choice of d. Our procedure depends on understanding and then exploiting the properties of $\hat{p}_2(d)$.

Figure 5 presents $\hat{p}_2(d); d = 0,\ldots,n-2$ for the same data set that was discussed in Figures 1-3. Notice how in the figure, $\hat{p}_2(d)$ starts out less than 1.0 for small values of d and slowing increases towards 1.0, as d increases. We examined many such figures of $\hat{p}_2(d)$ for a variety of bias shapes and a similar pattern was evident in most of the cases. For cases with high lag-1 correlation ($\rho \geq 0.5$), we began to see less of a monotonic pattern to the curve. We examine the effect of high correlation in our experiments. While we do not, as of yet, have a formal proof of this pattern, it is exactly this behavior that we will try to exploit in our procedure.

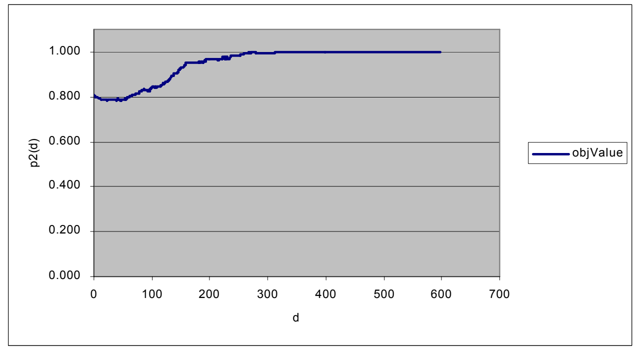

Figure 5: $\hat{p}_2(d)$ *vs. d*

In fact, for this particular case $d^* = 380$, we see that $\hat{p}_2(d)$ is very near 1.0 for $d > 280$. The values of d are based on the original ordering of the data. This indicates the $\hat{p}_2(d)$ can give us an indication of how to set d. Assuming that $\hat{p}_2(d)$ follows the monotonically increasing pattern as d increases allows us to utilize a basic search procedure for determining the recommended deletion point. We denote the recommended deletion point as d_r.

In our procedure, the user must specify the parameters for the EWMA control chart, λ and L, and the user must provide the desired proportion of the exponentially weighted data remaining after deletion that fall within the control limits. We denote this desired proportion as p_2'. This is a relatively easy criteria for users to understand. Obviously, the user would want this proportion to be high after the deletion. We examine the effect of setting this criteria on d_r within our experiments. Our procedure then becomes finding the smallest d such that $\hat{p}_2(d) \geq p_2'$. Why the smallest d? There is an explicit trade-off between the bias in using $\bar{Y}(d)$ to estimate μ_0 and the variance associated with using $\bar{Y}(d)$. As more data is deleted, the bias related to using $\bar{Y}(d)$ to estimate μ_0 decreases, but the variance of $\bar{Y}(d)$ increases. Thus, there is an intuitive desire to delete as little of the data as possible. Notice that because of the assumed monotonic property of $\hat{p}_2(d)$, the search criteria $\hat{p}_2(d) \geq p_2'$, can be recast as a root finding problem for $\hat{p}_2(d) - p_2' = 0$. Therefore, we do not have to evaluate the entire function, $\hat{p}_2(d)$. In our procedure, we use a simple binary search to find d_r.

Our procedure is as follows:

1. Collect $\{Y_i; i = 1,2,\ldots,n\}$ and reverse the data
2. Set p_2', λ, L
3. $a \leftarrow 0, b \leftarrow n$

4. $d \leftarrow \lfloor (a+b)/2 \rfloor$
5. if $\hat{p}_2(d) < p_2'$
   $a \leftarrow a, b \leftarrow d$
   else
   $a \leftarrow d, b \leftarrow b$
6. if $a = b$ return $d_r = a$
   else go to step 4

We note that reversing the data makes a difference when applying the procedure since the sequence of Z_i will be different depending on the direction of application of the control chart. With the direction of the data reversed, the early Z_i tend to be smoother because they tend to be based on unbiased data in the original data.

Notice that in step 5, we compute $\hat{p}_2(d)$ using equations (2-4), equations (7) and (8) with μ_0 and σ^2 estimated via equations (5) and (6). The procedure starts the search with an initial deletion point at $n/2$. At each iteration, the interval is reduced by half, depending on whether or not $\hat{p}_2(d) < p_2'$. The limits of the search interval are updated based on whether or not $\hat{p}_2(d) < p_2'$. When the interval has been reduced such that the lower limit equals the upper limit, we have found the recommended deletion point that meets the criteria. Notice we use integer values within the search. In addition to the above mentioned steps, one could also test the remaining data after deletion for initialization bias using Schruben's initialization bias test. See the description in Nelson(1992) for an easy to implement algorithmic representation of Schruben's test. In the following section, we discuss our experiments to examine the quality of our procedure. In addition, we discuss the results of our experiments.

4 EXPERIMENTS AND RESULTS

Our experiments involved two types of assessment. First, we examined the behavior of the method over a set of experimental design points using data generated from the aforementioned data generating functions. We then examined the performance of the procedure by applying it to the waiting times from a M/M/1 queue, a difficult test case for initialization bias detection procedures. Table 1 presents the experimental factors and their levels.

Table 1 Experimental Factors and Levels

Factor Name	Low Level	High Level
sample size	500	1000
d* ratio	0.2	0.8
shape	1: straight line	3: exponential
variance	1	10

In the experiments indicated in Table 1, we varied the sample size (the number of data points generated), the percentage of the points having bias, the shape of the bias deterioration, and the variance of the generated points. The shape indicated by 1 has a linearly decreasing bias until d^* is reached. The shape indicated by 3 has an exponentially decreasing bias until d^* is reached. In Table 2, we present the average bias, variance, and mean squared error of the estimator based on 10 replications of datasets with 20% of the data being biased . As indicated in Table 2, the bias has been reduced and d_r is near d^* ; however, d_r tends to be less than d^* .

Table 2: Results for $d^* = 0.2 \times n$

	$d^* = 0.2 \times n$			
	$n = 500,\ d^* = 100$		$n = 1000,\ d^* = 200$	
	$\sigma^2 = 1$	$\sigma^2 = 10$	$\sigma^2 = 1$	$\sigma^2 = 10$
	Avg. (s.e.)	Avg. (s.e.)	Avg. (s.e.)	Avg. (s.e.)
bias	0.019 (0.015)	0.054 (0.056)	0.004 (0.008)	0.019 (0.047)
var	1.029 (0.026)	9.910 (0.225)	1.009 (0.021)	10.348 (0.174)
MSE	1.031 (0.026)	9.941 (0.229)	1.009 (0.021)	10.368 (0.173)
d_r	85.60 (0.859)	67.400 (1.899)	211.8 (34.49)	193.0 (49.05)
bias(q1)	-0.011 (0.017)	-0.072 (0.054)	-0.006 (0.008)	-0.057 (0.048)
bias(q3)	-0.005 (0.019)	-0.062 (0.053)	0.003 (0.012)	-0.043 (0.056)
bias(q5)	-0.021 (0.024)	-0.007 (0.065)	0.013 (0.012)	-0.047 (0.063)

In Table 3, we present the average bias, variance, and mean squared error of the estimator based on 10 replications of datasets with 80% of the data being biased . As indicated in Table 8, the bias has been somewhat reduced but still remains significant. In addition, the trend for d_r to be less than d^* appears to be more strongly indicated.

Table 3: Results for $d^* = 0.8 \times n$

	$d^* = 0.8 \times n$			
	$n = 500,\ d^* = 400$		$n = 1000,\ d^* = 800$	
	$\sigma^2 = 1$	$\sigma^2 = 10$	$\sigma^2 = 1$	$\sigma^2 = 10$
	Avg. (s.e.)	Avg. (s.e.)	Avg. (s.e.)	Avg. (s.e.)
bias	0.135 (0.021)	0.620 (0.130)	0.027 (0.020)	0.338 (0.093)
var	1.031 (0.036)	10.421 (0.239)	1.027 (0.035)	10.380 (0.340)
MSE	1.053 (0.038)	10.957 (0.335)	1.031 (0.035)	10.573 (0.380)
d_r	361.7 (2.745)	302.6 (8.036)	749.9 (2.447)	662.0 (16.20)
bias(q1)	0.058 (0.020)	0.446 (0.130)	-0.028 (0.021)	0.184 (0.092)
bias(q3)	0.002 (0.017)	0.152 (0.138)	-0.049 (0.025)	0.068 (0.080)
bias(q5)	-0.008 (0.020)	0.031 (0.114)	-0.054 (0.030)	0.051 (0.070)

Figure 7 summarizes Tables 2 and 3 by providing 95% confidence intervals on the remaining bias for each of the factors/levels. From Figure 7, we concluded that, as the percentage of biased data points increases, our methodology is unable to adequately indicate the deletion point. We found that our procedure tends to consistently underestimate d^* and that this tendency is detrimental when there are a large amount of data points that are biased. As we would expect, the procedure performs slightly worse as the variance is increased and for shape 1. In shape 1, the bias lingers longer than for the exponentially decreasing bias function.

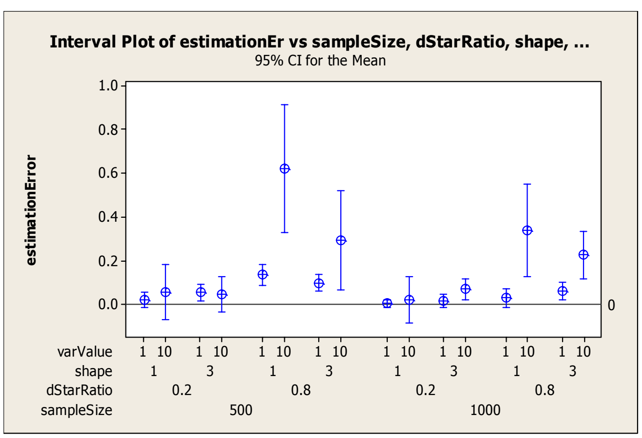

Figure 7: 95% Confidence Intervals on Bias

As we have indicated, d_r tends to be underestimated. In order to examine this effect, we shift d_r towards the right according to the formula $q\lfloor n - d_r \rfloor$, where the amount of shifting we tested was defined by *q= 0.1, 0.3, 0.5*. Tables 2 and 3 indicate the average bias remaining after these shifts. Figure 8 provides the 95% confidence intervals on the remaining bias for each of the factors/levels for the case of *q= 0.5*. As indicated in Figure 8, this shifting clearly improves the bias removal.

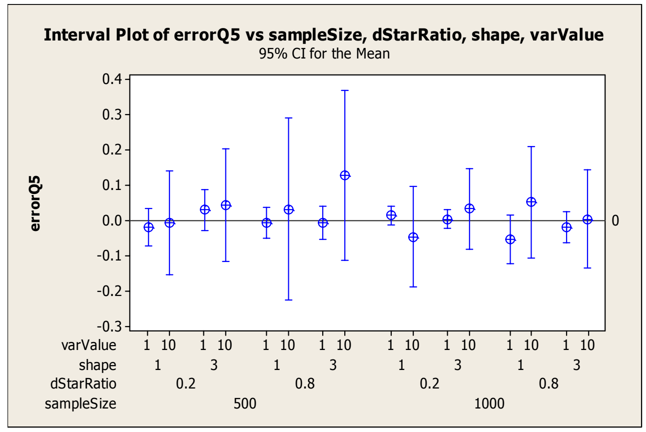

Figure 7: 95% C. I.'s on Bias for *q= 0.5*

We also performed additional sets of experiments similar to those already presented except that the total amount of data generated was n = 10000 and n = 20000. As indicated in Figure 8, our procedure has a difficult time removing the bias when a large portion of the data is biased. In this case, all factors/levels remained the same. Thus, we have significantly more data points that contain bias for these sample sizes.

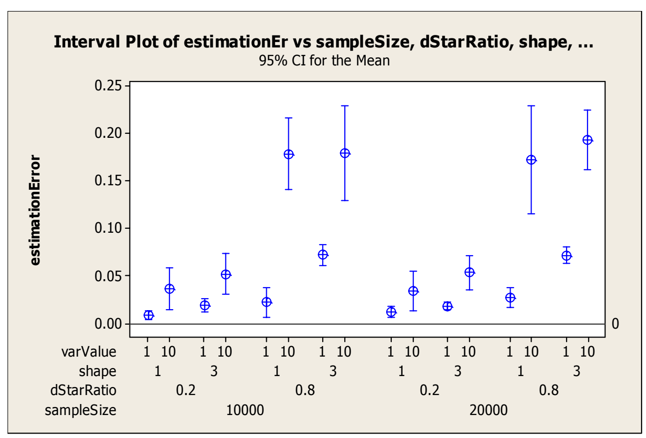

Figure 8: 95% C. I.'s on Bias for *n=10K, 20K*

In addition to the previous experiments, we utilized our procedure to determine the deletion point for the mean waiting time in the queue for the M/M/1 queue. We performed two experiments with the queue utilization, $\rho = 0.2, 0.8$. Each experiment was replicated 10 times for

20000 customers. It is well known that the customer waiting times are positively correlated for such queueing systems. Thus, we decided to apply our method to various batch sizes to examine the effect of reducing the correlation in the data. Figure 9 and 10 present the results for batch sizes of 10, 100, and 200.

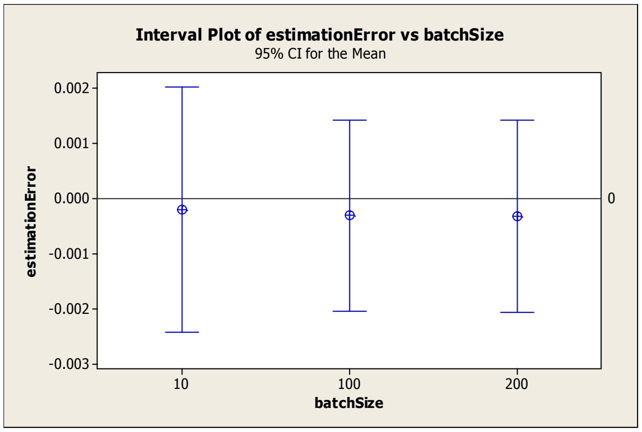

Figure 9: 95% C. I.'s on Bias, $\rho = 0.2$

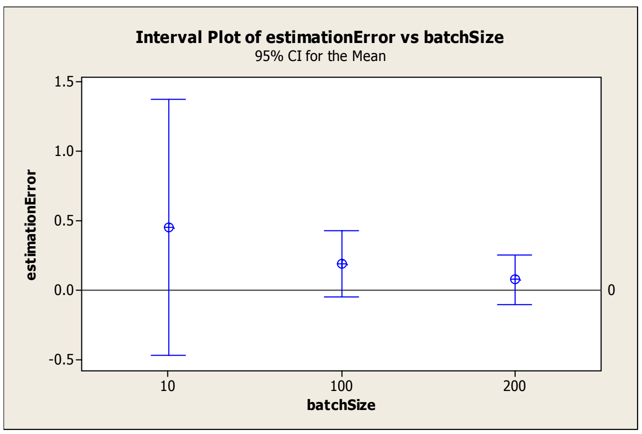

Figure 10: 95% C. I.'s on Bias, $\rho = 0.8$

In both cases show in Figures 9 and 10, the bias appears to have been mitigated and that the batching appears to help the procedure, especially in the case of higher utilization.

We performed additional experimentation that is not reported in this paper. From all of our experimentations, we found the following.

1. A large value for p_2' is always better as long as no significant autocorrelation exists. This is primarily due to the fact that $\hat{p}_2(d)$ increases very quickly towards 1.0.
2. When the proportion of the data that contains bias is reduced the procedure is better able to mitigate the bias. When the proportion of the data that contains bias is increased, the procedure is less able to mitigate the bias. This is due to the tendency of the procedure to consistently underestimate d^*. From our investigation, the procedure underestimates d^* probably due to the fact that $\hat{p}_2(d)$ tends to increase very quickly to 1.0.
3. We also performed some preliminary experimentation on λ and the results indicated that a smaller value of λ improves the estimation. Typical values for λ in EWMA charts are between 0.05 and 0.10 and we found that this was appropriate in our preliminary experiments.
4. For the larger values of σ^2 we found that the procedure was less able to mitigate the bias. This has some interaction with the percentage of data that was biased. As a larger portion of the data has bias, σ^2 tends to naturally be larger.
5. Shape does have an effect on the results and there may be interactions between shape and the batch size factor.
6. Autocorrelation in the data set is detrimental to the procedure; however, batching can be used to remove the autocorrelation. There was some indication that an "optimal" batch size may exist given the interaction with other factors, such as shape, variance, correlation, and λ.

While the results presented here may appear negative, we are actually quite pleased with the results. In fact, the most negative result, that the procedures tends to consistently underestimate d^*, has us the most intrigued. If in fact our recommended deletion point is close to but less than d^* then we might be able to exploit this fact as a lower bound on d^*. An obvious augmentation would be to reapply the method after deleting the initially recommended data points, especially since our procedure performs better when there are less biased data points. In addition, if we could find a good upper bound on d^* then it might be possible to combine the lower and upper bounds to provide a better deletion point. Although preliminary, the results for the M/M/1 are encouraging, especially since the shape of the sample path can vary significantly from what we have investigated to date.

5 SUMMARY AND FUTURE RESEARCH

In this paper, we presented a new method for determining the warm-up period in discrete-event simulation experiments. The method is based on the use of the exponentially weighted moving average statistical process control charts. This control charting technique is well known but has not been previously used for initialization bias detection. The control chart has useful capabilities for detecting small shifts in the process mean, which we felt would increase its effectiveness when used for initialization bias detection. We developed a criteria, $\hat{p}_2(d)$, the observed proportion of the exponentially weighted data remaining after deletion that fall within the control limits, that we exploited to determine the recommended deletion point.

While a more rigorous theoretical basis for our procedure is yet to be established, we show through our experiments that out method shows some promise as a deletion point determination procedure. It is easy to implement and relatively quick computationally. The experiments indicate that the procedure tends to under estimate d^*, the true deletion point; however, this is not necessarily a bad situation. We found, that as the procedure is currently applied, the procedure has difficulty discerning the bias when the bias is within the noise of the underlying process. This tends to cause d^* to be underestimated, but if the bias is indistinguishable from the noise we probably have a relatively unbiased estimate. Thus, we would prefer to have d^* to be small, so as to improve the variance of the estimator. This tendency might be useful in developing better procedures based on our approach.

Based on the results, we are excited about the possibilities for future work in this area. We plan to try to understand better some of the theoretical assumptions that are necessary to have $\hat{p}_2(d)$ behave monotonically. In addition, we have yet to exploit all of the theory that has been developed to set optimal parameter values for the EWMA chart. The many possibilities for future research that we are considering include:

1. Examining additional bias generating function shapes and their affects
2. Utilizing response surface methods to optimally tune the parameters (e.g. λ, L, p_2') of the procedure across a wide variety of test cases. These tuned parameter settings could be used as defaults when applying the procedure.
3. Exploring further the effect of batching to reduce correlation when applying the procedure
4. Comparing our procedure to other initialization bias detection procedures utilizing an experimental design and test case methodology
5. Utilizing additional criteria during the search to improve the sensitivity of detecting bias, i.e. to push d_r closer to d^*.
6. Building on our procedure to develop a within-run procedure for determining the length of the warm up period

We are especially interested in tackling items (5) and (6) and have already developed tentative procedures for those items.

APPENDIX: BIAS FUNCTIONS

We used the following two functions to construct two initial bias patterns for our experiments:

Shape 1:

$$f_1(x) = \begin{cases} a + bx, 0 \le x \le -\dfrac{a}{b} \\ 0, x > -\dfrac{a}{b} \end{cases}, \quad a > 0,\ b < 0$$

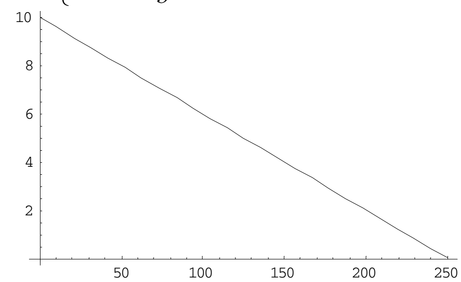

Figure A-1: Diagram of Monotonically Decreasing Function

We combined the function with an AR(1) process with 0 autocorrelation (white noise), the time series plot is shown in Figure A-2 below.

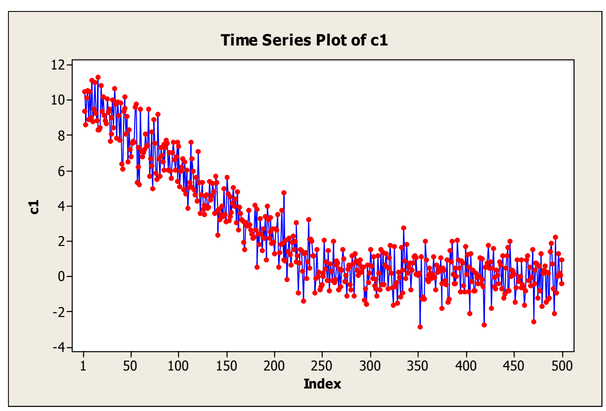

Figure A-2: Time Series Plot of Monotonically Decreasing Function

Shape 3:

$$f_3(x) = \begin{cases} a \times b^{-x} & a > 0,\ b > 1 \\ 0 & |f_3(x) \le 0.01| \end{cases},$$

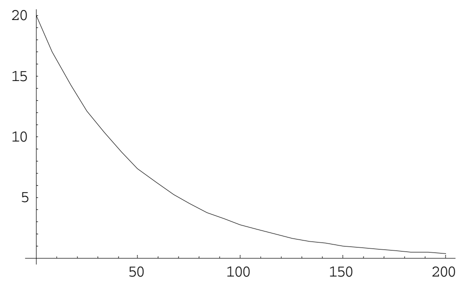

Figure A-3: Diagram of Monotonically Exponential Decreasing Function

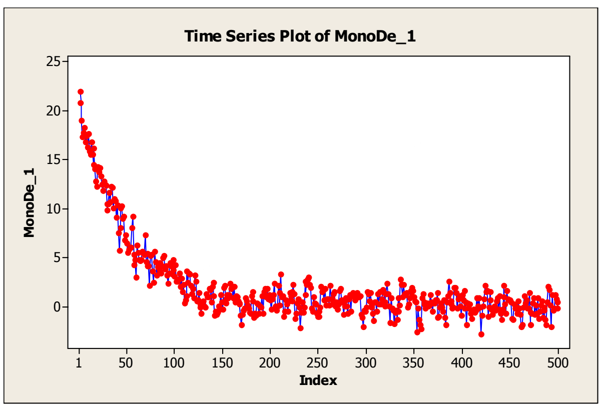

Figure A-4: Time Series Plot of Monotonically Exponential Decreasing Function

REFERENCES

Cash, C.R., D.G. Dippold, J.M. Long, B.L. Nelson, and W.P. Pollard. 1992. Evaluation of tests for initial conditions bias. *Proceedings of the 1992 Winter Simulation Conference*, eds., J.J. Swain, D. Goldsman, R.C. Crain, and J.R. Wilson, 577-585. I Piscataway, NJ: Institute of Electrical and Electronics Engineers.

Lada, E. K., Wilson, J. R., Steiger, N. M. 2003. A Wavelet-Based Spectral Method for Steady-State Simulation Analysis, *Proceedings of the 2003 Winter Simulation Conference*, eds. S. Chick, P. J. Sánchez, D. Ferrin, and D. J. Morrice, 422-430. Piscataway, NJ: Institute of Electrical and Electronics Engineers.

Litton, J. R. and Harmonosky, C. H. 2002. A Comparison of Selective Initialization Bias Elimination Methods,. *Proceeding of the 2002 Winter Simulation Conference*, eds. E. Yucesan, C.-H. Chen, J. L. Snowdon, and J. M. Charnes, 1951-1957. Institute of Electrical and Electronics Engineers, Piscataway, NJ.

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

Nelson, B.L. 1992. Initial-condition bias. In *Handbook of Industrial Engineering*, 2nd ed., ed., G. Salvendy. New York: John Wiley.

Montgomery, D.C. 2001. *Introduction to Statistical Quality Control*, John Wiley & Sons Inc.

Roberts, S.W. 1959. Control chart tests based on geometric moving averages, *Technometrics*, 1.

Robinson, S. 2002. A statistical process control approach for estimating the warm-up period. *Proceedings of the 2002 Winter Simulation Conference,* 439-446. Piscataway, NJ: Institute of Electrical and Electronics Engineers.

Rossetti, M. D., and P. J. Delaney. 1995. Control of initialization bias in queueing simulations using queueing approximations. *Proceedings of the 1995 Winter Simulation Conference*, ed. C. Alexopoulos, K. Kang, W. R. Lilegdon, and D. Goldsman. 322-329. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers.

Schruben, L. 1982. Detecting initialization bias in simulation output. *Operations Research*. 30: 569–590.

Schruben, L., H. Singh and L. Tierney. 1983. Optimal tests for initialization bias in simulation output. *Operations Research* 31: 1167–1 178.

Spratt, S. C. 1998. Heuristics for the startup problem. Unpublished M.S. Thesis, Department of Systems Engineering, University of Virginia.

Welch, P. D. 1983a. The statistical analysis of simulation results. In *Computer Performance Modeling Handbook*, ed. Lavenberg, 267-329.

Welch, P.D., 1983b, A graphical approach to the initial transient problem in steady state simulation, *10th IMACS World Congress on System Simulation and Scientific Computation*, 219-221.

White, K. P., Cobb, M. J., Spratt, S. C. 2000. A comparison of five steady-state truncation heuristics for simulatio. *Proceeding of the 2000 Winter Simulation Conference*, eds. J. A. Joines, R. R. Barton, K. Kang, and P. A. Fishwick, 755-760. Piscataway, NJ: Institute of Electrical and Electronics Engineers.

Wilson, J.R. and A.A.B. Pritsker. 1978. A survey of research on the simulation startup problem. *Simulation*, 31(2): 55-58.

Yucesan, E., Van Wassenhove, L. N., Papanikas, K., & Van Dijk, N. M., 2001. On-line error bounds for steady-state approximations: A potential solution to the initialization bias problem. Proceedings of the 2001 Winter Simulation Conference, 460-467. Piscataway, NJ: Institute of Electrical and Electronics Engineers.

AUTHOR BIOGRAPHIES

MANUEL D. ROSSETTI, Ph.D., P.E. is an Associate Professor in the Industrial Engineering Department at the University of Arkansas. He received his Ph.D. in Industrial and Systems Engineering from The Ohio State University. Dr. Rossetti has published over thirty-five journal and conference articles in the areas of transportation, manufacturing, health care and simulation and he has obtained over $1.5 million dollars in extra-mural research funding. His research interests include the design, analysis, and optimization of manufacturing, health care, and transportation systems using stochastic modeling, computer simulation, and artificial intelligence techniques. He was selected as a Lilly Teaching Fellow in 1997/98 and has been nominated three times for outstanding teaching awards. He is currently serving as Departmental ABET Coordinator. He serves as an Associate Editor for the International Journal of Modeling and Simulation and is active in IIE, INFORMS, and ASEE. He served as co-editor for the WSC 2004 conference. His email and web addresses are `rossetti@uark.edu` and `www.uark.edu/~rossetti`.

ZHE LI is a Ph.D. student in Industrial Engineering at University of Arkansas. He is currently completing his dissertation in the area of multi-criteria optimization applied to network optimization.

PENG QU, Ph.D. is an Industrial Engineer at Advanced Micro Devices. He received his M.S. and Ph.D. degrees in Industrial Engineering from the University of Arkansas. He is a member of IIE and INFORMS.