Proceedings of the 1992 Winter Simulation Conference
ed. J. J. Swain, D. Goldsman, R. C. Crain, and J. R. Wilson

CONFIDENCE INTERVALS FOR UNIVARIATE DISCRETE-EVENT SIMULATION OUTPUT USING THE KALMAN FILTER

Randall B. Howard
Mark A. Gallagher
Kenneth W. Bauer Jr.
Peter S. Maybeck

Air Force Institute of Technology
Wright-Patterson AFB, Ohio 45433-6583

ABSTRACT

The construction of confidence intervals for discrete-event simulation parameters must account for the correlated nature of simulation output. Through the determination of system equations and application of the Kalman filter to simulation output data, a new confidence interval construction technique has been developed. The technique uses Multiple Model Adaptive Estimation (MMAE) to obtain a nonsymmetric confidence interval for the mean estimator of a univariate output sequence. A Monte Carlo analysis of data generated from simulations of $M/M/1$ queues was used to compare the performance of the proposed techniques with other published techniques.

1 INTRODUCTION

This paper is organized in four sections. The first section discusses output analysis for discrete-event simulations, including a brief description of several published techniques for confidence interval construction. The next section develops a proposed Kalman filter confidence interval construction technique. Included are discussions on the Kalman filter, model formulation, Multiple Model Adaptive Estimation (MMAE), and the steps involved in the proposed confidence interval construction technique. The third section reports the results of a Monte Carlo analysis on data from simulations of $M/M/1$ queues. The final section provides a brief summary. The research presented in this paper draws heavily upon results found in Howard (1992) and Gallagher (1992).

1.1 Discrete-Event Simulations

In discrete-event simulation, a typical output analysis objective is to obtain estimates of various output parameters. Often, a confidence interval is constructed to give the analyst a better measure of an estimate's reliability. This research concentrates on constructing confidence intervals for parameters of discrete-event simulations. One class of discrete-event simulations is infinite-horizon simulations, in which one is interested in estimating steady-state parameters. Although it is possible that these parameters may be cyclic, we focus on output sequences that attain a stationary steady-state probability distribution. Often simulation output contains transient data which may bias parameter estimates. In this research, the effect of the start-up problem was diminished by truncating a large part of the simulation output.

Typically, the simulation output sequence is positively correlated, and classical statistics for independent observations do not apply. Several techniques, each of which deal with the correlation problem in different ways, have been proposed for estimating confidence intervals based on one long simulation run (Fishman 1971, Fishman 1978, Law 1983, Meketon and Schmeiser 1984, and Schruben 1983). The proposed Kalman filter technique offers a novel approach for addressing the correlation issue and constructing confidence intervals. Before discussing this approach, brief descriptions of four popular techniques are provided. These techniques will be used in the Monte Carlo analysis to provide comparisons for the developed technique.

1.2 Confidence Interval Construction Techniques

Four popular techniques for confidence interval construction were chosen for comparison in this study. These techniques are nonoverlapping batch means, overlapping batch means, standardized time series, and autoregressive time series. Let $\{Y_1, Y_2, \ldots, Y_n\}$ be a sequence of univariate simulation output. An estimate of the grand mean can be obtained using

$$\bar{Y}(n) = \frac{1}{n}\sum_{i=1}^{n} Y_i \qquad (1)$$

With an estimate of the variance of the sample mean, $\hat{\sigma}^2_{\bar{Y}}$, a confidence interval can be constructed using

$$\bar{Y}(n) \pm t_{d,1-\alpha/2}\sqrt{\hat{\sigma}^2_{\bar{Y}}} \tag{2}$$

where $t_{d,1-\alpha/2}$ is the upper $1-\alpha/2$ critical value for a t distribution with d degrees of freedom. The difference in the four confidence interval techniques is the estimation of $\hat{\sigma}^2_{\bar{Y}}$.

Nonoverlapping batch means (NOBM) seeks to eliminate the problem of correlated observations by dividing the simulation output into b batches, each of m sequential observations (Law 1983). Batch means are calculated using

$$Y_i(m) = \frac{1}{m}\sum_{j=1}^{m} Y_{(i-1)m+j}, \text{ where } i = 1, \ldots, b$$

and the grand mean is:

$$\bar{Y}(n) = \frac{1}{b}\sum_{i=1}^{b} Y_i(m)$$

The batch means are assumed uncorrelated and an estimate of the sample mean variance is found using

$$\hat{\sigma}^2_{\bar{Y}} = \frac{\hat{\sigma}^2_{Y_i(m)}}{b} \tag{3}$$

where

$$\hat{\sigma}^2_{Y_i(m)} = \sum_{i=1}^{b} \frac{(Y_i(m) - \bar{Y}(n))^2}{b-1}$$

A confidence interval is obtained using Equations (1), (2), and (3) with degrees of freedom d of $b-1$.

Overlapping batch means (OBM), introduced by Meketon and Schmeiser (1984), is very similar to nonoverlapping batch means. They contend that the number and size of batches is more important than the independence between batches. Therefore, they allow the batches to overlap and calculate $n-m-1$ batch means. The OBM estimator of the mean estimator variance is:

$$\hat{\sigma}^2_{\bar{Y}} = \frac{m}{n}\sum_{j=1}^{n-m+1} \frac{(\bar{Y}_j(m) - \bar{Y}(n))^2}{n-2m+1}$$

where $\bar{Y}_j(m) = \frac{1}{m}\sum_{i=0}^{m-1} Y_{j+i}$ is the batch mean of size m beginning with observation Y_j. In the same manner as nonoverlapping batch means, a confidence interval can be constructed. They suggest $d = 1.5\frac{n}{m} - 1$ is the appropriate number of degrees.

Schruben (1983) introduced several techniques for confidence interval construction based on standardized time series. We used Schruben's standardized sum methodology (STDS) in this paper. This method separates the run into b batches, exactly as in the method of batch means, and calculates individual batch means $Y_i(m)$ and a grand mean $\bar{Y}(n)$ in the same way. This presentation follows Law (1983). If n is large, then $\bar{Y}(n)$, will be unbiased and approximately normally distributed with variance τ^2/n, where $\tau^2 = \lim_{n\to\infty} n \text{ Var}[\bar{Y}(n)]$. Define A as:

$$A = \frac{12}{(m^3-m)}\sum_{j=1}^{b}\left\{\sum_{l=1}^{m}\sum_{i=1}^{l}[\bar{Y}_j(m) - Y_{i+(j-1)m}]\right\}^2$$

For a fixed number of batches b, A will be asymptotically (as $m \to \infty$) distributed as τ^2 times a chi-square random variable with b degrees of freedom and asymptotically independent of $\bar{Y}(n)$. The A statistic is calculated, and a variance estimate of the sample mean is:

$$\hat{\sigma}^2_{\bar{Y}} = \frac{A}{bn} \tag{4}$$

A confidence interval is obtained using Equations (1), (2), and (4) with degrees of freedom d of n.

Fishman's (1971, 1978) autoregressive methodology (AUTO) fits the simulation output to a pth order autoregressive process, AR(p):

$$\eta_i = \phi_1\eta_{i-1} + \phi_2\eta_{i-2} + \ldots + \phi_p\eta_{i-p} + w_d(t_i) \tag{5}$$

where $\{w_d(t_i)\}$ is a sequence of discrete white shocks that are normally-distributed with mean zero and variance of Q_d. Linear least squares is used to obtain estimates for the autoregressive coefficients ϕ_i and the variance Q_d. The method uses a statistical test, with user-specified significance, to determine the smallest order p that will adequately represent the data. The variance of the mean estimator is (Fishman 1978):

$$\hat{\sigma}^2_{\bar{Y}} = \frac{Q_d}{n(1-\phi_1-\phi_2-\ldots-\phi_p)^2} \tag{6}$$

A confidence interval is obtained using Equations (1), (2), and (6) where d is the degrees of freedom and is given by

$$d = \frac{n(1-\phi_1-\phi_2-\ldots-\phi_p)}{2\sum_{s=0}^{p}(2s-p)\phi_s}$$

2 PROPOSED METHODOLOGY

This section develops the proposed multiple model adaptive estimation (MMAE) confidence interval construction technique. First, we provide a basic discussion on the Kalman filter. Then, we provide a discussion on the formulation of a model for simulation output. Next, we discuss Multiple Model Adaptive Estimation (MMAE). Finally, this section concludes with the steps in the proposed confidence interval construction algorithm.

2.1 The Kalman Filter

A state vector is a set of variables such that the present state vector along with future inputs describe the system behavior completely. In a state-space representation, system equations depict the state dynamics and measurement relationships. The Kalman filter is a recursive, state-estimation algorithm based on these relationships. Beginning at a specific time, the state estimate is propagated to the next time index. Using the information from a measurement, the state estimate is updated. The steps in the Kalman filter algorithm iterate between the propagation and the measurement update stages.

This section's development of the Kalman filter follows Maybeck (1979, 1982). Alternative Kalman filter notation and descriptions include those given by Meinhold and Singpurwalla (1983) and Harvey (1989).

In this section, we assume a system which can be described or approximated by a linear stochastic difference equation,

$$\mathbf{x}(t_{i+1}) = \mathbf{\Phi}(t_{i+1}, t_i)\mathbf{x}(t_i) + \mathbf{G}(t_i)w_d(t_i) \quad (7)$$

where $\mathbf{x}$ is the state vector, $\mathbf{\Phi}$ is the transition matrix, $\mathbf{G}$ is the dynamics noise input matrix, and w_d is discrete-time dynamics noise. The sequence of dynamics noise $\{w_d(t_i)\}$ is assumed to be white (uncorrelated in time) and normally distributed, with mean zero and variance $Q_d(t_i)$.

Similarly, the measurement model is:

$$z(t_i) = \mathbf{H}(t_i)\mathbf{x}(t_i) + v(t_i) \quad (8)$$

where z is the measurement, $\mathbf{H}$ is the measurement matrix, and v is the measurement noise. The sequence of measurement noises $\{v(t_i)\}$ is assumed to be white and normally-distributed, with mean zero and covariance $R(t_i)$. In addition, the noise sequences $\{w_d(t_i)\}$ and $\{v(t_i)\}$ are assumed to be uncorrelated with each other. With appropriate system matrices choices, this Kalman filter formulation can include any autoregressive-moving average (ARMA) model. However, in the Kalman filter formulation, the states may be observed with the addition of measurement noise. This paper deals only with univariate simulation output sequences. Therefore, the dynamics noise and its associated variance, the measurement noise and its associated variance, and the measurements are all scalars. The notation throughout this paper reflects this univariate case, but the Kalman filter concepts can easily be extended to the multivariate case (Gallagher 1992).

The discrete-time Kalman filter algorithm is shown below for a time-invariant, linear system with no control inputs, and normally-distributed (Gaussian) zero-mean discrete dynamics noise and measurement noise (Maybeck 1979). Discrete-time implies that the propagation and measurement updates occur only at set intervals, and linear implies that the values of the filter-design system states $\mathbf{x}(t_i)$ do not affect the values of the transition matrix $\mathbf{\Phi}$, dynamics noise dispersion matrix $\mathbf{G}$, or the measurement matrix $\mathbf{H}$. Time-invariant means that the filter-design system matrices $\mathbf{\Phi}$, $\mathbf{G}$, and $\mathbf{H}$ do not change throughout the stochastic process. Since process stationarity is assumed for this application, the variances of the noises, Q_d and R, do not change throughout the stochastic process.

Before showing the two stages of the Kalman filter algorithm, some notation must be explained. The notation's use will be demonstrated in the following discussion of the propagation and measurement update stages. Estimates are indicated by "hat" over the variable, such as the estimated state vector $\hat{\mathbf{x}}$. At each point in time, two estimates of the state vector are encountered. The first is the estimate based on the propagated output of the dynamics equation, before the measurement information is incorporated at that sample time. These estimates from the propagation stage are labeled with a superscript minus sign, $\hat{\mathbf{x}}(t_i^-)$. In contrast, the state estimate resulting from incorporating the latest measurement information are labeled with a superscript plus sign, $\hat{\mathbf{x}}(t_i^+)$. The associated covariance matrices of these state estimates have similar notation: $\mathbf{P}(t_i^-)$ and $\mathbf{P}(t_i^+)$ respectively. At time t_0, the initial state estimate $\hat{\mathbf{x}}(t_0)$ and associated covariance $\mathbf{P}(t_0)$ must be available or assumed.

The first stage in the Kalman filter, the propagation stage, is comprised of two equations. The propagation equation takes the best state estimate at the previous time $\hat{\mathbf{x}}(t_{i-1}^+)$, and moves it through time by multipling by the transition matrix $\mathbf{\Phi}$. Therefore, the propagation equation is

$$\hat{\mathbf{x}}(t_i^-) = \mathbf{\Phi}\hat{\mathbf{x}}(t_{i-1}^+) \quad (9)$$

The associated covariance matrix $\mathbf{P}(t_i^-)$ is calculated with the following formula:

$$\mathbf{P}(t_i^-) = \mathbf{\Phi}\mathbf{P}(t_{i-1}^+)\mathbf{\Phi}^T + \mathbf{G}Q_d\mathbf{G}^T$$

These two equations, one for the state estimate and the other for the associated covariance, complete the propagation stage.

The second stage in the Kalman filter is the measurement update. The measurement update stage incorporates the information available from a measurement z_i. The Kalman filter gain $\mathbf{K}(t_i)$ is given by

$$\mathbf{K}(t_i) = \mathbf{P}(t_i^-)\mathbf{H}^T[\mathbf{H}\mathbf{P}(t_i^-)\mathbf{H}^T + R]^{-1}$$

The residual r_i is the difference between the current measurement and the best prediction of the measurement before it arrived:

$$r_i = z_i - \mathbf{H}\hat{\mathbf{x}}(t_i^-) \tag{10}$$

Multiplying the Kalman filter gain times the residual results in the new information which is added to the previous state estimate $\hat{\mathbf{x}}(t_i^-)$ in order to obtain the updated state estimate $\hat{\mathbf{x}}(t_i^+)$. Therefore, the updated state estimate is:

$$\hat{\mathbf{x}}(t_i^+) = \hat{\mathbf{x}}(t_i^-) + \mathbf{K}(t_i)[z_i - \mathbf{H}\hat{\mathbf{x}}(t_i^-)] \tag{11}$$

Along with an updated state estimate, an updated covariance matrix is also calculated. Since the Kalman filter gain $\mathbf{K}(t_i)$ is the relative weight based on comparing the variances of the estimate from the propagation stage and the measurement model, it is reasonable that this same gain is used to determine the reduction in variance resulting from incorporating the measurement information:

$$\mathbf{P}(t_i^+) = \mathbf{P}(t_i^-) - \mathbf{K}(t_i)\mathbf{H}\mathbf{P}(t_i^-)$$

Each measurement update stage is followed by a propagation stage, as the two-stage cycle of the Kalman filter begins again.

An important observation is that the state estimate covariances $\mathbf{P}(t_i^-)$ and $\mathbf{P}(t_i^+)$ do not depend on the actual measurements. If the system model $\{\mathbf{\Phi}, \mathbf{G}, \mathbf{H}\}$ are known, and the dynamics driving noise variance Q_d and the measurement noise variance R are known, then the Kalman filter gains $\mathbf{K}(t_i)$ and the covariances $\mathbf{P}(t_i^-)$ and $\mathbf{P}(t_i^+)$ are completely determined. Therefore, these values can be precomputed and stored prior to the actual running of the Kalman filter.

In addition, the Kalman filter gain $\mathbf{K}(t_i)$ and the covariance matrices $\mathbf{P}(t_i^-)$ and $\mathbf{P}(t_i^+)$ are functions of the initial estimate $\mathbf{P}(t_0)$ and the dynamics noise variance Q_d and the measurement noise variance R. Since Q_d and R are constant (as are $\mathbf{\Phi}$, $\mathbf{G}$, and $\mathbf{H}$) in these applications, $\mathbf{K}(t_i)$, $\mathbf{P}(t_i^-)$, and $\mathbf{P}(t_i^+)$ attain steady-state values as the contribution of $\mathbf{P}(t_0)$ decays. The steady-state values of $\mathbf{K}$, $\mathbf{P}^-$, and, $\mathbf{P}^+$ are used in this application.

2.2 Model Formulation

Before applying a Kalman filter to simulation output, the system dynamics and measurement equations must be determined. For this formulation, $\mathbf{\Phi}$, $\mathbf{G}$, Q_d, $\mathbf{H}$ and R are time-invariant and are estimated from the simulation output. The following discussion follows a methodology developed by Gallagher (1992).

Both Kelton and Law (1983) and Schruben (1982) model the steady-state output as a constant mean plus noise

$$Y_i = \mu_Y + \eta_i \text{ for observations } i = 1, 2, \ldots, n \tag{12}$$

where η_i is noise with $E[\eta_i] = 0$. Let the sequence of Kalman filter measurements be the simulation output minus a mean estimate. Using this model for the simulation output, the dynamics and measurement models must account for the correlation in the noise.

Assume that the noise correlation can be approximated with an AR(2) process, shown in Equation (5) with $p = 2$. If the state vector is defined to be the last two noise terms,

$$\mathbf{x} = \begin{bmatrix} \eta_i \\ \eta_{i-1} \end{bmatrix}$$

then equivalent dynamics model, Equation (7), matrices are

$$\mathbf{x}(t_i) = \begin{bmatrix} \eta_i \\ \eta_{i-1} \end{bmatrix}, \ \mathbf{\Phi} = \begin{bmatrix} \phi_1 & \phi_2 \\ 1 & 0 \end{bmatrix}, \ \mathbf{G} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Since the assumed model for the simulation output is given by Equation (12), the measurements of the first filter-design state, $x_1(t_i) = \eta_i$, are modeled by

$$z(t_i) = Y_i - \mu_Y \text{ for } i = 1, 2, \ldots, n \tag{13}$$

Therefore, the measurement matrix in Equation (8) is $\mathbf{H} = \begin{bmatrix} 1 & 0 \end{bmatrix}$. However, pure autoregressive models do not include measurement noise $v(t_i)$, so either $v(t_i) = 0$, or equivalently, since the measurement noise already has a mean of zero, its variance could be zero, $R = 0$. While not a pure autoregressive process, nonzero measurement noise $v(t_i)$ may be incorporated into model (Harvey 1989). In this application, since the simulation output Y_i is known exactly, the measurement noise $v(t_i)$ may represent a "lack of fit" from assuming the noise correlation is an AR(2) process. Inclusion of measurement noise requires that the variance R of the measurement noise $v(t_i)$ must also be estimated.

In order to implement this formulation of the Kalman filter, the system parameters must be estimated. The AR(2) plus measurement noise model formulation has five unknown parameters to estimate, which are the steady-state mean μ_Y, the autoregressive coefficients ϕ_1 and ϕ_2, the variance Q_d of the dynamics driving noise $w_d(t_i)$, and the variance R of the measurement noise $v(t_i)$. Since the Kalman filter gain is determined by the ratio of Q_d and R, simultaneous estimation of these two noise variances leads

to an indeterminate system. Therefore, the first term of the Kalman filter gain is estimated instead of Q_d and R .

Mehra (1971) reviews numerous schemes for estimating the necessary Kalman filter parameters. The applied estimation routine was based on a hybrid of least squares estimation and Mehra's (1971) correlation technique. The use of least squares estimation is suggested by the fact that the "best" Kalman filter should provide a minimum value for the sum of the squared residuals, calculated with Equation (10). In contrast, the correlation technique equates sample correlations with the theoretical values for the assumed system model. This method provides an estimate of the Kalman filter gain $\mathbf{K}$, the output mean μ_Y, and the autoregressive coefficients, ϕ_1, and ϕ_2. The filter-computed residual variance $\mathbf{HP}^-\mathbf{H}^T + R$ is estimated as the mean squared residual. Using $\mathbf{K}$ and $\mathbf{HP}^-\mathbf{H}^T + R$, the noise variances, Q_d and R, can be estimated. A complete description of the estimation routine can be found in Howard (1992).

2.3 Multiple Model Adaptive Estimation

Multiple Model Adaptive Estimation (MMAE) as described by Maybeck (1982, 1986) can be applied to obtain variance information of the mean estimator. MMAE estimation discretizes the parameter space and runs a Kalman filter for each combination of discrete parameters. Based on each filter's residuals, the probability of that filter's parameter values being correct can be calculated. In this application, the MMAE technique is applied with various estimates for the mean.

For application of MMAE with a single unknown parameter, let "a" denote the parameter to be estimated. The continuous range of values for a will be discretized into L representative values. After the discretization of the parameter space is complete, let the probability that the parameter a assumes the value a_j conditioned on the measurement history prior to and including time t_i be $p_j(t_i) =$ Prob $(\mathsf{a} = a_j | Z(t_i) = Z_i)$, where Z_i is the measurement history up to time t_i. Assuming a is limited to the L discrete values a_j, the probability is calculated as (Maybeck 1982):

$$p_j(t_i) = \frac{f_{z(t_i)|\mathsf{a},Z(t_{i-1})}(z_i|a_j, Z_{i-1}) \cdot p_j(t_{i-1})}{\sum_{l=1}^{L} f_{z(t_i)|\mathsf{a},Z(t_{i-1})}(z_i|a_l, Z_{i-1}) \cdot p_l(t_{i-1})} \tag{14}$$

The probabilities are calculated using the Kalman filter residuals $r_i^j = z_i^j - \mathbf{H}_j \hat{x}_j(t_i)$ and the filter-computed residual variance $A_j = \mathbf{H}_j \mathbf{P}_j^- \mathbf{H}_j^T + R_j$. Since these residuals are assumed jointly normally distributed and the measurements, z_i, are univariate, we have

$$f_{z(t_i)|\mathsf{a},Z(t_{i-1})}(z_i|a_j, Z_{i-1}) = (2\pi)^{-1/2}|A_j|^{-1/2} \cdot \exp\left\{-\frac{1}{2}(r_i^j)^2(t_i)A_j^{-1}\right\} \tag{15}$$

The residuals r_i^j for the "best" a_j should be small relative to filter-computed residual variance A_j, so the "best" value of a_j will have a high probability assigned by the preceding $p_j(t_i)$ computation. Similarly, the residuals for a "mismatched" model should be large and the associated probability should be small (Maybeck 1986).

Before implementing this technique and constructing a confidence interval, several tactical issues must be decided. These issues are all interrelated and can have a significant impact on the MMAE estimate of a parameter's mean and variance. These issues are:

1. The spread between the filters with the smallest and largest mean estimate.
2. The number of discretized levels of the parameter, which is the number of filters.
3. The spacing between the filters.
4. The initial or *a priori* MMAE filter probabilities.

In most engineering applications, an extensive and detailed model, known as the "truth model", can be generated. With the "truth model", several mathematical techniques provide answers to the tactical issues (Lainiotis 1971, 1976, Maybeck and Hentz 1987, Sheldon and Maybeck 1990). However, discrete-event simulation is most often applied when no analytical solution is apparent. Therefore, no "truth model" is obvious, and heuristics were used to answer these questions (Howard 1992).

2.4 Confidence Interval Construction Steps

In this application, the unknown parameter a varied in the MMAE bank of filters is the mean. The simulation output is processed through L Kalman filters, each with a different assumed mean, $a_j = \mu_Y^j$. Each state vector is initiated as a vector of zeros. The realizations of the measurements for each filter are $z_i^j = Y_i - \mu_Y^j$ based on Equation (13). In this application, the filter-computed residual variance based on the estimated Kalman filter parameters was used in all of the filters. For each simulation observation, the state estimates are propagated with Equation (9) and updated with Equations (10) and (11).

After processing the simulation output through each of the MMAE filters, a confidence interval is

Table 1: Actual Coverage Rates

Data	NOBM			OBM			STDS			AUTO	MMAE PI
Size	5	10	20	5	10	20	5	10	20		
1280	0.815	0.784	0.730	0.801	0.781	0.735	0.747	0.672	0.531	0.792	0.757
2560	0.857	0.843	0.814	0.852	0.826	0.802	0.820	0.765	0.674	0.830	0.842
5120	0.881	0.868	0.844	0.878	0.869	0.850	0.843	0.827	0.762	0.845	0.898

Note: With nominal rate of 0.9, estimation accuracy is $\approx \pm 0.016$ for 1000 runs.

Table 2: Average Half Widths

Data	NOBM			OBM			STDS			AUTO	MMAE PI
Size	5	10	20	5	10	20	5	10	20		
1280	1.520	1.219	1.009	1.378	1.177	0.992	1.187	0.888	0.595	1.455	1.301
2560	1.201	1.018	0.884	1.096	0.971	0.866	1.056	0.823	0.621	0.990	1.061
5120	0.898	0.776	0.713	0.830	0.755	0.700	0.806	0.710	0.576	0.700	0.822

constructed. To construct the confidence interval we used the final filter probabilities from Equation (14). These filter probabilities serve as a discrete approximation of the underlying distribution. Therefore, the confidence interval end points are calculated such that the sum of the filter probabilities in each tail is $\alpha/2$. We used a linear interpolation between the discrete filter locations, and call this new technique MMAE PI (Probabilistic Interval).

3 MONTE CARLO RESULTS

The baseline confidence interval construction techniques of nonoverlapping batch means (NOBM), overlapping batch means (OBM), standardized time series (STDS), and autoregressive time series (AUTO) were compared, against the new Kalman filter technique (MMAE PI) in a Monte Carlo analysis.

Numerous Monte Carlo analyses were run using data generated from discrete-event simulations of $M/M/1$ queues. These analyses involved the use of different strategies for the tactical issues involved in implementing the MMAE techniques (e.g., number of filters, *a priori* filter probabilities, etc.). A complete review of the results is given by Howard (1992). This section provides results from a representative simulation of an $M/M/1$ queue.

This representative case looked at the average waiting time in the queue for an $M/M/1$ queue with a traffic intensity of 0.8 and the first 5000 observations truncated to diminish the initial transient bias. 1000 confidence intervals were constructed for three sample sizes (1280, 2560, and 5120) and three batch sizes. For NOBM and STDS, the number of batches was 5, 10, and 20. For OBM, the same batch sizes, $\frac{1}{5}$, $\frac{1}{10}$, and $\frac{1}{20}$ of the sample size, were used, but more batches resulted from overlapping. MMAE PI used 24 evenly-spaced filters centered on the grand mean $\bar{Y}(n)$ with a total spread of $\pm$ 1.6 times $\hat{\sigma}_{Y_i}$. The variance of the output Y_i for an AR(2) with measurement noise model is (Gallagher 1992):

$$\sigma^2_{Y_i} = \left(\frac{1-\phi_2}{1+\phi_2}\right)\frac{Q_d}{\{(1-\phi_2)^2-\phi_1^2\}} + R$$

The actual coverage rates for the techniques are found in Table 1. The average half widths and the associated standard deviations are found in Table 2 and Table 3, respectively.

Table 1 indicates that, for a large sample size, the MMAE PI technique provides the highest actual coverage rate of any of the techniques. At a sample size of 2560, the MMAE PI technique offers coverage rates similar to those provided by the better baseline techniques (e.g., NOBM and OBM). At the smallest sample size, 1260, the coverage rates provided by MMAE PI were significantly below those provided by NOBM and OBM with large batches.

Table 2 indicates that the Kalman filter technique offers small average half widths for the levels of coverage it provides. In particular, with a sample size of 5120, MMAE PI had an average half width of 0.822. This half width is smaller than the average half width (0.898) provided by the baseline technique with the highest coverage (NOBM with 5 batches). Table 3 indicates that, for the levels of coverage provided by the Kalman filter technique, not only are the average half widths smaller, as seen in Table 2, but their standard

Table 3: Standard Deviation of Half Widths

Data	NOBM			OBM			STDS			AUTO	MMAE PI
Size	5	10	20	5	10	20	5	10	20		
1280	1.165	0.755	0.508	0.959	0.701	0.492	0.730	0.374	0.167	2.246	0.877
2560	0.807	0.592	0.429	0.648	0.526	0.405	0.677	0.398	0.193	0.804	0.607
5120	0.525	0.391	0.314	0.435	0.357	0.297	0.471	0.342	0.206	0.365	0.333

deviations are also smaller. Under some conditions, the MMAE PI was best by all three criteria.

Another advantage of the Kalman filter technique is that it does not require the selection of a batch size. Table 3 illustrates that the choice of a batch size is generally a tradeoff between obtaining high coverage with high variability or low coverage with low variability. The Kalman filter technique does not require the analyst to make this choice.

MMAE PI appears to provide the best results for large (5120) sample sizes. However, 5120 observations is not really large for a method based on one long simulation run. MMAE PI's performance degrades as sample size decreases. This may be due to assuming the state estimates can be modeled as an AR(2) process observed with additive measurement noise. Another possibility is that $M/M/1$ output is extremely noisy and the MMAE filters might simply require a large amount of data to stabilize.

4 SUMMARY

In summary, a new technique, MMAE PI, has been proposed for constructing confidence intervals using information provided by the Kalman filter. The results obtained by the MMAE PI method were excellent. However, these results may not be the best results obtainable with this method. As discussed in the previous section, the number of filters used, the spacing between the filters, and the total spread of the filters are three critical parameters affecting MMAE performance. The results demonstrate that the technique based on the Kalman filter offers a novel and efficient way to construct meaningful confidence intervals for steady-state parameters of discrete-event simulations.

ACKNOWLEDGMENTS

This research was partially sponsored by Air Force Office of Scientific Research grant AFOSR-PO-91-0003.

REFERENCES

Fishman, G.S.. 1971. Estimating Sample Size in Computing Simulation Experiments. *Management Science* **18**: 21-38.

Fishman, G.S. 1978. *Principles of Discrete Event Simulation.* New York: John Wiley & Sons.

Gallagher, M.A. 1992. Identification of the Initial Transient in Discrete-Event Simulation Output Using the Kalman Filter. PhD Dissertation, Department of Operational Sciences, Air Force Institute of Technology, Wright-Patterson AFB, Ohio.

Harvey, A.C. 1989. *Forecasting, Structural Time Series Models and the Kalman Filter.* New York: Cambridge University Press.

Howard, R.B. 1992 Confidence Interval Estimation for Output of Discrete-Event Simulations Using the Kalman Filter. Master's Thesis, Department of Operational Sciences, Air Force Institute of Technology, Wright-Patterson AFB, Ohio.

Kelton, W.D. and A.M. Law. 1983. A New Approach for Dealing with the Startup Problem in Discrete Event Simulation. *Naval Research Logistics Quarterly* **30**: 641–658.

Demetrios G. Lainiotis, D.G. 1971. Optimal Adaptive Estimation: Structure and Parameter Adaptation. *IEEE Transactions on Automatic Control* **AC-16**: 160-170.

Demetrios G. Lainiotis, D.G. 1976. Partitioning: A Unifying Framework for Adaptive Systems, I: Estimation. *Proceedings of the IEEE* **64**: 1126-1143.

Law, A.M. 1983. Statistical Analysis of Simulation Output Data. *Operations Research* **31**: 983-1029.

Maybeck, P.S. 1979. *Stochastic Models, Estimation, and Control: Volume 1.* New York: Academic Press, Inc.

Maybeck, P.S. 1982. *Stochastic Models, Estimation, and Control: Volume 2 & 3.* New York: Academic Press, Inc.

Maybeck, P.S. 1986. Adaptive Tracking of Dynamic Airborne Vehicles Based on (FLIR) Image Plane Intensity Data. *Proceedings of the Third Bad Honnef Conference on Stochastic Differential Systems*: 284-305.

Maybeck, P.S. and K.P. Hentz. 1987. Investigation of Moving-Bank Multiple Model Adaptive Algorithms. *AIAA Journal of Guidance, Control, and Dynamics* **10**: 90-96.

Mehra, R.K. 1971. On-line Identification of Linear Dynamic Systems with Applications to Kalman Filtering. *IEEE Transactions on Automatic Control* **AC-16**: 10-21.

Meinhold R.J. and N.D. Singpurwalla. 1983. Understanding the Kalman Filter. *Journal of the American Statistician* **37**: 123-127.

Meketon, M.S. and B. Schmeiser. 1984. Overlapping Batch Means: Something for Nothing? *Proceedings of the 1984 Winter Simulation Conference*: 227-230.

Schruben, L.W. 1982. Detecting Initialization Bias in Simulation Output. *Operations Research* **30**: 569-590.

Sheldon, S.N. and P.S. Maybeck. 1990. An Optimizing Design Strategy for Multiple Model Adaptive Estimation and Control. *Proceedings of the 29th IEEE Conference on Decision and Control*: 3522-3527.

Schruben, L.W. 1983. Confidence Interval Estimation Using Standardized Time Series. *Operations Research* **31**: 1090-1108.

AUTHOR BIOGRAPHIES

RANDALL B. HOWARD, Captain, USAF, is an air staff policy analyst in the Analysis Division, Directorate of Personnel Plans, Headquarters, United States Air Force. He received a B.S. in Mathematics and Chemistry from Birmingham Southern College in 1986 and an M.S. from the Air Force Institute of Technology (AFIT) in 1992. His mailing address is: AF/DPYO, Room 5C360, The Pentagon, Washington, D.C. 20330

MARK A. GALLAGHER, Captain, USAF, is a doctoral student at the Air Force Institute of Technology (AFIT). He received a B.S. in Operations Research and Computer Science from the US Air Force Academy in 1983 and an M.S. from the Air Force Institute of Technology (AFIT) in 1986. His mailing address is: AFIT/ENS, WPAFB, OH 45433-6583.

KENNETH W. BAUER, Jr., Major, USAF, is an assistant professor in the Department of Operational Sciences, School of Engineering, at the Air Force Institute of Technology (AFIT). He received a B.S. in Mathematics from Miami University (Ohio) in 1976, a M.E.A. from the University of Utah in 1980, an M.S. from AFIT in 1981, and a Ph.D. from Purdue University in 1987. His research interests are in the statistical aspects of simulation. His mailing address is: AFIT/ENS, WPAFB, OH 45433-6583.

PETER S. MAYBECK, is a full professor in the Department of Electrical Engineering, School of Engineering, at the Air Force Institute of Technology. He received his degrees, B.S. in 1968 and Ph.D. in 1972, in Aeronautics and Astronautics from the Massachuset Institute of Technology. His research interests include stochastic estimation and control and applications of Kalman filters. His mailing address is: AFIT/ENG, WPAFB, OH 45433-6583.