

## A SUBSET SELECTION PROCEDURE UNDER INPUT PARAMETER UNCERTAINTY

Canan G. Corlu

Metropolitan College  
Boston University  
Boston, MA 02215, USA

Bahar Biller

Tepper School of Business  
Carnegie Mellon University  
Pittsburgh, PA 15213, USA

### ABSTRACT

This paper considers a stochastic system simulation with unknown input distribution parameters and assumes the availability of a limited amount of historical data for parameter estimation. We investigate how to account for parameter uncertainty — the uncertainty that is due to the estimation of the input distribution parameters from historical data of finite length — in a subset selection procedure that identifies the stochastic system designs whose sample means are within a user-specified distance of the best mean performance measure. We show that even when the number of simulation replications is large enough for the stochastic uncertainty to be negligible, the amount of parameter uncertainty in output data imposes a threshold on the user-specified distance for an effective use of the subset selection procedure for simulation. We demonstrate the significance of this effect of parameter uncertainty for a multi-item inventory system simulation in the presence of short demand histories.

### 1 INTRODUCTION

Two important components of a stochastic simulation are input modeling and output analysis. Input modeling is the selection of appropriate probability distributions that characterize the stochastic behavior of the system inputs, while output analysis is the study of the simulation output data to estimate the distributional properties of the system performance measures. The simulation generally starts with the estimation of the unknown parameters of the input probability distributions from the historical input data of finite length. Then, the finite-sample estimates of the parameters are used to drive the simulation as if they were the true values. Finally, the output data obtained from this simulation are analyzed for predicting, for instance, the mean performance measure and constructing the confidence interval.

However, when the input distributions and their parameters are unknown and the available historical input data are limited, there are three main sources of uncertainty to represent in the output analysis: stochastic uncertainty (i.e., the uncertainty that is due to the dependence of the simulation output on random input processes) (Helton 1997), model uncertainty (i.e., the uncertainty that is due to the selection of a single input model from a set of alternative models), and parameter uncertainty (i.e., the uncertainty that is due to the estimation of the input-model parameters from limited data) (Raftery, Madigan, and Volinsky 1996). Stochastic uncertainty is inherent in every simulation and controlled by the number of simulation replications. The input model and parameter uncertainties are, on the other hand, often ignored as a result of driving the simulations with the probability distributions estimated from input data of finite length. Unfortunately, this practice often leads to inconsistent estimates for the performance measures as well as inconsistent coverage of the confidence intervals. Assuming known functional forms for all input models in this paper, we account for the stochastic and parameter uncertainties in the analysis of the simulation output data. However, we do this with the objective of selecting a subset of system designs whose sample means are within a user-specified distance of the best mean performance measure.

Cheng and Holland (1997) are the first to show the dependence of the simulation output on stochastic and parameter uncertainties using the delta and parametric bootstrap methods. These solution approaches are followed by the two-point method in Cheng and Holland (1998) as well as an alternative approach that particularly considers the simulation bias in Cheng and Holland (2004). Barton and Schruben (1993, 2001) and Barton (2007) characterize the parameter uncertainty in the simulation output by using the non-parametric bootstrap method, while Barton, Nelson, and Xie (2010) illustrate the metamodel-assisted bootstrapping in dealing with parameter uncertainty in the simulation of queuing systems. Following a parametric approach, Chick (2001), Zouaoui and Wilson (2003), and Ng and Chick (2006) develop computationally effective Bayesian models to represent the parameter uncertainty. Similarly, we use a Bayesian model to account for the parameter uncertainty. Differently, we use this model in a subset selection procedure when the parameters of the input distributions are unknown and there is only a limited amount of historical input data for parameter estimation.

A subset selection procedure is a screening device that aims to select a subset of alternative system designs including the best system with at least a pre-specified probability. Gupta (1965) proposed a single-stage procedure for this problem that is applicable when the output data from the simulations of the system designs are independent, balanced, and normally distributed with common (unknown) variance. Nelson, Swann, Goldsman, and Song (2001) developed a single-stage subset selection procedure that permits unequal and unknown variances. Wilson (2001) enhanced the lower bound on the probability of correct selection with a decomposition lemma for the screening-and-selection procedure of Nelson, Swann, Goldsman, and Song (2001). The screen-to-the-best procedure of Boesel, Nelson, and Kim (2003) extended this procedure further to allow unequal sample sizes. In the absence of parameter uncertainty, our procedures reduce to the screen-to-the-best procedure of Boesel, Nelson, and Kim (2003) with an additional error tolerance parameter.

In a single-stage selection procedure, the number of systems included in the subset is random. In other words, the subset can include only a single system corresponding to the one with the best mean performance measure, while it is also possible that the subset includes all the system designs; i.e., no system design with inferior performance is screened out. To address this issue, Gupta and Santner (1973) devised a restricted subset selection procedure that allowed the user to specify an upper bound on the subset size. This was followed by Santner (1976) that proposed a restricted subset selection rule with the goal of including at least one of a certain number of best system designs in a subset with a pre-specified maximum size. However, both of these subset selection procedures are only applicable to simulation outputs that are assumed to have a known common variance. The Gupta-Santner restricted subset selection rule was extended to unknown variances by Sullivan and Wilson (1989). We refer the reader to Hsu and Panchapakesan (2005) and Chen (2008, 2010) for more recent studies of restricted subset selection procedures in simulation.

Our goal is to understand the role of parameter uncertainty on the performance of the (unrestricted) subset selection procedure, especially when the number of simulation replications is large enough to deliver a single system as the best in the absence of parameter uncertainty.

We organize the remainder of the paper as follows. In Section 2, we present the well-known subset selection procedure under the assumption of known input parameters, and discuss the challenge in the implementation of this procedure when only a limited amount of historical input data are available. We present a Bayesian model to capture parameter uncertainty in Section 3 and introduce the subset selection procedure accounting for parameter uncertainty in Section 4. We apply this procedure to a multi-item inventory system simulation in Section 5, and conclude with a summary of findings and a discussion of several extensions in Section 6.

## **2 SUBSET SELECTION FOR STOCHASTIC SIMULATION**

Section 2.1 presents the subset selection procedure for stochastic simulations with known input parameters. Section 2.2 discusses the challenge arising from unknown input parameters with a newsvendor example.

**2.1 Subset Selection Formulation: Known Input Parameters**

We consider  $k$  alternative stochastic system designs with independent and normally distributed simulation output data. More specifically, we let  $X_{i\ell}$  represent the  $\ell$ th output from system  $i$  and  $\mathbf{X}_i = \{X_{i\ell}; \ell = 1, 2, \dots, n_i\}$  denote the output sequence of  $n_i$  replications from this system. Thus, we assume that  $X_{i\ell}, \ell = 1, 2, \dots, n_i$  are independent and identically distributed normal random variables with mean  $\mu_i := \mathbb{E}[X_{i\ell}]$  and variance  $\sigma_i^2 := \mathbb{V}[X_{i\ell}]$  for  $i = 1, 2, \dots, k$ ; i.e.,  $X_{i\ell}$  can be considered as the average of a large number of output variables or a batch mean in a steady-state simulation. We further assume that the processes  $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_k$  are mutually independent and are compared based on their true means  $\mu_i, i = 1, 2, \dots, k$ . Finally, we let the best system to be the one with the largest true mean and denote the ordered means by  $\mu_{[k]} \geq \mu_{[k-1]} \geq \dots \geq \mu_{[1]}$ .

The goal is to obtain a (random) subset  $I \subseteq \{1, 2, \dots, k\}$  of  $k$  alternatives such that the probability of correct selection,  $\mathbb{P}\{[k] \in I | \mu_{[k]} - \mu_{[k-1]} \geq \delta\}$  is at least  $1 - \alpha$ , where  $0 < \alpha < 1 - 1/k$  and  $\delta > 0$  is the error tolerance; i.e., a user-specified, practically significant difference between the mean performance measures of the best and second best system designs. Under the assumption of unknown variances  $\sigma_i^2, i = 1, 2, \dots, k$  and  $n_i$  simulation replications performed for the  $i$ th system design, the decision rule achieving  $1 - \alpha$  probability of correct selection takes the following form for  $i = 1, 2, \dots, k$  and all values of  $j \in \{1, 2, \dots, k\}$  that are different from  $i$ :

$$\bar{X}_i \geq \bar{X}_j - \left[ \left( t_{(1-\alpha)^{1/(k-1)}, n_i-1}^2 \frac{S_i^2}{n_i} + t_{(1-\alpha)^{1/(k-1)}, n_j-1}^2 \frac{S_j^2}{n_j} \right)^{1/2} - \delta \right]^+, \tag{1}$$

where  $\bar{X}_i := \sum_{\ell=1}^{n_i} X_{i\ell} / n_i$  is the sample mean of the simulation output data collected for system  $i$ ,  $S_i^2 := \sum_{\ell=1}^{n_i} (X_{i\ell} - \bar{X}_i)^2 / (n_i - 1)$  is the unbiased sample variance estimator of  $\sigma_i^2$ , and  $t_{(1-\alpha)^{1/(k-1)}, n_i-1}$  is the  $(1 - \alpha)^{1/(k-1)}$  quantile of the Student's t distribution with  $n_i - 1$  degrees of freedom. This is the generalization of Gupta's (1965) subset selection procedure for known variances to unknown variances and an error tolerance of  $\delta > 0$ , as a result of which the resulting subset contains those alternatives whose sample means are within  $\delta$  of the best with probability of at least  $1 - \alpha$ . When the simulation replications  $n_i, i = 1, 2, \dots, k$  are large enough for the stochastic uncertainty (e.g.,  $S_i^2/n_i$  for the  $i$ th system) to be negligible, the identification of the best system design reduces to the selection of the system with the largest mean.

**2.2 Challenge: Unknown Input Parameters**

The assumption of known input parameters rarely holds in practice and the input parameters are often estimated from very limited amounts of historical data. This is mainly because of the fact that the underlying input generating processes do not remain constant indefinitely, and, even if a long history of observations is available, it is common to consider only the most recent observations.

The basic premise of this paper is that the uncertainty around the input parameters estimated from historical data of finite length might be too large to identify the best system design, even with negligible stochastic uncertainty in the simulation output data. More specifically, if there is no parameter uncertainty, then increasing the number of replications diminishes the simulation output variance and reduces the comparison of different system designs to the comparison of their sample mean performance measures. However, if the input distribution parameters are unknown and estimated from a historical input data set of finite length, then the increasing number of simulation replications has no effect on the simulation output variance due to parameter uncertainty.

We illustrate this situation with a simple inventory example to demonstrate the role of demand parameter uncertainty on the estimation of the Type-1 service level; i.e., the probability that all customer orders arriving are completely met from the stock on hand. We consider the simulation of the following repeated newsvendor setting: (i) Historical demand data  $d_t, t = 1, 2, \dots, m$  of length  $m$  are available and no forecasting process is in place. (ii)  $F(\cdot; \Upsilon)$  is the demand's cumulative distribution function (cdf) with unknown (but stationary) parameter vector  $\Upsilon$ . We assume a normal demand distribution with (unknown) mean  $\mu_D$  and (unknown)

standard deviation  $\sigma_D$ ; thus,  $\Upsilon = (\mu_D, \sigma_D)'$ . (iii) A Bayesian approach is used to obtain an estimate  $\tilde{\Upsilon}$  of  $\Upsilon$  from the available historical demand data; see Biller, Corlu, and Tayur (2013a) for an example application of this approach for inventory control. (iv) The inventory target  $Q$  is one of the inputs to the simulation. For the given inventory target  $Q$  and historical demand data  $d_t, t = 1, 2, \dots, m$ , we use simulation to estimate the Type-1 service level and its 95% confidence interval.

Obviously, the inventory system of interest is not complex enough to warrant the use of simulation for service-level estimation. However, our objective is to understand the role of demand parameter uncertainty in the output analysis of a simple inventory simulation. Therefore, we represent the mean simulation output response function as  $F(Q; \Upsilon)$  in terms of the demand parameter vector  $\Upsilon$ , and build on the distributional properties of this response function together with maximum a posteriori estimates of the demand parameters in order to obtain an approximation for the amount of demand parameter uncertainty in the simulation output variance (Biller, Corlu, and Tayur 2013b):

**Corollary 1** In the presence of a historical demand data set of length  $m$ , the simulation output variance due to demand parameter uncertainty,  $\mathbb{V}_\Upsilon[F(Q; \Upsilon)]$  is given by

$$\phi^2 \left( \frac{(Q - \gamma_m)(v_m + 3)^{1/2}}{\zeta_m} \right) \left( \frac{1}{\kappa_m} + \frac{(Q - \gamma_m)^2}{2\zeta_m^2} \right),$$

where  $\phi(\cdot)$  is the probability density function of the standard normal random variable and  $\gamma_m := (\kappa_0 \gamma_0 + \sum_{t=1}^m d_t) / (\kappa_0 + m)$ ,  $\kappa_m := \kappa_0 + m$ ,  $v_m := v_0 + m$ , and

$$\zeta_m^2 := \zeta_0^2 + \frac{\kappa_0 m}{\kappa_0 + m} \left( \frac{1}{m} \sum_{t=1}^m d_t - \mu_0 \right)^2 + \sum_{t=1}^m \left( d_t - \frac{1}{m} \sum_{t=1}^m d_t \right)^2$$

are the posterior demand parameters defined in terms of the historical demand data  $d_t, t = 1, 2, \dots, m$  and the hyper-parameters  $\gamma_0, \kappa_0, v_0$ , and  $\zeta_0^2$  of the prior density function.

We refer the reader to Biller, Corlu, and Tayur (2013a) for inferring the values of the hyper-parameters  $\gamma_0, \kappa_0, v_0$ , and  $\zeta_0^2$  from expert opinion. In the remainder of the section, we use the above characterization to understand the effect of demand parameter uncertainty on the Type-1 service-level estimation via stochastic simulation.

When the historical data set contains 10 observations with a critical fractile of 90% (i.e.,  $m = 10$  and  $\varphi = 0.90$ ), we find that accounting for demand parameter uncertainty in the simulation leads to a length of 0.21 for the 95% confidence interval when the number of simulation replications is large enough for the stochastic uncertainty to be negligible. In other words, the 95% confidence interval for the estimated Type-1 service level, say 0.78, is obtained as  $[0.68, 0.89]$ ; i.e.,  $[0.78 - 0.21/2, 0.78 + 0.21/2]$ . Thus, the impact of the demand parameter uncertainty on the confidence interval of the mean Type-1 service level can be quite significant for the limited amounts of historical demand data to the extent that the resulting confidence interval is too wide for the simulation to guide the inventory manager towards setting any particular inventory target.

While a way to reduce the confidence-interval length of 0.21 is the collection of more demand data, a natural question to ask is how fast the additional demand observations reduce the length of the confidence interval. We achieve a reduction of 14.43% in the confidence-interval length with 5 additional data points and a reduction of 34.73% with 20 additional data points. Fifty additional data points, on the other hand, achieve a reduction of 46.60% in the confidence-interval length. That is, we need to increase the number of observations in the demand data set by at least five times to approximately halve the confidence-interval length. Thus, the effect of the parameter uncertainty in the simulation output variance disappears very slowly with the increasing number of observations in the historical input data set, even in a simple setting like the single-item repeated newsvendor setting. Our goal is to equip the subset selection procedure with

the ability to compare alternative system designs (e.g., the multi-item inventory system simulations that do not necessarily satisfy the assumptions of the repeated newsvendor) in the presence of input parameter uncertainty.

### 3 QUANTIFICATION OF THE PARAMETER UNCERTAINTY IN THE OUTPUT DATA

An excellent review of the methods of capturing parameter uncertainty in stochastic simulations is available in Barton (2012). In this section, we only review one of these methods, the asymptotic normality approximation (Cheng and Holland 1997, Ng and Chick 2006), for the decomposition of the simulation output variance into terms related to stochastic uncertainty and parameter uncertainty. We refer the reader to Section 6 for a brief discussion of extending the presentation to an alternative method of capturing parameter uncertainty. In this paper, however, we use the decomposition resulting from the asymptotic normality approximation to formulate the subset selection procedure under parameter uncertainty.

We start with denoting the output random variable  $X_{i\ell}$  of the  $\ell$ th simulation replication of system  $i$  with

$$X_{i\ell} = g_i(\Upsilon_i) + \sigma_i Z_{i\ell},$$

where  $g_i(\Upsilon_i)$  is the mean simulation response function with the unknown input parameter vector  $\Upsilon_i$ ,  $\sigma_i^2$  is the simulation stochastic variance, and  $Z_{i\ell}$ ,  $\ell = 1, 2, \dots, n_i$  are independent and identically distributed random variables, each with a mean of zero and a standard deviation of one. We assume the availability of historical input data of length  $m_i$ , which we denote by  $d_{i,t}$ ,  $t = 1, 2, \dots, m_i$  for system  $i$ . The key to the asymptotic normality approximation is the convergence of the density function of the parameter vector  $\Upsilon_i$  to a multivariate normal density function with a mean of  $\tilde{\Upsilon}_i$  (i.e., the maximum a posteriori estimate of the unknown parameter vector  $\Upsilon_i$ ) and a variance-covariance matrix of  $\Sigma_i$  with the increasing length  $m_i$ ; i.e., as  $m_i \rightarrow \infty$ ,  $\Sigma_i^{-1/2}(\Upsilon_i - \tilde{\Upsilon}_i) \rightarrow \mathcal{N}_{\omega_i}(\mathbf{0}_{\omega_i}, \mathbf{I}_{\omega_i})$ , where  $\mathbf{0}_{\omega_i}$  denotes the  $\omega_i \times 1$  vector of zeros and  $\mathbf{I}_{\omega_i}$  denotes the  $\omega_i \times \omega_i$  identity matrix for  $i = 1, 2, \dots, k$ . However, this asymptotic normality approximation only works under certain regularity conditions (Bernardo and Smith 2000). The satisfaction of these conditions is verified together with the continuous differentiability of  $g_i(\Upsilon_i)$ , the equivalence of  $\lambda_{i,\max}$  and  $O(\lambda_{i,\min})$  with  $\lambda_{i,\min}$  and  $\lambda_{i,\max}$  the smallest and largest eigenvalues of  $\Sigma_i$ , and the convergence of  $\tilde{\Upsilon}_i$  to  $\Upsilon_i$  in probability. Letting  $\Delta g_i(\tilde{\Upsilon}_i)$  represent the  $1 \times \omega_i$  vector of  $g_i$  evaluated at  $\tilde{\Upsilon}_i$  and utilizing the multivariate delta method (Hogg, McKean, and Craig 2005), we see that the response function  $g_i(\Upsilon_i)$  is asymptotically normal with a mean of  $g_i(\tilde{\Upsilon}_i)$  and a variance of

$$\tau_i^2 = \Delta g_i(\tilde{\Upsilon}_i) \Sigma_i [\Delta g_i(\tilde{\Upsilon}_i)]^T,$$

which quantifies the amount of input parameter uncertainty in the simulation output data.

We are now ready to decompose the variance of the  $i$ th system design into  $\sigma_i^2$ , the simulation stochastic variance and  $\tau_i^2$ , the simulation output variance due to parameter uncertainty:

$$\mathbb{V}[X_{i\ell}] = \mathbb{V}_{\Upsilon_i} \left[ \mathbb{E} \left[ g_i(\Upsilon_i) + \sigma_i Z_{i\ell} \mid \Upsilon_i \right] \right] + \mathbb{E}_{\Upsilon_i} \left[ \mathbb{V} \left[ g_i(\Upsilon_i) + \sigma_i Z_{i\ell} \mid \Upsilon_i \right] \right] = \mathbb{V}_{\Upsilon_i} [g_i(\Upsilon_i)] + \mathbb{E}_{\Upsilon_i} [\sigma_i^2] = \tau_i^2 + \sigma_i^2.$$

Consequently, for  $i = 1, 2, \dots, k$ , we obtain the mean performance measure,  $\mathbb{E}[\bar{X}_i]$  as  $g_i(\tilde{\Upsilon}_i)$  and the variance of the mean performance measure,  $\mathbb{V}[\bar{X}_i]$  as  $\tau_i^2 + \sigma_i^2/n_i$ . Thus, for the  $i$ th system design, the amount of parameter uncertainty in the variance of the mean performance measure,  $\tau_i^2$  can be reduced by the collection of additional input data (i.e., by increasing the value of  $m_i$ ), but the increase in the number of replications,  $n_i$  has no impact on this portion of the simulation output variance due to parameter uncertainty.

### 4 SUBSET SELECTION UNDER PARAMETER UNCERTAINTY

Building on the response-surface representation of the simulation output data in Section 3, we redesign the subset selection procedure of Section 2 under parameter uncertainty. As we do this, it is important to

distinguish between different sources of uncertainty (i.e., across all the  $k$  system designs, the input random variables and thus the input parameter vectors  $\Upsilon_i$ ,  $i = 1, 2, \dots, k$  are different and independent of each other) and identical sources of uncertainty (i.e.,  $\Upsilon_1 = \Upsilon_2 = \dots = \Upsilon_k$  with a single distribution function for the identical input random variables driving the simulation of each system design). We provide a subset selection procedure for stochastic simulations with different sources of uncertainty in this section, and discuss the case of identical sources of uncertainty in Section 6.

More specifically, we represent the response function of the simulation output data by  $X_{i\ell} = g_i(\Upsilon_i) + \sigma_i Z_{i\ell}$ ,  $\ell = 1, 2, \dots, n_i$ ,  $i = 1, 2, \dots, k$  with independent random vectors  $\Upsilon_i$ ,  $i = 1, 2, \dots, k$  and independent zero-mean normal random variables  $Z_{i\ell}$ ,  $\ell = 1, 2, \dots, n_i$ ,  $i = 1, 2, \dots, k$ . Proposition 1 provides the decision rule to identify a subset of the system designs within  $\delta$  of the best system design with probability of at least  $1 - \alpha$  in the presence of  $n_i$ ,  $i = 1, 2, \dots, k$  simulation replications and  $m_i$ ,  $i = 1, 2, \dots, k$  observations in the historical data set. The proof of this proposition can be found in Biller and Corlu (2013):

**Proposition 1** The decision rule ensuring the probability of correct selection to be at least  $1 - \alpha$  is

$$\bar{X}_i \geq \bar{X}_j - \left[ t_{(1-\alpha)^{1/(k-1)}, n_i-1}^2 \left( \frac{\chi^2(n_i)}{n_i} \tau_i^2 + \frac{S_i^2}{n_i} \right) + t_{(1-\alpha)^{1/(k-1)}, n_j-1}^2 \left( \frac{\chi^2(n_j)}{n_j} \tau_j^2 + \frac{S_j^2}{n_j} \right) \right]^{1/2} - \delta \Big]^+$$

for  $i = 1, 2, \dots, k$  and all values of  $j \in \{1, 2, \dots, k\}$  that are different from  $i$ , where  $\chi^2(n_i)$  is the chi-squared random variable with  $n_i$  degrees of freedom.

If there were no input parameter uncertainty in the system simulations (i.e.,  $\tau_i^2 = 0$  for  $i = 1, 2, \dots, k$ ), then the decision rule in Proposition 1 would reduce to the subset selection rule (1) in Section 2 and identify the system designs that are within  $\delta$  of the best system design with probability of at least  $1 - \alpha$ . However, in the case of a limited amount of demand data with  $\tau_i^2 \geq 0$ ,  $i = 1, 2, \dots, k$ , it is important for the subset selection procedure to account for the parameter uncertainty to deliver at least  $1 - \alpha$  probability of correct selection.

For a historical data length of  $m_i$ , the contribution of  $\tau_i^2$  to the decision rule in Proposition 1 is proportional to the ratio of  $\chi^2(n_i)$  to  $n_i$ . A natural question to ask is which value is to be chosen for the ratio  $\chi^2(n_i)/n_i$  as we implement our decision rule for subset selection. Since  $\chi^2(n_i)/n_i \rightarrow 1$  as  $n_i \rightarrow \infty$ , a reasonable choice might be one. We investigated the sensitivity of the subset selection to different quantiles of the random variable  $\chi^2(n_i)$  for the stochastic simulation of the multi-item inventory system in Section 5. We found that the selection of one for  $\chi^2(n_i)/n_i$ ,  $i = 1, 2, \dots, k$  works very well for this particular application. While we expect this result to extend to the simulations of other stochastic systems, further investigation of this issue would require focus on the specific applications.

Next, we consider the case of increasing the number of simulation replications as a result of which the simulation output variance due to stochastic uncertainty diminishes to zero:

**Corollary 2** As  $n_i \rightarrow \infty$  for  $i = 1, 2, \dots, k$ , the identification of the best system design reduces to the comparison of the sample output means only when  $\delta \geq z_{(1-\alpha)^{1/(k-1)}} \sqrt{\tau_i^2 + \tau_j^2}$  for  $i = 1, 2, \dots, k$  and all values of  $j \in \{1, 2, \dots, k\}$  that are different from  $i$ .

For large enough number of simulation replications in the absence of parameter uncertainty, the identification of the best system design reduces to the comparison of the mean performance measures. The take-away from Corollary 2 is that the value of the parameter  $\delta$  must be at least the maximum of  $z_{(1-\alpha)^{1/(k-1)}} (\tau_i^2 + \tau_j^2)^{1/2}$ ,  $i = 1, 2, \dots, k$ ,  $\forall j \neq i$  to apply the subset selection procedure in a similar manner under input parameter uncertainty. That is, the amount of parameter uncertainty in the simulation output imposes a threshold on  $\delta$  when the goal is to identify a system design based on the comparison of the sample means for sufficiently large number of simulation replications. Obviously, the lower bound on  $\delta$  increases with the amount of parameter uncertainty in the simulation output data. While we do not

address this issue here, the lower bound on  $\delta$  can also be used to identify the stochastic system designs for which more data are to be collected for the effective use of the subset selection procedure under parameter uncertainty. In Section 5, we quantify this lower bound on  $\delta$  for a multi-item inventory system, and discuss the effect of parameter uncertainty on this quantification when the goal is to identify the inventory system with the maximum joint demand fulfillment probability in the presence of limited amounts of historical demand data.

## 5 AN APPLICATION TO A MULTI-ITEM INVENTORY SYSTEM SIMULATION

This section illustrates the procedure of the previous section with an application to a multi-item inventory system simulation. In particular, we consider five different distribution centers, each of which holds two items, and all of these items and thus their probability distributions, are different from each other. Table 1 tabulates the characteristics of each inventory system in terms of the means and standard deviations of the item demands, the item base-stock levels, and the maximum joint demand fulfillment probability achievable under the assumption of known demand parameters. Our goal is to identify the distribution center that delivers the maximum joint demand fulfillment probability under demand parameter uncertainty. For the case of known demand parameters, the decision rule (1) with  $n \geq 100$ ,  $\alpha = 0.05$ , and  $\delta = 0$  selects the first system design as the only element of the resulting subset.

First, we utilize the functional form of the joint demand fulfillment probability and apply the decision rule in Proposition 1 to investigate the impact of the demand parameter uncertainty on the subset selection. Table 2 provides the probability of including the (optimal) first system design in the subset, first for the case in which parameter uncertainty is ignored (see the Case I column) and then for the case in which we account for the parameter uncertainty (see the Case II column).  $m$  represents the (equal) number of historical demand observations available for each item and  $n$  represents the number of simulation replications that is also equal across all system designs. We conduct each experiment for 1000 macro-replications and set the value of the ratio  $\chi^2(n)/n$  to one in the decision rule of Proposition 1.

Our results demonstrate the significant effect of the parameter uncertainty on the probability of including the optimal system design in the subset in the presence of short demand history. When the number of simulation replications is large enough for the simulation output variance due to stochastic uncertainty to be negligible (e.g., when  $n = 5000$ ), we identify the probability of including the optimal system design in the subset only as 0.519 for a historical data length of 10. This probability increases to 0.851 when the number of historical demand observations is 100 and 1.000 when the number of available data points is 1000. Accounting for parameter uncertainty, on the other hand, results in higher probability of including the optimal system in the subset; i.e., we observe this probability to always exceed 0.99 in Table 2. Nevertheless, the impact of representing the parameter uncertainty diminishes in the presence of more than 100 input data points. For the cases in which  $m = 500$  and  $m = 1000$ , probability of including the optimal system design in the subset is almost one regardless of accounting for the parameter uncertainty. To put it another way, while the importance of considering parameter uncertainty decreases with the increasing length of the demand history, the simulation user must be aware of the improved performance of the subset

Table 1: Characteristics of the inventory system designs with different sources of demand uncertainty

System Index	First Item		Second Item		Base-Stock Levels		Fulfillment Probability
	Mean	Std. Dev.	Mean	Std. Dev.	First Item	Second Item	
1	100.00	25.00	120.00	15.60	131	140	0.80
2	80.00	20.00	110.00	14.30	93	127	0.66
3	130.00	32.50	70.00	9.10	140	85	0.59
4	50.00	12.50	150.00	19.50	62	170	0.71
5	155.00	38.75	85.00	11.10	195	100	0.78

Table 2: The probability of including the optimal system design in the subset of selection: (i) Case I ignores the demand parameter uncertainty; (ii) Case II accounts for the demand parameter uncertainty.

$m$	Case I			Case II		
	$n = 100$	$n = 1000$	$n = 5000$	$n = 100$	$n = 1000$	$n = 5000$
10	0.778	0.568	<b>0.519</b>	0.994	0.990	<b>0.993</b>
30	0.937	0.770	0.690	0.993	0.991	0.993
50	0.952	0.857	0.751	0.993	0.997	0.994
100	0.984	0.911	<b>0.851</b>	0.997	0.999	<b>1.000</b>
500	0.997	0.994	0.981	0.995	0.999	1.000
1000	0.998	1.000	<b>1.000</b>	0.995	0.997	<b>1.000</b>

selection procedure as a result of accounting for parameter uncertainty when there is a limited amount of historical data.

Next, we consider the case of large enough number of simulation replications for which the output variance due to stochastic uncertainty is negligible. In the absence of parameter uncertainty, the decision rule (1) reduces to the comparison of the sample means of the performance measures. This coincides with the selection of the inventory system having the maximum joint demand fulfillment probability in the experimental setting of this section. As shown by Corollary 2, parameter uncertainty imposes a lower bound on the value of  $\delta$ , which would allow us to identify a (best) system design by simply comparing the simulation mean performance measures of the alternatives. This bound on the parameter  $\delta$  is tabulated in Table 3 as a function of the number of historical demand observations available for each item. We observe that when the length of the demand history is short, the bound on  $\delta$  can be high; e.g., when we have only 10 data points, we identify the bound on  $\delta$  as 0.338. This means that identifying the first system design as the only element of the subset requires the specification of the error tolerance to be at least 0.338. Considering that this parameter takes values only from the unit interval, which follows from the definition of a joint demand fulfillment probability, the amount of parameter uncertainty in the simulation output data hinders the ability of the subset selection procedure to screen out the inferior system designs. Currently, we are investigating how to use this bound to guide the simulation user towards data collection for selective system designs to perform effective subset selection. Only when the number of data points exceeds 100, the bound on  $\delta$  falls below 0.1.

## 6 CONCLUSION

A common purpose for the use of simulation in stochastic system design and analysis is to evaluate alternative decisions. Despite the examples demonstrating the importance of capturing the uncertainty around the input parameter estimates, accounting for parameter uncertainty in a stochastic simulation is still not the standard practice. Consequently, there is a lack of statistically-valid procedures that compare alternative simulated decisions by accounting for parameter uncertainty. However, with the growing use of stochastic simulation in business system design and analysis, it is imperative to develop data-driven modeling support for statistical comparison of alternative decisions in the absence of full knowledge about the systems' input distribution parameters.

Our paper addresses this limitation of the simulation design and analysis by restricting the focus on subset selection when the parameters of the input distributions are unknown and there is only a limited amount of historical data available for parameter estimation. Redesigning the subset selection procedure

Table 3: The threshold on  $\delta$  as a function of  $m$ , the number of observations in the historical demand data.

threshold on $\delta$	length of the historical data set					
	10	30	50	100	500	1000
	<b>0.338</b>	0.207	0.163	0.117	<b>0.053</b>	<b>0.037</b>

to account for parameter uncertainty, we provide insights into the effect of parameter uncertainty on the identification of the stochastic system designs whose sample means are within an error tolerance of the best. We demonstrate that the amount of parameter uncertainty in the simulation output data imposes a threshold on the error tolerance even if the number of simulation replications is sufficiently large for a negligible amount of stochastic uncertainty. In other words, the error tolerance is to take the value of this threshold to effectively identify a subset including the best system design with a probability of correct selection that is at least of some user-specified value. This threshold can potentially guide the simulation practitioner towards data collection in selective system designs for effective subset selection under parameter uncertainty. We illustrate our subset selection procedure with an application to a multi-item inventory simulation in the presence of limited amounts of historical demand data.

We conclude with a discussion of extensions to our procedure to broaden its use under input uncertainty:

**Alternatives to Asymptotic Normality Approximation** The motivation behind our focus on the asymptotic normality approximation in this paper is the functional form it provides for the subset selection procedure under parameter uncertainty, which leads to an analytical approximation demonstrating the significant role that parameter uncertainty plays on the identification of the best decision even when the number of simulation replications approaches infinity. Asymptotic normality approximation is also suitable for the analysis of inventory system simulations with focus on service-level measurement. However, the asymptotic normality approximation works only with restrictions on the simulation inputs as well as with a smooth, well-defined function for the simulation output response. As one of the alternative methods to the asymptotic normality approximation, Bayesian simulation replication algorithm, on the other hand, does not suffer from any of these limitations of the asymptotic normality approximation (Zouaoui and Wilson 2003, 2004). It further provides increasing flexibility in the quantification of the model uncertainty in the simulation output data. Thus, a follow-up is to build on more flexible methods of capturing parameter uncertainty to devise subset selection procedures independent of the restrictive conditions under which the asymptotic normality approximation works.

**Identical Sources of Uncertainty** Despite our focus on simulation system designs with different sources of uncertainty in this paper, it is also possible that the system designs are faced with identical sources of uncertainty. More specifically, we consider the case of  $Y_i = Y$  for  $i = 1, 2, \dots, k$  and thus, the representation of the response function of the simulation output data by  $X_{i\ell} = g_i(Y) + \sigma_i Z_{i\ell}$ ,  $\ell = 1, 2, \dots, n_i$ ,  $i = 1, 2, \dots, k$ . What is important to recognize is that  $Y$  is common to all system designs creates dependence among their output response functions, and this dependence should be factored into the development of a subset selection procedure under input parameter uncertainty. Preliminary analysis shows that the effect of such dependence in the simulation output data might decrease the impact of parameter uncertainty on the lower bound to  $\delta$ . A natural question to ask is whether such a reduction would be significant enough to eliminate the effect of parameter uncertainty from subset selection.

**Common Random Numbers** When the input random variables are the same for all system designs, another question to ask is whether the use of Common Random Numbers (CRN) would improve the performance of the decision rule to identify the subset of system designs within  $\delta$  of the best system design with probability at least  $1 - \alpha$ . A good resource for insights on using CRN in selection procedures is Chen (2012). It is shown that it is generally better to use CRN to increase the probability of correct selection for subset selection when common random numbers are properly synchronized. In our particular case, the use of CRN to drive the simulation leads to the representation of the response function of the simulation output data by  $X_{i\ell} = g_i(Y) + \sigma Z_{\ell}$ ,  $\ell = 1, 2, \dots, n$ ,  $i = 1, 2, \dots, k$  with an identical number of simulation replications across all system designs; i.e.,  $n_i = n$ ,  $i = 1, 2, \dots, k$ . The derivation of decision rules to capture this situation is the subject of ongoing work.

## REFERENCES

- Barton, R. R. 2007. "Presenting a more complete characterization of uncertainty: Can it be done?". In *Proceedings of the INFORMS Simulation Society Research Workshop*. Fontainebleau, France.
- Barton, R. R. 2012. "Tutorial: Input uncertainty in output analysis". In *Proceedings of the 2012 Winter Simulation Conference*, edited by C. Laroque, J. Himmelspach, R. Pasupathy, O. Rose, and A. M. Uhrmacher, 162 – 172. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.
- Barton, R. R. and L. W. Schruben. 1993. "Uniform and bootstrap resampling of input distributions". In *Proceedings of the 1993 Winter Simulation Conference*, edited by G. W. Evans, M. Mollaghasemi, W. E. Biles, and E. C. Russell, 503 – 508. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.
- Barton, R. R. and L. W. Schruben. 2001. "Resampling methods for input modeling". In *Proceedings of the 2001 Winter Simulation Conference*, edited by B. A. Peters, J. S. Smith, D. J. Medeiros, and M. W. Rohrer, 372 – 378. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.
- Barton, R. R., B. L. Nelson, and W. Xie. 2010. "A framework for input uncertainty analysis". In *Proceedings of the 2010 Winter Simulation Conference*, edited by B. Johansson, S. Jain, J. Montoya-Torres, J. Hugan, and E. Yucesan, 1189 – 1198. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.
- Bernardo, J. M. and A. F. M. Smith. 2000. *Bayesian Theory*. United Kingdom: Wiley & Sons.
- Biller, B. and C. G. Corlu. 2013. "Subset selection procedures for stochastic simulations with unknown input parameters". Tepper Working Paper, Carnegie Mellon University, Pittsburgh, PA.
- Biller, B., C. G. Corlu, and S. Tayur. 2013a. "Maximizing demand fulfillment probability under parameter uncertainty". Tepper Working Paper, Carnegie Mellon University, Pittsburgh, PA.
- Biller, B., C. G. Corlu, and S. Tayur. 2013b. "Service-level estimation in inventory simulations: The case of limited historical demand data". Tepper Working Paper, Carnegie Mellon University, Pittsburgh, PA.
- Boesel, J., B. L. Nelson, and S.-H. Kim. 2003. "Using ranking and selection to "clean up" after simulation optimization". *Operations Research* 51: 814 – 825.
- Chen, E. J. 2008. "Restricted subset selection". In *Proceedings of the 2008 Winter Simulation Conference*, edited by S. J. Mason, R. R. Hill, L. Manch, O. Rose, T. Jefferson, and J. W. Fowler, 281 – 289. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.
- Chen, E. J. 2010. "Restricted subset selection and its application". *Journal of Simulation* 4: 116 – 127.
- Chen, E. J. 2012. "Some insights of using common random numbers in selection procedures". *Discrete Event Dynamic Systems: Theory and Applications*. Published online on June 2012. Springer Science and Business Media, LLC.
- Cheng, R. C. H. and W. Holland. 1997. "Sensitivity of computer simulation experiments to errors in input data". *Journal of Statistical Computation and Simulation* 57: 327 – 335.
- Cheng, R. C. H. and W. Holland. 1998. "Two-point methods for assessing variability in simulation output". *Journal of Statistical Computation and Simulation* 60: 183 – 205.
- Cheng, R. C. H. and W. Holland. 2004. "Calculation of confidence intervals for simulation output". *ACM Transactions on Modeling and Computer Simulation* 14: 344 – 362.
- Chick, S. E. 2001. "Input distribution selection for simulation experiments: Accounting for input uncertainty". *Operations Research* 49: 744 – 758.
- Gupta, S. S. 1965. "On some multiple decision (ranking and selection) rules". *Technometrics* 7: 225 – 245.
- Gupta, S. S. and T. J. Santner. 1973. "On selection and ranking procedures – a restricted subset selection rule". *Proceedings of the 39th Session of the International Statistical Institute* 1: 478 – 486.
- Helton, J. C. 1997. "Uncertainty and sensitivity analysis in the presence of stochastic and subjective uncertainty". *Journal of Statistical Computation and Simulation* 57: 3 – 76.

- Hogg, R. V., J. W. McKean, and A. T. Craig. 2005. *Introduction to Mathematical Statistics*. 6th ed. Upper Saddle River, NJ: Pearson Prentice Hall.
- Hsu, L. and S. Panchapakesan. 2005. A restricted subset selection rule for selecting at least one of the  $t$  best normal populations in terms of their means: Common known variance case. *Advances in Ranking and Selection, Multiple Comparisons, and Reliability Statistics for Industry and Technology* 117 – 131.
- Nelson, B. L., J. Swann, D. Goldsman, and W. -M. T. Song. 2001. “Simple procedures for selecting the best system when the number of alternatives is large”. *Operations Research* 49: 950 – 963.
- Ng, S. H. and S. E. Chick. 2006. “Reducing parameter uncertainty for stochastic systems”. *ACM Transactions on Modeling and Computer Simulation* 16: 26 – 51.
- Raftery, A. E., D. Madigan, and C. T. Volinsky. 1996. “Accounting for model uncertainty in survival analysis improves predictive performance (with discussion)”. In *Bayesian Statistics*, edited by J. M. Bernardo, J. O. Berger, A. P. David, and A. F. M. Smith, 323 – 349. Oxford Press.
- Santner, T. J. 1976. “A generalized goal in restricted subset selection theory”. *Sankhya* B38: 129 – 143.
- Sullivan, D. W. and J. R. Wilson. 1989. “Restricted subset selection procedures for simulation”. *Operations Research* 37: 52 – 71.
- Wilson, J. R. 2001. “A multiplicative decomposition property of the screening-and-selection procedures of Nelson et al”. *Operations Research* 49: 964 – 966.
- Zouaoui, F. and J. R. Wilson. 2003. “Accounting for parameter uncertainty in simulation input modeling”. *IIE Transactions* 35: 781 – 792.
- Zouaoui, F. and J. R. Wilson. 2004. “Accounting for input-model and input-parameter uncertainties in simulation”. *IIE Transactions* 36: 1135 – 1151.

#### **AUTHOR BIOGRAPHIES**

**CANAN GUNES CORLU** is an Assistant Professor of Administrative Sciences at Boston University. Her research interests include the design of large-scale simulations with applications to inventory management, and the applications of Operations Research techniques (e.g., vehicle routing, inventory control) tailored to the problems of non-profit sectors.

**BAHAR BILLER** is an Associate Professor of Operations Management and Manufacturing at Carnegie Mellon University. Her primary research interest lies in the area of computer simulation experiments for stochastic systems and more specifically, in multivariate input modeling for dependent input processes with applications to operations management.