Uluslararası Mühendislik Araştırma ve Geliştirme Dergisi
International Journal of Engineering Research and Development

UMAGD, (2019) 11(2), 551-559.

10.29137/umagd.497045

Cilt/Volume:11 Sayı/Issue:2 Haziran/June 2019

Araştırma Makalesi / Research Article

A Mixed Integer Linear Programming Model for Finding Optimum Operating Conditions of Experimental Design Variables Using Computer-Aided Optimal Experimental Designs

Akın Özdemir*[1]

[1]*Department of Industrial Engineering, Faculty of Engineering, Bayburt University, 69000 Bayburt, TURKEY*

Başvuru/Received: *14/11/2018* ***Kabul/Accepted***: *14/05/2019* ***Son Versiyon/Final Version***: *30/06/2019*

Abstract

Computer-aided optimal experimental designs are an effective quality improvement tool that provides insights of information under various quality engineering problems. In the literature, considerable attention has been focused on maximizing the determinant of the information matrix in order to generate optimal design points. However, minimizing the average prediction based on the *I*-optimality criterion is more useful than commonly used *D*-optimality criterion for a number of situations. In this paper, special experimental design situations are explored where both qualitative and quantitative input variables are considered for an irregular design space with the pre-specified number of design points and the first-order polynomial model. In addition, this paper lays out the algorithmic foundations for the proposed *D*- and *I*-optimality criteria embedded mixed integer linear programming models in order to obtain optimal operating conditions using the first-order response functions. Comparative studies are also conducted. The results show that the proposed model for the *D*-optimality is able to identify approximately 30% smaller standard deviation value than the traditional counterpart. The proposed model for the *I*-optimality also provides a feasible solution whereas there is no optimum solution for the traditional counterpart.

Key Words

"Quality by design, computer-aided design, optimum operating condition, mixed integer linear programming, optimization"

***Sorumlu Yazar:** akinozdemir@bayburt.edu.tr*

Nomenclature

$\hat{\mu}(\mathbf{x},\mathbf{z})$	: Response function of the process mean
$\hat{\mu}(\mathbf{x},\mathbf{z})-\tau$	: The process bias
n	: A total number of experimental design points
n_a	: Additional experimental design points
n_n	: Necessary experimental design points
LB	: Lower bound
LSL	: Lower specification limit
$\hat{\sigma}^2(\mathbf{x},\mathbf{z})$	: Response function of the process variance
$\hat{\sigma}(\mathbf{x},\mathbf{z})$	: Response function of the process standard deviation
s_u	: The u^{th} run for the process standard deviation where u=1, 2, …, n
s_u^2	: The u^{th} run for the process variance where u=1, 2, …, n
τ	: The desired target value
UB	: Upper bound
USL	: Upper specification limit
$(\mathbf{x},\mathbf{z})$	: A vector of both real-valued and qualitative input variables
x_i^*	: The i^{th} real-valued input variable where i=1, 2, …, l
Y	: Response variable
$\bar{y}_u$	: The u^{th} mean value where u=1, 2, …, n
$z_j^{q_v*}$	: The j^{th} indicator variable for the different levels of the v^{th} qualitative input variable where j=1, 2, …, m

1. INTRODUCTION

The design of experiments is an effective continuous quality improvement tool in developing new processes, providing insights of the variation of information under various conditions, and enhancing process performance, including the variance reduction. For a number of practical situations, optimal design criteria may allow the best experimental design schemes to be generated based on the decision maker’s choices while the experimental design region is subject to physical restrictions and constraints using both qualitative and quantitative factors in production processes.

Computer-aided optimal experimental designs has been in the quality engineering literature for many years. The first study of the field of computer-aided optimal experimental design was conducted for prediction purposes by Smith (1918). Later, Wald (1943) introduced the *D*-optimality criterion. This criterion defines maximizing the determinant of the information matrix. Then, Kiefer and Wolfowitz (1959) proposed computational methods in order to find optimal experimental design points. Along the same lines, John and Draper (1975) provided brief reviews of the procedures of *D*-optimal experimental designs. In addition, Cook and Nachtsheim (1980) reviewed the algorithms in order to generate *D*-optimal experimental designs. On the contrary, Box and Draper (1959) was offered the *I*-optimality criterion. This criterion focuses on the integrated variance function over a studied design region. Draper (1982), Borkowski (2003) and Allen and Tseng (2011) conducted the further studies in the context of the *I*-optimality criterion. Furthermore, Toro Diaz et al. (2012) and Myers et al. (2016) reviewed the existing studies and provided theoretical aspects of computer-aided optimal experimental designs.

For optimization of design variables, Vining and Myers (1990) studied one of the first research attempt using a dual response model for quantitative input variables. Further, Del Castillo and Montgomery (1993) improved the dual response model. On the other hand, Lin and Tu (1995) proposed a mean-squared error model in order to enhance an optimum solutions of design variables. Copeland and Nelson (1996) conducted the further improvement of the MSE model. Furthermore, Steinberg and Bursztyn (1998), Robinson et al. (2004), Park et al. (2006) and Arvidsson and Gremyr (2008) provided comprehensive reviews of optimization models of design variables. Recent studies were conducted by Ozdemir and Cho (2016, 2017), Lu et al. (2017) and Chatterjee et al. (2018).

This paper was organized as follows. Research motivation and scope are presented in Section 2. The proposed model development methodology for the *D*- and *I*-optimality criteria is presented in Section 3 with a detailed description of each step. Then, a numerical example is performed in Section 4. Finally, concluding remarks and future studies are drawn in Section 5.

2. RESEARCH MOTIVATION AND SCOPE

There are no standard response surface designs, such as the traditional central composite designs and Box-Behnken designs, which are not suitable to exactly fit in irregular experimental design spaces. In addition, it is believed that a mixed integer linear programming (MILP) model to solve optimal experimental designs embedded robust design problems has not been adequately addressed in the literature in finding robust solutions for both qualitative and quantitative input variables. In this paper, a MILP model is formulated to solve an optimization problem based on optimal experimental designs, which can efficiently provide optimal operating conditions in irregular design spaces. The main purpose of this paper is five-fold. One, the design space may be constrained due to the physical infeasibility, safety concerns based constraints, and resource limitations. Therefore, optimal designs are suitable alternatives to conduct experimental designs under constrained design spaces. In addition, the *D*-optimality criterion works well in order to estimate model parameters while maximizing the determinant of the information matrix. On the other hand, the *I*-optimality criterion addresses prediction variance to generate a measure of prediction performance. Two, there are some special knowledge for some practical situations, such as the potential application of convex hull about the process being studied that may offer a nonstandard model. In these situations, optimal designs are effective to estimate or predict nonstandard model parameters with minimum variance and without experimental bias. Three, it may be desired to reduce the number of design points due to resource limitations and cost considerations. Therefore, the number of design points can be reduced with the *D*- and *I*-optimality criteria. Fractional factorial designs may be considered in order to reduce the number of observations for quantitative input variables; however, they may not be effective for second-order or nonstandard models. Four, traditional experimental designs may not be longer effective while considering both quantitative and qualitative input variables, especially more than two coded levels of qualitative input variables. Thus, this limitation may be eliminated using optimal experimental designs. Five, this paper lays out the theoretical foundation for transmitting the *D*- and *I*-optimality criteria into a MILP model in order to obtain optimum operating conditions.

3. METHODOLOGY DEVELOPMENT

The proposed methodology development consists of three main phases: (1) Experimental design selection phase, (2) modelling phase, and (3) optimization phase.

3.1. Experimental Design Selection Phase

Traditional response surface designs are not appropriate while dealing with an irregular design region. In this paper, the *D*- and *I*-optimality criteria are the alternatives. The *D*-optimality criterion is based on the following notion as follows:

$$D-\text{optimality} = \max \ \det\left[\mathbf{X}'\mathbf{X}\right] \tag{1}$$

where $\left[\mathbf{X}'\mathbf{X}\right]$ is the information matrix. Additionally, the *D*-optimality criterion is the best alternative to estimate model parameters while maximizing the determinant of the information matrix. On the contrary, the *I*-optimality criterion is based on prediction variance to produce a measure of prediction performance through an averaging process over the region of interest *R* as follows:

$$\begin{aligned} &\min_{\zeta} \int_R V[\mathbf{x},\mathbf{z}]d[\mathbf{x},\mathbf{z}] / \int_R d[\mathbf{x},\mathbf{z}] \\ &\Rightarrow \min_{\zeta} \int_R n[\mathbf{x},\mathbf{z}]^{(m)'} \left(\mathbf{X}'\mathbf{X}\right)^{-1} [\mathbf{x},\mathbf{z}]^{(m)} d[\mathbf{x},\mathbf{z}] / \int_R d[\mathbf{x},\mathbf{z}] \end{aligned} \tag{2}$$

where $V[\mathbf{x},\mathbf{z}] = n[\mathbf{x},\mathbf{z}]^{(m)'} \left(\mathbf{X}'\mathbf{X}\right)^{-1} [\mathbf{x},\mathbf{z}]^{(m)}$.

3.2. Modelling Phase

In general, a low-order polynomial model is a suitable choice for many cases (Myers et al., 2016). In this paper, a first-order polynomial model is used. A general formula of the first-order model is denoted for both qualitative and quantitative input variables as follows:

$$\hat{y} = \beta_0 + \sum_{i=1}^{l} \beta_i x_i + \sum_{j=1}^{m} z_j^{q_v} \varphi_j^{q_v} \tag{3}$$

where β_0 is the intercept, and β_i and φ_j represent regression coefficients of quantitative and qualitative input variables, respectively. Along the same lines, the fitted mean response function is found by

$$\begin{aligned} &\hat{\mu}\left(\mathbf{x},\mathbf{z}\right) = \mathbf{X}\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\bar{\mathbf{y}} \\ &\text{where } \mathbf{X} = \begin{pmatrix} 1 & x_{11} & x_{21} & \cdots & x_{i1} & z_{11}^{q_1} & z_{21}^{q_1} & \ldots & z_{m1}^{q_1} \ldots & z_{11}^{q_v} & z_{21}^{q_v} & \ldots & z_{m1}^{q_v} \\ 1 & x_{12} & x_{22} & \cdots & x_{i2} & z_{12}^{q_1} & z_{22}^{q_1} & \ldots & z_{i2}^{q_1} \ldots & z_{12}^{q_v} & z_{22}^{q_v} & \ldots & z_{i2}^{q_v} \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 1 & x_{1l} & x_{2l} & \cdots & x_{il} & z_{1m}^{q_1} & z_{2m}^{q_1} & \ldots & z_{im}^{q_1} \ldots & z_{1m}^{q_v} & z_{2m}^{q_v} & \ldots & z_{im}^{q_v} \end{pmatrix} \text{ and } \bar{\mathbf{y}} = \left[y_1,\ y_2,\ \ldots,\ y_n\right]' \end{aligned} \tag{4}$$

In addition, the fitted standard deviation and variance response functions are

$$\hat{\sigma}\left(\mathbf{x},\mathbf{z}\right) = \mathbf{X}\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{s} \text{ where } \mathbf{s} = \left[s_1,\ s_2,\ \ldots,\ s_n\right]' \tag{5}$$

$$\hat{\sigma}^2\left(\mathbf{x},\mathbf{z}\right) = \mathbf{X}\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{s}^2 \text{ where } \mathbf{s}^2 = \left[s_1^2,\ s_2^2,\ \ldots, s_n^2\right]' \tag{6}$$

3.3. Optimization Phase

One of the important goals in quality engineering problems is to minimize the process variance. Therefore, the objective function of the proposed MILP model is to minimize the process variance for quality improvement. The proposed model is subject to five constraints. The first constraint is related to the process mean requirements in order to provide the customer's satisfaction for the process. The second constraint is non-negative variance to prevent meaningless results. The third constraint is defined an indicator variable for levels of qualitative input variables. The design region is irregular; therefore, the fourth constraint is based on the non-standard design region. The last constraint is related to boundary requirements for both qualitative and quantitative input variables. Table 1 shows the proposed MILP model.

Table 1. Proposed MILP model

Objective	$\min\ \hat{\sigma}^2(\mathbf{x},\mathbf{z})$
Subject to	Constraints: (1) Process mean requirements: $LSL \le \hat{\mu}(\mathbf{x},\mathbf{z}) \le USL$ (2) Non-negative variance: $\hat{\sigma}^2(\mathbf{x},\mathbf{z}) \ge 0$ (3) Indicator variable: $\sum_{j=1}^{m} z_j^{q_v} \le 1$ for $j = 1,\ 2,\ ...,\ m$ (4) Non-standard design region: $h_u(\mathbf{x},\mathbf{z}) \le L_u$ for $u = 1,\ 2,\ ...,\ k$ (5) Bounds for both real-valued and qualitative input variables: $LB_i \le x_i \le UB_i$ for $x_i \in R$ and $i = 1,\ 2,\ ...,\ l$ $z_j^{q_v} \in \{0,1\}$ and $j = 1,\ 2,\ ...,\ m$
Approach	A branch-and-cut solution technique
Find	Optimal robust solution for x_i^* and $z_j^{q_v*}$

In Table 1, the proposed MILP model can be solved over the polyhedra and an optimal operating condition exists at the extreme point of the solution space if the optimization problem is feasible. In the ideal case, each extreme point is an integer point, so an integer linear programming model can be solved. Therefore, the geometric aspect of compact sets plays a significant role to obtain robust solutions to proposed MILP model. It is then defined as the convex hull of the feasible experimental design space, $conv(X)$, as follows:

$$conv(X) = \left\{ \mathbf{x},\mathbf{z} \left| \begin{array}{l} [\mathbf{x},\mathbf{z}] = \sum_{i=1}^{l} \lambda_i x_i + \sum_{j=1}^{m} \pi_j z_j^{q_v},\ \sum_{i=1}^{l} \lambda_i + \sum_{j=1}^{m} \pi_j = 1,\ \sum_{i=1}^{l} \lambda_i \ge 0,\ \sum_{j=1}^{m} \pi_j \ge 0 \\ \text{for } i = 1,\ 2,\ ...,\ l \text{ and } j = 1,\ 2,\ ...,\ m \end{array} \right. \right\} \tag{7}$$

In addition, two proposed propositions are expressed to the proposed MILP optimization as follows:

Proposition 1. $conv(X)$ is a polyhedron for the proposed MILP optimization model.

Proof. The fitted response functions of the mean and variance in the proposed MILP optimization are linear due to the first-order polynomial model. In addition, the extreme points are at the corner points because all constraints in the optimization model are linear. Therefore, the convex hull should be a polyhedron.

Proposition 2. The extreme points of $conv(X)$ all lie in a bounded polyhedral set, X.

Proof. The extreme points are not interior points of any line segment in the extreme points of the convex set. In addition, the extreme points of the convex hull are the subset of the feasible experimental design space. Therefore, the extreme points all lie in X.

Further, the proposed optimization model is also formulated with the convex hull of the problem in the context of the MILP framework as follows:

$$\text{MILP} = \left\{ \min\ \hat{\sigma}^2(\mathbf{x},\mathbf{z}) \left| \begin{array}{l} LSL \le \hat{\mu}(\mathbf{x},\mathbf{z}) \le USL, \hat{\sigma}^2(\mathbf{x},\mathbf{z}) \ge 0, \\ \sum_{j=1}^{m} z_j^{q_v} \le 1, h_u(\mathbf{x},\mathbf{z}) \le L_u \text{ and } \mathbf{x},\mathbf{z} \in X \end{array} \right. \right\} \tag{8}$$

where $X = \left\{ \mathbf{x}, \mathbf{z} \middle| LB_i \le x_i \le UB_i, z_j^{q_v} \in \{0,1\}, i = 1,\ 2,\ ...,\ l,\ j = 1,\ 2,\ ...,\ m, \text{ and } u = 1,\ 2,\ ...,\ k \right\}$. The MILP can be then replaced by the equivalent LP (linear programming) as follows:

$$\text{LP} = \left\{ \min\ \hat{\sigma}^2(\mathbf{x}, \mathbf{z}) \middle| \mathbf{x}, \mathbf{z} \in conv(X) \right\} \tag{9}$$

Note that the ideal solution of $conv(X)$ has the property, which is expressed as $X \subseteq conv(X) \subseteq P$ for all polyhedral (*P*) formulations. In addition, the ideal solution of $conv(X)$ is solved efficiently by using a branch-and-cut (BC) algorithm.

4. NUMERICAL EXAMPLE

Reconsider the experiment, conducted by Myers et al. (2016), in which the amount of extraction is modelled a function of type of solvent (with solvent A, B, and C). Temperature (x_1) and time (x_2) are also continuous-valued quantitative input variables taken as -1 to +1 on the coded variable scale. The amount of extraction with higher temperatures may give higher yields due to increased solubility. Therefore, it is also desired to obtain higher temperatures for solvent A in this paper. In terms of coded input variables, this leads to constraints on the design variables as follows:

$$\begin{aligned} & 1 \le x_1 + z_1^{q_1} \le 2 \\ & \text{where } z_1^{q_1} = \begin{cases} 1 \text{ if solvent A is the level} \\ 0 \text{ elsewhere} \end{cases} \\ & z_2^{q_1} = \begin{cases} 1 \text{ if solvent B is the level} \\ 0 \text{ elsewhere} \end{cases} \end{aligned} \tag{10}$$

In addition, the desired target value for the amount of extraction is 10 grams, where the allowable lower and upper bounds are 9.5 and 10.5 grams, respectively. Furthermore, the regression model assumed is first-order in temperature, time and the qualitative input variable (solvent) for MILP optimization models. The regression model is also written formally as follows:

$$\hat{y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varphi_1^{q_1} z_1^{q_1} + \varphi_2^{q_1} z_2^{q_1} \tag{11}$$

In this section, the numerical examples are provided using the proposed model development methodology in this paper for the *D*- and *I*-optimality criteria, respectively. Note that the computer, which has 2.3 GHz Intel Core i5 with 8 GB DDR4 memory, is used to run programs with JMP and AMPL software. The desired design point runs are 11 where $p = 5$ and $n_a = 6$. In addition, Table 2 shows global solutions of design points for the *D*- and *I*-optimality criteria.

Table 2. Globally *D*- and *I*-optimal designs

Design Criterion	*D*-optimality			*I*-optimality		
Input Variables	x_1	x_2	z^{q_1}	x_1	x_2	z^{q_1}
Design Matrix (**DM**)	0.00	-1.00	A	1.00	1.00	C
	-1.00	1.00	B	-1.00	1.00	B
	-1.00	-1.00	C	-1.00	-1.00	C
	1.00	1.00	A	0.00	-1.00	A
	1.00	-1.00	C	0.00	1.00	A
	1.00	1.00	C	1.00	1.00	B
	1.00	-1.00	A	1.00	-1.00	B
	1.00	-1.00	B	0.01	0.00	A
	1.00	1.00	B	-1.00	-1.00	B
	-1.00	1.00	C	-1.00	1.00	C
	-1.00	-1.00	B	1.00	-1.00	C
Number of iterations	100000			100000		
D-efficiency	94.164			91.569		
Average variance of prediction	0.370			0.353		
Running time (seconds)	62.22			81.48		
Solution type	Global			Global		

In Table 2, the *D*-efficiency is 94% for the *D*-optimality criterion and the *D*-efficiency of the *I*-optimality criterion is about 92%. Hence, the *D*-optimality criterion generates a better solution in order to estimate the model coefficients. In addition, the average variance of prediction for the *I*-optimality is 0.353 compared to 0.370 for the *D*-optimality. This explains confidence intervals for prediction is almost 5% longer for the *D*-optimality. Therefore, the *I*-optimality criterion is preferable for predicting the response in the irregular design space. In addition, the *D*- and *I*-optimal experimental designs and the data are shown in Tables 3 and 4, respectively.

Table 3. *D*-Optimal experimental design and data

Run	Coded levels			Observed data				$\bar{y}_u$	s_u	s_u^2
	x_1	x_2	z^{q_1}	y_{u1}	y_{u2}	y_{u3}	y_{u4}			
1	0.00	-1.00	A	8.70	11.54	10.07	7.90	9.55	1.600	2.561
2	-1.00	1.00	B	13.97	14.22	6.41	9.61	11.05	3.750	14.063
3	-1.00	-1.00	C	10.13	9.24	10.78	13.80	10.99	1.978	3.913
4	1.00	1.00	A	13.08	8.80	11.38	9.21	10.62	1.996	3.983
5	1.00	-1.00	C	10.21	8.13	9.25	11.50	9.77	1.432	2.0519
6	1.00	1.00	C	12.54	9.30	7.79	8.68	9.58	2.072	4.293
7	1.00	-1.00	A	9.20	8.23	11.09	9.56	9.52	1.188	1.411
8	1.00	-1.00	B	8.96	7.63	7.36	7.87	7.95	0.701	0.491
9	1.00	1.00	B	11.53	7.36	5.97	9.86	8.68	2.488	6.193
10	-1.00	1.00	C	10.22	13.55	11.81	8.97	11.14	1.986	3.944
11	-1.00	-1.00	B	9.09	11.23	10.59	10.08	10.25	0.902	0.813

Table 4. *I*-optimal experimental design and data

Run	Coded levels			Observed data				$\bar{y}_u$	s_u	s_u^2
	x_1	x_2	z^{q_1}	y_{u1}	y_{u2}	y_{u3}	y_{u4}			
1	1.00	1.00	C	10.68	10.89	11.96	10.92	11.11	0.576	0.332
2	-1.00	1.00	B	8.91	10.00	11.50	8.66	9.77	1.290	1.663
3	-1.00	-1.00	C	7.33	10.62	8.65	8.91	8.88	1.350	1.822
4	0.00	-1.00	A	8.78	10.81	8.18	5.83	8.40	2.051	4.205
5	0.00	1.00	A	5.55	11.17	12.28	9.38	9.60	2.950	8.701
6	1.00	1.00	B	6.82	12.19	12.35	13.59	11.24	3.013	9.079
7	1.00	-1.00	B	7.77	10.11	9.64	15.94	10.87	3.533	12.483
8	0.01	0.00	A	13.35	7.38	10.31	4.96	9.00	3.635	13.210
9	-1.00	-1.00	B	9.96	12.07	9.40	8.17	9.90	1.627	2.649
10	-1.00	1.00	C	8.65	8.90	8.06	10.48	9.02	1.033	1.067
11	1.00	-1.00	C	11.48	9.81	11.60	8.89	10.45	1.319	1.741

Tables 5 and 6 show Analysis of Variance (ANOVA) for the *D*- and *I*-optimal experimental designs.

Table 5. ANOVA for the *D*-optimal experimental design

Source	Degrees of Freedom	Sum of Squares	Mean Square	F Ratio
Model	4	8.542078	2.13552	7.0383
Error	6	1.820486	0.30341	P-value
C. Total	10	10.362564		**0.0188***

**R^2=0.824

Table 6. ANOVA for the *I*-optimal experimental design

Source	Degrees of Freedom	Sum of Squares	Mean Square	F Ratio
Model	4	8.7372425	2.18431	18.5122
Error	6	0.7079575	0.11799	P-value
C. Total	10	9.4452000		**0.0016***

**R^2=0.925

The values of the determination coefficient, R^2, are 0.824 and 0.925, respectively. Therefore, the models may predict 82.40% and 92.50% of the variability for the *D*- and *I*-optimal responses, respectively. The analysis of variance (ANOVA) is shown in Table 5, and the value of an F-test is 7.0383 where the *p*-value is 0.0188. In addition, the value of an F-test is found 18.5122 for the *I*-optimal experimental design where the *p*-value is obtained 0.0016. Therefore, it is concluded that the regression models are highly significant.

No *VIF* is larger than 10, so there is no severe multicollinearity for the *D*-optimal design in Table 3. The fitted response functions for the mean, standard deviation, and variance are also obtained using JMP software as follows:

$$\hat{\mu}(\mathbf{x},\mathbf{z}) = 10.14 - 0.84x_1 + 0.33x_2 + 0.43z_1^{q_1} - 0.66z_2^{q_1} \tag{12}$$

$$\hat{\sigma}(\mathbf{x},\mathbf{z}) = 1.93 - 0.27x_1 + 0.59x_2 + 0.04z_1^{q_1} + 0.03z_2^{q_1} \tag{13}$$

$$\hat{\sigma}^2(\mathbf{x},\mathbf{z}) = 4.41 - 1.29x_1 + 2.32x_2 - 0.12z_1^{q_1} + 0.98z_2^{q_1} \tag{14}$$

In addition, no *VIF* is larger than 10, so there is no severe multicollinearity for the *I*-optimal design in Table 4. The fitted response functions for the mean, standard deviation, and variance are also obtained using JMP software as follows:

$$\hat{\mu}(\mathbf{x},\mathbf{z}) = 9.76 + 0.76x_1 + 0.22x_2 - 0.77z_1^{q_1} + 0.68z_2^{q_1} \tag{15}$$

$$\hat{\sigma}(\mathbf{x},\mathbf{z}) = 2.10 + 0.39x_1 - 0.10x_2 + 0.77z_1^{q_1} + 0.26z_2^{q_1} \tag{16}$$

$$\hat{\sigma}^2(\mathbf{x},\mathbf{z}) = 5.46 + 2.06x_1 - 0.21x_2 + 3.22z_1^{q_1} + 1.00z_2^{q_1} \tag{17}$$

The proposed optimization models are also formulated with the convex hull of the *D*- and *I*-optimal design problems as follows:

$$\text{MILP}_D = \left\{ \min\ 4.41 - 1.29x_1 + 2.32x_2 - 0.12z_1^{q_1} + 0.98z_2^{q_1} \;\middle|\; \begin{array}{l} 9.50 \le 10.14 - 0.84x_1 + 0.33x_2 \\ +0.43z_1^{q_1} - 0.66z_2^{q_1} \le 10.50, \\ 4.41 - 1.29x_1 + 2.32x_2 \\ -0.12z_1^{q_1} + 0.98z_2^{q_1} \ge 0, \\ 1 \le x_1 + z_1^{q_1} \le 2,\ z_1^{q_1} + z_2^{q_1} \le 1 \\ \text{and } \mathbf{x},\mathbf{z} \in X \end{array} \right\} \tag{18}$$

$$\text{MILP}_I = \left\{ \min\ 5.46 + 2.06x_1 - 0.21x_2 + 3.22z_1^{q_1} + 1.00z_2^{q_1} \;\middle|\; \begin{array}{l} 9.50 \le 9.76 + 0.76x_1 + 0.22x_2 \\ -0.77z_1^{q_1} + 0.68z_2^{q_1} \le 10.50, \\ 5.46 + 2.06x_1 - 0.21x_2 \\ +3.22z_1^{q_1} + 1.00z_2^{q_1} \ge 0, \\ 1 \le x_1 + z_1^{q_1} \le 2,\ z_1^{q_1} + z_2^{q_1} \le 1 \\ \text{and } \mathbf{x},\mathbf{z} \in X \end{array} \right\} \tag{19}$$

where MILP_D and MILP_I represent *D*- and *I*-optimal designs embedded proposed MILP models, respectively and $X = \left\{ \mathbf{x},\mathbf{z} \middle| -1 \le x_i \le 1, z_j^{q_1} \in \{0,1\},\ i = 1,\ 2,\ \text{and}\ j = 1,\ 2 \right\}$. In addition, the results of the proposed model and traditional counterpart are summarized in Table 7.

Table 7. The results of the proposed model and traditional counterpart for MILP models

Model	Design	Optimal solution				Bias^*	$\hat{\sigma}^*(\mathbf{x},\mathbf{z})$	Solvent
		x_1^*	x_2^*	$z_1^{q_1 *}$	$z_2^{q_1 *}$			
VM	*D*-optimality	0.286	-1.000	1.000	0.000	0.000	1.303	A
VM	*I*-optimality	-	-	-	-	-	-	-
Proposed	*D*-optimality	0.881	-1.000	1.000	0.000	0.500	0.913	A
Proposed	*I*-optimality	1.000	-0.091	0.000	0.000	0.500	2.746	C

In Table 7, the proposed models are superior to traditional counterparts when attaining slight bias. For example, the proposed model for the *D*-optimality achieve approximately 30% more variability reduction than the *VM* (Vining and Myers, 1990) model. In addition, the *VM* model does not give an optimal operating condition for the *I*-optimality due to the strict equality constraint.

Furthermore, some traditional models in the literature may not be appropriate for MILP models. For example, Lin and Tu (1995) model is not appropriate due to $\left(\hat{\mu}(\mathbf{x},\mathbf{z})-\mu_{\tau}\right)^2$, so this model is not used for the comparison studies in this section.

5. CONCLUSIONS

In this paper, a MILP model is proposed using computer-aided optimal experimental designs in order to obtain an optimum operating condition of both quantitative and qualitative input variables for quality improvement. For the design selection phase, the *D*- and *I*-optimality criteria are preferred over the traditional response designs while considering a non-standard design region. For the estimation purpose, the *D*-optimality criterion is selected. On the other hand, the *I*-optimality criterion is preferred over the *D*-optimality criterion while the prediction performance is priority. For the modeling phase, the first-order polynomial model is used. The mean, the standard deviation and the variance response functions are also calculated for the optimization phase. An application of the convex hull is illustrated in the optimization phase. For the solution procedure, the branch-and-cut method is applied in order to find an optimum operating condition of both qualitative and quantitative input variables. The proposed optimization model achieves more variance reduction than the traditional counterparts. Finally, the proposed methodology is an effective technique to improve the production processes for quality improvement.

For future research studies, multiple quality characteristics could be implement the proposed methodology for multiple response optimization problems. Another research study could be to consider a second-order polynomial model for quadratic effects of quantitative input variables. Then, a general polynomial approximation model could be defined while considering quantitative input variables of a second-order model and qualitative input variables. For the optimization phase, a mixed integer nonlinear programming model could be proposed due to quadratic effects of quantitative input variables.

REFERENCES

Allen, T. T., & Tseng, S. H. (2011). Variance plus bias optimal response surface designs with qualitative factors applied to stem choice modeling. *Quality and Reliability Engineering International*, *27*(8), 1199-1210.

Arvidsson, M., & Gremyr, I. (2008). Principles of robust design methodology. *Quality and Reliability Engineering International*, *24*(1), 23-35.

Borkowski, J. J. (2003). A comparison of prediction variance criteria for response surface designs. *Journal of Quality Technology*, *35*(1), 70-77.

Box, G. E., & Draper, N. R. (1959). A basis for the selection of a response surface design. *Journal of the American Statistical Association*, *54*(287), 622-654.

Chatterjee, K., Drosou, K., Georgiou, S. D., & Koukouvinos, C. (2018). Response modelling approach to robust parameter design methodology using supersaturated designs. *Journal of Quality Technology*, *50*(1), 66-75.

Cook, R. D., & Nachtrheim, C. J. (1980). A comparison of algorithms for constructing exact D-optimal designs. *Technometrics*, *22*(3), 315-324.

Copeland, K. A., & Nelson, P. R. (1996). Dual response optimization via direct function minimization. *Journal of Quality Technology*, *28*(3), 331-336.

Del Castillo, E., & Montgomery, D. C. (1993). A nonlinear programming solution to the dual response problem. *Journal of Quality Technology*, *25*(3), 199-204.

Draper, N. R. (1982). Center points in second—order response surface designs. *Technometrics*, *24*(2), 127-133.

John, R. S., & Draper, N. R. (1975). D-optimality for regression designs: a review. *Technometrics*, *17*(1), 15-23.

Kiefer, J., & Wolfowitz, J. (1959). Optimum designs in regression problems. *The Annals of Mathematical Statistics*, *30*(2), 271-294.

Lin, D. K., & Tu, W. (1995). Dual response surface optimization. *Journal of Quality Technology*, *27*(1), 34-39.

Lu, Y., Wang, S., Yan, C., & Huang, Z. (2017). Robust optimal design of renewable energy system in nearly/net zero energy buildings under uncertainties. *Applied Energy*, *187*, 62-71.

Myers, R. H., Montgomery, D. C., & Anderson-Cook, C. M. (2016). Response Surface Methodology: Process and Product Optimization Using Designed Experiments (Wiley Series in Probability and Statistics).

Ozdemir, A., & Cho, B. R. (2016). A nonlinear integer programming approach to solving the robust parameter design optimization problem. *Quality and Reliability Engineering International*, *32*(8), 2859-2870.

Ozdemir, A., & Cho, B. R. (2017). Response surface-based robust parameter design optimization with both qualitative and quantitative variables. *Engineering Optimization*, *49*(10), 1796-1812.

Park, G. J., Lee, T. H., Lee, K. H., & Hwang, K. H. (2006). Robust design: an overview. *AIAA journal*, *44*(1), 181-191.

Robinson, T. J., Borror, C. M., & Myers, R. H. (2004). Robust parameter design: a review. *Quality and Reliability Engineering International*, *20*(1), 81-101.

SAS Institute, 2013. *Using JMP 11*. SAS Institute. Cary, NC USA.

Smith, K. (1918). On the standard deviations of adjusted and interpolated values of an observed polynomial function and its constants and the guidance they give towards a proper choice of the distribution of observations. *Biometrika*, *12*(1/2), 1-85.

Steinberg, D. M., & Bursztyn, D. (1998). Noise factors, dispersion effects, and robust design. *Statistica Sinica*, *8*(1), 67-85.

Toro Díaz, H. H., Chan, H. L., & Cho, B. R. (2012). Optimally designing experiments under non-standard experimental situations. *International Journal of Experimental Design and Process Optimisation*, *3*(2), 133-158.

Vining, G. G., & Myers, R. H. (1990). Combining Taguchi and response surface philosophies: a dual response approach. *Journal of Quality Technology*, *22*(1), 38-45.

Wald, A. (1943). On the efficient design of statistical investigations. *The Annals of Mathematical Statistics*, *14*(2), 134-140.