Journal of Industrial and Systems Engineering
Vol. 2, No. 4, pp 259-270
Winter 2009

JISE

Hybrid Probabilistic Search Methods for Simulation Optimization

Alireza Kabirian

Department of Industrial and Manufacturing Systems Engineering, Iowa State University, Black Engineering, Ames, IA 50011, USA

ABSTRACT

Discrete-event simulation based optimization is the process of finding the optimum design of a stochastic system when the performance measure(s) could only be estimated via simulation. Randomness in simulation outputs often challenges the correct selection of the optimum. We propose an algorithm that merges Ranking and Selection procedures with a large class of random search methods for continuous simulation optimization problems. Under a mild assumption, we prove the convergence of the algorithm in probability to a global optimum. The new algorithm addresses the noise in simulation outputs while benefits the proven efficiency of random search methods.

***Keywords*:** Simulation Optimization, Random Search, Ranking and Selection, Asymptotic Convergence.

1. INTRODUCTION

Simulation Optimization (SO) is the process of selecting the best system among a set of possible options when the performance measure(s) (objective function) and/or constraint(s) are estimated through simulation. SO problems have recently received considerable attention by many researches in simulation and optimization communities. However, application of SO methods is not limited to the cases where discrete event simulation is required. Indeed, wherever a costly experiment in terms of computation, danger, etc. evaluates the performance in an optimization problem, there is a potential need for SO methods (e.g. see Kabirian and Hemmati 2007 for an application in an energy model via fluids simulation).

In contrast to deterministic problems, SO problems in general have two distinguishing properties. First, the closed form of the objective function is not known; this property rules out direct application of powerful arsenal of the mathematical programming approaches. Secondly, since the values of the objective function could only be estimated through simulation, it involves noise that can mislead the optimization process.

Over the years, many approaches have been proposed in the literature for SO problems. Detailed review of available methods can be found in Andradóttir (1998, 2006), Olafsson and Kim (2002), Gosavi (2003) Olafsson (2006) and Fu et. al. (2005). When the number of design options is limited, statistical selection methods are usually appropriate (Kim and Nelson 2006). Metamodel based methods such as Response Surface Methodology fit regression or neural networks on the response surface locally or globally and use it to conduct new searches (Barton and Meckesheimer 2006). A major class of SO approaches is random search methods such as Stochastic Ruler method (Yan and

Mukai 1992) and Stochastic Comparison algorithms (Gong et. al. 1999). Gradient search methods such as Stochastic Approximation (Robins and Morono 1951 and Kiefer and Wolfowitz 1952) estimate the gradient of the objective function (Fu 2006) and then use gradient methods of mathematical programming. Metaheuristic methods such as Simulated Annealing (Krikpatrick et. al. 1983), Genetic Algorithm (Holland 1992), Tabu Search (Glover 1989, 1990) and Scatter Search (Glover 1997) are a rich set of deterministic optimization algorithms which are extensively used in practical SO problems and commercial software packages (Olafsson 2006).

Designing a new SO algorithm or selecting an SO method among available routines involve a tradeoff between two conflicting criteria: efficiency vs. convergence (Andradóttir 2006). A convergent method can guarantee that the randomness in simulation responses does not mislead the optimization process and global optimum could eventually be reached. This nice property, however, comes at a cost because rigorous mathematical proof of convergence requires periodic simulation of points that have already been simulated (Olafsson 2006) or are located in close proximity of old simulated points. These unwelcome consequences decrease the efficiency of convergent SO methods (Kabirian 2006, Kabirian and Olafsson 2007a and 2007b). Available methods in the literature usually have tendency to one of the two criteria. For instance, many heuristic methods (including metaheuristics) applied to SO problems are among the efficient class that have been implemented in commercial SO software packages (Fu et. al. 2005); in contrast, methods such as Sample Path Optimization and its variants provably have very nice convergence properties (Robinson 1996, Shapiro 1996), though criticized to be inefficient (Deng and Ferris 2007, Azadivar 1999).

In this paper, we bridge the gap between practical appeal of a large class of heuristic methods called "Probabilistic Search" techniques and theoretical convergence guarantee by proposing a methodology to link these methods with Ranking and Selection (R&S). The remainder of the paper is organized as follows. We formally define Probabilistic Search algorithms in section 2. In section 3, R&S methods are briefly introduced and a methodology for merging R&S with Probabilistic Search algorithms is presented. Section 4 contains numerical results. Finally, section 5 concludes the paper.

2. PROBABILISTIC SEARCH METHODS

Consider the deterministic continuous optimization problem below:

$$\min_{\theta \in \Theta} f(\theta) \tag{1}$$

where $f : \Theta \to \mathrm{R}$ and $\Theta \subset \mathrm{R}^n$ is the compact set of feasible points. Let $\theta^* \in \Theta$ be one of the possibly many global optimums of (1).

Random Search methods are a class of iterative optimization algorithms in whose kth iteration, a finite number of points $\theta_k^{(1)},...,\theta_k^{(H_k)} \in \Theta$ are selected via a specific sampling strategy Ψ_k and evaluated (see Andradottir 2006 for SO version).

Algorithm 1: Random Search Methods

Step 0: (Initialize). Choose the initial sampling strategy Ψ_1 *and let* $k = 1$.

Step 1: (Sample). Select points $\theta_k^{(1)},...,\theta_k^{(H_k)} \in \Theta$ *according to sampling strategy* Ψ_k

Step 2: (Evaluate). Compute $f(\theta_k^{(j)})$ *for* $j = 1,...,H_k$.

Step 3: (Update). Considering the quality of the evaluated points thus far, pick and introduce current optimum of the search process. If termination condition(s) of the algorithm hold(s), stop the algorithm, otherwise choose an updated strategy Ψ_{k+1}, *let* $k = k+1$ *and go to step 1.*

This class is broad enough to include many heuristic and metaheuristic methods. We narrow the definition a bit to what we call Probabilistic Search (PS) methods as follows:

Definition 1: *Let* W *be a Lebesgue measure. The probabilistic search methods are a subclass of random search methods for problem (1) such that for any arbitrary subset* $G \subseteq \Theta$ *with* $W(G) > 0$, *the followings hold:*

1) $\Pr\left\{\bigcup_{j=1}^{H_k}\left(\theta_k^{(j)} \in G\right)\right\} > 0 \quad \forall k = 1,2,...$

2) $\lim_{k\to\infty} \Pr\left\{\bigcup_{j=1}^{H_k}\left(\theta_k^{(j)} \in G\right)\right\} \neq 0$

Many Random Search based methods satisfy the conditions of Definition 1. Others can also be a PS method if, in addition to their core optimization process, a point is also selected randomly from Θ and evaluated in each iteration.

Now, we prove that every PS algorithm almost surely visits a point with objective function within an arbitrary neighborhood of a global optimum if a regularity condition is met.

Definition 2: *Call the interval* $O(\varepsilon) = [f(\theta^*), f(\theta^*) + \varepsilon]$ *where* $\varepsilon > 0$, *the* ε *-optimum interval.*

Definition 3: *Let* $B(\theta; r) = \left\{x \in \mathrm{R}^n \middle| \quad |\theta - x| \le r\right\}$ *be a ball of radius* $r > 0$ *centered at the point* θ.

Definition 4: *The discrete random variable* $X \sim NHG\left(\{p_k\}_{k=1,2,...}\right)$ *has non-homogenous geometric (NHG) distribution with success probability* p_k *in* k *th trial if*

$$pr(X = x) = \begin{cases} p_x \prod_{k=1}^{x-1}(1-p_k) & \forall x = 1,2,... \\ 0 & o.w. \end{cases} \tag{2}$$

Assumption 1: *A ball* $B(\theta^*;r)$ *exists such that* $W(B(\theta^*;r)\cap\Theta)>0$ *and* $f(\theta)$ *is continuous for all* $\theta\in B(\theta^*;r)\cap\Theta$.

Lemma 1: *For problem (1) under Assumption 1, there exists another ball* $B_\varepsilon(\theta^*;r')$ *for any given* $\varepsilon>0$ *such that:*

1) $B_\varepsilon(\theta^*;r')\subseteq B(\theta^*;r)$
2) $f(\theta)\in O(\varepsilon)$ *for all* $\theta\in B_\varepsilon(\theta^*;r')\cap\Theta$
3) $W(B_\varepsilon(\theta^*;r')\cap\Theta)>0$

Proof: The proof is a direct result of the continuity of f on $B(\theta^*;r)$.

Lemma 2: *Let* $X\sim NHG\left(\{p_k\}_{k=1,2,\dots}\right)$ *where* $p_k>0 \quad \forall k=1,2,\dots$ *and* $\lim_{k\to\infty}p_k\neq 0$. *Then* $\Pr\{X<\infty\}=1$.

Proof: Using equation 2, the proof is straightforward.

Theorem 1: *Let* $X(\varepsilon)$ *be the number of iterations required by a PS method to visit a point* θ *such that* $f(\theta)\in O(\varepsilon)$ *for a given* $\varepsilon>0$. *Under Assumption 1,* $\Pr\{X(\varepsilon)<\infty\}=1$.

Proof: By Lemma 1, a $B_\varepsilon(\theta^*;r')$ where $W(B_\varepsilon(\theta^*;r')\cap\Theta)>0$ exists for the given $\varepsilon>0$. Then, observe that $X(\varepsilon)\sim NHG\left(\{p_k\}_{k=1,2,\dots}\right)$ where $p_k=\Pr\left\{\bigcup_{j=1}^{H_k}\left(\theta_k^{(j)}\in B_\varepsilon(\theta^*;r')\cap\Theta\right)\right\}$. Since $W(B_\varepsilon(\theta^*;r')\cap\Theta)>0$, Definition 1 implies $p_k>0 \quad \forall k=1,2,\dots$ and $\lim_{k\to\infty}p_k\neq 0$. Hence, $\Pr\{X(\varepsilon)<\infty\}=1$ following Lemma 2.

Now let us consider the SO version of problem (1):

$$\min_{\theta\in\Theta}\left\{f(\theta)=\mathrm{E}\left(L(\theta)\right)\right\} \tag{3}$$

where L is a function of decision variables and a consistent estimator of f, also $\mathrm{E}(.)$ is the mathematical expectation operator. Indeed, we assume the closed form function $f(\theta)$ is not available and can only be numerically estimated by $L(\theta)$ through averaging a number of sample performance functions obtained via simulating design point θ.

If estimated values of the objective function are used, PS methods can be applied to SO problems too. However, in order to prove the convergence properties of the PS methods, the number of simulation replications for each solution point must be controlled. The real challenge is that although a PS method can visit a point in any ε-optimum interval by Theorem 1, it may not be able to actually introduce (select) the visited point or a better one as the optimum because of the random

noise in objective function values. A remedy that could be used for discrete and bounded decision variables is to guarantee (in probability or almost surely) that each single feasible point is simulated with infinite simulation runs as the number of iterations of the search algorithm goes to infinity. Then following the laws of large numbers, asymptotic convergence to a global optimum is easy to establish. However, we need to deal with infinite number of points in continuous-variable problems and the remedy for discrete settings may not work by its own. In the subsequent sections, we propose merging PS methods with specific R&S procedures. This way, as we will show later, the asymptotic convergence is guaranteed for continuous cases.

3. ANALYSIS OF HYBRID PROBABILISTIC SEARCH METHODS

In this section, we propose merging PS methods with R&S methods. Then the convergence properties of the integrated method are studied.

3.1. Indifference zone methods

Since an appropriate R&S method is required for the hybrid PS method of next section, we briefly introduce these methods in this section. Essentially, R&S procedures are statistical selection methods that evaluate exhaustively all members from a fixed and finite set of possible design options to determine ordering (Kim and Nelson 2006). Two major classes of R&S methods are subset selection and Indifference Zone (IZ) methods. We focus on IZ methods in this section.

Mathematically, assume there are ς different designs, the j th of which has mean objective function $\mu^{(j)}$ such that unknown to us $\mu^{(1)} \le \mu^{(2)} \le ... \le \mu^{(\varsigma)}$ (i.e. design one is the best). In IZ methods which only select one design, system one is selected with at least $(1-\alpha)\%$ chance if $\mu_2 - \mu_1 \ge \delta$ where the so called IZ parameter $\delta > 0$ is the smallest difference the experimenter feels is worth detecting. Indeed, IZ methods guarantee $\Pr\{\text{select one } | \mu_2 - \mu_1 \ge \delta\} \ge 1-\alpha$.

For a recent presentation of available IZ methods, see Benson et. al. (2006) and Kim and Nelson (2006). As a typical method, Bechhofer (1954) proposed the very first method under very strong assumptions of normality of responses with equal variances. This method determines the number of required simulation replications, λ, from each of competing designs via the following formula:

$$\lambda = \left\lceil 2\left(\frac{\sigma \mathrm{N}_{\varsigma-1,\rho}^{(1-p^*)}}{\delta} \right) \right\rceil$$

where σ is the known variance of simulation responses, $\mathrm{N}_{\varsigma-1,\rho}^{(1-p^*)}$ is the $(1-p^*)$ equicoordinate point of the $\varsigma-1$ dimensional multivariate standard normal distribution with off-diagonal correlation ρ. Values for $\mathrm{N}_{\varsigma-1,\rho}^{(1-p^*)}$ may be obtained from table lookup from Bechhofer et al. (1995).

3.2. Hybrid probabilistic search algorithm

Since simulation responses are noisy, the number of replications (samples) in a simulation run for each design point proposed by a PS can significantly affect two conflicting goals: convergence properties and efficiency of the algorithm. Efficiency recommends less replications in a simulation

run, but this means more noise in the estimated values of objective function, more chance of wrongly announcing the best solution and consequently poor convergence. We propose a statistical procedure to assess the number of replications of each design point that help the algorithm guarantee asymptotic convergence. Whenever the PS sends a new point to simulation module for the first time, only a constant number of replications denoted by η are initially used for estimation of the objective function; further replications are subject to necessity. We use an R&S method (specifically IZ technique) after a fixed number of iterations in order to determine the best point among already simulated points.

The statistical procedure merges with PS core optimizer as the algorithm below outlines:

Algorithm 2: Hybrid Probabilistic Search

Step 0: *Define two sequences called error rate denoted by* $\{\alpha_h\}_{h=1}^{\infty}$ *and IZ (parameter) denoted by* $\{\gamma_h\}_{h=1}^{\infty}$ *where* $\lim_{h\to\infty}\alpha_h = \lim_{h\to\infty}\gamma_h = 0$, $0 < \alpha_{h+1} \le \alpha_h \le 1$ *and* $0 < \gamma_{h+1} \le \gamma_h$. *Define the number of replications between the R&S implementations and denote it by* ϑ.

Step 1: *Choose the initial sampling strategy* Ψ_1 *and let algorithm iteration counter* $k = 1$ *and R&S implementation counter* $h = 1$.

Step 2: *Denote the introduced optimum of the algorithm after iteration* k *by* $\hat{\theta}_k^*$. *Set* $\hat{\theta}_1^* = \phi$.

Step 3: *Select new points* $\theta_k^{(1)},...,\theta_k^{(H_k)} \in \Theta$ *according to the sampling strategy* Ψ_k *and let* $Z_k = \left\{\theta_{k'}^{(j)} \quad j = 1,...,H_{k'}, k' = 1,2,...,k\right\}$.

Step 4: *For* $j = 1,...,H_k$, *simulate* $\theta_k^{(j)}$ *with* η *replications.*

Step 5: *Let* $\nu_k(\theta_{k'}^{(j)})$ *be the number of simulation replications done for* $\theta_{k'}^{(j)}$ *by the end of iteration* k. *Set replication counter* $\nu_k(\theta_k^{(j)}) = \eta$ *for* $j = 1,2,...,H_k$. *Set* $\nu_k(\theta_{k'}^{(j)}) = \nu_{k-1}(\theta_{k'}^{(j)})$ *for* $k' = 1,...,k-1$ *and* $j = 1,...,H_{k'}$.

Step 6: *Denote by* $L_r(\theta_k^{(j)})$, *the objective function of* $\theta_k^{(j)}$ *estimated in* r *th simulation replication. Set sample mean* $L(\theta_k^{(j)}) = \sum_{r=1}^{\nu_k(\theta_k^{(j)})} \frac{L_r(\theta_k^{(j)})}{\nu_k(\theta_k^{(j)})}$ *for* $j = 1,...,H_k$.

Step 7: *If* $k \neq \vartheta h$, *then let* $\hat{\theta}_k^*$ *be one of the possibly many design points in* Z_k *for which* $L(\hat{\theta}_k^*) = \min_{\theta_k^{(j)} \in Z_k} L(\theta_{k'}^{(j)})$ *and go to step 8; otherwise do:*

7.1. *Design an R&S method and apply it to* Z_k *such that a difference of* γ_h *or less in the mean objective functions of the best point in* Z_k *and all other points in* Z_k *is detected with*

probability $1-\alpha_h$ *or more. For* $k'=1,...,k$ *and* $j=1,...,H_{k'}$, *let* $\lambda_{hjk'}$ *denote the total number of new simulation replications consumed in the current R&S implementation for* $\theta_{k'}^{(j)} \in \mathrm{Z}_k$ *on top of* $\nu_k(\theta_{k'}^{(j)})$ *old simulation replications available for this point and let* $\nu_k(\theta_{k'}^{(j)}) = \nu_k(\theta_{k'}^{(j)}) + \lambda_{hjk'}$.

7.2. *Set* $L(\theta_{k'}^{(j)}) = \sum_{r=1}^{\nu_k(\theta_{k'}^{(j)})} \frac{L_r(\theta_{k'}^{(j)})}{\nu_k(\theta_{k'}^{(j)})}$ *for all* $\theta_{k'}^{(j)} \in \mathrm{Z}_k$. *Then let* $\hat{\theta}_k^*$ *be one of the possibly many design points in* Z_k *for which* $L(\hat{\theta}_k^*) = \min_{\theta_{k'}^{(j)} \in \mathrm{Z}_k} L(\theta_{k'}^{(j)})$.

7.3. *Set* $h = h+1$ *and go to step 8.*

Step 8: *If a termination condition holds, introduce* $\hat{\theta}_k^*$ *as the optimum and exit the algorithm. Otherwise choose an updated strategy* Ψ_{k+1}, *let* $k = k+1$ *and go to step 3.*

Generally, any IZ procedure that uses old simulation replications along with new replications and guarantees selection of the best with a given probability when the true objective function of the best and the rest of the designs are distanced by an IZ parameter could be used in Hybrid PS algorithm. Boesel et. al. (2003) proposes such procedures. Bayesian methods are other alternatives in which a posterior probability of correct selection is guaranteed (Chick and Inoue 2001a and 2001b).

However, as far as we are aware, all statistical selection of the best procedures in the literature (including in Boesel et. al. and Chick and Inoue 2001a and 2001b) assume that simulation outputs are normally distributed. Nelson et. al. (2001) and Nelson and Goldsman (2001) study the robustness of normality assumption and conclude that probability of correct selection could approximately be retained with mild departures from normality. The main justification for normal assumption in many simulation studies is that interesting simulation outputs are usually averages of a large number of observations; hence, central limit theorem suggests normality holds asymptotically. In addition, we need to assume that the second moment of the objective function for all feasible values of decision variable is finite.

3.3. Convergence of the hybrid PS methods

Theorem 2: *If a hybrid PS method is applied to SO problem of equation 1 under Assumption 1, the sequence* $\left\{f(\hat{\theta}_k^*)\right\}_{k=1}^{\infty}$ *converges in probability to* $f(\theta^*)$, *that is* $\lim_{k\to\infty} \Pr\left\{\left|f(\hat{\theta}_k^*) - f(\theta^*)\right| < \varepsilon\right\} = 1$ *for all* $\varepsilon > 0$.

Proof: Following Theorem 1, the number of iterations before the algorithm visits a point with mean objective function in $O(\varepsilon)$ would be finite with probability 1 (w.p.1); say for iterations $k \geq k'$ where $\Pr(k' < \infty) = 1$, a point θ' has already been simulated such that $f(\theta') \in O(\varepsilon)$. Since the IZ sequence $\{\gamma_h\}_{h=1}^{\infty}$ converges to zero, the number of iterations before the algorithm visits a point θ' with mean objective function in $O(\varepsilon)$ and IZ parameter becomes less than $f(\theta^*) + \varepsilon - f(\theta')$ is finite w.p.1; say these two events have happened for any iteration $k \geq k''$ where $\Pr(k'' < \infty) = 1$.

Let $h'' = \left\lceil \frac{k''}{\vartheta} \right\rceil$ where $\lceil . \rceil$ is round up function. Then for $h \geq h'$, the R&S procedure implementations guarantee introduction of θ' or another point with objective function in $O(\varepsilon)$ as the optimum with probability $1-\alpha_{h'}$ or more because at least there exist θ' in the set of simulated points such that it is $f(\theta^*) + \varepsilon - f(\theta') > \gamma_{h'}$ far away from all the other simulated points outside $O(\varepsilon)$. Hence: $\lim_{k\to\infty} \Pr\left\{ \left| f(\hat{\theta}_k^*) - f(\theta^*) \right| < \varepsilon \right\} \geq \Pr(k'' < \infty) \lim_{h\to\infty} (1-\alpha_h) = 1$ and consequently $\lim_{k\to\infty} \Pr\left\{ \left| f(\hat{\theta}_k^*) - f(\theta^*) \right| < \varepsilon \right\} = 1$.

4. EXAMPLE PROCEDURE

In this section, we showcase the effectiveness of the Hybrid PS methods. Of course, any PS method could potentially be used here; but we are more interested to see how well the wedding between R&S methods and PS procedures work. Therefore, we select the simplest possible PS method which is called Naïve Random Search (NRS). The sampling strategy of NRS picks one point ($H_k = 1$ for $k = 1,2,...$) uniformly randomly from the compact feasible region. The method is called "Naïve" because it ignores the information of past searches in future sampling strategies. When applied to SO problems, NRS introduces the point with lowest estimated objective function as the current optimum.

Algorithm 3: Naïve Random Search

Step 1: Let $k = 1$.

Step 2: Select one point from the feasible region uniformly randomly. Denote this point by $\theta_k^{(1)}$.

Step 3: Simulate $\theta_k^{(1)}$ *with* η *replications and let* $L(\theta_k^{(1)}) = \sum_{r=1}^{\eta} \frac{L_r(\theta_k^{(1)})}{\eta}$ *be the estimated objective function.*

Step 4: Let $\hat{\theta}_k^*$ *be one of the possibly many design points for which* $L(\hat{\theta}_k^*) = \min_{\theta_{k'}^{(1)}} \min_{\forall k'=1,2,...,k} L(\theta_{k'}^{(1)})$.

Step 5: If termination condition(s) of the algorithm hold(s), stop the algorithm, otherwise let $k = k+1$ *and go to step 2.*

We are interested in comparing the performance of NRS with the so called Hybrid NRS defined below.

Definition 5: *Hybrid NRS is a kind of Hybrid PS method in which the sampling strategy of each iteration selects one point uniformly randomly from the feasible region.*

In our experiments, we use the IZ procedure of Boesel et. al. (2003) as the R&S procedure required for the Hybrid NRS. Also, we set $\eta = 2$, $\vartheta = 100$, $\alpha_1 = 0.50$, $\gamma_1 = 10$, and for $h = 1,2,...$

$\alpha_{h+1} = 0.9\alpha_h$, $\gamma_{h+1} = 0.9\gamma_h$. To accelerate the experiments, we replace simulation with a noisy objective function. Specifically, we use a closed form 2-dimensional objective function with a unique global optimum. Whenever the simulation output is required for a point, we generate a zero-mean normal random variable with variance 10 and add it to the objective function value computed via the closed-form formula. Figure 1 shows the objective function of the problem we used in our experiment. The decision variables are both bounded between 0 and 10. The global optimum of the problem is (5,5) with the objective function 1. We terminated the optimization process of both algorithms when a budget of 1000 simulation replications was spent. Both methods were run 10000 times in order to get robust results. For each algorithm, we computed the average of the expected value of the objective function of the introduced optimum after each simulation replication. Figure 2 shows logarithmic values of these averages as the optimization process progresses.

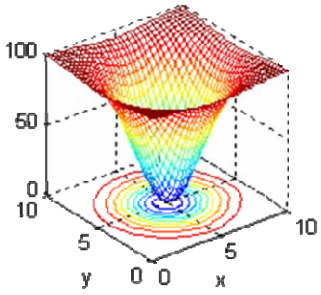

Figure 1 The objective function of the test problem

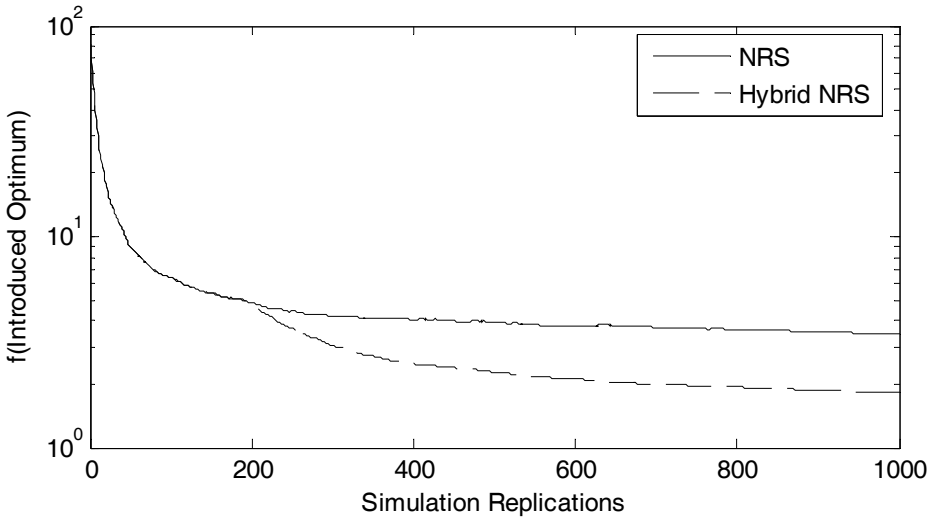

Figure 2 The performance curves of Naïve Random Search and Hybrid Random Search methods

Figure 2 suggests that when R&S procedure is applied for the first time in Hybrid NRS, convergence is accelerated. As we expected, the disadvantage of solely using NRS is that it may simulate a bad point which turn out to have a very good estimated objective function such that even near optimum points no longer could outperform the misleading estimated good quality of the bad point. In fact, R&S helps clean up the quality of the points NRS simulates.

5. REMARKS

In this paper, a new methodology for merging R&S and a class of random search methods called Probabilistic Search algorithms for continuous simulation optimization was proposed. We proved the asymptotic convergence in probability to a global optimum under a regular assumption and showed the performance of the new methodology thorough a simple example. The importance of the analysis of this paper roots in these facts:

- Discrete event Simulation based optimization methodologies are widely applicable in many engineering system design problems where evaluating the objective function and/or constraints are costly, requiring a submodel such as a simulation, an experiment, etc to run.

- The new methodology presented in this paper can simultaneously benefit the observed efficiency of a large class of random search methods and convergence properties of statistical selection procedures.

There are several open directions for future research. The feasible region in equation (1) was deterministic; however, stochastic constraints seem more realistic. Extension of the methodology presented here to problems with noisy constraints is an open research. We are exploring this area in Kabirian and Olafsson (2007c and 2008). The challenge in stochastic-constraint case is that given a point, we are never sure that the point is feasible or not because simulation runs are terminated in finite time. Unless unrealistic assumptions are considered, all we can hope in such cases is an asymptotic convergence based on laws of large numbers. In addition, the convergence properties of the methodology in finite time (not asymptotically) deserve more investigation.

ACKNOWLEDGEMENT

I would like to thank Professor Sigurdur Olafsson for supervising my Ph.D. dissertation at Iowa State University. The helpful comments of the anonymous referees of this paper are also greatly appreciated.

REFERENCES

[1] Andradóttir S. (2006), An overview of simulation optimization via random search. In Henderson and Nelson (eds.); *Handbooks in operations research and management science* 13; 617 – 632, Elsevier.

[2] Andradóttir S. (1998), A review of simulation optimization techniques. In D.J. Medeiros, E.F.Watson, J.S. Carson, and M.S. Manivannan (eds.); *Proceedings of the 1998 Winter Simulation Conference*; 151-158.

[3] Azadivar F. (1999), Simulation optimization methodologies; *Proceedings of the 1999 Winter Simulation Conference*; 93-100.

[4] Barton R.R., Meckesheimer M. (2006), Metamodel-based simulation optimization, In Henderson and Nelson (eds.), *Handbooks in operations research and management science* 13; 617 – 632, Elsevier.

[5] 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.

[6] Bechhofer R.E., Santner T.J., and Goldsman D.M. (1995), Design and analysis of experiments for statistical selection screening and multiple comparisons; John Wiley & Sons, Inc, New York.

[7] Benson K.C., Goldsman D., and Pritchett A.R. (2006), Ranking and election procedures for simulation, *proceedings of the 2006 winter simulation conference*; 179-185.

[8] Boesel J, Nelson B.L., and Kim S-H. (2003), Using ranking and selection to "clean up" after simulation optimization; *Operations Research* 51(5); 814–825.

[9] Chick S.E., Inoue K. (2001a), New two-stage and sequential procedures for selecting the best simulated system; *Operations Research* 49(5); 732-743.

[10] Chick S.E., Inoue K. (2001b), New procedures for identifying the best simulated system using common random numbers; *Management Science* 47(8); 1133–1149.

[11] Deng G., Ferris M.C. (2009), Variable-number sample-path optimization, *Mathematical Programming* 117(1-2); 81-109.

[12] Fu M.C. (2006), Gradient estimation. In Henderson and Nelson (eds.), *Handbook in operations research and management science* 13; 575-616, Elsevier.

[13] Fu M.C., Glover F.W., and April J. (2005), Simulation optimization: a review, new developments, and applications; *Proceedings of the 2005 Winter Simulation Conference*; 83-95.

[14] Glover F. (1997), A template for scatter search and path relinking; *Lecture Notes in Computer Science 1997, edited by Hao J.K., Lutton E., Ronald E., Schoenauer M., and Snyers D.*; 1363: 1-53.

[15] Glover, F. (1989), Tabu search – part I; *ORSA Journal on Computing* 1; 190-206.

[16] Glover, F. (1990), Tabu search –part II; *ORSA Journal on Computing* 2; 4-32.

[17] Goldsman D., Nelson B.L. (1998), Comparing systems via simulation. In Banks (ed), *Handbooks of simulation* (13); John Wiley and Sons.

[18] Gong W.B., Ho Y.C., and Zhai W. (1999), Stochastic comparison algorithm for discrete optimization with estimations; *SIAM Journal on Optimization* 10(2); 384-404.

[19] Gosavi A. (2003), Simulation-based optimization: parametric optimization techniques and reinforcement learning; Kluwer Academic Publishers.

[20] Holland J.H. (1992), *Genetic algorithms; Scientific American July 1992*; 66-72.

[21] Kabirian A. (2006), A review of current methodologies and developing a new heuristic method of simulation optimization; *Unpublished Master of Science Dissertation, Under supervision of Professor Hashem Mahlooji Jan. 2006*; Sharif University of Technology, Tehran, Iran.

[22] Kabirian A., Hemmati M.R. (2007), A strategic planning model for natural gas transmission networks, *Energy Policy* 35(11); 5656-5670, Elsevier.

[23] Kabirian A., Olafsson S. (2007a), Allocation of simulation runs for simulation optimization; *Proceedings of the 2007 Winter Simulation Conference*; PhD Colloquium & Poster Session, Washington D.C., December 2007.

[24] Kabirian A., Olafsson S. (2007b), Simulation based optimization via golden region search; *Working Paper*; Department of Industrial and Manufacturing Systems Engineering, Iowa State University, Ames, IA, USA.

[25] Kabirian A., Olafsson S. (2007c), Multiresponse golden partition search under noisy constraints; *Working Paper*, Department of Industrial and Manufacturing Systems Engineering; Iowa State University, Ames, IA, USA.

[26] Kabirian A., Olafsson S. (2008), Selection of the best with multiple stochastic constraints; *Working Paper*, Department of Industrial and Manufacturing Systems Engineering; Iowa State University, Ames, IA, USA.

[27] Kiefer J., Wolfoxitz J. (1952), Stochastic estimation of the maximum of a regression function. *Annals of Mathematical Statistics* 23; 462-466.

[28] Kim S., Nelson B.L. (2006), Selecting the best system. In Henderson and Nelson (eds.), *Handbooks in operations research and management science* 13; 501-534, Elsevier.

[29] Krikpatrick S., Gelatt Jr.C.D., Vecchi M.P. (1983), Optimization by simulated annealing; *Science* 220; 671-680.

[30] Nelson B.L., Goldsman L.D. (2001), Comparison with a standard in simulation experiments; *Management Science* 47; 449-463.

[31] Nelson B.L., Swann J., Goldsman D., and Song W. (2001), Simple procedures for selecting the best simulated system when the number of alternatives is large; *Operations Research* 49; 950–963.

[32] Olafsson S., Kim J. (2002), Simulation optimization; *Proceedings of the 2002 Winter Simulation Conference*; 1931-1936.

[33] Olafsson S. (2006), Metaheuristics. In Henderson and Nelson (eds.); *Handbook in operations research and management science* 13; 633–654, Elsevier.

[34] Robins H., Monro S. (1951), A stochastic approximation method; *Annals of Mathematical Statistics* 22; 400-407.

[35] Robinson S.M. (1996), Analysis of sample-path optimization; *Mathematics of Operations Research*; 21 (3); 513-528.

[36] Shapiro A. (1996), Simulation based optimization, *Proceedings of the 1996 Winter Simulation Conference, ed. J.M. Charnes, D.J. Morrice, D.T. Brunner, and J.J.Swain*; 332-336.

[37] Yan D., Mukai H. (1992), Stochastic discrete optimization; *SIAM Journal on Control and Optimization* 30; 594-612.