

2015

Kriging-Based Design of Experiments for Multi-Source Exposure-Response Studies in Nanotoxicology

Ying Pei

Follow this and additional works at: <https://researchrepository.wvu.edu/etd>

Recommended Citation

Pei, Ying, "Kriging-Based Design of Experiments for Multi-Source Exposure-Response Studies in Nanotoxicology" (2015). *Graduate Theses, Dissertations, and Problem Reports*. 6394.
<https://researchrepository.wvu.edu/etd/6394>

This Thesis is protected by copyright and/or related rights. It has been brought to you by the The Research Repository @ WVU with permission from the rights-holder(s). You are free to use this Thesis in any way that is permitted by the copyright and related rights legislation that applies to your use. For other uses you must obtain permission from the rights-holder(s) directly, unless additional rights are indicated by a Creative Commons license in the record and/ or on the work itself. This Thesis has been accepted for inclusion in WVU Graduate Theses, Dissertations, and Problem Reports collection by an authorized administrator of The Research Repository @ WVU. For more information, please contact researchrepository@mail.wvu.edu.

Kriging-Based Design of Experiments for Multi-Source
Exposure-Response Studies in Nanotoxicology

Ying Pei

Thesis submitted to the
College of Engineering and Mineral Resources
at West Virginia University
in partial fulfillment of the requirements
for the degree of

Master of Science
in
Industrial Engineering

Feng Yang, Ph.D., Chair
Wafik Iskander, Ph.D.
Majid Jaridi, Ph.D.

Department of Industrial and Management Systems Engineering

Morgantown, West Virginia
2015

Keywords: Design of experiments, Multi-source data, Exposure-response surface,
Stochastic kriging, Nanotoxicology

Copyright 2015 Ying Pei

ABSTRACT

Kriging-Based Design of Experiments for Multi-Source Exposure-Response Studies in Nanotoxicology

Ying Pei

One of the main challenges with toxicology studies of nanomaterials (NMs), compared to traditional materials or chemicals, lies in the large variety of NMs with various physicochemical properties. This work is concerned with the efficient design of multi-source biological experiments for the toxicity characterization of NMs in terms of their exposure-response profiles. The major contribution of this thesis is the development of a two-stage experimental design procedure based on the statistical model, stochastic kriging with qualitative factors (SKQ). With a given experimental budget, the SKQ-based design method aims at achieving the highest-quality SKQ, which synergistically models the exposure-response data from multiple sources (e.g., NM types). The method determines the experimental design (that is, the sampling location as well as allocation) in such a way that the resulting sampling data allow SKQ to realize its maximum potential to pool information across multiple sources for efficient modeling. Built in a two-stage framework, which enables a learning process of the target exposure-response relationships, the SKQ-based design procedure also inherits the general advantages of stochastic kriging in the sense that the design is particularly tailored to model the possibly nonlinear and complex relationships and heterogeneous data variances. Through simulation studies, the efficiency of the SKQ-based procedure for multi-source experiments is demonstrated over two alternative design methods.

ACKNOWLEDGMENTS

I would like to thank my advisor, Dr. Feng Yang for her great guidance, for giving me the opportunity to work on this project, which could not have been finished without her encouragement and support. I am also thankful to Dr. Majid Jaridi and Dr. Wafik Iskander for serving on my committee, and for their insightful ideas and assistance in preparing this thesis.

Contents

List of Figures	vi
List of Tables	vii
List of Acronyms	viii
1 Introduction	1
2 Statement of the Research Problem	4
3 Review of Stochastic Kriging with Qualitative Factors(SKQ)	6
4 SKQ-Based Two-Stage Design Procedure	11
4.1 Initial Design and Preliminary Modeling	12
4.2 Design Augmentation	13
5 Empirical Studies	17
5.1 Case 1	17
5.1.1 Simulation Model	18
5.1.2 Applying the Two-Stage Procedure	19
5.1.3 Evaluation and Comparison	21
5.2 Case 2	24
5.2.1 Simulation Model	24
5.2.2 Applying the Two-Stage Procedure	26
5.2.3 Evaluation and Comparison	27

6 Conclusion	31
References	32
A Derivation of equation (4.6)	36

List of Figures

5.1	The two-stage experimental design: the stars denote the Stage-1 design points and the dots denote the Stage-2 design points.	20
5.2	The design points in the dose-time region for each source (short or long nanobelts) in the traditional design	22
5.3	The design points of both sources in space-filling design: triangles denote the design points for short nanobelts and squares denote the design points for long nanobelts	23
5.4	The two box plots for each of the three design methods	24
5.5	The six box plots for each of the three design methods	30

List of Tables

5.1	SLHD space-filling design in Stage-1 for Case 2	26
5.2	Stage-2 design for Case 2	28
5.3	Design points in the space-filling design for Case 2.	29

List of Acronyms

NM Nanomaterials

DOE Design of Experiments

SKQ Stochastic Kriging with Qualitative Factors

SK Stochastic Kriging

DK Deterministic Kriging

MSE Mean Squared Error

IMSE Integrated Mean Squared Error

GA Genetic Algorithm

SLHD Sliced Latin Hypercube Design

Chapter 1

Introduction

With the rapid development of nanotechnology, nanomaterials (NMs) find increasing applications in energy [1, 2], biomedicine [3, 4, 5], and environment [6, 7, 8, 9]. The production, use, and disposal of NMs inevitably lead to their appearance in air, water, soils, etc., and hence raise the safety, health, and environmental concerns of NMs. For the safe and sustainable development of nanotechnology, the risk assessment of NMs plays a critical role.

One of the most fundamental steps in assessing the risk of a NM (or any other substance) is to characterize the NM by its exposure-response profile [10, 11], which describes the dependence of the adverse bioactivity effects (the responses) upon the NM exposure conditions [12]. The exposure condition is typically specified through the settings of two factors: the NM dosage and the time factor involved (e.g., post-exposure time for acute studies). To obtain the characteristic profile of a NM, biological experiments need to be performed at different exposure conditions to observe the corresponding bioactivity responses of animals. Such biological experiments are extremely expensive and time-consuming, and the particular challenge with the exposure-response studies of NMs, compared to traditional materials or chemicals, lies in the large NM variety caused by their various physico-chemical properties (e.g., chemical compositions, shape, size, and surface chemistry).

How to design exposure-response experiments across multiple sources (i.e., determine the experimental exposure conditions for each type of NMs and the sample allocations), for the efficient utilization of limited resources? To address this question, a kriging-based design of experiments (DOE) procedure is developed in this thesis. The DOE method is

built upon the stochastic kriging with qualitative factors (SKQ) model developed in [13]. Thus, it inherits the advantages of SKQ such as high flexibility and generality, and seeks to generate a design that allows SKQ to realize its maximum efficiency in modeling multi-source data. Specifically, the SKQ-based DOE aims at designing multi-source experiments to obtain a most informative data ensemble across multiple sources. The data ensemble is most informative in the sense that when being synergistically modeled by SKQ, information will be pooled from all the data sources leading to the highest-quality model (e.g., exposure-response model) for each source (e.g., NM type). When the target relationships are nonlinear or there is variance heterogeneity in the data, both of which are commonly encountered in nanotoxicology studies, the optimal design that optimizes the model estimation quality is dependent on the true underlying response surfaces and variance patterns, which are unknown. To circumvent this problem, the DOE procedure employs a two-stage paradigm: In the first stage, some preliminary experiments are carried out, on which SKQ modeling is performed to derive information regarding the target relationships and data variance; in the second stage, the information obtained from the previous stage is utilized to guide the Stage-2 design, which aims at optimizing the quality of the SKQ models fitted from both stages of data for the target response surfaces.

The SKQ-based two-stage design method represents a new addition to the existing literature of experimental design, which can be generally divided into two groups: model-independent versus model-based designs. In these two groups, the design methods most relevant to the current study are briefly reviewed in the following. Both “naive” designs commonly used by biology experimenters and space-filling designs [14, 15] fall into the category of model-independent designs, which have no bearing on the models for the target response surfaces. In the current DOE practice for toxicology studies, “naive” designs are generated based on empirical experiences in a somewhat arbitrary manner [16, 17]. They usually involve equally-spaced levels (on a linear or log scale) in dose and/or time range, and the same design is typically adopted across multiple sources (e.g., NM types). A space-filling

design seeks to provide an even coverage of the design regions of interest. In particular, in the presence of multiple sources, sliced space-filling designs [18, 19] have been developed. In contrast to disregarding the target response surfaces, model-based designs aim at optimizing the quality of the resulting estimated models for the target surfaces. Since our design method is based on SKQ, a kriging model, we focus on reviewing kriging-based designs. Depending on whether or not the response is stochastic, there is deterministic kriging (DK) versus stochastic kriging (SK), based on which respective design methods have been developed. The majority of the work has been on DK-based designs [20, 21, 22, 23, 24], which only need to determine the location of design points with one sample assigned to each point. Some research efforts have been devoted to SK-based designs [25, 26], which determine not only the design-point locations but also the sample size at each point; such SK models the effects of quantitative factors only. As pointed out in Wang et al. [13], SKQ represents the kriging model that models the variability across replications (randomness in responses) and the variability arising from quantitative as well as qualitative factors (e.g., source factors). Accordingly, our SKQ-based design method utilizes the information regarding the target response surfaces and the variance structures to find the design (that is, the design-point locations and sample allocations) for multi-source experiments that leads to the fitted SKQ model of the highest quality.

The remainder of this thesis is organized as follows. Chapter 2 describes in precise terms the problem of designing multi-source experiments for exposure-response studies. A brief review of SKQ and its advantages is given in Chapter 3. The SKQ-based two-stage design procedure is detailed in Chapter 4. In chapter 5, the design procedure is applied into two cases and its efficiency is illustrated via comparison with two other DOE methods: the “naive” design commonly used in toxicology studies and a space-filling design in the DOE literature. Chapter 6 summarizes the SKQ-based two-stage design procedure and its advantages.

Chapter 2

Statement of the Research Problem

For the exposure-response studies of an NM, biological experiments need to be carried out under a range of experimental conditions. An experimental condition is defined by the combination of a number of factors, which can be divided into two categories, quantitative and qualitative factors.

- Quantitative factors typically include but are not limited to the toxicant dosage administered to an animal, and the time factor. Depending on the time scope of the toxicology study, the time factor could be exposure time for long-term studies, or post-exposure time for acute studies. The vector \mathbf{x} is used to represent the quantitative factors considered.
- Qualitative factors mainly include the various source factors such as the NM type, the conducting laboratory for experiments. The qualitative factors are denoted by the vector \mathbf{z} .

The experimental condition is specified in terms of the factor vector $\mathbf{w} = (\mathbf{x}^\top, \mathbf{z}^\top)^\top$. The random response obtained from an animal subject at a \mathbf{w} can be generally written as

$$\mathcal{Y}(\mathbf{w}) = E[\mathcal{Y}(\mathbf{w})] + \varepsilon(\mathbf{w}) = Y(\mathbf{w}) + \varepsilon(\mathbf{w}), \quad (2.1)$$

where $Y(\mathbf{w}) = E[\mathcal{Y}(\mathbf{w})]$ represents the true expected response, and $\varepsilon(\mathbf{w})$ the random zero-mean error accounting for the variability across animal subjects.

A setting of the qualitative factors \mathbf{z} specifies a combination category, say c_q , representing one source of data. The total of Q data sources specified by the categories of \mathbf{z} are denoted as $\{c_q; q = 1, 2, \dots, Q\}$. The biological data collected at a range of \mathbf{w} settings are represented as

$$\{(\mathbf{w}_i, \mathcal{Y}_j(\mathbf{w}_i)); i = 1, 2, \dots, I; j = 1, 2, \dots, n(\mathbf{w}_i)\} \quad (2.2)$$

where I denotes the number of distinct design points (i.e., factor setting at which experiments are performed), \mathbf{w}_i the i^{th} design point, $\mathcal{Y}_j(\mathbf{w}_i)$ the response from the j^{th} replication at \mathbf{w}_i , and $n(\mathbf{w}_i)$ the number of replications performed at \mathbf{w}_i .

The goal of DOE is to determine the location of design points $\{\mathbf{w}_i; i = 1, 2, \dots, I\}$ and the sample allocation $\{n(\mathbf{w}_i); i = 1, 2, \dots, I\}$ to achieve the maximum experimental efficiency.

Chapter 3

Review of Stochastic Kriging with Qualitative Factors(SKQ)

The DOE method in this work is developed based on the SKQ modeling of multi-source data. Hence, a brief review of the SKQ developed in Wang et al. [13] is provided herein. As pointed out in Wang et al. [13], SKQ is the first kriging model that is able to accommodate the variability arising from quantitative as well as qualitative factors, and the variability across replications.

SKQ models the dependence of a continuous response upon the factors $\mathbf{w} = (\mathbf{x}^\top, \mathbf{z}^\top)^\top$, with $\mathbf{x} = (x_1, x_2, \dots, x_d)^\top \in \mathbb{R}^d$ and $\mathbf{z} = (z_1, z_2, \dots, z_L)^\top$. There are L qualitative factors, and each factor z_ℓ ($\ell = 1, 2, \dots, L$) has a number of category levels. The response at \mathbf{w} from the j th replication (animal subject) is modeled by SKQ as

$$\mathcal{Y}_j(\mathbf{w}) = Y(\mathbf{w}) + \varepsilon_j(\mathbf{w}) = \mathbf{f}(\mathbf{w})^\top \boldsymbol{\beta} + M(\mathbf{w}) + \varepsilon_j(\mathbf{w}), \quad (3.1)$$

The expectation $Y(\mathbf{w})$ consists of two parts: $\mathbf{f}(\mathbf{w})^\top \boldsymbol{\beta}$ and $M(\mathbf{w})$. $\mathbf{f}(\mathbf{w})$ is a vector of known functions of \mathbf{w} , and $\boldsymbol{\beta}$ a vector of unknown parameters of compatible dimension. In this work, we set $\mathbf{f}(\mathbf{w})^\top \boldsymbol{\beta} = \beta_0$, which has been widely accepted as sufficient for most applications. The term $M(\mathbf{w})$ represents a mean-zero stationary Gaussian process, and intends to capture the variability due to the factors \mathbf{w} , which is referred to as the extrinsic variability.

The intrinsic variability refers to the randomness of $\varepsilon(\mathbf{w})$. The random noise $\{\varepsilon_j(\mathbf{w}); j = 1, 2, \dots\}$ at \mathbf{w} is assumed to have mean zero, and be independent and identically distributed (i.i.d.) across replications. The error variance $\text{Var}[\varepsilon(\mathbf{w})]$ can be \mathbf{w} -dependent.

With the sample data (2.2), the sample average of the responses at \mathbf{w}_i across the $n(\mathbf{w}_i)$ replications is obtained as

$$\bar{\mathcal{Y}}(\mathbf{w}_i) = \frac{1}{n(\mathbf{w}_i)} \sum_{j=1}^{n(\mathbf{w}_i)} \mathcal{Y}_j(\mathbf{w}_i) = \beta_0 + \mathbf{M}(\mathbf{w}_i) + \frac{1}{n(\mathbf{w}_i)} \sum_{j=1}^{n(\mathbf{w}_i)} \varepsilon_j(\mathbf{w}_i).$$

For I distinct design points, the $I \times 1$ vector of sample averages is denoted as

$$\bar{\mathcal{Y}} = (\bar{\mathcal{Y}}(\mathbf{w}_1), \bar{\mathcal{Y}}(\mathbf{w}_2), \dots, \bar{\mathcal{Y}}(\mathbf{w}_I))^\top. \quad (3.2)$$

The vector of sample average errors is represented as

$$\bar{\varepsilon} = (\bar{\varepsilon}(\mathbf{w}_1), \bar{\varepsilon}(\mathbf{w}_2), \dots, \bar{\varepsilon}(\mathbf{w}_I))^\top, \quad (3.3)$$

with $\bar{\varepsilon}(\mathbf{w}_i) = n(\mathbf{w}_i)^{-1} \sum_{j=1}^{n(\mathbf{w}_i)} \varepsilon_j(\mathbf{w}_i)$, $i = 1, 2, \dots, I$.

The extrinsic variability is modeled in SKQ by $\mathbf{M}(\mathbf{w})$, which is specified by its covariance

$$\text{Cov}[\mathbf{M}(\mathbf{w}), \mathbf{M}(\mathbf{w}')] = \sigma^2 \cdot \text{Corr}[\mathbf{M}(\mathbf{w}), \mathbf{M}(\mathbf{w}')] = \sigma^2 \cdot \left[\prod_{\ell=1}^L \tau_{z_\ell, z'_\ell}^{(\ell)} \right] \cdot K(\mathbf{x}, \mathbf{x}'), \quad (3.4)$$

where σ^2 is the variance of the Gaussian process. The correlation $\text{Corr}[\mathbf{M}(\mathbf{w}), \mathbf{M}(\mathbf{w}')]$ is decomposed into two parts: $\prod_{\ell=1}^L \tau_{z_\ell, z'_\ell}^{(\ell)}$ and $K(\mathbf{x}, \mathbf{x}')$. For the estimation of a SKQ, specific functional forms need to be assumed for both parts. The correlation across the quantitative settings is represented by $K(\mathbf{x}, \mathbf{x}')$, for which a range of functional forms are provided in the literature [27, 28]. A most widely-adopted function is the exponential correlation function

$$K(\mathbf{x}, \mathbf{x}') = \exp \left\{ \sum_{h=1}^d -\theta_h |x_h - x'_h|^p \right\}, \quad (3.5)$$

where $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_d)$ is a vector of unknown parameters. It is required that $\theta_h > 0$ ($h = 1, 2, \dots, d$), and $\boldsymbol{\theta}$ determines the roughness of the response surface for a given combination category of \mathbf{z} . The parameter $p \in (0, 2]$ also needs to be estimated unless p is pre-specified as 2, [29]. In (3.4), the term $\prod_{\ell=1}^L \tau_{z_\ell, z'_\ell}^{(\ell)}$ models the correlations across different categories of qualitative factors. As noted in Qian et al.[28], $\tau_{z_\ell, z'_\ell}^{(\ell)}$ measures the correlation (similarity) at any two settings \mathbf{w} and \mathbf{w}' that differ only on the values of the ℓ th qualitative factor. For $\tau_{z_\ell, z'_\ell}^{(\ell)}$, a range of functional forms have been proposed in Qian et al. [28] and Zhou et al. [30]. In this paper, we use the exchangeable correlation function as follows.

Exchangeable correlation functions (**EC**):

$$\tau_{z_\ell, z'_\ell}^{(\ell)} = \exp\{-\phi^{(\ell)} I(z_\ell \neq z'_\ell)\}; \quad \ell = 1, 2, \dots, L \quad (3.6)$$

In (3.6), $\boldsymbol{\Phi} = \{\phi^{(\ell)}; \ell = 1, 2, \dots, L\}$ represents the set of unknown parameters to be estimated; and $I[A]$ is an indicator function that takes 1 if event A is true and 0 otherwise. Clearly, **EC** assumes that all the category levels of the ℓ th qualitative factor are of isotropic nature; that is, for a given ℓ , $\tau_{z_\ell, z'_\ell}^{(\ell)}$ is a constant as long as $z_\ell \neq z'_\ell$.

Given the data (2.2) collected at I distinct design points, the $I \times I$ variance-covariance matrix Σ_M is defined as

$$\Sigma_M = \sigma^2 \cdot \mathbf{R}(\boldsymbol{\theta}, \boldsymbol{\Phi}) = \sigma^2 \cdot \begin{pmatrix} 1 & \text{Corr}[\mathbf{M}(\mathbf{w}_1), \mathbf{M}(\mathbf{w}_2)] & \cdots & \text{Corr}[\mathbf{M}(\mathbf{w}_1), \mathbf{M}(\mathbf{w}_I)] \\ \text{Corr}[\mathbf{M}(\mathbf{w}_2), \mathbf{M}(\mathbf{w}_1)] & 1 & \cdots & \text{Corr}[\mathbf{M}(\mathbf{w}_2), \mathbf{M}(\mathbf{w}_I)] \\ \vdots & \vdots & \ddots & \vdots \\ \text{Corr}[\mathbf{M}(\mathbf{w}_I), \mathbf{M}(\mathbf{w}_1)] & \text{Corr}[\mathbf{M}(\mathbf{w}_I), \mathbf{M}(\mathbf{w}_2)] & \cdots & 1 \end{pmatrix} \quad (3.7)$$

where $\mathbf{R}(\boldsymbol{\theta}, \boldsymbol{\Phi})$ denotes the correlation matrix with each element representing a correlation. Each element correlation can be decomposed into two parts as explained above, and involves the unknown parameters $\boldsymbol{\theta}$ and $\boldsymbol{\Phi}$. For an arbitrary \mathbf{w}_0 , the $I \times 1$ vector $\Sigma_M(\mathbf{w}_0, \cdot)$ is defined

as

$$\Sigma_M(\mathbf{w}_0, \cdot) = \sigma^2 \mathbf{v}(\mathbf{w}_0, \boldsymbol{\theta}, \Phi) = \sigma^2 \begin{pmatrix} \text{Corr}[M(\mathbf{w}_0), M(\mathbf{w}_1)] \\ \text{Corr}[M(\mathbf{w}_0), M(\mathbf{w}_2)] \\ \vdots \\ \text{Corr}[M(\mathbf{w}_0), M(\mathbf{w}_I)] \end{pmatrix}, \quad (3.8)$$

with $\mathbf{v}(\mathbf{w}_0, \boldsymbol{\theta}, \Phi)$ being a correlation vector involving \mathbf{w}_0 , $\boldsymbol{\theta}$ and Φ .

Denote Σ_ε as the $I \times I$ variance-covariance matrix of vector $\bar{\boldsymbol{\varepsilon}}$ defined in (3.3). With i.i.d. random errors, Σ_ε is a $I \times I$ diagonal matrix

$$\Sigma_\varepsilon = \text{diag}\{\text{Var}[\varepsilon(\mathbf{w}_1)]/n(\mathbf{w}_1), \text{Var}[\varepsilon(\mathbf{w}_2)]/n(\mathbf{w}_2), \dots, \text{Var}[\varepsilon(\mathbf{w}_I)]/n(\mathbf{w}_I)\}. \quad (3.9)$$

For a data set (2.2), the SKQ estimation and inference procedure in Wang et al. [13] is summarized as follows.

1. Estimate Σ_ε as:

$$\widehat{\Sigma}_\varepsilon = \text{diag}\{\widehat{\text{Var}}[\varepsilon(\mathbf{w}_1)]/n(\mathbf{w}_1), \widehat{\text{Var}}[\varepsilon(\mathbf{w}_2)]/n(\mathbf{w}_2), \dots, \widehat{\text{Var}}[\varepsilon(\mathbf{w}_I)]/n(\mathbf{w}_I)\} \quad (3.10)$$

where

$$\widehat{\text{Var}}[\varepsilon(\mathbf{w}_i)] = \frac{1}{n(\mathbf{w}_i) - 1} \sum_{j=1}^{n(\mathbf{w}_i)} (\mathcal{Y}_j(\mathbf{w}_i) - \bar{\mathcal{Y}}(\mathbf{w}_i))^2, \quad i = 1, 2, \dots, I. \quad (3.11)$$

2. Replace Σ_ε by $\widehat{\Sigma}_\varepsilon$ in the maximum likelihood (ML) function(3.12), and obtain the ML estimates $(\widehat{\beta}_0, \widehat{\sigma}^2, \widehat{\boldsymbol{\theta}}, \widehat{\boldsymbol{\Phi}})$, which fully specify the SKQ model.

$$\begin{aligned} \ln \mathcal{L}(\sigma^2, \boldsymbol{\theta}, \boldsymbol{\Phi}) &= -\ln[(2\pi)^{I/2}] - \frac{1}{2} \ln[|\sigma^2 \mathbf{R}(\boldsymbol{\theta}, \boldsymbol{\Phi}) + \widehat{\Sigma}_\varepsilon|] \\ &- \frac{1}{2} (\bar{\mathcal{Y}} - \widehat{\beta}_0(\sigma^2, \boldsymbol{\theta}, \boldsymbol{\Phi}) \mathbf{1}_I)^\top [\sigma^2 \mathbf{R}(\boldsymbol{\theta}, \boldsymbol{\Phi}) + \widehat{\Sigma}_\varepsilon]^{-1} (\bar{\mathcal{Y}} - \widehat{\beta}_0(\sigma^2, \boldsymbol{\theta}, \boldsymbol{\Phi}) \mathbf{1}_I), \end{aligned} \quad (3.12)$$

3. For an arbitrary \mathbf{w}_0 , the expected response $\mathbf{Y}(\mathbf{w}_0)$ can be estimated by

$$\widehat{\mathbf{Y}}(\mathbf{w}_0) = \widehat{\beta}_0 + \mathbf{v}(\mathbf{w}_0, \widehat{\boldsymbol{\theta}}, \widehat{\boldsymbol{\Phi}})^\top [\widehat{\sigma}^2 \mathbf{R}(\widehat{\boldsymbol{\theta}}, \widehat{\boldsymbol{\Phi}}) + \widehat{\Sigma}_\varepsilon]^{-1} (\bar{\mathcal{Y}} - \widehat{\beta}_0 \mathbf{1}_I), \quad (3.13)$$

The mean squared error (MSE) of $\widehat{\mathbf{Y}}(\mathbf{w}_0)$ is obtained as:

$$\begin{aligned} \widehat{\text{MSE}}[\widehat{\mathbf{Y}}(\mathbf{w}_0)] &= \widehat{\sigma}^2 - \widehat{\sigma}^4 \mathbf{v}(\mathbf{w}_0, \widehat{\boldsymbol{\theta}}, \widehat{\boldsymbol{\Phi}})^\top [\widehat{\sigma}^2 \mathbf{R}(\widehat{\boldsymbol{\theta}}, \widehat{\boldsymbol{\Phi}}) + \widehat{\Sigma}_\varepsilon]^{-1} \mathbf{v}(\mathbf{w}_0, \widehat{\boldsymbol{\theta}}, \widehat{\boldsymbol{\Phi}}) \\ &+ \eta^2 (\mathbf{1}_I^\top [\widehat{\sigma}^2 \mathbf{R}(\widehat{\boldsymbol{\theta}}, \widehat{\boldsymbol{\Phi}}) + \widehat{\Sigma}_\varepsilon]^{-1} \mathbf{1}_I)^{-1}, \end{aligned} \quad (3.14)$$

where $\eta = 1 - \mathbf{1}_I^\top [\widehat{\sigma}^2 \mathbf{R}(\widehat{\boldsymbol{\theta}}, \widehat{\boldsymbol{\Phi}}) + \widehat{\Sigma}_\varepsilon]^{-1} \mathbf{v}(\mathbf{w}_0, \widehat{\boldsymbol{\theta}}, \widehat{\boldsymbol{\Phi}}) \widehat{\sigma}^2$.

Chapter 4

SKQ-Based Two-Stage Design Procedure

A SKQ-based two-stage procedure is developed to determine the design $\{\mathbf{w}_i, n(\mathbf{w}_i); i = 1, 2, \dots, I\}$ for multi-source experiments. The pre-specified inputs required by the procedure are given as follows.

- N : the total number of samples (animals) available, which depends on the experimental budget.
- N_1 : the number of samples assigned to Stage 1, which implies that the Stage-2 sample size is $N - N_1$. It is recommended to set N_1 as $1/4 - 1/2$ of N [31].
- \mathbf{H} : the feasible space of the variable vector \mathbf{w} . \mathbf{H} includes Q slices (Q sources) of the region χ , denoting the feasible region of the quantitative vector \mathbf{x} .
- I : the total number of distinct design points to be assigned. In biological experiments, it is typical to assign at least three samples to a design point, and thus $I \leq N/3$. Also, I is recommended to be set at least about ten times the dimension of χ [32], to provide an adequate coverage of the region χ .

An overview of the SKQ-based two-stage procedure is briefed as follows. In Stage 1, N_1 pilot experiments are performed following the initial design, which will be discussed in Section 4.1. Based on the preliminary data, two kriging models will be fitted: (i) One quantifies the dependence of the expected response $Y(\mathbf{w})$ on \mathbf{w} , and (ii) the other approximates the variance $\text{Var}[\mathcal{Y}(\mathbf{w})]$ as a function of \mathbf{w} . In Stage 2, the information contained in the two fitted SKQ models is utilized to design the remaining $N - N_1$ samples. The follow-up design

aims at optimizing the quality of the resulting response surfaces (exposure-response surfaces for multiple sources) estimated from both stages of data.

4.1 Initial Design and Preliminary Modeling

At Stage 1, N_1 samples are to be allocated. Denote the initial design as

$$\mathbf{D}_1 = \begin{pmatrix} \mathbf{w}_1 & \mathbf{w}_2 & \cdots & \mathbf{w}_{I_1} \\ n_1 & n_2 & \cdots & n_{I_1} \end{pmatrix} \quad (4.1)$$

where I_1 is the number of distinct experimental conditions \mathbf{w} .

Given the sample size N_1 , the initial design (4.1) is determined as follows. At this initial stage, n_i is set as n for $i = 1, 2, \dots, I_1$, with absent information on the variance pattern. The value of n is specified by the experimenter based on her experience of biological experiments with NMs, which is typically done to determine a naive design in the current practice. The number of distinct design points I_1 is then calculated as $I_1 = N_1/n$. The I_1 sampling points will then be allocated to the multiple slices (sources) of the quantitative region χ . The purpose of the initial design is to provide a fair coverage of the design space to gain some information on both the target surfaces and the variance structure throughout the multi-slice design regions. Hence, we adopt the Sliced Latin Hypercube Design (SLHD) [18] to determine the locations of the design points. SLHD is a special space-filling design that can be partitioned into slices of smaller Latin hypercube designs (LHD). Each slice of the design achieves maximum uniformity in any one-dimensional projection, and provides an even coverage of χ for each slice (source). When collapsed across the slices, all the sampling points in χ have the maximum stratification in any one-dimensional projection.

From the initial data, which are denoted as $\{(\mathbf{w}_i, \mathcal{Y}_j(\mathbf{w}_i)); i = 1, 2, \dots, I_1; j = 1, 2, \dots, n(\mathbf{w}_i)\}$, two preliminary models will be estimated. First, the SKQ estimation (Chapter 3) will be performed quantifying the relationship between the expected response $Y(\mathbf{w})$

and \mathbf{w} (that is, the multiple response surfaces across sources). Second, deterministic kriging [33] will be carried out based on the data pair $\{(\mathbf{w}_i, \widehat{\text{Var}}[\varepsilon(\mathbf{w}_i)]); i = 1, 2, \dots\}$, where $\widehat{\text{Var}}[\varepsilon(\mathbf{w}_i)]$ represents the sample variance as calculated in (3.11). The second kriging model approximates $\text{Var}[\varepsilon(\mathbf{w})]$ as a function of \mathbf{w} . These two kriging models, which contain the target surface and variance information derived from the initial data, will be utilized to guide the follow-up design in Stage 2.

4.2 Design Augmentation

At Stage 2, the task is to determine the design for the $N_2 = N - N_1$ samples, which are to be allocated to $I_2 = I - I_1$ distinct design points. The Stage-2 design is denoted as

$$\mathbf{D}_2 = \begin{pmatrix} \mathbf{w}_{I_1+1} & \mathbf{w}_{I_1+2} & \cdots & \mathbf{w}_I \\ n_{I_1+1} & n_{I_1+2} & \cdots & n_I \end{pmatrix}, \quad (4.2)$$

and is determined in such a way that the quality of the final SKQ, fitted from both stages of data $\mathbf{D} = \mathbf{D}_1 \cup \mathbf{D}_2$ to model the multi-source response surfaces, is optimized.

The quality of a model can be measured by a range of metrics [26] such as the integrated mean squared error(IMSE), the comparison index based criterion(Comp), etc. An appropriate performance measure can be selected depending on the modeling purpose. Herein, as an example, the integrated mean squared error (IMSE) is used as the design criterion, and the design optimization problem in Stage 2 can be formulated as:

$$\text{Minimize}_{\mathbf{D}^{(2)}} \quad \text{IMSE} = \sum_{c_q} \int_{\mathbf{x}_0 \in \mathcal{X}} \text{MSE}[\widehat{\mathbf{Y}}(\mathbf{x}_0)] d\mathbf{x}_0 \quad (4.3)$$

$$\text{Subject to} \quad \sum_{i=I_1+1}^I n_i = N_2, \quad (4.4)$$

where I_1 , I , and N_2 are all predetermined parameters. The objective criterion IMSE represents the integration of $\text{MSE}[\widehat{\mathbf{Y}}(\mathbf{w}_0)]$, the MSE of the response estimates $\widehat{\mathbf{Y}}(\mathbf{w}_0)$ over the

feasible region of \mathbf{H} ; and the estimate $\widehat{Y}(\mathbf{w}_0)$ at an arbitrary \mathbf{w}_0 is provided by the SKQ model fitted from both stages of data following the design $\mathbf{D} = \mathbf{D}_1 \cup \mathbf{D}_2$.

How is the criterion IMSE related to the design \mathbf{D} ? Recall that for a fitted SKQ model, the MSE of the response estimates $\text{MSE}[\widehat{Y}(\mathbf{w}_0)]$ can be estimated as (3.14), which is re-written as follows for convenience of discussion:

$$\begin{aligned} \widehat{\text{MSE}}[\widehat{Y}(\mathbf{w}_0)] &= \widehat{\sigma}^2 - \widehat{\sigma}^4 \mathbf{v}(\mathbf{w}_0, \widehat{\boldsymbol{\theta}}, \widehat{\boldsymbol{\Phi}})^\top [\widehat{\sigma}^2 \mathbf{R}(\widehat{\boldsymbol{\theta}}, \widehat{\boldsymbol{\Phi}}) + \widehat{\Sigma}_\varepsilon]^{-1} \mathbf{v}(\mathbf{w}_0, \widehat{\boldsymbol{\theta}}, \widehat{\boldsymbol{\Phi}}) \\ &\quad + \eta^2 (\mathbf{1}_I^\top [\widehat{\sigma}^2 \mathbf{R}(\widehat{\boldsymbol{\theta}}, \widehat{\boldsymbol{\Phi}}) + \widehat{\Sigma}_\varepsilon]^{-1} \mathbf{1}_I)^{-1}, \end{aligned} \quad (4.5)$$

where $\eta = 1 - \mathbf{1}_I^\top [\widehat{\sigma}^2 \mathbf{R}(\widehat{\boldsymbol{\theta}}, \widehat{\boldsymbol{\Phi}}) + \widehat{\Sigma}_\varepsilon]^{-1} \mathbf{v}(\mathbf{w}_0, \widehat{\boldsymbol{\theta}}, \widehat{\boldsymbol{\Phi}}) \widehat{\sigma}^2$.

As can be seen from (4.5), an MSE (and hence the IMSE) depends on three items: the SKQ parameters $(\beta_0, \sigma^2, \boldsymbol{\theta}, \boldsymbol{\Phi})$ for the target response surfaces, which have been estimated from Stage 1 and can be plugged into (4.5); the variance model specifying the relationship between $\text{Var}[\varepsilon(\mathbf{w})]$ and \mathbf{w} , which have been estimated from Stage 1 as well; and the design $\mathbf{D} = \mathbf{D}_1 \cup \mathbf{D}_2$, with \mathbf{D}_1 given from Stage 1 and \mathbf{D}_2 to be determined via the optimization. More specifically, in (4.5), both $\mathbf{R}(\widehat{\boldsymbol{\theta}}, \widehat{\boldsymbol{\Phi}})$ and $\mathbf{v}(\mathbf{w}_0, \widehat{\boldsymbol{\theta}}, \widehat{\boldsymbol{\Phi}})$, which are defined in (3.7) and (3.8) respectively, depend on the design \mathbf{D} and the SKQ parameters $(\beta_0, \sigma^2, \boldsymbol{\theta}, \boldsymbol{\Phi})$. The diagonal variance matrix $\widehat{\Sigma}_\varepsilon$, which is defined in (3.9), has each element as $\text{Var}[\varepsilon(\mathbf{w}_i)]/n(\mathbf{w}_i)$, $i = 1, 2, \dots, I$, and is dependent on \mathbf{D} ; the functional dependence of $\text{Var}[\varepsilon(\mathbf{w})]$ upon \mathbf{w} has been approximately established by the variance model.

Hence, for a candidate design of \mathbf{D}_2 , the corresponding IMSE criterion can be evaluated based on the prior obtained estimates for $(\beta_0, \sigma^2, \boldsymbol{\theta}, \boldsymbol{\Phi})$, the variance model, and \mathbf{D}_1 determined in Stage 1. This evaluation ability provides the necessary basis to perform a numeric search for the optimal values of \mathbf{D}_2 .

Solving (4.3) is challenging due to the high dimension of the decision variables involved in \mathbf{D}_2 (4.2), and the integer nature of n_i ($i = I_1 + 1, I_1 + 2, \dots, I$). We approach this problem

by utilizing the approximate relationships between the optimal values of $\{n(\mathbf{w}_i); i = 1, 2, \dots\}$ and the given design locations $\{\mathbf{w}_i; i = 1, 2, \dots\}$, and by an iterative procedure.

Given the remaining sample size N_2 , the first stage design \mathbf{D}_1 and the design-point locations of the second stage $\{\mathbf{w}_i; i = I_1 + 1, I_1 + 2, \dots, I\}$, the values of $\{n(\mathbf{w}_i); i = I_1 + 1, I_1 + 2, \dots, I\}$ that minimize the IMSE of the SKQ model, which is fitted from the N -sample data following the design $\{(\mathbf{w}_i, n(\mathbf{w}_i)); i = 1, 2, \dots, I\}$, can be approximated as:

$$n_i \approx N_2 \frac{\sqrt{V_i C_i}}{\sum_{j=1}^I \sqrt{V_j C_j}}. \quad (4.6)$$

The detailed derivation is given in Appendix A. In (4.6), V_i is the intrinsic variance, and $C_i = [\Sigma_{\mathbf{M}}^{-1} \mathbf{W} \Sigma_{\mathbf{M}}^{-1}]_{ii}$, \mathbf{W} is the $I \times I$ matrix with elements $\mathbf{W}_{ij} = \sum_{c_q} \int_{\mathbf{x}_0 \in \mathcal{X}} r_{0i} r_{0j} d\mathbf{x}_0$ and $r_{0i} = \text{Corr}[\mathbf{M}(\mathbf{w}_i), \mathbf{M}(\mathbf{w}_0)]$.

Utilizing (4.6), the iterative procedure is developed as follows to solve the design optimization (4.3).

Inputs: (a) \mathbf{D}_1 , the Stage-1 design;

(b) The SKQ model fitted from the Stage-1 data approximating the target multi-source response surfaces;

(c) The kriging model fitted from the Stage-1 data approximating the random error variance as a function of \mathbf{w} .

Step 0: Set $n_i \approx N_2 \frac{\sqrt{V_i C_i}}{\sum_{j=1}^I \sqrt{V_j C_j}}$ for $i = I_1 + 1, I_1 + 2, \dots, I$; solve (4.3) w.r.t. $\{\mathbf{w}_i; i = I_1 + 1, I_1 + 2, \dots, I\}$.

Step 1: Fix $\{\mathbf{w}_i; i = I_1 + 1, I_1 + 2, \dots, I\}$ at its most recently-obtained values, and solve (4.3) w.r.t. $\{n(\mathbf{w}_i); i = I_1 + 1, I_1 + 2, \dots, I\}$.

Step 2: Fix $\{n(\mathbf{w}_i); i = I_1 + 1, I_1 + 2, \dots, I\}$ at its most recently-obtained values, and solve (4.3) w.r.t. $\{\mathbf{w}_i; i = I_1 + 1, I_1 + 2, \dots, I\}$.

Step 3: Repeat Step 1-2 until there is no significant changes in $\{\mathbf{w}_i; i = I_1 + 1, I_1 + 2, \dots, I\}$ or $\{n(\mathbf{w}_i); i = I_1 + 1, I_1 + 2, \dots, I\}$.

Through the iterative procedure, the size of the optimization problems to be solved is reduced by half. In our empirical experience, it sufficed to perform one round of Steps 1-2 to achieve convergence in \mathbf{D}_2 . The global genetic algorithm (GA) in Matlab is employed to solve the optimization problems involved in this procedure.

Chapter 5

Empirical Studies

In this chapter, the SKQ-based two-stage design procedure is statistically evaluated and compared to two other design methods: the traditional design and space-filling design through two empirical case studies.

Case 1: A multi-source dose-time-response case with two quantitative factors and one qualitative factor.

Case 2: A multi-source dose-response case with one quantitative factor and one qualitative factor.

The evaluation of a statistical method such as the design procedure requires the availability of the true target relationships (e.g., the true dose-time-response and dose-response surfaces) as a benchmark and an extremely large amount of validation data, and thus is usually performed based on simulation, as opposed to real experiments. A simulation model includes the true benchmark as part of the model, and can be used to generate data that reflect any important features of real data via computer experiments [34].

5.1 Case 1

This case is derived from the toxicology study of TiO_2 NMs performed in Porter et al. [35]. The experimental condition is defined by two quantitative factors $\mathbf{x} = (x_1, x_2)$, and one qualitative factor z . The TiO_2 dosage administered to an animal is denoted by x_1 , with $x_1 \in [0, 15]\mu\text{g}$; and the post-exposure time is denoted by x_2 , with $x_2 \in [1, 112]$ days. The factor z

has two categories (sources) $\{c_1, c_2\}$, corresponding to the two different shapes of NMs: c_1 for short and c_2 for long TiO₂ nanobelts. Thus, the experimental condition is specified by the factor vector $\mathbf{w} = (\mathbf{x}^\top, z)^\top$, and the response of interest is BAL (bronchoalveolar lavage) PMNs measured in the units of 10³/mouse.

5.1.1 Simulation Model

A simulation model is developed to generate dose-time-response data mimicking the real experimental data in the TiO₂ toxicology study described above.

The simulation model is specified as follows. The true expected dose-time-responses for the two sources (short and long nanobelts) are represented as $\{\mathbf{Y}(\mathbf{x}, c_1), \mathbf{Y}(\mathbf{x}, c_2)\}$, with specific expressions given as Model (5.1-5.2):

$$\begin{aligned}
Y(\mathbf{x}, c_1) = & 67.55 + \frac{0.14}{\exp(-12.85 + 0.41x_1 + 0.06x_2)} + \frac{-33.23}{\exp(-1.39 + 0.12x_1 - 0.02x_2)} \\
& + \frac{-0.08}{\exp(-8.57 + 1.22x_1 - 0.2x_2)} + \frac{-76.00}{\exp(-0.21 + 0.06x_1 - 0.01x_2)} \\
& + \frac{-3.49}{\exp(-0.40 + 0.20x_1 - 0.04x_2)} + \frac{-76.99}{\exp(-0.51 - 0.05x_1 + 0.03x_2)} \\
& + \frac{5.15}{\exp(-5.52 + 0.16x_1 + 0.04x_2)} + \frac{-0.03}{\exp(-25.89 + 2.99x_1 + 0.28x_2)} \\
& + \frac{32.77}{\exp(-0.62 + 0.05x_1 + 0.03x_2)} + \frac{1.09}{\exp(5.76 + 0.26x_1 - 0.07x_2)} \quad (5.1)
\end{aligned}$$

$$\begin{aligned}
Y(\mathbf{x}, c_2) = & 71.74 + \frac{-104.84}{\exp(-0.66 + 0.09x_1 - 0.02x_2)} + \frac{-1.18}{\exp(7.66 - 0.45x_1 - 0.01x_2)} \\
& + \frac{0.04}{\exp(28.01 - 3.62 + 0.24x_2)} + \frac{-4.50}{\exp(5.02 - 0.16x_1 - 0.04x_2)} \\
& + \frac{82.50}{\exp(0.74 + 0.04x_1 - 0.03x_2)} + \frac{-27.73}{\exp(0.83 + 0.13x_1 - 0.02x_2)} \\
& + \frac{35.84}{\exp(-1.18 + 0.04x_1 + 0.03x_2)} + \frac{-13.58}{\exp(-0.85 - 0.18x_1 + 0.03x_2)} \\
& + \frac{-2.61}{\exp(0.32 - 0.27x_1 + 0.01x_2)} + \frac{2.59}{\exp(-8.86 - 0.05x_1 + 0.08x_2)}. \quad (5.2)
\end{aligned}$$

The error variance is dependent on \mathbf{w} and the true variance models are given as:

$$\mathbf{Var}[\varepsilon(\mathbf{x}, c_1)] = (0.33\mathbf{Y}(\mathbf{x}, c_1)^{0.26})^2 \quad (5.3)$$

$$\mathbf{Var}[\varepsilon(\mathbf{x}, c_2)] = (0.77\exp(\mathbf{Y}(\mathbf{x}, c_2) \times 0.0076))^2 \quad (5.4)$$

For an arbitrary subpopulation $c_q (q = 1, 2)$ and at an exposure level \mathbf{x}_0 , a random response y_0 is simulated as

$$y_0 = \mathbf{Y}(\mathbf{x}_0, c_q) + \sqrt{\mathbf{Var}[\varepsilon(\mathbf{x}_0, c_q)]} \cdot \epsilon; \quad q = 1, 2 \quad (5.5)$$

where ϵ is a random error provided by a standard normal random generator[34].

The above simulation models are blind in applying the two-stage procedure. They only serve two purposes in this study. First, they are used to generate dose-time-response data via computer experiments. Second, the true expected response surfaces $\{\mathbf{Y}(\mathbf{x}, c_1), \mathbf{Y}(\mathbf{x}, c_2)\}$, which are part of the simulation model, provide the true benchmark to evaluate the SKQ model fitted from data following a certain design, and hence to compare different design methods in terms of their efficiency.

5.1.2 Applying the Two-Stage Procedure

By using the simulation model (5.1.1) as the sampling approach for data generation, we applied the SKQ-based two-stage procedure to design and model the multi-source experiments for the TiO_2 toxicity study.

The inputs of the procedure (listed at the beginning of Chapter 4) are given as follows for this case: $N = 240$; $N_1 = 120$; $Q = 2$; $\chi = [0, 15] \times [1, 112]$; $I = 24$.

Stage 1: A total of $N_1 = 120$ samples are allocated to the initial stage. The number of replications at each design point is selected to be 10, and thus the number of distinct design points is $I_1 = 12$. The locations of these 12 points are determined by applying SLHD to the two slices of region χ . A realization of the SLHD space-filling design is plotted as

stars in Figure 5.1(a) and (b), with the number next to each star or dot representing the number of replications assigned to that design point. Each of the two sources (short and long nanobelts) has been assigned 6 points.

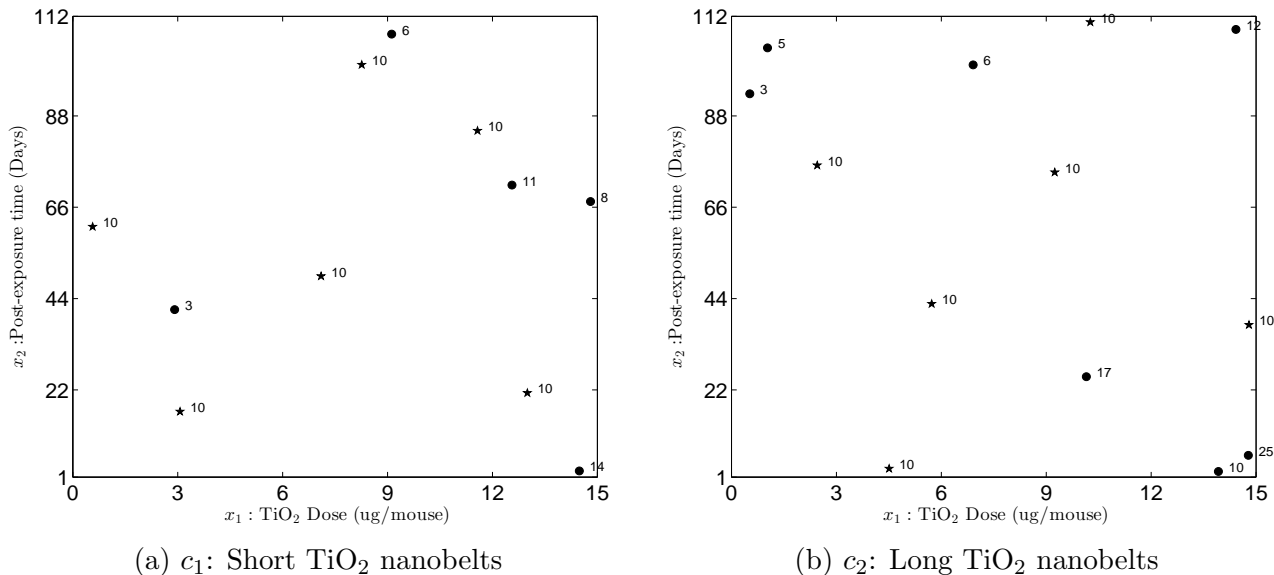


Figure 5.1: The two-stage experimental design: the stars denote the Stage-1 design points and the dots denote the Stage-2 design points.

Thus, the Stage-1 design \mathbf{D}_1 consists of 12 points (stars in Figure 5.1) with 10 replications assigned to each point. At each distinct design point, Model (5.5) is used to generate 10 i.i.d. random responses. The initial data set is denoted as $\{(\mathbf{w}_i, \mathcal{Y}_j(\mathbf{w})); i = 1, 2, \dots, 12; j = 1, 2, \dots, 10\}$, on which normalization is performed so that both the quantitative factors and responses fall within the range of $[0, 1]$. Based on the initial data, two models are fitted: The fitted SKQ model $\widehat{Y}(\mathbf{w})$ approximating the two target dose-time-response surfaces is obtained by applying the SKQ estimation procedure (Chapter 3). The variance model $\widehat{\text{Var}}(\mathbf{w})$ is fitted from the sample variance data $\{(\mathbf{w}_i, \widehat{\text{Var}}[\varepsilon(\mathbf{w})]); i = 1, 2, \dots, 12\}$ via deterministic kriging.

Stage 2: $N_2 = N - N_1 = 120$ samples are to be allocated to $I_2 = I - I_1 = 12$ distinct design points. The location of these 12 design points and the sample allocation are obtained by solving the design optimization problem (4.3) using the iterative procedure in Section 4.2. The inputs needed for the iterative procedure have been obtained from Stage 1: the Stage-1

design \mathbf{D}_1 , the SKQ model $\widehat{Y}(\mathbf{w})$ and the variance model $\widehat{\text{Var}}[\mathbf{w}]$ that are both fitted from the Stage-1 data.

The design points obtained from the design optimization for Stage 2 are plotted as dots in Figure 5.1(a) and (b), with the number next to each dot representing the sample size assigned to that design point. Among the 12 design points, 5 are assigned to the short nanobelts, and 7 to the long nanobelts. The samples are not evenly distributed to the 12 design points: the points with higher error variance tend to have more replications assigned to them.

Following the Stage-2 design, a new batch of experiments are carried out. From the two stages of data, SKQ will be performed to simultaneously model the exposure-response surfaces for both short and long TiO_2 nanobelts.

It is worthy of noting that the outcome (the resulting two-stage design, sample data and fitted SKQ from both stages of data) of applying the two-stage procedure is random: The stars in Figure 5.1 represent one possible SLHD space-filling design in the two slices of region χ ; the Stage-2 design is dependent on both the design and randomly sampled data in Stage 1; the final SKQ model is fitted from the randomly sampled data following the two-stage design.

It is important to point out that the outcome of applying the improved two-stage design above only represents one possible optimal design under the certain design criterion. Due to the random nature of responses, reapplying the procedure will lead to a different stage I design which will change the following stage II design results. Considering the randomness, a total of 200 macro-replications will be run to get an average for comparison results.

5.1.3 Evaluation and Comparison

Due to the stochastic nature of its outcome, the SKQ-based design procedure is statistically evaluated and compared to the other two design methods based on large macro-replications.

The two alternative design methods are described as follows. (i) The traditional design includes 12 evenly-spaced sampling points (Figure 5.2) in the dose-time region for each of the two sources. Ten replications are assigned to each design point, with a total sample size of $12 \times 2 \times 10 = 240$. (ii) The SLHD space-filling design is performed with 24 distinct design points assigned to the two slices of dose-time region, and the total sample size is also 240 with 10 replications at each point. An example of the generated SLHD design is given in Figure 5.3, with the triangles representing the design points for the short nanobelts, and the squares those for the long nanobelts. Clearly, all three design methods have the same sample budget.

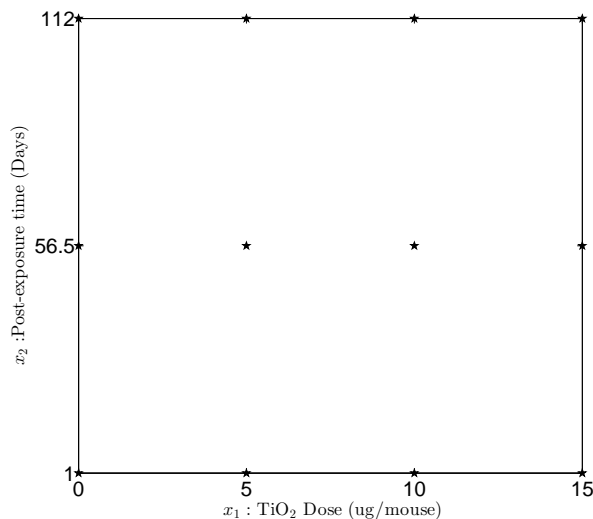


Figure 5.2: The design points in the dose-time region for each source (short or long nanobelts) in the traditional design

For each design method, $M = 200$ macro-replications (that is, independent applications of the method) have been performed as follows.

Step 0: Initialize the index $m = 1$.

Step 1: Apply the design method and follow the design to carry out the experiments; based on the collected data, fit the SKQ model approximating the exposure-response surfaces for both short and long nanobelts; denote the fitted SKQ as $\hat{Y}^{(m)}(\mathbf{w})$.

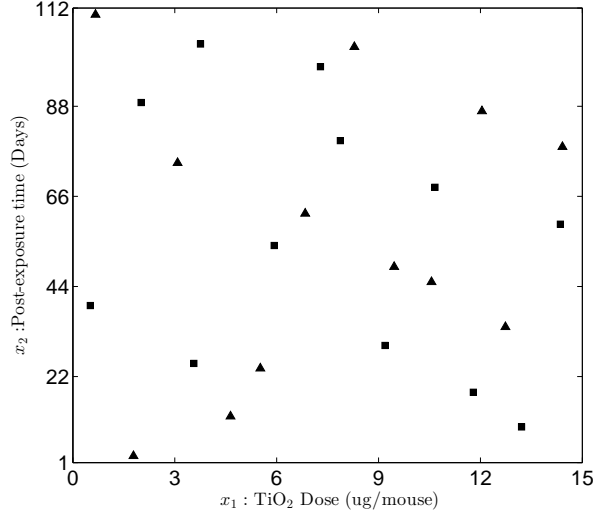


Figure 5.3: The design points of both sources in space-filling design: triangles denote the design points for short nanobelts and squares denote the design points for long nanobelts

Step 2: Evaluate the goodness of $\hat{Y}^{(m)}(\mathbf{w})$ by calculating the estimated root mean squared error (ERMSE)

$$\text{ERMSE}^{(m)}(\mathcal{C}_{c_q}) = \sqrt{\frac{1}{\#[\mathcal{C}_{c_q}]} \sum_{\mathbf{w} \in \mathcal{C}_{c_q}} \left(\hat{Y}^{(m)}(\mathbf{w}) - Y^{(m)}(\mathbf{w}) \right)^2}; \quad q = 1, 2 \quad (5.6)$$

In (5.6), \mathcal{C}_{c_q} represents the collection of 8,456 check points in the dose-time region for short ($q = 1$) or long ($q = 2$) nanobelts. These check points are evenly-spaced dense grids in the dose-time region. $\#[\mathcal{C}_{c_q}]$ denotes the total number of check points in the set \mathcal{C}_{c_q} , and is equal to 8,456 for both q values. ERMSE measures the average deviation of $\hat{Y}(\cdot)$ from its true correspondence $Y(\cdot)$ at the check points in \mathcal{C} . The true value $Y(\cdot)$ is available from the simulation model (5.1.1).

For each design method, $\{\text{ERMSE}^{(m)}(\mathcal{C}_{c_1}); m = 1, 2, \dots, 200\}$ and $\{\text{ERMSE}^{(m)}(\mathcal{C}_{c_2}); m = 1, 2, \dots, 200\}$ are obtained from 200 macro-replications, and used to generate the two box plots for short (c_1) and long (c_2) nanobelts respectively. Figure 5.4 includes the two box plots for each of the three design methods. In comparison, the boxes of our two-stage procedure are the lowest, illustrating its best performance in

terms of ERMSE among the three methods; the boxes are relatively narrow showing that the two-stage procedure leads to consistently low ERMSE. The ERMSE performance of the space-filling design is quite unstable, with the two widest boxes. The ERMSE of the traditional design is consistently high, corresponding to the highest and narrowest boxes in the figure.

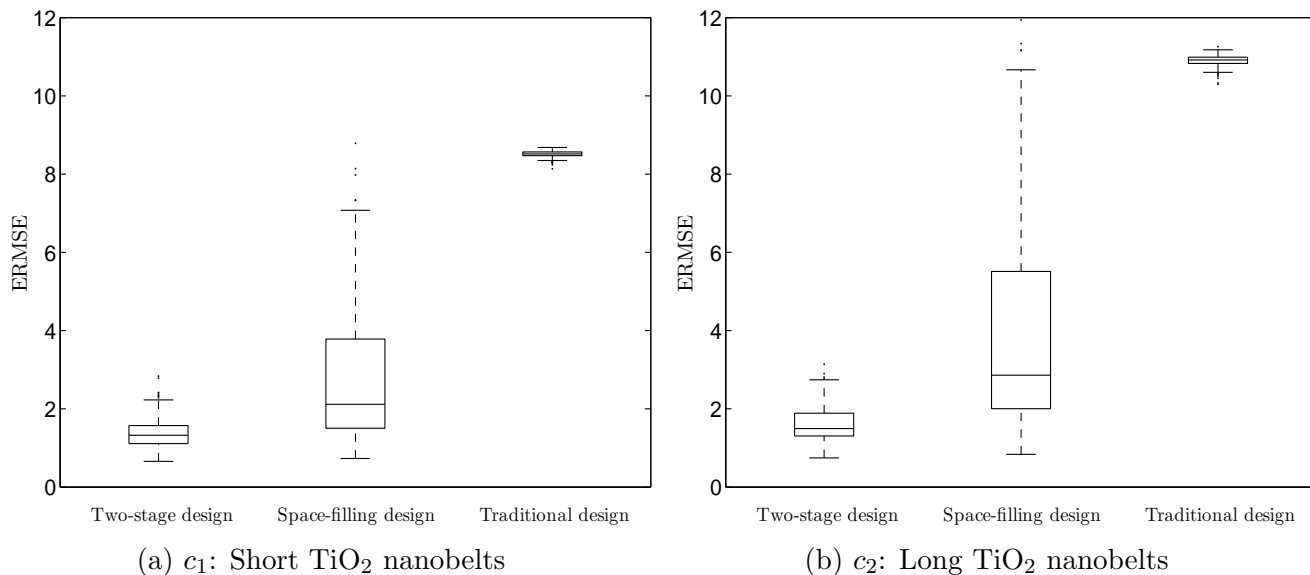


Figure 5.4: The two box plots for each of the three design methods

5.2 Case 2

This case is constructed based on the dose-response study of ZnO engineered nanomaterial(ENM) performed by Tian et al.[36]. There is one quantitative factor x representing the ZnO concentration with $x \in [0, 50]\mu\text{g}/\text{mL}$ and one qualitative factor z with six categories $\{c_1, c_2, c_3, c_4, c_5, c_6\}$. Each category corresponds to a laboratory where the experiments were carried out. The experimental condition is written as $\mathbf{w} = (x, z)$. The response of interest is Percent LDH(Lactate dehydrogenase) release.

5.2.1 Simulation Model

The true expected dose-response models are given in (5.7-5.12).

$$Y(x, c_1) = 17.56 + \frac{64.58}{1 + \exp(0.66 + 4.14x)} \quad (5.7)$$

$$Y(x, c_2) = 67.36 + \frac{-61.16}{1 + \exp(-0.95 - 3.92x)} \quad (5.8)$$

$$Y(x, c_3) = 13.45 \exp(x/28) \quad (5.9)$$

$$Y(x, c_4) = 16.16 \exp(x/34.5) \quad (5.10)$$

$$Y(x, c_5) = 20.59 \exp(x/33) \quad (5.11)$$

$$Y(x, c_6) = 39.40 \exp(x/54.6) \quad (5.12)$$

The error variance is dependent on \mathbf{w} and the true variance models are given as:

$$\mathbf{Var}[\varepsilon(x, c_1)] = \left(\frac{1}{0.15\sqrt{2\pi}} \exp\left(-\frac{(x-37)^2}{450}\right) \right)^2 \quad (5.13)$$

$$\mathbf{Var}[\varepsilon(x, c_2)] = (1.353 \exp(x/52.8))^2 \quad (5.14)$$

$$\mathbf{Var}[\varepsilon(x, c_3)] = \left(\frac{1}{0.06\sqrt{2\pi}} \exp\left(-\frac{(x-37)^2}{450}\right) \right)^2 \quad (5.15)$$

$$\mathbf{Var}[\varepsilon(x, c_4)] = \left(5 + \frac{4.2}{1 + \exp(6 - 0.3x)} \right)^2 \quad (5.16)$$

$$\mathbf{Var}[\varepsilon(x, c_5)] = \left(\frac{1}{0.083\sqrt{2\pi}} \exp\left(-\frac{(x-38)^2}{338}\right) \right)^2 \quad (5.17)$$

$$\mathbf{Var}[\varepsilon(x, c_6)] = \left(\frac{1}{0.076\sqrt{2\pi}} \exp\left(-\frac{(x-17)^2}{128}\right) \right)^2 \quad (5.18)$$

For an arbitrary subpopulation $c_q (q = 1, 2, 3, 4, 5, 6)$ and at an exposure level x_0 , a random response y_0 is simulated as

$$y_0 = \mathbf{Y}(\mathbf{x}_0, c_q) + \sqrt{\mathbf{Var}[\varepsilon(x_0, c_q)]} \cdot \epsilon; \quad q = 1, 2, 3, 4, 5, 6 \quad (5.19)$$

where ϵ is a random error provided by a standard normal random generator[34].

5.2.2 Applying the Two-Stage Procedure

As in Case 1, we use the simulation model 5.2.1 to generate data and then applied the SKQ-based two-stage procedure to design and model the multi-source experiments for the ZnO toxicity study.

The inputs are given as follows: $N = 120$; $N_1 = 60$; $Q = 6$; $\chi = [0, 50]$; $I = 24$.

Stage 1: A total of $N_1 = 60$ samples are allocated in the first stage. The number of replications at each design point is set to be 5, so the number of distinct design points turns out to be $I_1 = 12$ with 2 design points per category. Similarly, SLHD is employed to determine the location of these 12 design points to the six slices of region χ . One possible SLHD space-filling design is shown in Table 5.1.

Table 5.1: SLHD space-filling design in Stage-1 for Case 2

x : dose	c_q :subpopulation	n :replications
26.0028	c_1	5
3.6714	c_1	5
23.1572	c_2	5
42.6297	c_2	5
15.8945	c_3	5
37.5966	c_3	5
36.5581	c_4	5
10.5538	c_4	5
5.6975	c_5	5
46.6739	c_5	5
19.2036	c_6	5
32.1494	c_6	5

At each distinct design point $\{\mathbf{w}_i; i = 1, 2, \dots, 12\}$, 5 i.i.d. random responses are generated by Model (5.19). Thus, the initial data set is denoted as $\{(\mathbf{w}_i, \mathcal{Y}_j(\mathbf{w})); i = 1, 2, \dots, 12; j = 1, 2, \dots, 5\}$, on which normalization is also needed. SK and deterministic kriging (DK) modeling are applied to the initial data set and two Kriging models are obtained: the fitted SKQ model $\hat{Y}(\mathbf{w})$ approximating the six target dose-response surfaces

and the fitted deterministic kriging (DK) model $\widehat{\text{Var}}(\mathbf{w})$ from the sample variance data $\{(\mathbf{w}_i, \widehat{\text{Var}}[\varepsilon(\mathbf{w})]); i = 1, 2, \dots, 12\}$.

Stage 2: The remaining of $N_2 = N - N_1 = 60$ samples are to be allocated to $I_2 = I - I_1 = 12$ distinct design points. Following the design procedure in Section 4.2, the Stage-2 design \mathbf{D}_2 can be determined by solving the design optimization problem (4.3) using the iterative procedure in Section 4.2.

The inputs needed to solve the design optimization problem are as follows. \mathbf{D}_1 : the Stage-1 design; $\widehat{Y}(\mathbf{w})$: the fitted SKQ model estimating the target dose-response curve; $\widehat{\text{Var}}(\mathbf{w})$: the fitted deterministic kriging model approximating the variance model. All of these inputs have been derived from the first stage.

The design points and their corresponding replications obtained in Stage-2 are shown in Table 5.2. For the additional 12 design points, 3 are assigned to the Lab 1, 3 are assigned to the Lab 2, 2 are assigned to the Lab 3, 1 is assigned to Lab 4, 2 are assigned to Lab 5, and 1 is assigned to Lab 6. Also, the replications of these 12 design points are not equal. This result is reasonable because more design points and replications will be needed in the dosage scale with higher error variance.

Thus, the complete design $\mathbf{D} = \mathbf{D}_1 \cup \mathbf{D}_2$ is obtained. Based on \mathbf{D} , a total of 120 pairs of data will be generated by Model 5.19 denoted by $\{(\mathbf{w}_i, \mathcal{Y}_j(\mathbf{w})); i = 1, 2, \dots, 24; j = 1, 2, \dots, n(\mathbf{w}_i)\}$. From these data, SKQ is to be performed to simultaneously model the dose-response curves for all the six Labs.

For the same reasons as Case 1, a total of 200 macro-replications will also be run to get an fair results to compare with the other two design method.

5.2.3 Evaluation and Comparison

Similar to Case 1, the traditional design and space-filling design methods are performed under a same sample budget $N = 120$ as a comparison with the proposed design procedure.

Table 5.2: Stage-2 design for Case 2

x : dose	c_q :subpopulation	n :replications
8.0260	c_1	6
38.4289	c_1	5
19.9689	c_1	2
46.9781	c_2	2
6.3843	c_2	6
23.5860	c_2	9
43.9672	c_3	6
30.5763	c_3	3
40.7347	c_4	5
20.1781	c_5	3
42.5425	c_5	6
43.3273	c_6	7

For the traditional design, 4 equally spaced design points are selected over the dosage range $[x_L, x_U] = [0, 50]$ for each of the six sources. Five replications are assigned to each of the design point such that the total sample size is $4 \times 6 \times 5 = 120$. The traditional design per source is specified as (5.20).

$$\begin{pmatrix} 0 & 16.7 & 33.4 & 50 \\ 5 & 5 & 5 & 5 \end{pmatrix} \quad (5.20)$$

For the space-filling design, SLHD is performed within an equally total sample size of 120. $I = 24$ design points are assigned to the six slices of dosage range with five replications at each point. A possible locations of design points generated by SLHD is given in Table 5.3.

For each design method, $M = 200$ macro-replications are to be performed following the steps in Section 5.1.3. Thus, six box plots(Figure 5.5) for all sources will be generated based on the six groups of ERMSEs: $\{\text{ERMSE}^{(m)}(\mathcal{C}_{c_q}); m = 1, 2, \dots, 200; q = 1, 2, \dots, 6\}$. From the Figure (5.5), it is obvious that our two-stage procedure performs best in terms of

Table 5.3: Design points in the space-filling design for Case 2.

x : dose	2.5	22.1	41.4	27.0	38.8	8.0	12.5	27.8	49.2	32.0	11.2	17.1
c_q : sources	c_1				c_2				c_3			
x : dose	44.1	24.9	9.1	30.9	14.6	0.4	36.3	45.9	43.2	5.7	19.2	34.6
c_q : sources	c_4				c_5				c_6			

ERMSE compared to the other two design methods. First, the boxes of the proposed design procedure are lowest which means our design can get a best estimated curve among the three method. Second, the boxes of ours are relatively narrow showing that the two-stage design method has a stable good performance.

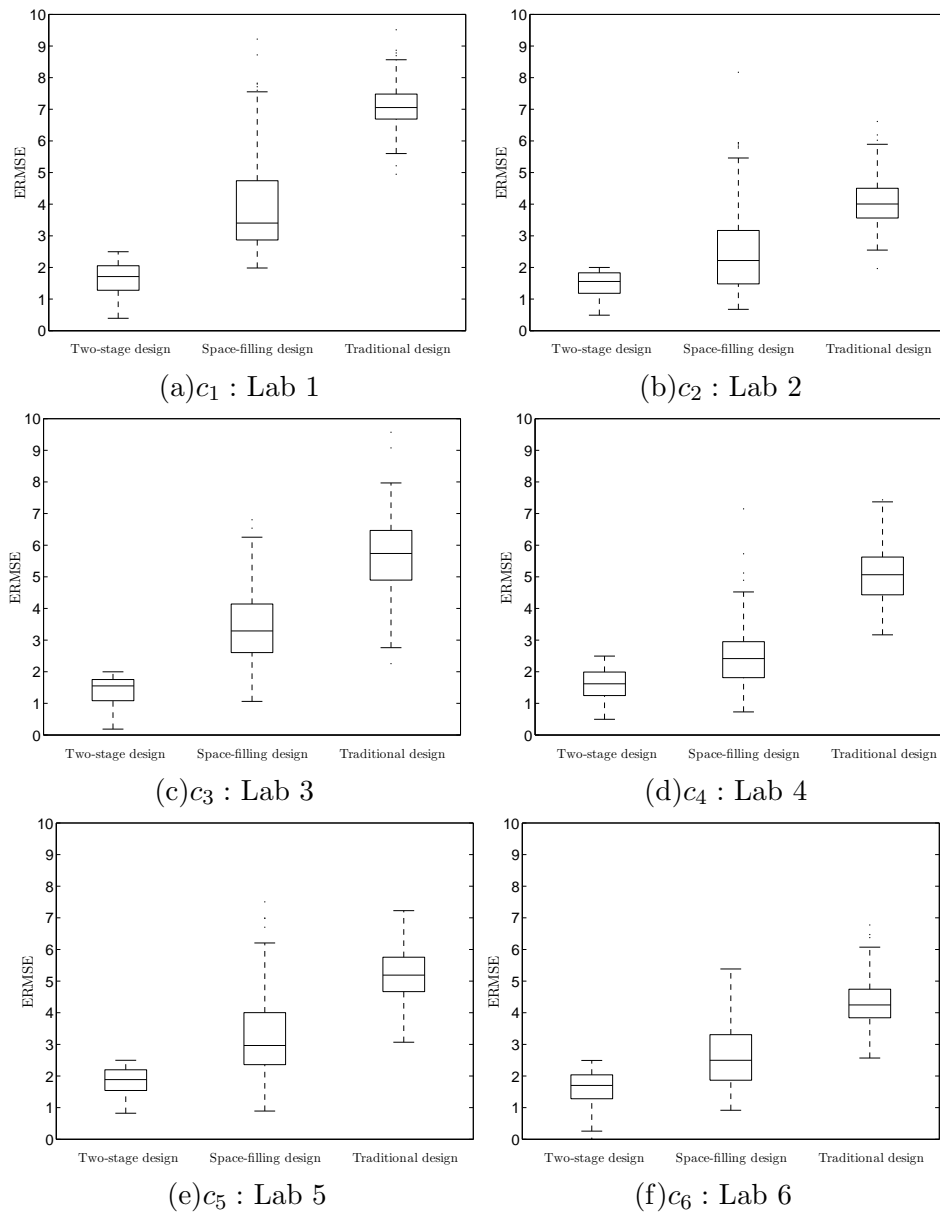


Figure 5.5: The six box plots for each of the three design methods

Chapter 6

Conclusion

A two-stage design of experiments (DOE) procedure was developed to efficiently design multi-source experiments in exposure-response studies of NMs. The design method is closely coupled with stochastic kriging with qualitative factors (SKQ), a statistical model which synergistically models multi-source data by pooling information across sources. Based on SKQ, the DOE procedure seeks to find the experimental design (the sampling location and allocation) that enables SKQ to maximize its information-pooling capability and thus to render fitted relationships (e.g., exposure-response) of the highest quality. The DOE is built in a two-stage framework to allow for the learning of the possibly nonlinear relationships and heterogeneous variance structures, and the learned information is utilized to guide the optimal experimental design.

Through the empirical simulation studies, the efficiency of the proposed SKQ-based two-stage DOE over the two alternative design methods which are widely used in biological experiments has been demonstrated. The DOE method developed in this paper is expected to substantially reduce the cost and time in exposure-response studies for nanotoxicology, and accelerate the progress toward quantifying the risk, safety and health effects of NM exposure.

References

- [1] Mingjia Zhi, Ayyakkannu Manivannan, Fanke Meng, and Nianqiang Wu. Highly conductive electrospun carbon nanofiber/mno 2 coaxial nano-cables for high energy and power density supercapacitors. *Journal of Power Sources*, 208:345–353, 2012.
- [2] Liming Dai, Dong Wook Chang, Jong-Beom Baek, and Wen Lu. Carbon nanomaterials for advanced energy conversion and storage. *Small*, 8(8):1130–1166, 2012.
- [3] Ming Li, Jianming Zhang, Savan Suri, Letha J Sooter, Dongling Ma, and Nianqiang Wu. Detection of adenosine triphosphate with an aptamer biosensor based on surface-enhanced raman scattering. *Analytical chemistry*, 84(6):2837–2842, 2012.
- [4] Na Ren, Rui Li, Limei Chen, Guancong Wang, Duo Liu, Yingjun Wang, Lin Zheng, Wei Tang, Xiaoqiang Yu, Huaidong Jiang, et al. In situ construction of a titanate–silver nanoparticle–titanate sandwich nanostructure on a metallic titanium surface for bacteriostatic and biocompatible implants. *Journal of Materials Chemistry*, 22(36):19151–19160, 2012.
- [5] Michael Giersig and Gennady B Khomutov. *Nanomaterials for application in medicine and biology*. Springer, 2008.
- [6] Ming Li, Qiaoyi Wang, Xiaodong Shi, Lawrence A Hornak, and Nianqiang Wu. Detection of mercury (ii) by quantum dot/dna/gold nanoparticle ensemble based nanosensor via nanometal surface energy transfer. *Analytical chemistry*, 83(18):7061–7065, 2011.
- [7] Nianqiang Wu, Minhua Zhao, Jian-Guo Zheng, Chuanbin Jiang, Ben Myers, Shuoxin Li, Minking Chyu, and Scott X Mao. Porous cuo–zno nanocomposite for sensing electrode of high-temperature co solid-state electrochemical sensor. *Nanotechnology*, 16(12):2878, 2005.
- [8] Nianqiang Wu, Zheng Chen, Jianhui Xu, Minking Chyu, and Scott X Mao. Impedance-metric pt/ysz/au–ga 2 o 3 sensor for co detection at high temperature. *Sensors and Actuators B: Chemical*, 110(1):49–53, 2005.

- [9] Marinella Farré, Krisztina Gajda-Schranz, Lina Kantiani, and Damià Barceló. Ecotoxicity and analysis of nanomaterials in the aquatic environment. *Analytical and Bioanalytical Chemistry*, 393(1):81–95, 2009.
- [10] L Edler, K Poirier, M Dourson, J Kleiner, B Mileson, H Nordmann, A Renwick, W Slob, K Walton, and G Würtzen. Mathematical modelling and quantitative methods. *Food and Chemical Toxicology*, 40(2):283–326, 2002.
- [11] Frank Bretz, Jason Hsu, Jose Pinheiro, and Yi Liu. Dose finding—a challenge in statistics. *Biometrical Journal*, 50(4):480–504, 2008.
- [12] US EPA. *Benchmark dose technical guidance*. Washington, DC: U.S. Environmental Protection Agency, 2012.
- [13] Kai Wang, Xi Chen, Feng Yang, Dale W Porter, and Nianqiang Wu. A new stochastic kriging method for modeling multi-source exposure–response data in toxicology studies. *ACS sustainable chemistry & engineering*, 2(7):1581–1591, 2014.
- [14] Peter ZG Qian, Boxin Tang, and CF Jeff Wu. Nested space-filling designs for computer experiments with two levels of accuracy. *Statistica Sinica*, 19(1):287, 2009.
- [15] Peter ZG Qian, Mingyao Ai, and CF Jeff Wu. Construction of nested space-filling designs. *The Annals of Statistics*, pages 3616–3643, 2009.
- [16] R.F. Hamilton, N.Q. Wu, D. Porter, M. Buford, M. Wolfarth, and A. Holian. Particle length-dependent titanium dioxide nanomaterials toxicity and bioactivity. *Particle and Fibre Toxicology*, 6(35), 2009.
- [17] D.W. Porter, A.F. Hubbs, R.R. Mercer, N.Q. Wu, M.G. Wolfarth, K. Sriram, S. Leonard, L. Battelli, D. Schwegler-Berry, S. Friend, M. Andrew, B.T. Chen, S. Tsuruoka, M. Endo, and V. Castranova. Mouse pulmonary dose- and time course-responses induced by exposure to multi-walled carbon nanotubes. *Toxicology*, 269(2-3):136–147, 2010.
- [18] Peter ZG Qian. Sliced latin hypercube designs. *Journal of the American Statistical Association*, 107(497):393–399, 2012.
- [19] Peter ZG Qian and CF Jeff Wu. Sliced space-filling designs. *Biometrika*, 96(4):945–956, 2009.
- [20] Jerome Sacks, William J Welch, Toby J Mitchell, and Henry P Wynn. Design and analysis of computer experiments. *Statistical science*, pages 409–423, 1989.

- [21] Jay D Martin and Timothy W Simpson. On the use of kriging models to approximate deterministic computer models. In *ASME 2004 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, pages 481–492. American Society of Mechanical Engineers, 2004.
- [22] Timothy W Simpson, Jesse Peplinski, Patrick N Koch, and Janet K Allen. On the use of statistics in design and the implications for deterministic computer experiments. *Design Theory and Methodology-DTM'97*, pages 14–17, 1997.
- [23] Jack PC Kleijnen and Wim CM Van Beers. Application-driven sequential designs for simulation experiments: Kriging metamodelling. *Journal of the Operational Research Society*, 55(8):876–883, 2004.
- [24] Deng Huang, TT Allen, WI Notz, and RA Miller. Sequential kriging optimization using multiple-fidelity evaluations. *Structural and Multidisciplinary Optimization*, 32(5):369–382, 2006.
- [25] Bruce Ankenman, Barry L Nelson, and Jeremy Staum. Stochastic kriging for simulation metamodeling. *Operations research*, 58(2):371–382, 2010.
- [26] Xi Chen and Qiang Zhou. Sequential experimental designs for stochastic kriging. In *Proceedings of the 2014 Winter Simulation Conference*, pages 3821–3832. IEEE Press, 2014.
- [27] Thomas J. Santner, Brian J. Williams, and William Notz. *The design and analysis of computer experiments*. Springer, 2003.
- [28] Peter ZG Qian, Huaqing Wu, and CF Jeff Wu. Gaussian process models for computer experiments with qualitative and quantitative factors. *Technometrics*, 50(3), 2008.
- [29] Carl Edward Rasmussen and K. I. Williams Christopher. *Gaussian processes for machine learning*. The MIT Press, 2006.
- [30] Qiang Zhou, Peter ZG Qian, and Shiyu Zhou. A simple approach to emulation for computer models with qualitative and quantitative factors. *Technometrics*, 53(3), 2011.
- [31] S. Sand, D. von Rosen, P. Eriksson, A. Fredriksson, H. Viberg, K. Victorin, and A.F. Filipsson. Dose-response modeling and benchmark calculations from spontaneous behavior data on mice neonatally exposed to 2,2',4,4',5-pentabromodiphenyl ether. *Toxicological Sciences*, 81:491–501, 2004.

- [32] Donald R Jones, Matthias Schonlau, and William J Welch. Efficient global optimization of expensive black-box functions. *Journal of Global optimization*, 13(4):455–492, 1998.
- [33] Jay D Martin and Timothy W Simpson. A study on the use of kriging models to approximate deterministic computer models. In *ASME 2003 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, pages 567–576. American Society of Mechanical Engineers, 2003.
- [34] Averill M Law, W David Kelton, and W David Kelton. *Simulation modeling and analysis*, volume 2. McGraw-Hill New York, 1991.
- [35] Dale W Porter, Nianqiang Wu, Ann Hubbs, Robert Mercer, Kathleen Funk, Fanke Meng, Jiangtian Li, Michael Wolfarth, Lori Battelli, Sherri Friend, et al. Differential mouse pulmonary dose-and time course-responses to titanium dioxide nanospheres and nanobelts. *Toxicological Sciences*, page kfs261, 2012.
- [36] Tian Xia, Raymond F Hamilton, James C Bonner, Edward D Crandall, Alison Elder, Farnoosh Fazlollahi, Teri A Girtsman, Kwang Kim, Somenath Mitra, Susana A Ntim, et al. Interlaboratory evaluation of in vitro cytotoxicity and inflammatory responses to engineered nanomaterials: the niehs nano go consortium. *Environmental health perspectives*, 121(6):683–690, 2013.

Appendix A

Derivation of equation (4.6)

We shall show how to derive the approximation formula between optimal allocation and the locations of design points. First, we need to note that

$$\begin{aligned}
 \widehat{\text{MSE}}[\widehat{Y}(\mathbf{w}_0)] &= \Sigma_{\mathbf{M}}(\mathbf{w}_0, \mathbf{w}_0) - \Sigma_{\mathbf{M}}(\mathbf{w}_0, \cdot)(\Sigma_{\mathbf{M}} + \Sigma_{\varepsilon})^{-1}\Sigma_{\mathbf{M}}(\mathbf{w}_0, \cdot) \\
 &= \sigma^2 - \sigma^4 \cdot \mathbf{v}(\mathbf{w}_0, \boldsymbol{\theta}, \Phi)^\top (\Sigma_{\mathbf{M}} + \Sigma_{\varepsilon})^{-1} \mathbf{v}(\mathbf{w}_0, \boldsymbol{\theta}, \Phi) \\
 &= \sigma^2 - \sigma^4 \cdot \sum_{i,j=1}^I \{[(\Sigma_{\mathbf{M}} + \Sigma_{\varepsilon})^{-1}] \cdot r_{0i}r_{0j}\} \tag{A.1}
 \end{aligned}$$

where, σ^2 is the variance of the Gaussian process. $\Sigma_{\mathbf{M}}$ is the $I \times I$ variance-covariance matrix,

$$\Sigma_{\mathbf{M}} = \sigma^2 \mathbf{R}(\boldsymbol{\theta}, \Phi) = \sigma^2 \begin{pmatrix} 1 & \text{Corr}[\mathbf{M}(\mathbf{w}_1), \mathbf{M}(\mathbf{w}_2)] & \dots & \text{Corr}[\mathbf{M}(\mathbf{w}_1), \mathbf{M}(\mathbf{w}_I)] \\ \text{Corr}[\mathbf{M}(\mathbf{w}_2), \mathbf{M}(\mathbf{w}_1)] & 1 & \dots & \text{Corr}[\mathbf{M}(\mathbf{w}_2), \mathbf{M}(\mathbf{w}_I)] \\ \vdots & \vdots & \ddots & \vdots \\ \text{Corr}[\mathbf{M}(\mathbf{w}_I), \mathbf{M}(\mathbf{w}_1)] & \text{Corr}[\mathbf{M}(\mathbf{w}_I), \mathbf{M}(\mathbf{w}_2)] & \dots & 1 \end{pmatrix}. \tag{A.2}$$

$\mathbf{v}(\mathbf{w}_0, \boldsymbol{\theta}, \Phi)$ denotes the correlation vector with each component being a correlation function dependent on \mathbf{w}_0 , and the unknown parameters $\boldsymbol{\theta}$ and Φ ,

$$\mathbf{v}(\mathbf{w}_0, \boldsymbol{\theta}, \Phi) = \begin{pmatrix} \text{Corr}[\mathbf{M}(\mathbf{w}_0), \mathbf{M}(\mathbf{w}_1)] \\ \text{Corr}[\mathbf{M}(\mathbf{w}_0), \mathbf{M}(\mathbf{w}_2)] \\ \vdots \\ \text{Corr}[\mathbf{M}(\mathbf{w}_0), \mathbf{M}(\mathbf{w}_I)] \end{pmatrix}, \tag{A.3}$$

Σ_ε is the $I \times I$ variance-covariance matrix of vector $\bar{\varepsilon}$. Under the i.i.d. assumption of random errors, Σ_ε is a $I \times I$ diagonal matrix

$$\widehat{\Sigma}_\varepsilon = \text{diag}\{\widehat{\text{Var}}[\varepsilon(\mathbf{w}_1)]/n(\mathbf{w}_1), \widehat{\text{Var}}[\varepsilon(\mathbf{w}_2)]/n(\mathbf{w}_2), \dots, \widehat{\text{Var}}[\varepsilon(\mathbf{w}_I)]/n(\mathbf{w}_I)\}, \quad (\text{A.4})$$

r_{0i} is the correlation between \mathbf{w}_0 and \mathbf{w}_i , for $i = 1, 2, \dots, I$.

$$r_{0i} = \text{Corr}[\mathbf{M}(\mathbf{w}_0), \mathbf{M}(\mathbf{w}_i)] \quad (\text{A.5})$$

Next step, IMSE is computed across the whole design space of interest \mathbf{H} .

$$\begin{aligned} \text{IMSE} = \\ \sigma^2 - \sigma^4 \cdot \sum_{i,j=1}^I \{[(\Sigma_M + \Sigma_\varepsilon)^{-1}] \cdot \sum_{c_q} \int_{\mathbf{x}_0 \in \mathcal{X}} r_{0i} r_{0j} d\mathbf{x}_0\} \end{aligned} \quad (\text{A.6})$$

Define $\mathbf{W}_{I \times I}$ as the matrix with elements $\mathbf{W}_{ij} = \sum_{c_q} \int_{\mathbf{x}_0 \in \mathcal{X}} r_{0i} r_{0j} d\mathbf{x}_0$ and $(\Sigma_M + \Sigma_\varepsilon)^{-1} = \Sigma^{-1}$, then

$$\text{IMSE} = \sigma^2 - \sigma^4 \cdot \mathbf{1}_I^\top [\Sigma^{-1} \circ \mathbf{W}_{I \times I}] \mathbf{1}_I \quad (\text{A.7})$$

Taking the derivative w.r.t. n_i for both sides of (A.7), we get

$$\frac{\partial \text{IMSE}}{\partial n_i} = -\sigma^4 \cdot \mathbf{1}_I^\top \left[\frac{\partial \Sigma^{-1}}{\partial n_i} \circ \mathbf{W}_{I \times I} \right] \mathbf{1}_I \quad (\text{A.8})$$

Then, the problem reduces to computing the $\frac{\partial \Sigma^{-1}}{\partial n_i}$. According to the matrix derivative operation,

$$\begin{aligned} \frac{\partial \Sigma^{-1}}{\partial n_i} &= -\Sigma^{-1} \frac{\partial \Sigma}{\partial n_i} \Sigma^{-1} \\ &= -\Sigma^{-1} \left(-\frac{\text{Var}(\mathbf{w}_i)}{n_i^2} \mathbf{L}^{ii} \right) \Sigma^{-1} \end{aligned} \quad (\text{A.9})$$

where, \mathbf{L}^{ii} is $I \times I$ matrix with 1 in position (i,i) and 0 in other places.

Plugging (A.9) into (A.8) , then (A.8) can be re-written as:

$$\begin{aligned}
\frac{\partial \text{IMSE}}{\partial n_i} &= -\sigma^4 \cdot \mathbf{1}_I^\top \left\{ [-\Sigma^{-1} \left(-\frac{\text{Var}(\varepsilon(\mathbf{w}_i))}{n_i^2} \mathbf{L}^{ii} \right) \Sigma^{-1}] \circ \mathbf{W}_{I \times I} \right\} \mathbf{1}_I \\
&= -\sigma^4 \cdot \frac{\text{Var}(\varepsilon(\mathbf{w}_i))}{n_i^2} \cdot \mathbf{1}_I^\top [(\Sigma^{-1} \mathbf{L}^{ii} \Sigma^{-1}) \circ \mathbf{W}_{I \times I}] \mathbf{1}_I \\
&= -\sigma^4 \cdot \frac{\text{Var}(\varepsilon(\mathbf{w}_i))}{n_i^2} [\Sigma^{-1} \mathbf{W}_{I \times I} \Sigma^{-1}]_{ii}
\end{aligned} \tag{A.10}$$

To be simple, define $C_i = [\Sigma^{-1} \mathbf{W}_{I \times I} \Sigma^{-1}]_{ii}$ and $\text{Var}(\varepsilon(\mathbf{w}_i)) = V_i$. When N is large, $\Sigma^{-1} = (\Sigma_M + \Sigma_M)^{-1} \approx \Sigma_M^{-1}$, hence, $C_i \approx \Sigma_M^{-1} \mathbf{W}_{I \times I} \Sigma_M^{-1}$. Finally, (A.10) can be rewritten as the following:

$$\frac{\partial \text{IMSE}}{\partial n_i} = -\sigma^4 \frac{V_i}{n_i^2} C_i \tag{A.11}$$

Therefore, $n_i \propto \sqrt{V_i C_i}$.