Computational Statistics and Data Analysis 157 (2021) 107155

Contents lists available at ScienceDirect

Computational Statistics and Data Analysis

journal homepage: www.elsevier.com/locate/csda

A mapping-based universal Kriging model for order-of-addition experiments in drug combination studies

Qian Xiao [a,*], Hongquan Xu [b]

[a] Department of Statistics, University of Georgia, Athens, GA 30602, USA
[b] Department of Statistics, University of California, Los Angeles, CA 90095, USA

ARTICLE INFO

Article history:
Received 8 March 2020
Received in revised form 6 December 2020
Accepted 8 December 2020
Available online 23 December 2020

Keywords:
Chemotherapy
Experimental design
Gaussian process modeling
Response surface modeling

ABSTRACT

In modern pharmaceutical studies, treatments may include several drugs added sequentially, and the drugs' order-of-addition can have significant impacts on their efficacy. In practice, experiments enumerating all possible drug sequences are often not affordable, and appropriate statistical models which can accurately predict all cases using only a small number of experimental trials are required. A novel mapping-based universal Kriging (MUK) model and its simplified variant are proposed for analyzing such order-of-addition experiments with blocking. They can provide accurate predictions and have robust performances under various experimental designs. The MUK model can also incorporate available domain knowledge to enhance its interpretation. The superiority of the proposed methods is illustrated via a real five-drug experiment on lymphoma and two simulation examples.

1. Introduction

In pharmaceutical sciences, there are quite a few studies on the order-of-addition effects where the sequence of arranging drug components plays a significant role to affect the response. In an experiment analyzing protein function within an in-vitro transport system (Preuss et al., 2009), the sequence of actions for two components was found to be significant on the response. In another drug combination experiment on oral cancer (Ding et al., 2015), the authors showed that the sequence of adding three drugs (bortezomib, camptothecin, and doxorubicin) played a vital role in ensuring the efficacy of treatments. We can also see the order-of-addition effects in many other scientific disciplines, such as chemical science (Fuleki and Francis, 1968), political science (Miller and Krosnick, 1998), bio-chemistry (Shinohara and Ogawa, 1998), food science (Jourdain et al., 2009), nutritional science (Karim et al., 2000), manufacturing (Cheng and Wang, 2000) and behavioral science (Dupagne et al., 1999).

To show the characteristics of order-of-addition experiments, we illustrate an example in Yang et al. (2020) where three anti-tumor drugs, denoted as A, B and C, were added into cell cultures either sequentially or simultaneously. There are in total 6 possible sequences for arranging the three drugs which were added one by one every six hours. The responses were the percentages of tumor inhibition measured 12 h after the addition of last drug. The researchers found that adding drugs following the sequence $B \to C \to A$ would achieve the best response among all possible arrangements including the case of all drugs added together. Such an order-of-addition experiment is different from a crossover trial (Jones and Kenward, 2014). The former only measures the endpoint response after all drugs are added, and the drug effects are believed to be dependent on their orders of addition. While, the latter measures the response after adding each drug

* Corresponding author.
E-mail address: qx69137@uga.edu (Q. Xiao).

https://doi.org/10.1016/j.csda.2020.107155
0167-9473/

(3 responses per run in this example), and each drug has a fixed effect which may carry over to the next period but does not depend on its order of addition.

In drug combination studies, some researchers would judiciously decide the drug sequences (without doing any experiment) based on the drug mechanisms and the corresponding pathway information of diseases, which may not be possible in many cases where the mechanisms are unknown (Ding et al., 2015). Thus, the order-of-addition experiments are often needed. There are $k!$ possible sequences for arranging k drugs, and an experiment enumerating all of them may not be affordable in practice (Yang et al., 2020). Appropriate statistical models are needed for accurately predicting the outcomes of all sequences using only a small number of runs. The current literature on modeling order-of-addition experiments is limited, and only two types of linear models are proposed: the pair-wise ordering (PWO) model (Van Nostrand, 1995; Voelkel, 2019; Mee, 2020) and the component-position (CP) model (Yang et al., 2020). Linear models have their advantages in modeling order-of-addition problems, such as clear interpretations. Yet, their prediction accuracy may not always be satisfactory; see Sections 4 and 5 for examples.

Originally from geosciences, Kriging models, aka. Gaussian process models, have been used as a core tool for modeling physical and computer experiments (Krige, 1951; Sacks et al., 1989; Fang et al., 2005; Kleijnen, 2009; Ginsbourger et al., 2009). Specifically, in a drug combination experiment on lung cancer, the researchers showed that Kriging model can provide more accurate predictions than some popular linear and non-linear models (Xiao et al., 2019). The Kriging methods are widely accepted as accurate surrogates for measuring complex response surfaces, and can provide reliable uncertainty quantification (Williams and Rasmussen, 2006; Sullivan, 2015). Standard Kriging models can only take quantitative inputs, and some recent research extend their applications to both quantitative and qualitative inputs (Qian et al., 2008; Deng et al., 2017; Zhang et al., 2020).

In this paper, we propose to apply the universal Kriging (UK) method for modeling a new type of inputs: the order-of-addition inputs. We further develop a novel mapping-based universal Kriging (MUK) model to improve the predictions and enhance the interpretations. The MUK model can incorporate available domain knowledge into its model structure and is flexible enough to address various practical concerns. Compared to the existing methods, the proposed Kriging models may have the following three advantages. First, they can provide accurate predictions in some real data analysis. Second, they can work for various experimental designs and have robust performances. Third, they are parsimonious. When many drugs are involved, the proposed models have much less parameters and thus require much less runs compared to the existing methods introduced in Section 2.

The remainder of this paper is organized as follows. In Section 2, we review current methods for modeling order-of-addition experiments. In Section 3, we first introduce the application of the UK model, then develop the MUK model, and finally illustrate the model estimations. In Section 4 we present a real data analysis on lymphoma treatment and in Section 5 we include two simulation studies to show the superiority of the proposed methods. Section 6 concludes and discusses some future research.

2. Existing methods

In the current literature, two classes of linear models are proposed for order-of-addition problems: the pair-wise ordering (PWO) model (Van Nostrand, 1995; Voelkel, 2019; Mee, 2020) and the component-position (CP) model (Yang et al., 2020); see Lin and Peng (2019) for a review. In this section, we introduce their generalized versions as benchmark methods, which can take blocking into account.

Consider a drug order-of-addition experiment with n runs, k drugs and m blocks. For simplicity, we illustrate the case of $m = 2$ here, and it is straightforward to generalize the methods for $m > 2$. Denote each run in the experiment as $w_i = (a_i, b_i)$ where $a_i = (a_{i,1}, \ldots, a_{i,k})$ is a vector containing the drug sequence and $b_i \in \{0, 1\}$ is the block level. If drugs are added sequentially, a_i will be a permutation of numbers from 1 to k. The PWO model adopts the precedence patterns between all pairs of drugs to represent the features of a drug sequence. Explicitly, let S be the set of all pairs (p, q) for $1 \le p < q \le k$, and for each $(p, q) \in S$, define the PWO indicator as:

$$z_{p,q}(a) = \begin{cases} 1 & \text{if } p \text{ precedes } q \text{ in the sequence } a, \\ -1 & \text{if } q \text{ precedes } p \text{ in the sequence } a. \end{cases}$$

As an illustration, for $k = 3$ and $a = (C, A, B) = (3, 1, 2)$, we have $S = \{(1, 2), (1, 3), (2, 3)\}$, then $z_{1,2}(a) = 1$, $z_{1,3}(a) = -1$ and $z_{2,3}(a) = -1$.

The main-PWO model (Van Nostrand, 1995; Voelkel, 2019) which considers main effects of PWO indicators is defined as

$$y(w_i) = \beta_0 + \beta_1 b_i + \sum_{(p,q)\in S} z_{p,q}(a_i)\beta_{p,q} + \epsilon_i, \tag{1}$$

where β_0, β_1 and $\beta_{p,q}$ are coefficients to be estimated via the least squares method. The residual ϵ_i follows the standard assumptions of linear models.

To capture two-factor interactions of PWO indicators, Mee (2020) proposed the triplet-PWO model, which is defined as

$$y(w_i) = \beta_0 + \beta_1 b_i + \sum_{(p,q)\in S} z_{p,q}(a_i)\beta_{p,q} + \sum_{(l,p,q)\in S'} \left[\beta^{(1)}_{l,p,q} z_{l,p}(a_i) z_{l,q}(a_i) + \beta^{(2)}_{l,p,q} z_{l,p}(a_i) z_{p,q}(a_i)\right] + \epsilon_i, \tag{2}$$

where S' is the set of all tuplets (l, p, q) for $1 \le l < p < q \le k$, and $\beta^{(1)}_{l,p,q}$ and $\beta^{(2)}_{l,p,q}$ are coefficients for interaction effects.

Another class of linear models is the component-position (CP) model (Yang et al., 2020), which is defined as

$$y_i = \beta_0 + \beta_1 b_i + \sum_{j=1}^{k-1}\sum_{c=1}^{k-1} x^{(j)}_{i,c}\beta_{j,c} + \epsilon_i, \tag{3}$$

where $x^{(j)}_{i,c}$ equals to 1 if $a_{i,j} = c$ and 0 otherwise. In other words, $x^{(j)}_{i,c}$ is an indicator of whether drug c is used in the ith run and jth position. Simply speaking, the CP method is a multivariate linear regression treating each position as a factor with k levels.

3. Mapping-based universal Kriging models for order-of-addition experiments

3.1. Universal Kriging for order-of-addition experiments

Consider a drug order-of-addition experiment with n runs, k drugs and m blocks. Denote the ith $(i = 1, \ldots, n)$ run as $w_i = (o_i, b_i)$ where $o_i = (o_{i,1}, \ldots, o_{i,k})$ is an order vector and $b_i \in \{0, \ldots, m-1\}$ is a block level. The order vector o_i contains the order-indexes of the elements in the sequence vector a_i defined in Section 2. As an illustration, if drugs are added following the sequence $C \to A \to B$ (i.e. $a_i = (C, A, B) = (3, 1, 2)$), the vector containing the order-indexes for drugs A, B and C is $o_i = (2, 3, 1)$.

Due to the existence of random errors, e.g., measurement errors, in drug experiments, we propose the following universal Kriging (UK) model for order-of-addition inputs:

$$y(w_i) = \mu(b_i) + Z(o_i) + \epsilon_i. \tag{4}$$

The trend part $\mu(b_i) = B_i^T\beta$ is a linear model for block levels, where B_i is the dummy coding for the block level b_i and $\beta = (\beta_1, \ldots, \beta_m)^T$ is a vector of coefficients. In Eq. (4), $Z(o_i)$ is a Gaussian process with zero mean and stationary covariance function. The random error term $\epsilon_i \sim N(0, \tau_i^2)$ and is stochastically independent of $Z(o_i)$. When assuming homogeneous variances, we have $\tau_1^2 = \cdots = \tau_n^2 = \tau^2$. The covariance function ϕ for the Gaussian process $Z(o_i)$ is defined as:

$$\phi(o_i, o_j) = \mathrm{cov}(Z(o_i), Z(o_j)) = \sigma^2 \prod_{l=1}^{k} K(h_l; \theta_l), \tag{5}$$

where σ^2 is the variance parameter and $K(\cdot\,; \theta_l)$ is a chosen kernel function with a positive correlation parameter θ_l $(l = 1, \ldots, k)$.

Following Eqs. (4) and (5), the UK model adopts the distance measure $h_l = |o_{i,l} - o_{j,l}|/k$ in $K(h_l; \theta_l)$ for the lth drug's order-of-addition, where $o_{i,l}$ and $o_{j,l}$ are the lth elements in order vectors o_i and o_j, respectively. Note that h_l is within 0–1 range. There are two popular types of kernel functions:

$$\text{Gaussian:}\quad K(h_l; \theta_l) = \exp(-(\theta_l h_l)^2/2),$$

$$\text{Matérn}(\nu):\quad K(h_l; \theta_l) = \exp(-\sqrt{2\nu}\theta_l h_l)\frac{p!}{(2p)!}\sum_{i=0}^{p}\frac{(p+i)!}{i!(p-i)!}(\sqrt{8\nu}\theta_l h_l)^{p-i},$$

where $\nu = p + 1/2$, p is a positive integer, and θ_l is the correlation parameter which scales the correlation length. The sample path of a Gaussian process with Gaussian kernel has derivatives at all orders, which may be too smooth in practice. A recommended one for practitioners by Williams and Rasmussen (2006) is the Matérn family with parameter $\nu = 5/2$:

$$K(h_l; \theta_l) = \left(1 + \sqrt{5}\theta_l h_l + 5(\theta_l h_l)^2/3\right)\exp\left(-\sqrt{5}\theta_l h_l\right). \tag{6}$$

The sample path of a Gaussian process under Eq. (6) is twice differentiable, thus being less smooth than that under the Gaussian kernel. Another choice of Matérn kernel is to use parameter $\nu = 3/2$, which will lead to even rougher path than that with $\nu = 5/2$. In this paper, we use Eq. (6) which offers a good balance of smoothness. We plot it with different values of θ in Fig. 1(a). We can see that the correlation decreases as the distance h increases and larger θ corresponds to faster decreasing rate.

Finally, based on Eqs. (4)–(6), the covariance between responses $y(w_i)$ and $y(w_j)$ in the UK model can be written as

$$\mathrm{Cov}(y(w_i), y(w_j)) = \sigma^2 \prod_{l=1}^{k}\left[\left(1 + \sqrt{5}\theta_l h_l + \frac{5(\theta_l h_l)^2}{3}\right)\exp\left(-\sqrt{5}\theta_l h_l\right)\right] + \tau^2\delta_{ij}, \tag{7}$$

where $h_l = |o_{i,l} - o_{j,l}|/k$, and $\delta_{ij} = 1$ if $i = j$ and 0 otherwise. For appropriate model inference, the covariance matrix $C = \left(\mathrm{Cov}(y(w_i), y(w_j))\right)_{n\times n}$ defined by Eq. (7) needs to be positive definite.

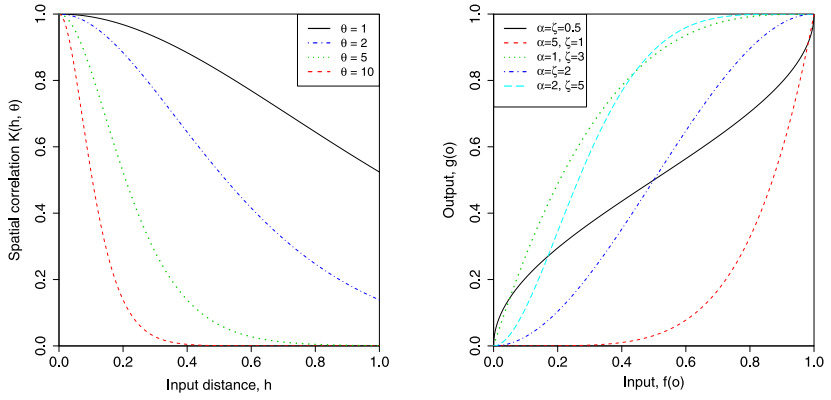

Fig. 1. Examples of (a) Matérn 5/2 correlation functions and (b) cumulative beta mapping functions.

Theorem 1. *For any set of inputs, the covariance matrix of outputs, i.e.* $C = \left(Cov(y(w_i), y(w_j))\right)_{n\times n}$, *induced by the covariance function in (7) with* $\tau^2 \neq 0$ *under the UK model is positive definite.*

Theorem 1 can be proved as follows. Since the Matérn kernel with $\nu = 5/2$ is valid and the product of valid kernels is also a valid kernel, the covariance matrix $\Phi = \left(\phi(o_i, o_j)\right)_{n\times n}$ induced by Eq. (5) is positive semidefinite (Williams and Rasmussen, 2006). When $\tau^2 \neq 0$, the diagonal matrix $(\tau^2 I_n)$ is positive definite. Since the covariance matrix $C = \Phi + \tau^2 I_n$, C is positive definite, which completes the proof. In addition, when all order-of-addition inputs o are distinct (i.e., no replicates are included), the matrix Φ is positive definite. Thus, for any set of distinct order-of-addition inputs, regardless of the τ^2 values, the covariance matrix C induced by (7) is positive definite.

In drug combination experiments, one key reason explaining the order-of-addition effects is the sophisticated interactions among drugs. The Kriging models consider all ways of factors interact with each other and are widely used to model complex response surfaces (Gramacy, 2020, Ch. 5). Another possible reason explaining the order-of-addition effects is the drugs' time-effect, since different orders are essentially different timings for adding drugs. In the UK model, the orders are viewed as numerical inputs, which fixes the distances among them. As an illustration, in a three-drug experiment, the distance between the first and second orders ($h = |2-1|/3 = 1/3$) is always equal to that between the second and third orders ($h = |3-2|/3 = 1/3$) for every drug. A similar idea is discussed in Peng et al. (2019) from an experimental design perspective. In the Kriging methods, similarities between inputs are measured by their distances. Thus, the UK model does not consider different patterns of time-effects. Next, we propose a method to improve this.

3.2. Mapping-based universal kriging for order-of-addition experiments

The MUK model relaxes the constraint in the UK model that all distances between orders are fixed, and adopts a data-driven approach of deciding the pairwise distances. In the MUK model, flexible mapping functions are used to take different patterns of time-effects into account.

The MUK model follows the same Eqs. (4)–(7) as the UK model, but defines a general distance measure $h_l = |g_l(o_{i,l}) - g_l(o_{j,l})|$ in $K(h_l; \theta_l)$ for the lth drug's order-of-addition, where $l = 1, \ldots, k$, $i, j = 1, \ldots, n$, $o_{i,l}$ and $o_{j,l}$ are the lth elements in the order vectors o_i and o_j, respectively. The mapping function $g_l(\cdot)$ is specific to the lth drug, and different drugs can have different mapping functions reflecting their specific patterns. Clearly, the UK model is a special case of the MUK model via using a simple mapping function $g_l(o_{i,l}) = (o_{i,l} - c)/k$ where $c \in (0, 1)$ is a constant. Various types of mapping functions can be considered, e.g., polynomials and Sigmoid curves. Practitioners may choose a type following the available domain knowledge of study.

As drug effects often appear to be cumulative over time, a useful mapping function in the MUK is the cumulative beta function

$$g_l(o_{i,l}) = \frac{1}{B(\alpha_l, \zeta_l)} \int_0^{f(o_{i,l})} t^{\alpha_l - 1}(1-t)^{\zeta_l - 1} dt, \tag{8}$$

where α_l and ζ_l are parameters to be estimated via MLE and the normalizing constant $B(\alpha_l, \zeta_l) = \int_0^1 t^{\alpha_l - 1}(1-t)^{\zeta_l - 1} dt$. The function $f(o_{i,l}) = (o_{i,l} - 0.5)/k$ is used to standardize the experimental design (with levels $o_{i,l} \in \{1, \ldots, k\}$) to 0–1

range. Some examples of cumulative beta functions with different parameters α and ζ are shown in Fig. 1(b), where the distances between successive orders depend on parameters (α, ζ). The MUK model addresses drugs' different patterns on timing (or positions) through parameters α and ζ via a data-driven approach.

Similarly, we can show that Theorem 1 also hold for the MUK model with the mapping function defined in Eq. (8). Additionally, when all inputs o are distinct (i.e., no replicates are included), the covariance matrix of responses in the MUK model is positive definite regardless of the τ^2 values.

When assuming known homogeneous variances (τ^2), the total number of parameters in the UK and MUK models are $k+m+1$ and $3k+m+1$, respectively. As a comparison, the numbers of parameters in the main-PWO, triplet-PWO and CP models are $k(k-1)/2+m$, $k(k-1)(k-2)/3+k(k-1)/2+m$ and $(k-1)^2+m$, respectively. In this paper, we recommend practitioners using designs with run sizes no less than the numbers of parameters in the models. Clearly, when moderate or large numbers of drugs are involved, the UK and MUK models require much less runs compared to the existing methods. Note that if practitioners follow the rule-of-thumb for choosing run size $n = 10k$ (Loeppky et al., 2009), both the UK and MUK models can fit well, but none of the existing methods will work for large k.

Compared to the UK model, the MUK model includes more parameters and thus requires more data. In practice, the UK model generally performs better for small experiments. Yet, when enough data are available, the MUK model is often preferred due to its better interpretation and flexibility.

3.3. Model estimation

When assuming known homogeneous variance τ^2, the MUK model contains parameters $\beta = (\beta_i)^T_{1\times m}$, σ^2, $\theta = (\theta_j)^T_{1\times k}$, $\alpha = (\alpha_l)^T_{1\times k}$ and $\zeta = (\zeta_l)^T_{1\times k}$. Its likelihood takes the form

$$L(\beta, \sigma^2, \theta, \alpha, \zeta) = \frac{1}{(2\pi)^{n/2}|C|^{1/2}} \exp\left(-\frac{1}{2}(y-B\beta)^T C^{-1}(y-B\beta)\right), \tag{9}$$

where $y = (y_1, \ldots, y_n)^T$ is the response vector of n observed values, $B = (B_1, B_2, \ldots, B_n)^T$ is an $n \times m$ matrix, B_i is the column vector of dummy coding for block level b_i, covariance matrix $C = C(\sigma^2, \theta, \alpha, \zeta)$ is defined in Eq. (7) with $h_l = |g_l(o_{i,l}) - g_l(o_{j,l})|$ and $g_l(\cdot)$ defined in Eq. (8). The MLE of β has an analytical expression:

$$\widehat{\beta} = (B^T C^{-1} B)^{-1} B^T C^{-1} y. \tag{10}$$

After dropping some constants and taking Eq. (10) into account, maximizing the log-likelihood is equivalent to solving the following minimization problem:

$$[\sigma^2, \theta, \alpha, \zeta] = \text{argmin } \log|C| + (y - B\widehat{\beta})^T C^{-1}(y - B\widehat{\beta}). \tag{11}$$

This minimization can be solved via standard optimization algorithms in R or MATLAB, which leads to the MLEs of parameters σ^2, θ, α and ζ. In this paper, we adopt the "rgenoud" package in R (Mebane Jr. and Sekhon, 2011) which combines evolutionary search algorithms with derivative-based quasi-Newton methods. It adopts numerical gradients by default, which suffices for small or moderate experiment sizes. For large experiments, we derive the analytical gradients for the MUK model to further speedup the estimations.

Based on Eqs. (10) and (11), for any parameter in $C(\sigma^2, \theta, \alpha, \zeta)$, the expression of the analytical gradient is:

$$\frac{\partial f}{\partial \bullet} = tr(C^{-1}\frac{\partial C}{\partial \bullet}) - (y - B\widehat{\beta})^T C^{-1}\frac{\partial C}{\partial \bullet} C^{-1}(y - B\widehat{\beta}), \tag{12}$$

where $\frac{\partial C}{\partial \bullet} = \left(\frac{\partial \phi(o_i, o_j)}{\partial \bullet}\right)_{n\times n}$ is an $n \times n$ matrix. For any $i, j = 1, \ldots, n$ and $l = 1, \ldots, k$, we can derive

$$\frac{\partial \phi(o_i, o_j)}{\partial \sigma^2} = \prod_{l=1}^{k} K(h_l; \theta_l) = \prod_{l=1}^{k}\left[(1+\sqrt{5}\theta_l h_l + \frac{5(\theta_l h_l)^2}{3})\exp(-\sqrt{5}\theta_l h_l)\right],$$

$$\frac{\partial \phi(o_i, o_j)}{\partial \theta_l} = -\frac{5\theta_l h_l^2 \sigma^2}{3}(1+\sqrt{5}\theta_l h_l)\exp(-\sqrt{5}\theta_l h_l) \prod_{i=1,\ldots,k, i\neq l} K(h_i; \theta_i).$$

Next, we consider the analytical gradients for the mapping parameters α_l (or ζ_l in a similar way). Then, we can derive:

$$\frac{\partial \phi(o_i, o_j)}{\partial \alpha_l} = \sigma^2 \frac{\partial K(h_l; \theta_l)}{\partial \alpha_l} \prod_{i=1,\ldots,k, i\neq l} K(h_i; \theta_i),$$

$$\frac{\partial K(h_l; \theta_l)}{\partial \alpha_l} = c_1(\sqrt{5}\theta_l + \frac{10\theta_l^2 h_l}{3})\exp(-\sqrt{5}\theta_l h_l) + c_2(1 + \sqrt{5}\theta_l h_l + \frac{5(\theta_l h_l)^2}{3}),$$

where

$$c_1 = \frac{\partial h_l}{\partial \alpha_l} = \frac{g_l(o_{i,l}) - g_l(o_{j,l})}{|g_l(o_{i,l}) - g_l(o_{j,l})|}(\frac{\partial g_l(o_{i,l})}{\partial \alpha_l} - \frac{\partial g_l(o_{j,l})}{\partial \alpha_l}),$$

Table 1
Data for the five-drug order-of-addition experiment with blocking.

Bl	Seq	Y	Bl	Seq	Y	Bl	Seq	Y	Bl	Seq	Y
1	DBACE	4.93	1	BCAED	23.85	2	DBCEA	5.53	2	BEACD	20.4
1	BACDE	13.63	1	DECBA	25.23	2	BADEC	7.72	2	ACBED	22.06
1	DABEC	15.57	1	ECBDA	25.62	2	ABDCE	10.96	2	CBEAD	22.35
1	DCEAB	18.47	1	CBDEA	26.08	2	BDCAE	12.09	2	CABDE	23.37
1	EDABC	19.5	1	EADCB	26.75	2	DAECB	13.84	2	DEBAC	23.4
1	ABEDC	20.23	1	BEDAC	28.38	2	ADEBC	16.25	2	EBADC	24.31
1	BDECA	21.47	1	CDBAE	29.43	2	AECDB	16.37	2	BCEDA	24.65
1	AEBCD	21.59	1	CEADB	30.52	2	DCABE	17.97	2	CDAEB	25.99
1	ACDBE	23.55	1	CAEBD	31.27	2	ECDAB	19.71	2	CEDBA	26.3
1	ADCEB	23.61	1	EBCAD	31.96	2	EDBCA	20.35	2	EACBD	26.49

$$c_2 = \frac{\partial\left(\exp(-\sqrt{5}\theta_l h_l)\right)}{\partial\alpha_l} = -\sqrt{5}\theta_l \exp(-\sqrt{5}\theta_l h_l)c_1.$$

When using $g_l(\cdot)$ defined in Eq. (8), the derivatives $\partial g_l(\cdot)/\partial\alpha_l$ in the incomplete Beta function can be easily solved using the method in Boik and Robison-Cox (1998).

Finally, given all estimated parameters, the mean prediction and variance of the response at target input $w_* = (o_*, b_*)$ are given by

$$\widehat{y}(w_*) = B_*^T\widehat{\beta} + \gamma^T C^{-1}(y - B\widehat{\beta}),$$

$$s^2(w_*) = \sigma^2 - \gamma^T C^{-1}\gamma + (B_*^T - \gamma^T C^{-1}B)(B^T C^{-1}B)^{-1}(B_*^T - \gamma^T C^{-1}B)^T,$$

where γ is the covariance vector $(\phi(o_*, o_i))_{n\times 1}$ for $i = 1, \ldots, n$ and B_* is the dummy coding for b_*. Refer to Roustant et al. (2012) for more details on the derivation of $\widehat{y}(w_*)$ and $s^2(w_*)$.

In practice, when background information is available, we can set the noise variance τ^2 according to the measurement accuracy of the responses. Different choices of small τ^2 often do not affect the prediction much in practice (Xiao et al., 2019). When there is no such prior knowledge, we can estimate τ^2 via MLE. In such cases, the covariance matrix becomes $C_1 = C(\tau^2, \sigma^2, \theta, \alpha, \zeta) = (\phi(o_i, o_j))_{n\times n} + \tau^2 I_n$ where I_n is an identity matrix of size n. Similarly as above, we are to solve $[\tau^2, \sigma^2, \theta, \alpha, \zeta] = \text{argmin } \log|C_1| + (y - B\widehat{\beta})^T C_1^{-1}(y - B\widehat{\beta})$, where the MLE of $\widehat{\beta} = (B^T C_1^{-1}B)^{-1}B^T C_1^{-1}y$. This minimization can be implemented by the same optimization algorithm as above. For the analytical gradients, Eq. (12) still applies after replacing matrix C with C_1, where we have $\partial C_1/\partial\tau^2 = I_n$.

In the current literature, all methods assume homogeneous noise variances for the order-of-addition experiments (Van Nostrand, 1995; Voelkel, 2019; Mee, 2020; Yang et al., 2020; Lin and Peng, 2019; Zhao et al., 2020). Another potential advantage of the proposed UK and MUK model is that they can also be generalized to deal with heterogeneous variances for replicated designs. The variance of replicates for each sequence can be used as the unbiased estimate of the heterogeneous variance, and the covariance matrix becomes $C_2 = C(\sigma^2, \theta, \alpha, \zeta) = (\phi(o_i, o_j))_{n\times n} + V$ where V is a diagonal matrix with the heterogeneous variances $(\hat{\tau}_1^2, \ldots, \hat{\tau}_n^2)$ as the diagonal elements. After replacing matrix C with C_2, Eqs. (10) and (11) still apply, and the MLEs of parameters can be solved by the same optimization algorithm. As the UK model is a special case of the MUK model, it follows a similar estimation procedure.

4. Case study: a five-drug order-of-addition experiment on lymphoma treatment

Lymphoma is cancer that begins in infection-fighting cells of the immune system. In a real order-of-addition experiment in vitro on lymphoma treatment (Mee, 2020; Yang et al., 2020; Wang et al., 2020), five FDA approved chemotherapeutics for clinical trials were studied. Different from traditional drug combination experiments which targeted on finding optimal drug doses, this study focused on identifying drugs' order-of-addition impacts, where all drug doses were fixed at appropriate levels. Raji cell, a human lymphoma cell line, was used as an in vitro model to evaluate various drug sequences. In each run of this experiment, drugs were added into the cell culture one by one every 3 h. The responses (Y) were the percentages of cell inhibition (ranging from 0–100) which were measured 12 h after the addition of last drug. In this study, larger responses (Y) are better. Table 1 shows the data consisting of 40 runs in two blocks, where "Bl" gives the block levels and "Seq" shows the drug sequences. Denote the design matrix consisting of all these 40 runs as D_{full}.

First, we target on comparing the prediction accuracy of the proposed models and the existing methods with small designs. We split the full data in Table 1 into various training and test sets, and calculate three popular performance measures: R_1: the correlation between the actual and predicted responses of all observations, R_2: the correlation between the actual and predicted responses of observations in the test set, RMSE: the root mean square error of predicted responses of observations in the test set. In pharmaceutical studies, correlation is a common way to visualize performances (Ning

Table 2
Results of R_1/R_2(RMSE) for different models using the PWO D-efficient designs.

	PWO	UK	MUK
12-run	0.66/0.55(7.34)	0.78/0.66(5.43)	NA
16-run	0.69/0.55(6.17)	0.81/0.69(5.03)	NA
20-run	0.70/0.50(6.38)	0.85/0.61(5.01)	0.87/0.68(4.76)
24-run	0.74/0.54(6.22)	0.88/0.54(4.88)	0.89/0.56(4.71)
28-run	0.75/0.76(5.29)	0.91/0.74(5.06)	0.92/0.79(4.61)
32-run	0.89/0.77(8.02)	0.93/0.73(5.31)	0.95/0.84(4.51)
36-run	0.97/0.76(4.36)	0.99/0.99(2.75)	0.99/0.95(2.90)

Note: The PWO is the triplet-PWO for the 32 and 36-run cases and the main-PWO for smaller cases.

Table 3
Estimated parameters for the UK and MUK models in the five-drug experiment.

UK model:									
Parameter	β_1	β_2	σ^2	θ_1	θ_2	θ_3	θ_4	θ_5	
MLE	22.95	19.00	38.64	1.88	0.57	2.5	1.46	4.64	
MUK model:									
Parameter	β_1	β_2	σ^2	θ_1	θ_2	θ_3	θ_4	θ_5	α_1
MLE	24.26	20.29	33.66	2.23	2.13	6.64	0.59	3.34	0.82
Parameter	ζ_1	α_2	ζ_2	α_3	ζ_3	α_4	ζ_4	α_5	ζ_5
MLE	2.35	1.94	4.56	0.9	4.82	4.10	4.88	1.53	3.75

et al., 2014), where larger correlations (approaching 1) represent more accurate predictions. In statistics, RMSE is a more rigorous performance measure, which is defined as:

$$\text{RMSE} = \sqrt{\frac{1}{n_t}\sum_{i=1}^{n_t}(\widehat{y}(w_i) - y(w_i))^2},$$

where $\hat{y}(w_i)$ and $y(w_i)$ are the predicted and actual responses of input w_i, respectively, and n_t is the number of observations in the test set. Smaller RMSE represents more accurate predictions.

In Table 2, we compare the proposed UK and MUK models to the PWO models, where the results are listed in the form of R_1/R_2(RMSE). As the main-PWO and triplet-PWO models include 12 and 32 parameters for this experiment, the PWO model in Table 2 is the triplet PWO for the 32 and 36-run cases and the main-PWO for the smaller cases. The UK and MUK models include 8 and 18 parameters, respectively. Thus, we mark "NA" for the MUK model in the 12 and 16-run cases in Table 2. For an n-run case here, we choose the most D-efficient n-run design (subset of D_{full}) out of 10 000 random ones for the corresponding PWO model used. We fit the UK, MUK and PWO models with the training data selected by these D-efficient designs, and then evaluate their performances via R_1, R_2 and RMSE measures. According to the measurement accuracy in this real experiment, we set $\tau^2 = 1$ in the UK and MUK models.

From Table 2, it is seen that the MUK model outperforms the PWO models for all cases in terms of all criteria, and the UK model outperforms the PWO models except for the 28 and 32-run cases where the UK model gives better R_1 and RMSE but worse R_2 values. Compared to the PWO models, the UK and MUK models require less runs to achieve similar prediction accuracy. When focusing on the RMSE measure, we can see that the UK and MUK models using the 20-run design can perform better than the PWO models using the designs with 32 runs or less.

We would like to remark that the optimal designs for the PWO models may not be optimal for the UK and MUK models. For Kriging-based models, space-filling designs can be better choices (Gramacy, 2020). Here, the UK and MUK models can perform very well when not using their optimal designs. Not surprisingly, if random designs are used rather than D-efficient designs (for PWO models), the performances of PWO models may deteriorate dramatically, while the proposed UK and MUK models suffer less. The CP model does not perform as well as the PWO models for this particular experiment (Mee, 2020), although one cannot make a generalization from this outcome. Please refer to the Supplementary Material for a detailed numerical study on the comparison of these models, including the CP model, with random designs.

Next, we consider the performances and interpretations for fitting the proposed models with the full data. Under the 40-run design D_{full}, the UK and MUK models give the $R_1 = 1.00$. The main-PWO and triplet-PWO give the R_1 values of 0.78 and 0.98, respectively. Note that there is no R_2 or RMSE when using the full data as the training set. We show the estimated parameters for the UK and MUK models in Table 3. The trend parameter for block 1 (β_1) is larger than that for block 2 (β_2) by roughly 4 under both models. This difference on blocks is consistent with the analysis presented in Mee (2020).

When five drugs are added sequentially, there are in total $5! = 120$ possible sequences. Using the fitted UK and MUK models with parameters in Table 3, we can predict the responses for all 120 drug sequences. In Fig. 2, we show the results in drug-order-effect plots. In such plots, the horizontal axis denotes the order at which a drug is added. The vertical axis

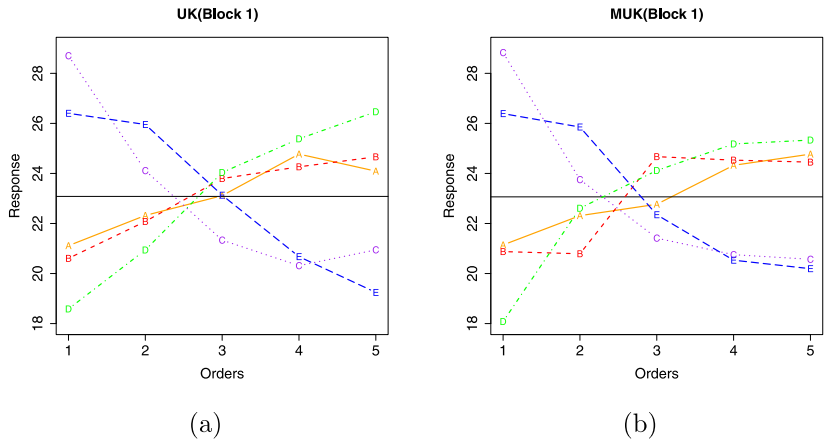

Fig. 2. The drug-order-effect plots for the UK and MUK models (block 1).

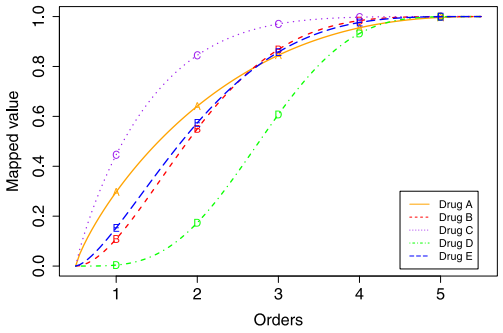

Fig. 3. Plots of mapping functions for the five drugs.

denotes the mean response, and each dot denotes the mean response corresponding to the drug at a given order. For each drug, the five dots corresponding to five orders are lined together. We use a horizontal line to represent the overall mean. Except for a difference of roughly 4 in overall means, the drug-order-effect plots for blocks 1 and 2 are the same. Thus, we only show plots for block 1 in Fig. 2. The drug-order-effect plot can be interpreted similarly as the usual main-effect plot. The general patterns in Fig. 2 are similar under both models. Roughly speaking, when the orders increase, effects of drugs C and E decrease whereas effects of drugs A, B and D increase. The drug-order-effect plot can help explain desirable sequences with large responses. As an illustration, consider the order 1 in Fig. 2. It is clear that the two largest mean responses are from drugs C and E, which helps explain why the drug sequences with the largest three responses (in either block) in Table 1 start with either drug C or drug E.

The drug-order-effect plot is used for visualization and interpretation, but may not be sufficient for making decisions on optimal sequences, as it does not capture interactions. We should directly identify desirable sequences from model predictions. Given the robust performance of the MUK model in Table 2, we focus on the MUK fitted with all 40 runs here. With its estimated parameters in Table 3, the best ten drug sequences predicted from the MUK model include five observed ones in this experiment: EBCAD, CAEBD, CEADB, EACBD and CEDBA, and another five uncovered ones: CAEDB, EBCDA, EBACD, CEDAB and EABCD. Researchers may want to test these newly discovered sequences in follow-up experiments to confirm the optimal solution.

In Fig. 3, we show mapping functions $g_l(\cdot)$ in the MUK model based on the estimated parameters in Table 3. Fig. 3 can be informative given certain domain knowledge. As an illustration, we can see that the distances between consecutive orders for drug C are decreasing. As the similarities between responses are measured by distances between inputs in the proposed Kriging models, larger differences in cell inhibition are expected when drug C is used in the first three orders compared to the last three orders. The shape of mapping function for each drug reflects its specific pattern, and one potential reason is their different time-course effects (Al-Sallami et al., 2009). For example, if drug C has significant delayed effect (the effect appears delayed with respect to the concentration–time profile) and is added in the last several orders, it is possible that the effect of drug C is not fully exerted before measuring the responses. This may be a reason explaining the shape of the curve for drug C. The pathological mechanism behind this is worth future studies when more domain knowledge is available.

Table 4
Comparison of results (R_1/R_2) for different models in Example 1.

	CP	Main-PWO	Triplet-PWO	UK	MUK
Case 1	1/1	0.89/0.89	0.86/0.84	0.98/0.98	0.99/0.99
Case 2	0.99/0.99	1/1	1/1	0.99/0.99	0.99/0.99
Case 3	0.72/0.70	0.73/0.72	1/1	0.92/0.90	0.91/0.90

Table 5
Comparison of results (RMSE) for different models in Example 1.

	CP	Main-PWO	Triplet-PWO	UK	MUK
Case 1	0	2.57	3.22	1.14	0.64
Case 2	0.79	0	0	0.64	0.69
Case 3	4.58	4.43	0	2.82	2.92

5. Simulation studies

To further illustrate the accuracy and robustness of the proposed models, we discuss two simulations in this section. In Example 1, we revisit the case study in Section 4 and assume the true model to be the CP, main-PWO and triplet-PWO, respectively. We aim to evaluate the performances of other models under each of the assumed true ones. In Example 2, we discuss a single machine scheduling problem (Emmons, 1969; Townsend, 1978; Allahverdi et al., 1999; Zhao et al., 2020) to compare the performances of different models under various optimal order-of-addition experimental designs in the literature.

Example 1. For the 40-run and 5-drug order-of-addition experiment in Section 4, consider the following three cases of simulations. In Case 1, we fit a CP model with the full data (D_{full}) in Table 1, and assume it as the true model. Under this true model, we simulate the true responses for all $5! = 120$ possible sequences. We use the true responses selected by D_{full} as the new training data, and use the rest as the new test data. Then, we fit all models with the new training data, and use them to predict the responses of all 120 sequences. Finally, we measure models' performances via comparing their predicted responses with the simulated true responses. Similarly, in Cases 2 and 3, we assume the main-PWO and the triplet-PWO as the true models, respectively.

The assumed true model will be exactly accurate in each corresponding case, and we focus on comparing the performances of other models. We show the results in the form of R_1/R_2 in Table 4, and list the RMSEs in Table 5. From Tables 4 and 5, we can see that the proposed UK and MUK models perform better than other models (excluding the assumed true ones) in all cases. Specifically, when the triplet-PWO is assumed to be the true model in Case 3, the proposed UK and MUK models perform much better than the CP and main-PWO models in terms of all criteria, which shows their superior capabilities to model interactions. Note that the main-PWO is a sub-model of the triplet-PWO. Thus, when assuming the main-PWO as the true model in Case 2, the triplet-PWO after step-wise selection will be exactly accurate in prediction, which should not be compared with other models. For all cases here, both the UK and MUK models achieve R_1 and R_2 values higher than 90%, and thus they are accurate and robust surrogates for all true models.

Example 2. Consider a single machine scheduling problem (Emmons, 1969; Allahverdi et al., 1999; Zhao et al., 2020) consisting of 6 jobs, indexed by Jobs $1, \ldots, 6$. These jobs need to be processed on a single machine one after another, and Job i takes a fixed processing time x_i. Denote the job-sequence on this single machine as $\alpha = (\alpha_1, \ldots, \alpha_6)$ and its order-sequence as $o = (o_1, \ldots, o_6)$, where Job α_i has the processing order $o_{\alpha_i} = i$. The completion time of Job α_i is $T(\alpha_i) = \sum_{j=1}^{i} x_{\alpha_j}$. Now, consider the quadratic cost function of completion time by Townsend (1978), which is defined as

$$C(\alpha) = \sum_{i=1}^{6} w_i T^2(\alpha_i) + \epsilon, \tag{13}$$

where w_i $(i = 1, \ldots, 6)$ are weights and the noise term $\epsilon \sim N(0, \tau^2)$. Here we randomly draw processing time x_i from chi-squared distribution with 1 degree of freedom and randomly draw weight w_i from uniform distribution within range 0–1. Specifically, we set $(x_1, \ldots, x_6) = (0.27, 0.33, 0.065, 1.04, 2.51, 0.72)$ and $(w_1, \ldots, w_k) = (0.23, 0.97, 0.53, 0.33, 0.30, 0.48)$. We consider two scenarios for the true noise variance: (1) a moderate case of $\tau^2 = 1$ and (2) a large case of $\tau^2 = 10$ (considering the range of the responses).

In this example, there are $6! = 720$ possible sequences for arranging the six jobs which will lead to different responses $C(\alpha)$. To generate the training data, we consider four optimal order-of-addition designs in the literature; they are OOFA_{24}: 24-run order-of-addition orthogonal array by Voelkel (2019), COA_{30}: 30-run component orthogonal array by Yang et al. (2020), and MFD_{25} and MFD_{30}: 25- and 30-run D-efficient designs for order-of-addition experiments by Chen et al. (2020), respectively. For each corresponding case, we use the rest of sequences to form the test data. Note that the OOFA and MFD

Table 6
Mean results of R_1/R_2(RMSE) for different models under various designs when the true noise variance $\tau^2 = 1$ in Example 2.

	CP	Main PWO	UK	MUK
$OOFA_{24}$	0.91/0.91(4.08)	0.97/0.97(2.27)	0.99/0.99(1.84)	0.99/0.99(1.54)
MFD_{25}	0.89/0.89(4.57)	0.94/0.94(3.11)	0.99/0.99(1.74)	0.98/0.98(1.96)
MFD_{30}	0.76/0.76(8.54)	0.96/0.96(2.61)	0.99/0.99(1.61)	0.98/0.98(1.85)
COA_{30}	0.98/0.98(1.97)	0.97/0.97(2.41)	0.99/0.99(1.87)	0.99/0.99(1.50)

Table 7
Mean results of R_1/R_2(RMSE) for different models under various designs when the true noise variance $\tau^2 = 10$ in Example 2.

	CP	Main PWO	UK	MUK
$OOFA_{24}$	0.75/0.75(7.71)	0.94/0.94(3.33)	0.97/0.97(2.45)	0.94/0.94(3.12)
MFD_{25}	0.73/0.73(8.61)	0.89/0.88(4.59)	0.97/0.97(2.50)	0.94/0.94(3.23)
MFD_{30}	0.75/0.75(8.93)	0.91/0.91(4.00)	0.97/0.97(2.29)	0.95/0.95(2.98)
COA_{30}	0.93/0.93(3.69)	0.94/0.94(3.42)	0.97/0.97(2.28)	0.96/0.96(2.75)

are developed for the PWO model and the COA is developed for the CP model. Suppose we do not have any background knowledge on the noise variance τ^2 in this example. Thus, we use the MLE of τ^2 in the proposed UK and MUK models.

First, we study the case when there is a moderate true noise variance $\tau^2 = 1$. In Table 6, we show the mean results of 50 replications in the form of "R_1/R_2(RMSE)" for the CP, main-PWO, UK and MUK models under various designs. It is seen that the proposed UK and MUK models outperform the CP and main-PWO models in all cases. Specifically, they give R_1 and R_2 values higher than 98% and RMSE values less than 2 under all designs here, which are very good performances in practice. The CP model works well when using its specific optimal design COA_{30}, but performs badly when using design MFD_{30} of the same size. The main-PWO model performs well in all cases, but is still inferior to the UK and MUK models. Note that the triplet-PWO model requires at least 56 runs in this simulation, and thus we fit it with the D-efficient 60-run design MFD_{60}. Yet, its predictions are not satisfactory with the mean $R_1 = 0.85$, $R_2 = 0.85$ and RMSE $= 5.61$, which are worse than those under the main-PWO model fitted with MFD_{30}. Thus, we did not show it in the table.

Next, we look at the case when there is a large true noise variance $\tau^2 = 10$, and we show the mean results of 50 replications in Table 7. From Table 7, it is seen that the proposed UK and MUK models outperform the CP and main-PWO models in all cases. The CP model works best with its optimal design COA_{30}, and the main-PWO model works best with its optimal design $OOFA_{24}$. The triplet-PWO model with MFD_{60} gives the mean $R_1 = 0.73$, $R_2 = 0.71$ and RMSE $= 8.72$, which is again worse than the main-PWO model with MFD_{30}. Compared to the results in Table 6, all results in Table 7 are worse. Large noise variances will cause challenges for all models. Compared to the UK model, the impact of large noise variance to the MUK model is larger. Intuitively, as all Kriging-based methods incline to interpolate observations, the UK and MUK models are more suitable for modeling experiments with small or moderate errors. For the cases with very large error variances, further investigations may be needed in practice.

6. Conclusions and discussions

In traditional drug combination studies, researchers only focus on drug doses. In several recent drug experiments, scientists discover the significant impacts of drugs' order-of-addition on the efficacy of treatments. Yet, enumerations of all possible drug sequences are often not affordable. In this paper, we propose the UK and MUK models for analyzing drugs' order-of-addition experiments. Compared to existing methods, the proposed models can have stronger prediction power and are more robust under various experimental designs. They include less parameters, require less runs and can incorporate available domain knowledge to enhance interpretations. Their superiority is illustrated via a real case study on lymphoma treatment and two follow-up simulation studies.

In the current literature, all methods assume homogeneous noise variances in the order-of-addition experiments (Van Nostrand, 1995; Voelkel, 2019; Mee, 2020; Yang et al., 2020; Lin and Peng, 2019). The proposed UK and MUK models can be generalized to deal with heterogeneous variances in replicated experiments. Yet, how to enable the proposed models working with heterogeneous variances in single-replicate cases is challenging, as there will be more parameters in the model than the number of runs. A potential way is to use the Bayesian approach (Berger et al., 2001; Gramacy, 2007). For example, Gramacy (2020) discusses the form of the posterior under the noise variances following inverse gamma distributions with different hyper-parameters. For such cases, meaningful inference priors need to be identified. This will be our future research.

In this paper, we adopt the cumulative beta mapping function for the in vitro experiment where drug effects often appear to be cumulative over time. An important topic for future research is to study the choices of mapping functions for various kinds of experiments under their specific domain knowledge. In addition, desirable designs may further improve the performances of the proposed UK and MUK models. Space-filling designs (Zhou and Xu, 2014; Chen et al., 2014, 2015, 2019) are shown to be desirable based on the theoretical prediction variance properties for Kriging models (Silvestrini

et al., 2013). An interesting future work will be how to construct efficient designs for Kriging-based models dealing with order-of-addition problems.

In modern statistics and pharmaceutical studies, sequential experiments, aka. active learning, are found to be more effective in identifying optimal solutions compared to traditional non-adaptive experiments (Christen and Sansó, 2011; Bartroff et al., 2013). The use of UK and MUK models can easily allow reliable uncertainty quantification (Williams and Rasmussen, 2006) for data with order-of-addition inputs, and consequently enable exploration in sequential experiments (Bartroff et al., 2013; Scarpa et al., 2007). It will be another interesting topic for the future research.

Acknowledgments

The authors thank a co-editor, an associate editor and two referees for their valuable comments, which have led to the improvement of the paper.

This research of Xiao was supported by a grant (JFSGFY20) from the University of Georgia Research Foundation, Inc, USA.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.csda.2020.107155.

References

Al-Sallami, H.S., Pavan Kumar, V.V., Landersdorfer, C.B., Bulitta, J.B., Duffull, S.B., 2009. The time course of drug effects. Pharm. Stat.: J. Appl. Stat. Pharm. Ind. 8 (3), 176–185.
Allahverdi, A., Gupta, J.N., Aldowaisan, T., 1999. A review of scheduling research involving setup considerations. Omega 27 (2), 219–239.
Bartroff, J., Lai, T.L., Shih, M.-C., 2013. Sequential Experimentation in Clinical Trials: Design and Analysis. Springer Science Business Media, New York, NY.
Berger, J.O., De Oliveira, V., Sansó, B., 2001. Objective Bayesian analysis of spatially correlated data. J. Amer. Statist. Assoc. 96 (456), 1361–1374.
Boik, R.J., Robison-Cox, J.F., 1998. Derivatives of the incomplete beta function. J. Stat. Softw. 3 (1), 1–20.
Chen, R.-B., Chang, S.-P., Wang, W., Tung, H.-C., Wong, W.K., 2015. Minimax optimal designs via particle swarm optimization methods. Stat. Comput. 25 (5), 975–988.
Chen, J., Han, X., Lin, D.K.J., Yang, L., Zhou, Y., 2020. Design and analysis of ordering problems. Manuscript.
Chen, R.-B., Hsu, Y.-W., Hung, Y., Wang, W., 2014. Discrete particle swarm optimization for constructing uniform design on irregular regions. Comput. Statist. Data Anal. 72, 282–297.
Chen, R.-B., Li, C.-H., Hung, Y., Wang, W., 2019. Optimal noncollapsing space-filling designs for irregular experimental regions. J. Comput. Graph. Statist. 28 (1), 74–91.
Cheng, T.E., Wang, G., 2000. Single machine scheduling with learning effect considerations. Ann. Oper. Res. 98 (1–4), 273–290.
Christen, J.A., Sansó, B., 2011. Advances in the sequential design of computer experiments based on active learning. Comm. Statist. Theory Methods 40 (24), 4467–4483.
Deng, X., Lin, C.D., Liu, K.-W., Rowe, R.K., 2017. Additive Gaussian process for computer models with qualitative and quantitative factors. Technometrics 59 (3), 283–292.
Ding, X., Matsuo, K., Xu, L., Yang, J., Zheng, L., 2015. Optimized combinations of bortezomib, camptothecin, and doxorubicin show increased efficacy and reduced toxicity in treating oral cancer. Anticancer Drugs 26 (5), 547–554.
Dupagne, M., Salwen, M.B., Paul, B., 1999. Impact of question order on the third-person effect. Int. J. Public Opin. Res. 11 (4), 334–345.
Emmons, H., 1969. One-machine sequencing to minimize certain functions of job tardiness. Oper. Res. 17 (4), 701–715.
Fang, K.-T., Li, R., Sudjianto, A., 2005. Design and Modeling for Computer Experiments. Chapman and Hall/CRC, Boca Raton, FL.
Fuleki, T., Francis, F., 1968. Quantitative methods for anthocyanins. 1. Extraction and determination of total anthocyanin in cranberries. J. Food Sci. 33 (1), 72–77.
Ginsbourger, D., Dupuy, D., Badea, A., Carraro, L., Roustant, O., 2009. A note on the choice and the estimation of kriging models for the analysis of deterministic computer experiments. Appl. Stoch. Models Bus. Ind. 25 (2), 115–131.
Gramacy, R.B., 2007. Tgp: an R package for Bayesian nonstationary, semiparametric nonlinear regression and design by treed Gaussian process models. J. Stat. Softw. 19 (9), 6.
Gramacy, R.B., 2020. Surrogates: Gaussian Process Modeling, Design, and Optimization for the Applied Sciences. Chapman and Hall/CRC, Boca Raton, FL.
Jones, B., Kenward, M.G., 2014. Design and Analysis of Cross-Over Trials. Chapman and Hall/CRC, Boca Raton, FL.
Jourdain, L.S., Schmitt, C., Leser, M.E., Murray, B.S., Dickinson, E., 2009. Mixed layers of sodium caseinate+ dextran sulfate: influence of order of addition to oil-water interface. Langmuir 25 (17), 10026–10037.
Karim, M., McCormick, K., Kappagoda, C.T., 2000. Effects of cocoa extracts on endothelium-dependent relaxation. J. Nutr. 130 (8), 2105S–2108S.
Kleijnen, J.P., 2009. Kriging metamodeling in simulation: a review. European J. Oper. Res. 192 (3), 707–716.
Krige, D.G., 1951. A statistical approach to some basic mine valuation problems on the witwatersrand. J. South. Afr. Inst. Min. Metall. 52 (6), 119–139.
Lin, D.K.J., Peng, J., 2019. Order-of-addition experiments: a review and some new thoughts. Qual. Eng. 31 (1), 49–59.
Loeppky, J.L., Sacks, J., Welch, W.J., 2009. Choosing the sample size of a computer experiment: a practical guide. Technometrics 51 (4), 366–376.
Mebane Jr., W.R., Sekhon, J.S., 2011. Genetic optimization using derivatives: the rgenoud package for R. J. Stat. Softw. 42 (11), 1–26.
Mee, R.W., 2020. Order-of-addition modeling. Statist. Sinica 30, 1543–1559.
Miller, J.M., Krosnick, J.A., 1998. The impact of candidate name order on election outcomes. Public Opin. Q. 62 (3), 291–330.
Ning, S., Xu, H., Al-Shyoukh, I., Feng, J., Sun, R., 2014. An application of a hill-based response surface model for a drug combination experiment on lung cancer. Stat. Med. 33 (24), 4227–4236.
Peng, J., Mukerjee, R., Lin, D.K.J., 2019. Design of order-of-addition experiments. Biometrika 106 (3), 683–694.
Preuss, M.L., Weidman, P., Nielsen, E., 2009. How we study protein transport. In: Trafficking Inside Cells: Pathways, Mechanisms and Regulation. Springer-Verlag New York, pp. 15–41.

Qian, P.Z.G., Wu, H., Wu, C.F.J., 2008. Gaussian process models for computer experiments with qualitative and quantitative factors. Technometrics 50 (3), 383–396.
Roustant, O., Ginsbourger, D., Deville, Y., 2012. DiceKriging, DiceOptim: two R packages for the analysis of computer experiments by kriging-based metamodeling and optimization. J. Stat. Softw. 51 (1), 1–55.
Sacks, J., Welch, W.J., Mitchell, T.J., Wynn, H.P., 1989. Design and analysis of computer experiments. Statist. Sci. 4 (4), 409–423.
Scarpa, R., Campbell, D., Hutchinson, W.G., 2007. Benefit estimates for landscape improvements: sequential Bayesian design and respondents' rationality in a choice experiment. Land Econom. 83 (4), 617–634.
Shinohara, A., Ogawa, T., 1998. Stimulation by rad52 of yeast rad51-mediated recombination. Nature 391 (6665), 404.
Silvestrini, R.T., Montgomery, D.C., Jones, B., 2013. Comparing computer experiments for the Gaussian process model using integrated prediction variance. Qual. Eng. 25 (2), 164–174.
Sullivan, T.J., 2015. Introduction to Uncertainty Quantification. Springer, Cham.
Townsend, W., 1978. The single machine problem with quadratic penalty function of completion times: a branch-and-bound solution. Manage. Sci. 24 (5), 530–534.
Van Nostrand, R., 1995. Design of experiments where the order of addition is important. In: ASA Proceedings of the Section on Physical and Engineering Sciences. American Statistical Association, Alexandria, VA, pp. 155–160.
Voelkel, J.G., 2019. The design of order-of-addition experiments. J. Qual. Technol. 51 (3), 230–241.
Wang, A., Xu, H., Ding, X., 2020. Simultaneous optimization of drug combination dose-ratio sequence with innovative design and active learning. Adv. Therapeutics 3 (4), 1900135.
Williams, C.K., Rasmussen, C.E., 2006. Gaussian Processes for Machine Learning. MIT press, Cambridge, MA.
Xiao, Q., Wang, L., Xu, H., 2019. Application of kriging models for a drug combination experiment on lung cancer. Stat. Med. 38 (2), 236–246.
Yang, J., Sun, F., Xu, H., 2020. A component-position model, analysis and design for order-of-addition experiments. Technometrics http://dx.doi.org/10.1080/00401706.2020.1764394.
Zhang, Y., Tao, S., Chen, W., Apley, D.W., 2020. A latent variable approach to Gaussian process modeling with qualitative and quantitative factors. Technometrics 62 (3), 291–302.
Zhao, Y., Lin, D.K., Liu, M.-Q., 2020. Designs for order-of-addition experiments. J. Appl. Stat. http://dx.doi.org/10.1080/02664763.2020.1801607.
Zhou, Y.-D., Xu, H., 2014. Space-filling fractional factorial designs. J. Amer. Statist. Assoc. 109 (507), 1134–1144.