

Analysis of Replicated Order-of-Addition Experiments

Jianbin Chen^{1,2}, Xue-Ru Zhang¹ and Dennis K.J. Lin²

¹ School of Statistics and Data Science, LPMC & KLMDASR
Nankai University, Tianjin 300071, China

² Department of Statistics, Purdue University, West Lafayette, IN 47907, U.S.A

Received: 7 February 2021; Revised: 18 April 2021; Accepted: 19 April 2021

Abstract

The objective of the order-of-addition (OofA) problem is to find the optimal (addition) order. Existing literature concentrated on the responses of different orders with homoscedasticity. Study was made here for the cases of heteroscedasticity, where the dispersion effects for replicated OofA experiments should be considered. This paper proposes some approaches to speculate optimal orders for the replicated OofA experiment. Based on the pair-wise-order (PWO) model, the obtained orders from the proposed methodologies not only achieve the goal of OofA experiment, but also minimize the standard deviation within the OofA framework. Theoretical support is given under the specific setups. Simulation studies are used to illustrate these methodologies. It is shown that the proposed methods perform well for replicated OofA experiments.

Key words: Constrained optimization; Dual response; Mean square error; Pair-wise-order model.

AMS Subject Classifications: 62K05

1. Introduction

The order-of-addition (OofA) experiment has been found to have wide applications in many areas, such as, bio-chemistry, nutritional science and scheduling problems. The goal of the OofA experiment is to find the optimal order. Suppose the OofA experiment involves m (≥ 2) components. There will be $m!$ intrinsic different orders of adding sequences in the system. Each order π is a permutation of $\{1, \dots, m\}$. It is not affordable to test all the $m!$ orders, especially when m is large (*e.g.*, when $m = 10$, $m! = 10! \approx 3.6$ millions). A (relatively small) subset is desirable to explore the optimal order. Empirical studies show that a random selection is rather inefficient (Zhao, *et al.*, 2020). Thus the design problem arises to choose a subset of all possible orders for searching the optimal order.

For any pair of components i and j , Van Nostrand (1995) proposed the pair-wise-order (PWO) factor. Define

$$I_{ij} = \begin{cases} 1, & \text{if } i \text{ precedes } j; \\ -1, & \text{if } j \text{ precedes } i, \end{cases} \quad (1)$$

as the PWO factor, where i and j are different components. Clearly, there are $q = \binom{m}{2}$ PWO factors, corresponding to all pairs of component orders, and such factors are arranged according to the lexicographic ordering of the components' indices. Denote β_{ij} as the effect to response caused by I_{ij} , then the PWO model is a linear model of

$$y(\pi) = \beta_0 + \sum_{i < j} \beta_{ij} I_{ij}(\pi) + \epsilon, \quad \pi \in \Pi, \quad (2)$$

where y is the response of interest, ϵ is a random error from independent normal distribution $N(0, \sigma^2)$, and Π is the set of all of $m!$ possible orders. There are $p = q + 1$ parameters to be estimated. The PWO model is used in most recent literature. For a comprehensive discussion on OofA experiments, one may refer to Lin and Peng (2019), Peng, *et al.* (2019b), Voelkel (2019), Chen, *et al.* (2020a, 2020b, 2020c), Mee (2020), Winker, *et al.* (2020), Yang, *et al.* (2020), Zhao, *et al.* (2020) and the references therein.

In general, a good solution (optimal order) should be reproducible under various (mostly uncontrollable) environments. It is thus critical to study the stability of the optimal order. Ding, *et al.* (2015) investigated the sequence of drug administration that can impact clinical outcomes. They also showed that the setting of the time interval influences the final cell livability. Hence, the sequence of drug and time interval should be considered simultaneously in the experiment. However, time intervals are not easy to be controlled in practice. An alternative way is treating such factors as noise factors. Assume the relationship between factors (variables) and response y is formulated by $y = f(\pi, \mathbf{z})$, where π is the ordering factor and \mathbf{z} is the vector of noise factors. Specifically, the variation of y stems from both the random error ϵ and the variation of \mathbf{z} . The purpose of this problem is then to find the optimal addition order for achieving the accepted targets, while simultaneously minimizing its standard deviation. The existing analytical OofA methods are not appropriate for this case. The dispersion effect for each OofA experiment should be considered.

For a replicated OofA experiment, the location and the dispersion effects of the OofA experiment are the two interested responses. Solving the replicated OofA experiment can then be considered as solving a dual response surface problem. For such a dual response problem, a good solution is to achieve at some compromise involving these two responses. The dual response surface approach employs two separate models to the mean and standard deviation of response. Vining and Myers (1990) utilized the dual response approach by Myers and Carter (1973), and demonstrated to optimize one response with an acceptable constraint on the value of another response. For more details please refer to Lin and Tu (1995), Copeland and Nelson (1996) and Kim and Lin (1998).

This paper proposes two modified dual response approaches to speculate optimal orders of OofA experiment for balancing two objective functions: (a) achieve the goal of location effect and (b) minimize the dispersion effect. The rest of the paper is organized as follows. Section 2 discusses the dual response surface approach for OofA experiments. A case study for job scheduling problems is discussed in Section 3. The corresponding theoretical support is given in Section 4. Section 5 introduces a simulation study with $m = 10$. The conclusion and discussion are given in Section 6.

2. Proposed Methods

2.1. Review of the analyzing approach for OofA experiment

The PWO model (2) is used as the assumed model. The sign of the true parameter β_{ij} shows the order of components i and j , *i.e.*, the positive sign of β_{ij} represents $i \rightarrow j$ (component i shall proceed component j), when the goal of the objective is “the larger the better”, while the negative sign of β_{ij} represents $j \rightarrow i$. The hypothesis $H_0 : \beta_{ij} = 0$ is used to identify the significant parameters. Each component is regarded as a node in the directed graph. The arranged order “ $i \rightarrow j$ ” indicates an edge from i to j , and an insignificant $\hat{\beta}_{ij}$ (namely, $\beta_{ij} = 0$) implies no edge between i and j . One can generate the corresponding directed graph by sequentially connecting the nodes according to the all significant parameters. Finally, all feasible paths (a small candidate pool of optimal orders) can be obtained from the specifically directed graph. The details of finding optimal order(s) can be found in Chen, *et al.* (2020c).

2.2. Dispersion effects for OofA experiments

We next consider the OofA experiments involves heterogeneous standard deviations. Suppose each OofA experiment with $t \geq 2$ replicates. Let y_{kl} denote the l th replicated experiment (or response) at the k th treatment (or design point), where $l = 1, \dots, t$ and $k = 1, \dots, n$. Define

$$m_k = \frac{1}{t} \sum_{l=1}^t y_{kl}, \quad k = 1, \dots, n,$$

and

$$s_k = \sqrt{\frac{1}{t-1} \sum_{l=1}^t (y_{kl} - m_k)^2}, \quad k = 1, \dots, n.$$

Let $\hat{y}_\mu = (m_1, \dots, m_n)$ and $\hat{y}_\sigma = (s_1, \dots, s_n)$ be the estimator of the local effect (y_μ) and the dispersion effect (y_σ). Here, we apply a dual response surface approach to solve the replicated OofA problem under the PWO model (2). Suppose the fitted location response function is

$$y_\mu = \hat{\alpha}_0 + \sum_{i < j} (\hat{\alpha}_{ij}^c I_{ij}^c + \hat{\alpha}_{ij}^\mu I_{ij}^\mu) + \epsilon_\mu, \quad (3)$$

and the fitted dispersion response is

$$y_\sigma = \hat{\beta}_0 + \sum_{i < j} (\hat{\beta}_{ij}^c I_{ij}^c + \hat{\beta}_{ij}^\sigma I_{ij}^\sigma) + \epsilon_\sigma, \quad (4)$$

where I_{ij}^c denote the variables in both y_μ and y_σ ; I_{ij}^μ and I_{ij}^σ denote the variables only in y_μ and y_σ , respectively. Therefore $I_{ij} = (I_{ij}^c, I_{ij}^\mu, I_{ij}^\sigma)^T$. For the location function, there are three basic situations to be considered: (a) “the target is the best”; (b) “the larger the better”; and (c) “the smaller the better”. For simplicity, we only focus on the case (a), the other two cases can be conducted in a similar manner. For the dispersion function, only the situation for the smaller the better needs to be considered.

2.3. Proposed methods

Here, we propose two methods to tackle the replicated OofA problem: (a) the two-step approach: first minimize the dispersion model, and next achieve the location model closer to the target T ; and (b) the mean square error (MSE) approach.

The PWO model (2) is commonly used as the assumed model. The location model y_μ and dispersion model y_σ are fitted based on the obtained data from the OofA design. For the two-step approach, we choose the appropriate levels for some factors (I_{ij}^μ , not in y_σ) to make y_μ closer to the target value T , and select the level of factor (I_{ij}^c, I_{ij}^σ) to minimize y_σ . Combining these factors setting together, the optimal order(s) is finally obtained. (One could also first select the level of factor (I_{ij}^c, I_{ij}^σ) to minimize y_σ), and then choose the appropriate levels for factor (I_{ij}^μ , not in y_σ) to make y_μ closer to the target value T).

The two-step approach is powerful when the location model y_μ does not have any common factors in the dispersion model y_σ . Otherwise, adjusting the factor level in y_μ may make y_σ undesirable, *i.e.*, the two-step approach may obtain undesirable results when two interesting responses have common factors (see Section 5). In this case, the MSE approach (see below) should be considered.

The MSE criterion allows the location effect closer to the target T , while keeps the minimum standard deviation. The MSE criterion is defined as

$$\text{MSE} = \hat{y}_\sigma^2 + (\hat{y}_\mu - T)^2, \quad (5)$$

where T is the target value, \hat{y}_μ and \hat{y}_σ are the estimates of y_μ and y_σ , respectively. The MSE approach are formally presented in Algorithm 1.

Algorithm 1 The MSE approach

- Step 1** Based on the PWO model, select the best OofA design and conduct OofA experiment with $t \geq 2$ replicates;
- Step 2** Based upon the obtained data, the location model y_μ and dispersion model y_σ are fitted;
- Step 3** Evaluate the MSE value (5) and regard it as the response to be optimized;
- Step 4** Construct the corresponding directed graph according to the active factors in the MSE (5).
- Step 5** Obtain the optimal orders as the output.
-

In Step 3 of Algorithm 1, the MSE value for OofA experiment is considered as the response. In Step 4, we construct the corresponding directed graph according to the significant parameters in MSE (5). Based on the constructed directed graph, the optimal sequences are thus obtained.

3. An Illustrative Example

3.1. Problem formulation

Inspired by the drug experiment of Ding, *et al.* (2015), an illustrative example of OofA problems with heteroscedasticity is considered in this section. The noise factor z is subject to a uniform distribution over $[-1, 1]$. Let $m = 3$, the true relationship between the two kind variables and response $y \in [0, 1]$ be

$$y = 0.5 - 0.1(4I_{12}(\pi) + I_{13}(\pi))z + 0.1I_{23}(\pi) + \epsilon,$$

where $\epsilon \sim N(0, 0.09^2)$. The optimal order (unknown to us) is $2 \rightarrow 1 \rightarrow 3$.

The main purpose of this problem is finding the optimal (stable) order to make y_μ as close to $T = 1$ as possible. For each order, we conduct three replicated experiments and denote the obtained response values by Y_1, Y_2, Y_3 . The order can be converted into PWO factor I_{ij} . Take the first run (row) of Table 1 as an example, $3 \rightarrow 2 \rightarrow 1$, the component 2 precedes component 1, then $I_{12} = -1$. Similarly, we have $I_{13} = -1, I_{23} = -1$. All possible ($3! = 6$) experiments, their resulting replications Y_1, Y_2, Y_3 , the corresponding values of I_{12} , I_{13} and I_{23} for each order, as well as the estimated \hat{y}_μ and \hat{y}_σ are displayed in Table 1.

Table 1: The design and responses of the drug problem

Run	Order	I_{12}	I_{13}	I_{23}	Y_1	Y_2	Y_3	\hat{y}_μ	\hat{y}_σ
1	$3 \rightarrow 2 \rightarrow 1$	-1	-1	-1	0.199	0.010	0.508	0.239	0.251
2	$3 \rightarrow 1 \rightarrow 2$	1	-1	-1	0.731	0.137	0.307	0.392	0.306
3	$2 \rightarrow 3 \rightarrow 1$	-1	-1	1	0.223	0.605	0.385	0.404	0.192
4	$2 \rightarrow 1 \rightarrow 3$	-1	1	1	0.363	0.391	0.423	0.392	0.030
5	$1 \rightarrow 2 \rightarrow 3$	1	1	1	0.573	0.350	0.709	0.544	0.181
6	$1 \rightarrow 3 \rightarrow 2$	1	1	-1	0.332	0.085	0.053	0.157	0.153

3.2. Conventional approach

In the conventional method, the dispersion response y_σ is considered to be a constant; and the local response y_μ is considered as the only response. Without consideration of y_σ , the existing conventional OofA methods aim to find the optimal order based on y_μ . The fitting model for the location effect (y_μ) is,

$$y_\mu = 0.355 + 0.092I_{23}(\pi) + \epsilon_\mu. \quad (6)$$

This problem is essentially an unconstrained optimization problem that aims to make y_μ in (6) as close to $T = 1$ as possible. For the fitting model (6), the sign of the significant parameter β_{23} is positive (“+”), hence, the possible optimal orders are subject to $2 \rightarrow 3$. Model (6) provides no information on the order relative to the component 1. Consequently, three possible orders are: (1) $1 \rightarrow 2 \rightarrow 3$; (2) $2 \rightarrow 1 \rightarrow 3$; and (3) $2 \rightarrow 3 \rightarrow 1$. All orders have the prediction $\hat{y}_\mu = 0.447$ via (6). With the conventional approach, any of these three orders can be considered as the optimal order. Note that these three orders have the same \hat{y}_μ , but different standard deviations \hat{y}_σ (as will be shown next). To obtain the stable order(s), new approach should be employed.

3.3. Proposed approaches

According the data in Table 1, the fitting models for y_μ and y_σ of the replicated response are

$$\begin{aligned} y_\mu &= 0.355 + 0.092I_{23}(\pi) + \epsilon_\mu, \quad \text{and} \\ y_\sigma &= 0.185 + 0.055I_{12}(\pi) - 0.082I_{13}(\pi) + \epsilon_\sigma. \end{aligned} \quad (7)$$

For the two-step approach, the first step sets the level of the adjusting variables $I_{2,3}(\pi)$ as “+1” for closing \hat{y}_μ to $T = 1$. Thus, the possible orders should satisfy the arranged order “2 \rightarrow 3”. The second step is to determine the arranged orders of significant parameters to minimize \hat{y}_σ in (7). Given the constraint $\hat{y}_\sigma \geq 0$, the recommend levels of PWO factors are set as $I_{1,2}(\pi) = -1$ and $I_{1,3}(\pi) = 1$ for minimizing \hat{y}_σ . Hence, the arranged orders are “2 \rightarrow 1” and “1 \rightarrow 3”. Combining those arranged orders together, the optimal order 2 \rightarrow 1 \rightarrow 3 is resulted. This is indeed the true optimal order.

Next, the MSE approach in Algorithm 1 is used to solve the problem. The MSE criterion $(\hat{y}_\mu - T)^2 + \hat{y}_\sigma^2$ is used to find the optimal order. In Step 3, the objective function of this example becomes

$$\min (0.092I_{23}(\pi) - 0.645)^2 + (0.185 + 0.055I_{12}(\pi) - 0.082I_{13}(\pi))^2$$

For Step 4 of Algorithm 1, the level of PWO factors are found to be $I_{1,2}(\pi) = -1$, $I_{1,3}(\pi) = 1$ and $I_{2,3}(\pi) = 1$ for minimizing MSE. Thus, the possible orders should achieve “2 \rightarrow 1”, “1 \rightarrow 3” and “2 \rightarrow 3”. Based on the generated directed graph, the optimal order 2 \rightarrow 1 \rightarrow 3 is obtained. That is identical to the true optimal order.

3.4. Discussion

Using the conventional approach, three orders are resulted. For any OofA problem, y_σ should be taken into account to explore an optimal order. The dispersion effect y_σ mainly stems from the noise factor \mathbf{z} (hard to be controlled accurately in practice). Therefore, the location model and dispersion model are respectively built based on both the control factors and ordering factors (such as PWO factors). The experimental goal is to find an optimal order such that (a) location effect y_μ is closer to the target T ; and (b) dispersion effect y_σ is minimized. The proposed methods (both the two-step method and the MSE approach) yield the same optimal order 2 \rightarrow 1 \rightarrow 3. This may not be the case in general.

Compared with the conventional approach which only considers y_μ (assuming y_σ is a constant), the proposed methods not only focus on the location model y_μ , but also consider the dispersion model y_σ . The dispersion model y_σ adds more restrictions on the possible orders decided by the location model y_μ . For this case, the resulting orders of the MSE approach achieve the target of location model (y_μ), while perform well on dispersion model (y_σ).

4. Theoretical Supports

Denote S_1 and S_2 are the candidate pool of optimal orders for \hat{y}_μ (3) and \hat{y}_σ (4), respectively. Let S be the candidate pool of optimal orders for \hat{y}_μ (3) and \hat{y}_σ (4). Suppose there exist orders $\pi_1 \in S_1$, $\pi_2 \in S_2$ and $\pi \in S$, we have the following theorem.

Theorem 1: For replicated OofA experiments, if we have orders $\pi_1 \in S_1$, $\pi_2 \in S_2$ and $\pi \in S$, then $MSE(\pi) \leq MSE(\pi_1)$ and $MSE(\pi) \leq MSE(\pi_2)$.

Proof: For simplicity, we only discuss the case of “the smaller, the better”. The proofs of the remaining two cases (“the larger, the better” and “the target, the better”) are similar and thus omitted. Recall that the fitted location model is

$$y_\mu = \hat{\alpha}_0 + \sum_{i < j} (\hat{\alpha}_{ij}^c I_{ij}^c + \hat{\alpha}_{ij}^\mu I_{ij}^\mu) + \epsilon_\mu,$$

and the fitted dispersion model is

$$y_\sigma = \hat{\beta}_0 + \sum_{i < j} (\hat{\beta}_{ij}^c I_{ij}^c + \hat{\beta}_{ij}^\sigma I_{ij}^\sigma) + \epsilon_\sigma,$$

where I_{ij}^c denote the variables in both y_μ and y_σ ; I_{ij}^μ and I_{ij}^σ denote the variables only in y_μ and y_σ , respectively.

Case 1. Suppose location model y_μ (3) and dispersion model y_σ (4) have common factors, *i.e.*, $I_{ij}^\mu = I_{ij}^\sigma = 0$ for all $i, j \in \{1, \dots, m\}$. In this case, the optimal order π_1 of y_μ (3) is also the optimal order for y_σ (4). Hence, for $\pi_1 \in S_1$ and $\pi \in S$, we have $MSE(\pi) = MSE(\pi_1)$.

Case 2. Suppose there exist an factor making $I_{kl}^\mu(\pi_1) = -I_{kl}^\sigma(\pi_2) \neq 0$. Note that

$$\begin{aligned} MSE(\pi_1) &= \hat{y}_\mu^2(\pi_1) + \hat{y}_\sigma^2(\pi_1) \\ &= \left[\hat{\alpha}_0 + \sum_{i < j, i \neq k, l; j \neq k, l} (\hat{\alpha}_{ij}^c I_{ij}^c + \hat{\alpha}_{ij}^\mu I_{ij}^\mu) + \hat{\alpha}_{kl}^\mu I_{kl}^\mu \right]^2 (\pi_1) + \\ &= \left[\hat{\beta}_0 + \sum_{i < j, i \neq k, l; j \neq k, l} (\hat{\beta}_{ij}^c I_{ij}^c + \hat{\beta}_{ij}^\sigma I_{ij}^\sigma) + \hat{\beta}_{kl}^\sigma I_{kl}^\sigma \right]^2 (\pi_1) \\ &= C + (\hat{\alpha}_{kl}^\mu + \hat{\beta}_{kl}^\sigma)B, \end{aligned}$$

where B and C are constant with other active standard deviations. Similarly, we have

$$\begin{aligned} MSE(\pi) &= \hat{y}_\mu^2(\pi) + \hat{y}_\sigma^2(\pi) \\ &= C + (\hat{\alpha}_{kl}^\mu - \hat{\beta}_{kl}^\sigma)B. \end{aligned}$$

Obviously, we have $MSE(\pi) < MSE(\pi_1)$. In this case, one seeks the optimal order π_1 making the target optimal regardless of standard deviation. Theses π_1 making the standard deviation large.

Case 3. For $\{k, l, s, t\} \neq \{i, j\}$, we have two active standard deviations $I_{kl}^\mu(\pi_1) \neq 0$ and $I_{st}^\sigma(\pi_2) \neq 0$. Note that

$$\begin{aligned} MSE(\pi_1) &= \hat{y}_\mu^2(\pi_1) + \hat{y}_\sigma^2(\pi_1) \\ &= \left[\hat{\alpha}_0 + \sum_{i < j, i \neq k, l; j \neq k, l} (\hat{\alpha}_{ij}^c I_{ij}^c + \hat{\alpha}_{ij}^\mu I_{ij}^\mu) + \hat{\alpha}_{kl}^\mu I_{kl}^\mu \right]^2 (\pi_1) + \\ &= \left[\hat{\beta}_0 + \sum_{i < j, i \neq k, l; j \neq k, l} (\hat{\beta}_{ij}^c I_{ij}^c + \hat{\beta}_{ij}^\sigma I_{ij}^\sigma) \right]^2 (\pi_1). \end{aligned}$$

Similarly, we have

$$\begin{aligned} MSE(\pi) &= \hat{y}_\mu^2(\pi) + \hat{y}_\sigma^2(\pi) \\ &= \left[\hat{\alpha}_0 + \sum_{i < j, i \neq k, l; j \neq k, l} (\hat{\alpha}_{ij}^c I_{ij}^c + \hat{\alpha}_{ij}^\mu I_{ij}^\mu) + \hat{\alpha}_{kl}^\mu I_{kl}^\mu \right]^2 (\pi) + \\ &\quad \left[\hat{\beta}_0 + \sum_{i < j, i \neq k, l; j \neq k, l} (\hat{\beta}_{ij}^c I_{ij}^c + \hat{\beta}_{ij}^\sigma I_{ij}^\sigma) + \hat{\beta}_{kl}^\sigma I_{kl}^\sigma \right]^2 (\pi). \end{aligned}$$

The solution π want to keep $\hat{y}_\sigma^2(\pi)$ smaller. Obviously, $\hat{y}_\mu^2(\pi_1) = \hat{y}_\mu^2(\pi)$ and $\hat{y}_\sigma^2(\pi) = \hat{y}_\sigma^2(\pi_1)$, thus $MSE(\pi) < MSE(\pi_1)$.

The other cases can be similarly proved. Hence, the proof is completed. \square

Theorem 1 shows that the obtained order by the proposed approach is optimal compared to the conventional method. This indicates that the proposed approaches have a smaller MSE value in the replicated experiments.

5. A Numerical Simulation

Here, we provide a distinct example with $m = 10$ for illustrating the effectiveness of the proposed method for OofA experiments with heteroscedasticity. The underlying true model is

$$\begin{aligned} y &= 50 + 2I_{13}(\pi) - 5I_{14}(\pi) - 2I_{26}(\pi) + 3I_{2(10)}(\pi) + 3I_{35}(\pi)z_1 + 3I_{39}(\pi) - I_{45}(\pi) \\ &\quad - 2I_{5(10)}(\pi) - 2I_{67}(\pi) + 4I_{68}(\pi)z_2 + 7I_{78}(\pi) + 5I_{7(10)}(\pi)z_1 + 4I_{8(10)}(\pi)z_2 + \epsilon, \end{aligned} \quad (8)$$

where z_1 and z_2 are two noise factors with $z_1 \sim N(0, 0.5^2)$, $z_2 \sim N(0, 1)$, and $\epsilon \sim N(0, 0.05^2)$. Note that $I_{2(10)}$ (for example) is the PWO variable between components 2 and 10. The purpose of this experiment is to find optimal order making y_μ close to $T = 23$, while minimizing y_σ . The optimal order is in fact $9 \rightarrow 3 \rightarrow 1 \rightarrow 5 \rightarrow 4 \rightarrow 10 \rightarrow 2 \rightarrow 6 \rightarrow 8 \rightarrow 7$, whose resulting expectation is 23 with standard deviation 1.001.

A D -optimal OofA design (from Winker *et al.*, 2020), with 46 $\left(= 1 + \binom{10}{2}\right)$ runs, is chosen the OofA design. For each run, five replicated experiments are conducted and their responses Y_1, \dots, Y_5 are obtained. Two responses (\hat{y}_μ and \hat{y}_σ) are evaluated by those five replicated responses. This is displayed in the Appendix (Table A.1). Via stepwise regression method, the location model and the dispersion model are respectively fitted as

$$\begin{aligned} y_\mu &= 49.583 + 2.231I_{13}(\pi) - 4.513I_{14}(\pi) - 1.286I_{18}(\pi) - 0.843I_{26}(\pi) - 1.085I_{27}(\pi) + 3.318 \\ &\quad \times I_{2(10)}(\pi) + 3.849I_{39}(\pi) + 0.959I_{3(10)}(\pi) - 1.145I_{45}(\pi) - 1.354I_{49}(\pi) - 3.406I_{5(10)}(\pi) \\ &\quad + 7.843I_{78}(\pi) + \epsilon_\mu \end{aligned} \quad (9)$$

and

$$\begin{aligned} y_\sigma &= 5.051 + 1.697I_{14}(\pi) + 0.764I_{29}(\pi) - 2.423I_{35}(\pi) + 2.735I_{45}(\pi) + 1.618I_{57}(\pi) \\ &\quad - 1.226I_{67}(\pi) + \epsilon_\sigma. \end{aligned} \quad (10)$$

We first employ the two-step approach to solve this example. To make y_μ close to $T = 23$, the recommend levels of PWO factors in y_μ (9) are $I_{13}(\pi) = I_{2(10)}(\pi) = I_{39}(\pi) = I_{49}(\pi) = I_{78}(\pi) = -1$ and $I_{14}(\pi) = I_{18}(\pi) = I_{26}(\pi) = I_{27}(\pi) = I_{3(10)}(\pi) = I_{45}(\pi) = I_{5(10)}(\pi) = 1$. According to those active factors, the orders of “3 \rightarrow 1, 10 \rightarrow 2, 9 \rightarrow 3, 9 \rightarrow 4, 8 \rightarrow 7, 1 \rightarrow 4, 1 \rightarrow 8, 2 \rightarrow 6, 2 \rightarrow 7, 3 \rightarrow 10, 4 \rightarrow 5, 5 \rightarrow 10” are obtained. Given the constraint $y_\sigma \geq 0$, the adjusted factors are set as $I_{35}(\pi) = I_{57}(\pi) = I_{67}(\pi) = 1$ and $I_{29}(\pi) = -1$. From those active factors, the optimal orders satisfy “3 \rightarrow 5, 5 \rightarrow 7, 6 \rightarrow 7, 9 \rightarrow 2”. Based on those arranged orders, all optimal orders can be found in the Appendix (Table A.2). As an example, one of the orders is 9 \rightarrow 3 \rightarrow 1 \rightarrow 4 \rightarrow 5 \rightarrow 10 \rightarrow 2 \rightarrow 6 \rightarrow 8 \rightarrow 7 ($\hat{y}_\mu = 22.377$, $\hat{y}_\sigma = 6.688$, $MSE = 45.118$ via (9), (10) and (5), respectively).

To avoid the above drawback of two-step method, the MSE method (Algorithm 1) is used here to find the optimal order. The objective function of MSE method in Step 3 is Equation (5) $(\hat{y}_\mu - T)^2 + \hat{y}_\sigma^2$, where now \hat{y}_μ and \hat{y}_σ represent the location model (9) and dispersion model (10), respectively. Via Step 4 of Algorithm 1, the level of PWO factors are set as $I_{14}(\pi) = I_{18}(\pi) = I_{26}(\pi) = I_{27}(\pi) = I_{35}(\pi) = I_{3(10)}(\pi) = I_{57}(\pi) = I_{5(10)}(\pi) = I_{67}(\pi) = 1$ and $I_{13}(\pi) = I_{29}(\pi) = I_{45}(\pi) = I_{49}(\pi) = I_{2(10)}(\pi) = I_{39}(\pi) = I_{78}(\pi) = -1$.

Thus, the possible orders are “1 \rightarrow 4, 1 \rightarrow 8, 2 \rightarrow 6, 2 \rightarrow 7, 3 \rightarrow 5, 3 \rightarrow 10, 5 \rightarrow 7, 5 \rightarrow 10, 6 \rightarrow 7, 3 \rightarrow 1, 9 \rightarrow 2, 5 \rightarrow 4, 10 \rightarrow 2, 9 \rightarrow 4, 9 \rightarrow 3, 8 \rightarrow 7”. Each component is regarded as a node, the possible order “ $i \rightarrow j$ ” implies a directed edge from i to j . One can generate the directed graph by connecting all directed edges (see Figure 1). According to Figure 1, all optimal orders can be found in Table 2. For example, one possible order is 9 \rightarrow 3 \rightarrow 1 \rightarrow 5 \rightarrow 4 \rightarrow 10 \rightarrow 2 \rightarrow 6 \rightarrow 8 \rightarrow 7 ($\hat{y}_\mu = 24.667$, $\hat{y}_\sigma = 1.218$ and $MSE = 4.262$ via (9), (10) and (5), respectively). This is to make \hat{y}_μ close to the target $T = 23$ while keeping \hat{y}_σ relatively small. As compared to the solution from the two-step approach, the order obtained by the MSE approach has a much smaller \hat{y}_σ while \hat{y}_μ is also close to the target. For confirmation, those two orders obtained by two-step as well MSE approaches were evaluated via the true model (8). It is shown that the expectations are 21 (for two-step order), and 23 (for MSE order), respectively; with identical standard deviations of 1.001.

Table 2: The optimal orders by MSE approach

Run	Order
1	9 \rightarrow 3 \rightarrow 1 \rightarrow 8 \rightarrow 5 \rightarrow 4 \rightarrow 10 \rightarrow 2 \rightarrow 6 \rightarrow 7
2	9 \rightarrow 3 \rightarrow 1 \rightarrow 5 \rightarrow 8 \rightarrow 4 \rightarrow 10 \rightarrow 2 \rightarrow 6 \rightarrow 7
3	9 \rightarrow 3 \rightarrow 1 \rightarrow 5 \rightarrow 4 \rightarrow 8 \rightarrow 10 \rightarrow 2 \rightarrow 6 \rightarrow 7
4	9 \rightarrow 3 \rightarrow 1 \rightarrow 5 \rightarrow 4 \rightarrow 10 \rightarrow 8 \rightarrow 2 \rightarrow 6 \rightarrow 7
5	9 \rightarrow 3 \rightarrow 1 \rightarrow 5 \rightarrow 4 \rightarrow 10 \rightarrow 2 \rightarrow 8 \rightarrow 6 \rightarrow 7
6	9 \rightarrow 3 \rightarrow 1 \rightarrow 5 \rightarrow 4 \rightarrow 10 \rightarrow 2 \rightarrow 6 \rightarrow 8 \rightarrow 7
7	9 \rightarrow 3 \rightarrow 5 \rightarrow 1 \rightarrow 8 \rightarrow 4 \rightarrow 10 \rightarrow 2 \rightarrow 6 \rightarrow 7
8	9 \rightarrow 3 \rightarrow 5 \rightarrow 1 \rightarrow 4 \rightarrow 8 \rightarrow 10 \rightarrow 2 \rightarrow 6 \rightarrow 7
10	9 \rightarrow 3 \rightarrow 5 \rightarrow 1 \rightarrow 4 \rightarrow 10 \rightarrow 8 \rightarrow 2 \rightarrow 6 \rightarrow 7
10	9 \rightarrow 3 \rightarrow 5 \rightarrow 1 \rightarrow 4 \rightarrow 10 \rightarrow 2 \rightarrow 8 \rightarrow 6 \rightarrow 7
11	9 \rightarrow 3 \rightarrow 5 \rightarrow 1 \rightarrow 4 \rightarrow 10 \rightarrow 2 \rightarrow 6 \rightarrow 8 \rightarrow 7

Note: All orders have $\hat{y}_\mu = 24.667$ and $\hat{y}_\sigma = 1.218$ and $MSE=4.262$.

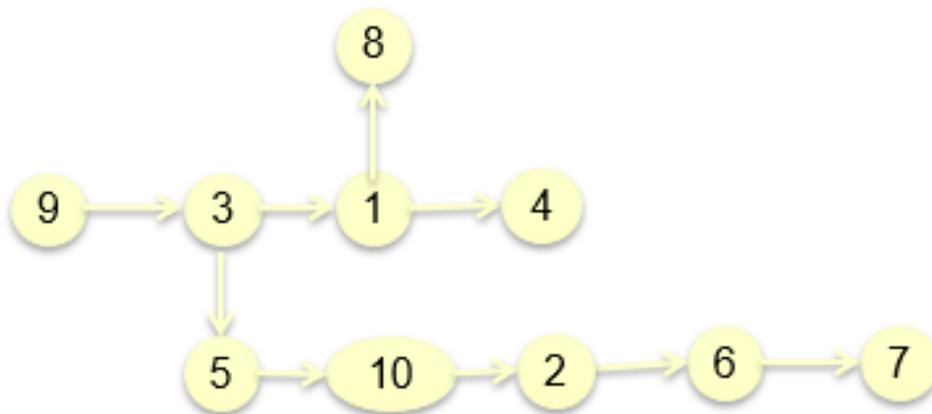


Figure 1: Directed graph

6. Conclusion and Discussion

The goal for an OofA experiment is to find the optimal order. The current research focuses on the case where the dispersion effect for each order is a constant ($y_\sigma \equiv c$). For the OofA experiments with heteroscedasticity, the dual response approach is employed. The location and the dispersion effects are the two interested responses. The two-step and MSE approaches are proposed to speculate optimal orders. When the location model and the dispersion model do not share any common factors, the commonly used two-step method, is able to find optimal order. The location model and dispersion model typically share some common active factors. The two-step approach may be misleading in this case. Lin and Tu (1995) showed that the optimization problem based on MSE is more appropriate to solve the dual response problem. Motivated by their idea, this paper proposes an MSE approach to find the optimal orders for the OofA experiments. Based on the MSE, the obtained orders not only achieve the goal of the OofA experiment, but also minimize the standard deviation within the OofA framework. Some theoretical supports are given (Section 4) to illustrate the effectiveness of the proposed method for OofA experiments with heteroscedasticity. Simulation studies confirmed that the proposed approaches perform well for searching the optimal order(s) in replicated OofA experiments.

The proposed method can be easily extended to the OofA problem with some prior information, the objective function of MSE criterion can be rewritten as

$$\text{MSE} = \omega \hat{y}_\sigma^2 + (1 - \omega)(\hat{y}_\mu - T)^2, \quad (11)$$

where ω measures the relative importance of \hat{y}_σ and \hat{y}_μ with $0 \leq \omega \leq 1$. Especially, for $\omega > 0.5$, the experimenter is inclined to “risk lover”; for $\omega = 0.5$, the experimenter tends to “risk-neutral”; and for $\omega < 0.5$, the experimenter is more likely to “risk averter”. Naturally, if there is no prior information, we suggest setting $\omega = 0.5$. The MSE (11) can be used in Step 3 for Algorithm 1 to tackle the replicated OofA experiment with prior information.

Acknowledgements

This article is in memory of our best friend Alope Dey. Alope was a great researcher. His research is deeply important to us. The first two authors contributed equally to the work. The work of Lin is supported by the *National Science Foundation* via Grant DMS-18102925. The work of the Chen is supported by the National Natural Science Foundation of China (Grant No. 11771220), National Ten-Thousand Talents Program, Tianjin Development Program for Innovation and Entrepreneurship (Grant No.C02055). The work of Zhang was supported by China Scholarship Council.

References

- Box, G. E. and Wilson, K. B. (1951). On the experimental attainment of optimum conditions. *Journal of the Royal Statistical Society: Series B (Methodological)*, 1–45.
- Chen, J. B., Han, X. X., Yang, L. Q., Lin, D. K. J. and Zhou, Y. D. (2020a). Robust designs for ordering problem. Submitted manuscript.
- Chen, J. B., Mukerjee, R. and Lin, D. K. J. (2020b). Construction of optimal fractional Order-of-Addition designs via block designs. *Statistics & Probability Letters*, forthcoming.
- Chen, J. B., Peng, J. Y. and Lin, D. K. J. (2020c). A statistical perspective on NP-Hard problems: making uses of design for order-of-addition experiment. Submitted manuscript.
- Copeland, K. A. and Nelson P. R. (1996) Dual response optimization via direct function minimization. *Journal of Quality Technology*, **28**, 331–336.
- Ding, X., Matsuo, K., Xu, L., Yang, J. and Zheng, L. (2015). Optimized combinations of bortezomib, camptothecin, and doxorubicin show increased efficacy and reduced toxicity in treating oral cancer. *Anti-Cancer Drugs*, **26**, 547–554.
- Kim, K. and Lin, D. K. J. (1998). Dual response surface optimization: A fuzzy modeling approach. *Journal of Quality Technology*, **30**, 1–10.
- Lin, D. K. J. and Peng, J. Y. (2019). The order-of-addition experiments: a review and some new thoughts (with discussion). *Quality Engineering*, **31**(1), 49–59.
- Lin, D. K. J. and Tu, W. Z. (1995). Dual response surface optimization. *Journal of Quality Technology*, **27**(1), 34–39.
- Mee, R. W. (2020). Order of addition modeling. *Statistica Sinica*, **30**, 1543–1559..
- Myers, R. H. and Carter, W. H. (1973). Response surface techniques for dual response systems. *Technometrics*, **15**, 301–317.
- Peng, J. Y., Barton, R. and Lin, D. K. J. (2019a). Swap-r algorithm for the construction of order-of-addition designs. Submitted manuscript.
- Peng, J. Y., Mukerjee, R. and Lin, D. K. J. (2019b). Design of order-of-addition experiments. *Biometrika*, **106** (3), 683–694.
- Van Nostrand, R. C. (1995). Design of experiments where the order of addition is important. *ASA Proceedings of the Section on Physical and Engineering Sciences*. **155–160**. American Statistical Association, Alexandria, Virginia.
- Vining, G. G., and Myers, R. H. (1990). Combining Taguchi and response surface philosophies: a dual response approach. *Journal of Quality Technology*, **22**(1), 38–45.
- Voelkel, J. G. (2019). The design of order-of-addition experiments. *Journal of Quality Technology*, 51, 230–241.

- Winker, P., Chen, J. B. and Lin, D. K. J. (2020). The construction of optimal design for order-of-addition experiment via threshold accepting. In *Contemporary Design of Experiments, Multivariate Analysis and Data mining—In Celebration of Prof. Kai-Tai Fang's 80th Birthday* (pp. 93–109). Springer.
- Yang, J. F., Sun, F. S. and Xu, H. Q. (2020). A component-position model, analysis and design for order-of-addition experiments. *Technometrics*, forthcoming.
- Zhao, Y. N., Lin, D. K. J. and Liu, M. Q. (2020). Designs for order of addition experiments. *Journal of Applied Statistics*, forthcoming.

ANNEXURE

Table A.1: The design and responses for $m = 10$ in Section 5

Run	Order	Y_1	Y_2	Y_3	Y_4	Y_5	\hat{y}_μ	\hat{y}_σ
1	4 → 5 → 10 → 2 → 3 → 8 → 9 → 6 → 7 → 1	21.730	51.429	37.559	37.766	20.475	33.792	12.883
2	8 → 3 → 4 → 10 → 6 → 5 → 9 → 2 → 7 → 1	46.328	47.872	47.869	47.793	48.517	47.676	0.808
3	3 → 1 → 8 → 10 → 6 → 2 → 9 → 5 → 7 → 4	40.144	38.301	37.612	40.160	40.079	39.259	1.214
4	9 → 6 → 4 → 10 → 8 → 7 → 3 → 5 → 2 → 1	39.611	42.334	39.546	40.570	41.057	40.624	1.152
5	8 → 2 → 10 → 6 → 4 → 7 → 9 → 1 → 3 → 5	44.863	46.482	46.901	46.364	45.377	45.997	0.845
6	9 → 2 → 7 → 3 → 6 → 5 → 10 → 1 → 8 → 4	48.396	48.322	48.050	51.828	51.925	49.704	1.988
7	3 → 9 → 1 → 7 → 8 → 5 → 4 → 10 → 2 → 6	48.768	50.391	50.193	49.121	50.767	49.848	0.859
8	1 → 5 → 6 → 2 → 9 → 3 → 10 → 4 → 7 → 8	52.615	55.621	59.494	53.982	58.514	56.045	2.924
9	6 → 10 → 2 → 4 → 3 → 9 → 1 → 8 → 7 → 5	47.376	48.119	46.674	46.319	46.005	46.899	0.852
10	10 → 6 → 7 → 2 → 1 → 8 → 5 → 4 → 3 → 9	55.107	58.215	62.524	60.011	52.709	57.713	3.890
11	8 → 9 → 7 → 5 → 1 → 10 → 3 → 4 → 2 → 6	33.430	33.170	32.657	33.080	31.654	32.798	0.697
12	10 → 8 → 2 → 6 → 4 → 3 → 5 → 7 → 1 → 9	44.626	35.094	58.733	37.707	39.556	43.143	9.387
13	10 → 2 → 1 → 7 → 8 → 5 → 9 → 3 → 6 → 4	61.054	37.452	40.934	44.154	47.954	46.310	9.112
14	7 → 4 → 9 → 1 → 3 → 10 → 6 → 5 → 8 → 2	66.904	60.267	56.888	60.988	57.257	60.461	4.027
15	5 → 6 → 8 → 2 → 1 → 4 → 10 → 9 → 3 → 7	39.443	42.287	26.276	39.532	19.579	33.423	9.936
16	9 → 3 → 7 → 8 → 6 → 4 → 2 → 1 → 5 → 10	59.972	62.076	63.205	55.378	64.840	61.094	3.653
17	7 → 1 → 6 → 4 → 2 → 8 → 3 → 5 → 10 → 9	64.383	63.005	62.900	47.747	48.806	57.368	8.328
18	8 → 5 → 9 → 2 → 10 → 1 → 3 → 6 → 7 → 4	30.842	36.799	40.545	36.269	30.181	34.927	4.361
19	8 → 6 → 4 → 9 → 3 → 10 → 1 → 2 → 7 → 5	41.884	42.204	40.577	41.517	39.555	41.147	1.079
20	3 → 8 → 2 → 7 → 1 → 4 → 5 → 6 → 9 → 10	41.364	40.458	43.810	30.455	42.614	39.740	5.343
21	10 → 4 → 1 → 3 → 2 → 6 → 9 → 7 → 5 → 8	62.047	61.259	60.860	60.301	60.776	61.049	0.654
22	4 → 2 → 8 → 1 → 7 → 3 → 9 → 10 → 5 → 6	57.020	59.270	53.590	57.238	57.690	56.961	2.080
23	4 → 6 → 7 → 8 → 2 → 9 → 5 → 3 → 10 → 1	45.022	63.261	46.962	63.306	60.828	55.876	9.104
24	10 → 3 → 7 → 6 → 1 → 9 → 5 → 4 → 2 → 8	56.225	56.508	57.817	57.147	57.051	56.950	0.617
25	5 → 2 → 9 → 8 → 4 → 3 → 1 → 6 → 10 → 7	45.928	37.625	47.268	42.823	44.084	43.546	3.721
26	3 → 6 → 8 → 1 → 10 → 7 → 4 → 5 → 2 → 9	48.220	41.063	31.275	33.908	40.185	38.930	6.638
27	3 → 5 → 8 → 9 → 4 → 10 → 1 → 7 → 6 → 2	48.260	50.106	48.836	48.423	50.625	49.250	1.056
28	10 → 9 → 2 → 7 → 4 → 6 → 3 → 8 → 5 → 1	54.791	54.607	55.921	54.604	55.423	55.069	0.583
29	7 → 10 → 5 → 8 → 3 → 9 → 1 → 2 → 6 → 4	66.056	51.830	67.087	61.377	64.349	62.140	6.156
30	1 → 6 → 3 → 5 → 4 → 2 → 7 → 10 → 9 → 8	56.312	58.569	51.123	65.361	56.110	57.495	5.171
31	10 → 3 → 7 → 4 → 1 → 8 → 9 → 5 → 6 → 2	61.557	59.821	53.028	53.603	59.097	57.421	3.858
32	4 → 6 → 1 → 3 → 10 → 8 → 9 → 5 → 7 → 2	50.623	52.384	51.387	50.144	52.042	51.316	0.939
33	3 → 9 → 4 → 1 → 2 → 5 → 10 → 6 → 8 → 7	43.954	44.504	44.216	44.102	45.845	44.524	0.766
34	6 → 1 → 5 → 7 → 8 → 9 → 10 → 3 → 2 → 4	45.547	50.965	38.488	56.200	52.814	48.803	6.937
35	1 → 3 → 10 → 7 → 2 → 4 → 9 → 5 → 8 → 6	66.057	60.607	57.330	44.683	64.239	58.583	8.466
36	1 → 3 → 7 → 2 → 6 → 8 → 4 → 9 → 10 → 5	56.218	59.921	50.645	70.229	60.365	59.475	7.163
37	8 → 2 → 9 → 3 → 7 → 4 → 10 → 6 → 1 → 5	48.917	41.185	47.054	54.306	47.560	47.805	4.689
38	4 → 9 → 6 → 8 → 10 → 7 → 1 → 2 → 5 → 3	57.758	48.841	41.355	44.317	48.609	48.176	6.202
39	7 → 10 → 2 → 8 → 3 → 1 → 6 → 4 → 9 → 5	51.210	58.677	50.375	57.681	49.846	53.558	4.261
40	10 → 5 → 7 → 8 → 6 → 4 → 3 → 2 → 9 → 1	72.109	63.371	74.888	57.994	62.421	66.157	7.071
41	9 → 5 → 6 → 3 → 8 → 7 → 10 → 4 → 2 → 1	41.493	34.877	37.153	36.896	38.706	37.825	2.462
42	8 → 6 → 7 → 4 → 5 → 10 → 1 → 2 → 3 → 9	47.518	46.557	46.132	46.314	49.209	47.146	1.271
43	1 → 4 → 8 → 6 → 9 → 7 → 2 → 10 → 3 → 5	43.905	46.309	43.034	33.763	38.586	41.119	4.972
44	7 → 8 → 9 → 1 → 4 → 6 → 10 → 2 → 3 → 5	56.602	49.770	47.541	49.172	50.307	50.678	3.470
45	9 → 3 → 8 → 5 → 6 → 7 → 10 → 1 → 2 → 4	29.556	28.567	24.450	26.749	31.075	28.079	2.566
46	10 → 8 → 1 → 9 → 4 → 7 → 6 → 5 → 3 → 2	23.215	37.266	61.518	42.248	34.875	39.824	13.994

Table A.2: The optimal orders by two-step approach in Section 5

Run	Order
1	9 \rightarrow 3 \rightarrow 1 \rightarrow 8 \rightarrow 4 \rightarrow 5 \rightarrow 10 \rightarrow 2 \rightarrow 6 \rightarrow 7
2	9 \rightarrow 3 \rightarrow 1 \rightarrow 4 \rightarrow 8 \rightarrow 5 \rightarrow 10 \rightarrow 2 \rightarrow 6 \rightarrow 7
3	9 \rightarrow 3 \rightarrow 1 \rightarrow 4 \rightarrow 5 \rightarrow 8 \rightarrow 10 \rightarrow 2 \rightarrow 6 \rightarrow 7
4	9 \rightarrow 3 \rightarrow 1 \rightarrow 4 \rightarrow 5 \rightarrow 10 \rightarrow 8 \rightarrow 2 \rightarrow 6 \rightarrow 7
5	9 \rightarrow 3 \rightarrow 1 \rightarrow 4 \rightarrow 5 \rightarrow 10 \rightarrow 2 \rightarrow 8 \rightarrow 6 \rightarrow 7
6	9 \rightarrow 3 \rightarrow 1 \rightarrow 4 \rightarrow 5 \rightarrow 10 \rightarrow 2 \rightarrow 6 \rightarrow 8 \rightarrow 7

Note: All orders have $\hat{y}_\mu = 22.377$ and $\hat{y}_\sigma = 6.688$ with MSE=45.118.