RECENT ADVANCES IN RANKING AND SELECTION

Seong-Hee Kim

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

Barry L. Nelson

Department of Industrial Engineering and Management Sciences
Northwestern University
Evanston, IL 60208-3119, U.S.A.

ABSTRACT

This tutorial provides an overview on recent advances made in ranking and selection (R&S) for selecting the best simulated system and discusses challenges that still exist in the field. We focus on indifference-zone R&S procedures that provide a guaranteed probability of correct selection when the best system is at least a user-specified amount better than the other systems.

1 INTRODUCTION

Ranking-and-selection (R&S) procedures are to compare a finite number of simulated alternatives. R&S procedures were first developed in the 1950's by the statistics community. In the early 1980's, R&S drew the attention of the simulation community due to its potential usefulness in *stochastic* simulation output analysis. There exist a number of simulation issues that make simulation R&S more difficult than classical R&S. There are also opportunities available in simulation experiments that are not present in physical experiments. Fortunately, a number of researchers in that arena have been successful in developing new R&S procedures that are useful for simulation environments.

Our goal is to review these simulation issues, how they have been overcome, and what challenges still lie in the field. We present three R&S procedures: one to illustrate simulation issues of classical R&S procedures and the other two as example procedures useful for simulation outputs. See Goldsman and Nelson (1998) and Law and Kelton (2000) for detailed "how to" guides, Bechhofer et al. (1995) for

a comprehensive survey of R&S procedures; and Kim and Nelson (2006b) for how to construct R&S procedures.

There exist at least four classes of comparison problems that arise in simulation studies: selecting the system with the largest or smallest expected performance measure (selection of the best), comparing all alternatives against a standard (comparison with a standard), selecting the system with the largest probability of actually being the best performer (multinomial selection), and selecting the system with the largest probability of success (Bernoulli selection). For all of these problems, a requirement is imposed either on the probability of correct selection (PCS) or on the simulation budget. Some procedures find a desirable system with a guarantee on the PCS, while other procedures maximize the PCS under the budget constraint. Our focus is on selection-of-the-best problems with a PCS constraint. A good procedure is one that delivers the desired PCS efficiently (with minimal simulated data) and is robust to modest violations of its underlying assumptions. In this tutorial "best" means maximum expected value of performance, such as expected throughput or profit.

The paper is organized as follows: In Section 2 we define notation, present definitions and provide one classical R&S procedure to illustrate problems in classical R&S. Section 3 reviews issues and opportunities encountered in simulation problems, along with how they have been overcome. Two specific procedures that are useful in simulation environments are presented in Section 4, followed by a numerical illustration in Section 5. Section 6 discusses challenges that still lie in the field and directions to approach them. Finally,

Section 7 closes the paper by describing other approaches of the R&S problem and giving appropriate references.

2 BACKGROUND

In this section, we define notation, provide definitions and give a classical R&S procedure due to Rinott (1978).

Rinott's procedure guarantees, with confidence level at least $1 - \alpha$, that the system ultimately selected has the largest true mean when the true mean of the best system is at least δ better than the second best. When there are inferior systems whose means are within δ of the true best, then the procedure guarantees to find one of these "close enough" systems with the same probability. The parameter δ, which defines the *indifference zone*, is set by the experimenter to the smallest absolute difference in expected performance that is considered important to detect. Differences of less than δ are considered practically insignificant. Procedures of this type are known as *indifference-zone(IZ) R&S procedures*. Comprehensive reviews of R&S can be found in Bechhofer et al. (1995) and Goldsman and Nelson (1998). The procedure below from Rinott (1978) is sequential, by which we mean it typically requires more than one stage of simulation.

Suppose that there are $k \geq 2$ systems, and let X_{ij} denote the jth independent observation from system i. Both procedures assume that the $X_{ij} \sim \mathrm{N}(\mu_i, \sigma_i^2)$, with μ_i and σ_i^2 unknown, and that the data across systems are independent. Also let $\bar{X}_i(r) = r^{-1} \sum_{j=1}^{r} X_{ij}$ denote the sample mean of the first r observations from system i.

Procedure R due to Rinott (1978) requires at most two stages of simulation; it is one of the simplest and most well-known R&S procedures.

Procedure R

1. *Setup:* Select confidence level $1 - \alpha$, IZ parameter $\delta > 0$ and first-stage sample size $n_0 \geq 2$.
2. *Initialization:* Obtain Rinott's constant $h = h(n_0, k, 1 - \alpha)$ from the tables in Wilcox (1984) or Bechoffer et al. (1995). See also Table 8.3 in Goldsman and Nelson (1998).
   Obtain n_0 observations $X_{ij}, j = 1, 2, \ldots, n_0$, from each system $i = 1, 2, \ldots, k$.
   For $i = 1, 2, \ldots, k$ compute

$$S_i^2 = \frac{1}{n_0 - 1} \sum_{j=1}^{n_0} \left(X_{ij} - \bar{X}_i(n_0) \right)^2,$$

the sample variance of the data from system i. Let

$$N_i = \max \left\{ n_0, \left\lceil \frac{h^2 S_i^2}{\delta^2} \right\rceil \right\}$$

where $\lceil \cdot \rceil$ indicates rounding up any fractional part to the next larger integer. Here N_i is the number of observations that will be taken from system i.

3. *Stopping Rule:* If $n_0 \geq \max_i N_i$ then stop and select the system with the largest $\bar{X}_i(n_0)$ as the best. Otherwise, take $N_i - n_0$ additional observations $X_{i,n_0+1}, X_{i,n_0+2}, \ldots, X_{i,N_i}$ from each system i for which $N_i > n_0$.
   Select the system with the largest $\bar{X}_i(N_i)$ as the best.

This procedure is a good example of classical R&S procedures that work—in the sense that it satisfies the PCS requirement—under assumptions such as independence and normality. Specifically, the procedure guarantees

$$\Pr\{\text{select } k | \mu_k - \mu_{k-1} \geq \delta\} \geq 1 - \alpha \qquad (1)$$

where $1/k < 1 - \alpha < 1$.

To give a brief proof sketch of the statistical validity of the procedure, assume that variances are known and equal across all systems. Then the procedure does not need two stages, becomes a one-stage procedure, and N_i becomes n such that

$$n = \left\lceil \frac{2c^2 \sigma^2}{\delta^2}, \right\rceil \qquad (2)$$

where c is a constant. We also need to define a configuration called *least-favorable configuration* (LFC), the configuration of system means under which it is most difficult to correctly select the best. The Slippage Configuration (SC), $\mu_i = \mu_k - \delta$ for $i = 1, 2, \ldots, k-1$, is known to be the LFC for many procedures and thus minimizes the PCS. In this paper we break from the statistics literature in that we will not be concerned with identifying the LFC; our only interest is insuring that (1) is met.

Then the proof for known and equal variances goes as follows: assuming $\mu_k - \mu_{k-1} \geq \delta$,

$$
\begin{aligned}
&\Pr\{\text{select } k\} \\
&= \Pr\left\{ \bar{X}_k(n) > \bar{X}_i(n), \forall i \neq k \right\} \\
&= \Pr\left\{ \frac{\bar{X}_i(n) - \bar{X}_k(n) - (\mu_i - \mu_k)}{\sigma \sqrt{2/n}} < -\frac{(\mu_i - \mu_k)}{\sigma \sqrt{2/n}}, \forall i \neq k \right\} \\
&\geq \Pr\left\{ \frac{\bar{X}_i(n) - \bar{X}_k(n) - (\mu_i - \mu_k)}{\sigma \sqrt{2/n}} < \frac{\delta}{\sigma \sqrt{2/n}}, \forall i \neq k \right\} \\
&\geq \Pr_{\mathrm{LFC}}\left\{ \frac{\bar{X}_i(n) - \bar{X}_k(n) - (\mu_i - \mu_k)}{\sigma \sqrt{2/n}} < c, \forall i \neq k \right\} \\
&= \Pr\{Z_i < c, i = 1, 2, \ldots, k-1\} = 1 - \alpha,
\end{aligned}
$$

where $(Z_1, Z_2, \ldots, Z_{k-1})$ has a multivariate normal distribution with means 0, variances 1, and common pairwise correlations $1/2$, implying c needs to be the $1 - \alpha$ quantile of the maximum of such a multivariate normal random vector.

For the R procedure, additional attention is required for unknown and unequal variances. The Rinott's constant h is similar to c in the sense that it is also the $1 - \alpha$ quantile of a distribution, but the distribution is more complicated than a multivariate normal distribution. Kim and Nelson (2006b) give theorems useful to handle unknown and unequal variances. *Notice that the proof is closely related to the collection of random variables*

$$\bar{X}_i(n) - \bar{X}_k(n) - (\mu_i - \mu_k), i = 1, 2, \ldots, k - 1, \qquad (3)$$

and this is true for many other IZ R&S procedures.

3 SIMULATION ISSUES

Indifference-zone R&S procedures first try to satisfy the PCS requirement but this is difficult in general unless observations satisfy some assumptions. For example, the R procedure needs independent and identically distributed (i.i.d.) basic observations from a normal distribution, which may not hold in simulation. We review such issues and review how those issues have been overcome.

3.1 Non-normality and Dependence of Output Data

Raw output data from industrial and service simulations are rarely normally distributed nor independent. Surprisingly, non-normality and dependence are usually not concerns in simulation experiments that (a) are designed to make multiple independent replications, and (b) use a within-replication average of a large number of raw simulation outputs as the basic summary measure. This is frequently the situation for so-called "terminating simulations" in which the initial conditions and stopping time for each replication are an inherent part of the definition of the system. A standard example is a store that opens empty at 6 AM, then closes when the last customer to arrive before 9 PM leaves the store. If the output of interest is the average customer delay in the checkout line over the course of the day, and comparisons will be based on the expected value of this average, and the average is over many individual customer delays, then the Central Limit Theorem suggests that the replication averages will be approximately normally distributed. Moreover, if each replication is independent, the replication averages will be independent, too. Whether or not each replication is independent relies on the capability of a pseudo-random number generator to generate a long sequence of numbers that look like i.i.d. uniform(0,1). See L'Ecuyer (2006) for a number of reliable generators with good statistical properties.

Difficulties arise in so-called "steady-state simulations" where the parameter of interest is defined by a limit as the time index of a stochastic process approaches infinity (and therefore forgets its initial conditions). Some steady-state simulations are amenable to multiple replications of each alternative and within-replication averages as summary statistics, in which case the preceding discussion applies. Unfortunately, severe estimator bias due to residual effects of the initial conditions sometimes force an experiment design consisting of a single, long replication from each alternative. The raw outputs within each replication are typically neither normally distributed nor independent. For example, waiting times of individual customers in a queueing system are usually dependent because a long delay for one customer tends to increase the delays of the customers who follow. The best we can hope for is an approximately stationary output process from each system, but not normality or independence.

The most common approach for dealing with this problem is to transform the raw data into *batch means*, which are averages of large number of raw outputs. The batch means are often far less dependent and non-normal than the raw output data. Many algorithms are available that determine a batch size to achieve approximate i.i.d. normal batch means. For example, see Schmeiser (1982). There are problems with the batching approach for R&S, however. They are discussed in Section 3.5.

3.2 Common Random Numbers

The R procedure assumes that data across the k alternative systems are independent. In simulation experiments this assumption can be made valid by using different sequences of random numbers to drive the simulation of each system. However, since we are making comparisons, there is a potential advantage to using common random numbers (CRN) to drive the simulation of each system because

$$\text{Var}[X_{ij} - X_{\ell j}] = \text{Var}[X_{ij}] + \text{Var}[X_{\ell j}] - 2\text{Cov}[X_{ij}, X_{\ell j}].$$

If implemented correctly (see, for instance, Banks, et al. 2001), CRN tends to make $\text{Cov}[X_{ij}, X_{\ell j}] > 0$ thereby reducing the variance of the difference.

As pointed out in Section 2, R&S procedures often need to make probability statements about the collection of random variables

$$\bar{X}_i(n) - \bar{X}_k(n) - (\mu_i - \mu_k), i = 1, 2, \ldots, k - 1. \qquad (4)$$

The appearance of the common term $\bar{X}_k(n)$ causes dependence among these random variables, but it is often easy to model or tightly bound. The introduction of CRN induces dependence between $\bar{X}_i(n)$ and $\bar{X}_k(n)$ as well. Even though the sign of the induced covariance is believed known, its value is not, making it difficult to say anything about the dependence among the differences (4).

Two approaches are frequently used. The first is to replace the basic data $\{X_{ij}; i = 1, 2, \ldots, k; j = 1, 2, \ldots, n\}$

with pairwise differences $\{X_{ij} - X_{\ell j}; i \neq \ell; j = 1, 2, \ldots, n\}$ because the variance of the sample mean of the difference includes the effect of the CRN-induced covariance. The second is to apply the Bonferroni inequality to break up joint statements about (4) into statements about the individual terms. Recall that for events $\mathcal{E}_1, \mathcal{E}_2, \ldots, \mathcal{E}_{k-1}$, the Bonferroni inequality states that

$$\Pr\left\{\bigcap_{i=1}^{k-1} \mathcal{E}_i\right\} \geq 1 - \sum_{i=1}^{k-1} \Pr\{\mathcal{E}_i^c\}. \tag{5}$$

Approaches based on the Bonferroni inequality make no assumption about the induced dependence, and therefore are very conservative. A more aggressive approach is to assume some structure for the dependence induced by CRN. One standard assumption is that all pairwise correlations $\rho = \mathrm{Corr}[X_{ij}, X_{\ell j}]$ are positive, but identical, and all variances are equal; this is known as *compound symmetry*. Nelson and Matejcik (1995) extended the R procedure in conjunction with CRN under a more general structure called *sphericity*. The specific assumption is

$$\mathrm{Cov}[X_{ij}, X_{\ell j}] = \begin{cases} 2\beta_i + \tau^2, & i = \ell \\ \beta_i + \beta_\ell, & i \neq \ell \end{cases} \tag{6}$$

with $\tau^2 > 0$, which is equivalent to assuming that $\mathrm{Var}[X_{ij} - X_{\ell j}] = 2\tau^2$ for all $i \neq \ell$, a type of variance balance. This particular structure is useful because there exists an estimator $\widehat{\tau}^2$ of τ^2 that is independent of the sample means and has a χ^2 distribution. Nelson and Matejcik (1995) showed that procedures based on this assumption are robust to departures from sphericity when the number of systems k is no more than 20, at least in part because assuming sphericity is like assuming that all pairwise correlations equal the average pairwise correlation.

3.3 The Sequential Nature of Simulation

The R procedure is sequential in that the procedure requires two stages of simulation. There are procedures that take only a single basic output from each alternative still in contention at each stage. Such procedures are *fully sequential*.

Suppose an IZ ranking procedure is applied in the study of k new blood pressure medications. Then "replications" correspond to patients, and the idea of using a fully-sequential procedure (assign one patient at a time to each drug, then wait for the results before recruiting the next patient) seems absurd. In simulation experiments, however, data are naturally generated sequentially, at least within each simulated alternative, making multi-stage procedures much more attractive. However, there are some issues:

- In multiple-replication designs, sequential sampling is particularly attractive. All that needs to be retained to start the next stage of sampling is the ending random number seeds from the previous stage. In single-replication designs it can be more difficult to resume sampling from a previous stage, since the entire state of the system must be retained and restored.

- A hidden cost of using multi-stage procedures is the computational overhead in switching among the simulations of the k alternatives. On a single-processor computer, switching can involve saving output, state and seed information from the current system; swapping the program for the current system out of, and for the next system into, active memory; and restoring previous state and seed information for the next system. Thus, the overall computation effort includes both the cost of generating simulated data and the cost of switching. Hong and Nelson (2005b) look at sequential IZ procedures that attempt to minimize the total computational cost. Their procedures require only $k - 1$ number of switches.

- If k processors are available, then an attractive option is to assign each system to a processor and simulate in parallel. This is highly effective in conjunction with R&S procedures that require little or no coordination between the simulations of each system, such as subset-selection procedures—that divide systems into the maybe-the-best group and the clearly-not-the-best group—or IZ-ranking procedures that use only variance information (and not differences among the sample means). Unfortunately, a fully-sequential procedure with elimination would defeat much of the benefit of parallel processing because communication among the processors is required after generating each output.

Many fully-sequential procedures are based on results for Brownian motion processes. Let $\mathscr{B}(t; \Delta)$ be a standard Brownian motion process with drift Δ. Consider the partial sum of the pairwise difference $D_i(r) = \sum_{j=1}^{r}(X_{kj} - X_{ij})$, $r = 1, 2, \ldots$. If the X_{ij} are i.i.d. normal and $\mu_k - \mu_i = \delta$, then $\{D_i(r), r = 1, 2, \ldots\} \stackrel{\mathscr{D}}{=} \{\sigma\mathscr{B}(t; \delta/\sigma), t = 1, 2, \ldots\}$, where $\sigma^2 = \mathrm{Var}[X_{kj} - X_{ij}]$ (with or without CRN). In other words, $D_i(r)$ is a Brownian motion process with drift observed only at discrete (integer) points in time. A great deal is known about the probability of Brownian motion processes crossing boundaries in various ways (see, for instance, Siegmund 1985 or Jennison and Turnbull 2000). Thus, it seems natural to design R&S procedures for $\sigma\mathscr{B}(t; \delta/\sigma)$ and apply them to $D_i(r)$ after taking care of unknown and unequal variances.

Here we provide one specific result. Let $c(t)$ be a symmetric (about 0) continuation region for $\sigma\mathscr{B}(t; \delta/\sigma)$, and let an incorrect selection correspond to the process

exiting the region in the wrong direction (down, when the drift is positive). If $T = \inf\{t \geq 0 : |\sigma\mathscr{B}(t;\delta/\sigma)| > c(t)\}$, then

$$\Pr\{\text{ICS}_i\} = \Pr\{\sigma\mathscr{B}(T;\delta/\sigma) < 0\}.$$

Of course $\sigma\mathscr{B}(t;\delta/\sigma)$ is only an approximation for $D_i(r)$. However, Jennison et al. (1980) show that under very general conditions, $\Pr\{\text{ICS}_i\}$ is no greater if the Brownian motion process is observed at discrete times; thus, procedures designed for $\sigma\mathscr{B}(t;\delta/\sigma)$ provide an upper bound on the probability of incorrect selection for $D_i(r)$. In conjunction with a decomposition into pairwise comparisons, this result can be used to derive R&S procedures for $k \geq 2$.

Fabian (1974) tightened the triangular continuation region used by Paulson (1964), and this was exploited by Hartmann (1988, 1991), Kim and Nelson (2001, 2006a) and Hong and Nelson (2005b).

Theorem 1 (Fabian 1974) *Let $\{\mathscr{B}(t,\Delta), t \geq 0\}$ be a standard Brownian motion with drift $\Delta > 0$. Let*

$$\begin{aligned} l(t) &= -a + \lambda t \\ u(t) &= a - \lambda t \end{aligned}$$

for some $a > 0$ and $\lambda = \Delta/(2b)$ for some positive integer b. Let $c(t)$ denote the continuation region $(l(t), u(t))$ and let T be the first time that $\mathscr{B}(t,\Delta) \notin c(t)$. Then

$$\begin{aligned} \Pr\{\mathscr{B}(T,\Delta) < 0\} &\leq \sum_{j=1}^{b} (-1)^{j+1} \left(1 - \frac{1}{2}\mathscr{I}(j=b)\right) \\ &\times \exp\{-2a\lambda(2b-j)j\}. \end{aligned}$$

Batur and Kim (2006) develop R&S procedures with parabolic boundaries. Their procedures are based on Ferebee (1982) which studies behaviors of $\sigma\mathscr{B}(t;\delta/\sigma)$ crossing a parabolic boundary.

3.4 Large Number of Alternatives

The number of alternatives of interest in simulation problems can be quite large, with up to 100 being relatively common. However, classical IZ procedures such as the R procedure were developed for relatively small numbers of alternatives, say no more than 20. They can be inefficient when the number of alternatives is large because they were developed to protect against the LFC—to free the procedure from dependence on the true differences among the means.

When the number of systems is large we rarely encounter anything remotely like the SC configuration, because large numbers of alternatives typically result from taking all feasible combinations of some controllable decision variables. Thus, the performance measures of the systems are likely to be spread out, rather than all clustered near the best. Fully-sequential procedures with elimination might seem to be a cure for this ill because they eliminate clearly inferior systems immediately if there is any evidence that they are. But the inequalities used to decompose the problem of k systems into paired comparisons with system k are typically quite conservative and become much more so with increasing k although Kim and Nelson's (2001) fully-sequential procedure KN, described in the next section, has been shown to work well for up to $k = 500$ systems (see Section 6).

To overcome the inefficiency of IZ approaches for large numbers of alternatives, one idea is to try to gain the benefits of screening, as in fully-sequential procedures, but avoid the conservatism required to compensate for so many looks at the data. Nelson, et al. (2001) proposed spending some of the α for incorrect selection on an initial screening stage (using a subset-selection procedure), and spending the remainder on a second ranking stage (using a Rinott-type IZ procedure). Additive and multiplicative α spending is possible, depending on the situation (see Nelson, et al. 2001 and Wilson 2001). The resulting procedure, named NSGS, is presented in the next section.

This so-called "α-spending" approach—spreading the probability of incorrect selection across multiple stages—is a general-purpose tool, and there is no inherent reason to use only a single split. See Jennison and Turnbull (2000) for a thorough discussion.

3.5 Steady-State Simulation

As we discussed in Section 3.1, procedures for i.i.d. normal data are applicable to steady-state simulation experiments if we make multiple replications of each alternative or use batch means of many raw outputs from a single long replication as basic observations.

Unfortunately, both of these solutions for dependent data have disadvantages. If we make multiple replications, then we have to discard raw outputs collected during the so-called warm-up period from each replication. This will be inefficient if the warm-up period is long and a large number of observations need to be discarded. Batching within a replication may also be inefficient. If a "stage" is defined by batch means rather than raw output, then the simulation effort consumed by a stage is a multiple of the batch size. When a large batch size is required to achieve approximate independence—and batch sizes of several thousand are common—then the selection procedure is forced to make decisions at long intervals, wasting outputs and time. This inefficiency becomes serious when fully-sequential procedures are employed because the elimination decisions for clearly inferior systems must wait for an entire batch to be formed. Therefore, for steady-state simulations, selection procedures that use individual raw outputs as basic observations are desirable.

The difficulty is proving that R&S procedures provide a correct-selection guarantee for a finite sample. Nevertheless, well designed procedures have shown good empirical performance. *Asymptotic analysis* can provide theoretical support for this observation. Asymptotic analysis typically means analysis as the simulation effort (run length, number of replications, or perhaps both) increases (conceptually) without bound. The power of asymptotic analysis is that many of the problem-specific details that thwart mathematical analysis in the finite-sample case wash out in the limit. Asymptotic analysis, done appropriately, can establish conditions under which we can expect procedures to work, rather than just relying on limited empirical evidence that they do; it can also establish the asymptotic superiority of one procedure over another.

Goldsman et al. (2001) extended the R procedure for use in steady-state simulation. Nakayama (1997) presented single-stage multiple-comparison procedures, and Damerdji and Nakayama (1999) developed two-stage multiple-comparison procedures, to provide inference on the best system for steady-state simulation. The extended R procedure is heuristic but the procedures in the latter two papers were shown to be asymptotically valid. The procedures in all three papers keep the number of sampling stages small (1 or 2), do not eliminate any alternatives until the final stage of sampling is completed, and thus are not appropriate when the number of systems is large.

Goldsman et al. (2001) and Kim and Nelson (2006a) proposed fully-sequential procedures—called KN+ and KN++—by extending the KN procedure to steady-state simulation. Their procedures are effective in eliminating inferior systems and thus more efficient than the procedures mentioned in the three papers above.

The key to extend procedures for i.i.d. normal data for use in steady-state simulation is to replace the usual sample variance S_i^2 with a variance estimator that works for stationary but dependent non-normal data. Basic observations X_{ij} are now raw observations from a single replication, the steady-state means are estimated by $\bar{X}_i(r)$ for some suitably large r, and we need a good estimator for the sample mean's variance. Rather than directly estimating the $\text{Var}[\bar{X}_i(r)]$, we can instead seek a good estimator of the *variance parameter* (or *asymptotic variance constant*), $v_i^2 \equiv \lim_{r\to\infty} r\text{Var}[\bar{X}_i(r)]$. Kim and Nelson (2006a) give requirements that variance estimators should satisfy for use in their procedures. Goldsman and Nelson (2006) give extensive review of variance estimators and their statistical properties.

4 EXAMPLE PROCEDURES

In this section we present two specific procedures to illustrate the concepts described in earlier sections. The NSGS procedure, due to Nelson, et al. (2001), and the KN procedure, due to Kim and Nelson (2001), are appropriate for terminating simulations or for steady-state simulations when multiple replications are employed.

The NSGS procedure requires that the output data from each system are i.i.d. normal, and that outputs across systems are independent, which leaves out CRN. The NSGS procedure is the combination of a subset-selection procedure, to reduce the number of alternatives still in play after the first stage of sampling, and a ranking procedure applied to the systems in the subset. The procedure uses α-spending between the subset selection and ranking to control the overall PCS.

Procedure NSGS

1. Specify the overall desired probability of correct selection $1 - \alpha$, the IZ parameter δ, a common initial sample size from each system $n_0 \geq 2$, and the initial number of competing systems k. Further, set

$$t = t_{n_0-1,1-(1-\alpha/2)^{\frac{1}{k-1}}}$$

and obtain Rinott's constant $h = h(n_0, k, 1 - \alpha/2)$ from the tables in Wilcox (1984) or Bechoffer et al. (1995). See also Table 8.3 in Goldsman and Nelson (1998).

2. Take n_0 outputs from each system. Calculate the first-stage sample means $\bar{X}_i(n_0)$ and marginal sample variances

$$S_i^2 = \frac{1}{n_0 - 1} \sum_{j=1}^{n_0} \left(X_{ij} - \bar{X}_i(n_0)\right)^2,$$

for $i = 1, 2, \ldots, k$.

3. *Subset Selection.* Calculate the quantity

$$W_{i\ell} = t \left(\frac{S_i^2 + S_\ell^2}{n_0}\right)^{1/2}$$

for all $i \neq \ell$. Form the screening subset I, containing every alternative i such that $1 \leq i \leq k$ and

$$\bar{X}_i(n_0) \geq \bar{X}_\ell(n_0) - (W_{i\ell} - \delta)^+ \quad \text{for all } \ell \neq i.$$

4. If $|I| = 1$, then stop and return the system in I as the best. Otherwise, for all $i \in I$, compute the second-stage sample sizes

$$N_i = \max \left\{n_0, \lceil (hS_i/\delta)^2 \rceil\right\},$$

where $\lceil \cdot \rceil$ is the ceiling function.

5. Take $N_i - n_0$ additional outputs from all systems $i \in I$.

6. Compute the overall sample means $\bar{X}_i(N_i)$ for all $i \in I$. Select the system with the largest $\bar{X}_i(N_i)$ as best.

Nelson et al. (2001) showed that any subset-selection procedure and any two-stage IZ ranking procedure that satisfy certain mild conditions can be combined in this way while guaranteeing the overall probability of correct selection. The NGSG procedure can handle a relatively large number of systems because the first-stage screening is pretty tight. Nelson et al. (2001) provide a revised version of the NGSG procedure, the Group-Screening procedure, in which one can avoid simulating all the systems simultaneously. Boesel et al. (2003) extended the Group-Screening procedure for "clean up" after optimization via simulation.

The KN procedure is *fully sequential*. Also, if there exists clear evidence that a system is inferior, then it will be eliminated from consideration immediately—unlike the NSGS procedure, where elimination occurs only after the first stage. The KN procedure also requires i.i.d. normal data, but does allow CRN. The KN procedure exploits the ideas of using paired differences, and controlling the Pr{ICS} on pairs to control it overall. Fabian's result is used to bound the error of a Brownian motion process that approximates each pair.

Procedure KN

1. *Setup.* Select confidence level $1 - \alpha$, IZ parameter δ and first stage sample size $n_0 \geq 2$. Set

$$\eta = \frac{1}{2} \left[\left(\frac{2\alpha}{k-1} \right)^{-2/(n_0-1)} - 1 \right].$$

2. *Initialization.* Let $I = \{1, 2, \ldots, k\}$ be the set of systems still in contention, and let $h^2 = 2\eta(n_0 - 1)$. Obtain n_0 outputs X_{ij} ($j = 1, 2, \ldots, n_0$) from each system i ($i = 1, 2, \ldots, k$) and let $\bar{X}_i(n_0) = n_0^{-1} \sum_{j=1}^{n_0} X_{ij}$ denote the sample mean of the first n_0 outputs from system i.
   For all $i \neq \ell$ compute

$$S_{i\ell}^2 = \frac{1}{n_0 - 1} \sum_{j=1}^{n_0} \left(X_{ij} - X_{\ell j} - [\bar{X}_i(n_0) - \bar{X}_\ell(n_0)] \right)^2,$$

   the sample variance of the difference between systems i and ℓ. Set $r = n_0$.

3. *Screening.* Set $I^{\text{old}} = I$. Let

$$I = \left\{ i : i \in I^{\text{old}} \text{ and} \right.$$
$$\left. \bar{X}_i(r) \geq \bar{X}_\ell(r) - W_{i\ell}(r), \forall \ell \in I^{\text{old}}, \ell \neq i \right\},$$

where

$$W_{i\ell}(r) = \max \left\{ 0, \frac{\delta}{2r} \left(\frac{h^2 S_{i\ell}^2}{\delta^2} - r \right) \right\}.$$

4. *Stopping Rule.* If $|I| = 1$, then stop and select the system whose index is in I as the best.
   Otherwise, take one additional output $X_{i,r+1}$ from each system $i \in I$, set $r = r + 1$ and go to *Screening*.

The KN procedure requires simulation of all systems simultaneously and a lot of switching among them. As discussed in Section 3.3, the switching cost can overwhelm the sampling cost, but this has become less of an issue in modern computing environments.

5 APPLICATION

This section illustrates the R, NSGS, and KN procedures using an (s, S) inventory system with the five inventory policies as described in Koenig and Law (1985). The goal of this study is to compare the five polices given in Table 1 and find the one with the smallest expected average cost per month for the first 30 months of operation. Table 1 also contains the expected cost (in thousands of dollars) of each policy, which can be analytically computed in this case. We set $\delta = \$1$ thousand, $n_0 = 10$ initial replications, and $1 - \alpha = 0.95$.

Table 2 shows the total number of outputs taken for each procedure. Procedure R needed 1556 observations. In the NSGS procedure, policies 3, 4, and 5 were eliminated after the first stage of sampling, so only policies 1 and 2 received second-stage samples. In the KN procedure, only policies 4 and 5 were eliminated after the first stage, but the elimination of policies 3 and 1 occurred after they received 16 and 98 observations, respectively. This illustrates the value of the tighter initial screen in the NSGS procedure, which takes only one look at the data, and the potential savings from taking many looks, as the KN procedure does. All three procedures chose policy 2 as the best (which is in fact correct). Since δ is smaller than the true difference, the R, NSGS, and KN procedures will choose the true best with 95% confidence. However, in general we do not have any information about the true differences; therefore, the best we can conclude without prior knowledge is that policy 2 is either the true best, or has expected cost per month within $1 thousand of the true best policy, with 95% confidence.

6 EXISTING CHALLENGES

Although these recent developments have brought a number of new R&S procedures that are useful in simulation environments, there still exist some challenges in the field.

Table 1: The Five Alternative Inventory Policies

Policy i	s	S	Expected Cost
1	20	40	114.176
2	20	80	112.742
3	40	60	130.550
4	40	100	130.699
5	60	100	147.382

Table 2: The Number of Outputs Taken from Each System of the (s, S) Inventory Policy Example

Policy i	R	NSGS	KN
1	209	209	98
2	349	349	98
3	492	10	16
4	378	10	10
5	129	10	10
Total	1556	588	232

In this section, we discuss the challenges with directions to approach them.

6.1 Improving Efficiency Further

Although new R&S procedures such as the NSGS and KN procedures have shown an impressive improvement on the number of observations used until termination, their performances in terms of actual PCS are still a bit conservative. Especially, when the number of system is larger than 20, actual PCS are often significantly larger than the nominal PCS. There are a number of reasons for this. We discuss two major reasons for inefficiency whose remedies are not known yet and then give three methods helpful in increasing efficiency.

The biggest factor that contributes to conservatism of IZ R&S procedures is the fact that many procedures are developed to protect PCS against *worst-case* scenarios known as the LFC. As the number of systems increases, this conservatism becomes more serious as the true mean configurations deviate from the LFC. Elimination by subset selection as in the NSGS procedure or fully-sequential approach as in the KN procedure helps reduce this conservatism, but higher actual PCS than the nominal value clearly implies that there is more room for efficiency improvement. Ideally, this problem would go away if we can replace δ in the design of procedures with $\max\{\delta, \mu_k - \mu_i\}$ for each system i. Unfortunately, this is impossible because mean performance measures are what we want to estimate by simulation. Instead, Chen and Kelton (2005) suggest that system i uses $\max\{\delta, \bar{X}_{[k]}(n_0) - \bar{X}_i(n_0)\}$ where $\bar{X}_{[k]}(n_0)$ denotes the largest first-stage sample mean among k systems. The performance is better in terms of observations compared to corresponding procedures with constant δ, but the procedures are heuristic and highly dependent on the quality of first-stage sample means.

Another factor that contributes to conservatism of R&S procedures is the use of the Bonferroni inequality. This is inherent to many R&S procedures, especially when the procedures need to account for dependence across systems due to the use of CRN. The Bonferroni inequality is an easy way to handle dependence across systems but provides a loose bound for PCS. Nelson and Matejcik (1995) avoided the Bonferroni inequality by employing the sphericity assumption, but the procedure is not recommended for use when k is large as the sphericity assumption tends to be violated for large k. Finding a way to avoid the Bonferroni inequality when k is large is currently an open research problem.

Now we discuss three methods that can possibly bring meaningful improvement in efficiency. Variance reduction techniques have been shown to be effective in simulation estimation problems and CRN is successfully exploited in simulation comparisons. Recently there have been efforts to introduce the control-variate (CV) technique to R&S procedures and use CV in combination with CRN. See Tsai and Nelson (2006). This technique has been adapted to subset selection and two-stage procedures so far.

Like other statistical procedures, the quality of variance estimators often affects the performance of R&S procedures and it is more so in steady-state simulation where a single replication is made and raw observations are taken as basic observations. We discussed in Section 3.5, such procedures are desirable but require estimators for the asymptotic variance constant v_i^2. Estimating v_i^2 is a difficult problem. A number of estimators have been proposed and there are still efforts to develop new variance estimators with better statistical properties such as smaller variance and smaller bias. Experimental results in Healey et al. (2007) show that variance estimators with smaller variance and smaller bias can result in meaningful (sometimes huge) savings in the number of observations for the KN+ and KN++ procedures.

Lastly, updating variance estimators helps make procedures more effective. Goldsman et al. (2001), Kim and Nelson (2006a), and Malone et al. (2005) showed that variance updating leads to huge savings in observations for procedures both for i.i.d. normal data and for steady-state simulation.

6.2 Optimization via Simulation

Most R&S procedures require the number of alternatives to be small enough to simulate all the alternative systems (a couple of hundred is considered small enough). When the search space is too large to simulate all the alternatives,

the problem requires different solution techniques such as heuristic algorithms or algorithms for optimization via simulation (OvS). Heuristic algorithms include tabu search, scatter search, integer programming, and neural networks; and algorithms for OvS with discrete decision variables include random search, stochastic ruler (Andradóttir 2006), nested partitions methods (Shi and Ólafsson 2000), and COMPASS (Hong and Nelson 2005a). Heuristic algorithms give no statistically meaningful estimates and provide no information about how close the chosen solution is to the true best solution. On the other hand, some algorithms such as random search algorithms for discrete OvS possess the global convergence property, but the convergence rate is often very slow. New developments in R&S open the possibility of improving discrete OvS.

R&S procedures can be used for neighborhood search in OvS to provide faster convergence to either local or global optimal and a probability of correct selection guarantee at algorithm termination. For example, Pichitlamken et al. (2006) developed an efficient selection procedure called Sequential Selection with Memory (SSM) for use in neighborhood search of OvS, and Pichitlamken and Nelson (2003) develop a global guidance system with the nested partitions method that provides an irreducible Markov chain over subsets of the feasible space. Then, they embed aggressive local-improvement schemes and enhance the local-improvement schemes with SSM.

Embedding R&S procedures to OvS raises a number of issues that need to be addressed before combining R&S with OvS. Boesel et al. (2003) and Pichitlamken et al. (2006) discuss why R&S procedures can not be directly used in the neighborhood search of OvS algorithms and that the extension of R&S procedures for use within an OvS algorithm is not straightforward. During the search step in OvS, solutions previous visited are likely to be visited again. When simulation is costly and partial or complete past information on alternatives previously visited is maintained, it is desirable to use the available past information. Most R&S procedures assume that none of the alternatives have already been sampled, and that they have all been sampled equally. Such R&S procedures are not applicable when alternatives retain past information (e.g., observations or sample means) because it results in different initial samples available across alternatives.

6.3 Multiple Performance Measures

R&S and OvS have focused on optimization problems with a single performance measure and deterministic constraints only. For example, consider an agile-worker allocation problem that is to find the optimal allocation policy of *M* flexible workers among *m* stations to maximize throughput. Then, the condition that the sum of the workers assigned at each station should be equal to *M* acts as a deterministic

constraint on decision variables, and it constrains the search space for the allocation problem. However, if a requirement on the expected number of waiting jobs at each station is imposed, then the problem becomes optimization with constraints on secondary performance measures as well as with respect to deterministic constraints. We call R&S with stochastic constraints *constrained R&S* and discrete optimization problem with stochastic constraints *stochastic constrained OvS*. Very little work has been done in handling stochastic constraints—although we often meet stochastic constrained optimization problems in practice.

Simply adding another requirement on a second performance measure changes the nature of the comparison problem a great deal. Butler et al. (2001) and Santner and Tamhane (1984) proposed two-stage selection procedures in consideration of two more more performance measures. But their two-stage procedures are either difficult to apply in practice or handle only a special case. Another problem with these procedures is that they become inefficient for 20 or more systems. Recently, Andradóttir et al. (2005) and Andradóttir and Kim (2007) have developed R&S procedures for a constrained R&S problem where a constraint on a secondary performance is considered. Batur and Kim (2005, 2007) present R&S procedures that identify a set of feasible systems for multiple constraints, but they do not solve the problem of finding the best feasible system in the presence of multiple stochastic constraints.

Therefore, for comparison problems with stochastic constraints, we need statistical procedures that provide an overall correct selection guarantee for both the primary and multiple secondary measures, yet are efficient and easy to implement so that they can eventually be embedded into OvS where the space of discrete decision variables is large.

7 CONCLUSION

Throughout this paper we have focused on the problem of finding the best when the best is defined as the system with the largest expected performance measure. As discussed in Section 1, there exist other types of comparison problems. For other types of R&S problems, see Kim and Nelson (2006b).

Instead of providing a PCS guarantee, Bayesian approaches attempt to allocate a finite data budget to maximize the posterior PCS of the selected system. Chick (2006) is a good reference.

REFERENCES

Andradóttir, S. 2006. An overview of simulation optimization via random search. In *Handbooks in Operations Research and Management Science: Simulation*, ed. S. G. Henderson and B. L. Nelson, Chapter 20, 617–631. Oxford, UK: Elsevier Science.

Andradóttir, S., D. Goldsman, and S.-H. Kim. 2005. Finding the best in the presence of a stochastic constraint. In *Proceedings of the 2005 Winter Simulation Conference*, ed. M.E. Kuhl, N.M. Steiger, F.B. Armstrong, and J.A. Joines, 732–738. Piscataway, NJ:IEEE.

Andradóttir, S., and S.-H. Kim. 2007. Fully sequential procedures for comparing constrained systems via simulation. Technical Report, The H. Milton Stewart School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, Georgia.

Banks, J., J. S. Carson, B. L. Nelson, and D. Nicol. 2001. *Discrete-Event System Simulation*. Upper Saddle River, NJ: Prentice Hall.

Batur, D. and S.-H. Kim. 2005. Procedures for feasibility detection in the presence of multiple constraints. In *Proceedings of the 2005 Winter Simulation Conference*, ed. M.E. Kuhl, N.M. Steiger, F.B. Armstrong, and J.A. Joines, 692–698. IEEE, Piscataway, NJ.

Batur, D., and S.-H. Kim. 2006. Fully sequential selection procedures with parabolic boundary. *IIE Transactions* 38:749–764.

Batur, D. and S.-H. Kim. 2007. Finding feasible systems in the presence of constraints on multiple performance measures. Technical Report, The H. Milton Stewart School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, Georgia.

Bechhofer, R. E., T. J. Santner, and D. Goldsman. 1995. *Design and Analysis of Experiments for Statistical Selection, Screening and Multiple Comparisons*. New York: John Wiley & Sons.

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, E.J., and W.D. Kelton. 2005. Sequential selection procedures: using sample means to improve efficiency. *European Journal of Operational Research* 166:133–153.

Chick, S. 2006. Subjective probability and Bayesian methodology. In *Handbooks in Operations Research and Management Science: Simulation*, ed. S. G. Henderson and B. L. Nelson, Chapter 9, 225–257. Oxford, UK: Elsevier Science.

Damerdji, H., and M. K. Nakayama. 1999. Two-stage multiple-comparison procedures for steady-state simulation. *ACM TOMACS* 9:1–30.

Fabian, V. 1974. Note on Anderson's sequential procedures with triangular boundary. *Annals of Statistics* 2:170–176.

Ferebee, B. 1982. Tests with parabolic boundary for the drift of a Wiener process. *Annals of Statistics* 10:882-894.

Glynn, P. W., and D. L. Iglehart. 1990. Simulation output analysis using standardized time series. *Mathematics of Operations Research* 15:1–16.

Goldsman, D., S.-H. Kim, W. Marshall, and B. L. Nelson. 2001. Ranking and selection procedures for steady-state simulation: Perspectives and procedures. *INFORMS Journals on Computing* 14:2–19.

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

Goldsman, D., and B. L. Nelson. 2006. Correlation-based methods for output analysis. In *Handbooks in Operations Research and Management Science: Simulation*, ed. S. G. Henderson and B. L. Nelson, Chapter 15, 455 – 475. Oxford, UK: Elsevier Science.

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.

Healey, C., D. Goldsman, and S.-H. Kim. 2007. Ranking and selection techniques with overlapping variance estimators. In *Proceedings of the 2006 Winter Simulation Conference*, ed. S. G. Henderson, B. Biller, M.-H. Hsieh, J. Shortle, J. D. Tew, and R. R. Barton, forthcoming. Pscataway, New Jersey:IEEE.

Hong, L. J., and B. L. Nelson. 2005a. Discrete optimization via simulation using COMPASS. *Operations Research* 54:115–129.

Hong, L. J., and B. L. Nelson. 2005b. The tradeoff between sampling and switching: New sequential procedures for indifference-zone selection. *IIE Transactions* 37:623–634.

Jennison, C., I. M. Johnstone, and B. W. Turnbull. 1980. Asymptotically optimal procedures for sequential adaptive selection of the best of several normal means. Technical Report, Dept. of ORIE, Cornell Univ., Ithaca, NY.

Jennison, C., and B. W. Turnbull. 2000. *Group Sequential Methods with Applications to Clinical Trials*. New York: Chapman & Hall.

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. 2006a. On the asymptotic validity of fully sequential selection procedures for

steady-state simulation. *Operations Research*, 54:475–488.

Kim, S.-H., and B. L. Nelson. 2006b. Selecting the best system. In *Handbooks in Operations Research and Management Science: Simulation*, ed. S. G. Henderson and B. L. Nelson, Chapter 17, 501 – 534. Oxford, UK: Elsevier Science.

Koenig, L. W., and A. M. Law. 1985. A procedure for selecting a subset of size m containing the ℓ best of k independent normal populations, with applications to simulation. *Communications in Statistics—Simulation and Computation* B14:719–734.

Law, A. M., and W. D. Kelton. 2000. *Simulation modeling and analysis*, 3d ed. New York: McGraw-Hill.

L'Ecuyer, P. 2006. Uniform random number generation. In *Handbooks in Operations Research and Management Science: Simulation*, ed. S. G. Henderson and B. L. Nelson, Chapter 3, 55–81. Oxford, UK: Elsevier Science.

Malone, G. J., S.-H. Kim, D. Goldsman, and D. Batur. 2005. Performance of variance updating ranking and selection procedures. In *Proceedings of the 2005 Winter Simulation Conference*, ed. M. E. Kuhl, N. M. Steiger, F. B. Armstrong, and J. A. Joines, 732–738. Piscataway, New Jersey: IEEE.

Nakayama, M. K. 1997. Multiple-comparison procedures for steady-state simulations. *The Annals of Statistics* 25:2433–2450.

Nelson, B. L., and F. J. Matejcik. 1995. Using common random numbers for indifference-zone selection and multiple comparisons in simulation. *Management Science* 41:1935–1945

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. 2003. A combined procedure for optimization via simulation. *ACM Transactions on Modeling and Computer Simulation* 13:155–179.

Pichitlamken, J., B. L. Nelson, and L. J. Hong. 2006. A sequential procedure for neighborhood selection-of-the-best in optimization via simulation. *European Journal of Operational Research* 173:283–298.

Rinott, Y. 1978. On two-stage selection procedures and related probability-inequalities. *Communications in Statistics—Theory and Methods* 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. New York: Marcel-Dekker.

Schmeiser, B. W. 1982. Batch size effects in the analysis of simulation output. *Operations Research* 30:556–568.

Shi, L., and S. Ólafsson. 2000. Nested partitions method for global optimization. *Operations Research*, 48:390–407.

Siegmund, D. 1985. *Sequential Analysis: Tests and Confidence Intervals*. New York: Springer-Verlag.

Sullivan, D. W., and J. R. Wilson. 1989. Restricted subset selection for simulation. *Operations Research* 37:52–71.

Tsai, S. C., and B. L. Nelson. 2006. Combined ranking and selection with control variates. In *Proceedings of the 2006 Winter Simulation Conference*, ed. L. F. Perrone, F. P. Wieland, J. Liu, B. G. Lawson, D. M. Nicol, and R.M. Fujimoto, 733–740. Pscataway, New Jersey:IEEE.

Wilcox, R. R. 1984. A table for Rinott's selection procedure. *Journal of Quality Technology* 16:97–100.

Wilson, J. R. 2001. A multiplicative decomposition property of the screening-and-selection procedures of Nelson et al. *Operations Research* 49:964–966.

AUTHOR BIOGRAPHIES

SEONG-HEE KIM is an Associate Professor in the H. Milton Stewart School of Industrial and Systems Engineering at the Georgia Institute of Technology. Her research interests include optimization via simulation, quality control, and simulation output analysis. She serves on several editorial boards including *IIE Transactions*, *Operations Research* and *The American Statistician*. Her e-mail and web addresses are ⟨skim@isye.gatech.edu⟩ and ⟨www.isye.gatech.edu/~skim/⟩, respectively.

BARRY L. NELSON is the Charles Deering McCormick Professor of Industrial Engineering and Management Sciences at Northwestern University. His research centers on the design and analysis of computer simulation experiments on models of stochastic systems. He has published numerous papers and two books. He has served the profession as the Editor in Chief of *Naval Research Logistics*, the Simulation Area Editor of *Operations Research*, and President of the INFORMS (then TIMS) College on Simulation. He has held many positions for the Winter Simulation Conference, including Program Chair in 1997 and current membership on the Board of Directors. His e-mail and web addresses are ⟨nelsonb@northwestern.edu⟩ and ⟨www.iems.northwestern.edu/~nelsonb/⟩.