Nonlinear Dynamics and Systems Theory, **21** (1) (2021) 13–26

Stopping Rules for Selecting the Optimal Subset

Mohammad H. Almomani [1]* and Mahmoud H. Alrefaei [2] and Mohanad H. Almomani

[1] *Department of Mathematics, Faculty of Science, The Hashemite University, P.O. box 330127, Zarqa 13133, JORDAN.*
[2] *Department of Mathematics and Statistics, Jordan University of Science and Technology, Irbid, JORDAN.*
[3] *The Ministry of Education, Ajloun, JORDAN.*

Received: September 25, 2020; Revised: December 26, 2020

Abstract: Selecting the best of a finite set of alternatives is a very important area of research. In this paper, we discuss the stopping rules of the procedure of selecting the optimum subset out of a very large alternative dynamic system. A combined procedure with two stages is studied. The first stage employs the ordinal optimization to select a subset that overlaps with the set of actual best $k\%$ designs with high probability. After that, the optimal computing budget allocation is used in the second stage to select the best m designs from the selected subset. The efficiency of selection procedures with two different stopping rules is studied by implementing them on two test problems to see the efficiency of the procedure in the context of the most effective stopping rule. The first problem is a generic example and the second one is a buffer allocation problem.

Keywords: *large scale problems; simulation optimization; ordinal optimization; stopping rules; optimal computing budget allocation.*

1 Introduction

Statistical selection procedures are designed to answer the question "*which treatment can be considered the best?*", where the *best* refers to the design that has the maximum or minimum expected performance measure. Different sampling assumptions, approximations, parameters and stopping rules were combined to define a procedure. Due to the increasing demands that are being placed upon simulation optimization algorithms together with having many differences between the statistical selection procedures, it is getting important to find out which of these procedures is the most convenient one to

* Corresponding author: mailto:mh_momani@hu.edu.jo

use. However, evaluating selection procedures can be done in several ways, including the theoretical, empirical, and practical perspectives.

Simulation is often used by managers to help make a decision by exploring different options. Creating different physical models and conducting experiments on them in order to choose the best model usually cost a lot. Simulation is a relatively cheap alternative for collecting information, since it provides some estimates about the performance of a system that does not exist physically. The experiments are conducted as statistical experiments; the number of times the simulation runs is determined, the experiments are performed, and the output data are analyzed. A stopping rule that decides when the experiment ends -i. e., how many replicates of the simulation are made - must be chosen by the manager. This paper studies how different stopping rules affect the simulation output analysis.

Selection procedures used simulation to estimate the performance measure. Since the simulation methods are used to indicate the performance measure for each alternative, an incorrect selection is possible. Thus, some measures are needed to determine the selection quality. Two measures of selection quality exist; the first one is the Probability of Correct Selection ($P(CS)$) and the second measure is the Expected Opportunity Cost ($E(OC)$) of a potentially incorrect selection, see He et al. [1]. Traditional selection procedures identify the best system with high probability of correct selection, by maximizing the $P(CS)$. However, the $E(OC)$ is applied in business, engineering and many other applications. This led recently to a new selection procedure that reduces the cost of a potentially incorrect selection.

The measures of selection quality can be used to decide when to stop the sampling process. In particular, Brank et al. [2] proposed the following stopping rules:

1. Sequential (S): Repeat sampling while $\sum_{i=1}^{n} T_i < T$, for some specified total budget T and T_i is the number of samples allocated to design i, where $i = 1, 2, \ldots, n$.

2. Expected opportunity cost (EOC): Repeat sampling while $E(OC) > \varepsilon$, for a specified expected opportunity cost target $\varepsilon > 0$.

3. Probability of good selection ($P(GS)_{\delta^*}$): Repeat sampling while $P(GS)_{\delta^*} < 1 - \varphi^*$, for a specified probability target $1 - \varphi^* \in [1/n, 1)$ and given $\delta^* \geq 0$. In the Indifference Zone (IZ) procedure, see Bechhofer et al. [3], δ^* is the difference between the favorable design and the best design and is called the indifference zone. It represents the smallest difference that one wants to achieve.

In many practical problems, selecting a set of m best solutions out of n solutions is more convenient than selecting only one solution. This is done based on the simulation output from each design. In case of having a small size of the feasible solution set, the best design or a subset of the best designs can be selected using Ranking and Selection procedures, see Bechhofer et al. [3], Law and Kelton [4], and Kim and Nelson [5], [6]. Ranking and Selection procedures for large alternatives require a very big computational time. Thus, such procedures might not be feasible for large scale problems. For comprehensive reviews of the Ranking and Selection procedures, see Gibbons et al. [7] and Gupta and Panchapakesan [8].

The Ordinal Optimization (OO) that was proposed by Ho et al. [9] relaxes the objective to finding good enough designs, rather than estimating the performance of the designs accurately. In fact, the OO procedure seeks to isolate a subset of solutions with high goodness probability. After that, the optimal solution(s) can be located from the isolated set by using any simulation optimization procedure. Previously, many procedures have been proposed to select a good design (designs) to solve this selection problem, see for example Alrefaei and Almomani [10], Almomani and Alrefaei [11], Almomani and Abdul Rahman [12], Al-Salem et al. [13], Almomani et al. [14], Almomani et al. [15], Alrefaei et al. [16].

Consider the problem of distributing an available budget computation in simulating the different solutions. Instead of distributing this budget evenly on different alternatives, the available budget can be distributed in a way that maximizes the probability of correct selecting the good solutions.Therefore, the idea of the Optimal Computing Budget Allocation $OCBA_m$ has been proposed by Chen et al. [17] for selecting the best m designs.

Recently, a sequential selection procedure was considered by Almomani and Alrefaei [18] for selecting a good subset of solutions from a large size problem. The procedure combines the $OCBA_m$ and the OO procedures. In the first stage, an isolation of a subset of good enough designs with high probability is made by the OO procedure. This reduces the feasible solution set size and makes it appropriate to apply the $OCBA_m$ procedure. In the second stage, a maximization problem, that seeks to maximize the probability of selecting all best m designs correctly from the subset found in the first stage, is formulated using the $OCBA_m$. A constraint on the total number of available simulation replications is considered for this maximization problem. The procedure starts by simulating each alternative in the set and considers by initial simulation replications of size t_0. After that, a fixed increment of replications Δ is added and distributed among all solutions in the set. The process is repeated until all available computations are consumed.

In this paper, the effect of the stopping rules is studied; we study two stopping rules and implement them in our proposed algorithm; these are the sequential S and the expected opportunity cost EOC. However, the third stopping rule that uses the probability of good selection $P(GS)_{\delta^*}$ is not applicable in the proposed algorithm. These stopping rules are tested and compared by different examples using different measures to study their effect on the final solution of the selection procedure. These rules give the chance to stop the procedure earlier whenever the evidence for correct selection is high enough, and allow for additional sampling when it is not, which gives a kind of flexibility. Furthermore, we apply a numerical illustration for this approach to display the advantages and the disadvantages for each stopping rule, and to determine the most effective stopping rule that works better with this selection procedure.

The rest of this paper is organized as follows. In Section 2, we provide the background of the problem statement, OO procedure, and $OCBA_m$ procedure. In Section 3, the sequential selection procedure is presented with two different stopping rules. The performance of the selection approach under these stopping rules is illustrated with a series of numerical examples in Section 4. Finally, Section 5 includes concluding remarks.

2 Background

2.1 Problem statement

Consider the following simulation optimization problem

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

where Θ is an arbitrary, large and finite feasible solution. Let $Y(\theta) = E[L(\theta, X)]$ be the expected performance measure in a specific complex stochastic design, where θ represents the system parameters as a vector, X represents the random effects on the system and L is a deterministic function of θ and X.

Simulation is used to infer the set of the best m designs, which are the designs with the m smallest means out of the n designs we have in the feasible solution set. Let Y_{ij} (observation) represent the j^{th} sample of $Y(i)$ for the design i. We assume that Y_{ij} are independent and identically distributed ($i.i.d.$) normally distributed with unknown means $Y_i = E(Y_{ij})$ and variances $\sigma_i^2 = Var(Y_{ij})$, i.e. $Y_{i1}, Y_{i2}, \ldots, Y_{iT_i}$ are $i.i.d.$ $N(Y_i, \sigma_i^2)$. There is no a problem with the normality assumption here since it holds for sure. This can be shown using the Central Limit Theorem, regarding that simulation outputs are acquired through an average performance or from batch means. In practice, we estimate the variance σ_i^2 using the sample variance s_i^2 for Y_{ij} because it is unknown. Our aim is to select a set, S_m, containing the best m designs. The word "best", in a minimization problem, refers to the one with the smallest sample mean. Define $\overline{Y}_{[r]}$ to be the r-th smallest (statistic order) of $\{\overline{Y}_1, \overline{Y}_2, \ldots, \overline{Y}_n\}$, i.e., $\overline{Y}_{[1]} \leq \overline{Y}_{[2]} \leq \ldots \leq \overline{Y}_{[n]}$, where $\overline{Y}_i = \frac{1}{T_i}\sum_{i=1}^{T_i} Y_{ij}$ is the sample mean for the design i. After that, let $S_m = \{[1], [2], \ldots, [m]\}$, which gives the correct selection that contains all of the m designs with smallest means, i.e., $CS_m = \{\max_{i \in S_m} \overline{Y}_i \leq \min_{i \notin S_m} \overline{Y}_i\}$.

2.2 Ordinal optimization

For large scale selection problems, the Ordinal Optimization (OO) procedure was proposed by Ho et al. [9]. Due to the high cost of accurate estimating the design performance values in the optimization process, it would be more practical to select a subset of the feasible set containing some of the best designs with high probability. This means that ordinal optimization is first used to isolate a subset of good enough designs, then cardinal optimization is applied on this isolated set. The main objective is to reduce the required simulation time for the discrete event simulation. A review of the OO procedure can be found in Ho et al. [19], Horng and Lin [20] and Ma et al. [21].

2.3 The selection procedure

Minimizing the total computational time for different designs in the simulation process is important to make the OO procedure more effective. Therefore, it is necessary to allocate the simulation samples cleverly, where a greater number of samples is applied to the designs that are more effective in identifying the best design. In this case, noncritical designs with smaller effect on discovering the best designs are not given much simulation samples. Chen et al. [22] have proposed the Optimal Computing Budget Allocation ($OCBA$) procedure which focuses on selecting the best design. On the other hand, an efficient allocation procedure, ($OCBA_m$), was also proposed by Chen et al. [17] for selecting the top m designs.

The problem is formulated by Chen et al. [17] so as to maximize the probability of selecting the best m designs $P(CS_m)$ correctly, subject to a constraint on the available number of samples. In mathematical notation, the problem can be written as

$$
\begin{aligned}
&\max_{T_1,\ldots,T_n} \; P(CS_m) \\
&\text{s.t.} \; \sum_{i=1}^{n} T_i = T,
\end{aligned}
\tag{2}
$$

where T is the number of available simulation samples, and for the design i we set T_i simulation samples. n is the total number of designs, m is the size of the optimal subset, S_m, that contains the best designs. The goal is to allocate the simulation samples we have in a way that maximizes the probability of the correct selection, $P(CS_m)$, given the total number of samples as $\sum_{i=1}^{n} T_i$. By this formulation, the computational cost of each sample is implicitly assumed to be constant across designs. Chen and Lee [23] suggested approximating $P(CS_m)$ by a lower bound of it, $APCS_m$, which determines an asymptotic approximation for the samples $T_i,\ i = 1, \ldots, n$ that maximize $APCS_m$. The following theorem of Chen and Lee [23] gives the estimates of these $T_i's$.

Theorem 2.1 *Given a total number of simulation samples T to be distributed to n competing designs whose performance is represented by random variables with means $Y_1, Y_2, \ldots, Y_n$, and finite variances $\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2$, respectively. The Approximate Probability of Correct Selection for m best ($APCS_m$) as $T \longrightarrow \infty$ can be asymptotically maximized when $\frac{T_i}{T_j} = \left(\frac{\sigma_i/\delta_i}{\sigma_j/\delta_j}\right)^2$; for any $i, j \in \{1, 2, \ldots, n\}$ and $i \neq j$, where $\delta_i = \bar{Y}_i - c$, for some constant c.*

3 The Selection Procedure with Stopping Rules

We consider the sequential selection procedure that consists of two stages, see Almomani and Alrefaei [18]. In the first stage, out of the search space, a subset G is selected randomly using the OO procedure, such that G overlaps the set containing the actual best $k\%$ designs with high probability. After that, in the second stage, the best m designs are identified from the subset G using the $OCBA_m$ procedure.

In this section, we present the sequential selection procedure for selecting the optimal subset, then we discuss the different stopping rules used to stop the procedure. These are the sequential, the expected opportunity cost and the probability of good selection.

3.1 A selection procedure for selecting the best m designs

The run length of a simulation experiment is often determined using sequential stopping rules. These rules can be used within confidence interval procedures for simulation output analysis. In the sequential stopping rule if the total number of samples is exceeded (i.e., if $\sum_{i=1}^{g} T_i^l \geq T$), then the algorithm stops. In fact, most traditional selection procedures use this stopping rule. Since the target is selecting the best design with minimum elapsed time, we can control this by increasing or decreasing the total budget T.

We first present the sequential algorithm proceedure and then we discuss the stopping rules used to stop the algorithm.

Setup: Determine the precision level p_0 and let $|G| = g$, where G is defined as the required subset from Θ, that satisfies $P(G$ contains at least m of the best $k\%$ designs$) \geq p_0$. Determine the number of initial simulation samples $t_0 \geq 5$.

Determine the total computing budget T, and the value of m (best top m). Let $l = 0$ and let $T_1^l = T_2^l = \ldots = T_g^l = t_0$, where l is the iteration number.

Select a subset G of g alternatives randomly from Θ. Take random samples of t_0 observations Y_{ij} $(1 \leq j \leq t_0)$ for each design i in G, where $i = 1, 2, \ldots, g$.

Initialization: For each $i \in G$, find an estimate of the sample mean and the sample variance as $\bar{Y}_i = \frac{1}{T_i^l} \sum_{j=1}^{T_i^l} Y_{ij}$ and $s_i^2 = \frac{1}{T_i^l - 1} \sum_{j=1}^{T_i^l} (Y_{ij} - \bar{Y}_i)^2$.

Order the sample means $\bar{Y}_{[1]} \geq \bar{Y}_{[2]} \geq \ldots \geq \bar{Y}_{[g]}$. Then select the set of top m designs, S_m.

Stopping Rule: Test the stopping rule, if it is satisfied, then stop, return S_m as the required subset. Otherwise, select randomly a subset S_z of $g - m$ alternatives from $\Theta - S_m$. Take random samples of t_0 observations Y_{ij} $(1 \leq j \leq t_0)$ for each design i in S_z. Compute $\bar{Y}_i$ and s_i as in the **Initialization** step above. Let $G = S_m \bigcup S_z$.

Simulation Budget Allocation: Increase the computing budget by Δ and compute the new budget allocation, $T_1^{l+1}, T_2^{l+1}, \ldots, T_g^{l+1}$ using $\frac{T_1^{l+1}}{\left(\frac{s_1}{\delta_1}\right)^2} = \frac{T_2^{l+1}}{\left(\frac{s_2}{\delta_2}\right)^2} = \cdots = \frac{T_g^{l+1}}{\left(\frac{s_g}{\delta_g}\right)^2}$, where $\delta_i = \bar{Y}_i - c$ and $c = \frac{\hat{\sigma}_{i_{m+1}} \bar{Y}_{i_m} + \hat{\sigma}_{i_m} \bar{Y}_{i_{m+1}}}{\hat{\sigma}_{i_m} + \hat{\sigma}_{i_{m+1}}}$ with $\hat{\sigma}_i = s_i / \sqrt{T_i^l}$, for all $i = 1, 2, \ldots, g$, see Chen and Lee [23].

Perform additional $\max\{0, T_i^{l+1} - T_i^l\}$ simulations samples for each designs i, where $i = 1, 2, \ldots, g$, let $l \longleftarrow l + 1$. Go to **Initialization**.

We consider two **Stopping Rules** for the proposed algorithm.

3.2 The sequential cost stopping rule

In this rule, a total number of samples T is predetermined. If the number of samples used in the algorithm reaches T, stop the algorithm, otherwise, continue. Therefore, the stopping rule becomes:

Stopping Rule: If $\sum_{i=1}^{g} T_i^l \leq T$, for a specified total number of samples T, then stop. Otherwise, proceed.

3.3 The expected opportunity cost stopping rule

This rule uses the expected opportunity cost EOC to stop the algorithm. In fact, the EOC stopping rule is recommended when the goal is to select the best design with the minimum $E(OC)$, especially, in business applications. Therefore, the stopping rule becomes:

Stopping Rule: If $E(OC) \leq \varepsilon$, for a specified expected opportunity cost significant $\varepsilon > 0$, then stop. Otherwise, proceed.

As we stated before, there is another stopping rule that was used in the literature but it is not applicable in our algorithm. This stopping rule is called the probability of good selection stopping rule. A selected design within δ^* from the best design is called the "good" design. However, since the objective in this paper is to select good enough designs from the actual best designs, but we are not concerned with the difference between the selected design and the actual best design(s), this stopping rule is not applicable. For more details about these three stopping rules, please, refer to Brank et al. [2] who provide an illustration about the difference between these three stopping rules through $KN{+}{+}$ procedure, see also Goldsman et al. [24].

4 Numerical Examples

In this section, we present two numerical examples: a generic monotone increasing mean example and a queuing model example. In both examples, the algorithm is applied under different experiment settings and different stopping rules.

4.1 Example 1

Consider n different designs, each one is normally distributed $N(\mu_i, \sigma^2)$ with mean μ_i and variance σ^2 for $i = 1, 2, \ldots, n$. Such problem is called the Monotone Increasing Mean (MIM), which aims to find the best m designs with the minimum mean. If we let Θ be the feasible solution set, then $\Theta = \{\theta_1, \theta_2, \ldots, \theta_n\}$ and $f(\theta_i)$ represents the index of θ_i in the feasible region, the optimization problem is

$$\min_{i=1,\ldots,n} f(\theta_i). \tag{3}$$

The proposed algorithm is applied on this example where $n = 1,000$, using the proposed selection procedure by implementing the first two stopping rules. It is assumed that for each $i \in \Theta, \mu_i = 10 + \frac{(i-1)}{10}$ and variance $\sigma_i^2 = 1$. Let $\theta_{[1]}, \theta_{[2]}, \ldots, \theta_{[n]}$ be the order of alternatives and we seek to select a subset of $m = 5$ solutions from the best 10% designs that have minimum means. If the selected design is in $\{\theta_{[1]}, \theta_{[2]}, \ldots, \theta_{[100]}\}$, then it is considered as a correct selection. The selection algorithm is applied using $g = 100$ solutions in G in order to study the effect of the simulation parameters, such as t_0 and Δ, on the performance of the algorithm. Furthermore, to achieve the normality assumption we use multiple replications method, where the number of multiple replications for each alternative equals M.

In the first experiment, we implement the proposed algorithm using the first stopping rule- the sequential S stopping rule. Here $n = 1,000$, $g = 100$, $k\% = 10\%$, $\Delta = 40$, $t_0 = 10$ and the total budget $T = 10,000$ (these settings are chosen arbitrary). In the second experiment, we implement the algorithm using the second stopping rule- the expected opportunity cost stopping rule with the same parameters setting as in the first experiment. The total budget condition is removed and replaced with the expected opportunity cost condition such that $E(OC) \leq 0.05$ (i.e., the significance level $\varepsilon = 0.05$). Table 1 contains the average performance of the algorithm over 100 replications for selecting 5 of the best 10% designs, for the first and the second experiment. In Table 1, $\overline{T}$ represents the average total sample size used in the algorithm $\sum_{i=1}^{g} T_i$ over the 100 replications. $\overline{P}$ represents the average probability of correctly selecting the best m designs; $P(CS_m)$ over the 100 replications, $\overline{E}$ represents the average expected opportunity cost for selecting the best m designs, $E(OC_m)$ over the 100 replications.

From Table 1, we note that the performance of the algorithm under the two stopping rules, the sequential S and the expected opportunity cost EOC, are almost the same with a preference of EOC on S when the measure is the expected opportunity cost used.

Table 1: The performance of the proposed procedure under different stopping rules for $n = 1000, g = 100, \Delta = 40, t_0 = 10, k\% = 10\%$ over 100 replications.

	Sequential Stopping Rule	EOC Stopping Rule
$\overline{T}$	6913	6995
$\overline{P}$	93%	93%
$\overline{E}$	0.010549	0.000523

To see the effect of the two stopping rules over different values of simulation budget on the probability of correct selection $P(CS_m)$, the results are depicted in Figure 1. It is clear from Figure 1 that the proposed algorithm with these two stopping rules produces a high $P(CS_m)$ quickly. Moreover, the performance of the two stopping rules is almost the same.

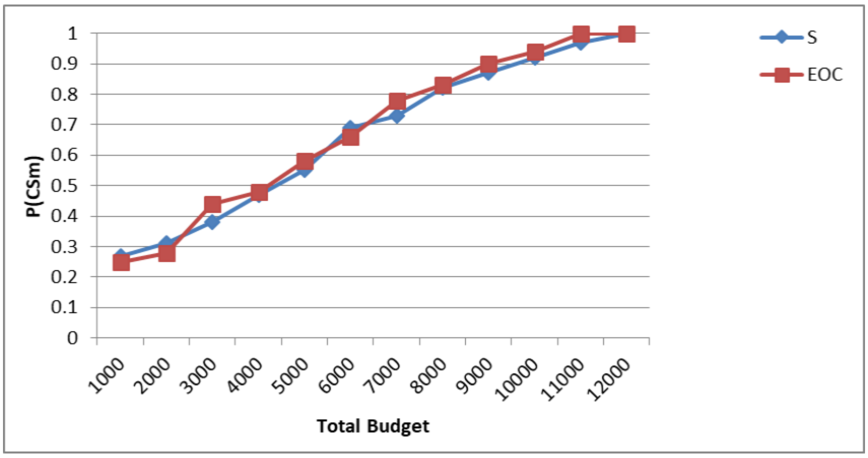

Figure 1: Comparison of the $P(CS_m)$ of S and EOC stopping rules over T budget for (MIM).

To study the effect of the two stopping rules on the proposed algorithm from the prospective of $E(OC_m)$, the results are depicted in Figure 2. Figure 2 shows the $E(OC_m)$ for the proposed algorithm using these two stopping rules over different values of simulation budget. From Figure 2, it is clear that the second stopping rule that uses the EOC gives better performance over the sequential procedure and that $E(OC_m)$ becomes close to 0 using this stopping rule under reasonable number of samples.

To enhance the performance of the algorithm, we increase the number of samples used in the multiple replication simulation method in order to get better estimates of the sample mean. Figure 3 shows the performance of the algorithm on the $E(OC_m)$ performance measure using the two stopping rules. Here M represents the samples used for each alternative. Obviously, it shows that the increase in the number of multiple replications M decreases the $E(OC_m)$ in both of the stopping rules. Moreover, the algorithm gives better performance when the EOC stopping rule is used. This is because when we increase the value of M we get better estimate value of the mean, therefore,

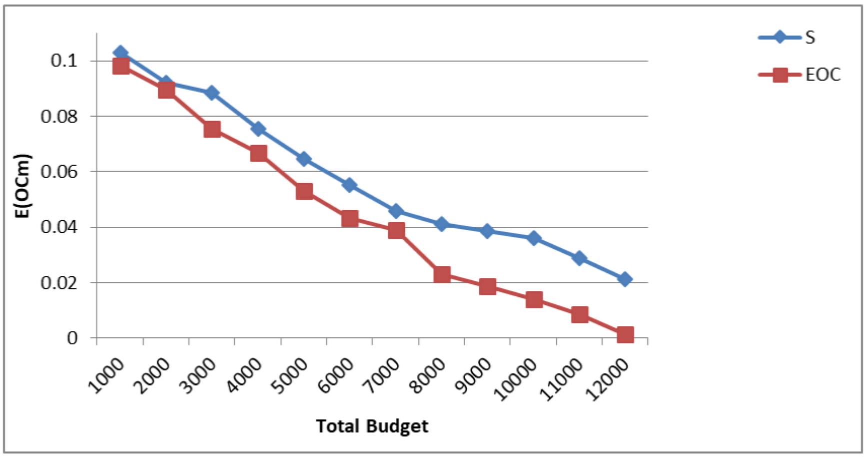

Figure 2: Comparison of the $E(OC_m)$ of S and EOC stopping rules over T budget for (MIM).

the difference between the estimated mean and the actual mean will be very small. In particular, when M is increased, then the $E(OC_m)$ approaches zero.

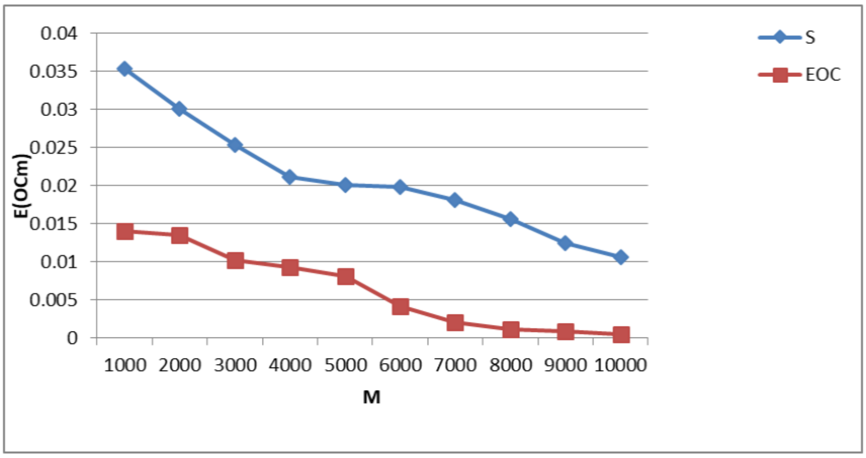

Figure 3: Comparison of the $E(OC_m)$ of S and EOC stopping rules over M replications for (MIM).

4.2 The buffer allocation problem (BAP)

We consider the Buffer Allocation Problem (BAP). The BAP consists of $q + 1$ machines and q intermediate buffers in between. The question is how to distribute the available Q buffer slots over the q buffers in a way that meets a specific purpose.

Each station is modeled as a single server queuing model with $q+1$ machines $M_0, M_1, \ldots, M_q$ and q intermediate buffers $B_1, B_2, \ldots, B_q$ in a production line as shown in Figure 4. Suppose that there are limits neither on the jobs in front of machine M_0, nor on the space for completed jobs after machine M_q. The service time at each machine is assumed to be independent and exponentially distributed random variable with rate $\mu_i,\ i = 1, 2, \ldots, q$. After the service is finished in machine M_i, the job tries to enter the queue of machine $M_{(i+1)}$. This cannot be done if the queue is full, and this prevents machine M_i from receiving new jobs to serve until the current job leaves it. Our goal is to maximize the production rate (throughput) by allocating the available spaces optimally on the intermediate buffers. Let Θ be the solution set, then Θ contains $\binom{Q+q-1}{Q}$ different solutions, see Almomani et al. [25], Papadopoulos et al. [26], Yuzukirmizia and Smith [27] and Alrefaei and Andradóttir [28].

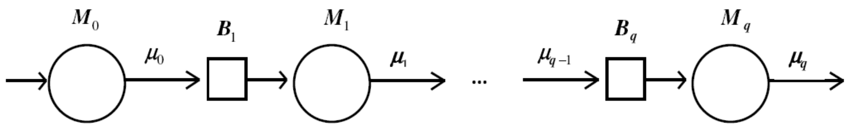

Figure 4: A production line with $q+1$ machines, q buffers, no limits either on the jobs in front of the machine M_0, or on the space for completed jobs after the machine M_q

The proposed algorithm is applied here on the specific type of BAP, with two stopping rules under some assumptions. In fact, there are different classifications for BAP problems. The first one is according to the length of the production line, which was presented by Papadopoulos et al. [26]. A production line is considered as “short” if the number of the machines is up to 6 with no more than 20 buffer spaces. Otherwise, the line is “large”. Another point of view defines the BAP according to whether it has a balanced line, with equal mean service time at each machine, or an unbalanced one. Moreover, production lines can be seen as reliable (no machine fails) or unreliable. For more information about these classifications, see Almomani et al. [25].

Suppose that there are $Q = 15$ slots to be allocated over $q = 5$ buffers. Thus, we have 6 workstation and Θ contains 3,876 different designs ($n = 3,876$). In addition, assume that $\mu_0 = \mu_1 = \mu_2 = \mu_3 = 5$ and $\mu_4 = \mu_5 = 10$, which means that we assumed an unbalanced production line in this example. Furthermore, let the size of the set G be $g = 80$, the number of initial simulation samples $t_0 = 20$, the total computing budget $T = 100,000$, and the increment in simulation samples $\Delta = 50$. Suppose that our objective is to select the two design from the best 5% designs of Θ. This means that the correct selection here is to select the 2 designs that belong to the set $\{\theta_{[1]}, \theta_{[2]}, \ldots, \theta_{[193]}\}$, where $\theta_{[i]}, i = 1, 2, \ldots, 193$ represents the set of the actual top 5% designs with the maximum throughput in the set Θ.

We apply the proposed algorithm using the two stopping rules, the sequential S and the expected opportunity cost EOC rules. The experiment is repeated for 100 replications and the results are summarized in Table 2. In Table 2, $\overline{T}$ represents the average total sample size used in the algorithm $\sum_{i=1}^{g} T_i$ over the 100 replications. $\overline{P}$ represents the average probability of correct selecting the best m designs; $P(CS_m)$ over the 100 replications, $\overline{E}$ represents the average expected opportunity cost for selecting the best m designs, $E(OC_m)$ over the 100 replications. Clearly, the proposed algorithm selected the best buffer profile with high $P(CS_m)$ and small $E(OC_m)$. At the same time, there is a relatively small number of simulation samples needed.

Table 2: The performance of the proposed procedure under different stopping rules for $n = 3876, g = 80, \Delta = 50, t_0 = 20, k\% = 5\%$ over 100 replications.

	The S Stopping Rule	The EOC Stopping Rule
$\overline{T}$	241334	245921
$\overline{P}$	85%	89%
$\overline{E}$	0.002032	0.0000876

Figure 5 shows the average $P(CS_m)$ for selecting the 2 designs of the best 5% designs, with the two stopping rules, the sequential S and the expected opportunity cost EOC over different values of simulation budget. It is clear that the proposed algorithm selects the best designs with hight $P(CS_m)$ for the two stopping rules S and EOC and the two stopping rules give almost the same results.

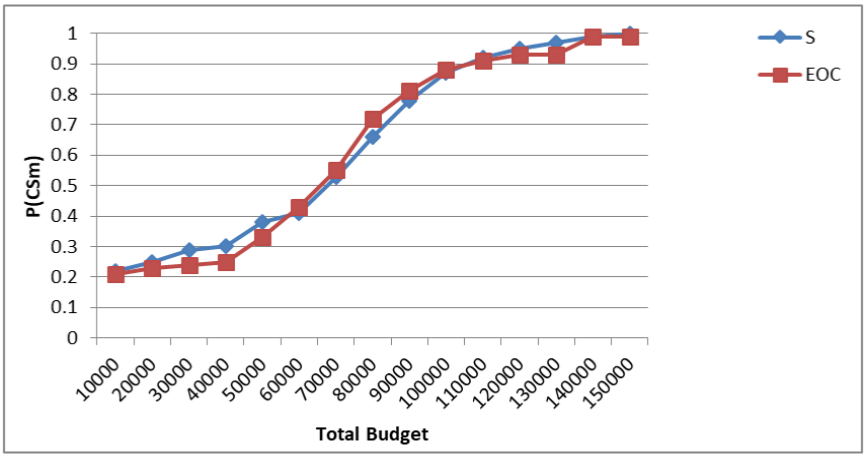

Figure 5: Comparison of the $P(CS_m)$ of S and EOC stopping rules over T budget for (BAP).

Figure 6 gives the $E(OC_m)$ performance measure for the proposed algorithm with two stopping rules, S and EOC, over different values of simulation budget. Clearly, the algorithm produces a very small $E(OC_m)$ under the two stopping rules with a little preferance of the EOC stopping rule over the sequential stopping rule. Also, with high value of total budget, the algorithm gives a very small value of $E(OC_m)$ which is close to 0, especially when the expected opportunity cost EOC stopping rule is used.

To enhance the performance of the algorithm, we increase the number of samples used in the multiple replication simulation method in order to get better estimates of the sample mean. Figure 7 gives the value of the $E(OC_m)$ against the number of samples M in the multiple replications method, for the two stopping rules S and EOC. Obviously, it shows that the increase in M gives a smaller value of $E(OC_m)$. It is clear again that the second stopping rule that uses EOC gives a slight better performance over the sequential stopping rule.

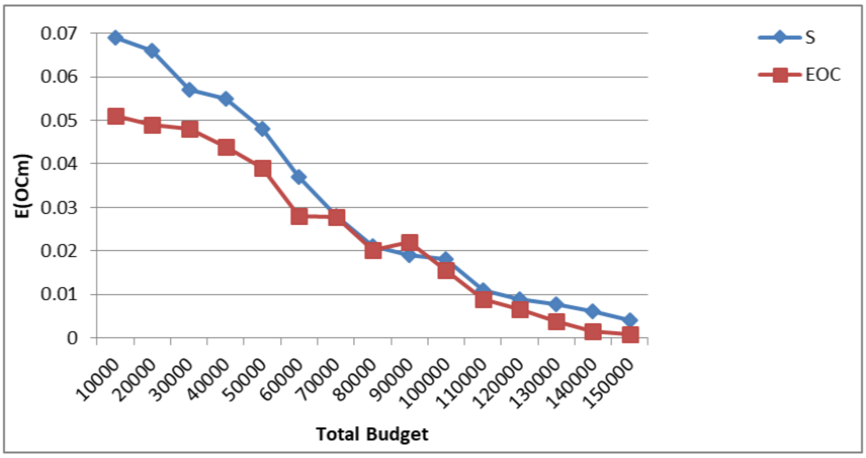

Figure 6: Comparison of the $E(OC_m)$ of S and EOC stopping rules over T budget for (BAP).

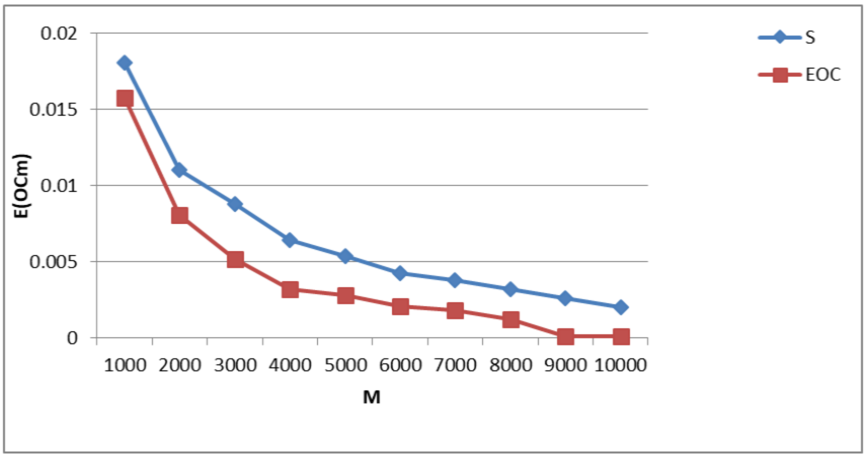

Figure 7: Comparison of the $E(OC_m)$ of S and EOC stopping rules over M replications for (BAP).

From MIM and BAP examples, it is clear that if objective is to select the best designs with high $P(CS_m)$ and minimum elapsed time, then the algorithm with the two stopping rules gives almost the same results. On the other hand, if the objective is to select the best design with minimum $E(OC_m)$, then the algorithm behaves better when the expected opportunity cost EOC is used as a stopping rule. Moreover, increasing the samples in the multiple replications M increases the performance of the algorithm under the second stopping rule. However, it is clear that when the value of the significant level ε is decreased, then the $E(OC_m)$ will decrease, but we get the optimal value for the $E(OC_m)$ which is 0 when the $\varepsilon = 0$. In this case the mean of the selected design will be equal to the mean of the actual best design. Nevertheless, we cannot take too small value of ε since this will require that the number of multiple replications M to be increased, and, of course, this leads to a huge computational time.

5 Conclusion

In this paper, we have discussed the effect of two stopping rules on the performance of a sequential selection procedure that is used to select a set of good enough simulated designs when the number of alternatives is very large. These two rules include the sequential S stopping rule and the expected opportunity cost EOC stopping rule. We have applied these rules on two different examples.

The results obtained from the numerical applications of the procedure using the two stopping rules indicate that to improve the efficiency of the approach using the EOC stopping rule, we need to increase the number of multiple replications M. We conclude that if the objective of the experiment is to select the best designs with high $P(CS_m)$, then both stopping rules give almost the same performance. However, if the objective is to select the best designs with minimum $E(OC_m)$, then the second stopping rule that uses EOC gives better performance.

References

[1] D. He, S.E. Chick, and C.H. Chen, Opportunity Cost and OCBA Selection Procedures in Ordinal Optimization for a Fixed Number of Alternative Systems. *IEEE Transactions on Systems* **37** (2007) 951–961.

[2] J. Branke, S.E. Chick and C. Schmidt. Selecting a Selection Procedure. *Management Science* **53** (12) (2007) 1916–1932.

[3] R.E. Bechhofer, T.J. Santner and D.M. Goldsman. *Design and Analysis of Experiments for Statistical Selection, Screening, and Multiple Comparisons.* New York, Wiley, 1995.

[4] A.M. Law and W.D. Kelton. *Simulation Modeling and Analysis.* New York, McGraw-Hill, 2000.

[5] S.H. Kim and B.L. Nelson. Selecting the best system. In: *Handbooks in Operations Research and Management Science*, Chapter 17. Elsevier, 2006, 501–534.

[6] S.H. Kim and B.L. Nelson. On the asymptotic validity of fully sequential selection procedures for steady-state simulation. *Operations Research* **54** (3) (2006) 475–488.

[7] J.D. Gibbons, I. Olkin and M. Sobel. *Selecting and Ordering Populations: a New Statistical Methodology.* Society for Industrial and Applied Mathematics, 1999.

[8] S.S. Gupta and S. Panchapakesan. *Multiple Decision Procedures: Theory and Methodology of Selecting and Ranking Populations*, Wiley, 1979.

[9] Y.C. Ho, R.S. Sreenivas and P. Vakili. Ordinal optimization of DEDS. *Discrete Events Dynamic Systems* **2** (1992) 61–88.

[10] M.H. Alrefaei and M.H. Almomani. Subset selection of best simulated systems. *Journal of the Franklin Institute* **344** (2007) 495–506.

[11] M.H. Almomani and M.H. Alrefaei. A three-stage procedure for selecting a good enough simulated system. *Journal of Applied Probability and Statistics* **6** (1&2) (2012) 13–27.

[12] M.H. Almomani and R. Abdul Rahman. Selecting a good stochastic system for the large number of alternatives. *Communications in Statistics-Simulation and Computation* **41** (2012) 222–237.

[13] M. Al-Salem, M.H. Almomani, M.H. Alrefaei and Ali Diabat. On the optimal computing budget allocation problem for large scale simulation optimization. *Simulation Modelling Practice and Theory* **71** (2017) 149–159.

[14] M.H. Almomani, M.H. Alrefaei and S. Al Mansour. A combined statistical selection procedure measured by expected opportunity cost. *Arabian Journal for Science and Engineering* **43** (6) (2018) 3163–3171, DOI: 10.1007/s13369-017-2865-8.

[15] M.H. Almomani, M.H. Alrefaei and S. Al Mansour. A method for selecting the best performance systems. *International Journal of Pure and Applied Mathematics* **120** (2) (2018) 191–202, DOI: 10.12732/ijpam.v120i2.3.

[16] M.H. Alrefaei, M.H. Almomani and S.N. Alabed Alhadi. Asymptotic approximation of the probability of correctly selecting the best systems. In: *AIP Conference Proceedings* **1991** (1) (2018) 020022, AIP Publishing LLC, DOI: https://doi.org/10.1063/1.5047895.

[17] C.H. Chen, D. He, M. Fu and L.H. Lee. Efficient simulation budget allocation for selecting an optimal subset. *Informs Journal on Computing* **20** (4) (2008) 579–595.

[18] M.H. Almomani and M.H. Alrefaei. Ordinal optimization with computing budget allocation for selecting an optimal subset. *Asia-Pacific Journal of Operational Research* **33** (2) (2016), DOI: 10.1142/S0217595916500093.

[19] Y.C. Ho, Q.C. Zhao and Q.S. Jia. *Ordinal Optimization: Soft Optimization for Hard Problems*. Springer, 2007.

[20] S.C. Horng and S.S. Lin. Bat algorithm assisted by ordinal optimization for solving discrete probabilistic bicriteria optimization problems. *Mathematics and Computers in Simulation* **166** (2019) 346–364.

[21] T. Ma, F. Tian and B. Dong. Ordinal optimization-based performance model estimation method for HDFS. *IEEE Access* **8** (2019) 889–899.

[22] C.H. Chen, E. Yucesan and S.E. Chick. Simulation Budget Allocation for Further Enhancing the Efficiency of Ordinal Optimization. *Discrete Event Dynamic Systems* **10** (3) (2000) 251–270.

[23] C.H. Chen and L.H. Lee. Stochastic Simulation Optimization: An Optimal Computing Budget Allocation, Chapter 5. World Scientific Publishing Company, 2010.

[24] D. Goldsman, S.H. Kim, W.S. Marshall and B.L. Nelson. Ranking and selection for steady-state simulation: Procedures and perspectives. *INFORMS Journal on Computing* **14** (1) (2002) 2–19.

[25] M.H. Almomani, R. Abdul Rahman, A. Baharum and M.H. Alrefaei. A Selection Approach for Solving the Buffer Allocation Problem. *International Journal of the Physical Sciences* **7** (3) (2012) 413–422.

[26] C.T. Papadopoulos, M.E.J. O'kelly, M. J. Vidalis and D. Spinellis. *Analysis and Design of Discrete Part Production Lines*. Springer, 2009.

[27] M. Yuzukirmizi and J.M. Smith. Optimal buffer allocation in finite closed networks with multiple servers. *Computers & Operations Research* **35** (2008) 2579–2598.

[28] M.H. Alrefaei and S. Andradóttir. Discrete stochastic optimization using variants of the stochastic ruler method. *Naval Research Logistics* **52** (4) (2005) 344–360.