

BEST-SUBSET SELECTION PROCEDURE

Yu Wang
Louis Luangkesorn
Larry J. Shuman

Department of Industrial Engineering
University of Pittsburgh
Pittsburgh, PA 15261, U.S.A

ABSTRACT

We propose an indifference-zone approach for a ranking and selection (R&S) problem with the goal of finding the best-subset from a finite number of competing simulated systems given a level of correct-selection probability. Here the “best” system refers to the system with the largest or smallest performance measures. We present a best-subset selection procedure that can effectively eliminate the non-competitive systems and return only those alternatives as the selection result where statistically confident conclusions hold. Numerical experiments document that our procedure works well by selecting the correct best-subset with very high probability.

1 INTRODUCTION

A common goal of optimization-via-simulation (OvS) studies is to choose the “best” system design from a set of competing alternatives, where the “best” refers to the system with the largest or smallest expected performance measure. This can be accomplished by comparing output from different alternative systems using the appropriate statistical methods. Ranking and Selection (R&S) procedures are specifically developed to select the best system or a subset of systems that contains the best from competing alternatives (Goldsman, Nelson, and Schmeiser 1991). Over the last decade, there have been fruitful efforts in developing statistically valid R&S procedures. In general, these procedures can be classified into two large categories: Frequentist procedures and Bayesian procedures.

Compared with Bayesian-based approaches (Chen et al. 2000, Chick and Inoue 2001), Frequentist procedures, such as Rinott (1978) and Kim and Nelson (2001), are relatively conservative, but they can provide statistical guarantees of correct-selection, which is preferred in many application cases. Therefore, the focus of this paper is on frequentist procedures, and the procedure proposed here does guarantee a pre-specified level of correct-selection probability.

Within R&S, two formulations of the problem of comparing alternative systems are indifference-zone formulations and subset selection formulations. Indifference-zone formulations provide a guarantee of selecting the single best system, where an indifference-zone parameter δ is defined at the range where the experimenter is “indifferent” to alternatives within δ of the best system. Subset-selection formulations choose a subset of the available alternatives so that there is a defined probability guaranteeing that the subset includes the best system.

2 LITERATURE REVIEW

2.1 Indifference-Zone Formulation

A large set of R&S procedures are based on indifference-zone approaches. they are characterized by two parameters, $\{\delta, P^*\}$, where δ is known as indifference-zone, which indicates a region in which the

experimenter would not discriminate between competing systems. The value P^* denotes the threshold of desired probability of correctly selecting the best alternative $\Pr\{\mathbf{CS}\}$, it is expected that $\Pr\{\mathbf{CS}\} \geq P^*$.

The original indifference-zone R&S procedure was proposed by Bechhofer (1954) as a single stage procedure. From a given $\{\delta, P^*\}$, the procedure can determine the number of required observations from each of competing systems. A major disadvantage of Bechhofer's procedure is its assumption for common, known variance across all systems, which is difficult to justify for simulations.

To address this issue, Dudewicz and Dalal (1975) presented a two-stage procedure (D-D), in which variances are estimated at the end of first stage and are used to calculate the number of observations required at the second stage. A weighted average of the first and second stage sample means is then used to select the best system. Rinott (1978) modified the D-D procedure to the R procedure, which yields a greater $\Pr\{\mathbf{CS}\}$ in some cases, but may require a larger total number of observations. For this reason, it is not appropriate to use the R procedure when the number of competing systems is large, especially when run time is an issue.

In order to handle the case involving a large number of alternatives, Nelson et al. (2001) presented the NSGS procedure, which uses the data from the first stage sampling to screen out alternatives that are not competitive, and thereby avoid the (typically much larger) second-stage sample for these systems.

The two-stage procedures with screening can be extended to more than two stages or to sequential-stage procedures, where a screening procedure is applied at each stage until only the best alternative is left, such as the fully-sequential procedures KN (Kim and Nelson 2001) and KN+/KN++ (Kim and Nelson 2006). These procedures are effective in eliminating inferior systems and thus more efficient than the R procedure.

In recent years, researchers put their attentions on improving the applicability and efficiency of Indifference-zone Selection (IZS) procedures. Hong and Nelson (2005) proposed sequential procedures (HN) that attempt to balance the cost of sampling and switching to minimize the total computational cost. Hong and Nelson (2007) also presented procedures that are capable of selecting the best alternative in the situations when the alternatives are revealed (generated) sequentially during the experiment. Osogami (2009) proposed a two-stage indifference-zone approach (TSSD) with the goal of reducing both the number of simulated samples of the performance and the frequency of configuration changes. Tsai and Nelson (2010) applied the Control Variates (CV) technique in fully-sequential indifference-zone selections to develop a more efficient R&S procedure.

As a common characteristic, almost all existing procedures (KN/KN+/KN++, HN, TSSD) are designed to select only a single system (which will then be claimed as the best) whose mean performance measure (μ_b) is within an indifference-zone (δ) to the true-best system's mean (μ_B , unknown). All others will be screened out in the early stage or be disregarded in the final stage (due to exceeding the computation budget limitation).

2.2 Subset Selection Formulation

The other major type of R&S is Subset Selection, which is first presented by Gupta (1965). The goal of subset selection procedures is to identify a subset of random size that contains the best system, with user-specified probability P^* and without the specification of an indifference-zone (i.e., $\delta = 0$). Like Bechhofer's indifference-zone procedures, Gupta's subset selection procedure requires equal and known variances among competing alternatives. To solve this issue, Sullivan and Wilson (1989) developed two subset selection procedures which extends Gupta's work by allowing unknown and unequal variance, and specification of a non-zero indifference-zone.

Compared with indifference-zone procedures, subset selections are less popular in practice. Existing subset selection procedures provide no guarantee on the performance of all alternative systems among the selected subset, they only claim that their selection subset contains the best but do not claim that the whole subset selected is "good". This lack of a guarantee makes it difficult to use the results from subset-selection procedures in practice.

2.3 Discussion

However, in practice, decision makers want a method that examines alternatives and provides them with a subset of alternatives that are close to the best, so that they can choose the final decision from the “best-subset”, instead of unconditionally trusting the best solution provided by a computer. Such a requirement is not only due to the fact that people are unwilling to leave their decision making responsibilities to computers, but also has its practical reason - it is usually neither possible nor necessary to include all constraints into a computer model when evaluating system designs by simulation. Further, some of constraints may not be negligible when performing practical implementations. Indeed, the simulation model is only an abstraction of the real system, but not a full replication. Consequently, the best system selected by computer for the abstraction may not be best for the real system, and could even be infeasible or simply unrealistic in practice. Hence, we suggest that a best-subset involving multiple potential alternatives should be more useful than the only one choice, as it allows the decision maker to choose among a set of alternatives based on criteria not in the model, such as social or political feasibility.

This work is motivated by a practical policy selection problem from a local government agency. We developed an agent-based simulation model of emergency response and are using the model to examine a set of alternative emergency response policies. The ideal deliverable for the customer is a methodology that can select a subset of alternative policies which all demonstrate good performance in the simulation model, so the customer can choose among them using other factors to select the policy to implement in practice.

However, our review of the literature suggests that there has not yet a solution for the problem selecting statistically guaranteed best-subset. The results from indifference-zone methods that select only a single best alternative were not acceptable (because of skepticism of the output of computer models) and the results of subset selection methods were not useful (because of the lack of guarantees on the performance of all systems included in the subset). Therefore, we developed this methodology to select the best-subset from a finite number of competing alternatives while guaranteeing a pre-specified correct-selection probability level.

For relevant research, Kim (2005) developed a fully sequential procedure for comparison systems with a standard. Andradóttir, Goldsman, and Kim (2005), Andradóttir and Kim (2010), Batur and Kim (2005) and Batur and Kim (2010) considered the problem of finding a set of feasible or near-feasible systems among a finite number of simulated systems in the presence of stochastic constraints. However, none of these research efforts addresses the best-subset selection problem. The problem is difficult because the actual best system is unknown, which makes it hard to recognize and eliminate the inferior systems during the screening stage as well as to establish appropriate stopping criteria.

The contribution of this paper is the development of a fully sequential R&S procedure to select the best-subset while satisfying the pre-specified correct-selection probability requirement. According to Osogami (2009), the KN series are still the most efficient algorithms in terms of the number of samples needed. So we extend KN in order to develop this new procedure that selects the best-subset by efficiently screening out obviously inferior alternatives in the early stages.

The rest of the article is organized as follows. The mathematical formulation of the problem is provided in Section 3. In Section 4, we introduce a fully sequential procedure named as best-subset selection procedure. Section 5 provides a numerical example which illustrates the effectiveness of the procedure. And the paper ends with conclusions and possible future research directions in Section 6.

3 PROBLEM FORMULATION

In this section, we formulate the best subset-selection problem and define notations. Assuming that there are in total $k \geq 2$ competing simulation systems, let X_{ij} be a univariate real-valued output data from replication (or batch) j of system i , and the performance measures of different systems are defined as

$x_i = \mathbf{E}[X_{ij}]$ ($i = 1, \dots, k$). We will assume that a larger mean is better, and we let $x_{[1]} \geq x_{[2]} \geq \dots \geq x_{[k]}$, where system [1] is the best system (unknown to us).

Our problem is to find the best-subset I , here we refer the best system as the system with the largest mean, and any system whose performance measure is within λ -distance to the best will be considered as item belonging to the best-subset, so we can define I as

$$I = \left\{ i : x_i \in \left[\max_{i=1, \dots, k} x_i - \lambda, \max_{i=1, \dots, k} x_i \right] \right\}$$

For solving this problem, we assume that $X_{ij} \sim \mathbf{N}(x_i, \sigma_i^2)$ ($i = 1, 2, \dots, k$), that is, X_{ij} 's are distributed as normal distributions with mean of x_i and variance of σ_i^2 .

It should be mentioned that this assumption is quite mild since it requires neither common variance nor independent sampling (which implies that our selection procedure allows using common random numbers (CRN) to increase the precision when comparing two or more alternative configurations by simulation). Besides that, the assumption of normality is generally plausible when the basic observations of system performance are either within-replication averages (from a transient or steady-state simulation) or batch means with a large batch size (from a steady-state simulation). The non-normality of basic observations may be a problem. Kim and Nelson (2001) show that fully sequential R&S procedures tend to be robust to non-normality. Also, any non-normality can be mitigated by using batches of non-normal data as basic observations (as in Kim and Nelson (2001)). In future work, we may test the robustness of our procedures to non-normal data, both with and without batching.

For stochastic systems, it is not always possible to guarantee that we select the best-subset which contains all systems that within λ -distance to the best. Instead, we apply the idea of indifference-zone again to the λ -boundary to find a set of best systems. Specifically, we adopt a similar approach that was used in Andradóttir, Goldsman, and Kim (2005) and Andradóttir and Kim (2010) by asking a decision maker to specify a range $[\lambda^-, \lambda^+]$ around the λ -boundary such that $\lambda^+ > \lambda^-$. Then three regions can be defined:

- $x_B - x_i \leq \lambda^-$: This is the definitely best region (S_B). Any system in this range should be retained in the final best-subset. We call any system inside this range as “desired”.
- $\lambda^- < x_B - x_i \leq \lambda^+$: This is the transition region (S_T). For any system inside this region, it does not matter whether it is selected into the final best-subset or not. In other words, it is all right to exclude a system in this region from the best-subset (even if $x_B - x_i \leq \lambda$), or to accept a system into the best-subset as long as the system i is within this region. Therefore we call any system inside this region as an “acceptable” system.
- $\lambda^+ < x_B - x_i$: This is the elimination region (S_E). Any system in this region should be screened out in the screen phase and should not be contained in the final best-subset. We use the term “undesired” to refer to any system in this region.

For convenience, we chose the parameters λ^- and λ^+ so that $\lambda = (\lambda^- + \lambda^+)/2$, and define $\varepsilon = (\lambda^+ - \lambda^-)/2$. Essentially, λ is a target value that behaves as a cutoff point between desired and undesired systems, and ε is the level of precision to the specification of λ , that is, ε specifies how much we are willing to be off from λ . In other words, ε defines an indifference-zone around λ and plays a similar role as the δ defined in Kim and Nelson (2001) and Kim and Nelson (2006).

4 BEST SUBSET SELECTION PROCEDURE

In this section, we present a procedure that eliminates all undesired systems and to return a resultant set that contains all the desired systems plus, possibly, some acceptable systems. This method is also described

more completely along with proof in Wang et al. (2011).

Setup: Select the overall desired probability of correct selection (PCS) (confident level) $P^* = 1 - \alpha$ ($0 < P^* < 1$), boundary parameter $\lambda > 0$ and common first-stage sample size $n_0 \geq 2$. Choose a small value for the indifference-zone parameter ε ($\varepsilon > 0$), which indicates the half-width of transition region. Calculate η as described below:

$$g(\eta) \equiv \sum_{\ell=1}^c (-1)^{\ell+1} \left(1 - \frac{1}{2} \mathcal{I}(\ell = c)\right) \left(1 + \frac{2\eta(2c - \ell)\ell}{c}\right)^{-(n_0-1)/2} = \frac{2\alpha}{k(k-1)}$$

where \mathcal{I} is the indicator function. In the special case that $c = 1$, we have the closed-form solution

$$\eta = \frac{1}{2} \left[\left(\frac{4\alpha}{k(k-1)} \right)^{-2/(n_0-1)} - 1 \right]$$

Initialization: Let $I = \{1, 2, \dots, k\}$ be the set of systems in contention, N be the set of best systems, and $h^2 = 2c\eta(n_0 - 1)$.

Obtain n_0 outputs X_{ij} ($j = 1, 2, \dots, n_0$) from each system i ($i = 1, 2, \dots, k$) and let $\bar{X}_i(n_0) = (\sum_{j=1}^{n_0} X_{ij})/n_0$ denote the sample mean of the first n_0 outputs from system i .

For all $i \neq \ell$ calculate the sample variance of the difference between system i and ℓ .

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

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

Screening: Define

$$\bar{Y}_{i\ell}(r) = \bar{X}_i(r) - \bar{X}_\ell(r)$$

For each system combination (i, ℓ) (where $i \in I, \ell \in I, i \neq \ell$), if $\bar{Y}_{i\ell}(r) - \lambda \geq +R_{i\ell}(r)$ ($i \in I, \ell \in I$), then eliminate system ℓ from set I , where

$$R_{i\ell}(r) = \max \left\{ 0, \frac{\varepsilon}{2cr} \left(\frac{h^2 S_{i\ell}^2}{\varepsilon^2} - r \right) \right\}$$

Stopping Rule: If $\forall i \in I$ and $\forall \ell \in I, \bar{Y}_{i\ell}(r) - \lambda \leq -R_{i\ell}(r)$, then stop and return I as the best subset N . Otherwise, take one additional output $X_{i,r+1}$ from each system $i \in I$, set $r = r + 1$ and go to **Screening**.

From the experimental result shown in the following section, we can observe that the procedure returns subset N containing all desired systems, plus some acceptable ones, and without any undesired systems, with very high probability (probability of correct selection - PCS). A proof of the result (Theorem 1) can be found in Wang et al. (2011)

Theorem 1 Assume that $X_{ij}, i = 1, 2, \dots, k$ are normally distributed, the best subset selection procedure guarantees

$$\Pr\{\mathbf{CS}\} = \Pr\{S_B \subseteq I \subseteq (S_B \cup S_T)\} \geq 1 - \alpha$$

5 EXPERIMENTAL RESULTS

In this section we illustrate the performance of the best-subset selection procedure by two numerical experiments.

The first experiment is to demonstrate that the best-subset selection procedure can correctly select the best-subset with required probability. In the setup stage, we choose the confidence level as $P^* = 1 - \alpha = 0.95$

and take $n_0 = 10$ samples from each alternative. The boundary/indifference-zone parameters are chosen as $(\lambda^-, \lambda^+) = (5, 5.5)$ so that the transition region width $2\epsilon = \lambda^+ - \lambda^- = 0.5$, which means, any system whose mean value is within 5 unit distance of the true best system's should be selected for inclusion ("desired"), and it is also acceptable for the final selection set to include any system whose mean differs from the true best greater than 5 but less than 5.5. For example, suppose the best system's mean is 100, then all systems whose means are greater or equal to 95 should be selected, and any system whose mean is greater or equal to 94.5 but less than 95 could be selected, but all others (mean less than 94.5) should be excluded from the final selection set.

We use a set of normal random variables X_{ij} with different means and variances to represent the competitive systems ($X_{ij} \sim \mathbf{N}(x_i, \sigma_i^2)$). The number of total alternative systems $k = 3$. The mean and variance values are shown in Table 1. Such a configuration is known as the slippage configuration (SC) since system 2, 3 are just at the edge of the desired boundary. According to the region definitions (see Section 3) and parameters setting for this case, we know that both system 2 and 3 should be included in the final set and the selected best-subset should contains all alternatives $\{1, 2, 3\}$, but it is difficult to make the correct selection.

Table 1: Mean and variance configurations for the first experiment (SC).

No.	1	2	3
x_i	100	95	95
σ_i^2	1.0	2.0	4.0

Table 2: Replication statistics of the first experiment (SC).

Mean (Number of Samples)	Standard Deviation (Number of Samples)	Pr{CS}
2803.42	1672.78	0.9546

We run 10,000 independent replications. The percentage of correct selection (Pr{CS}) and the number of samples required in the replications are summarized in Table 2. From Table 2 we can see that the percentage of correct selection $\text{Pr}\{\mathbf{CS}\} = 0.9546$ which satisfies the requirement of $\mathbf{PCS} \geq 0.95$ very well.

Now let us consider a more complex scenario in the second experiment, where we have $k = 16$ competitive systems ($X_{ij} \sim \mathbf{N}(x_i, \sigma_i^2)$). The mean and variance configurations are shown in Table 3. All other parameters are the same as in the first experiment ($P^* = 1 - \alpha = 0.95$, $(\lambda^-, \lambda^+) = (5, 5.5)$, $2\epsilon = \lambda^+ - \lambda^- = 0.5$). Here we only test configurations with a common variance for all configurations ($\sigma_i^2 = 1.0$). Wang et al. (2011) details the performance of the best-subset selection procedure where the alternative systems have a range of different variances.

Table 3: Mean and variance configurations for the second experiment.

No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
x_i	100	95.3	95.2	95.1	94.95	94.9	94.85	94.8	94.75	94.7	94.65	94.6	94.55	94.4	94.3	94.2
σ_i^2	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0

For this experiment, we run 30 batches where each batch consists of 500 replications (15,000 in total, independent replications). The benefit of using batches is to enable plotting box-plot of correct selection frequency for each alternative system, as shown in the Figure 1. The number of samples required in the replications is summarized in Table 4.

From Figure 1 we can see that our best-subset procedure works well. With a very high probability, all desired systems ($x_i \geq 95$) were selected, and no undesired systems were kept in the final selection. For the systems within the transition region, it is obvious that the frequency-of-being-selected drops very quickly as the distance to the best system increasing, which means those non-competitive system can be effectively eliminated during the screening procedures. The resultant percentage of correct selection $\text{Pr}\{\mathbf{CS}\} = 0.99973$ is very high, which may suggest certain conservation existing when handling a large set of alternatives.

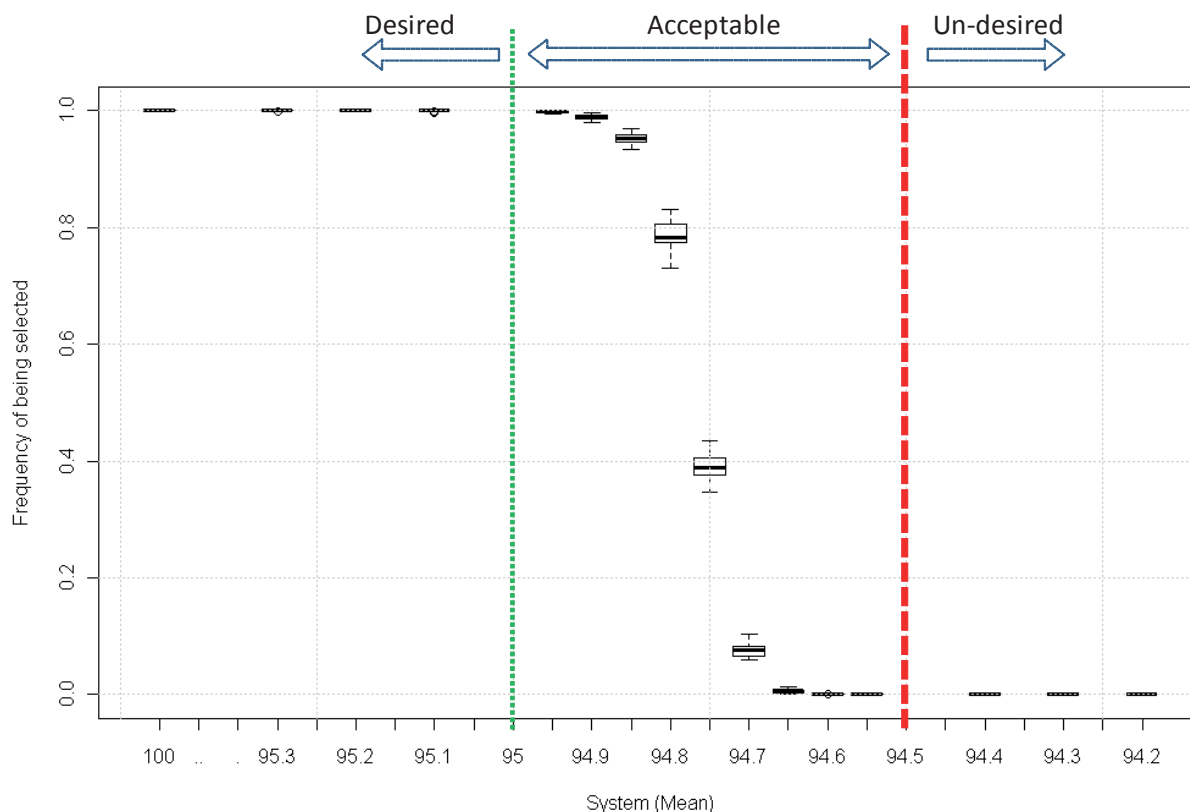


Figure 1: Numerical results of the best-subset selection procedure.

Table 4: Replication statistics of the second experiment.

Mean(Number of Samples)	Standard Deviation(Number of Samples)	Pr{CS}
13665.73	4650.98	0.99973

The possible source of conservatism may come from the application of Bonferroni inequality to control the overall incorrect selection probability by combining pairwise-comparison results in together (i.e. the $\frac{2\alpha}{k(k-1)}$ term when calculating η), and possible techniques to reduce this conservatism will be studied in future work.

6 CONCLUSION

In this paper, we consider a new R&S problem which requires choosing the best-subset from a set of alternative systems. We extended the existing KN procedure to a new best-subset selection procedure to solve this problem and demonstrated its ability to select all systems that are close enough to the best system so that the decision maker is indifferent to the difference.

As this work was motivated by the problem of policy selection via an emergency response simulation model, we will apply this method to the model to examine the effects of a range of alternative emergency response policies. We believe that this procedure can provide results that are more useful to policy makers than the current selection of the best system or subset selection procedures.

For future work, we are continuing research into how to improve the efficiency of the procedure, in particular in the case where there is a large set of alternatives, while maintaining the desired probability of

correct selection. We will also will investigate the effects of variance reduction techniques such as common random numbers (CRN) and test the procedure's robustness to non-normality.

ACKNOWLEDGMENTS

The authors thank Dr. Barry L. Nelson for early discussions on the nature of the best-subset selection problem and potential approaches. This work has been funded by PEMA (Pennsylvania Emergency Management Agency) support to the University of Pittsburgh Center for National Preparedness under grant number (PO# 4300212171).

REFERENCES

- Andradóttir, S., D. Goldsman, and S.-H. Kim. 2005, December. "Finding the Best in the Presence of a Stochastic Constraint". In *Proceedings of the 2005 Winter Simulation Conference*, edited by M. E. Kuhl, N. M. Steiger, F. B. Armstrong, and J. A. Joines, 732–738. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.
- Andradóttir, S., and S.-H. Kim. 2010. "Fully Sequential Procedures for Comparing Constrained Systems via Simulation". *Naval Research Logistics* 57 (5): 403–421.
- Batur, D., and S.-H. Kim. 2005, December. "Procedures for Feasibility Detection in the Presence of Multiple Constraints". In *Proceedings of the 2005 Winter Simulation Conference*, edited by M. E. Kuhl, N. M. Steiger, F. B. Armstrong, and J. A. Joines, 692–698. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.
- Batur, D., and S.-H. Kim. 2010. "Finding Feasible Systems in the Presence of Constraints on Multiple Performance Measures". *ACM Transactions on Modeling and Computer Simulation* 20 (3): 1–26.
- Bechhofer, R. E. 1954. "A Single-Sample Multiple Decision Procedure for Ranking Means of Normal Populations with Known Variances". *The Annals of Mathematical Statistics* 25 (1): 16–39.
- Chen, C.-H., J. Lin, E. Yücesan, and S. E. Chick. 2000. "Simulation Budget Allocation for Further Enhancing the Efficiency of Ordinal Optimization". *Discrete Event Dynamic Systems* 10 (3): 251–270.
- Chick, S. E., and K. Inoue. 2001. "New Two-Stage and Sequential Procedures for Selecting the Best Simulated System". *Operations Research* 49 (5): 732–743.
- Dudewicz, E. J., and S. R. Dalal. 1975. "Allocation of Observations in Ranking and Selection with Unequal Variances". *Sankhyā: The Indian Journal of Statistics, Series B* 37 (1): 28–78.
- Goldsman, D., B. L. Nelson, and B. Schmeiser. 1991, December. "Methods for Selecting the Best System". In *Proceedings of the 1991 Winter Simulation Conference*, edited by B. L. Nelson, W. D. Kelton, and G. M. Clark, 177–186. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.
- Gupta, S. S. 1965. "On Some Multiple Decision (Selection and Ranking) Rules". *Technometrics* 7 (2): 225–245.
- Hong, L. J., and B. L. Nelson. 2005. "The Tradeoff between Sampling and Switching: New Sequential Procedures for Indifference-zone Selection". *IIE Transactions* 37 (7): 623–634.
- Hong, L. J., and B. L. Nelson. 2007. "Selecting the Best System When Systems Are Revealed Sequentially". *IIE Transactions* 39 (7): 723–734.
- Kim, S.-H. 2005. "Comparison with a Standard via Fully Sequential Procedures". *ACM Transactions on Modeling and Computer Simulation* 15 (2): 155–174.
- Kim, S.-H., and B. L. Nelson. 2001. "A Fully Sequential Procedure for Indifference-Zone Selection in Simulation". *ACM Transactions on Modeling and Computer Simulation* 11 (3): 251–273.
- Kim, S.-H., and B. L. Nelson. 2006. "On the Asymptotic Validity of Fully Sequential Selection Procedures for Steady-State Simulation". *Operations Research* 54 (3): 475–488.
- Nelson, B. L., J. Swann, D. Goldsman, and W. Song. 2001. "Simple Procedures for Selecting the Best Simulated System When the Number of Alternatives Is Large". *Operations Research* 49 (6): 950–963.

- Osogami, T. 2009. "Finding Probably Best Systems Quickly via Simulations". *ACM Transactions on Modeling and Computer Simulation* 19 (3): 1–19.
- Rinott, Y. 1978. "On Two-Stage Selection Procedures and Related Probability-Inequalities". *Communications in Statistics - Theory and Methods* 7 (8): 799–811.
- Sullivan, D. W., and J. R. Wilson. 1989. "Restricted Subset Selection Procedures for Simulation". *Operations Research* 37 (1): 52–71.
- Tsai, S. C., and B. L. Nelson. 2010. "Fully Sequential Selection Procedures with Control Variates". *IIE Transactions* 42 (1): 71–82.
- Wang, Y., L. Luangkesorn, and L. Shuman. 2011. "A Fully Sequential Procedure for Best-Subset Selection via Simulation". Working paper, Dept of Industrial Engineering, University of Pittsburgh, Pittsburgh, PA, USA.

AUTHOR BIOGRAPHIES

YU WANG is currently a Ph.D. candidate in the Department of Industrial Engineering, University of Pittsburgh, Pittsburgh, PA. He earned a B.S. and M.S. in Mechanical Engineering both from Tsinghua University, Beijing, China, and an M.S. in Industrial Engineering from University of Pittsburgh, Pittsburgh, PA. His research interests include large-scale complex systems modeling, simulation and optimization. He is a member of INFORMS. This paper is a part of his research work supported by the Center for National Preparedness, University of Pittsburgh. His e-mail address is yuw9@pitt.edu and his web page is <http://www.pitt.edu/~yuw9/>

LOUIS LUANGKESORN is a Research Assistant Professor in the Department of Industrial Engineering at the University of Pittsburgh. He received a B.S. in General Engineering and a B.A. in Political Science from the University of Illinois-Urbana, an M.A. in Science, Technology and Public Policy from The George Washington University, and a Ph.D. in Industrial Engineering and the Management Sciences from Northwestern University. His research interests include logistics and resource management in health and emergency response settings, as well as the use of simulation and models for policy analysis and evaluation. His e-mail address is loll1@pitt.edu

LARRY J. SHUMAN is Senior Associate Dean for Academics and Professor of Industrial Engineering at the University of Pittsburgh. His research focuses on improving the engineering educational experience with an emphasis on assessment of design and problem solving, the study of the ethical behavior of engineers and engineering managers; he is also studying responses to natural and man-made disasters. A former senior editor of the *Journal of Engineering Education*, Dr. Shuman is the founding editor of *Advances in Engineering Education*. He received his Ph.D. from The Johns Hopkins University in Operations Research and the BSEE from the University of Cincinnati.