QUALITY ENGINEERING
https://doi.org/10.1080/08982112.2017.1320726

Taylor & Francis
Taylor & Francis Group

Design and analysis of computer experiment via dimensional analysis

Weijie Shen[a], Dennis K. J. Lin[a], and Chia-Jung Chang[b]

[a]Department of Statistics, The Pennsylvania State University, State College, Pennsylvania; [b]Department of Industrial and Manufacturing Engineering, The Pennsylvania State University, State College, Pennsylvania

ABSTRACT
Design and analysis techniques for computer experiments have been largely developed. However, empirical emulators generated based on experimental data usually fail to incorporate physical principles and dimensional constraints the computer model follows. In this article, we propose a new design and analysis framework based on dimensional analysis (DA), a widely used reduction technique in physics and engineering. We show that implementing DA in computer experiments is efficient, scalable, interpretable, robust and costless. We demonstrate the benefits and improvement in details by theoretical derivations and numerical examples of the borehole model and the damped harmonic oscillation.

KEYWORDS
design of experiments; dimensional constraints; factor space; Kriging; transformation

Introduction

The significance of computer experiments in scientific and technological development has been recognized and well studied in recent years. The advancement of computing power has made simulations of complicated systems feasible. Due to the substantial savings and other practical concerns, it is common to perform studies using computer simulations before physical experiments in modern research. Because computer simulations can be time-consuming, computer experiments are conducted to produce surrogate models. The problem of how to design and analyze computer experiments efficiently has thus been put forward and studied in great detail: see, for example, Santner et al. (2003). Space-filling designs such as Latin hypercube designs (LHDs) are chosen instead of factorial designs because the redundancy of design points in projective subspaces provides little information in computer experiments. Gaussian process models (such as Kriging) are often implemented as surrogate because of the deterministic feature of computer experiments and the need for interpolation.

However, the empirical emulators generated through experimental data usually do not follow the fundamental physical principles the computer experiment follows. For example, in the computer simulated aerodynamic tests of plane wings, the emulated lifting force by empirical emulators may not be a monotone function of the wind speed as it is in computer experiments. Even though the derived emulator fits the data well, it may generate predictions that do not satisfy the physical constraints when interpolating and extrapolating. This is a problematic issue in the following ways. First, such predictions may not be interpretable for the scientific purpose. Second, such lack-of-fit becomes unnecessarily severe when the response surface is complex and not smooth, while the complexity can be explained by physical knowledge. Third, even assuming very dense design points, an emulator that fits the data perfectly and a relatively smooth response surface, the emulator will possibly generate poor predictions when extrapolating into different scales (especially for Kriging), while the scaling is governed by the physical laws as in the accelerated life testing. Finally, the efficiency of the modeling deteriorates without utilizing the physical information that helps guide and restrict the directions of modeling. In this article, we consider utilizing the dimensional information through dimensional analysis.

In physics and engineering applications, it has been a common practice to conduct dimensional analysis (DA), which was first introduced by Buckingham

CONTACT Weijie Shen jayshenwei@gmail.com Department of Statistics, 327 Thomas Building, University Park, PA 16802-2111.
Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/lqen.

(1914) and Rayleigh (1915). For example, Reynolds number has been used a lot in fluid mechanics to predict flow patterns, laminar or turbulent, in different flow situations. Buckingham (1914) formulates the theoretical justification of dimensional analysis procedure by introducing the dimensional homogeneity principle of any physical equation and the Buckingham's Π-theorem. The methodology covered the most general forms of physical equations, leading to its great utility and efficacy in scientific problems. DA received more attentions in statistics in recent researches. Albrecht et al. (2013) discussed the general features of applying dimensional analysis in designs of experiments. Shen et al. (2014) proposed the procedure to incorporate dimensional analysis in statistical analysis and discussed the resulting benefits. The aforementioned works demonstrated both the power and the ease of the adjustment by DA in statistical problems. See also Albrecht et al. (2013), Davis (2013), Lin and Shen (2013), Frey (2013), Jones (2013), Piepel (2013), Plumlee et al. (2013), and the rejoinder Albrecht et al. (2013) for a comprehensive discussion.

Variables in many computer experiments have physical meanings according to the background scenario and obey the dimensional constraints. However, few published works have taken it as a critical component in the statistical meta-modeling. Methodologies are developed either from pure physical derivations, or from pure statistical concerns of uncertainty. Dimensional validation is often ad hoc. In this article, DA is used to determine the factor space on which statistical meta-models should be built. The fact that the generated models satisfy the dimensional constraints greatly increases their performance in extrapolation. Due to the reduction of number of variables, more consistent and efficient models can be produced. Fewer runs are needed for a predetermined tolerance of errors, which relieves the computational burden especially when the computer simulations are expensive. Our contribution is to propose and validate a systematic paradigm to incorporate DA into computer experiments, and to show the advantages of DA factor space using both theoretical and numerical studies. Compared with previous works on physical experiments and optimal designs, this article focuses more on space-filling designs and prediction performance. With comparisons to procedures without using DA, we demonstrate that DA produces great improvement for extrapolation on complex response surfaces. We conclude that incorporating the DA factor space is a general and promising framework for building statistical emulators of computer experiments that obey dimensional constraints.

The remaining paper is organized as follows. In the Methodology section, the methodology of DA for computer experiments is proposed, including introduction of dimensional analysis based factor space (DA factor space) and a Conditional LHD strategy to accommodate irregular design spaces. Furthermore, some theoretical properties are derived in general. In the Case Studies section, we illustrate the procedure and study the impact of DA via two popular examples of computer experiment: the borehole model and damped harmonic oscillation. In the Conclusion section, we draw conclusions about the significant advantages and improvements of incorporating DA, with a discussion of limitations.

Methodology

We first introduce the general setting of the prediction problem in computer experiments. Then we propose a systematic framework of how to apply DA in computer experiments. After formulating DA factor space and the Conditional LHD, some theoretical properties from both physical and statistical perspectives are discussed.

Settings of computer experiment

The problem concerned here is the prediction of a function $y(t)$ on a domain T, where $T \subseteq \mathbb{R}^k$ and $y(t) \in \mathbb{R}^1, \forall t \in T$. Our main goal is to estimate function $y(t)$ based on the dataset $\Xi = \{(t_i, y(t_i)) : t_i \in T, i = 1, \ldots, n\}$ of size n. In the context of computer experiments, t is the k-dimensional input vector to a computer model of an unknown physical or behavioral system and $y(t)$ is a scalar response. In the design phase, we assign values of t_i's to optimize the design $D_T = \{t_i : t_i \in T, i = 1, \ldots, n\}$ in some sense. After the design of D_T, $t_i \in D_T$ is fixed and $y(t_i)$ are obtained from the computer model. In the analysis phase, the corresponding responses $y_i = y(t_i)$ are used to estimate the function and to predict $y(t)$ at any given $t \in T$. Several features are commonly shared by computer models (see also Santner et al., 2003).

1. Function $y(t)$ is mostly deterministic: if the computer experiment is run with the same values of t, the same values of y will be generated. Exceptions are usually due to practical reasons such as follows. In some cases, factors with

small and/or high-order effects are excluded from the model. Also, some factors may imitate random variation due to noise in the environment. When considering the subset of modeling variables of interest, such excluded factors are considered as random effects and contribute to the unreproducible random errors upon the lack of fit of the model. In this sense, $y(t)$ is considered random.

2. Computer models are numerical (not analytical) approaches, thus they are typically computationally intense and time consuming. The feasible number of runs n is limited.
3. The functional dependence of y on t is usually complex, and thus simple interpolation is not sufficient.

Resulting from the above features, a space filling design D_T on the original space T seems appropriate. Adjusting the factor space to cope with the domain knowledge should also be part of statistical modeling. Naively using original factor space T is possibly unjustifiable and inefficient. Here, we propose to use the DA factor space Q.

Formulation of dimensional analysis-based factor space

The dimensional knowledge behind computer experiments often implies clues of functional dependence between y_i and t_i. Dimensional analysis suggests that the better way to associate y_i with t_i is through dimensionless variables. The details are as follows.

Suppose the input $t = (x_1, \ldots, x_k)$, where x_j are the observed physical quantities with unit U_j, $j = 1, \ldots, k$. The physical response we intend to model is y with unit U_0. Among x_j's, select b *basis quantities*, and denote them as $x_1, \ldots, x_b$ with corresponding units $U_1, \ldots, U_b$, such that they have the following two properties: (a) representativity: they can express all other units $U_0, U_{b+1}, \ldots, U_k$, i.e., $U_j = f_j(U_1, \ldots, U_b)$, $j = 0, b+1, \ldots, k$; (b) independence: they cannot represent each other, $\nexists f_j$ such that $U_j = f_j(U_1, \ldots, U_{i\neq j}, \ldots, U_b)$, $j = 1, \ldots, b$. For example, suppose y has dimension length and unit meter, x_1 has dimension length and unit meter, x_2 has dimension time and unit second, x_3 has dimension speed and unit meter/second. Then we can choose x_1 and x_2 to be basis quantities because their units (meter and second) have (a) representativity: can express y's unit meter and x_3's unit meter/second; (b) independence: meter and second cannot represent each other, $\nexists f$ such that $meter = f(second)$. Notice the set of basis quantities cannot include x_3 in addition to $\{x_1, x_2\}$ because that violates independence. It cannot be just $\{x_1\}$ because that violates representativity. But it can be $\{x_1, x_3\}$ or $\{x_2, x_3\}$. The set is not unique but the size b is.

Due to representativity, non-basis quantities can be transformed into dimensionless forms using basis quantities. To accomplish this transformation, let $q_0 = y/f_0(x_1, \ldots, x_b)$; $q_j = x_j/f_j(x_1, \ldots, x_b)$, $j = b+1, \ldots, k$, where f_j's are shown in property (a). Then q_j's are dimensionless and all have unit 1. Note that f_j's are known by physics. Due to property (b) independence, the representations f_j's are unique. Buckingham's Π theorem shows that, rather than original functional dependence $y = y(t) = g(x_1, \ldots, x_k)$, or equivalently

$$q_0 = h(x_1, \ldots, x_b, q_{b+1}, \ldots, q_k), \qquad [1]$$

it should be

$$q_0 = h(q_{b+1}, \ldots, q_k), \qquad [2]$$

to satisfy the dimensional homogeneity, mainly because $x_1, \ldots, x_b$ are basis quantities and cannot represent the dimensionless unit. The *DA factor space* is denoted as $Q = \{(q_{b+1}, \ldots, q_k) : q_j = x_j/f_j(x_1, \ldots, x_b),\ j = b+1, \ldots, k;\ (x_1, \ldots, x_k) \in T\}$ which is a subset of $\mathbb{R}^{k-b}$. The response y can then be recovered by

$$y = f_0(x_1, \ldots, x_b)h(x_{b+1}/f_{b+1}(x_1, \ldots, x_b), \ldots, x_k/f_k \times (x_1, \ldots, x_b)). \qquad [3]$$

We call a space filling design on Q, a *dimensional analysis based design (DA design)*, and denote it by D_Q. Due to the reduced dimension, points in D_Q map to multiple correspondences in T. h is the only unknown function. A model is called a *dimensional analysis based model (DA model)*, if it is able to estimate h in [2] and to recover y through [3].

Here, we briefly illustrate the DA technique using the example of computer simulated aerodynamic tests of plane wings mentioned in the introduction. Suppose the input variables are wind attributes, speed v, density ρ, temperature T and plane wings geometries (of length), $L_1, \ldots, L_s$. The output variable is the lifting force F. According to SI unit system, v has m/s, ρ has kg/m^3, T has K, $L_1, \ldots, L_s$ have m and F has $kg \cdot m/s^2$ as their units, respectively. Without DA, the model is

$F = g(v, \rho, T, L_1, \ldots, L_s)$. With DA, basis quantities are selected to be v, ρ, T, L_1. Other quantities are transformed into dimensionless forms: $q_F = F/(\rho v^2 L_1^2)$ the dimensionless force which describes how F is related to the environment of the wind and $q_{L_j} = L_j/L_1$ for $j = 2, \ldots, s$ which characterizes the geometry of wings. Following the same notation as above, the model is given by $q_F = h(v, \rho, T, L_1, q_{L_2}, \ldots, q_{L_s})$. By dimensional homogeneity requirement, $h(v, \rho, T, L_1, q_{L_2}, \ldots, q_{L_s})$ should be dimensionless. However, v, ρ, T, L_1 are basis quantities that cannot represent dimensionless quantities, and thus should be removed. $q_F = h(q_{L_2}, \ldots, q_{L_s})$, that is $F/(\rho v^2 L_1^2) = h(L_2/L_1, \ldots, L_s/L_1)$, and T is dropped because no other quantities have temperature dimension. The DA model is thus $F = \rho v^2 L_1^2 h(L_2/L_1, \ldots, L_s/L_1)$. Therefore, it turns out that $F \propto v^2$ the lifting force should be a monotone (actually quadratic) function of the wind speed, just as required. In fact, the actual approximate force of lift equation is $F = \rho v^2 S C_L/2$, where $S = L_1 \times L_2$ is the wing area and C_L is the lift coefficient (Anderson, 2004). The lift coefficient C_L that is universal to various conditions is essentially $2q_F/q_{L_2}$. The function h we want to estimate is simply $h(x_2, \ldots, x_s) = x_2/2$. As a result, the DA model utilizes physical information to reveal the relationship between F and ρ, v, and focuses on estimating the effects from wings geometries, which is the benefit that the model without DA, i.e., $F = g(v, \rho, T, L_1, \ldots, L_s)$, cannot achieve.

From DA procedure, it can be seen that the choice of basis quantities is not unique, just as the basis of a linear space. As long as the functional relationship is dimensionally homogeneous, DA is valid and such choice should not affect the results. However, different choice of basis quantities yields different functional forms to be estimated that are not equally attractive in statistical perspective. Intuitively, the basis quantities that yield a simpler functional relationship are more desirable, such as a linear one. The statistical implications of different choice of basis quantities is out of the scope of this article.

Framework of design and analysis for computer experiment via dimensional analysis

By incorporating the DA technique, a design and analysis framework tailored to the dimensional constraints of computer experiment is proposed, which proceeds as follows.

Step 1. Examine the physical dimensions of the variables in the system.

Step 2. Apply dimensional analysis procedures to both input variables and response. The variables are hence transformed into dimensionless. The DA factor space is achieved, as a result.

Step 3. Based on the transformed variables, reformulate the prediction problems. Construct the design, such as Latin Hypercube Design, on the transformed variables. The DA design is achieved, as a result.

Step 4. Generate the responses of the design points via computer experiments.

Step 5. Implement statistical models such as regression and Kriging to emulate the predictions. The DA model is achieved, as a result.

Step 6. Compare the resulting estimates and/or variances to the computer model through confirmation runs and evaluate the prediction performance. Return the estimated function as the final meta-model to the computer model.

Step 1 usually requires domain knowledge of the physical aspect of the system. Details on Step 2 can be found in the previous section. Standard design approaches can be used in Step 3. However, here we propose a new class of designs (in the next section) to comply with features of DA. Computer experiments generate response values in Step 4. Analyses can be carried out according to the context in Step 5. Step 6 evaluate its validation, and finally the prediction problem is resolved. Two case studies will be discussed in detail (in the Case Studies section) to demonstrate the procedure.

Conditional Latin Hypercube Design

Suppose that the relationship between inputs and output is completely unknown. A space filling design on those input variables is recommended. The most popular one is perhaps the Latin Hypercube Design (LHD), which enjoys many desirable properties (McKay et al., 1979; Park, 1994). Generally, the design points for original variables on T can be created straightforwardly by scaling LHD to fit the ranges when T is a hypercube. However, even if T is hypercube, the domain of the DA factor space Q is often irregular

(see Figure 3a in Lin and Shen, 2013, for an example with detailed discussion there). One way is to use the rejection method (RLHD): LHD points are sampled from a hybercube $[0, 1]^{k-b}$ containing the transformed domain Q and the LHD points are kept only if they fall in the domain Q. However, Q is sometimes so irregular that the efficiency for retraining design points is low. There are a few methods for generating a space filling design for an arbitrary irregular space such as Fast Flexible Space-Filling designs (Lekivetz and Jones, 2015) and optimal U-type designs (Lin et al., 2010). Such approaches require an explicit domain specification of Q (cumbersome to derive from T) and often need the convexity assumption to yield good results (usually not applicable for Q). They also need more computation and more expensive to streamline for computer experiments. Furthermore, given the design on Q, the corresponding working inputs t in T are not unique. Take the lifting force example in the Methodology section. Given a design of $q_{L_2} = L_2/L_1$, multiple values of L_1 and L_2 are valid. Plus, eventually the design need to be carried out in terms of the working density ρ and velocity v, which are not defined in Q. Finding the specific values of working inputs t's that both induce the given design on Q and satisfy the domain constraints in T is an additional effort. Finally, although in theory, choice of basis quantities and their values given design Q should not affect the result, but in reality $y(t)$ is seldom invariant to the particular decomposition of q onto t. We need a design that helps us justify this invariance property before we blindly trust DA and think all decompositions work the same.

To generate DA designs on irregular domains in an efficient manner, we propose a "Conditional Latin Hypercube Design" (CLHD). It is essentially a design for T but has good coverage over Q. This design is called "conditional" because the basic idea is to generate a few smaller LHD's conditioning on the values of basis quantities. This can be achieved because the domains are regular (hybercube) after conditioning on basis quantities. This method could be modified straightforwardly to generate other types of space filling designs (in addition to LHD) for irregular DA space in general. There are three major advantages for using CLHD to generate DA designs on irregular domains: (a) efficiency: acceptance rate is 100%; (b) good coverage for dimensionless quantities to model the nonlinear trend; (c) high power for testing linear effect in basis quantities and justifying the validity of DA.

Consider Model (1): $Q = \{(q_{b+1}, \ldots, q_k) : q_j = x_j/f_j(x_1, \ldots, x_b),\ j = b+1, \ldots, k;\ (x_1, \ldots, x_k) \in T\}$. Suppose $T = R_1 \bigoplus \cdots \bigoplus R_k$ is a k-dimensional hybercube, where R_i is the range of x_i. Thus, given the value of $x_1, \ldots, x_b$, the range of q_j is actually just a rescaled version of range of x_j by a factor of $1/f_j(x_1, \ldots, x_b)$. Denote conditional domain as $Q_{x_1,\ldots,x_b} = \{(q_{b+1}, \ldots, q_k) : q_j = x_j/f_j(x_1, \ldots, x_b),\ j = b+1, \ldots, k;\ (x_{b+1}, \ldots, x_k) \in R_{b+1} \bigoplus \cdots \bigoplus R_k\}$. Then $Q_{x_1,\ldots,x_b}$ is a $(k-b)$-dimensional rescaled hybercube. The overall design space is the union of all conditional domains $Q = \bigcup_{x_1 \in R_1, \ldots, x_b \in R_b} Q_{x_1,\ldots,x_b}$. CLHD is composed of the following four steps.

1. Generate a d level (fractional) factorial design P of size d^b on basis quantities $x_1, \ldots, x_b$. Denote as $P = \{P_r; r = 1, \ldots, d^b\}$ where each P_r is a design point $(x_{1r}, \ldots, x_{br})$.
2. For each point P_r, we assign a weight such that $w_r \geq 0$, $\sum_{r=1}^{d^b} w_r = 1$.
3. For each design point P_r, conditioning on design values $(x_{1r}, \ldots, x_{br})$, generate a space filling design Q_r of size nw_r on $q_{b+1}, \ldots, q_k$ in space $Q_{x_{1r},\ldots,x_{br}}$.
4. Combine all d^b designs $\bigcup_{r=1}^{d^b}(P_r, Q_r)$, a design on $x_1, \ldots, x_b, q_{b+1}, \ldots, q_k$ of size n can be obtained.

Notice that all design points lie within the workable domain in original scales. The acceptance rate is guaranteed to be 100%, regardless of how irregular the DA domain is. Also, the space filling design Q_r does not have to be a LHD, in which case it is more appropriate to be called as a "Conditional Space Filling Design".

Parameters d and w_r need to be determined for CLHD. In DA, b is usually small (3∼4). From Buckingham's Π theorem, we know that $x_1, \ldots, x_b$ do not have effects on the model. Thus, level d is taken to be small, say 2. The weights w_r serve as a tuning parameter providing some freedom in choosing the most appropriate design according to the context. The naive choice will be the equal weights $w_r = 1/d^b$. A better choice will be weights proportional to the volume of conditional domains $w_r \propto Vol(Q_r)$. For larger domain areas, more design points are assigned. A third alternative would be to take the weights that optimize the discrepancy criterion of the resulting design. It reduces clusters due to overlaps in the conditional domains. The resulting weights depend on the shape of Q and

function f_j, $j = b+1, \ldots, k$ to yield a "conditional" optimal design. The third approach is the best in terms of the uniformity, but requires extensive computations. We implement the equal weights in the later examples because this simple version is sufficient to show the unique performance of CLHD.

Note that this approach designs the experiments based upon both $q_{b+1}, \ldots, q_k$ and $x_1, \ldots, x_b$. Thus, besides modeling dimensionless $q_{b+1}, \ldots, q_k$, it is possible and effective to test the effect of basis quantities $x_1, \ldots, x_b$ empirically. Considering Eq. [1] as the model under the null hypothesis and Eq. [2] as the one under the alternative, we can test statistically whether Buckingham's Π theorem is reasonable for a specific data set. Violation of Buckingham's Π theorem can be viewed as a sign of missing key physical quantities. Such a test is a model validation step. It justifies whether DA modeling is appropriate.

Table 1 displays an illustrative example of the Conditional LHD. Suppose the original variables $x_1, \ldots, x_4$ have hypercube domains. x_1 is the basis quantity and q_2, q_3, q_4 are DA transformed variables correspondingly. We first have a two-level design on x_1. Conditioning on $x_1 = 1$, runs 1–3 form an LHD on q_2, q_3, q_4; conditioning on $x_1 = 2$, runs 4–6 form another LHD. To facilitate comprehension, the structures of two LHDs (runs 1–3 and runs 4–6, respectively) are kept the same: only the scaling is different which is determined by values and forms of x_1 in respective q_i's. However, it is not necessary to keep the same design structure: one can use different LHDs or even different types of space filling designs for each conditional design. Primarily for modeling purpose, we focus on the design of dimensionless q_2, q_3, q_4. As shown from Table 1, the three right most columns contain few repetitions but good coverage of the domain, providing adequate information for modeling. On the other hand, one may be interested in testing the effect of x_1 to check whether Buckingham's Π theorem holds and whether the key variables are missing. Then a linear effect seems adequate. End points are repeated to yield a high power of detection.

Table 1. An illustrative example of the Conditional LHD. $x_1, \ldots, x_4$ have ranges [1,2]. $q_2 = x_2 \times x_1$, $q_3 = x_3 \times x_1$, $q_4 = x_4/x_1$.

Runs	Basis quantities	Dimensionless quantities		
	x_1	q_2	q_3	q_4
1	1	1	3/2	2
2	1	3/2	2	1
3	1	2	1	3/2
4	2	2	3	1
5	2	3	4	1/2
6	2	4	2	3/4

Because $Q = \bigcup_{x_1 \in R_1, \ldots, x_b \in R_b} Q_{x_1, \ldots, x_b}$, all regions in Q are represented if basis quantities $x_1, \ldots, x_b$ are enumerated. When $x_1, \ldots, x_b$ take discrete values, unions of $Q_{x_1, \ldots, x_b}$ may not cover the entire Q, especially when the ranges $R_1, \ldots, R_b$ of the basis quantities are relatively large. Denote $R_i = [u_i, v_i]$ with $v_i > u_i > 0$ and v_i/u_i as its range ratio. When $v_i/u_i \gg v_j/u_j$ for $i \in \{1, \ldots, b\}$ $j \in \{b+1, \ldots, k\}$, $Q_{x_1, \ldots, x_b}$ will be far apart. There are two solutions: (a) choose basis quantities with smallest ranges ratios v_i/u_i, which is always possible; and (b) choose a multi-level design (or space filling design) for basis quantities, such as LHDs. On the other hand, when $v_i/u_i \ll v_j/u_j$ for $i \in \{1, \ldots, b\}$ and $j \in \{b+1, \ldots, k\}$, $Q_{x_1, \ldots, x_b}$ are not disjoint. The overlapping regions are overrepresented. They are mainly in the center of the domain where same values of dimensionless $q_{b+1}, \ldots, q_k$ can be obtained by multiple settings of basis quantities $x_1, \ldots x_b$. Such a repetition is useful in averaging out effects of $x_1, \ldots, x_b$ when Buckingham's Π theorem does not hold exactly due to numerical errors and random seed variations in computer simulations. Overall, we propose the range ratios v_i/u_i as the guidance for balancing between designs on basis quantities and dimensionless quantities to yield a full uniform coverage of Q. If uniformity is a strong concern, one can choose an appropriate uniform criterion such as MaxPro (Joseph et al., 2015) and optimize with respect to w_r and Q_r.

The Conditional LHD provides a more uniform space filling of domain Q, compared to designs on original space T. Figure 1 shows the comparison between the coverage performance of the CLHD in Table 1 and a transformed Maximin LHD. The transformed Maximin LHD is constructed by generating a Maximin LHD (using package "lhs" in R) on $x_1, \ldots, x_4$ and transform it to x_1, q_2, q_3, q_4. Figure 1 shows that the points of CLHD are well separated without many repetitions under each projection. However, the transformed Maximin LHD generates clusters of points and its projection to individual dimensions creates many repetitions that are not desirable for modeling computer experiments. Therefore, a space filling design on Q, such as CLHD and RLHD, is necessary. CLHD is more efficient than RLHD in irregular DA domains, in terms of the acceptance rate.

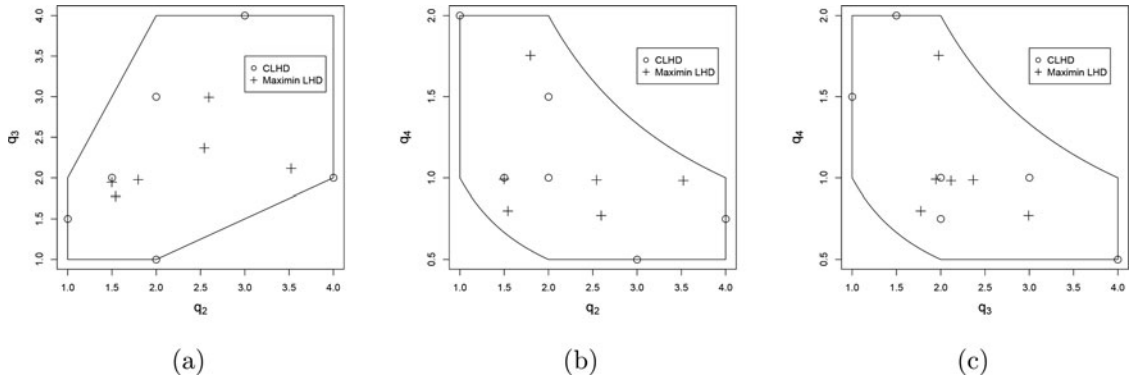


Figure 1. Coverage comparison between Conditional LHD (circle) and transformed Maximin LHD (plus) for the illustrative example. (a) Displays design points projected to q_2, q_3; (b) displays design points projected to q_2, q_4; (c) displays design points projected to q_3, q_4.

There are cases when T is far from hybercube (regular) but Q is close. In fact, our experiences indicate that this is more likely in practice. Under such situations, it is most appropriate to choose DA factor space and implement a space-filling design on Q.

It is worthwhile to mention that the Conditional LHD has a similar feature as those in the Sliced Space-filling Design (Qian, 2012). However, the main purpose of sliced space-filling design is to incorporate categorical variables, while Conditional LHD is for quick generation of DA designs. The different objectives lead to different designs.

Theoretical properties

Some theoretical discussions are given below to ensure the physical and statistical properties for the proposed method. The models described below are used because of their simplicity and interpretation. However, similar results can also be derived for other models. To facilitate comparisons, assume the non-DA model is $q_0 = h(x_1, \ldots, x_b, q_{b+1}, \ldots, q_k, \epsilon)$, while the DA model is $q_0 = h(q_{b+1}, \ldots, q_k, \epsilon)$. ϵ is the error/lack-of-fit.

From a physical perspective, a DA model is better than a non-DA model in the following aspects. First, since the variables are dimensionless, the functional dependence is free from units. Dimensionless quantities are conventionally preferred in describing the characteristics of a system. Second, the power law form of the transformed variables, described in Buckingham (1914), enables scalability in the original variables. Extrapolations in original variables may lead to interpolations in their dimensionless form. This is particularly favorable to scaling problems and acceleration experiments. Third, running experiments with b less variables could reduce sources of variations and provide consistency. The variations and measurement errors of $x_1, \ldots, x_b$ will not propagate to the response. Finally, according to Buckingham's Π-theorem, there will be no information loss if the function is physically legitimate (Buckingham, 1914).

From a statistical perspective, the main advantage of DA model is a reduced factor space without losing empirical information. It generates the following benefits. First, the uncertainty in the data may result in significant $x_1, \ldots, x_b$, but by the Buckingham's Π-theorem they should not be in the model. Without considering $x_1, \ldots, x_b$, we minimize the risk of falsely concluding the trivial significance because of randomness. Second, the problem of collinearity is now less severe. Less correlated structures in variables alleviate confounding issues, and facilitate extrapolations. Third, the extra degrees of freedom of data (based on a model with less parameters) provide more consistent estimation of random errors, and thus a more reliable inference.

The efficiency of estimating parameters corresponding to $q_{b+1}, \ldots, q_k$ increases, as shown in the following three scenarios.

1. Linear case.
   Rewrite $q_0 = \beta_1 x_1 + \cdots + \beta_b x_b + \beta_{b+1} q_{b+1} + \cdots + \beta_k q_k + \epsilon$ in a vector form $q = X\beta + \epsilon$ with $\epsilon \sim N(0, \Sigma)$. Suppose $X = (B, Q)$ is of full rank, $B = (x_1, \ldots, x_b)$ and $Q = (q_{b+1}, \ldots, q_k)$. Also, $\beta^T = (\beta_B^T, \beta_Q^T)$, where $\beta_B = (\beta_1, \ldots, \beta_b)^T$ and $\beta_Q = (\beta_{b+1}, \ldots, \beta_k)^T$. This representation covers both regression and Kriging models, by treating rows of ϵ as from repeated observations or from a realization of Gaussian process respectively. For simplicity, we

can use ordinary least square estimates $\widehat{\beta}$ and proceed with regression settings with $\Sigma = \sigma^2 I$. Notice that Kriging estimates are merely generalized least squares and can be transformed to the above regression case. Then for model (1),

$$\mathrm{Var}(\widehat{\beta}) = \hat{\sigma}^2 (X^T X)^{-1} = \hat{\sigma}^2 \begin{pmatrix} B^T B & B^T Q \\ Q^T B & Q^T Q \end{pmatrix}^{-1}, \text{ and}$$

$$\mathrm{Var}(\widehat{\beta}_Q) = \hat{\sigma}^2 \left(Q^T Q - Q^T B (B^T B)^{-1} B^T Q\right)^{-1}.$$

For the same design on Q with model (2),

$$\mathrm{Var}(\widehat{\beta}'_Q) = \hat{\sigma}^2 \left(Q^T Q\right)^{-1}.$$

Now, $B(B^T B)^{-1} B^T$, as well as $I - B(B^T B)^{-1} B^T$, is idempotent and thus semi-positive definite. Therefore, $Q^T Q \geq Q^T Q - Q^T B (B^T B)^{-1} B^T Q \geq 0$ and $(Q^T Q)^{-1} \leq (Q^T Q - Q^T B (B^T B)^{-1} B^T Q)^{-1}$ if invertible, with equal sign if and only if Q, B are mutually orthogonal. Thus, $\mathrm{Var}(c^T \widehat{\beta}'_Q) \leq \mathrm{Var}(c^T \widehat{\beta}_Q), \forall c \in \mathbb{R}^{k-b}$. In other words, estimates of coefficients have smaller variances by DA procedure. In particular, the estimated responses are $\widehat{q}_0 = Q\widehat{\beta}_Q$ for model (1), $\widehat{q}'_0 = Q\widehat{\beta}'_Q$ for model (2), and $\mathrm{Var}(\widehat{q}'_0) \leq \mathrm{Var}(\widehat{q}_0)$.

2. Fisher information and maximum likelihood estimator.
As a generalization to the above case, suppose $\chi = (q_0, x_1, \ldots, x_b, q_{b+1}, \ldots, q_k)$ follows some distribution with density f and parameter θ. Then the Fisher information matrix is

$$\mathcal{I}(\theta) = E\left[\left(\frac{\partial}{\partial \theta} \log f(\chi; \theta)\right) \left(\frac{\partial}{\partial \theta} \log f(\chi; \theta)\right)^T \middle| \theta \right].$$

Suppose $\theta = (\theta_B^T, \theta_Q^T)^T$, where θ_B and θ_Q characterize effects of basis quantities $\tilde{B} = (x_1, \ldots, x_b)^T$ and transformed variables $\tilde{Q} = (q_0, q_{b+1}, \ldots, q_k)^T$, respectively. Under the true DA model, assume θ_B is fixed at values indicating the zero effects of basis quantities. Therefore, $\mathcal{I}(\theta)$ can be expressed as a block matrix:

$$\mathcal{I}(\theta) = \begin{pmatrix} \mathcal{I}_B(\theta) & \mathcal{I}_{BQ}(\theta) \\ \mathcal{I}_{QB}(\theta) & \mathcal{I}_Q(\theta) \end{pmatrix}.$$

Only $\mathcal{I}_Q$ (the information to Q) is of main interest:

$$\mathcal{I}_Q(\theta) = E\left[\left(\frac{\partial}{\partial \theta_Q} \log f(\chi; \theta)\right) \times \left(\frac{\partial}{\partial \theta_Q} \log f(\chi; \theta)\right)^T \middle| \theta \right].$$

Following model (1), suppose $\hat{\theta}$ is the maximum likelihood estimator. Under mild regularity conditions, $E(\hat{\theta}) \to \theta$ and $\mathrm{Var}(\hat{\theta}) \approx \mathcal{I}^{-1}(\theta)$. Therefore, its sub-vector $E(\hat{\theta}_Q) \to \theta_Q$ and $\mathrm{Var}(\hat{\theta}_Q) \approx (\mathcal{I}_Q(\theta) - \mathcal{I}_{QB}(\theta)\mathcal{I}_B^{-1}(\theta)\mathcal{I}_{BQ}(\theta))^{-1}$.

Following model (2), θ_B is known to take values indicating zero effects and $\hat{\theta}'_Q$ is the maximum likelihood estimator. Under mild regularity conditions, $E(\hat{\theta}'_Q) \to \theta_Q$ and $\mathrm{Var}(\hat{\theta}'_Q) \approx \mathcal{I}_Q^{-1}(\theta)$. $\mathrm{Var}(\hat{\theta}'_Q) \leq \mathrm{Var}(\hat{\theta}_Q)$ with equal sign if and only if $\mathcal{I}_{QB}(\theta) = 0$.

3. Dimension reduction.
Suppose $Y, X_1, \ldots, X_k$ are random variables, and so are $q_0, q_{b+1}, \ldots, q_k$. We want to determine the relationship between Y and $X_1, \ldots, X_k$ and equivalently between q_0 and $X_1, \ldots, X_k$. The model (2) and DA principle indicate that

$$q_0 \perp\!\!\!\perp X_1, \ldots, X_k | q_{b+1}, \ldots, q_k,$$

which means that q_0 and $X_1, \ldots, X_k$ are independent given $q_{b+1}, \ldots, q_k$. $q_{b+1}, \ldots, q_k$ is a sufficient dimension reduction of $X_1, \ldots, X_k$ with regard to the response q_0. q_j's actually take the power law form: $q_j = X_j X_1^{a_{j1}} \cdots X_b^{a_{jb}}$. By taking a logarithm, $\log(q_j) = \log(X_j) + \sum_{i=1}^{b} a_{ji} \log(X_i) = \sum_{i=1}^{k} a_{ji} \log(X_i)$. Denote $\log(X_i)$ as $\tilde{X}_i$ and $\tilde{X} = (\tilde{X}_1, \ldots, \tilde{X}_k)$. Then

$$q_0 \perp\!\!\!\perp \tilde{X}_1, \ldots, \tilde{X}_k | A\tilde{X},$$

where $A = (a_{ji})$ is a $(k-b) \times k$ matrix. As a result, DA can be viewed as a predetermined physics-originated procedure to identify the linear sufficient dimension reduction space. Design and analysis via DA study the relationship of interest in a confined linear space, and do not lose information. Similar theoretical benefits can be derived as an analogy to statistical dimension reduction. Conversely, statistical dimension reduction approaches can be viewed as the data-driven versions of DA. The natural connection shown above proposes a general perspective on unifying DA as a scientific

approach and statistical dimension reduction as an empirical one, under the framework of sufficient dimension reduction.

The reduction in dimensions of variables also yields benefits under the Bayesian framework, which will not be discussed here. The theoretical properties from physical and statistical views can be further illustrated by the numerical examples in the following section.

Case studies

In this section, comparisons between designs and analyses with and without DA are performed on two examples, the borehole model and the damped harmonic oscillator. The case on borehole model serves as a detail illustration on how to perform DA in computer experiments. The case on damped harmonic oscillator emphasizes on showing how DA greatly improves predictions and why it is more powerful.

For both case studies, the responses can be explicitly expressed as known equations of the inputs, which are typically unavailable for computer models of interest. However, the simplicity leads to a quick assessment of the accuracy of predictions by comparing to the ground truth. The true theoretical solution is unknown during the design/analysis phase and only used for generating response in data collection and model assessment. Designs and analyses are conducted to estimate the functional relationship between the response and the input variables. Our objective is to compare the prediction performances of procedures with and without dimensional analysis.

To be specific, we apply random LHDs as non-DA designs. Then their rejected versions are taken as RLHDs and their conditional versions with equal weights as CLHDs. The reason of choosing random designs, rather than optimal designs, as kernels for all design strategies is to eliminate effects of optimal criterion and domain regularity on the performance, and to induce a fair comparison between the DA factor space and the original space. If an optimal design criterion is used, different domain shapes possibly lead to incomparable designs. The purpose here is to provide a comprehensive view on the "pure" effect of DA procedure, regardless of the specific type of LHD used in practice. Furthermore, random LHDs also provide us information about the robustness of DA modeling to the random variation of initial design. After assigning the design strategies for the input variables, the theoretical model is used to generate responses given the design of input variables, emulating certain complicate computer models. Two kinds of statistical analysis techniques, regression and Kriging, are used as examples. These two models are well-accepted and widely implemented. We use the regression as a benchmark for its parsimonious representation. Kriging is adopted due to its ability to capture complex nonlinearity. See Cook and Weisberg (1982), Santner et al. (2003), and Williams (1998) for details on the model assumptions and procedures.

After fitting the models, their respective response predictions are obtained on a set of validation points, which consists of 10^4 LHD points for the original variables. This stands for the evaluation of their respective performances on new random inputs. The criteria on which we base our evaluations are Mean Square Error (MSE $= \sum_{i=1}^{n}(y_i - \hat{y}_i)^2/n$) and Maximum Absolute Deviation (MAD $= \max_{i=1\ldots n}\{|y_i - \hat{y}_i|\}$), where y_i's are the theoretical values and $\hat{y}_i$'s are the predicted values from the estimated model. Finally, to assess DA over a wide range of designs, we generate the empirical distributions of MSE and MAD over a Monte Carlo sample of the initial training LHD. From that, we remove the randomness and assess the reliability and robustness of the DA procedure to the choice of designs.

Case Study - I: Borehole model

Borehole model

Illustrated in Figure 2 is a model widely used to describe the flow of water through a borehole from the ground surface through two aquifers. Previous

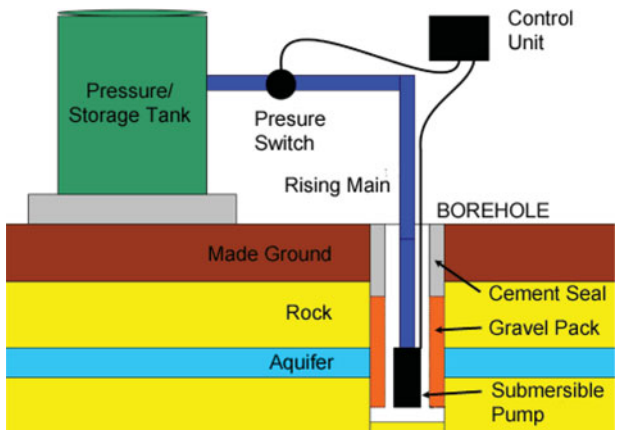


Figure 2. Illustration of borehole and water flow. Picture from http://www.waterseekers.co.uk/borehole-drilling-faq.php#&panel1-1.

Table 2. Table of physical meanings, ranges and dimensional symbols of variables in borehole model.

Variable	Physical meaning	Range	Dimension
y_0	Borehole flow rate		**[L³/T]**
r_w	Radius of borehole	0.05 to 0.15 m	**[L]**
r	Radius of influence	100 to 50,000 m	**[L]**
T_u	Transmissivity of upper aquifer	63,070 to 115,600 m²/yr	**[L²/T]**
H_u	Potentiometric head of upper aquifer	990 to 1,110 m	**[L]**
T_l	Transmissivity of lower aquifer	63.1 to 116 m²/yr	**[L²/T]**
H_l	Potentiometric head of lower aquifer	700 to 820 m	**[L]**
L	Length of borehole	1,120 to 1,680 m	**[L]**
K_w	Hydraulic conductivity of borehole	1,500 to 15,000 m/yr (extended)	**[L/T]**

studies on this model include Harper and Gupta (1983), Worley (1987), Morris et al. (1993), and Fang and Lin (2003). The model is based on assumptions of (i) no groundwater gradient; (ii) steady-state flow from upper aquifer into the borehole and from the borehole into the lower aquifer; and (iii) laminar, isothermal flow through the borehole. In this model, the flow rate through borehole y_0 as response is given by

$$y_0 = \frac{2\pi T_u(H_u - H_l)}{\ln(r/r_w)[1 + \frac{2LT_u}{\ln(r/r_w)r_w^2 K_w} + \frac{T_u}{T_l}]}. \quad [4]$$

Table 2 shows the physical meanings, ranges, units and dimensional symbols of the eight inputs and the response. The form of function seems highly nonlinear. Particularly, if the range of K_w is extended to $[1500, 15000]$ from $[9855, 12045]$, Figure 3 shows a nonadditive trend between the response y_0 and the inputs r_w and K_w when the remaining inputs are at their respective lowest values. Here, we focus on the extended range case. The detail analyses for the original range case is given in Appendix A.

Dimensional analysis on borehole model

The dimensional analysis is applied according to physical dimensions summarized in Table 2. Only two dimensions are involved: length **L** and time **T**. We need two quantities to represent these two dimensions and choose H_u and T_u as basis quantities based on smallest range ratios. Of course, this choice is not unique. Given that the borehole Eq. [4] is dimensionally homogeneous, the choice of basis quantities and their working values given dimensionless design Q should not affect results.

Other quantities are transformed into dimensionless forms as follows:

$$q_0 = \frac{y_0}{H_u T_u},\ q_1 = \frac{r_w}{H_u},\ q_2 = \frac{r}{H_u},\ q_5 = \frac{T_l}{T_u},$$
$$q_6 = \frac{H_l}{H_u},\ q_7 = \frac{L}{H_u},\ q_8 = \frac{K_w H_u}{T_u}. \quad [5]$$

Designs and analyses on DA factor space

The next step is to reformulate the prediction function. Without dimensional analysis, function g needs to be estimated, where

$$y_0 = g(r_w, r, T_u, H_u, T_l, H_l, L, K_w). \quad [6]$$

With dimensional analysis, our model is

$$\frac{y_0}{H_u T_u} = h\left(\frac{r_w}{H_u}, \frac{r}{H_u}, \frac{T_l}{T_u}, \frac{H_l}{H_u}, \frac{L}{H_u}, \frac{K_w H_u}{T_u}\right). \quad [7]$$

And the recovery function will be $\hat{y_0} = H_u T_u \cdot \hat{h}(r/H_u, r_w/H_u, T_l/T_u, H_l/H_u, L/H_u, K_w H_u/T_u)$. Note that, by rewriting model (4), the true h is of the form:

$$\frac{y_0}{H_u T_u} = 2\pi\left(1 - \frac{H_l}{H_u}\right)\left[\left(\ln\left(\frac{r}{H_u}\right) - \ln\left(\frac{r_w}{H_u}\right)\right) \times \left(1 + \frac{T_u}{T_l}\right) + 2\frac{L}{H_u}\left(\frac{H_u}{r_w}\right)^2 \frac{T_u}{K_w H_u}\right]^{-1}. \quad [8]$$

The simulation is done as follows: first, 80 design points are generated on Q and T, respectively. The non-DA design is (a) the random LHD; the DA designs are (b) LHD with rejection method (RLHD); and (c) the proposed conditional LHD (CLHD) as discussed in the Methodology section. The responses are then obtained by Eqs. [4]. Analyses including stepwise regression and

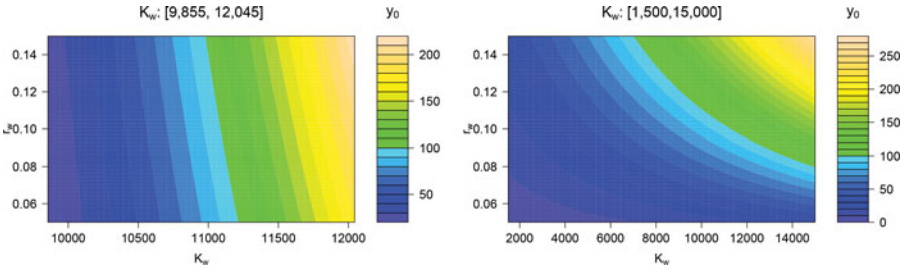


Figure 3. Contour plot of response y_0 with different ranges of K_w.

Kriging method are implemented to train the prediction model for both non-DA model (6) and DA model (7). The stepwise regression starts from a full quadratic model with pairwise interactions and selects variables forward and backward based on AIC. Kriging is done by first fitting an exponential correlation function on standardized covariates and response with no drift and nugget terms. Separate scale parameters are estimated by maximizing the likelihood. Then the MLEs of the scale parameters are fed to the exponential correlation function by "Krig" function in R package "fields." The model on standardized covariates with a linear drift trend and no nugget terms is fitted to produce predictions. Finally, the evaluation measures, MSE and MAD, are achieved on 10^4 validation points generated by maximin-LHD on original variables.

Results

Table 3 summarizes the performance measures of each combination of procedures. 100 simulations of designs are generated, evaluated with the same validation points. The average MSEs and MADs across 100 simulations are listed according to the designs and analyses. The corresponding standard errors are in the parentheses. If a modeling procedure is robust, we would expect a low MSE/MAD with little variance.

Table 3. Table of performance measures on LHD validation points in borehole example with extended range of K_w.

	MSE		MAD	
	Regression	Kriging	Regression	Kriging
Non-DA	32.03(0.89)	48.15(1.34)	48.04(0.76)	80.79(1.62)
DA-RLHD	35.46(0.98)	40.88(1.27)	51.73(0.81)	69.32(1.35)
DA-CLHD	25.27(0.58)	44.11(1.24)	47.82(0.60)	84.63(1.59)

Note that the response is rather nonlinear. Table 3 shows that DA procedures generally perform better than procedures without DA (except DA-RLHD for Regression). DA with CLHD performs well for regression method. It reduces MSE of prediction errors from non-DA model greatly. DA with RLHD performs well for Kriging method. It reduces both the MSE and MAD significantly. Compared with naive CLHD, RLHD is more uniform and puts more weights on those extreme edges. This is particularly good for Kriging to capture the exotic values, but bad for regression. A polynomial tends to fit these outliers, leading to an unstable predictive model. This is the reason why DA with RLHD performs worse than stepwise regression without DA variables. Overall, modeling on DA factor space has a potential in reducing the MSE of the prediction errors.

In additional to the means, Figure 4 shows the empirical distributions of MSE and MAD. Distributions of errors for procedures with DA are sharper in some cases. However, the difference is not that large.

Note that regression has a better performance in terms of predictions than Kriging in this case. The reason for lack of fit for regression method may be due to the overfitting and correlated interactions. The resulting models selected by stepwise regression often have an R^2 close to 1. Its better performance over Kriging

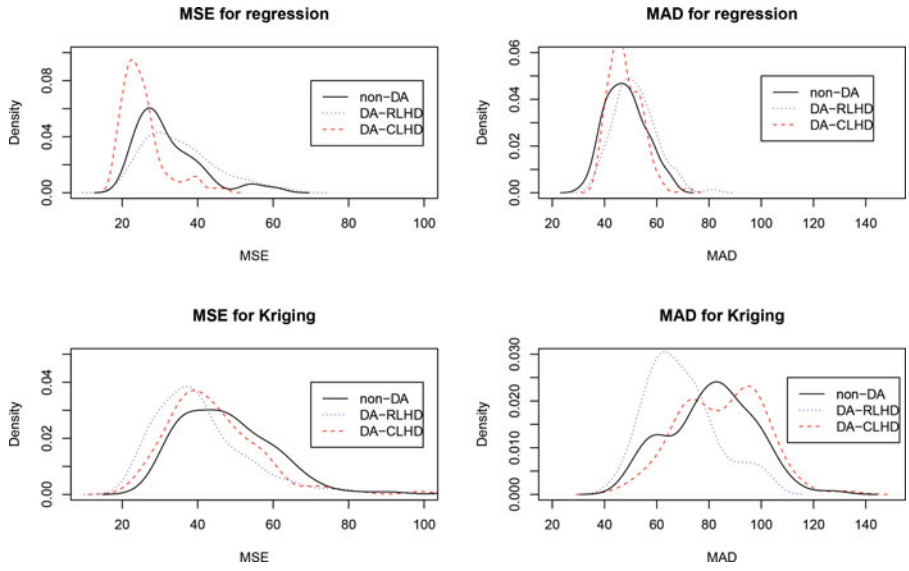


Figure 4. The empirical distribution of performance measures for different methods with and without DA in the case of extended K_w range.

Table 4. Table of variances of coefficients in borehole example with extended range of K_w. Values are in the scale of $\times 10^{-2}$.

Component	β_3	β_4	β_5	β_6	β_7	β_8
$\text{Var}(\widehat{\beta}_Q)$	1.42(0.01)	1.36(0.01)	5.07(0.02)	3.48(0.02)	1.69(0.01)	1.86(0.01)
$\text{Var}(\widehat{\beta}'_Q)$	1.37(0.01)	1.33(0.01)	1.62(0.01)	1.52(0.01)	1.51(0.01)	1.57(0.01)

may be attributed to the small size of data set. We use 80 points for the original problem with 8 variables for both DA and non-DA methods. Kriging provides too many degrees of freedom in modeling, leading to overfitting and lack of stability. Regression methods, on the other hand, restrain its predictive values to a hyperplane and recover the general trend. However, it is well known that Kriging is better in capturing the nonlinear complex dependence than regressions. Thus, we expect the performances will be comparable, or favorable to Kriging, as sample size gets larger.

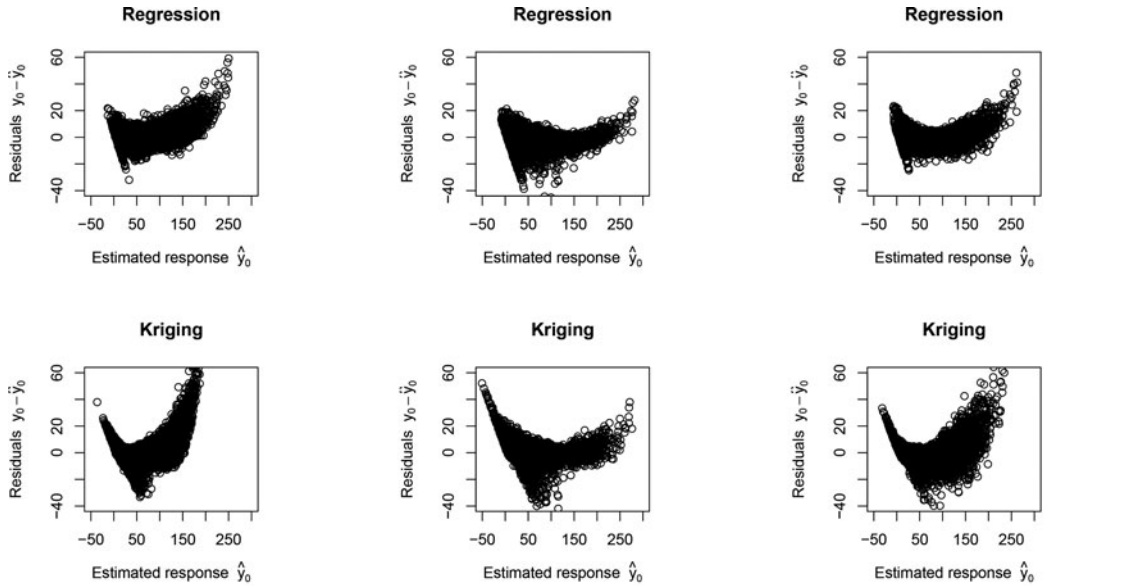
The variances $\text{Var}(\widehat{\beta}_Q)$ and $\text{Var}(\widehat{\beta}'_Q)$ can be numerically calculated. In particular, we generate 100 CLHDs and standardize the covariates. The diagonal of $\text{Var}(\widehat{\beta}_Q)$ and $\text{Var}(\widehat{\beta}'_Q)$ are displayed in Table 4. It can be seen that the results are consistent to the theories derived in the Theoretical Properties section. $\text{Var}(\widehat{\beta}'_Q)$ are smaller, especially for β_5 and β_6. $\text{Var}(\widehat{\beta}'_Q)$ is close to the theoretical variance $(n-1)^{-1} \approx 0.01266$ when covariates are orthogonal, but $\text{Var}(\widehat{\beta}_Q)$ is not. This occurs when Q has nearly orthogonal columns but is not orthogonal to B. Therefore, CLHD generates rather uncorrelated columns for dimensionless quantities but they may be correlated with basis quantities. Great reduction of variance can be achieved when applying DA and only modeling dimensionless quantities.

Figure 5 shows the prediction errors against estimated responses using different metamodels out of one run. Figure 5a features a clear quadratic pattern of residuals with non-DA procedures. Significant systematic bias is generated especially when using Kriging. By incorporating DA, we notice a sign of bias correction in Figures 5b and 5c. Outliers for regression become less severe. The residuals by Kriging seem to be more even and the bias pattern that is extreme on larger estimated responses for Kriging is drawn towards 0.

Case Study - II: Damped harmonic oscillator

Damped harmonic oscillator

In the classic mechanics, a damped harmonic oscillator is a system that, when displaced from its equilibrium position, experiences a restoring force F which is proportional to the displacement x, and a frictional force f, proportional to the velocity v. Harmonic oscillators are very important in physics because any object in stable equilibrium acts like a harmonic oscillator for small vibrations. Extensive researches have been done associated with damped harmonic oscillator in mechanical engineering, control engineering, structural engineering, electrical engineering, among others. Reference



(a) procedures without DA (b) procedure of DA-RLHD (c) procedure of DA-CLHD

Figure 5. Plot of prediction errors against true responses for various procedures.

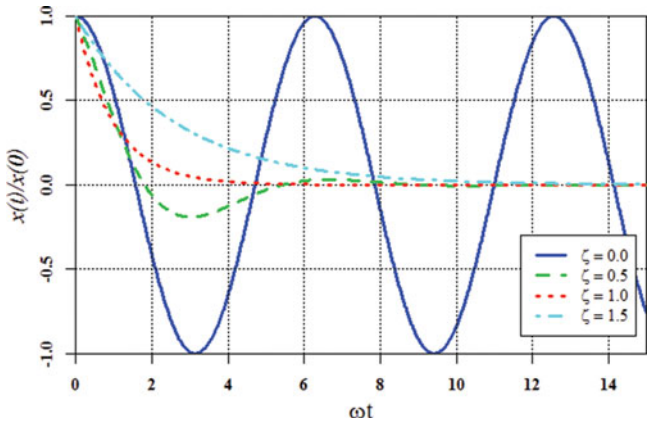


Figure 6. Behavior of damped harmonic oscillator with different damping modes.

of computer-aided experiment on damped harmonic oscillator could be found in McInerney (1985).

The Newton's Law gives the second order ordinary differential equation for characterizing this system:

$$m\frac{\mathrm{d}^2x}{\mathrm{d}t^2} + c\frac{\mathrm{d}x}{\mathrm{d}t} - kx = 0, \qquad [9]$$

where x is the displacement, t is the time, m is the mass, k is the proportional constant of restoring force F to the displacement x, called *spring constant*, and c is the proportional constant of frictional force f to velocity v, called *viscous damping coefficient*. Eq. [9] can also be written as:

$$\frac{\mathrm{d}^2x}{\mathrm{d}t^2} + 2\zeta\omega_0\frac{\mathrm{d}x}{\mathrm{d}t} + \omega_0^2x = 0, \qquad [10]$$

where $\omega_0 = \sqrt{k/m}$ is the undamped angular frequency and $\zeta = c/(2\sqrt{mk})$ is the damping ratio. Given the initial condition of $x(0) = x_0$; $x'(0) = 0$, we can solve this differential equation analytically (see Appendix B). The solution for the displacement is

$$x = x_0\mathbf{Re}[\mathrm{e}^{-\omega_0 t(\zeta - \sqrt{\zeta^2-1})}] \qquad [11]$$

From the above equation, we can characterize the oscillation modes into the following four types by different damping ratios. See Figure 6 for the illustration of the system behaviors with different oscillation modes.

- Undamped ($\zeta = 0$). There is no damping effect. The motion is periodic, repeating itself in a sinusoidal fashion with constant amplitude.
- Underdamping ($0 < \zeta < 1$). The damping effect is small. The system oscillates with the amplitude gradually decreasing to zero with angular frequency $\omega_1 = \omega_0\sqrt{1-\zeta^2}$.
- Critical damping ($\zeta = 1$). The system returns to steady state as quickly as possible without oscillating.
- Overdamping ($\zeta > 1$). The damping effect is strong. The system decays exponentially to steady state without oscillating with rate $\omega_0(\zeta - \sqrt{\zeta^2-1})$.

Here, we focus on the relatively more complex mode: underdamping oscillation. Further investigations on overdamping oscillation mode are also provided in Appendix C.

Dimensional analysis on damped harmonic oscillator

The main issue here is to predict the displacement x at an arbitrary time and physical configuration. We assume the initial velocity $x'(0) = 0$. To apply dimensional analysis, we first examine the physical dimensions of the variables in the system as the beginning step. These variables are summarized in Table 5. There are three dimensions involved: length **L**, mass **M**, and time **T**. We choose 3 basis quantities and reduce the number of variables from 5 to 2. The oscillator equation [11] is dimensionally homogeneous, therefore different choices yield same results theoretically.

Next, we choose m, k, and x_0 as basis quantities. The other quantities x, t, and c are transformed into dimensionless ones:

$$\pi_0 = \frac{x}{x_0}, \quad \pi_1 = \sqrt{\frac{k}{m}}t = \omega_0 t, \quad \pi_2 = \frac{c}{2\sqrt{km}} = \zeta \qquad [12]$$

As a result, π_0, π_1, π_2 are three new variables to be analyzed.

Designs and analyses on DA factor space

Without dimensional analysis, we would like to estimate function g, where

$$x = g(m, k, x_0, t, c). \qquad [13]$$

Table 5. Table of physical meanings, ranges, and dimensional symbols of variables in underdamped harmonic oscillator system.

Variable	Physical meaning	Range	Unit	Dimension
x	Displacement		meter	**[L]**
m	Mass	[1,5]	kilo.	**[M]**
t	Time	[1,10]	sec.	**[T]**
k	Spring constant	[1,5]	kilo./sec.^{2}	**[M/T^2]**
c	Damping coefficient	[0,2]	kilo./sec.	**[M/T]**
x_0	Initial displacement	[5,20]	meter	**[L]**

With dimensional analysis, we would estimate h, where

$$\frac{x}{x_0} = h\left(\sqrt{\frac{k}{m}}t, \quad \frac{c}{2\sqrt{km}}\right). \qquad [14]$$

The ranges of input variables are specified in Table 5, where $0 < \zeta < 1$ is satisfied and underdamping is guaranteed. We generate around 80 design points in the training step for each approach. The corresponding responses are then generated via theoretical model of Eq. [11]. (Note that this can also be done via numerical solutions to the differential Eq. [9], which is more realistic.)

We implement statistical models to estimate appropriate functions for representing our dataset. Stepwise regression and Kriging are used for estimating models (13) and (14). For non-DA model, the stepwise regression selects significant variables up to a full model of quadratic polynomials with interactions by AIC. The total number of predictors is 20. The predictors of the stepwise regression for DA model consist of all fifth-order polynomial terms with all lower order interactions. The total number of predictors is also 20. (There are only two independent variables in DA model.) Kriging is fitted in two step. First, the MLEs of separate scale parameters for exponential correlation function are obtained for a simple Guassian process on standardized covariates and response. Then, we fit a predictive model of Kriging with linear drift, no nugget terms and the above MLEs of scale parameters in the exponential correlation function, by implementing R package "fields". Respective performances are assessed by MSE and MAD criteria on 10^4 maximin-LHD points. Empirical distribution of prediction errors are obtained based on 100 simulations of the training LHDs.

Results

The values of performance measures are displayed in Table 6. It is clear that the methods with dimensional analysis are significantly better than those methods without DA, especially for MSE comparisons. DA with CLHD reduces MSE to approximately a half (∼60%) for regression and to 1/14 (∼7%) for Kriging. DA with RLHD also performs well in reducing MSE of Kriging, to 1/5 (∼20%) of the original method. The MAD of DA procedures increases in regression case but decreases in Kriging case. Such observations suggest that DA notably improves predictions of periodic-oscillating type of response. The number of variables is reduced from 5 to 2, which is one of the reasons for an improved performance. Kriging method is significantly better than stepwise regression for this underdamping case.

Table 6. Table of performance measures in underdamping case. Standard errors are in parentheses.

	MSE		MAD	
	Regression	Kriging	Regression	Kriging
Non-DA	28.35(0.39)	22.02(0.21)	26.31(0.42)	21.86(0.17)
DA-RLHD	25.55(0.68)	5.26(0.22)	30.80(0.70)	16.80(0.20)
DA-CLHD	17.68(0.24)	1.54(0.04)	35.62(1.39)	15.96(0.30)

In additional to the means, Figure 7 shows the empirical distributions of MSE and MAD. DA methods yield significantly smaller MSE than non-DA methods. The reduction of MAD is not as large, but is observed for Kriging approach. DA with CLHD seems to be particularly good. In additional to its overall smaller MSEs, the performance is very robust and consistent, resulting in sharp peaks in the empirical distributions of MSEs.

Figure 8 is the heat map of predictions from Kriging in this case. The above four panels are from Kriging without DA under four different initial designs (different random seeds for LHDs). The middle four panels are from Kriging with DA under four different RLHDs. The result of DA method with CLHD is displayed in the bottom four panels. From the plots, Kriging surface without DA seem to vary across different designs. The predictive surfaces are not robust to the designs. This inconsistency of estimated model associated with designs leads to somehow unreliable conclusions. Furthermore, the oscillating pattern in the horizontal axis and damping trend on the vertical axis can not be discovered. On the other hand, the results with DA are more consistent across random seeds. Different design points affect the final model only marginally, but the overall patterns of damped oscillations are always unveiled. In this case, Kriging with dimensional analysis is preferred and we can confidently conclude that the oscillation structure is captured.

Figure 9 shows the trace plot of prediction MSEs with respect to run sizes, achieved by Kriging with and without DA. Solid line is the mean of 500 Monte Carlo samples of MSE's, each based on 400 validation points. The dash lines characterize the approximate prediction intervals of MSE. It is shown that procedure with DA outperforms that without DA across all run sizes. The margin of the difference is large and statistically significant for moderate size $n > 30$. As

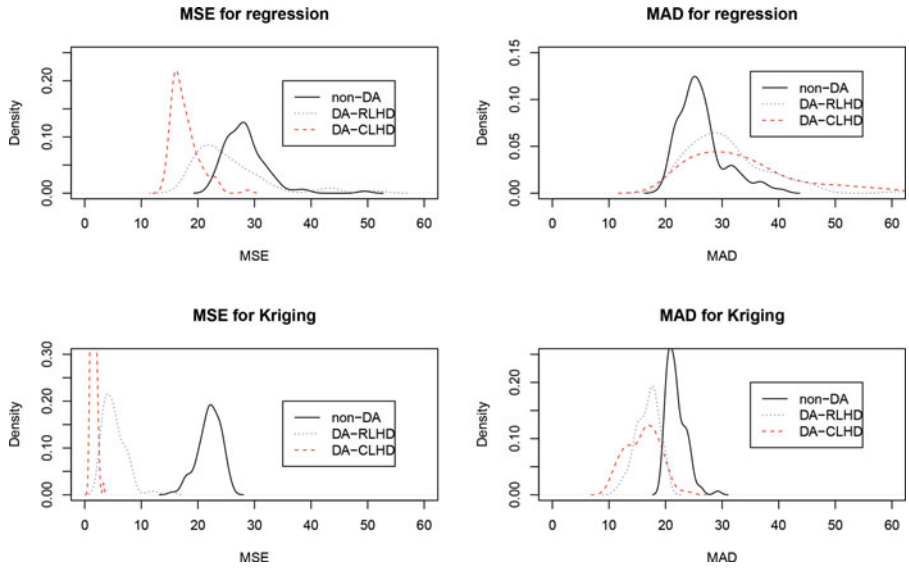


Figure 7. The empirical distribution of performance measures for different methods with and without DA in underdamping case.

an example, if the MSE is to be kept below 5 (solid horizontal line in Figure 9), the necessary number of runs for DA procedure is at most $n = 40$; while the number should be greater than 100 for non-DA procedure. We notice that, when the training design is nearly saturate ($n \approx p$; i.e., 20), procedure with DA is not stable. This is mostly because of the CLHD incorporates a full factorial for the basis quantities, leaving few degrees of freedom for the rest important quantities. Therefore, we recommend a smaller design (fractional factorial designs, Plackett-Burman Designs, or LHD) for basis quantities in CLHD if the number of runs is close to the number of variables in the model.

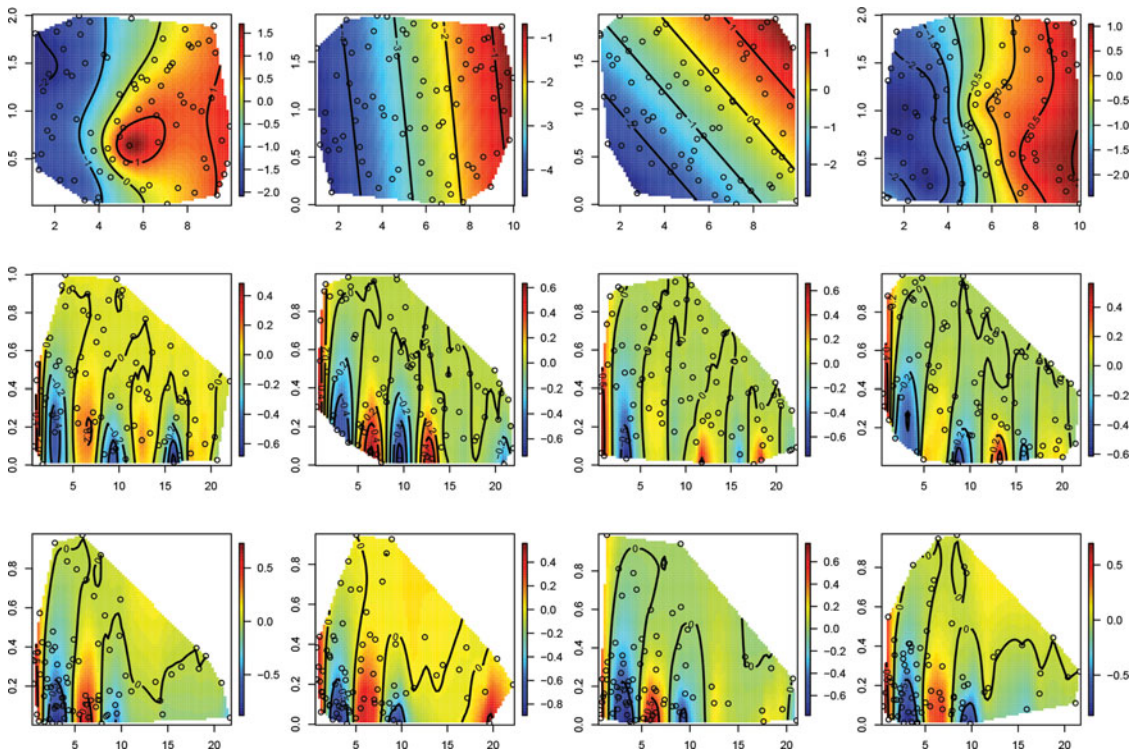


Figure 8. The predictions from Kriging with various designs non-DA (top row), DA-RLHD (middle row) and DA-CLHD (bottom row) in underdamping case. Four realizations are generated for each setting. The color stands for the value of the response. The horizontal axis refers to the phase of motion. The vertical axis refers to the levels of damping effect.

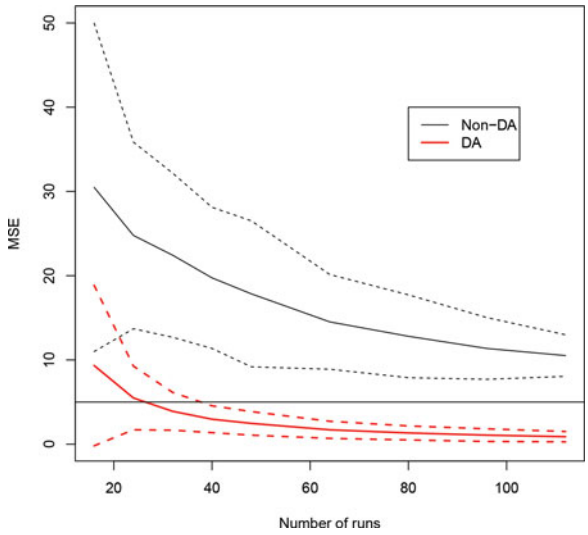


Figure 9. The trace plot of prediction MSE with respect to run size in solid lines. The dash lines are prediction intervals for prediction MSE. Kriging is adopted as metamodel and underdamping case is used. Procedure with DA has lower MSE. A reference line is drawn at MSE = 5. Note that there are $p = 5$ variables.

Across 100 CLHDs with standardized covariates, the average variances of coefficients $\mathrm{Var}(\widehat{\beta}_Q)$ and $\mathrm{Var}(\widehat{\beta}'_Q)$ are (0.0248 and 0.0127) for β_4 and (0.0210 and 0.0127) for β_5, respectively (with standard errors <0.0001). Compared with the theoretical variance $(n-1)^{-1} \approx 0.01266$ when covariates are orthogonal, we can see that β_4 and β_5 are nearly orthogonal. But they may be correlated with β_1, β_2, and β_3, leading to inflated $\mathrm{Var}(\widehat{\beta}_Q)$. The reduction in variance of β_4 (effect of time t) is remarkable. This is the case when B and Q are not orthogonal and great reductions can be achieved.

We observe the same phenomena for the predictions in the overdamping case shown in Appendix C. The prediction errors are even smaller. Indeed, the nonlinear behavior of the underdamped oscillator shown here is more complicated.

Summary

The above two examples demonstrate that, design and analysis via DA models can be of great use to computer experiments. In the case of borehole, the implementation of DA is demonstrated in detail, and the results indicate that DA designs are potentially more effective than non-DA designs for both regression and Kriging, especially when the relationship between response and predictors is nonlinear. In the case of damped harmonic oscillator, where DA is performed for a differential equation system, procedure without DA could not correctly identify the oscillating modes while procedure with DA still leads to good performance. In terms of prediction capability, DA models are better when specified properly. For a fixed design size, the reductions in MSE are remarkable. If a certain level of MSE is to be achieved, a smaller design size is required for a properly specified DA model. The consistent behavior across training designs further justifies its robustness in addition to its effectiveness and efficiency.

As illustrated by the two examples, the analytical nature of the DA model proves itself useful, feasible and accessible in physical computer experiments. The two case studies shed light on the advantages of DA modeling procedure in terms of (a) efficiency, (b) extrapolation, (c) interpretability, (d) robustness, and (e) feasibility. Compared to the borehole example, the impact of DA is more significant in damped harmonic oscillator example. We notice that DA reduces predictors from 8 to 6 in the previous one, while the reduction in the latter one is from 5 to 2. This suggests the intuition that if DA reduces more variables, it yields better improvement. However, reductions in dimensionality are of course not arbitrary, but rather dictated by the underlying physical processes.

Conclusion

In this article, we propose a new type of design and analysis strategy via DA to incorporate the physical knowledge in the computer experiments. The derived model is guaranteed to satisfy the dimensional constraints, which increases the accuracy when interpolating and extrapolating the empirical emulator. The procedure is justified through both theories and case studies. In theoretical aspects, the reduction of factor space without losing information produces less variance in estimations. From case studies, the impact of DA on improving the prediction performance has been quantified and demonstrated. The following advantages for applying DA in computer experiment can be concluded: (a) The number of experimental factors is reduced. The efficiency is thus increased, and fewer experimental runs are needed. With the same number of runs, better estimation can be achieved. (b) DA is effective in retrieving nonlinear dependence. It is common in engineering problem to have complex functional forms following physical laws. DA models incorporate the dimensional information and increase

accuracy when extrapolating. (c) The factor effects and models become more interpretable. Dimensionless variables are often used as indices to characterize and determine features of a system. The dimensionally homogeneous model complies with physical laws and is easier to understand by engineers, providing more physical insights. (d) Procedures with DA are often more robust to the designs chosen. They provide consistent performance under different designs. (e) Incorporating DA is fairly straightforward and costless. This methodology is practical and general to be applied to most engineering problems.

We also notice that, although modeling via DA yields good results for nonlinear problems, its effect is limited if the reduction is small as in the borehole case. Moreover, applying DA fully relies on the underlying physics. It is legitimate to question whether such physics as Buckingham's Π-theorem is reasonable in the context. Finally, the non-unique transformation from DA factor space Q to original space T may lead to possible predicament. However, incorporating DA in computer experiment seems promising.

About the authors

Weijie Shen is currently a Quantitative Analyst at Google, Inc. He received his a degree in Statistics at the Pennsylvania State University. Dr. Shen's main interests focus on industrial and engineering statistics, dimensional analysis, design of experiments, data mining, computer experiments, optimization, and time series. He has received awards including ASQ Brumbaugh award, DAE 2015 travel award, Q&P JSM student travel award, 2013 FTC student grants, QPRC 2013 student scholarship, Vollmer-Kleckner scholarship ,and university graduate fellowship from Penn State University. He is currently a member of the American Statistical Association (ASA) and Institute of Mathematical Statistics (IMS).

Dennis K. J. Lin is currently a Distinguished Professor of Statistics and Supply Chain in Department of Statistics at Pennsylvania State University. He received a Ph.D. degree in Statistics at University of Wisconsin-Madison. Dr. Lin's general interests focus on statistical methodology in business, industrial and government (BIG) applications. Much of his work has been in the area of data mining, experimental design, response surface methodology, quality engineering, statistical process control and reliability. These areas are heavily involved with statistical tools, such as statistical modeling, Bayesian inference, optimal design theory, optimization, and time series. He has published over 100 papers in a wide variety of journals, including *Technometrics, Journal of the Royal Statistical Society, Journal of Quality Technology, IEEE Transaction on Reliability*, and *Statistica Sinica*. Dr. Lin has received the 2015 Shewhart Medal from the American Society for Quality, the Outstanding Presentation Award (SPES, ASA), and the Allen Keally Outstanding Teaching Award (CBA,University of Tennessee). He is an elected member of the International Statistical Institute (ISI), elected Fellow of the American Statistical Association (ASA), senior member of the American Society for Quality (ASQ), fellow of the Royal Statistical Society (RSS), and lifetime member of the International Chinese Statistical Association (ICSA). He is also the recipient of the 2004 Faculty Scholar Medal Award at Penn State University.

Chia-Jung Chang is currently an Assistant Professor in Harold and Inge Marcus Department of Industrial and Manufacturing Engineering at Pennsylvania State University. She received a Ph.D. degree in Industrial Engineering and M.S. degree in Statistics at Georgia Institute of Technology, and her B.S. and M.S. degrees in Industrial Engineering at National Tsing-Hua University in Taiwan. Dr. Chang's research interests focus on the fusion of advanced statistics, domain knowledge, and control theory to develop methodologies for modeling, monitoring, diagnosis, and improvement for complex systems in data rich environments. She has served as the Associate Editor of *Asian Journal of Industrial and Systems Engineering* since 2012. She is a member of the Institute of Industrial Engineering (IIE) and Institute of Operations Research and Management (INFORMS), American Statistical Association (ASA), and American Society of Mechanical Engineers (ASME).

References

Albrecht, M. C., C. J. Nachtsheim, T. A. Albrecht, and R. D. Cook 2013. Experimental Design for Engineering Dimensional Analysis. *Technometrics* 55(3):257–270; with Rejoinder 292–295.

Anderson, J. D. 2004. *Introduction to flight* 5th ed. McGraw-Hill.

Buckingham, E. 1914. On physically similar systems; illustrations of the use of dimensional equations. *Physical Review* 4(4):345–376.

Cook, R. D., and S. Weisberg 1982. *Residuals and influence in regression*. London: Chapman & Hall.

Davis, T. P. 2013. Comment: Dimensional Analysis in Statistical Engineering. *Technometrics* 55(3):271–274.

Fang, K. T., and D. K. J. Lin 2003. Uniform experimental design and its applications in industry. In *Handbook of Statistics in Industry* eds. C. R. Rao and R. Khattree, Ch. 4., New York: North Holland.

Frey, D. 2013. Comments: dimensional analysis and experimentation as a catalyst to learning from data. *Technometrics* 55(3):271–274.

Harper, W. V., and S. K. Gupta 1983. Sensitivity/uncertainty analysis of a borehole scenario comparing Latin Hypercube Sampling and deterministic sensitivity approaches. Technical Report BMI/ONWI-516, Office of Nuclear Waste Isolation, Columbus, OH: Battelle Memorial Institute.

Jones, B. 2013. Comments: enhancing the search for compromise designs. *Technometrics* 55(3):278–280.

Joseph, V. R., E. Gul, and S. Ba 2015. Maximum projection designs for computer experiments. *Biometrika* 102(2):371–380.

Lekivetz, R., and B. Jones 2015. Fast flexible space-filling designs for nonrectangular Regions. *Quality and Reliability Engineering International* 31, 829–837.

Lin, D. K., C. Sharpe, and P. Winker 2010. Optimized U-type designs on flexible regions. *Computational Statistics and Data Analysis* 54, 1505–1515.

Lin, D. K. J., and W. Shen 2013. Comments: experimental design for engineering dimensional analysis. *Technometrics* 55(3):281–285.

McInerney, M. F. 1985. Computer-aided experiments with the damped harmonic oscillator. *American Journal of Physics* 53(10):991–996.

McKay, M. D., R. J. Beckman, and W. J. Conover 1979. Comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics* 21(2):239–245.

Morris, M. D., T. J. Mitchell, and D. Ylvisaker 1993. Bayesian design and analysis of computer experiments: use of derivatives in surface prediction. *Technometrics* 35(3).

Park, J.-S. 1994. Optimal Latin-hypercube designs for computer experiments. *Journal of Statistical Planning and Inference* 39(1):95–111.

Piepel, G. F. 2013. Comments: spurious correlation and other observations on experimental design for engineering dimensional analysis. *Technometrics* 55(3):286–289.

Plumlee, M., V. R. Joseph, and C. F. J. Wu 2013. Comments: alternative strategies for experimental design. *Technometrics* 55(3):289–292.

Qian, P. Z. G. 2012. Sliced Latin Hypercube Designs. *Journal of the American Statistical Association* 107(497):393–399.

Rayleigh, L. 1915. The principle of similitude. *Nature* 95(2373):202–203.

Santner, T. J., B. J. Williams, and W. I. Notz 2003. *The Design and Analysis of Computer Experiments*. Berlin: Springer.

Shen, W., T. Davis, D. K. J. Lin, and C. J. Nachtsheim 2014. Dimensional analysis and its applications in statistics. *Journal of Quality Technology* 46(3):185–198.

Williams, C. K. 1998. Prediction with Gaussian processes: From linear regression to linear prediction and beyond. *Learning in Graphical Models*, 599–612.

Worley, B. A. 1987. Deterministic uncertainty analysis. Technical Report ORNL-6428, National Technical Information Service, Springfield, VA.