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

FINDING THE BEST IN THE PRESENCE OF A STOCHASTIC CONSTRAINT

Sigrún Andradóttir
David Goldsman
Seong-Hee Kim

School of Industrial and Systems Engineering
Georgia Institute of Technology
Atlanta, GA 30332-0205, U.S.A.

ABSTRACT

Our problem is that of finding the best system—i.e., the system with the largest or smallest primary performance measure—among a finite number of simulated systems in the presence of a stochastic constraint on a secondary performance measure. In order to solve this problem, we first find a set that contains feasible or near-feasible systems (Phase I) and then choose the best among those systems in the set (Phase II). We present a procedure for Phase I; and then we propose another procedure that performs Phases I and II sequentially to find the best feasible system.

1 INTRODUCTION

Our goal is to select the best or near-best system from a set of competing systems, where the term "best" is with respect to a primary performance measure among the systems, which we want to maximize or minimize. The fly in the ointment is that we also want the selected system to satisfy a stochastic constraint on a secondary performance measure.

Due to randomness in output data, one needs to be careful when comparing a number of simulated systems. Over the last decade, there have been fruitful efforts in developing statistically valid ranking and selection (R&S) procedures that find the best among a finite number of simulated alternatives. Boesel, Nelson, and Kim (2003), Nelson et al. (2001), Kim and Nelson (2001, 2005), Goldsman et al. (2000, 2002), Chen (1996), Chen et al. (1997, 2000), Chick (1997), and Chick and Inoue (2001a, 2001b) developed different types of statistically valid or heuristic selection procedures that are useful in simulation, when the goal is to find a system with the minimum or maximum expected performance measure among a finite number of simulated systems. They showed that their procedures are a great more efficient than classical R&S procedures such as that due to Rinott (1978). A limitation of those procedures lies in the fact that they consider only one performance measure. In reality, we often face constraints on performance measures other than a primary performance measure due to physical or managerial limits placed on a system. When a stochastic constraint is present, R&S procedures lose their statistical guarantee about a correct selection, defeating the biggest advantage of using R&S procedures over heuristic procedures.

There is not a rich literature on solving discrete optimization problems with stochastic constraints using simulation. Butler, Morrice, and Mullarkey (2001) combined multiple attribute utility theory from economics with the Rinott (1978) procedure to handle multiple performance measures. Their procedure finds a system that gives the highest utility. But it does not handle constraints; and finding an appropriate attribute utility function itself is a very difficult problem. Santner and Tamhane (1984) proposed a two-stage procedure that is specially designed to find a system with the largest mean among those whose variances are smaller than a constant when the data under consideration are are independent and identically distributed (IID) normal. This two-stage procedure handles only a special case and can not be applied for a general stochastic constraint. Another problem with these procedures is that they become inefficient for 20 or more systems since they adopt Rinott's procedure—see Boesel, Nelson, and Kim (2003), who explain where the inefficiency comes from.

Fully-sequential procedures take only a single basic observation from each alternative still in play at each stage of sampling and apply a decision strategy at every stage to eliminate apparently inferior alternatives early in the experimentation process. For this reason, such procedures are expected to reduce the overall simulation effort required to find the best system. Paulson (1964) and Hartmann (1988, 1991) presented fully sequential procedures in the case of IID normal data with equal variances; and Kim and Nelson (2001) extended their procedures in a direction to be more appropriate

for simulation environments where it makes sense to assume unequal variances across systems and dependence across systems due to the use of common random numbers (CRN). In this paper, we propose an efficient fully-sequential procedure that handles a general stochastic constraint.

This paper is organized as follows: In Section 2, we formulate the problem of interest. In Section 3, we present a procedure that finds a set of systems that satisfy a stochastic constraint, and then Section 4 gives a new procedure for the finding-the-best problem with a constraint. We present experimental results from the new procedure in Section 5, and conclude the paper with Section 6.

2 PROBLEM

In this section, we formulate our problem and define notation for the paper. We assume that output data from a system are IID normally distributed and that all systems are simulated independently. IID normality is plausible as long as basic observations are either within-replication averages or batch means with a large batch size. Let X_{ij} be an observation associated with a primary performance measure from replication j of system i and Y_{ij} be an observation associated with a constraint (a secondary performance measure) from replication j of system i. The expected primary and secondary performance measures are defined as $x_i = E[X_{ij}]$ and $y_i = E[Y_{ij}]$, respectively. Then our problem amounts to finding

$$\begin{aligned} \text{argmax}_{i=1,\ldots,k} \quad & x_i \\ \text{s.t.} \quad & y_i \le Q, \end{aligned}$$

where Q is a constant. We allow for dependence between X_{ij} and Y_{ij}—actually, they are likely to be correlated. For instance, the throughput and downtime of a production system are usually negatively correlated.

For stochastic systems, it is not possible to guarantee that we identify all systems satisfying the stochastic constraint. Instead we adopt an idea similar to that of the *indifference-zone* approach of Bechhofer (1954) to find a set of feasible or near-feasible systems. A decision maker will be asked to give a range around the constant Q, say, (Q_ℓ, Q_u) for $Q_\ell < Q_u$. Then three regions are defined:

- $y_i \le Q_\ell$: This is the *desirable region*. Any system in this range is feasible.
- $Q_\ell < y_i < Q_u$: This is the *acceptable region*. It is possible that a feasible system in this range is declared infeasible and vice versa. We assume that one is willing to accept a system in this range as the best if it is declared "feasible" and our procedure chooses it as the best.
- $y_i \ge Q_u$: This is the *unacceptable region*. A system in this range is *infeasible* and should be declined.

We can also define the following three sets for the constraint.

$$\begin{aligned} S_D &= \text{the set of all desirable systems;} \\ S_A &= \text{the set of all acceptable systems;} \\ S_U &= \text{the set of all unacceptable systems.} \end{aligned}$$

With this setup, there are two types of risk:

- If a feasible system is in the acceptable region, then it may be declared infeasible and declined. We will call this Risk I.
- If an infeasible system is in the acceptable region, then it may be declared feasible and accepted. We will call this Risk II.

The concepts of Risk I or Risk II are very similar to those of Type I or Type II errors from classical hypothesis testing. Values of Q_ℓ and Q_u will be determined depending on which risk a decision maker feels is more important. For example, suppose that a constraint is imposed on the expected number of delayed demands per month for an inventory system and that it should be no more than 100. If the decision maker defines $(Q_\ell, Q_u) = (100, 110)$, it means that she feels Risk I should be minimized and Risk II is relatively less important. On the other hand, if she defines $(Q_\ell, Q_u) = (90, 100)$, this implies that Risk II should be minimized and Risk I is not that critical. By choosing $Q_\ell < Q$ and $Q_u > Q$, one can strike a balance between Risk I and Risk II. In any case, for given Q_ℓ and Q_u, we define q and ϵ as $q = (Q_\ell + Q_u)/2$ and $\epsilon = (Q_u - Q_\ell)/2$, respectively, and henceforth our procedures will be presented in terms of q and ϵ rather than Q_ℓ and Q_u. Roughly speaking, q is a *target value* that behaves as a cut-off point between acceptable and unacceptable systems and ϵ is a *tolerance level*, i.e., how much we are willing to be below or above q. "Greater than or equal to" constraints can be transformed to "less than or equal to" constraints by multiplying both sides by minus one.

In addition to q and ϵ, the decision maker needs to choose the *indifference-zone* parameter δ for the primary performance measure. This is the smallest absolute difference in the primary measure that the decision maker feels is important to detect. Thus, only "practical difference" matters, and we are indifferent among systems whose primary performance measures are within δ of the true best.

To solve the problem, we first find a set of systems that contains all desirable systems, possibly with some acceptable systems (Phase I), and then we choose the best among the systems in that set (Phase II). A procedure that can be used in Phase I is presented in Section 3. in Section 4, we propose procedure $\mathcal{AGK}$, where Phases I and II are performed sequentially—that is, Phase I is performed first and then Phase II is applied to the survivors from Phase I.

3 A PROCEDURE FOR FEASIBILITY DETERMINATION

In this section, we provide a procedure for Phase I that eliminates all unacceptable systems and returns a set of all desirable systems, possibly including some acceptable systems.

In addition to the IID normal assumption on basic observations from each system, we also assume that systems are simulated independently; CRN are not considered in this paper.

First, we define some notation. Let $i \in \{1, 2, \ldots, k\}$ and

$$\begin{aligned} n_0 &= \text{initial first-stage sample size}, n_0 \geq 2; \\ r &= \text{number of observations taken so far}; \\ S_i^2 &= \text{the sample variance of } Y_{ij}, \\ &\quad j = 1, 2, \ldots, n_0; \\ W(r; v, w, z) &= \max\left\{0, \frac{v}{2}\left(\frac{wz}{v^2} - r\right)\right\}, \\ &\quad \text{for any } v, w, z \in \mathbb{R}, v, z \neq 0. \end{aligned}$$

Now we give an algorithm that finds a set of feasible systems.

Algorithm I: Feasibility Determination Procedure

Setup: Select $n_0 \geq 2$ and nominal probability of correct selection (PCS) $1/k < 1 - \alpha_1 < 1$. For a constraint, choose ϵ. Compute

$$\eta_1 = \frac{1}{2}\left\{2\left[1 - (1-\alpha_1)^{\frac{1}{k}}\right]^{-\frac{2}{n_0-1}} - 1\right\}. \qquad (1)$$

Initialization: Let $R = \{1, 2, \ldots, k\}$ and $F = \emptyset$ be the set of undetermined systems and the set of feasible systems, respectively. Let $h^2 = 2c\eta_1 \times (n_0 - 1)$.

Obtain n_0 observations $Y_{ij}, j = 1, 2, \ldots, n_0$, from each system i. Compute S_i^2.

Set the observation counter $r = n_0$ and go to **Feasibility Check**.

Feasibility Check: For each system $i \in R$, if

$$\sum_{j=1}^{r}(Y_{ij} - q) \leq -W(r; \epsilon, h^2, S_i^2),$$

then move i from R to F; else if

$$\sum_{j=1}^{r}(Y_{ij} - q) \geq +W(r; \epsilon, h^2, S_i^2),$$

then eliminate i from R.

Stopping Rule: If $|R| = 0$, then return F as the set of feasible systems.

Otherwise, take one additional observation $Y_{i,r+1}$ from each system $i \in R$ and set $r = r + 1$, and go to **Feasibility Check**.

Remark: Algorithm I can be performed without simulating systems simultaneously, so we can avoid the overhead of switching between systems—which is a typical disadvantage of fully-sequential procedures.

With high probability, the returned set F will contain all desirable systems, plus possibly some acceptable systems, but no unacceptable systems. This event defines a correct selection (CS) for Algorithm I.

Now we present the main result from Algorithm I.

Theorem 1 *Suppose that $Y_{ij}, j = 1, 2, \ldots,$ are IID normally distributed, and Y_{ij} and $Y_{\ell j}$ are independent for $i \neq \ell$. Then the proposed procedure guarantees*

$$\Pr\{\text{CS}\} \equiv \Pr\{S_D \subseteq F \subseteq (S_D \cup S_A)\} \geq 1 - \alpha_1.$$

We refer the proof to Andradóttir, Goldsman, and Kim (2005).

Algorithm I is closely related to the procedures for comparison with a standard presented by Kim (2003, 2005). The comparison with a standard problem often turns up in the context of simulation (see Goldsman and Nelson 1998 for more details). It turns out that the procedures due to Kim (2003, 2005) can be interpreted as special cases of Algorithm I.

In the next section, we show how our work can be used for a discrete optimization problem with one stochastic constraint when the number of alternatives under consideration is finite (and not too large), so that we can simulate all the scenarios.

4 A PROCEDURE FOR COMPARING CONSTRAINED SYSTEMS

In this section, we present procedure $\mathcal{AGK}$ where Phases I and II are performed sequentially. Algorithm

I from Section 3 will be used in Phase I where we determine the feasibility of each system. For Phase II of $\mathcal{AGK}$, we adopt the Sequential Selection with Memory ($\mathcal{SSM}$) procedure due to Pichitlamken and Nelson (2001) and Pichitlamken, Nelson, and Hong (2005). They extend the Kim and Nelson (2001) procedure for use within an optimization-via-simulation algorithm when simulation is costly, and partial or complete information on solutions previously visited is maintained. Since we assume that simulation is costly, we want to obtain X_{ij} as well as Y_{ij} when system i is simulated in Phase I and then use those partial data in Phase II. However, saving X_{ij} for the later use in Phase II brings up a memory space issue, especially when Phase I is long and the number of survivors from Phase I is large. Procedure $\mathcal{SSM}$ requires saving only sample means and thus removes the memory space issue.

Before we present the new procedure, we need some more notation in addition to that defined in Section 3.

$$\begin{aligned} S^2_{i\ell} &= \text{sample variance of the difference, } X_{ij} - X_{\ell j} \\ &\quad \text{between systems } i \text{ and } \ell,\ j = 1, 2, \ldots, n_0; \\ \bar{X}_i(r) &= \text{sample average of } X_{ij},\ j = 1, 2, \ldots, r. \end{aligned}$$

4.1 Procedure $\mathcal{AGK}$

$\mathcal{AGK}$ simply applies Algorithm I first to find a set of feasible systems and then Phase II to find the best among the survivors from Phase I.

Algorithm II: Procedure $\mathcal{AGK}$

Setup: Select the overall confidence level $1 - \alpha$ and choose α_1 and α_2, where $\alpha_1 + \alpha_2 = \alpha$, for Phases I and II. Choose ϵ. Also select the indifference-zone parameter $\delta > 0$ and first-stage sample size $n_0 \geq 2$. Calculate η_1 from Equation (1).

Initialization for Phase I: Let $R = \{1, 2, \ldots, k\}$ and $F = \emptyset$ be the set of undetermined systems and the set of feasible systems, respectively. Let $h_1^2 = 2c\eta_1 \times (n_0 - 1)$.

Obtain n_0 observations X_{ij} and Y_{ij}, $j = 1, 2, \ldots, n_0$, from each system $i = 1, 2, \ldots, k$.

For all i and $\ell \neq i$, compute the estimators S_i^2 and $S_{i\ell}^2$.

For each system i, compute $\bar{X}_i(n_0)$ and set $n_i = n_0$, the number of observations taken so far from system i in *Phase I*.

Set the observation counter $r = n_0$ and go to **Feasibility Check**.

Feasibility Check: Same as in Algorithm I in Section 3.

Stopping Rule for Phase I: If $|R| = 0$, then

- if $|F| = 0$, stop and return "no feasible system",
- if $|F| = 1$, stop and return the system as the best,
- otherwise, go to **Initialization for Phase II**.

Otherwise, take one additional observation $X_{i,r+1}$ and $Y_{i,r+1}$. Set $r = r + 1$ and $n_i = n_i + 1$ for each system $i \in R$. Then, update $\bar{X}_i(r)$ and go to **Feasibility Check**.

Initialization for Phase II: Set $m = |F|$. Compute

$$\eta_2 = \frac{1}{2}\left\{\left(\frac{2\alpha_2}{m-1}\right)^{-\frac{2}{n_0-1}} - 1\right\},$$

and $h_2^2 = 2c\eta_2 \times (n_0 - 1)$.

Let $R = F$ be the set of systems still in contention.

Set $r = n_0$ and go to **Comparison**.

Comparison: Set $R^{\text{old}} = R$. Let

$$\begin{aligned} R = \Big\{ & i : i \in R^{\text{old}} \text{ and} \\ & \sum_{j=1}^{r} \mathcal{X}_{ij} \geq \sum_{j=1}^{r} \mathcal{X}_{\ell j} - W(r; \delta, h_2^2, S_{i\ell}^2), \\ & \forall \ell \in R^{\text{old}},\ \ell \neq i \Big\} \end{aligned}$$

where

$$\mathcal{X}_{ij} = \begin{cases} X_{ij} & \text{if } n_i < r, \\ \bar{X}_i(n_i) & \text{if } n_i \geq r. \end{cases}$$

Stopping Rule for Phase II: If $|R| = 1$, then stop and select the system whose index is in I as the best.

Otherwise, take one additional observation $X_{i,r+1}$ from each system $i \in R$ with $n_i < r$, and set $r = r + 1$. Then go to **Comparison**.

We have not yet been successful in proving the statistical validity of $\mathcal{AGK}$, although our experiments support that its validity. Andradóttir, Goldsman, and Kim (2005) provide additional procedures that perform Phases I and II sequentially. They are statistically valid but less efficient than $\mathcal{AGK}$.

The correct selection event for $\mathcal{AGK}$ is defined as the event of selecting a system i such that $x_i > x_{[b]} - \delta$ and $i \in (S_D \cup S_A)$, where $[b]$ is the identity of the system with the largest primary performance measure among the systems in S_D.

5 EXPERIMENTAL RESULTS

In this section we illustrate the performance of $\mathcal{AGK}$ based on experiments in which we use bivariate normal random variables for $(X_{i\ell}, Y_{i\ell})$. For replication ℓ of system $i = 1, 2, \ldots, k$, we assume that the vector $(X_{i,\ell}, Y_{i,\ell})$ is bivariate normal with correlation ρ. We test $k = 2, 15, 25, 101$ and $\rho = -0.9, -0.6, 0, 0.3, 0.6, 0.9$. We can assume that the target value q for the secondary measure is $q = 0$ without loss of generality. We set $\delta = 1/\sqrt{n_0}$ and $\epsilon = 1/\sqrt{n_0}$.

For simplicity, we assume that no system is in the acceptable region. Therefore, all feasible systems are in the desirable region and all infeasible systems are the unacceptable region. Let f be the number of feasible systems. Then mean and variance configurations for the various systems under consideration are defined as follows:

- We consider two mean configurations: The difficult means (DM) configuration and the monotonically increasing means (MIM) configuration. In the DM configuration,

$$y_i = \begin{cases} -\epsilon, & i = 1, 2, \ldots, f \\ \epsilon, & i = f+1, \ldots, k \end{cases}$$

and

$$x_i = \begin{cases} 0, & i = 1, 2, \ldots, f-1 \\ \delta, & i = f \\ (i-1)\delta, & i = f+1, \ldots, k. \end{cases}$$

In DM, all feasible systems have y_i exactly ϵ smaller than q and all infeasible systems have y_i exactly ϵ larger than q. Therefore, it is difficult to distinguish between feasible and infeasible systems. In addition, all feasible systems have x_i very close to that of the true best feasible system, which makes it difficult to detect inferior systems. All infeasible systems have much larger primary performance measures x_i than that of the true best feasible system, which increases the chance for an infeasible system mistakenly declared feasible in Phase I to eliminate the true best feasible system. The DM configuration is used to test the validity of our new procedure.

To investigate the effectiveness of the procedures in eliminating infeasible or non-competitive systems, we use the MIM configuration. In the MIM configuration,

$$y_i = \begin{cases} -(f-i+1)\epsilon, & i = 1, 2, \ldots, f \\ (i-f)\epsilon, & i = f+1, \ldots, k \end{cases}$$

and

$$x_i = (i-1)\delta, \quad i = 1, \ldots, k.$$

In the MIM configuration, we only test the case where infeasible systems have larger x_i than feasible systems. This is difficult, since we should eliminate all infeasible systems in Phase I. If we fail to eliminate any infeasible system in Phase I, then the system is likely to be chosen as the best due to large its x_i, and this will increase the probability of an incorrect selection.

- We test three different variance configurations: constant (CONST) variances, increasing (INC) variances, and decreasing (DEC) variances. Let σ_{y_i} and σ_{x_i} denote the variances of observations associated with the secondary measure and primary measure of system i, respectively. In the CONST variances case, all variances are set to one. In INC, $\sigma^2_{y_i} = 1 + (i-1)\epsilon$ or $\sigma^2_{x_i} = 1 + (i-1)\delta$. In DEC, $\sigma^2_{y_i} = 1/(1 + (i-1)\epsilon)$ or $\sigma^2_{x_i} = 1/(1 + (i-1)\delta)$.

An example of mean and variance configurations when $k = 5$ is given in Table 1.

For the validity of the proposed procedures, we test the DM configuration with five different combinations of variance settings: CONST $\sigma^2_{y_i}$ and CONST $\sigma^2_{x_i}$, INC $\sigma^2_{y_i}$ and INC $\sigma^2_{x_i}$, INC $\sigma^2_{y_i}$ and DEC $\sigma^2_{x_i}$, DEC $\sigma^2_{y_i}$ and INC $\sigma^2_{x_i}$, and DEC $\sigma^2_{y_i}$ and DEC $\sigma^2_{x_i}$. For the effectiveness of the procedures, we test the MIM configuration with the same variance settings defined above. We make 10,000 macro-replications. Our required confidence level is $1 - \alpha = 0.95$.

5.1 Main Results

- Effect of Correlation: Correlation between observations of the primary and secondary performance measures does not seem to have any effect on the performance of the procedures. Table 2 shows the estimated total number of replications (REP) and estimated PCS when there are twenty-five systems with thirteen feasible systems and the variances of the primary and secondary measures are all increasing. As one can see, the REP and PCS do not change that much under various values of ρ. We observed the same tendency for all configurations we tested. This is expected since the derivation of $\mathcal{AGK}$ is based on a Bonferroni-type inequality which dissolves the effect of dependence—although the procedure itself is not statistically valid. Due to this insensitivity of the procedures to dependence, we will focus on the $\rho = 0$ case only from now on.

- The validity of $\mathcal{AGK}$: Table 3 shows the average total number of replications and estimated PCS when the DM configuration is employed with $f = 1 + k - 1/2$, $k = 5, 15, 25, 101$, and various variance configurations. The estimated PCS for

the procedure is well over 0.95 under all configurations tested. Although we did not include tables for $\rho > 0$, we indeed obtained similar results for negative or positive ρ (as mentioned before).

- The effectiveness of $\mathcal{AGK}$: Table 4 shows the average total number of replications and estimated PCS in the MIM configuration with various variance configurations. To test how effectively our procedure can eliminate infeasible or non-competitive systems, the MIM configuration is used instead of the DM configuration. One can notice that the estimated PCS is typically much larger than that under the DM configuration. This is expected since it is much easier to detect non-competitive systems when the differences between the true best and the other systems are large. Table 4 shows dramatic reduction in average total number of replications compared to Table 3. For example, for $k = 101$, $f = 51$, and increasing $\sigma^2_{Y_i}$ and increasing $\sigma^2_{X_i}$, $\mathcal{AGK}$ spent a total of 232,786 replications on average under the DM configuration, while it spent only 31,484 replications—more than 85% decrease—on average under the MIM configuration.

6 CONCLUSION

In this paper, we consider a simple problem where there is only one constraint. This shows the possibility that R&S procedures can be extended to a general discrete optimization problem with constraints in a small search space.

One can perform Phases I and II simultaneously instead of performing them sequentially as in $\mathcal{AGK}$. We call a procedure that performs Phase I and II simultaneously $\mathcal{AGK}+$ and it is presented in Andradóttir, Goldsman, and Kim (2005).

REFERENCES

Andradóttir, S, D. Goldsman, and S.-H. Kim. 2005. Fully sequential procedures for comparing constrained systems via simulation. Technical Report, School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA.

Bechhofer, R. E. 1954. A single-sample multiple decision procedure for ranking means of normal populations with known variances. *Annals of Mathematical Statistics*, 25:16–39.

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.

Butler, J., D. J. Morrice, and P. W. Mullarkey. 2001. A multiple attribute utility theory approach to ranking and selection. *Management Science*, 47:800–816.

Chen, C.-H. 1996. A lower bound for the correct-selection probability and its application to discrete event simulations. *IEEE Transactions on Automatic Control*, 41:1227-1231.

Chen, H. C. C. H. Chen, and E. Yücesan. 2000. Computing efforts allocation for ordinal optimization and discrete event simulation. *IEEE Transactions on Automatic Control*, 45: 960–964

Chen, H.-C., C.-H. Chen, L. Dai, and E. Yücesan. 1997. New development of optimal computing budget allocation for discrete event simulation, In *Proceedings of the 1997 Winter Simulation Conference*, ed. S. Andradóttir, K. J. Healy, D. H. Withers, and B. K. Nelson, 334–341. Institute of Electrical and Electronics Engineers, Piscataway, New Jersey.

Chick, S. E. 1997. Selecting the best system: A decision-theoretic approach. In *Proceedings of the 1997 Winter Simulation Conference*, ed. S. Andradóttir, K. J. Healy, D. H. Withers, and B.K. Nelson, 326–333. Institute of Electrical and Electronics Engineers, Piscataway, New Jersey.

Chick, S. E., and K. Inoue. 2001a. New two-stage and sequential procedures for selecting the best simulated system. *Operations Research* 49:732–743.

Chick, S. E., and K. Inoue. 2001b. New procedures for identifying the best simulated system using common random numbers. *Management Science* 47:1133–1149.

Goldsman, D., S.-H. Kim, W. S. Marshall and B. L. Nelson. 2000. Ranking and selection for steady-state simulation. In *Proceedings of the 2000 Winter Simulation Conference*, ed. J. Joines, R. R. Barton, P. Fishwick, and K. Kang, 544–553, Institute of Electrical and Electronics Engineers, Piscataway, New Jersey.

Goldsman, D., S.-H. Kim, W. Marshall and B. L. Nelson. 2002. Ranking and selection procedures for steady-state simulation: Perspectives and procedures. *INFORMS Journal on Computing*, 14:2–19.

Goldsman, D. and B. L. Nelson. 1998. Comparing systems via simulation. In *Handbook of Simulation: Principles, Methodology, Advances, Applications, and Practice*, ed. J. Banks, 273–306. New York: John Wiley & Sons.

Hartmann, M. 1988. An improvement on Paulson's sequential ranking procedure. *Sequential Analysis*, 7:363–372.

Hartmann, M. 1991. An improvement on Paulson's procedure for selecting the population with the largest mean from k normal populations with a common unknown variance. *Sequential Analysis*, 10:1–16.

Kim, S.-H.. 2003. Comparison with a standard via fully sequential procedures. In *Proceedings of the 2003 Winter Simulation Conference*, ed. S. chick, P. J. Sanchez, D. Ferrin, and D. J. Morrice, 528–534, Institute of Electrical and Electronics Engineers, Piscataway, New Jersey.

Kim, S.-H.. 2005. Comparison with a standard via fully sequential procedures. *ACM TOMACS*, 15:1–20.

Kim, S.-H. and B. L. Nelson. 2001. A fully sequential procedure for indifference-zone selection in simulation. *ACM TOMACS*, 11:251–273.

Kim, S.-H., and B. L. Nelson. 2005. On the asymptotic validity of fully sequential selection procedures for steady-state simulation. *Operations Research*, forthcoming.

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.

Paulson, E. 1964. A sequential procedure for selecting the population with the largest mean from k normal populations. *Annals of Mathematical Statistics* 35:174–180.

Pichitlamken, J. and B. L. Nelson. 2001. A selection-of-the best procedure for optimization via simulation. In *Proceedings of the 2001 Winter Simulation Conference*, ed. B.A. Peters, J.S. Smith, D.J. Medeiros, and M.W. Rohrer, 401–407. Institute of Electrical and Electronics Engineers, Piscataway, New Jersey.

Pichitlamken, J., B. L. Nelson, and L. J. Hong. 2005. A sequential procedure for neighborhood selection-of-the-best in optimization via simulation. *European Journal of Operational Research*, forthcoming.

Rinott, Y. 1978. On two-stage selection procedures and related probability-inequalities. *Communications in Statistics*, A7:799–811.

Santner, T. J. and A. C. Tamhane 1984. Designing experiments for selecting a normal population with a large mean and a small variance, In *Design of Experiments – Ranking and Selection: Essays in Honor of Robert E. Bechhofer*, ed. T. J. Santer and A. C. Tamhane, 179–198, Marcel-Dekker, New York.

AUTHOR BIOGRAPHIES

SIGRÚN ANDRADÓTTIR is a Professor in the School of Industrial and Systems Engineering at the Georgia Institute of Technology. She received a B.S. in Mathematics from the University of Iceland in 1986, an M.S. in Statistics from Stanford University in 1989, and a Ph.D. in Operations Research from Stanford University in 1990. Her research interests include stochastic optimization, simulation, and stochastic processes. Her e-mail address is <sa@isye.gatech.edu>.

DAVID GOLDSMAN is a Professor in the School of Industrial and Systems Engineering at the Georgia Institute of Technology. He received his Ph.D. in Operations Research and Industrial Engineering from Cornell University. His research interests include simulation output analysis and ranking and selection. He is an active participant in the Winter Simulation Conference, having been Program Chair in 1995, and having served on the WSC Board of Directors since 2002. His e-mail address is <sman@isye.gatech.edu>, and his web page is <www.isye.gatech.edu/~sman/>.

SEONG-HEE KIM is an Assistant Professor in the School of Industrial and Systems Engineering at the Georgia Institute of Technology. She received her Ph.D. in Industrial Engineering and Management Sciences from Northwestern University. Her research interests include ranking and selection, quality control, and simulation output analysis. Her e-mail and web addresses are <skim@isye.gatech.edu> and <www.isye.gatech.edu/~skim/>.

Table 1: Example MIM Mean and Associated Variance Configurations when $k = 5$ and $f = 3$

	X				Y			
System	Mean		Variance		Mean		Variances	
i	DM	MIM	INC	DEC	DM	MIM	INC	DEC
1	0	0	1	1	$-\epsilon$	-3ϵ	1	1
2	0	δ	$1+\delta$	$1/(1+\delta)$	$-\epsilon$	-2ϵ	$1+\epsilon$	$1/(1+\epsilon)$
3	δ	2δ	$1+2\delta$	$1/(1+2\delta)$	$-\epsilon$	$-\epsilon$	$1+2\epsilon$	$1/(1+2\epsilon)$
4	3δ	3δ	$1+3\delta$	$1/(1+3\delta)$	ϵ	ϵ	$1+3\epsilon$	$1/(1+3\epsilon)$
5	4δ	4δ	$1+4\delta$	$1/(1+4\delta)$	ϵ	2ϵ	$1+4\epsilon$	$1/(1+4\epsilon)$

Table 2: Average Total Number of Replications for $\rho = -0.9, 0, 0.9$ when $k = 25$ and $f = 13$ with the DM Configuration, INC $\sigma^2_{y_i}$, and INC $\sigma^2_{x_i}$

$\rho = -0.9$		$\rho = 0$		$\rho = 0.9$	
REP	PCS	REP	PCS	REP	PCS
14101	0.976	14110	0.974	13982	0.974

Table 3: Total Number of Replications and Estimated PCS of $\mathcal{AGK}$ when the DM Configuration is Employed, $\rho = 0$, and $f = 1 + (k-1)/2$

	CONST, CONST		INC, INC		DEC, INC		INC, DEC		DEC, DEC	
k	REP	PCS	REP	PCS	REP	PCS	REP	PCS	REP	PCS
5	576	0.969	802	0.967	588	0.965	669	0.972	431	0.972
15	2233	0.971	5433	0.976	2957	0.971	4120	0.983	1083	0.973
25	4063	0.973	14110	0.974	6819	0.978	10892	0.987	1527	0.975
101	20566	0.974	232786	0.979	88043	0.989	191525	0.989	3669	0.986

Table 4: Average Total Number of Replications and Estimated PCS of $\mathcal{AGK}$ when the MIM Configuration is Employed, $\rho = 0$, and $f = 1 + (k-1)/2$

	CONST, CONST		INC, INC		DEC, INC		INC, DEC		DEC, DEC	
k	REP	PCS	REP	PCS	REP	PCS	REP	PCS	REP	PCS
5	459	0.981	652	0.981	491	0.979	521	0.981	337	0.983
15	1024	0.994	2615	0.994	1695	0.993	1763	0.997	496	0.995
25	1387	0.996	5015	0.996	3141	0.997	3420	0.998	613	0.997
101	3326	0.999	31484	0.999	18624	1.000	22497	1.000	2031	0.999