ESTIMATION OF EXPECTATIONS IN TWO-LEVEL NESTED SIMULATION EXPERIMENTS

David Fernando Muñoz

Departamento de Ingeniería Industrial y Operaciones, Instituto Tecnológico Autónomo de México, Río Hondo #1, 01080 Ciudad de México, México

davidm@itam.mx

ABSTRACT
Two-level nested simulation methods have been recently applied for the analysis of simulation experiments under parameter uncertainty. On the outer level of the nested run, we generate (n) observations of the parameters, while on the inner level; we fix the parameter on its corresponding value and generate (m) observations using a simulation model. In this paper, we focus on the output analysis of two-level stochastic simulation experiments for the case where the observations of the inner level are independent, showing how the variance of the simulated observations can be decomposed in the sum of parametric and stochastic components. Furthermore, we derive central limit theorems that allow us to compute asymptotic confidence intervals to assess the accuracy of the simulation-based estimators for the point forecast and the variance components. Theoretical results are validated through experiments using a forecasting model for sporadic demand, where we have obtained analytical expressions for the point forecast and the variance components.

Keywords: Bayesian forecasting, stochastic simulation, parameter uncertainty, two-level simulation

1. INTRODUCTION

Simulation is widely recognized as an effective technique for producing forecasts, evaluating risk, animating and illustrating the performance of a system over time (see, e.g., Kelton et al. 2012). When there is uncertainty in some components of a simulation model, these random components are modeled using probability distributions and/or stochastic processes that are generated during the simulation run, in order to produce a stochastic simulation.

The method that is usually applied to estimate a performance measure θ (e.g., an expectation) in transient simulation, consists in computing an estimator θ_n from the output of n independent replications of the simulation experiment. This estimator must be consistent, i.e., it must satisfy $\theta_n \Rightarrow \theta$, as $n \to \infty$ (where "$\Rightarrow$" denotes weak convergence of random variables). Consistency guarantees that the estimator approaches the parameter as the number of replications n increases, and the accuracy of the simulation-based estimator θ_n is typically assessed by an asymptotic confidence interval (ACI) for the parameter. The expression for an ACI is usually obtained through a Central Limit Theorem (CLT) for the estimator θ_n (see, for example, chapter III of Asmussen and Glynn 2007).

For i = 1 to n:
  For j = 1 to m:
    a. Generate (independently) a value θ_j by sampling from $p(\theta|x)$.
    b. Run a simulation experiment with $\Theta = \theta_j$ to obtain an independent replication W_{ij} of $W = g(Y(s), 0 \le s \le T; \Theta)$
  End Loop
End Loop
Compute $\hat{\alpha}_i = \frac{1}{m}\sum_{i=1}^{m} W_{ij}, i = 1,\ldots,m, \hat{\alpha} = \frac{1}{n}\sum_{i=1}^{n}\hat{\alpha}_i.$

Figure 1: Two level algorithm for calculating a point estimator under parameter uncertainty

In contrast to the estimation of performance measures, input parameters of a simulation experiment are usually estimated from real-data observations (x) and, while the majority of applications covered in the relevant literature assume that no uncertainty exists in the value of these parameters, the uncertainty can actually be significant when little data is available. In these cases, Bayesian statistics can be used to incorporate this uncertainty in the output analysis of simulation experiments via the use of a posterior distribution $p(\theta|x)$. A Bayesian approach using simulation as a forecasting tool has been reported in diverse areas; for example, healthcare (see, e.g., Santos et al. 2013) or software development (see, e.g., Lee et al. 2009). A methodology currently proposed for the analysis of simulation experiments under parameter uncertainty and, in particular, for the estimation of expected values, is a two-level nested simulation method (see, e.g., Zouaoui and Wilson 2003, L'Ecuyer 2009; Andradóttir and Glynn 2016). In the outer level, we simulate (n) observations for the parameters from a posterior distribution $p(\theta|x)$, while in the inner level we simulate (m) observations for the response variable with the parameter fixed at the value generated in the outer level (see Figure 1). In this paper, we focus on the output analysis of two-level simulation experiments, for the particular case when the observations of the inner level are independent, showing how the variance of a simulated observation can be decomposed into parametric and stochastic variance components. Afterwards, we derive a CLT for both the estimator of the point forecast and the estimators of the variance components. Our CLT´s allows us to compute an ACI for each estimator. Our results are validated through experiments with a forecasting model for sporadic demand reported in Muñoz and Muñoz (2011).

Following this introduction, we present the proposed methodology for the construction of an ACI for the point forecast and the variance components in a two level simulation experiment. Afterwards, we present an illustrative an example that has an analytical solution for the parameters of interest in this paper. This example is used in the following section to illustrate the application and validity of our proposed methodologies for the construction of an ACI. Finally, in the last section, we present conclusions and directions for future research.

2. METHODOLOGY

In order to identify the variance components in each observation W_{ij} of the algorithm illustrated in Figure 1, let $\mu(\Theta) = E[W_{11}|\Theta]$, and $\sigma^2(\Theta) = E[W_{11}^2|\Theta] - \mu(\Theta)^2$. Under this notation, the point forecast is $\alpha = E[\mu(\Theta)]$, and the variance of W_{ij} is

$$
\begin{aligned}
V[W_{ij}] &\overset{def}{=} E[W_{ij}^2] - E[W_{ij}]^2 \\
&= E\left[E[W_{ij}^2|\Theta] - \mu(\Theta)^2\right] + E\left[\mu(\Theta)^2\right] - E[\mu(\Theta)]^2 \qquad (1)\\
&= \sigma_1^2 + \sigma_2^2,
\end{aligned}
$$

for $i = 1,\ldots,n$; $j = 1,\ldots,m$, where $\sigma_1^2 = E[\sigma^2(\Theta)]$, and $\sigma_2^2 = V[\mu(\Theta)] \overset{def}{=} E\left[\mu(\Theta)^2\right] - E[\mu(\Theta)]^2$. It is worth mentioning that, in relevant literature, σ_1^2 is commonly referred to as stochastic variance and σ_2^2 as parametric variance.

2.1. Point Estimators

In this paper, we are interested in both the estimation of the point forecast $\alpha = E[\mu(\Theta)]$ and the estimation of the variance components of the observations generated by the algorithm of Figure 1 and defined in (1), thus we first define the natural point estimators

$$
\hat{\alpha}(n) = \frac{1}{n}\sum_{i=1}^{n}\hat{\alpha}_i, \qquad (2)
$$

$$
\hat{\sigma}_T^2(n) = \frac{1}{n-1}\sum_{i=1}^{n}(\hat{\alpha}_i - \hat{\alpha}(n))^2, \hat{\sigma}_2^2(n) = \frac{1}{n}\sum_{i=1}^{n}S_i^2, \qquad (3)
$$

where $\hat{\alpha}_i = m^{-1}\left(\sum_{j=1}^{m} W_{ij}\right)/m$, and $S_i^2 = \left[\sum_{j=1}^{m}(W_{ij} - \hat{\alpha}_i)^2\right]/(m-1)$, for i = 1,...,m. Note that the $\hat{\alpha}_i$ are independent and identically distributed (i.i.d.) with expectation $E[\hat{\alpha}_1] = \alpha$ and variance

$$
\begin{aligned}
V[\hat{\alpha}_1] &= E\left[(\hat{\alpha}_1 - \alpha)^2\right] \\
&= m^{-2}\left(mE\left[(W_{11} - \alpha)^2\right] + m(m-1)E[(W_{11} - \alpha)(W_{12} - \alpha)]\right) \qquad (4)\\
&= m^{-1}\left(\sigma_1^2 + \sigma_2^2\right) + m^{-1}(m-1)\sigma_2^2 = \sigma_2^2 + m^{-1}\sigma_1^2
\end{aligned}
$$

On the other hand, the S_i^2 are i.i.d. with expectation $E[S_1^2] = s_2^2$. Thus, the next proposition follows from the Law of Large Numbers (Theorem 5.4.2 of Chung 2001).

Proposition 1. Given $m \geq 1$, if $E[|W_{11}|] < \infty$ then $\hat{\alpha}(n)$ and $\hat{\sigma}_T^2(n)$ are unbiased and consistent estimators for α and $\sigma_2^2 + m^{-1}\sigma_1^2$ (as $n \to \infty$), respectively. Furthermore, if $m \geq 2$, then $\hat{\sigma}_2^2(n)$ is an unbiased and consistent estimator for σ_2^2 (as $n \to \infty$).

2.2. Accuracy of the Point Estimators

As we established in the previous Section, the point estimators proposed in (2) and (3) are consistent, and thus converge to the corresponding parameters values as $n \to \infty$. Nonetheless, to establish the level of accuracy of these estimators, we must establish a CLT for each estimator that allows us to calculate the corresponding ACI. Note that both $\hat{\alpha}(n)$ and $\hat{\sigma}_2^2(n)$ are averages of i.i.d observations, thus the next proposition follows from the classic CLT for i.i.d. observations.

Proposition 2. Given $m \geq 1$, if $E[W_{11}^2] < \infty$ then

$$
\frac{\sqrt{n}(\hat{\alpha}(n) - \alpha)}{\sqrt{V(\hat{\alpha}_1)}} \Rightarrow N(0,1),
$$

as $n \to \infty$. Furthermore, if $m \geq 2$ and $E[W_{11}^4] < \infty$, then

$$
\frac{\sqrt{n}\left(\hat{\sigma}_2^2(n) - \sigma_2^2\right)}{\sqrt{V\left(S_1^2\right)}} \Rightarrow N(0,1),
$$

where $N(0,1)$ denotes the standard normal distribution, $\hat{\alpha}(n), \hat{\sigma}_2^2(n), \hat{\alpha}_1, S_1^2$ are defined in (2) and (3), $V[\hat{\alpha}_1]$ is defined in (4), and $V[S_1^2] = E\left[\left(S_1^2 - \sigma_2^2\right)^2\right]$.

Since we have consistent estimators for $V[\hat{\alpha}_1]$ and $V[S_1^2]$, the next corollary follows from Proposition 1 and Slutsky's Theorem (see, e.g., Serfling 2009).

Corollary 1. Under the same notation and assumptions from Proposition 2, we have

$$
\frac{\sqrt{n}(\hat{\alpha}(n) - \alpha)}{\sqrt{\hat{V}(\hat{\alpha}_1)}} \Rightarrow N(0,1), \text{ and } \frac{\sqrt{n}\left(\hat{\sigma}_2^2(n) - \sigma_2^2\right)}{\sqrt{\hat{V}\left(S_1^2\right)}} \Rightarrow N(0,1),
$$

for $m \geq 1$ and $m \geq 2$, respectively, as $n \to \infty$, where

$$\hat{V}[\hat{\alpha}_1] = \frac{1}{n-1} \sum_{i=1}^{n} (\hat{\alpha}_i - \hat{\alpha}(n))^2 ,$$

$$\hat{V}\left[S_1^2\right] = \frac{1}{n-1} \sum_{i=1}^{n} \left(S_i^2 - \bar{S}^2\right)^2, \bar{S}^2 = \frac{1}{n} \sum_{i=1}^{n} S_i^2 .$$

To obtain a CLT for $\hat{\sigma}_T^2(n)$, note that this estimator is the variance of a set of i.i.d. observations, thus we can use the following Lemma. We omit the proof of this Lemma, nonetheless, it can be proven by applying the Delta Method (see, e.g., Proposition 2 of Muñoz and Glynn 1997).

Lemma 1. If $X_1, X_2, \ldots$ is a sequence of i.i.d. random variables with $E\left[X_1^4\right] < \infty$, then

$$\frac{n^{1/2}\left(S^2(n) - \sigma^2\right)}{\sigma_S} \Rightarrow N(0,1),$$

as $n \to \infty$, where $\sigma^2 = \mu_2 - \mu_1^2$,

$$\sigma_S = \sqrt{8\mu_1^2\mu_2 - 4\mu_1^4 - 4\mu_1\mu_3 + \mu_4 - \mu_2^2} ,$$

$\mu_k = E\left[X_1^k\right]$, k = 1, 2, 3, 4;

$S^2(n) = (n-1)^{-1} \sum_{i=1}^{n} (X_i - \hat{\mu}_1)^2$, $\hat{\mu}_1 = n^{-1} \sum_{i=1}^{n} X_i$.

Corollary 2. Under the same assumptions as in Lemma 1 we have

$$\frac{n^{1/2}\left(S^2(n) - \sigma^2\right)}{\hat{\sigma}_S} \Rightarrow N(0,1),$$

as $n \to \infty$, where

$$\hat{\sigma}_S = \sqrt{8\hat{\mu}_1^2\hat{\mu}_2 - 4\hat{\mu}_1^4 - 4\hat{\mu}_1\hat{\mu}_3 + \hat{\mu}_4 - \hat{\mu}_2^2} ,$$

$$\hat{\mu}_k = n^{-1} \sum_{i=1}^{n} X_i^k .$$

Corollary 2 follows from the fact that $\hat{\mu}_k$ is an unbiased and consistent estimator of μ_k, and the next corollary follows from the fact that $\hat{\sigma}_T^2(n)$ is the sample variance of the $\hat{\alpha}_i$.

Corollary 3. Given $m \geq 1$, if $E\left[W_{11}^4\right] < \infty$ then

$$\frac{\sqrt{n}\left(\hat{\sigma}_T^2(n) - \left(\sigma_2^2 + m^{-1}\sigma_1^2\right)\right)}{\sqrt{\hat{V}_T}} \Rightarrow N(0,1),$$

as $n \to \infty$, where

$$\hat{V}_T = 8\bar{\alpha}_1^2\bar{\alpha}_2 - 4\bar{\alpha}_1^4 - 4\bar{\alpha}_1\bar{\alpha}_3 + \bar{\alpha}_4 - \bar{\alpha}_2^2 ,$$

$$\bar{\alpha}_k = n^{-1} \sum_{i=1}^{n} \hat{\alpha}_i^k .$$

Let $0 < \beta < 1$, and using corollaries 1 and 3 we can establish a 100β% ACI for the point forecast α, and variance components σ_2^2 and $\sigma_T^2 = \sigma_2^2 + m^{-1}\sigma_1^2$; each ACI is centered in the corresponding point estimator, $\hat{\alpha}(n), \hat{\sigma}_2^2(n)$ or $\hat{\sigma}_T^2(n)$, and the halfwidths are given by

$$H_n(\alpha) = z_\beta \frac{\sqrt{\hat{V}(\hat{\alpha}_1)}}{\sqrt{n}} , \tag{5}$$

$$H_n\left(\sigma_2^2\right) = z_\beta \frac{\sqrt{\hat{V}\left(S_1^2\right)}}{\sqrt{n}} , \tag{6}$$

$$H_n\left(\sigma_2^2\right) = z_\beta \frac{\sqrt{\hat{V}\left(S_1^2\right)}}{\sqrt{n}} , \tag{7}$$

for α, σ_2^2 and σ_T^2, respectively, where z_β is the $(1-\beta/2)-$quantile of a $N(0,1)$ distribution, $\hat{V}(\hat{\alpha}_1)$ and $\hat{V}\left(S_1^2\right)$ are defined in Corollary 1, and $\hat{V}_T$ is defined in Corollary 3.

Note that the ACI proposed in (5), (6) and (7) assume that the value of *m* from the algorithm of Figure 1 is fixed and the accuracy of the estimator improves as *n* (the number of observations in the outer level) increases (in turn, the halfwidth of the ACI gets smaller). Given that we can build a valid ACI for any value of *m*, a relevant question is how to find an adequate value of *m* to get an acceptable level of accuracy in a reasonable amount of running time. To answer this question for the case of the point estimator of α, let us fix the total number of iterations in the algorithm of Figure 1 to *k* = *nm*, and note from (3) that

$$n^{-1}V[\hat{\alpha}_1] = k^{-1}\left(m\sigma_2^2 + \sigma_1^2\right) \tag{8}$$

takes its minimal value for *m* = 1, suggesting that the point estimator $\hat{\alpha}(n)$ defined in (2) is more accurate as *m* approaches the value of 1. Note that for *m* = 1, a fixed number of iterations *k* = *nm* is convenient (from the point of view of running time), when the computation of W_{ij} requires the same or more computation time as Θ_i, as suggested in the relevant literature (see, for example, Andradóttir and Glynn 2016). Furthermore, if we allow *m* to increase with *n*, we can obtain the following proposition. We omit the proof, but mention that it follows from Lindeberg-Feller Theorem (Theorem 7.2.1 of Chung 2001).

Proposition 3. Given $0 < p \leq 1$, if $m = \lfloor n^{-1+1/p} \rfloor$ and $E\left[W_{11}^2\right] < \infty$ then

$$\frac{\sqrt{n}(\hat{\alpha}(n) - \alpha)}{\sqrt{V[\hat{\alpha}_1]}} \Rightarrow N(0,1),$$

as $n \to \infty$, where $V[\hat{\alpha}_1]$ is defined in (3).

If, once again, we set the total number of iterations in the algorithm of Figure 1 to *k* = *nm*, we let $n \approx k^p$, $m \approx k^{1-p}$ and *nm* = *k*, it follows from Proposition 3 that the asymptotic variance of $\hat{\alpha}(n)$ is $n^{-1}V[\hat{\alpha}_1]$ for every $0 < p \leq 1$. Note that this metric reaches its minimum value when $p = 1$, that is, when *n* = *k* and *m* = 1. However, note that we need $m \geq 2$ to estimate σ_2^2.

In the following section, we report some experiments that confirm our theoretical results.

3. AN EXAMPLE WITH ANALYTICAL SOLUTION

The following model (reported in Muñoz and Muñoz 2011) has been proposed to forecast sporadic demand. We can use this model to find analytical expressions for the parameters considered in this paper. This model is used in the following section to illustrate the validity of the ACI's proposed in the previous section.

Client arrivals for a particular item in a shop follows a Poisson process, yet there is uncertainty in the arrival rate , so that given , interarrival times between clients are i.i.d. with exponential density:

$$f(y|\theta_0) = \begin{cases} \theta_0 e^{-\theta_0 y}, & y > 0, \\ 0, & \text{otherwise}, \end{cases} \tag{9}$$

where $\theta_0 \in S_{00} = (0, \infty)$. Every client can order *j* units of this item with probability Θ_{1j}, $j = 1, \ldots, q$, $q \geq 2$. Let $\Theta_1 = (\Theta_{11}, \ldots, \Theta_{1(q-1)})$ and $\Theta_{1q} = 1 - \sum_{j=1}^{q-1} \Theta_{1j}$, then $\Theta = (\Theta_0, \Theta_1)$ is the parameter vector, and $S_0 = S_{00} \otimes S_{01}$ is the parameter space, where total demand during a period with length *T* is

$$D = \begin{cases} \sum_{i=1}^{N(T)} U_i, & N(T) > 0, \\ 0, & \text{otherwise}, \end{cases} \tag{10}$$

where $N(s)$ is the number of client arrivals during the interval $[0, s]$, $s \geq 0$, and $U_1, U_2 \ldots$ are the individual demands (conditionally independent relative to Θ). The information about Θ consists of i.i.d. observations $v = (v_1, \ldots, v_r)$, $u = (u_1, \ldots, u_r)$ of past clients, where v_i is the interarrival time between client *i* and client (*i* - 1), and u_i is the number of units ordered by client *i*. By using Jeffrey's non-informative prior as the prior density for Θ, we obtain the posterior density (see Muñoz and Muñoz 2011 for details) $p(\theta|x) = p(\theta_0|v) p(\theta_1|u)$, where $x_i = (v_i, u_i)$, $i = 1, \ldots, r$, $x = (x_1, \ldots, x_r)$, $\theta = (\theta_0, \theta_1)$,

$$p(\theta_0|v) = \frac{\theta_0^{r-1} \left(\sum_{i=1}^{r} v_i \right)^r e^{-\theta_0 \sum_{i=1}^{r} v_i}}{(r-1)!}, \tag{11}$$

$$p(\theta_1|u) = \frac{\left(1 - \sum_{j=1}^{q-1} \theta_{1j} \right)^{c_q - 1/2} \prod_{j=1}^{q-1} \theta_{1j}^{c_j - 1/2}}{B(c_1 + 1/2, \ldots, c_q + 1/2)}, \tag{12}$$

where $c_j = \sum_{i=1}^{r} I[u_i = j]$, and

$$B(a_1, \ldots, a_q) = \prod_{j=1}^{q} \Gamma(a_j) / \Gamma\left(\sum_{j=1}^{q} a_j \right),$$

for $a_1, \ldots, a_q > 0$. Using this notation, we can prove that

$$\alpha = E[T\Theta_0] \sum_{j=1}^{q} j p_j, \tag{13}$$

$$\sigma_1^2 = \frac{E[T^2 \Theta_0^2]}{(q_0 + 1)} \sum_{j=1}^{q} j^2 p_j + \frac{E[T\Theta_0]^2 [(q_0 / n) - 1]}{(q_0 + 1)} \left(\sum_{j=1}^{q} j p_j \right)^2, \tag{14}$$

$$\sigma_2^2 = E[T\Theta_0] \sum_{j=1}^{q} j^2 p_j, \tag{15}$$

where $p_j = q_j / q_0, q_j = c_j + 1/2, j = 1, \ldots, q,\ q_0 = \sum_{j=1}^{q} q_j$

$$E[T\Theta_0] = Tr \left(\sum_{i=1}^{r} v_i \right)^{-1},\ E[T^2 \Theta_0^2] = T^2 r (1 + r) \left(\sum_{i=1}^{r} v_i \right)^{-2}.$$

4. EXPERIMENTAL RESULTS

To validate the ACI proposed in (5), (6) and (7), we conducted some experiments with the example from the previous section to illustrate the estimation of α, σ_2^2 and σ_T^2. We considered the values $T = 15$, $r = 20$, $\sum_{i=1}^{r} x_i = 10$, $q = 5$, $c_1 = 5$, $c_2 = 3$, $c_3 = 2$, $c_4 = 3$, $c_5 = 7$. With this data, the point forecast is $\alpha \approx 95.333$, and the variance components are $\sigma_1^2 \approx 380.667$, $\sigma_2^2 \approx 568.598$.

Table 1: Performance of the 90% ACI based on 1000 replications with *m* = 2

	n = 100, *m* = 2			*n* = 1000, m = 2			*n* = 10000, *m* = 2		
	α	σ_2^2	σ_T^2	α	σ_2^2	σ_T^2	α	σ_2^2	σ_T^2
Coverage	0.896	0.893	0.868	0.892	0.902	0.892	0.898	0.910	0.897
Mean HW	4.493	91.327	183.384	1.423	29.616	60.703	0.450	9.419	19.386
St Dev HW	0.352	18.481	45.769	0.036	2.092	5.606	0.004	0.215	0.613
R.M.S.E.	2.716	53.794	117.903	0.875	17.901	38.067	0.278	5.673	11.995
Bias	-0.015	-1.547	1.918	0.011	-0.326	-0.226	0.008	0.116	-0.048

In tables 1 and 2 we report the results of 1000 independent replications of the algorithm of Figure 1 for every 90% ACI, with values *nm* = 200, 2000, 20000, and *m* = 2, 5. Finally, with the objective of comparing the choice of *m* = 1 (which we consider optimal for the estimation of the point forecast), we also report the results of similar experiments with *nm* = 100, 10000, *m* = 1 and $m \approx (nm)^{1/3}$ (suggested by Andradóttir and Glynn, 2016, as an adequate choice for *m* in the case of biased estimators in the inner level of the algorithm of Figure 1). In Table 1, we present the coverage (fraction of ACI's that cover the parameter value), average, and standard deviation of the corresponding halfwidth defined in (5), (6) and (7), square root of the mean square error (R.M.S.E.) and empirical bias based on 1000 replications for each ACI. As observed in the table, the coverage are acceptable (very close to the nominal value of 0.9, even for m = 100). This validates the ACI defined in (5), (6) and (7). Furthermore, every performance measure of the ACI (average, standard deviation of the halfwidth, R.M.S.E., and bias) improves as the number of replications n increases.

Table 2: Performance of the 90% ACI based on 1000 replications with m = 5

	$n = 40, m = 5$			$n = 400, m = 5$			$n = 4000, m = 5$		
	α	σ_2^2	σ_T^2	α	σ_2^2	σ_T^2	α	σ_2^2	σ_T^2
Coverage	0.894	0.862	0.818	0.898	0.892	0.891	0.890	0.891	0.899
Halfwidth Mean	6.553	75.712	235.617	2.077	24.695	81.722	0.655	7.829	25.879
Halfwidth Std Dev	0.830	15.714	95.501	0.082	1.706	11.997	0.008	0.178	1.191
R.M.S.E.	3.978	48.081	164.462	1.248	15.262	50.276	0.405	4.911	15.557
Bias	0.134	-2.111	10.234	-0.028	-0.262	3.827	-0.024	-0.176	-0.049

In Table 2, we present the results with m = 5. We can observe that, while the coverage is close to the nominal value of 0.9, all of the performance measures for the ACI (average, standard deviation of the halfwidth, R.M.S.E., and bias) are worse (larger) tan the ones reported in Table 1, for the estimation of α and σ_T^2, and better for the estimation of σ_2^2, suggesting that, for the same number of observations nm, a lower value of m is better for estimating α and a greater value of m is better for estimating σ_2^2.

Table 3: Performance of the 90% ACI based on 1000 replications with m = 1 and $m \approx (nm)^{1/3}$

	$nm = 100, m = 1$		$nm = 100, m = 5$		$nm = 10000, m = 1$		$nm = 10000, m = 20$	
	α	σ_T^2	α	σ_T^2	α	σ_T^2	α	σ_T^2
Coverage	0.913	0.869	0.890	0.781	0.889	0.899	0.887	0.872
Halfwidth Mean	5.029	232.305	9.203	306.107	0.504	24.351	1.765	65.799
Halfwidth Std Dev	0.392	61.479	1.655	157.539	0.004	0.752	0.065	8.873
R.M.S.E.	2.933	147.120	5.507	234.267	0.308	14.570	1.115	42.196
Bias	0.020	1.367	0.004	11.555	-0.015	-0.194	0.034	-0.890

Finally, in Table 3 we show the results for the estimation of α and σ_T^2 for the cases m = 1 and $m \approx (nm)^{1/3}$. We again find that all performance measure for the ACI (average, standard deviation of the halfwidth, R.M.S.E., and bias) are worse (larger) for $m \approx (nm)^{1/3}$, confirming our finding that, for the same number of replications nm, m = 1 produces better point estimators for α than $m \approx (nm)^{1/3}$.

5. METHODOLOGY

In this paper, we propose methodologies to calculate point estimators (and their corresponding halfwidths), for both the point forecast and the variance components in two-level nested stochastic simulation experiments. This method can be applied to the construction of Bayesian forecasts using simulation models.

Both theoretical and experimental results confirm that the proposed point estimators and their corresponding halfwidths are asymptotically valid, i.e., the point estimators converge to the corresponding parameter values and the halfwidths converge to the nominal coverage as the number of replications (n) of the outer level increases.

Furthermore, given a fixed number of total observations (nm), we show that the choice of only one replication in the inner level (m = 1) provides more accurate estimators for both the point forecast (α), and the variance of the point forecast ($\sigma_1^2 + \sigma_2^2$). However, $m \geq 2$ is required for the estimation of σ_2^2.

Directions for future research on this topic includes experimentation with other point estimators, such as, quasi Monte Carlo or Simpson integration, with the objective of finding more accurate point estimators for the parameters considered in this paper.

ACKNOWLEDGMENTS

This research is supported by the Asociación Mexicana de Cultura AC.

REFERENCES

Andradóttir S. and Glynn P. W., 2016. Computing Bayesian means using simulation. ACM Transactions on Modeling and Computer Simulation, 26(2), paper 10.

Asmussen S. and Glynn, P. W., 2007. Stochastic Simulation Algorithms and Analysis. New York: Springer.

Chung K. L., 2001. A Course in Probability Theory. New York: Academic Press.

L'Ecuyer P., 2009. Quasi-Monte Carlo methods with applications in finance. Finance and Stochastics, 13(3), 307-349.

Kelton W. D., Smith J. S. and Sturrock D. T., 2013. Simio and Simulation: Modeling, Analysis, Aplications. 3rd ed. Sewickley: Simio LLC.

Lee S., Celik N. And Son Y. J., 2009. An integrated simulation modelling framework for decision aids in enterprise software development process. International Journal of Simulation and Process Modelling, 5(1), 62-76.

Muñoz D. F. and Glynn P. W. 1997. A batch means methodology for estimation of a nonlinear function of a steady-state mean. Management Science, 43(8), 1121-1135.

Muñoz D. F. and Muñoz D. F., 2011. Bayesian forecasting of spare parts using simulation. In: N. Altay and L. A. Litteral, eds. Service Parts Management: Demand Forecasting and Inventory Control. New York: Springer, 105-123.

Santos Jr E., Joo Kim K., Yu F., Li D. and Rosen J., 2013. Bayesian knowledge modelling for healthcare practices. International Journal of Simulation and Process Modelling, 8(1), 52-65.

Serfling R. J., 2009. Approximation Theorems of Mathematical Statistics. New York: John Wiley & Sons.

Zouaoui F. and J. R. Wilson, 2003. Accounting for parameter uncertainty in simulation input modeling. IIE Transactions, 35(9), 781-792.

AUTHOR BIOGRAPHY

DAVID F. MUÑOZ is Professor and Head of the Department of Industrial and Operations Engineering at the Instituto Tecnológico Autónomo de México. He holds a BS in Statistics from the National Agrarian University of Peru, a M.S. in Mathematics from Catholic University of Peru, and M.S. and Ph.D. degrees in Operations Research from Stanford University. He has been recognized as an “Edelman Laureate” for the Institute for Operations Research and the Management Sciences, for participating in the project "Indeval Develops a New Operating and Settlement System Using Operations Research", winner of the 2010 Edelman Prize. His research interests include input and output analysis for stochastic simulations and the applications of Operations Research.