Proceedings of the 2023 Winter Simulation Conference
C. G. Corlu, S. R. Hunter, H. Lam, B. S. Onggo, J. Shortle, and B. Biller, eds.

TOP-M FACTOR SCREENING FOR STOCHASTIC SIMULATION: MULTI-ARMED BANDIT AND SEQUENTIAL BIFURCATION COMBINED

Wen Shi
Xiang Xie

Business School
Central South University
No. 932 South Lushan Road
Changsha 410083, CHINA

Hong Wan

Department of Industrial and Systems Engineering
North Carolina State University
915 Partners Way
Raleigh, NC 27607, USA

ABSTRACT

We propose a novel screening framework (abbreviated to Top$m$SB) to identify the top m key factors affecting the system performance. The new framework extends the standard SB's screening mechanism while in each stage marrying with an adaptive multi-armed bandit (MAB) procedure in order of the largest group. Compared to SB, Top$m$SB avoids specifying perplexing (un)importance threshold parameters, while providing desired computational efficiency and statistical precision guarantee. Numerical experiments demonstrate the efficiency and effectiveness of the proposed method.

1 INTRODUCTION

Factor screening, also known as screening, is the methodological process of identifying significant factors that have a notable effect on the output of a model of interest (Kleijnen 2015). Among the many screening methods, group screening methods are frequently employed in situations where a large number of factors are present.

The present study focuses on a widely used group-based screening technique called *sequential bifurcation* (SB), which is considered to be highly effective both statistically and computationally in deterministic and stochastic simulation experiments when certain assumptions are met (Ankenman et al. 2014). The theoretical foundation for SB in the deterministic simulation environment was first introduced by Bettonvil and Kleijnen (1997), while the majority of subsequent literature has focused on its application in the context of stochastic simulation experiments. Early versions of SB in stochastic simulation experiments used the Student t-statistic (Cheng 1997; Kleijnen et al. 2006) to assess the significance of each group, employing a fixed simulation budget for each test due to the presence of simulation noise. Controlled SB (CSB), on the other hand, utilizes a sequential testing mechanism in conjunction with the underlying SB process, allowing the user to simultaneously control Type-I and Type-II error rates (1-power) under conditions of heterogenous variance (Wan et al. 2006; Wan et al. 2010). To relax the assumptions underlying SB and enhance computational efficiency, CSB is integrated with fractional factorial and controlled sequential factorial designs (Sanchez et al. 2009; Shen et al. 2010). More recent research has investigated the issue of splitting within SB via a Bayesian dynamic programming framework (Frazier et al. 2012), as well as the effects of dispersion (Ankenman et al. 2014) and multiple responses (Shi et al. 2014; Wang and Wan 2014). Additionally, robust SB methods have been proposed by Xie et al. (2023) and Liu et al. (2019) in response to data contamination.

Although the state-of-the-art SB has the capability to provide accurate error control for identifying a single group or a factor, it is subject to two limitations. First, from a user's point of view, specifying the value of (top) m is more convenient than specifying (Δ_0, Δ_1), the unimportance and importance thresholds.

979-8-3503-6966-3/23/$31.00 ©2023 IEEE

The latter requires prior knowledge and an understanding of the specific problem at hand; see e.g., Wan et al. (2006), Wan et al. (2010), Shi et al. (2014), and Kleijnen (2015). Second, the testing procedure employed by existing SB primarily focuses on individual group effects, and therefore, does not provide error control guarantees for the overall screening procedure. Unfortunately, the absence of such a procedure can lead to poor screening identification results. In this regard, we propose a novel screening framework, referred to as Top$m$SB, which integrates stochastic multi-armed bandit (MAB) and SB. The objective of Top$m$SB is to effectively identify the top m influential factors through the utilization of group screening. The merit of Top$m$SB is that it allows users to avoid setting complex parameters in simulation factor screening while ensuring desired error control properties. Nonetheless, the abilities require Top$m$SB pairing with (i) a new step-down bifurcation mechanism, and (ii) a testing procedure for various group effects generated. Our main contribution is to fully address these two aspects of the research question.

The rest of this paper is organized as follows. Section 2 reviews the standard SB procedure. Section 3 presents the proposed SB screening method. Section 4 describes the instance sets and provides the numerical results. Section 5 concludes this paper and suggests future research directions.

2 A REVIEW ON SB

In this section, we provide a brief review on sequential bifurcation (SB), including its underlying setups and error control mechanism for the stochastic simulation.

2.1 Basic Setup

Suppose that a simulation model involves a total of K factors, $\mathbf{x} = (x_1, x_2, ..., x_K)^\top$, that affects the simulation output $\mathcal{Y}(\mathbf{x})$. Due to the model complexity, simulation analysts often treat their model as a *black box* and only input/output (I/O) data can be observed. This simulation can define an implicit and complex mathematical I/O function, which is often approximated by a simpler and explicit mathematical function, colloquially referred to as a *metamodel* (*surrogate* or *emulator*) (Kleijnen and van Beers 2022).

SB is a metamodel-based screening method and it assumes the following lower-order polynomial implied by $\mathbf{x}$ and $\mathcal{Y}(\mathbf{x})$ (Borgonovo and Plischke 2016):

$$\mathcal{Y}(\mathbf{x}) = \beta_0 + \sum_{j=1}^{K} \beta_j x_j + \sum_{j=1}^{K} \sum_{j'=j}^{K} \beta_{j;j'} x_j x_{j'} + \varepsilon(\mathbf{x}), j = 1, 2, ..., K. \tag{1}$$

where β_j is the main effect of factor j, and $\beta_{j;j'}$ is the interaction effect between factor j and j', for $j = 1, 2, ..., K$ and $j < j'$; $\varepsilon(\mathbf{x})$ denotes the random error inherent in the stochastic simulation model, which is typically assumed to follow a normal distribution with zero mean and unknown variance $\sigma_\varepsilon^2(\mathbf{x})$, probably depending on $\mathbf{x}$ (Wan et al. 2006; Wan et al. 2010; Shi et al. 2014); x_j is standardized to take values in $[-1, 1]$ such that the effects of different factors on the output $\mathcal{Y}$ are on the same scale.

SB is regarded as the most computational efficient screening method for stochastic simulation when the following assumption is satisfied (Ankenman et al. 2014).

Assumption 1 The signs of first-order effects on the output are known; i.e., either $\beta_j \geq 0$ or $\beta_j \leq 0$ is known for factor $j = 1, 2, ..., K$.

This assumption is a fundamental assumption for group-screening methods (including SB) (Shi et al. 2014). Its purpose is to avoid the cancellation of main effects when aggregating different factors into a group. In practice, this assumption may indeed hold. For example, in a $M/M/1$ queuing system, the arrival rate (x_1) has a positive effect on the average customer waiting time (i.e., $\beta_1 \geq 0$), while the service rate (x_2) has a negative effect on the waiting time (i.e., $\beta_2 \leq 0$). That is, an increase in customer arrival rate (service rate) makes the average waiting time for customers increase (decrease) or remain the same.

Definition 1 Let $\beta_{j\frown j'}$ denote the *group effect* (*aggregated main effect*) of factor j through j'

$$\beta_{j\frown j'} = \sum\nolimits_{\jmath=j}^{j'} \beta_\jmath,$$

which is the sum of main effects of the group.

Lemma 1 If Assumption 1 holds, the following unbiased estimator can be used to approximate the group effect $\beta_{j\frown j'}$:

$$\begin{aligned}\overline{\widehat{\beta}}_{j\frown j'} &= \frac{1}{R}\sum_{r=1}^{R}\widehat{\beta}_{j\frown j',r} \\ &= \frac{1}{R}\sum_{r=1}^{R}\frac{\left(w_{j',r}-w_{-j',r}\right)-\left(w_{(j-1),r}-w_{-(j-1),r}\right)}{4}, r=1,2,...,R,\end{aligned} \tag{2}$$

where R is the number of simulation replications; $w_{j',r}$ denotes the output of the rth simulation replication when factors 1 through j' are at their high levels and the remaining factors are at their low levels; $w_{-j',r}$ denotes the rth repeated output when factor 1 through j' are at their low levels and the left are at their high levels.

$w_{j',r}$, $w_{-j',r}$, $w_{(j-1),r}$ and $w_{-(j-1),r}$ are observable noisy outputs of a simulation model in replication r across different factor settings. When factor j is at a high level, it means $x_j = 1$; otherwise, it means $x_j = -1$. $w_{-j',r}$ is always called the *mirror* observation of $w_{j',r}$ (see e.g., Kleijnen 2015, pp. 148-149), which applies the well-known *foldover* design, originally developed for physical experiments (see Montgomery 2013, pp.141-143). It is not difficult to see that applying the foldover design, the second-order effects, $\beta_{j,j'}$'s in Equation (1), do not bias the first-order estimator in Equation (2).

The standard SB consists of a finite number of stages, the objective of which is to identify the significant factors with important main effects. To initialize, SB aggregates all K factors into a group. Then, SB tests whether the aggregation effect of this largest group, $\beta_{1\frown K}$, is important; if the test shows that $\beta_{1\frown K}$ is unimportant, which means that any individual factor in the group has no important effect on the output, then SB stops; otherwise, it indicates that individual important factor(s) may be present and SB splits the group into two subgroups; this splitting is called *bifurcation*. For each subgroup, SB tests whether the two subgroups are important, and removes the unimportant subgroup(s) and further splits the important one(s). In each stage, SB repeats two steps—tests the importance of existing groups attributable to current stage and conducts bifurcation for important groups tested—until the number of factors in a group decreases to size one, which we can determine the importance of individual factors.

During the implementation of SB, it is of particular importance to test whether a group $\beta_{j\frown j'}$ is important in the early stages and a single factor β_j is important in the last stage. This is because in the presence of simulation random noise, SB may make two types of errors: (1) Type-I error: truly unimportant groups (or factors) are identified important, and (2) Type-II error: those truly important groups (or factors) are identified unimportant. If an error is made at any iterative stage, the final identification results of SB may be biased. To address this, the advanced SB employ a sequential testing procedure (Wan et al. 2006; Wan et al. 2010; Shi et al. 2014; Liu et al. 2019). One distinct advantage of this SB is that, when testing a group or a factor, it can automatically determines the simulation cost to use while providing Type-I and -II error control guarantees. This SB specifies the Type-I and -II error rates, denoted by $(\alpha_{\mathrm{SB}}, \beta_{\mathrm{SB}})$, and the combination of unimportance and importance thresholds, denoted by (Δ_0, Δ_1).

In essence, the sequential testing procedure is *fixed-confidence* type. However, in practice, the simulation analyst using fixed-confidence SB may have difficulty in specifying a suitable value for (Δ_0, Δ_1) due to lack of sufficient industrial experience and understanding of the model. As a result, some extreme situations are likely to be encountered where either quite a lot of factors or no factor are identified. To overcome this challenge, this study turns to the SB employing a *fixed-budget* procedure, with the aim to free simulation analysts from setting complex parameters while ensures a high statistical accuracy of identifying important factors over the whole SB procedure. We notice that both fixed-confidence and fixed-budget procedures are equally important in stochastic simulation area and fixed-budget SB is ignored. Borrowing the idea of best arm identification of stochastic multi-armed bandit (MAB) developed in reinforcement learning, we propose a novel SB-based framework for identifying the top-m factors with important main effects. Obviously, the specification of $m \in \{1, 2, ..., K\}$ is more intuitive and convenient than (Δ_0, Δ_1).

3 THE FRAMEWORK OF TOPMSB

In this section we propose a SB-based factor screening framework, aiming at identifying the top-m important factors (abbreviated to Top$m$SB). Top$m$SB inherits the top-down sequential property of standard SB, but at each stage it adopts a new identification procedure that incorporates stochastic multi-armed bandit (MAB) of reinforcement learning.

3.1 Introduction to the Top$m$SB Framework

For notation convenience, denote the index set of the K factors by $\mathcal{I}$. Top$m$SB focuses on identifying the largest $m \in \{1, 2, ..., K\}$ factors of main effects out of the K factors, where m is pre-specified by the simulation analyst. Further denote the set of the top m factors by $\mathcal{I}_m \subseteq \mathcal{I}$. As in the standard SB, Top$m$SB is composed of a series of iterative stages, in which the important group is divided into two subgroups at each stage. It differs from the standard SB in that in a stage, Top$m$SB only identifies the group with the largest aggregated main effect from the set of *active* groups and only splits that group; the remaining groups, along with the newly split (sub)groups, are considered active and retained to next stage. In Top$m$SB, the set of active groups (*active group set*) is constantly updated, and the group with the largest aggregated main effect is identified from that set. As the experiment proceeds, it is certain that there would be a group with size only one (i.e., a single factor) identified as the largest group from the set of active groups. This factor is then categorized into $\mathcal{I}_m$. Top$m$SB repeats this procedure until the m-th factor is determined. For easy understanding, we provide the following illustrative example.

Example: Suppose that $K = 5$ and $\{\beta_1, \beta_2, \beta_3, \beta_4, \beta_5\} = \{5, 4, 3, 2, 1\}$. When m is specified to 3, the simulation analysts expect to find the first three important factors; i.e., $\mathcal{I}_m = \{\beta_1, \beta_2, \beta_3\}$. Table 1 illustrates the iterative process of the Top$m$SB procedure in this example. Since there are five factors in total, stage 1 contains two subgroups (i.e., $\{\beta_1, \beta_2, \beta_3\} = \{5, 4, 3\}$ and $\{\beta_4, \beta_5\} = \{2, 1\}$, and the active set is $\{\{\beta_1, \beta_2, \beta_3\}, \{\beta_4, \beta_5\}\}$. Top$m$SB now needs to identify the subgroup with the larger group effect. Since $\beta_{1\sim3} = 12 > \beta_{4\sim5} = 3$, Top$m$SB splits the group $\{\beta_1, \beta_2, \beta_3\}$ into two subgroups, $\{\beta_1, \beta_2\}$ and $\{\beta_3\}$. Then at stage 2, Top$m$SB encounters the active group set containing three subgroups, $\{\beta_1, \beta_2\} = \{5, 4\}$, $\beta_3 = 3$ and $\{\beta_4, \beta_5\} = \{2, 1\}$, where $\{\beta_1, \beta_2\}$ and β_3 are the two groups newly generated by the bifurcation conducted in Stage 1, and $\{\beta_4, \beta_5\}$ is the un-split subgroup inherited from previous stage. Similar to Stage 1, Top$m$SB determines the largest group from $\{\{\beta_1, \beta_2\}, \beta_3, \{\beta_4, \beta_5\}\}$. Because of $\beta_{1\sim2} = 9 > \beta_3 = \beta_{1\sim2} = 3$, Top$m$SB performs a bifurcation on $\{\beta_1, \beta_2\}$ and obtain two subgroups, β_1 and β_2. At stage 3, Top$m$SB further reaches a active group set $\{\beta_1, \beta_2, \beta_3, \{\beta_4, \beta_5\}\}$, and repeats the same procedure as it did in the previous stages and determines the subgroup β_1 is the largest. Given that β_1 is a single factor that cannot be split, it is categorized into the set $\mathcal{I}_m$, and the set is updated to $\{\beta_2, \beta_3, \{\beta_4, \beta_5\}\}$ for next stage. Following the similar operations, Top$m$SB finds β_2 and β_3 at stage 4 and 5, respectively.

In order to tease out the largest group, at each stage Top$m$SB requires simulation overhead. Let $\mathcal{C} \in \mathbb{Z}^+$ denote the simulation budget expended to identify the largest group at each stage. A complete algorithmic framework of Top$m$SB is provided in Algorithm 1. According to Algorithm 1, each stage of Top$m$SB contains two implementation steps. Two technical challenges are involved in Step 1, which is at the core of the algorithm. First, since the simulation models are stochastic in nature, the group effects observed are random; therefore, finding the group with the highest mean value from a set of stochastic group effects is

Table 1: Example of an iterative process for the Top$m$SB when m is specified to 3.

Iterative stage	Set of active groups	Top-m factors
Stage 1	$\{\beta_{1\sim 3}, \beta_{4\sim 5}\}$	—
Stage 2	$\{\beta_{1\sim 2}, \beta_3, \beta_{4\sim 5}\}$	—
Stage 3	$\{\beta_1, \beta_2, \beta_3, \beta_{4\sim 5}\}$	β_1
Stage 4	$\{\beta_2, \beta_3, \beta_{4\sim 5}\}$	β_2
Stage 5	$\{\beta_3, \beta_{4\sim 5}\}$	β_3
End		

challenging. Second, since Top$m$SB consists of numerous iterations, how can we ensure a relatively fair experimental budget for each group throughout Top$m$SB? In this study, we cast these two technical issues as a best-arm identification of MAB, and further considers two simulation budget allocation strategies: uniform allocation (UA) and adaptive allocation (AA) with the framework. The details is deferred to Subsection 3.2.

Algorithm 1: Algorithm framework for Top$m$SB

Input: $m \in \{1, 2, ..., K\}$, stage-level budget $\mathcal{C}$;
Initialization: $\mathcal{I}_m = \oslash$**, and the set of groups for** $i = 1$ **is** $\mathcal{A}_i := \left\{\beta_{1\backsim\lfloor\frac{1+k}{2}\rfloor}, \beta_{\lfloor\frac{1+k}{2}\rfloor+1\backsim K}\right\}$
  While $t \leqslant m$
    Step 1: Determine the simulation budget allocation strategy for the group with the largest effect value in $\mathcal{A}_i$ **at the** i**-th stage**
      (1) **Uniform allocation (UA)**: Allocate the simulation $\mathcal{C}$ budget equally to the newly generated groups at stage i;
      (2) **Adaptive allocation (AA)**: Allocate the simulation $\mathcal{C}$ budget adaptively (unevenly) to the newly generated groups at stage i.
    End
    Step 2: Perform a bifurcation on the group with the largest effect value β_i^***.**
      if β_i^* *only contains a factor* **then**
        $\mathcal{I}_m = \mathcal{I}_m \cup \{\beta_i^*\}$, $\mathcal{A}_i = \mathcal{A}_i \backslash \beta_i^*$, $m \leftarrow m + 1$;
        break;
      end
      else
        Split β_i^* into two smaller subgroups and add them to $\mathcal{A}_i$, $i \leftarrow i + 1$
      end
    End
  End while
  Output: $\mathcal{I}_m$
End algorithm

3.2 Mathematical Details

3.2.1 Determining the Largest Group

The purpose of this subsection is to answer the first technical question raised at the end of Subsection 3.1. Recall that one of the key issues in implementing Top$m$SB at each stage is finding the group with the largest aggregated effect. Without loss of generality, suppose that the number of active groups at stage i

is $N_i \in \mathbb{Z}^+$, and the set made up with these N_i groups is defined as

$$\mathcal{B}_i := \{\beta_{i,1}, \beta_{i,2}, ..., \beta_{i,N_i-2}, \beta_{i,N_i-1}, \beta_{i,N_i}\},$$

where the last two groups, $\{\beta_{i,N_i-1}, \beta_{i,N_i}\}$, represent the two subgroups newly generated by the bifurcation at stage $i-1$. At stage i, Top$m$SB needs to identify the group with the largest aggregated effect, namely, $\beta_i^* := \max\{\beta_{i,1}, \beta_{i,2}, ..., \beta_{i,N_i}\}$, and split it into two subgroups. Because of the randomness of the estimation of the group effect $\beta_{i,n}$ ($n = 1, 2, ..., N_i$), the primary problem of Top$m$SB is to distinguish β_i^* from the N_i noisy groups.

In this paper, we consider it a *best arm identification* issue, a variant of multiarmed bandits (MAB) problems. With a predetermined simulation (sampling) budget, we identify the largest group from the current active group set in order to minimize the likelihood of false selections. The traditional MAB problem involves sampling one arm at a time and receiving a reward (group effect). The key of MAB is to make a tradeoff between exploration (i.e., sampling different active groups) and exploitation (concentrating on the largest group). The standard objective of MAB problems is to minimize the cumulative sum of regret values (see, e.g., Agrawal 1995; Auer et al. 2002), but we focus on identifying the best arm (i.e., largest active group). In this study, we use the successive rejections (SR) procedure from Audibert and Bubeck (2010) as the underlying block of our the MAB-based method. Our reason for selecting SR is that, SR is parameter-free and, with the $\mathcal{C}$ used, it can guarantee that the largest group will be found with fast convergence rate. Specifically, as the experimental budget $\mathcal{C}$ increases, the probability that SR will wrongly find the largest group converges exponentially to 0. The second advantage of SR is that it is easy to implement and does not require much prior information about $\{\beta_{i,1}, \beta_{i,2}, ..., \beta_{i,N_i}\}$.

For stage i containing N_i active groups, SR proceeds $(N_i - 1)$ experimental *phases*, each of which allocates a certain amount of simulation budget from $\mathcal{C}$ to the groups in the current active group set, to run simulation experiments. At the end of each phase, the group with the lowest estimated aggregated effect, and therefore the least likely to be β_i^*, is removed from the active group set. There is only one group remaining after $(N_i - 1)$ iterative phases, and it is claimed to be the largest group of stage i.

In phase p, the number of active groups which have not been removed yet is $(N_i + 1 - p)$. SR provides a budget allocation strategy that assigns a portion of stage-wise $\mathcal{C}$ to the $(N_i + 1 - p)$ groups attributed to phase p, $p = 1, 2, ..., N_i - 1$. Let C_p denote the cumulative budget allocated to a group up to phase p, so the budget allocated to a group in phase p is $(C_p - C_{p-1})$, where C_p is specified as (see Audibert and Bubeck 2010)

$$C_p = \left\lceil \frac{1}{\overline{\log}\,(N_i)} \frac{\mathcal{C} - N_i}{N_i + 1 - p} \right\rceil, \; p = 1, 2, ..., N_i - 1, \tag{3}$$

where $\overline{\log}\,(N_i) = 1/2 + \sum_{n=2}^{N_i} 1/n$ and $C_0 = 0$. It is easy to show that $\sum_{p=1}^{N_i-1} (C_p - C_{p-1})(N_i + 1 - p) \lesssim \mathcal{C}$ where $(C_p - C_{p-1})(N_i + 1 - p)$ is the total budget allocated to the $(N_i + 1 - p)$ active groups in phase p of stage i.

3.2.2 Specification of the Two Budget Allocation Strategies

This subsection addresses the second technical question raised in Section 3.1, namely, the allocation of $\mathcal{C}$ to N_i groups at stage i.

As a straightforward approach, Top$m$SB could adopt the SR allocation strategy, which would allocate $(C_p - C_{p-1})$ to each of the $(N_i + 1 - p)$ groups in phase p. However, this allocation mechanism is inefficient in Top$m$SB's implementation framework. This is due to the top-down sequential procedure of Top$m$SB. Compared to old groups that have accumulated a certain amount of simulation budget over the previous stages, the newly generated two groups has no budget allocated at the outset of stage i. In this way, the old groups will always have more budget to estimate their aggregated effects than the new ones, making it easier to misidentify the largest group. We continue the example given in Table 1. Stage 2 consists of three groups: $\{\beta_1, \beta_2\}$, $\{\beta_3\}$, and $\{\beta_4, \beta_5\}$, and $\{\beta_3\}$ and $\{\beta_4, \beta_5\}$ are kept at stage 3 because their aggregated effects are smaller than that of $\{\beta_1, \beta_2\}$. When accessing stage 3, $\{\beta_3\}$ and $\{\beta_4, \beta_5\}$ have allocated budget, while the two newly groups (factors) $\{\beta_1\}$ and $\{\beta_2\}$ have zero budget. As a result, using the budget allocation strategy of SR would be more easy to make correct estimation for old groups but less accurate for the new groups.

Formally, denote the allocated budget for the old (N_i-2) groups $\{\beta_{i,1}, \beta_{i,2}, ..., \beta_{i,N_i-2}\}$ by $\{C_{i,1}, C_{i,2}, ..., C_{i,N_i-2}\}$, and therefore the two newly generated groups $\{\beta_{i,N_i-1}, \beta_{i,N_i}\}$ are $\{0, 0\}$, at the outset of stage i. Top$m$SB is now faced with a unique budget allocation problem of how to allocate the available $\mathcal{C}$ upon observing $\{C_{i,1}, C_{i,2}, ..., C_{i,N_i-2}, 0, 0\}$, with the aim of identifying β_i^*. Top$m$SB adopts two budget allocation strategies to overcome this problem. In the rest of this subsection, we elaborate on technical details of these two strategies.

- *Uniform allocation strategy* (UA), in which case the simulation budget $\mathcal{C}$ is equally partitioned to the two newly generated groups (i.e., $\{\mathcal{C}/2, \mathcal{C}/2\}$), with allotting no budget to the old groups. Clearly, this strategy is a relatively fair because it provides the exactly same simulation budget $\mathcal{C}/2$ to compares the groups' effects for all groups across Top$m$SB iterative stages. Even though UA is intuitive and simple, as we will show it can ensure adequate statistical validity in terms of identifying β_i^*.
- *Adaptive allocation strategy* (AA). Compared with UA strategy, AA strategy seeks to allocate the simulation budget $\mathcal{C}$ to the groups most needed, namely, the group that is more likely to be β_i^*, in the hope of achieving better statistical validity. For this purpose, Top$m$SB resorts to a modified version of SR's budget strategy. Recall that in phase p of stage i, SR applies an equal allocation to the (N_i+1-p) groups, and the total budget used for the groups is $(C_p - C_{p-1})(N_i+1-p)$. In contrast, for this $(C_p - C_{p-1})(N_i+1-p)$, AA follows an unequal budgeting fashion.

Since the groups have already accumulated a certain amount of simulation budget up to phase p ($p = 2, ..., N_i - 1$), for convenience, we denote the sequence of their current budget possessed in a non-increasing order as

$$C_{i,(1)} \geqslant C_{i,(2)} \geqslant ... \geqslant C_{i,(j)} \geqslant ... \geqslant C_{i,(N_i-p)} \geqslant C_{i,(N_i-p+1)}. \tag{4}$$

At phase p, Top$m$SB begins by determining a group j^*, and then assigns simulation budget to the groups that are ranked after $C_{i,(j^*)}$ (including $C_{i,(j^*)}$) in Equation (4), where

$$j^* := \arg\min_j \left(\frac{(C_p - C_{p-1})(N_i+1-p) + \sum_{l=j}^{N_i+1-p} C_{i,(l)}}{N_i - p - j + 2} < C_{i,(j-1)} \right), \tag{5}$$

where $(N_i - p - j + 2)$ in the denominator denotes the final $(N_i - p - j + 2)$ groups starting from j to the last $(N_i + 1 - p)$, and $\sum_{l=j}^{N_i+1-p} C_{i,(l)}$ is the sum of cumulative simulation budgets of the final $(N_i - p - j + 2)$ groups up to phase p. Thus, the left-hand side of the inequality represents the budget for each of the last $(N_i - p - j + 2)$ groups that would be, if the total phase-wise budget $(C_p - C_{p-1})(N_i + 1 - p)$ only allocates to these groups. The rationale behind Equation (5) is twofold. First, according to the original budgeting scheme of SR and the top-down nature of SB, those groups with larger aggregated effects (but not the largest so they do not have a split) will always receive more budget than those with smaller aggregated effects. The severity of this budget inequality increases over stages. In practice, it is always the case that such groups have accumulated sufficient budget over the course of their previous iterative stages and phases. In addition, we need to ensure that the budget allocated to the newly split groups is not too small. Based on the Equation (5), AA strategy will adaptively update the budget of the last $(N_i - p - j^* + 2)$ groups across distinct phases until it finds the optimal group at stage i. Following the allocation of a new simulation budget during phase p, additional simulation replications are performed for the last $(N_i - p - j^* + 2)$ groups. Subsequently, the group effect $\hat{\beta}_{i,l}$'s for $l \in \{j^*, j^*+1, ..., N_i\}$ are re-estimated using Equation (2). Algorithm 2 illustrates the main procedure of Top$m$SB's AA strategy.

Algorithm 2: Procedure for Top$m$SB-AA

Input: groups $\mathcal{A}_i := \{\beta_{i,1}, \beta_{i,2}, ..., \beta_{i,N_i}\}$, stage-level simulation budget $\mathcal{C}$, cumulative simulation budget $\{C_{i,1}, C_{i,2}, ..., C_{i,N_i-2}, 0, 0\}$;
For $p \leftarrow 1$ **to** $N_i - 1$**.**
  (1) Determine j^* via Equation (5);
  (2) Allocate simulation budget to the last $(N_i - p - j^* + 2)$ groups in sequence defined as Equation (4) with a total of $(C_p - C_{p-1})\,(N_i + 1 - p)$, where C_p is specified as Equation (3);
  (3) For $l \in \{j^*, j^*+1, ..., N_i + 1 - p\}$, the simulation budget is allocate to group l is
  $\frac{(C_p - C_{p-1})(N_i+1-p) + \sum_{l=j}^{N_i+1-p} C_{i,(l)}}{N_i - p - j + 2} - C_{i,(l)}$;
  (4) Update the estimated group's effect $\hat{\beta}_{i,l}$'s based on the newly assigned budget for $l \in \{j^*, j^*+1, ..., N_i\}$ in sequence defined as Equation (4);
  (5) Remove the empirical optimal group and update $\mathcal{A}_{i,p+1} = \mathcal{A}_{i,p} \setminus \arg\min_{l \in \mathcal{A}_{i,p}} \hat{\beta}_{i,l}$.
End
Output: $\mathcal{A}_{i,N_i}$

4 NUMERICAL STUDY

In this section we compare the computational efficiency and statistical efficacy of the proposed Top$m$SB to SB, and SAR on detecting important factors. Note that, although SAR is proposed in the context of MAB to find the largest arms in terms of expected mean value, one can directly apply it to our factor screening problem.

Top$m$SB is a black-box method; it runs a factor combination of K inputs $(x_1, x_2, \ldots, x_K)^\top$, and observes the resulting simulation output. Our experiments are built on white box, namely, we set specific values for the coefficients of the second-order polynomial given in Equation (1) and the variance of simulation random noise. We notice that similar large-scale factor screening problems are studied in Wan et al. (2006), Wan et al. (2010), Kleijnen (2015), and recently Liu et al. (2019), Liu et al. (2022). We study a screening problem with $K = 100$ factors. The simulation output generated at a factor combination $\mathbf{x} = (x_1, x_2, \ldots, x_{100})^\top$ satisfying (1), where $\mathbf{x} \in [-1, 1]^{100}$, and the random noise ε is assumed to be a random variable following $\mathcal{N}(0, \sigma_\epsilon^2)$ with σ_ϵ^2 being a constant that controls the variability of simulation errors. Further, the interaction effect coefficients $\beta_{j,j'}$'s are independently sampled from the normal distribution $\mathcal{N}(0, 0.1)$.

We study the performance of Top$m$SB with three budget allocation strategies (i.e., SR (Audibert and Bubeck 2010) and our proposed strategies UA and AA), together with two competitive methods, namely, fixed-confidence SB (Wan et al. 2010), and MAB-based SAR (Bubeck et al. 2013). Inspired by Kleijnen (2015), we consider the following problem characteristics:

- *Variability of important main effects*, in which case important factors take varying values: Important β_j's take small values from $\{2, 3, 4, 5\}$ or large values from $\{2, 4, 6, 8\}$, and unimportant ones take zaro;
- *Clustering of important main effects*: Important main effects are either clustered or scattered;
- *Signal–noise ratio*: The variability of the simulation errors σ_ε is set at 5 or 10; notice that σ_ε is relatively large compared to the magnitudes of main effects;
- *Number of factors to identify*: The number of the largest factors to identify is set to either $m = 4$ or $m = 8$.

With the problem characteristics specified, we consider 16 experimental scenarios as shown in Table 2. Since the running of SAR is based on a fixed computational budget, we set the available total budget allocated to SAR under each experimental scenario is the same as the average runs of Top$m$-AA for convenience of comparison.

To implement the fixed-confidence SB such that only top-m β_j's are expected to be important, we set the unimportance and importance threshold parameters according to the lowest value of top-m β_j's. The choice of Δ_0 and Δ_1 herein is to comprehensively examine the performance of SB, which can be found in similar work such as Shi et al. (2019) and Wan et al. (2010). Specifically, we set $\Delta_0 = 0$ and $\Delta_1 = 2$ for these scenarios with $m = 8$, while $\Delta_0 = 4$ (respectively, $\Delta_0 = 2$) and $\Delta_1 = 6$ (respectively, $\Delta_1 = 4$) for scenarios 3-4 and 11-12 (resp., scenarios 7-8 and 15-16). Besides, we use an initial sample size $N_0 = 25$ specified by Wan et al. (2010) for all subgroups, and set the Type I error and the power requirement respectively at $\alpha_{\mathrm{SB}} = 0.01$ and $\beta_{\mathrm{SB}} = 0.05$.

Each method is applied to independent 1,000 macro-replications under each of the 16 experimental combinations, and the computational efficiency of a method is calculated by their respective average number of simulation runs (i.e., group effects observed) (Wan et al. 2010), which are recorded in the last five columns of Table 2. The following observations are made from Table 2: (1) Both Top$m$SB and SB are heavily depend on the first two characteristics of the simulation model; i.e., the higher and clustered $|\beta_j|$'s are, the less simulation budget is required by Top$m$SB and SB to reach a screening result. (2) SB is more sensitive to random noise than Top$m$SB; compare, for example, the simulation budget used by Top$m$SB and SB in the scenarios 5-6. (3) Top$m$SB and SB need less simulation cost when $m = 4$ compared with $m = 8$. The possible explanation for this is that Top$m$SB goes through fewer stages, and fewer groups are classified important by SB because of the higher threshold. (4) Top$m$SB-UA requires less simulation budget than Top$m$SB-AA, because Top$m$SB-UA only allocates budget to those subgroups generated newly at each stage.

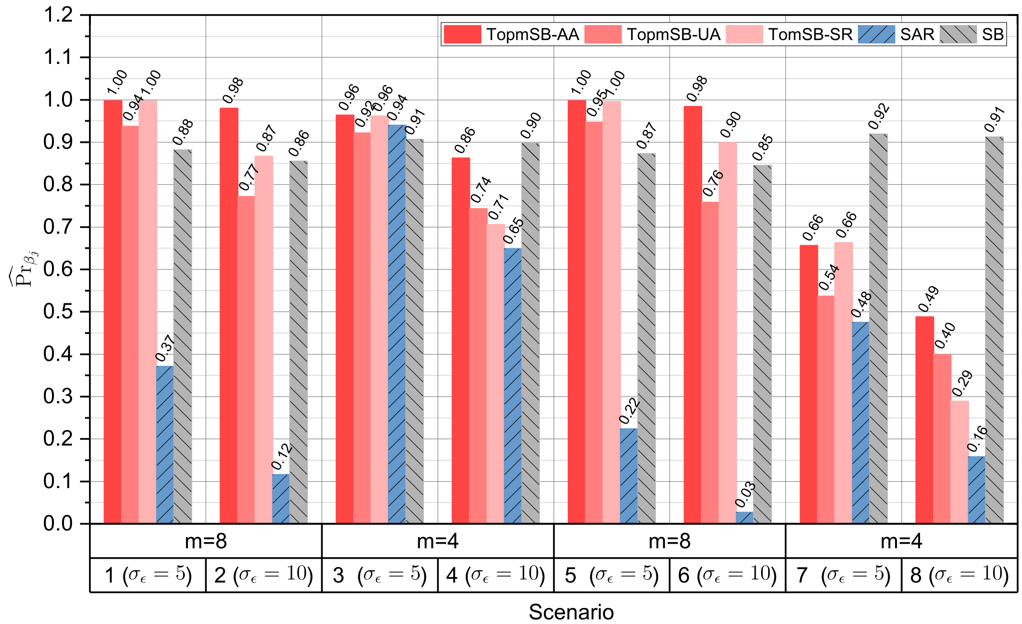
Further, we evaluate the efficacy of each method in identifying top-m main effects for a given experimental combination as the proportion of times that all true top-m main effect are declared important; that is,

$$\widehat{\Pr}_{\beta_j} = \frac{\#\{\{\text{set of top-}m\text{ factors identified}\} = \{\text{set of true top-}m\text{ factors}\}\}}{1,000 \text{ macro-replications}}.$$

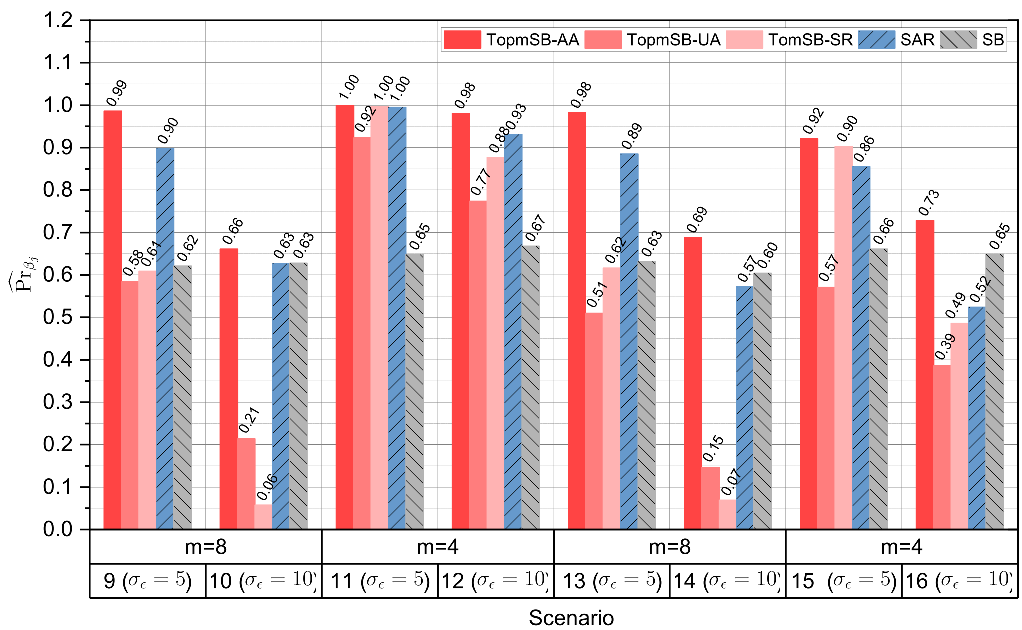
The resulting $\widehat{\Pr}_{\beta_j}$'s obtained by Top$m$SB-AA, Top$m$SB-UA, Top$m$SB-SR, SAR, and SB under the 16 experimental scenarios are shown in Figure 1(a)–(b). The following observations can be made from the figure. (1) In the framework of Top$m$SB, AA strategy significantly outperforms SR and UA strategy by virtue of adaptive budget allocation. (2) Top$m$-AA clearly performs better than SAR constrained by a fixed total budget, especially in the scenarios with scattered important main effects (i.e., scenario 1-8). (3) The smaller simulation error is, the more clustered important factors are, and the greater the difference among the values of the top-m β_j's is, it's easier to obtain a higher $\widehat{\Pr}_{\beta_j}$ by Top$m$SB framework; compare, for example, the $\widehat{\Pr}_{\beta_j}$'s obtained by Top$m$SB-AA, Top$m$SB-UA, Top$m$SB-SR in scenarios 1, 2, and 10. (4) Top$m$SB is slightly superior to SB by comparing efficiency and efficacy synthetically in identifying the top-m factors. Specifically, Top$m$SB can obtain higher $\widehat{\Pr}_{\beta_j}$'s by using basically the same or less simulation budget in the majority of scenarios; see, for example, the scenarios 14-16. Although Scenario 7-8 also shows the benefits of SB's sequential testing procedure, Top$m$SB-AA does not rely on experience-based and custom thresholds. Therefore, we conclude that Top$m$SB provides an ideal statistical performance guarantee for identifying the top-m factors, and even has higher efficacy under the AA strategy.

5 CONCLUSIONS

The present study introduces a novel framework for factor screening that relies on sequential bifurcation (SB) to identify the top-m influential factors (Top$m$SB) that impact the performance of a stochastic simulated system. To achieve this, we have implemented an improved version of the multi-armed bandit (MAB) algorithm that facilitates the identification of the most promising group of factors in each stage while keeping the probability of erroneous identification under control. Future research directions may include



(a) Scenarios 1-8



(b) Scenarios 9-16

Figure 1: The resulting $\widehat{\Pr}_{\beta_j}$'s obtained by Top$m$SB-AA, Top$m$SB–UA, Top$m$SB-SR, SAR, and SB under the 16 experimental scenarios.

Table 2: Scenarios of problem characteristics and simulation costs consumed by Top$m$SB and competitive methods.

				Simulation runs				
				TomSB			MAB	Traditional
Scen.	Important main effects	m	σ	SR	AA	UA	SAR	SB
1	#(1,97)=2 #(2,98)=4 #(3,99)=6 #(4,100)=8	8	5	1,284	1,276	808	1,221	961
2	#(1,97)=2 #(2,98)=4 #(3,99)=6 #(4,100)=8	8	10	1,415	1,319	986	1,262	1,459
3	#(1,97)=2 #(2,98)=4 #(3,99)=6 #(4,100)=8	4	5	744	741	504	691	818
4	#(1,97)=2 #(2,98)=4 #(3,99)=6 #(4,100)=8	4	10	740	742	508	691	1,114
5	#(1,97)=2 #(2,98)=3 #(3,99)=4 #(4,100)=5	8	5	1,280	1,274	801	1,221	986
6	#(1,97)=2 #(2,98)=3 #(3,99)=4 #(4,100)=5	8	10	1,398	1,310	986	1,260	1,548
7	#(1,97)=2 #(2,98)=3 #(3,99)=4 #(4,100)=5	4	5	763	764	528	711	906
8	#(1,97)=2 #(2,98)=3 #(3,99)=4 #(4,100)=5	4	10	762	764	548	711	1,382
9	#(1,70)=2 #(10,80)=4 #(20,90)=6 #(30,100)=8	8	5	4,410	4,372	3,474	4,319	2,417
10	#(1,70)=2 #(10,80)=4 #(20,90)=6 #(30,100)=8	8	10	4,325	4,305	3,345	4,256	4,072
11	#(1,70)=2 #(10,80)=4 #(20,90)=6 #(30,100)=8	4	5	1,487	1,478	1,295	1,423	1,313
12	#(1,70)=2 #(10,80)=4 #(20,90)=6 #(30,100)=8	4	10	1,633	1,518	1,349	1,471	2,042
13	#(1,70)=2 #(10,80)=3 #(20,90)=4 #(30,100)=5	8	5	4,396	4,363	3,389	4,307	2,493
14	#(1,70)=2 #(10,80)=3 #(20,90)=4 #(30,100)=5	8	10	4,338	4,338	3,243	4,283	4,392
15	#(1,70)=2 #(10,80)=3 #(20,90)=4 #(30,100)=5	4	5	1,614	1,607	1,403	1,556	1,527
16	#(1,70)=2 #(10,80)=3 #(20,90)=4 #(30,100)=5	4	10	1,891	1,712	1,485	1,660	2,638

exploring optimal strategies for budget allocation across the different stages of the screening process, as well as the development of more sophisticated MAB procedures to further enhance efficiency.

ACKNOWLEDGMENTS

The work of the first author is partially supported by the National Natural Sciences Foundation of China under Grants No. 71402048 and 72293574, the Hunan Provincial Science Fund for Distinguished Young Scholars No. 2022JJ10084. The work of the corresponding author is partially supported by the Fundamental Research Funds for the Central Universities of Central South University No. 2023ZZTS0263.

REFERENCES

Agrawal, R. 1995. "Sample Mean Based Index Policies by $O(\log n)$ Regret for the Multi-Armed Bandit Problem". *Advances in Applied Probability* 27(4):1054–1078.

Ankenman, B. E., Russell C. H. Cheng, and S. M. Lewis. 2014. "Screening for Dispersion Eeffects by Sequential Bifurcation". *ACM Transactions on Modeling and Computer Simulation* 100(100):1–27.

Audibert, J.-Y., and S. Bubeck. 2010. "Best Arm Identification in Multi-Armed Bandits". In *COLT - 23th Conference on Learning Theory - 2010*.

Auer, P., N. Cesa-Bianchi, and P. Fischer. 2002. "Finite-time Analysis of the Multiarmed Bandit Problem". *Machine learning* 47(2-3):235–256.

Bettonvil, B., and J. P. C. Kleijnen. 1997. "Searching for Important Factors in Simulation Models with Many Factors: Sequential Bifurcation". *European Journal of Operational Research* 96(1):180–194.

Borgonovo, E., and E. Plischke. 2016. "Sensitivity analysis: A Review of Recent Advances". *European Journal of Operational Research* 248(3):869–887.

Bubeck, S., T. Wang, and N. Viswanathan. 2013. "Multiple Identifications in Multi-Armed Bandits". In *Proceedings of the 30th International Conference on Machine Learning*, edited by S. Dasgupta and D. McAllester, Volume 28 of *Proceedings of Machine Learning Research*, 258–265. Atlanta, Georgia, USA: PMLR.

Cheng, R. C. H. 1997. "Searching for Important Factors: Sequential Bifurcation Under Uncertainty". In *Proceedings of the 1997 Winter Simulation Conf*, edited by D. H. Withers, B. L. Nelson, S. Andradóttir, and K. J. Healy, 275–280. Piscataway, NJ: Institute of Electrical and Eletronics Engineers.

Frazier, P. I., B. Jedynak, and L. Chen. 2012. “Sequential Screening: A Bayesian Dynamic Programming Analysis of Optimal Group-splitting”. In *Proceedings of the 2012 Winter Simulation Conference*, edited by C. Laroque, J. Himmelspach, R. Pasupathy, O. Rose, and A. M. Uhrmacher, 1–12. Institute of Electrical and Electronics Engineers, Inc.

Kleijnen, J. P. C. 2015. *Design and Analysis of Simulation Experiments*. 2 ed. New York: Springer US.

Kleijnen, J. P. C., B. Bettonvil, and F. Persson. 2006. “Screening for the Important Factors in Large Discrete-Event Simulation Models: Sequential Bifurcation And Its applications”. In *Screening: Methods for Experimentation in Industry, Drug Discovery, and Genetic*, edited by A. Dean and S. Lewis, 287–307. New York: Springer.

Kleijnen, J. P. C., and W. C. M. van Beers. 2022. “Statistical Tests for Cross-Balidation of Kriging Models”. *INFORMS Journal on Computing* 34(1):607–621.

Liu, L., J.-H. Byun, C. Park, and Y. Ma. 2022. “Modified Sequential Bifurcation for Simulation Factor Screening Under Skew-Normal Response Model”. *Computers & Industrial Engineering* 169:108274.

Liu, L., Y. Ma, C. Park, and J.-H. Byun. 2019. “Robust Sequential Bifurcation for Simulation Factor Screening Under Data Contamination”. *Computers & Industrial Engineering* 129:102–112.

Montgomery, D. 2013. *Design and Analysis of Experiments*. 8 ed. New York: John Wiley & Sons Inc.

Sanchez, S. M., H. Wan, and T. W. Lucas. 2009. “A Two-phase Screening Procedure for Simulation Experiments”. *ACM Transactions on Modeling and Computer Simulation* 19(2):1–24.

Shen, H., H. Wan, and S. M. Sanchez. 2010. “A Hybrid Method for Simulation Factor Screening”. *Naval Research Logistics (NRL)* 57(1):45–57.

Shi, W., X. Chen, and J. Shang. 2019. “An Efficient Morris Method-based Framework for Simulation Factor Screening”. *INFORMS Journal on Computing* 31(4):745–770.

Shi, W., J. P. Kleijnen, and Z. Liu. 2014. “Factor Screening for Simulation with Multiple Responses: Sequential Bifurcation”. *European Journal of Operational Research* 237(1):136 – 147.

Shi, W., J. Shang, Z.-X. Liu, and X.-L. Zuo. 2014. “Optimal Design of the Auto Parts Supply Chain for JIT Operations: Sequential Bifurcation Factor Screening and Multi-Response Surface Methodology”. *European Journal of Operational Research* 236(2):664–676.

Wan, H., B. E. Ankenman, and B. L. Nelson. 2006. “Controlled Sequential Bifurcation: A New Factor-Screening Method for Discrete-Event Simulation”. *Operations Research* 54(4):743–755.

Wan, H., B. E. Ankenman, and B. L. Nelson. 2010. “Improving the Efficiency And Efficacy of Controlled Sequential Bifurcation for Simulation Factor Screening”. *INFORMS Journal on Computering* 22(3):482–492.

Wang, W., and H. Wan. 2014. “Sequential Procedures for Multiple Responses Factor Screening”. In *Proceedings of the 2014 Winter Simulation Conference*, edited by A. Tolk, S. D. Diallo, I. O. Ryzhov, L. Yilmaz, S. Buckley, and J. A. Miller, 745–756: Institute of Electrical and Eletronics Engineers.

Xie, E., C. Park, L. Liu, J. Zhou, and Y. Ma. 2023. “Improving the Efficiency And Efficacy of Robust Sequential Bifurcation Under Data Contamination”. *Communications in Statistics - Simulation and Computation*:In advance.

AUTHOR BIOGRAPHIES

WEN SHI is a Full Professor in Business School at Central South University. He holds a Ph.D. in management science and engineering from Huazhong University of Science and Technology (2013). His research focused on simulation modeling, simulation experiments design and analysis and product recalls. His research work appeared in many international journals including EJOR, INFORMS JOC, NRL, RESS, and TRA. His email address is wenshi@csu.edu.cn.

HONG WAN is an associate professor in the Department of Industrial and Systems Engineering at North Carolina State University. She received her Ph.D. in industrial engineering and management sciences from Northwestern University. Her research focuses on the areas of simulation data analysis, complex system simulation, and blockchain modeling, mechanism design, and application. She is the director of the ISE blockchain lab, and serves as the editor in chief of Journal of Blockchain Research and the associate editor of ACM TOMACS. She is a member of INFORMS, IISE, POMS and IEEE blockchain society. Her email address is hwan4@ncsu.edu and her website is https://www.ise.ncsu.edu/people/hwan4/.

XIANG XIE is a doctoral student in Business School at Central South University. He also obtained his master degree in Business School at Central South University (2022). His main research interests are simulation modeling and sensitivity analysis. His research work was presented orally at many academic conferences including INFORMS 2021 annual meeting and CNAIS 2021 annual meeting. His email address is xiangxiehn@csu.edu.cn.