Int J Adv Manuf Technol (2009) 41:631–641
DOI 10.1007/s00170-008-1509-2

ORIGINAL ARTICLE

Another view of dual response surface modeling and optimization in robust parameter design

A. B. Shaibu · Byung Rae Cho

Received: 30 January 2008 / Accepted: 1 April 2008 / Published online: 27 May 2008

Abstract Robust parameter design (RPD) based on the concept of building quality into a design has received much attention from researchers and practitioners for years, and a number of methodologies have been studied in the research community. There have been many attempts to integrate RPD principles with well-established statistical techniques, such as response surface methodology, in order to model the response directly as a function of control factors. In this paper, we reinvestigate the dual response approach based on quadratic models Vinning and Myers (J Qual Technol 22:38–45), which is often referred to in the RPD literature and demonstrate that higher-order polynomial models may be more effective in finding better RPD solutions than the commonly-used quadratic model. We also propose optimization models for each of the three classes of quality characteristics (i.e., nominal-the-best, larger-the-better, and smaller-the-better). The optimal solutions obtained using the proposed models are compared with the solutions obtained using the RPD techniques in the current literature.

Keywords Higher-order polynomial models · L-type · model selection · N-type · Optimal solutions · Optimization model · Prediction ability · S-type

1 Introduction

The robust parameter design (RPD) methodology focuses on building quality into the design of products and processes through the determination of the optimum operating conditions in order to minimize performance variability and deviation from the target value of interest. Since it was first introduced by Taguchi [1,2], this approach has come under serious criticism due to the statistical analysis methods and optimization approaches utilized. Taguchi advocates minimizing signal-to-noise ratios to determine the best overall combination of design parameter settings and identifying adjustment factors which are used to tune a mean to a desired target. However, Nair and Shoemaker [3] argue that by simply collapsing the experimental data into signal-to-noise ratios, much of the information concerning the system's behavior is lost. Additionally, Taguchi gives no real justification for the use of these ratios, and the details surrounding the use of adjustment factors to achieve the desired target value are sketchy at best. To address these issues, there have been several attempts in the literature to improve the analysis and optimization phases of the RPD methodology. These improvement efforts have focused on the dual response surface approach, which was first incorporated into the RPD methodology by Vining and Myers [4]. The dual response approach is an extension of the work based on ridge analysis by Myers and Carter [5]. Ridge analysis has been studied by numerous authors including Hoerl [6], Draper [7], Myers [8], Box and Draper [9], and also Khuri and Cornell [10].

The dual response approach utilizes response surfaces in modeling process relationships by separately estimating the response functions for the process mean and variance of the system under investigation. Then, based on the optimization strategy chosen, these functions are optimized simultaneously over the region of interest to determine the system's optimum operating conditions. For nominal-the-best type (*N*-Type) quality characteristics, Vining and Myers [4] proposed minimizing the variance while main-

A. B. Shaibu (✉) · B. R. Cho
Advanced Quality Engineering Laboratory,
Department of Industrial Engineering, Clemson University,
Clemson, SC 29634, USA
e-mail: ashaibu@clemson.edu

B. R. Cho
e-mail: bcho@ces.clemson.edu

taining the mean on the desired target, thus referring to the variance as the primary response and the mean as the secondary response. For the larger-the-better type (*L*-Type) and smaller-the-better type (*S*-Type) quality characteristics, they proposed setting the mean as the primary response to be optimized, while setting the variance as the secondary response to be maintained on a target. Umland and Smith [11] and Myers and Carter [5] used the Lagrange multipliers as a solution method in solving those problems.

Del Castillo and Montgomery [12] observed that the Vining and Myers model [4] does not always yield global optimal RPD solutions, and proposed the generalized reduced gradient method with inequality constraints as an optimization method. Cho [13] and Lin and Tu [14] proposed minimizing the mean squared error which allows some bias in the primary response with substantial reduction in variability. Lin and Tu [14] demonstrated superiority in their methodology over the proposals of Vining and Myers [4], and Del Castillo and Montgomery [12]. As an extension, Copeland and Nelson [15] proposed minimizing the standard deviation subject to a constraint that bounds the bias in the definition of the mean squared error (i.e., $(\widehat{\mu}-\tau)^2 \leq \Delta^2$, where $\widehat{\mu}$ is an estimate of the target mean τ, and Δ is a desired upper bound for the bias). Several other approaches to solving the dual response problem have been proposed. For example, Kim and Lin [16] proposed a fuzzy optimization methodology, and Del Castillo et al. [17] and Fan [18] proposed different computational methods for global optimization.

Kim and Cho [19], and Tang and Xu [20] proposed the use of goal programming, and Köskoy and Dogamaksoy [21] suggested treating the secondary response as another primary response and generating Pareto optimal solutions. Further, Lam and Tang [22] proposed the use of a graphical approach, with compromise programming to generate a set of weakly Pareto optimal solutions to the dual response problem as an aid to studying trade-offs involved in selecting optimal design settings.

Numerous authors have considered the subject of designed experiments in the context of robust parameter design (RPD). Recent developments include Vining et al. [23], who considered industrial experiments involving hard-to-change factors in the context of split-plot design and modified the standard central composite design to apply to such cases. Kowalski et al. [24] proposed a further modification of the proposal by Vining et al. [23] who integrated the dual response methodology with the split-plot structure. More recently, Kunert et al. [25] used the experiments with the product array and combined array and determined that the product array was better at finding the effect on the variance than the combined array. They argue that the reason for their observation may be attributed to the fact that the effect on the variance could not be attributed to low-order interactions. Del Castillo et al. [26] proposed a criterion for designs used in RPD based on the mean square error of fitted models and the range of noise factors. Although it is not the focus of this paper, it is worth mentioning that RPD is being studied in the setting of generalized linear models. Some references of such studies include Nelder and Lee [27], Lee and Nelder [28, 29], Engel and Huele [30], Myers et al. [31], and Robinson et al. [32].

In what follows, we present the abbreviations and notation of this paper and state our research objective. The proposed model development is presented next followed by a numerical example. Finally, the paper ends with concluding remarks.

1.1 Abbreviatiosns and notation

The abbreviations and notations we use in this paper are as follows:

VM	Method of Vining and Myers [4]
LT	Method of Lin and Tu [14]
SC	Proposed method by Shaibu and Cho shown in this paper
T	Target value for mean response
T_S	Target value for standard deviation
S-type	The smaller-the-better type quality characteristic
L-type	The larger-the-better type quality characteristic
N-Type	The nominal-the-best type quality characteristic
$\mathbf{x}$	Vector of the levels of controllable factors
$\mathbf{y}$	Vector of observed responses
$\bar{y}$	Mean of observed responses
$\widehat{\mu}_{SC}(x)$	Estimated mean response surface obtained using the proposed model
$\mu_{MV}(x)$	Estimated mean response surface function using the Vining and Myers approach [4]
$\widehat{\mu}_{SC}(x)$	Estimated mean response surface function using the Lin and Tu approach [14]
$\widehat{\sigma}_{SC}(x)$	Estimated standard deviation response surface function using the proposed model
$\widehat{\sigma}_{MV}(x)$	Estimated standard deviation response surface function using the Vining and Myers approach [4]
$\widehat{\sigma}_{LT}(x)$	Estimated standard deviation response surface function using the Lin and Tu approach [14]
Ω	Feasible or experimental region of interest.

2 Research motivation

There are four main steps in RPD methodology:

1. Design an experiment from which a set of data on a quality characteristic of interest is generated
2. Model the mean and variability responses using the data in Step (1)

3. Formulate models to optimize the mean and variability response functions
4. Obtain solutions to the optimization models

We observe that the second step relies on the first, since a good experimental design yields reliable data, which is essential in obtaining reliable models for the mean and variability response functions. Similarly, obtaining realistic solutions from the optimization models depends on whether the constituents of the optimization models (i.e., the mean and variability response functions) represent the mean and variability of the quality characteristics as accurately as possible. In the RPD literature, much work has been done to ensure sound experimental designs under various conditions, but not nearly as much has been done with regards to response surface modeling. However, the importance of response surface models has been highlighted by LT who reiterated the suggestion by VM that “The fitness or the prediction ability of the mean and variability models is an extremely important consideration when optimizing a dual response problem.” This suggestion cannot be overemphasized since different models of the mean and variability of the same process or product characteristic can be used in solving the same optimization model, all achieving the objective of the optimization model (i.e., minimizing or maximizing the objective function), but with different optimal settings. We believe that better RPD solutions can be found by obtaining accurate response surface models in terms of prediction. Therefore, the second step above deserves a great deal of attention in order to make sure that the most accurate models are obtained before proceeding to the optimization step. Achieving accurate response surface models will serve to reduce, as much as possible, the disparity between solutions to optimization models and the results obtained by actually applying those solutions to the processes or systems of interest. To this end, in this paper, we select models based on statistical model selection techniques, and directly compare the models obtained to previous models. We will also compare the models by applying them to various optimization problems proposed in the context of RPD.

3 Proposed model development

The model development procedure we propose consists of three steps, namely the experimental phase, the model selection phase, and the optimization phase. We describe each of the phases in what follows.

3.1 Experimental phase

In this phase, an experimental design (e.g., full or fractional factorial designs, central composite designs, etc) is selected and the response of interest is measured under the selected design. That is, for n factors, the measurements are taken at various design points, where each design point consists of a combination of the levels of the control variables (see Table 1). We refer to the x_i's as the basic variables.

3.2 Model selection phase

When the response of interest (Y) is influenced by a set of factors $\{x_1, x_2, \ldots, x_n\}$, the functional relationship is often not known, but can be estimated to a reasonable degree of accuracy. If the relationship is polynomial in nature, then besides the linear terms in the basic variables, various powers as well as products of various forms also contain information about Y. Most RPD problems are analyzed by obtaining the second order estimated response surface functions. In this paper, we propose the use of higher-order polynomial functions in modeling the response. As we shall show in the numerical example, the prediction ability of the response surface models obtained using the proposed model is higher than that of the second order models.

The model selection phase consists of two stages. In the first stage, we form a set of variables (or factors) made up of the powers and cross-products of the basic variables and augment it with the set of basic variables to form a pool to choose from using statistical model selection techniques. The composition of the pool depends on the order of the model desired. For example, if we are interested in a third order model, we will construct a pool of the form

$$P = \left\{x_i, x_i^3, x_i^2, x_i x_j, x_i x_j x_k, x_i^2 x_j; i \neq j \neq k\right\}. \quad (1)$$

It is easy to show that for n variables, this set will consist of $\frac{n}{6}\left(n^2 + 3n + 14\right)$ elements. For example, if n=3 basic variables, we will have a pool of 16 elements to choose from. The standard statistical techniques used include stepwise regression, all possible subset regression, the coefficient of determination (R^2), the adjusted R-square (R_a^2), the predicted R-square (R_{pred}^2), the prediction error sum of squares (PRESS), the root mean square error (RMSE), Mallows C_p, and the variance inflation factor (VIF). Intuitively, it is obvious that models obtained through this proposed procedure cannot perform any worse than the second order models being used presently. We assume that the functional relationship between the variables in the pool and the response variable of interest (y) is of the form

$$y = X\beta + \varepsilon, \quad (2)$$

where X is a matrix with first column of ones, and the elements of the set P (in Eq. (1)) in the rest of its columns. is a column vector of parameters, and is a vector of random

Table 1 The printing process data [9]

Design point u	Control factors			Observations			Mean	Std. dev.
	x_1	X_2	x_3	y_{u1}	y_{u2}	y_{u3}	$\overline{y}_u$	s_u
1	−1	−1	−1	34	10	28	24.0	12.49
2	0	−1	−1	115	116	130	120.3	8.39
3	1	−1	−1	192	186	263	213.7	42.83
4	−1	0	−1	82	88	88	86.0	3.46
5	0	0	−1	44	188	188	140.0	83.14
6	1	0	−1	322	350	350	340.7	16.17
7	−1	1	−1	141	110	86	112.3	27.57
8	0	1	−1	259	251	259	256.3	4.62
9	1	1	−1	290	280	245	271.7	23.63
10	−1	−1	0	81	81	81	81.0	0.00
11	0	−1	0	90	122	93	101.7	17.67
12	1	−1	0	319	376	376	357.0	32.91
13	−1	0	0	180	180	154	171.3	15.01
14	0	0	0	372	372	372	372.0	0.00
15	1	0	0	541	568	396	501.7	92.50
16	−1	1	0	288	192	312	264.0	63.50
17	0	1	0	432	336	513	427.0	88.61
18	1	1	0	713	725	754	730.7	21.08
19	−1	−1	1	364	99	199	220.7	133.82
20	0	−1	1	232	221	266	239.7	23.46
21	1	−1	1	408	415	443	422.0	18.52
22	−1	0	1	182	233	182	199.0	29.44
23	0	0	1	507	515	434	485.3	44.64
24	1	0	1	846	535	640	673.7	158.21
25	−1	1	1	236	126	168	176.7	55.51
26	0	1	1	660	440	403	501.0	138.94
27	1	1	1	878	991	1161	1010.0	142.45

errors with constant variance and zero mean. The least squares estimator of is given by

$$\widehat{\beta} = (X'X)^{-1}X'y, \tag{3}$$

and the least square predicted model is of the form

$$\widehat{y} = X\widehat{\beta}. \tag{4}$$

For a model with p parameters, R^2, R_a^2, and RMSE are, respectively, defined as

$$R_a^2 = 1 - \left(1 - R^2\right)\left(\frac{n-1}{n-p}\right), \tag{5}$$

$$R^2 = \frac{\widehat{\beta}'X'X - n\overline{y}^2}{y'y - n\overline{y}^2}, \tag{6}$$

and

$$RMSE = \sqrt{\frac{y'y - \widehat{\beta}'X'X}{n-p}} \tag{7}$$

Models with large values of R^2, and R_a^2, and small values of RMSE are sought. It is well known that R^2 is an increasing function of the number of predictors in the model. That is, it increases with additional predictor variables regardless of how significant or insignificant the variables are. On the contrary, R_a^2 may decrease if additional predictors do not contribute significantly to explaining the variability in the response. Thus, it is important to observe both statistics rather than R^2 alone.

The PRESS and R^2_{pred} are useful in assessing the prediction ability of models. If $e_i = y_i - \widehat{y}_i$ represents the i^{th} residual, and, the ith diagonal of the hat matrix (see [33]), which is defined by, then

$$PRESS = \sum_{i=1}^{n}\left(\frac{e_i}{1-h_{ii}}\right)^2, \tag{8}$$

and

$$R^2_{pred} = 100\left[1 - \frac{PRESS}{y'y - n\overline{y}^2}\right]\%. \tag{9}$$

Lower values of PRESS and higher values of indicate a model of high prediction ability.

Since the proposed model of this work recommends consideration of the possibility of adding more variables in

Table 2 ANOVA for the mean response

Predictor	Coef.	SE coef.	T	P	VIF
Constant	339.47	25.90	13.11	0.000	
x1	177.000	8.936	19.81	0.000	1.0
x2	147.00	15.48	9.50	0.000	3.0
x3	115.53	11.99	9.64	0.000	1.8
x11	−3.72	26.81	−0.14	0.892	3.0
x22	−58.11	26.81	−2.17	0.055	3.0
x33	−10.53	20.77	−0.51	0.623	1.8
x12	47.67	18.96	2.51	0.031	3.0
x13	55.00	18.96	2.90	0.016	3.0
x23	43.58	10.94	3.98	0.003	1.0
x233	−56.36	18.96	−2.97	0.014	3.0
x123	82.79	13.40	6.18	0.000	1.0
x1122	80.53	38.85	2.07	0.065	7.0
x1223	30.71	23.22	1.32	0.215	3.0
x1233	27.54	23.22	1.19	0.263	3.0
x11223	35.43	17.98	1.97	0.077	1.8
x112233	−41.26	31.15	−1.32	0.215	3.8
S=37.9142 R-Sq=98.9% R-Sq(adj)=97.2% PRESS=78791.4 R-Sq(pred)=94.14%					
Source	DF	SS	MS	F	P
Regression	16	1331166	83198	57.88	0.000
Residual error	10	14375	1437		
Total	26	1345541			

Table 3 ANOVA for the standard deviation

Predictor	Coef. SE	Coef.	T	P	VIF
Constant	34.208	8.770	3.90	0.005	
X1	36.493	8.320	4.39	0.002	3.0
X2	35.55	10.74	3.31	0.011	5.0
X3	−19.25	14.41	−1.34	0.218	9.0
X11	3.900	8.320	0.47	0.652	1.0
X33	16.930	8.320	2.03	0.076	1.0
X12	−18.83	10.19	−1.85	0.102	3.0
X13	29.02	10.19	2.85	0.022	3.0
X23	29.81	10.19	2.93	0.019	3.0
X112	−22.68	10.19	−2.23	0.057	3.0
X113	61.26	17.65	3.47	0.008	9.0
X122	−37.45	10.19	−3.68	0.006	3.0
X223	56.60	17.65	3.21	0.012	9.0
X233	−7.67	10.19	−0.75	0.473	3.0
X123	29.566	7.205	4.10	0.003	1.0
X1123	−23.59	12.48	−1.89	0.095	3.0
X1223	−35.86	12.48	−2.87	0.021	3.0
X1233	39.83	12.48	3.19	0.013	3.0
X11223	−68.13	21.62	−3.15	0.014	9.0
S=20.3790 R-Sq=94.5% R-Sq(adj)=82.0% PRESS=21219.9 R-Sq(pred)=64.64%					
Source	DF	SS	MS	F	P
Regression	18	56682.3	3149.0	7.58	0.003
Residual error	8	3322.4	415.3		
Total	26	60004.7			

Table 4 Comparison of the mean response models

Model	PRESS	RMSE	R^2	R_a^2	R_{pred}^2
$\hat{\mu}_{SC}(x)$	78791	37.9142	98.9	97.2	94.14
$\hat{\mu}_{VM}(x)$	337545	76.0429	92.7	88.8	74.91
$\hat{\mu}_{LT}(x)$	127072	55.0496	95.7	94.1	90.56

Table 5 Comparing the standard deviation models

Model	PRESS	RMSE	R^2	R_a^2	R_{pred}^2
$\hat{\sigma}_{SC}(x)$	21219.9	20.3790	94.5	82.0	64.64
$\hat{\sigma}_{VM}(x)$	93691.4	44.0414	45.0	16.0	0.00
$\hat{\sigma}_{LT}(x)$	46809.5	37.6679	48.0	38.5	21.99

modeling the response surfaces, we emphasize the inclusion of VIF and Mallows C_p in the selection criteria. This is because the VIF is an important diagnostic for multi-collinearity, while the C_p criterion is used to diagnose bias or over-fit models, i.e., situations where more variables than necessary are added to the model. The VIF corresponding to the i^{th} variable is defined as

$$VIF_i = \frac{1}{1 - R_i^2}, \quad (10)$$

where is the coefficient of determination obtained by regressing the i^{th} variable against the rest of the variables in the model. In practice, VIF values not exceeding ten are tolerated. The Mallows C_p statistic [34] is defined as

$$C_p = \frac{SSE_p}{\hat{\sigma}^2} - n + 2p, \quad (11)$$

where SSE_p is the residual sum of squares of the full model, and is an unbiased estimate of the error variance. Models with values of the C_p statistic that are close to p are considered as the least bias models. Detailed discussions on the use of all these selection criteria are available in Montgomery and Peck [33].

3.3 Optimization phase

In this phase, optimization models are formulated and solved for optimum values of the mean and standard deviation of the response of interest in terms of the control variables. We will briefly summarize the models by VM and LT, followed by our proposed models.

3.4 Models by VM and LT

As mentioned earlier, VM proposed optimization models for the *N*-Type, *L*-Type, and *S*-Type quality characteristics. These models are given in Eqs. (12) through (14), where T_m is the target mean response and T_s is the target standard deviation. The VM models for the three types of quality characteristics are

$$\begin{aligned} &\max\ \hat{\mu}(x) \\ &s.t.\ \hat{\sigma}(x) = T_5 \end{aligned} \quad (12)$$

$$\begin{aligned} &\min\ \hat{\mu}(x) \\ &s.t.\ \hat{\sigma}(x) = T_5 \end{aligned} \quad (13)$$

$$\begin{aligned} &\min\ \hat{\sigma}(x) \\ &s.t.\ \hat{\mu}(x) = T_m \end{aligned} \quad (14)$$

LT observed that and "are only approximations of the 'true' responses (subject to certain random errors)", and that "restricting the optimization to equality constraints will inevitably exclude globally preferred values." Hence, they proposed minimizing the mean square error (MSE) instead

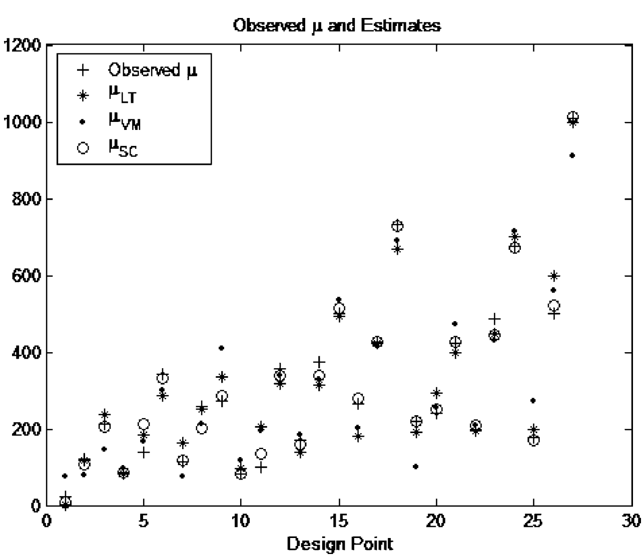

Fig. 1 Comparing models for the mean response

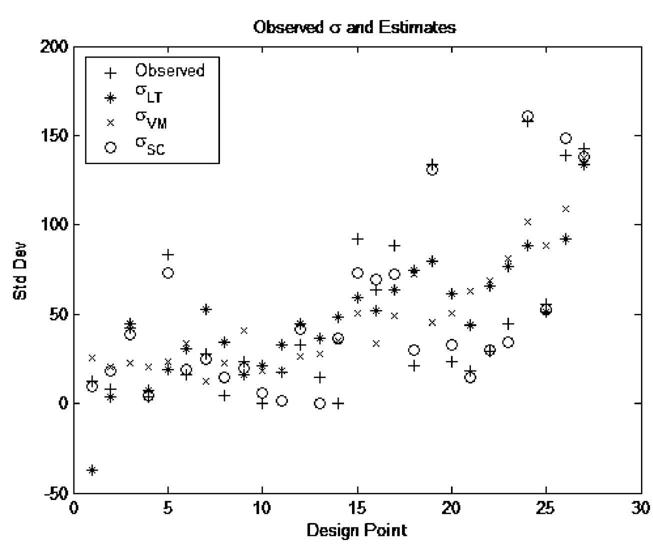

Fig. 2 Comparing models for the standard deviation

Table 6 Comparing solutions to the MSE optimization problem (15) with T=500

Models of mean and standard deviation	Optimal settings x*	$\hat{\mu}(x^*)$	$\hat{\sigma}(x^*)$	f(x*) (MSE)
$\hat{\mu}_{LT}(x)$ and $\hat{\sigma}_{LT}(x)$	(1.000, 1.000, -0.525)	492.231	44.136	2008.309
$\hat{\mu}_{VM}(x)$ and $\hat{\sigma}_{VM}(x)$	(1.000, 0.0599, -0.242)	494.651	44.599	2017.668
$\hat{\mu}_{SC}(x)$ and $\hat{\sigma}_{SC}(x)$	(1.000, 1.000, -0.561)	499.884	12.106	146.557

(see Eq. (15)), arguing that allowing some bias in the mean response results in greater reduction in variability.

$$\min \quad f(x) = (\hat{\mu}(x) - T)^2 + \hat{\sigma}^2(x) \tag{15}$$

All the optimization models are constrained to the experimental region.

3.5 Proposed optimization models

The L-type optimization model of Eq. (13) (proposed by VM) is equivalent to the model

$$\begin{aligned} &\min -\hat{\mu}(x) \\ &s.t. \quad (\hat{\sigma}(x) - T_s)^2 = 0 \end{aligned} \tag{16}$$

We observe that $(\hat{\sigma}(x) - T_s)^2$ is positive for values of T_S not equal to $\hat{\sigma}(x)$, thus the equality constraint forces $\hat{\sigma}(x)$ to equal T_S. Also, assuming that the mean response is positive, the smallest value of $f(x) = -\hat{\mu}(x) + (\hat{\sigma}(x) - T_s)^2$ is achieved when $-\hat{\mu}(x)$ is minimized and $(\hat{\sigma}(x) - T_s)^2$ is as small as possible (i.e., zero at best). In other words, $\hat{\mu}(x)$ is maximized and $\hat{\sigma}(x)$ is as close to S as possible. Hence we propose the optimization model

$$\begin{aligned} &\min f(x) = \hat{\mu}(x) + (\hat{\sigma}(x) - T_s)^2 \\ &s.t. \quad x \in \Omega \end{aligned} \tag{17}$$

where Ω denotes the experimental region of interest. This model relaxes the equality constraint in Eq. (16) in the same way that the MSE optimization model in Eq. (15) relaxes the equality constraint in Eq. (14). Since the smallest variability is always desired, we can consider the target S as an upper bound and seek the solution to the problem

$$\begin{aligned} &\min f(x) = -\left[\hat{\mu}(x) + (\hat{\sigma}(x) - T_s)^2\right] \\ &s.t. \quad \hat{\sigma}(x) \le S \qquad x \in \Omega \end{aligned} \tag{18}$$

This model clearly seeks to maximize $\hat{\mu}(x)$ and simultaneously find the $\hat{\sigma}(x)$ that is at most equal to T_S.

Now we consider an S-type problem, where the objective is to minimize the mean response. In this case, the standard deviation is still the secondary response, and since smaller values are desired, we assume that an upper bound (T_s') is set for the standard deviation. Thus, an analogous optimization model to model (14) is

$$\begin{aligned} &\min f(x) = \hat{\mu}(x) - (\hat{\sigma}(x) - T_s)^2 \\ &s.t. \ \hat{\sigma}(x) \le T_s \qquad x \in \Omega \end{aligned} \tag{19}$$

However, if we only consider a target standard deviation (T_s) and not an upper bound, we will have the less restrictive model

$$\begin{aligned} &\min f(x) = \hat{\mu}(x) + (\hat{\sigma}(x) - T_s)^2 \\ &s.t. \quad x \in \Omega \end{aligned} \tag{20}$$

Finally, we propose an N-Type optimization model, assuming target values T and T_S for the mean and standard deviation respectively. An appropriate optimization model to solve in this case is

$$\begin{aligned} &\min f(x) = (\hat{\mu}(x) - T)^2 + (\hat{\sigma}(x) - T_s)^2 \\ &s.t. \quad x \in \Omega \end{aligned} \tag{21}$$

4 Numerical example

Box and Draper [9] describe an experiment that was conducted to determine the effect on the quality of a printing process of three control variables, namely speed (x_1), pressure (x_2), and distance (x_3). VM used the data of this experiment to illustrate their proposed dual response methodology. In

Table 7 Comparing solutions to optimization problem (12) (VM) with T=500

Models of mean and standard deviation	Optimal settings x*	$\hat{\mu}(x^*)$	$\hat{\sigma}(x^*)$	f(x*) (MSE)
$\hat{\mu}_{LT}(x)$ and $\hat{\sigma}_{LT}(x)$	(1.000, 1.000, -0.502)	500	45.503	2070.529
$\hat{\mu}_{VM}(x)$ and $\hat{\sigma}_{VM}(x)$	(1.000, 0.104, -0.250)	500	45.241	2046.718
$\hat{\mu}_{SC}(x)$ and $\hat{\sigma}_{SC}(x)$	(1.000, 0.105, -0.561)	500	12.107	146.570

Table 8 Comparing solutions to the VM's *L*-Type optimization problem (13) with S=60

Models of mean and standard deviation	Optimal settings x*	$\hat{\mu}\left(x^*\right)$	$\hat{\sigma}\left(x^*\right)$
$\hat{\mu}_{LT}(x)$ and $\hat{\sigma}_{LT}(x)$	(1.000, 1.000, -0.254)	582.355	60
$\hat{\mu}_{VM}(x)$ and $\hat{\sigma}_{VM}(x)$	(1.000, 1.000, -0.278)	616.994	60
$\hat{\mu}_{SC}(x)$ and $\hat{\sigma}_{SC}(x)$	(1.000, 1.000, 0.384)	856.962	60

order to have a fair basis for comparison with the results in VM, LT used the same data to illustrate their proposition of alternative optimization models and improved models of the mean and standard deviation. Similarly, for the purpose of fair comparison, we will use the same data here, which is given in Table 1. We will first use our proposed methodology and find response surface models for the mean and standard deviation, and then compare the models obtained with the models of VM and LT. Next, we will consider solving the optimization models in VM and LT using their response surface models and the models obtained in this work. Finally, we will apply the three sets of response surface models (i.e., ours, VM's, and LT's) to each of our proposed optimization models and compare the optimal solutions we obtain.

4.1 Model selection

For the model selection, we will first create a pool of variables. Because of the coding system chosen for the factor levels (i.e., -1, 0, and 1),

$$x_i^3 = x_i \text{ and } x_i^4 = x_i^2; \quad i = 1, 2, 3. \tag{22}$$

Therefore, we exclude all x_i^3 and all x_i^4 from the pool of variables in the selection of predictor variables. Thus, the pool of variables is given by the set

$$\left\{ \begin{array}{l} x_i, x_i^2, \ x_i x_j, \ x_i^2 x_j, \ x_i^2 x_j^2, \ x_i^2 x_j x_k, \ x_i^2 x_j^2 x_k, \\ x_1^2 x_2^2 x_3^2; \quad i \neq j \neq k; \ i, j, k = 1, 2, 3 \end{array} \right\}.$$

This gives a total of 31variables to consider in the model selection. Using the methods mentioned in the previous section, we obtain the response surfaces in Eqs. (23) and (24) for the mean and the standard deviation respectively. In Tables 2 and 3, we show the analysis of variance (ANOVA) results for the models as obtained using Minitab.

$$\begin{aligned} \hat{\mu}_{sc}(x) = {} & 339.47 + 177.0x_i + 147.0x_2 + 115.53x_3 - 3.72x_1^2 \\ & - 58.11x_2^2 - 10.53x_3^2 + 47.67x_1x_2 + 55.00x_1x_3 \\ & + 43.58x_2x_3 - 56.36x_2x_3^2 + 82.79x_1x_2x_3 + 80.53x_1^2x_2^2 \\ & + 30.71x_1x_2^2x_3 + 24.54x_1x_2^2x_3^2 + 35.43x_1^2x_2^2x_3 \\ & - 41.26x_1^2x_2^2x_3^2 \end{aligned} \tag{23}$$

$$\begin{aligned} \hat{\sigma}_{sc}(x) = {} & 34.208 + 36.49x_1 + 35.55x_2 - 19.25x_3 + 3.9x_1^2 \\ & + 16.93x_3^2 - 18.83x_1x_2 + 29.02x_1x_3 + 29.81x_2x_3 \\ & - 22.68x_1^2x_2 + 61.26x_1^2x_3 - 37.45x_1x_2^2 + 56.60x_2^2x_3 \\ & - 7.67x_2x_3^2 + 29.57x_1x_2x_3 - 23.59x_1^2x_2x_3 \\ & - 35.86x_1x_2^2x_3 + 39.83x_1x_2x_3^2 - 68.13x_1^2x_2^2x_3 \end{aligned} \tag{24}$$

4.2 Comparison with previous models

In modeling the mean response for the data of Table 1, VM and LT reported the models in Eqs. (22) and (26) respectively. Model (26) is an improvement upon (24), and was found using model selection criteria on the terms of the full third order model.

$$\begin{aligned} \hat{\mu}_{VM}(x) = {} & 327.6 + 177.0x_1 + 109.4x_2 + 131.5x_3 + 32.0x_1^2 \\ & - 22.4x_2^2 - 29.1x_3^2 + 66.0x_1x_2 + 75.5x_1x_3 + 43.6x_2x_3 \end{aligned} \tag{25}$$

$$\begin{aligned} \hat{\mu}_{LT}(x) = {} & 314.67 + 117.0x_1 + 109.426x_2 + 131.463x_3 \\ & + 66.028x_1x_2 + 75.472x_1x_3 + 43.583x_2x_3 + 82.792x_1x_2x_3 \end{aligned} \tag{26}$$

Table 9 Comparing solutions to VM's *S*-type optimization problem (14) with S=60

Models of mean and standard deviation	Optimal settings x*	$\hat{\mu}\left(x^*\right)$	$\hat{\sigma}\left(x^*\right)$
$\hat{\mu}_{LT}(x)$ and $\hat{\sigma}_{LT}(x)$	(−1.0000, -1.0000, 0.6604)	157.5408	60
$\hat{\mu}_{VM}(x)$ and $\hat{\sigma}_{VM}(x)$	(−1.0000, -0.3810, 1.0000)	172.8955	60
$\hat{\mu}_{SC}(x)$ and $\hat{\sigma}_{SC}(x)$	(−1.0000, −1.0000, 0.5582)	149.9620	60

Table 10 Comparing solutions to the proposed *L*-type optimization model (17) with T_S=60

Models of mean and standard deviation	Optimal settings x*	$\hat{\mu}\left(x^*\right)$	$\hat{\sigma}\left(x^*\right)$	*f*(x*) (MSE)
$\hat{\mu}_{LT}(x)$ and $\hat{\sigma}_{LT}(x)$	(1.000, 1.000, -0.206)	598.491	62.840	−590.423
$\hat{\mu}_{VM}(x)$ and $\hat{\sigma}_{VM}(x)$	(1.000, 1.000, -0.200)	637.747	63.156	−627.790
$\hat{\mu}_{SC}(x)$ and $\hat{\sigma}_{SC}(x)$	(1.000, 1.000, 0.399)	861.624	61.518	−859.320

In Table 4, we compare these two models with the model we obtained in Eq. (23) based on PRESS, RMSE, R^2, R_a^2, and R_{pred}^2. We observe in terms of all these criteria that the model $\hat{\mu}_{SC}(x)$ of Eq. (23) is superior to the models of Eqs. (25) and (26). Thus, Eq. (23) better describes the mean response and also has a greater ability to predict the mean response than the other previous models.

We also note that the decrease in the RMSE achieved by $\hat{\mu}_{SC}(x)$ is about 50% relative to $\hat{\mu}_{VM}(x)$ and about 31% relative to $\hat{\mu}_{LT}(x)$. The reductions achieved in the PRESS values are 77% and 38% relative to $\hat{\mu}_{VM}(x)$ and $\hat{\mu}_{LT}(x)$, respectively. In Fig. 1, we present a graph of the observed mean values (μ) and the values obtained from the response surfaces $\hat{\mu}_{SC}(x)$, $\hat{\mu}_{VM}(x)$, and $\hat{\mu}_{LT}(x)$. The closeness of the values obtained from $\hat{\mu}_{SC}(x)$ to the observed mean values as compared to values obtained from $\hat{\mu}_{VM}(x)$ and $\hat{\mu}_{LT}(x)$ is evident from the graph, which indicates that the performance of is indeed an improvement upon the performance of $\hat{\mu}_{VM}(x)$, and $\hat{\mu}_{LT}(x)$.

For the same data, VM modeled the standard deviation by the full quadratic model

$$\hat{\sigma}_{VM} = 34.9 + 11.5x_1 + 15.3x_2 + 29.2x_3 + 4.2x_1^2 - 1.3x_2^2 + 16.8x_3^2 + 7.7x_1x_2 + 5.1x_1x_3 + 14.1x_2x_3 \quad (27)$$

In an attempt to improve upon this model, LT considered the full third order model and obtained the model in Eq. (28) via model selection techniques.

$$\hat{\sigma}_{LT} = 47.994 + 11.527x_1 + 15.323x_2 + 29.190x_3 + 29.566x_1x_2x_3 \quad (28)$$

Table 5 compares the standard deviation models of Eqs. (27) and (28) with the model $\hat{\sigma}_{SC}(x)$ in Eq. (24) that we proposed in this paper. Again, we observe that our proposed model, $\hat{\sigma}_{SC}(x)$, gives the least values of PRESS and RMSE, and also the highest values of R^2, R_a^2, and R_{pred}^2. A graphical comparison of the three models is shown in Fig. 2, and just as in the case of the means, we observe that the values obtained from $\hat{\sigma}_{SC}(x)$ are more likely to be closer to the observed standard deviation values than the values obtained from $\hat{\mu}_{LT}(x)$ and $\hat{\mu}_{VM}(x)$.

4.3 Optimization

In this part of the example, we will solve the various optimization models presented above, based on the response surface models obtained through our proposed methodology and those of VM and LT. We solve the optimization models of Eqs. (12), (13) and (15) using the *fmincon* routine in MATLAB, and compare the optimum solutions.

Table 6 shows the solutions to the MSE optimization model (i.e., Eq. (15)) with a target value of 500 for the mean response. We observe that the smallest MSE value is obtained from using the higher order models of this paper (i.e., $\hat{\mu}_{SC}(x)$ and $\hat{\sigma}_{SC}(x)$). Also, $\hat{\mu}_{SC}(x)$ and $\hat{\sigma}_{SC}(x)$give the smallest value of standard deviation and the closest mean value to the target mean. The observations here support the observations made in the previous section when the various response surface models were compared.

In Table 7, we display the solutions to the optimization problem in Eq. (12) for the various response surface Models of mean and standard deviation. Again, the results based on $\hat{\mu}_{SC}(x)$ and $\hat{\sigma}_{SC}(x)$ give the least values of MSE and standard deviation. By examining Tables 6 and 7 together, we observe that $\hat{\mu}_{SC}(x)$ and $\hat{\sigma}_{SC}(x)$ are relatively more robust in the sense that the MSE value at optimality is about the same for both optimization methods.

VM used the method of Lagrange multipliers and solved the *L*-type optimization problem in Eq. (13) for the data of Table 1, assuming target standard deviation values (T_S) of 60, 75, and 90. For T_S=60, we solve the same problem for the three sets of response surface models. Table 8 compares the solutions obtained. The solution based on $\hat{\mu}_{SC}(x)$ and

Table 11 Comparing solutions to the proposed *S*-type optimization model (20) with T_S=60

Models of mean and standard deviation	Optimal settings x*	$\hat{\mu}\left(x^*\right)$	$\hat{\sigma}\left(x^*\right)$	*f*(x*)
$\hat{\mu}_{LT}(x)$ and $\hat{\sigma}_{LT}(x)$	(−1.0000, −1.0000, 0.6466)	156.8879	59.1920	156.8879
$\hat{\mu}_{VM}(x)$ and $\hat{\sigma}_{VM}(x)$	(−1.0000, −0.4822, 1.0000)	162.1393	59.4685	167.5959
$\hat{\mu}_{SC}(x)$ and $\hat{\sigma}_{SC}(x)$	(−1.0000, −1.0000, 0.5542)	149.3975	57.6641	149.6801

Table 12 Comparing Solutions to the Proposed *N*-Type Optimization Model (21) with T=500 and T_S=60

Models of mean and standard deviation	Optimal settings x*	$\hat{\mu}\left(x^*\right)$	$\hat{\sigma}\left(x^*\right)$
$\hat{\mu}_{LT}(x)$ and $\hat{\sigma}_{LT}(x)$	(0.575, 0.900, -0.192)	500.000	60.000
$\hat{\mu}_{VM}(x)$ and $\hat{\sigma}_{VM}(x)$	(0.344, 0.504, 0.274)	500.000	60.000
$\hat{\mu}_{SC}(x)$ and $\hat{\sigma}_{SC}(x)$	(0.477, 0.545, -0.017)	500.000	60.000

$\hat{\sigma}_{SC}(x)$ gives a larger mean response value (at the target standard deviation) than the solutions based on the response surface models of MV and LT.

In Table 9, we show the solutions to VM's *S*-type optimization problem in (14). Again, the solution based on $\hat{\mu}_{SC}(x)$ and $\hat{\sigma}_{SC}(x)$ yields the most desired result, i.e., the smallest mean response value.

It is worth noting from Tables 7, 8, 9 and 13 that all the response surfaces considered gave optimal solutions that achieved the desired target in each case. However, the optimal settings in each case are not exactly the same. This accentuates the need for obtaining more effective response surface models in terms of prediction ability.

In what follows, we will solve the proposed optimization models of this work for all the response surface models we have been considering, and compare the results in a similar manner. We solved our proposed *L*-Type model in Eq. (17) for the data shown in Table 1 and obtained exactly the same solution set in Table 8. However, we recommend that the model in Eq. (18) be considered, since it has the flexibility of considering smaller values of the standard deviation in the optimization process. Table 10 compares the solutions to Eq. (17) obtained by using the various response surface models. Clearly, by allowing a little bias in the standard deviation, larger mean response values are obtained as compared to the solution of VM's model in Eq. (13), which allows no bias in the standard deviation. We again observe that $\hat{\mu}_{SC}(x)$ and $\hat{\sigma}_{SC}(x)$ give the smallest optimum objective function value, the smallest optimum standard deviation, and the largest optimum mean response value.

For our proposed *S*-type model in (19), we use the target standard deviation T_S=60 and solve it obtaining exactly the same solutions for VM's *S*-Type model in Eq. (14) shown in Table 9. Table 11 shows the optimal solutions to model (20) (i.e the alternative *S*-type model to (19)) for the same data when T_S=60. We observed here that by relaxing the equality constraint, the standard deviation actually dropped slightly and the mean response is further decreased. Also in the case also, the solutions based on and $\hat{\sigma}_{SC}(x)$ are more desirable in terms of the objectives, i.e., smaller standard deviation and smaller mean.

Finally, we solve our proposed *N*-Type problem in Eq. (21) for the data of Table 1 using the sets of target values (T=500, T_S=60) and (T=600, T_S=20). Table 12 shows the optimum solutions of this model for the first set of target values, where all the response surfaces used achieved the desired targets.

In Table 13, we display the optimum solutions for the second set of target values. We observe in this case that only the solutions based on $\hat{\mu}_{SC}(x)$ and $\hat{\sigma}_{SC}(x)$ achieved the desired targets, and therefore give the smallest possible objective function value (to three decimals). The solutions based on the response surfaces of VM and LT give much larger standard deviation values and smaller means than *T*.

5 Concluding remarks

In the context of robust parameter design optimization problems, we have addressed the need to consider higher order polynomial response surface models for the mean and standard deviation of quality characteristics as a way of increasing the predictive ability of the response surface models. A numerical example was used to illustrate the increased accuracy of the response surface models obtained in this work relative to existing response surface models for the same example. Significant increases in *R*-square, adjusted *R*-square, and predicted *R*-square were achieved by the models of this work, as well as significant decreases in root mean square error, and predicted error sum of squares. For example, relative to the best of the existing response surface models for the example considered, the mean response model obtained in this paper is shown to be higher by 3.95% in predicted R-square, while the standard

Table 13 Comparing solutions to the proposed *N*-type optimization model (16) with T=600 and T_S=20

Models of mean and standard deviation	Optimal settings x*	$\hat{\mu}\left(x^*\right)$	$\hat{\sigma}\left(x^*\right)$	*f*(x*)
$\hat{\mu}_{LT}(x)$ and $\hat{\sigma}_{LT}(x)$	(1.000, 1.000, -0.223)	592.640	61.811	1802.29
$\hat{\mu}_{VM}(x)$ and $\hat{\sigma}_{VM}(x)$	(0.344, 0.504, 0.274)	595.085	56.947	1389.24
$\hat{\mu}_{SC}(x)$ and $\hat{\sigma}_{SC}(x)$	(0.943, 0.997, -0.289)	600.000	20.000	0.00

deviation model is higher by 193.95%. The improvement in the modeling of the standard deviation is particularly important since modeling variability has generally been problematic, as observed in the literature. Optimization models were proposed and solved for the larger-the-better type, the smaller-the-better type, and the nominal-the-best type quality characteristics assuming in each case that there is a target value for the standard deviation.

We believe that considering higher order response surface models and using the well-known statistical model selection techniques properly will enhance the quality of solutions to robust parameter design problems. In fact, where the most powerful models are of lower order, such models will be sought out by the model selection procedure. Therefore, the proposed procedure of this paper is very likely to yield response surface models that are at least as powerful as the existing lower order models in the literature.

A natural extension of this work would be in the consideration of multiple quality characteristics, and modeling involving both controllable and noise factors. Finally, we recommend a great deal of caution when using the kinds of response surface models proposed in this paper and in VM and LT, as it is possible for such models to give results of no practical meaning or significance. For example, LT's standard deviation model, $\hat{\sigma}_{LT}(x)$ in Eq. (7), gives a negative standard deviation ((−37.3) when x=[−1, −1, −1]. Therefore, even with very powerful response surfaces, we recommend including constraints in optimization models that will serve to prevent such results from occurring (e.g., $0 \leq \sigma(x) \leq T_S'$, where T_S' is the upper bound value for the standard deviation). Another option is to consider functional forms such as exponential and logistic functions, which do not allow negative output values.

References

1. Taguchi G (1986) Introduction to quality engineering. UNIPBUK/Krans International: White Plains, NY.
2. Taguchi G (1987) Systems of experimental design: engineering methods to optimize quality and minimize cost. UNIPBUK/Krans International: White Plains, NY
3. Nair VN, Shoemaker AC (1990) The role of experimentation in quality engineering: a review of Taguchi's contribution. Design and analysis of industrial experiments. NY: Marcel Dekker, pp. 247–277
4. Vining GG, Myers RH (1990) Combining Taguchi and response surface philosophies: a dual response approach. J Qual Technol 22:38–45
5. Myers RH, Carter WH Jr (1973) Response surface techniques for dual response systems. Technometrics 15:301–317
6. Hoerl AE (1959) Optimum solution of many variables equations. Chem Eng Prog 55:67–78
7. Draper NR (1963) Ridge analysis of response surfaces. Technometrics 5:469–479
8. Myers RH (1976) Response surface methodology. Department of Statistics, Virginia Polythechnic Institute and State University, Blacksburg
9. Box GEP, Draper NR (1987) Empirical model building and response surface. Wiley, New York
10. Khuri AI, Cornell JA (1996) Response surfaces: design and analyses. Decker, New York
11. Umland AW, Smith WN (1959) The use of Lagrange multipliers with response surfaces. Technometrics 1(3):289–292
12. Del Castillo E, Montgomery DC (1995) A nonlinear programming solution to the dual response problem. J Qual Technol 25:199–204
13. Cho BR (1994) Optimization issues in quality engineering. Unpublished PhD Dissertation, School of Industrial Engineering, University of Oklahoma, USA
14. Lin DKJ, Tu W (1995) Dual response surface optimization. J Qual Technol 27:34–39
15. Copeland KAF, Nelson PR (1996) Dual response optimization via direct function minimization. J Qual Technol 28:331–336
16. Kim K, Lin DKJ (1998) Response surface optimization: a fuzzy modeling approach. J Qual Technol 30:1–10
17. Del Castillo E, Fan SK, Semple J (1997) Computation of global optima in dual response systems. J Qual Technol 29:347–353
18. Fan SKS (2000) A generalized global optimization algorithm for dual response systems. J Qual Technol 32:444–456
19. Kim YJ, Cho BR (2002) Development of priority-based robust design. Qual Eng 14:355–363
20. Tang LC, Xu K (2002) Unified approach for dual response optimization. J Qual Technol 34(4):437–447
21. Köskoy O, Doganaksoy N (2003) Joint optimization of mean and standard deviation using response surface methods. J Qual Technol 35:239–252
22. Lam SW, Tang LC (2005) A hraphical spproach to thedual response robust design approach. Reliability and maintainability symposium. Proceedings. Annual. pp. 200–206
23. Vining GG, Kowalski SM, Montgomery DC (2005) Response surface designs within a split-plot structure. J Qual Technol 37(2):115–129
24. Kowalski SM, Vining GG, Montgomery DC, Borror CM (2006) Modifying a central composite design to model the process mean and variance when there are hard-to-change factors. Appl Stat 55 (5):615–630
25. Kunert J, Auer C, Erdbrugge M, Ewers R (2007) An experiment to compare Taguchi's product array and the combined array. J Qual Technol 39(1):17–34
26. Del Castilllo E, Alvarez MJ, Ilzarbe L, Viles E (2007) A new design criterion for robust parameter experiments. J Qual Technol 39(3):279–295
27. Nelder JA, Lee Y (1991) Generalized linear models for the analysis of Taguchi-type experiments. Appl Stoch Models and Data Analysis 7:107–120
28. Lee Y, Nelder JA (1998) Robust design via generalized linear models. Can J Stat 8(26):95–105
29. Lee Y, Nelder JA (2003) Robust design via generalized linear models. J Qual Technol 35:2–12
30. Engel J, Huele AF (1996) A generalized linear model approach to robust design. Technometrics 25:365–373
31. Myers WR, Brenneman WA, Myers RH (2005) A dual-response approach to robust parameter design for a generalized linear model. J Qual Technol 37(2):130–138
32. Robinson TJ, Wulff SS, Montgomery DC (2006) Robust parameter design using generalized linear mixed models. J Qual Technol 38 (1):65–75
33. Montgomery DC, Peck EA (1992) Introduction to Linear Regression Analysis, 2nd Ed., Wiley, New York
34. Mallows CL (1964) Choosing variables in a linear regression: a graphical aid. Central regional meeting of the Institute of Mathematical Statistics, Manhattan, Kansas