ESKİŞEHİR TECHNICAL UNIVERSITY JOURNAL OF SCIENCE AND TECHNOLOGY
A- APPLIED SCIENCES AND ENGINEERING

2019, 20(4), pp. 458 - 473, DOI: 10.18038/estubtda.515702

OPTIMIZATION AND DECISION MAKING STAGES FOR ANALYSIS OF DUAL RESPONSE PROBLEM: MEDIAN-MAD, NSGA-II, TOPSIS

Özlem TÜRKŞEN[1,*], Ahmet KOCATÜRK[2], Gözde KARAKOÇ[3]

[1] Department of Statistics, Faculty of Science, Ankara University, Ankara, Turkey
[2] Department of Statistics, Faculty of Science, Gazi University, Ankara, Turkey
[3] Department of Mathematics, Faculty of Science, Ankara University, Ankara, Turkey

ABSTRACT

The purpose of this study is to obtain a proper set of input variables for replicated response measured data set by applying dual response strategy in robust framework with multi-objective perspective. The replicated response measures were transformed to dual response by using robust statistics, median (*MED*) and median absolute deviation (*MAD*), instead of mean-standard deviation statistics which were very commonly used in many existing studies. A compromise solution of proposed robust dual response model was obtained via a multi-criteria decision making approach since the optimization was achieved in multi-objective point of view. In this study, well-known two methods, NSGA-II and TOPSIS, were preferred for optimization and decision making stages, respectively. Quality of printing ink data set was used an application from the literature. It is seen from the analysis results that the performance of the proposed robust dual response model was encouraging with the most satisfactory input settings.

Keywords: Replicated response measures, Dual response problem, Robust statistics, Multiobjective optimization, Multi-criteria decision making

1. INTRODUCTION

Most of the product or process problems need a proper experimental design which is commonly composed of replicated response measures. One of the popular approach to model replicated response measured data set is dual response strategy. In the dual response strategy, central tendency and spread of replicates are considered as dual responses. The idea of dual response approach (DRA) was firstly introduced by [1] and popularized by [2]. Basically, the DRA builds two empirical response models and then optimizes these predicted dual response functions.

Mean and standard deviation statistics of replicates have been preferred to use as dual responses so far in various studies, e.g. [3–12]. The detailed literature studies can be found in the study of [13]. However, sometimes, mean and standard deviation of replicated response values may not be proper to consider as dual responses. In case of the replicated values have extreme values, robust statistics should be preferred to use. It is seen from the literature survey that even there have been many researches about modeling of the mean and the standard deviation as dual responses, the modeling studies about median (*MED*) and median absolute deviation (*MAD*) are fewer. Recently, the *MED* and the *MAD* statistics have been used to transform the replicated response measures to robust responses by [14, 15].

The main aim of this study is to focus that the mean and standard deviation may not always applicable for analysis of the replicated response measured data set as dual response problem. For this purpose, the *MED* and the *MAD* of replicated response measures were used as dual responses in robust framework. In this study, analysis of replicated response measured data set was achieved in two basic stages: (i) modeling, and (ii) optimization. In the modeling stage, second order polynomial probabilistic response functions were used for fitting observed the *MED* and the *MAD* responses with considering that the errors were independent with constant variance and zero mean. The dual response

*Corresponding Author: turksen@ankara.edu.tr
Received: 21.01.2019 Published:

model parameters were estimated by using least squares method. However, if the errors are not normally distributed the parameter estimation should be done according to the distribution of errors by using maximum likelihood (ML) method. The model parameter estimation methods in dual response studies are summarized in the study of [13]. In the optimization stage, robust responses were optimized simultaneously in multi-objective framework by using a multi-objective optimization (MOO) method, called NSGA-II (Non-dominated Sorting Genetic Algorithm-II). The NSGA-II, developed by [16], is one of the well-known multi-objective algorithm. The application of NSGA-II for response surface studies can be seen in the recent studies of [17] and [18]. By using NSGA-II as a MOO tool, it is possible to obtain Pareto solution set in a single run. And also, the NSGA-II gives many non-dominated solution in a short time. A compromise solution is obtained by using a well-known multi-criteria decision making approach, called TOPSIS (Technique for Order Preference by Similarity to an Ideal Solution) which is presented in [19] with reference to [20].

The paper was organized as follows. In Section 2, detailed description about modeling of replicated response measures as dual response problem through robust statistics was given and performance metrics of predicted responses were presented briefly. Section 3, contains evaluation of the robust dual response problem as MOO problem. The NSGA-II and TOPSIS methods were also shortly explained in Section 3. In Section 4, a real data set from the literature was used to illustrate the proposed analysis approach of replicated response measures as robust dual responses in multi-objective framework. Finally, conclusion was given in Section 5.

2. MODELING OF DUAL RESPONSE PROBLEM IN ROBUST FRAMEWORK

2.1. Modeling of Replicated Response Measures as Dual Response Problem

Let consider a system involving a response *Y* that depends on the level of *k* control factors or input variables, $\left(X_1, X_2, ..., X_k\right)$ with the assumption that the levels of X_i, $i = 1, 2, ..., k$ are quantitative and continuous, and also can be controlled by the experimenter. Suppose that *m* replicates are taken at each of the design points. The experimental format is illustrated in Table 1. Here, Y_{ij} represent the *j*th replicates at the *i*th design point where $i = 1, 2, ..., n$ and $j = 1, 2, ..., m$.

Table 1. Experimental design with replicated response values

Unit numbers	Control Factors				Replicates of Response			
	X_1	X_2	$\cdots$	X_k	Y_1	Y_2	$\cdots$	Y_m
1	x_{11}	x_{12}	$\cdots$	x_{1k}	y_{11}	y_{12}	$\cdots$	y_{1m}
2	x_{21}	x_{22}	$\cdots$	x_{2k}	y_{21}	y_{22}	$\cdots$	y_{2m}
.	.			.	.			.
.	.			.	.			.
.	.			.	.			.
n	x_{n1}	x_{n2}	$\cdots$	x_{nk}	y_{n1}	y_{n2}		y_{nm}

Generally, in order to apply dual response strategy to the experimental design, given in Table 1, two most popular statistics which are mean and standard deviation have been used as center and spread estimators of replicates so far. At the *i*th design point, the sample mean and sample standard deviation of replicates are calculated as follows:

$$\bar{Y}_i = \frac{1}{m}\sum_{j=1}^{m} Y_{ij} \quad \text{and} \quad S_i = \left(\frac{1}{m-1}\sum_{j=1}^{m}\left(Y_{ij} - \bar{Y}_i\right)^2\right)^{1/2}, \quad i = 1, 2, ..., n, \quad j = 1, 2, ..., m \tag{1}$$

However, these estimators are inappropriate for small sized non-normal data. And also, $\bar{Y}_i$ and S_i, $i=1,2,...,n$, are very sensitive to outliers of replicates that may present unexpected values in the experiment. In this case, it will be better to use robust statistics, e.g. the *MED* and the *MAD* as alternatives to the mean and standard deviation. The *MED* and the *MAD* of the experimental design can be defined as

$$MED_i = median\left(Y_{i1}, Y_{i2}, ..., Y_{im}\right),\ i=1,2,...,n, \tag{2}$$

and

$$MAD_i = \underset{1\le j\le m}{median}\left\{\left|Y_{ij} - MED_i\right|\right\},\ i=1,2,...,n. \tag{3}$$

The Table 1 is adapted to apply the dual response strategy through the robust statistics as given in Table 2.

Table 2. Replicated response measured experimental design with robust dual responses

Unit number	Control Factors				Replicates of Response				Robust Dual Responses	
	X_1	X_2	$\cdots$	X_k	Y_1	Y_2	$\cdots$	Y_m	*MED*	*MAD*
1	x_{11}	x_{12}	$\cdots$	x_{1k}	y_{11}	y_{12}	$\cdots$	y_{1m}	MED_1	MAD_1
2	x_{21}	x_{22}	$\cdots$	x_{2k}	y_{21}	y_{22}	$\cdots$	y_{2m}	MED_2	MAD_2
. . .	. . .			. . .	. . .			. . .	. . .	. . .
n	x_{n1}	x_{n2}	$\cdots$	x_{nk}	y_{n1}	y_{n2}		y_{nm}	MED_n	MAD_n

In Table 2, the *MED* and the *MAD* columns are considered as dual responses. One of the main purpose in dual response problem is to define the functional relationship between input variables $\left(X_1, X_2, ..., X_k\right)$ and the dual responses, *MED* and *MAD*, with minimum error. It should be noted here that the *MED* and the *MAD* are uncorrelated.

In many of the response surface studies, second order polynomial regression functions are used for fitting the data. The detailed information can be seen in the study of [21]. The response functions for the *MED* and the *MAD* can be given as

$$\mathbf{Y}_{MED} = \eta\left(\boldsymbol{\beta}, \mathbf{X}\right) + \boldsymbol{\varepsilon}_{MED} \tag{4}$$

and

$$\mathbf{Y}_{MAD} = \eta\left(\boldsymbol{\gamma}, \mathbf{X}\right) + \boldsymbol{\varepsilon}_{MAD} \tag{5}$$

in which

$$\eta\left(\boldsymbol{\beta}, \mathbf{X}\right) = \beta_0 + \sum_{t=1}^{k} \beta_t X_{ti} + \sum_{t=1}^{k}\sum_{w=1}^{k} \beta_{tw} X_{ti} X_{wi}\ ,\ i=1,2,...,n, \tag{6}$$

and

$$\eta\left(\boldsymbol{\gamma}, \mathbf{X}\right) = \gamma_0 + \sum_{t=1}^{k} \gamma_t X_{ti} + \sum_{t=1}^{k}\sum_{w=1}^{k} \gamma_{tw} X_{ti} X_{wi}\ ,\ i=1,2,...,n. \tag{7}$$

To obtain predicted response functions, the least squares estimates of unknown model parameters can be given as

$$\hat{\boldsymbol{\beta}} = \left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{Y}_{MED} \tag{8}$$

and

$$\hat{\boldsymbol{\gamma}} = \left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{Y}_{MAD}. \tag{9}$$

In order to apply the statistical test to the model parameters, $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$, the errors must be normally distributed. Otherwise, the statistical inference for the model parameters will be invalid. In the event of non-normal distributed errors, the model parameters of the *MED* and the *MAD* responses should be estimated by taking into consideration the distribution of errors through ML method.

2.2. Performance Metrics for Predicted Dual Response Models

The predicted performance of the dual response models are calculated by using the coefficient of determination (R^2), the adjusted coefficient of determination (R^2_{adj}), the root mean square error (*RMSE*), the mean absolute error (*MAE*) and the prediction error sum of squares (*PRESS*) are defined as, respectively,

$$R^2 = \frac{\hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{Y} - n\bar{Y}^2}{\mathbf{Y}'\mathbf{Y} - n\bar{Y}^2}, \tag{10}$$

$$R^2_{adj} = 1 - \left(1 - R^2\right)\left(\frac{n-1}{n-p}\right), \tag{11}$$

$$RMSE = \sqrt{\frac{\mathbf{Y}'\mathbf{Y} - \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{Y}}{n-p}}, \tag{12}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|Y_i - \hat{Y}_i\right|, \tag{13}$$

and

$$PRESS = \sum_{i=1}^{n}\left(\frac{Y_i - \hat{Y}_i}{1-h_{ii}}\right)^2 \tag{14}$$

in which p is the number of model parameters, $\hat{Y}_i$ is the ith predicted response value, h_{ii} is the ith diagonal element of the hat matrix, defined by $H = \mathbf{X}\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'$. Generally, the predicted dual response models with large R^2 and R^2_{adj} are preferred. The *RMSE*, *MAE* and *PRESS* are useful in assessing the prediction ability of the models. The lower values of these metrics indicate that the predicted dual response models have high prediction ability.

3. OPTIMIZATION OF DUAL RESPONSE PROBLEM WITH MULTI-OBJECTIVE PERSPECTIVE

The main goal of the dual response optimization (DRO) stage is to identify appropriate values of input variables. It is possible to obtain the appropriate settings of input variables by minimizing the deviations and remaining the central tendency at the preferred value. There have been many approaches in the statistics literature for DRO problems. A simple and straight forward approach for DRO is converting the dual responses into a single response function without constraints [3, 22] and with constraints [2, 6, 8, 23–29]. In multi-objective perspective, [30] studied mean-standard deviation DRO by using multi-objective Genetic Algorithm. In the field of DRO, most of these recent works have only focused on the mean-standard deviation dual response. In this study, robust dual response problem, composed with the *MED* and the *MAD* statistics, was considered. There have been too few studies about the optimization of robust dual responses, e.g. [14, 15, 31]. For optimization stage, the main contribution of the study is optimizing the robust dual response problem with multi-objective perspective. For this purpose, the robust dual response problem is considered as MOO problem, given below:

$$\begin{aligned} \min\ f_1(\mathbf{X}) &= \left\{\left|\hat{Y}_{MED}(\mathbf{X}) - T\right|\right\} \\ \min\ f_2(\mathbf{X}) &= \left\{\hat{Y}_{MAD}(\mathbf{X})\right\} \\ \mathbf{X} &\in S \end{aligned} \tag{15}$$

where T is the target value of central tendency and S is the domain of input variables. To optimize the Equation (15), a well-known population based multi-objective optimization algorithm, the NSGA-II was used. The optimization result was a set of non-dominated input values, called as Pareto solution set. The solutions in the Pareto set can be considered as alternative solutions. In this case, it is necessary to apply a decision making approach to define a compromise solution among the many alternative solutions. The TOPSIS method was preferred to use as multi-criteria decision making approach in the scope of this study. Some brief descriptions about the NSGA-II and the TOPSIS are given in the following sections.

3.1. NSGA-II

The NSGA-II finds a set of non-dominated solutions in a single run without requiring any preference information. The main working principle of the algorithm is based on a fast non-dominated sorting mechanism and a crowding distance to construct the Pareto front in a non-dominated order. In order to implement the NSGA-II to the robust dual response problem, firstly a random initial population is created in the search space with the size n_{pop} and called parent population. Then offspring population is generated from the parent population by using genetic operators e.g. a selection operator, a crossover operator with a proper crossover probability ($\Pr_{cross}$), a mutation operator with a proper mutation probability ($\Pr_{mut}$). By combining current and offspring populations, the next population is constructed and the non-dominated fronts are generated according to the non-dominated sorting and crowding distance operators. The search process of the algorithm continuous until the number of generations, n_{gen} , is reached.

3.2. TOPSIS

The TOPSIS method selects the most preferred alternative solution which is closest to the positive ideal solution (PIS) and farthest from the negative ideal solution (NIS). In order to apply the TOPSIS to the Pareto solution set of the problem, given in Equation (15), decision matrix (*D*) was constructed and the weight vector of criteria (*W*) was determined. The normalized decision matrix (*N*) was calculated since the criteria may have different units. Then, weighted normalized decision matrix was calculated. Provided that the sum of weights was equal to 1. The PIS and NIS were identified. The PIS is the solution that maximizes the benefit criteria and minimizes the cost criteria whereas the NIS maximizes the cost criteria and minimizes the benefit criteria. Afterwards, the separation measures from the PIS and the NIS are calculated by using the Euclidean distance. It should be noted here that a number of distance metric can be applied in the TOPSIS method. The relative closeness of the alternative solutions to the PIS were calculated and denoted as $R_i,\ i = 1, 2, ..., n_{pop}$ where n_{pop} is the number of Pareto solutions. Finally, a set of alternatives were ranked by the descending order of the value of $R_i,\ i = 1, 2, ..., n_{pop}$. Generally, the first order $R_i,\ i = 1, 2, ..., n_{pop}$, value was considered as the most compromise solution among the many alternatives. However, it is possible to use the decision rules defined by [32]. The decision rules are given in Table 3 for the $R_i,\ i = 1, 2, ..., n_{pop}$, values.

Table 3. Approval status of decision rules for relative closeness , R_i [32]

Relative closeness (R_i)	**Assessment status**
$R_i \in [0, 0.2)$	Do not recommend
$R_i \in [0.2, 0.4)$	Recommend with high risk
$R_i \in [0.4, 0.6)$	Recommend with low risk
$R_i \in [0.6, 0.8)$	Approved
$R_i \in [0.8, 1)$	Approved and preferred

4. APPLICATION

In this section, a real data set application is given in order to illustrate the modeling and optimization stages of replicated response measures as robust dual response via robust statistics with multi-objective perspective. The data set is, called printing process data, firstly described in the study of [33]. The purpose of the printing process data set is to determine the effects of three input variables, called speed (X_1), pressure (X_2) and distance (X_3) on the quality of a printing process, which is the response of experiment. The experiment was conducted in 3^3 factorial design with three replications in each *i*th experimental unit, $i = 1, 2, ..., 27$. The experimental data set can be seen in Table 4 with the coded values of input variables.

Before transforming the replicated response values to dual response, it is necessary to analyze the replicated values of responses statistically for each run of the experiment. For this purpose, box-plots of replicates were obtained. The box plots of replicates were presented in Figure 1.

It can be easily seen from Figure 1 that the most of the experimental units of replicates have skewed distribution. In this case, it was suggested to use robust statistics to compose the dual responses. The replicated response values were transformed to dual responses by using the *MED* and the *MAD* statistics. The experimental data set with robust dual response values are given in Table 5.

Table 4. The printing process data set [33]

Unit	**Input variables**			**Response**		
Number	**X_1**	**X_2**	**X_3**	**Rep1**	**Rep2**	**Rep3**
1	-1	-1	-1	34	10	28
2	0	-1	-1	115	116	130
3	1	-1	-1	192	186	263
4	-1	0	-1	82	88	88
5	0	0	-1	44	178	188
6	1	0	-1	322	350	350
7	-1	1	-1	141	110	86
8	0	1	-1	259	251	259
9	1	1	-1	290	280	245
10	-1	-1	0	81	81	81
11	0	-1	0	90	122	93
12	1	-1	0	319	376	376
13	-1	0	0	180	180	154
14	0	0	0	372	372	372
15	1	0	0	541	568	396
16	-1	1	0	288	192	312
17	0	1	0	432	336	513
18	1	1	0	713	725	754
19	-1	-1	1	364	99	199
20	0	-1	1	232	221	266
21	1	-1	1	408	415	443
22	-1	0	1	182	233	182
23	0	0	1	507	515	434
24	1	0	1	846	535	640
25	-1	1	1	236	126	168
26	0	1	1	660	440	403
27	1	1	1	878	991	1161

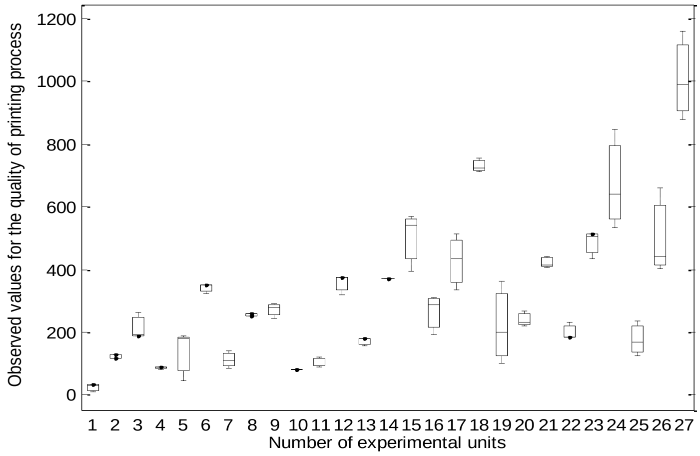

Figure 1. Box-plots of replicated response values for each experimental units

Table 5. The printing process data set with robust dual response values

Unit Number	Input variables			Observations			*MED*	*MAD*
	X_1	X_2	X_3	Rep1	Rep2	Rep3		
1	-1	-1	-1	34	10	28	28	6
2	0	-1	-1	115	116	130	116	1
3	1	-1	-1	192	186	263	192	6
4	-1	0	-1	82	88	88	88	0
5	0	0	-1	44	178	188	178	10
6	1	0	-1	322	350	350	350	0
7	-1	1	-1	141	110	86	110	24
8	0	1	-1	259	251	259	259	0
9	1	1	-1	290	280	245	280	10
10	-1	-1	0	81	81	81	81	0
11	0	-1	0	90	122	93	93	3
12	1	-1	0	319	376	376	376	0
13	-1	0	0	180	180	154	180	0
14	0	0	0	372	372	372	372	0
15	1	0	0	541	568	396	541	27
16	-1	1	0	288	192	312	288	24
17	0	1	0	432	336	513	432	81
18	1	1	0	713	725	754	725	12
19	-1	-1	1	364	99	199	199	100
20	0	-1	1	232	221	266	232	11
21	1	-1	1	408	415	443	415	7
22	-1	0	1	182	233	182	182	0
23	0	0	1	507	515	434	507	8
24	1	0	1	846	535	640	640	105
25	-1	1	1	236	126	168	168	42
26	0	1	1	660	440	403	440	37
27	1	1	1	878	991	1161	991	113

In order to calculate the effects of and X_1, X_2 and X_3 on Y, the second order polynomial regression models, given in Equations (6)-(7) were used for fitting. The calculations were conducted in Minitab14 and Matlab7.9 programs. The analysis of variance (ANOVA) results for the predicted robust dual response models were given in Tables 6-7.

From Table 6, it can be said that the predicted surface model for the *MED* is meaningful for $\alpha = 0.05$ nominal significance level ($p < \alpha$) and can be written as

$$\hat{Y}_{MED}(\mathbf{X}) = 345.89 + 177X_1 + 108.9X_2 + 120.7X_3 + 32X_1^2 - 36.1X_2^2 - 44.5X_3^2 + 62.9X_1X_2 + 75.1X_1X_3 + 36.6X_2X_3. \quad (16)$$

Table 6. ANOVA for the *MED* response

Predictor	Unstandardized Coef.		Standardized Coef.	t	p
	β	Std.Error	β		
Constant	345.89	40.103		8.625	0
X1	177	18.564	0.658	9.535	0
X2	108.9	18.564	0.405	5.869	0
X3	120.7	18.564	0.449	6.503	0
X1^2	32	32.154	0.069	0.995	0.334
X2^2	-36.1	32.154	-0.078	-1.125	0.276
X3^2	-44.5	32.154	-0.096	-1.384	0.184
X1X2	62.9	22.736	0.191	2.767	0.013
X1X3	75.1	22.736	0.228	3.302	0.004
X2X3	36.6	22.736	0.111	1.613	0.125
Source	**Sum of Squares**	**df**	**Mean Square**	**F**	**p**
Regression	1197050.611	9	133005.623	21.442	0
Residual	105454.056	17	6203.180		
Total	1302504.667	26			

In order to check the least squares assumptions of the predicted median model, residual analysis is done. The obtained results are presented in Figure 2. It can be said from Figure 2 that the least squares assumptions on the errors are satisfied. Besides, it is clear from Figure 3 that the normality assumption of errors is also satisfied for statistical inference of model parameters.

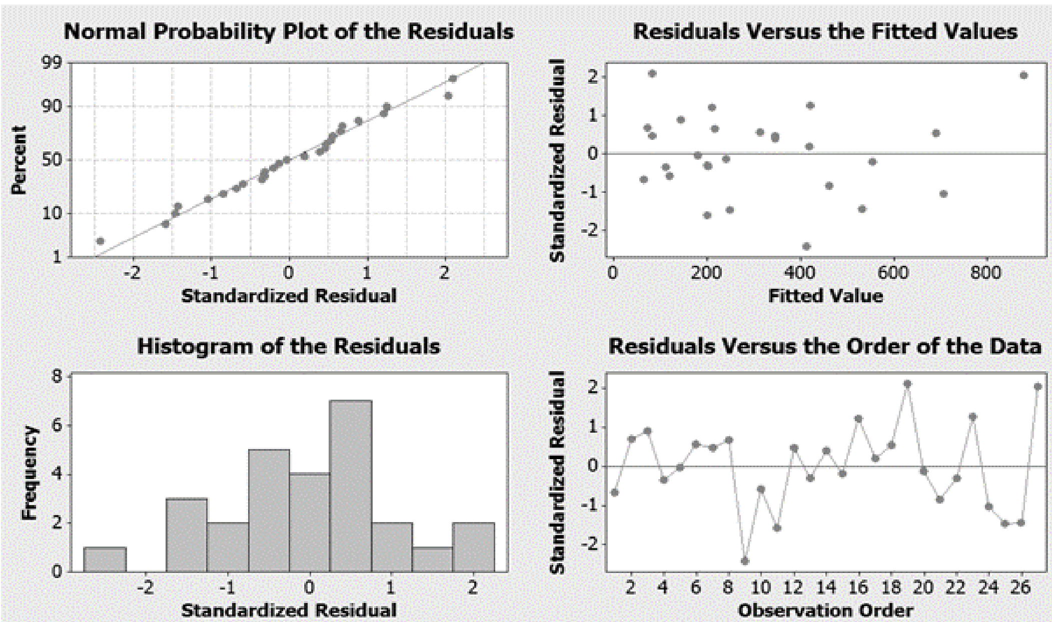

Figure 2. Plots of residuals for the *MED* response

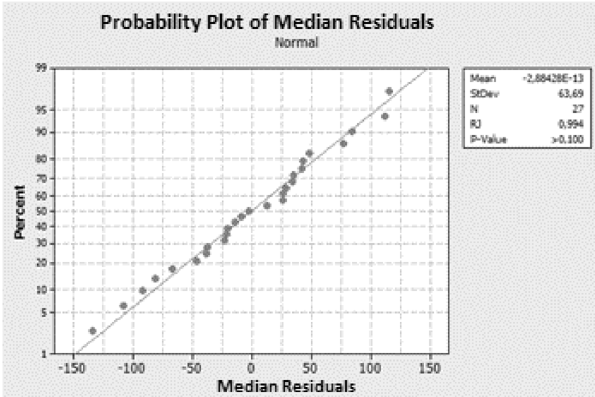

Figure 3. Normal probability plot of residuals for the *MED* response

However, as can be seen from Table 7 that the predicted function for the *MAD* response is meaningless for $\alpha = 0.05$ nominal significance level ($p = 0.18 > \alpha$). The residual analysis for the *MAD* response is presented in Figures 4-5. It is possible to say from the Figures 4-5 that the least squares assumptions of errors are not satisfied and the errors are not normally distributed.

Table 7. ANOVA for the *MAD* response

	Unstandardized Coef.		Standardized Coef.		
Predictor	**γ**	**Std.Error**	**γ**	**t**	**p**
Constant	3.333	16.017		0.208	0.838
X1	4.667	7.414	0.112	0.629	0.537
X2	11.611	7.414	0.278	1.566	0.136
X3	20.333	7.414	0.486	2.742	0.014
X1^2	9.667	12.842	0.133	0.753	0.462
X2^2	9.833	12.842	0.136	0.766	0.454
X3^2	10.333	12.842	0.143	0.805	0.432
X1X2	11.5	9.081	0.225	1.266	0.222
X1X3	8.083	9.081	0.158	0.890	0.386
X2X3	4.417	9.081	0.086	0.486	0.633
Source	**Sum of Squares**	**df**	**Mean Square**	**F**	**p**
Regression	14647.389	9	1627.488	1.645	0.180
Residual	16821.278	17	989.487		
Total	31468.667	26			

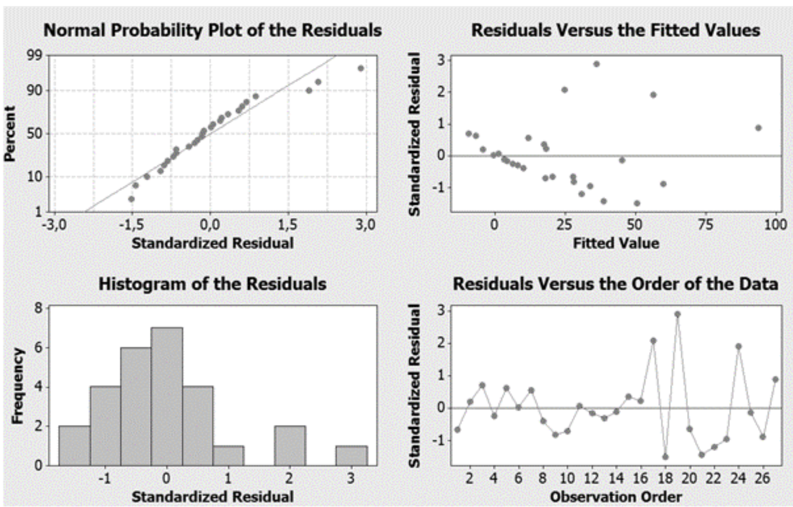

Figure 4. Plots of residuals for the *MAD* response

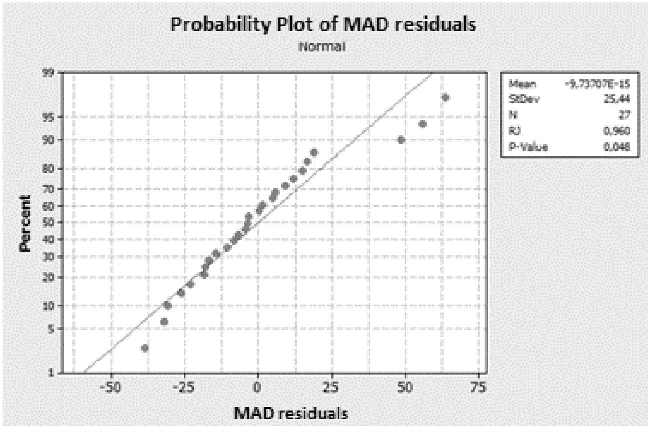

Figure 5. Normal probability plot of residuals for the *MAD* response

To fit a model for the *MAD* response, it was aimed to determine the probability distribution of errors. By using EasyFit 5.6 program, the distribution of errors was defined as three parameterized log-logistic distribution. The probability density function of the log-logistic distribution of errors can be written as

$$f\left(\varepsilon_i;\theta_1,\theta_2,\theta_3\right)=\frac{\theta_1}{\theta_2}\left(\frac{\varepsilon_i-\theta_3}{\theta_2}\right)^{\theta_1-1}\left(1+\left(\frac{\varepsilon_i-\theta_3}{\theta_2}\right)^{\theta_1}\right)^{-2},\ \ \varepsilon_i\in\left[0,\infty\right),\ \theta_1>0,\ \theta_2>0,\ \theta_3\in R \qquad (17)$$

where ε_i 's, $i=1,2,...,27$, are the errors of the *MAD* response, given in Equation (5). The goodness of fit result is presented in Figure 6 with the corresponding parameter values of log-logistic distribution obtained as $\theta_1=4.40$, $\theta_2=57.06$, and $\theta_3=-61.47$.

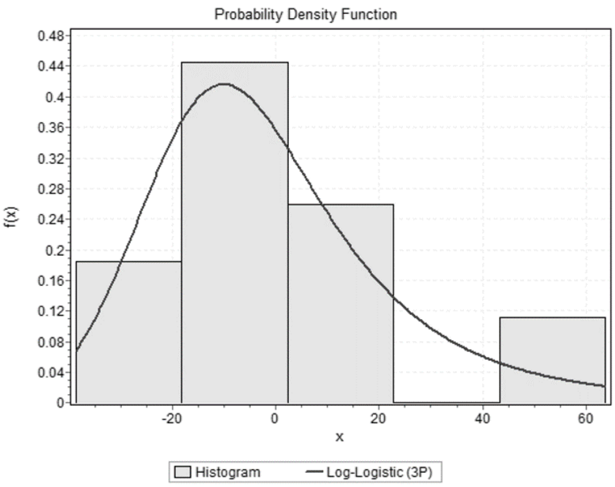

Figure 6. The plot for goodness of fit between residuals and log-logistic distribution

It is possible to use ML method to estimate the unknown model parameters of the *MAD* response. According to the log-logistic distribution, the likelihood function can be written as

$$L(\boldsymbol{\gamma}) = \prod_{i=1}^{27} f\left(\varepsilon_{MAD_i}; 4.40, 57.06, -61.47\right)$$
$$= \left(\frac{4.40}{57.06}\right)^{27} \prod_{i=1}^{27} \left(\frac{\varepsilon_{MAD_i} + 61.47}{57.06}\right)^{91.8} \left(1 + \left(\frac{\varepsilon_{MAD_i} + 61.47}{57.06}\right)^{4.40}\right)^{-54} \tag{18}$$

in which $\varepsilon_{MAD_i} = Y_{MAD_i} - \eta(\boldsymbol{\gamma}, \mathbf{X})$, $i = 1, 2, ..., 27$. The functional form of $\eta(\boldsymbol{\gamma}, \mathbf{X})$ was considered as given in Equation (5). It is possible to calculate the ML estimates of model parameters, $\boldsymbol{\gamma}$, by maximizing the likelihood function given as

$$L(\boldsymbol{\gamma}) = \left(\frac{4.40}{57.06}\right)^{27} \prod_{i=1}^{27} \left(\frac{\left(Y_{MAD_i} - \eta(\boldsymbol{\gamma}, \mathbf{X})\right) + 61.47}{57.06}\right)^{91.8} \left(1 + \left(\frac{\left(Y_{MAD_i} - \eta(\boldsymbol{\gamma}, \mathbf{X})\right) + 61.47}{57.06}\right)^{4.40}\right)^{-54} \tag{19}$$

which seems hard since the $L(\boldsymbol{\gamma})$ is nonlinear in model parameters. In this case, the GANMS hybrid algorithm, defined in the study of [34], is used as an optimization tool to maximize the likelihood function. The tunable parameters of the GANMS are given in Table 8.

Table 8. Tunable parameters values of GANMS hybrid algorithm

Methods	Parameters	Values
	Population size	100
	Maximum number of generation	100
	Probability of crossover	0.80
GA	Probability of mutation	0.01
	Selection operator	Roulette wheel
	Crossover operator	Single point
	Mutation operator	Bit flip
	Reflection	1
	Expansion	2
NMS	Contraction	0.5
	Shrinkage	0.5
	Stopping Criteria	10^{-5}

The predicted *MAD* response is obtained as below:

$$\hat{Y}_{MAD}(\mathbf{X}) = 10.53 + 20.13X_1 + 10.16X_2 + 10.21X_3 + 10.49X_1^2 + 13.98X_2^2 + 8.20X_3^2$$
$$+ 12.14X_1X_2 + 13.43X_1X_3 + 13.78X_2X_3. \tag{20}$$

For the purpose of fair comparison, the second order predicted dual response functions, $\hat{Y}_{MED}$ and $\hat{Y}_{MAD}$ were obtained as the same functional forms for mean-standard deviation dual responses given in the study of [8]. The comparison result of the *MED-MAD* robust dual responses and the mean-standard deviation dual responses were given in Table 9.

Table 9. Comparison of the predicted dual responses

Predicted Model	R^2	R^2_{adj}	***RMSE***	***MAE***	***PRESS***
$\hat{Y}_{MED}$	0.92	0.87	78.76	51.08	338831.38
$\hat{Y}_{\mu}$	0.92	0.88	76.74	50.57	337738.43
$\hat{Y}_{MAD}$	0.60	0.39	27.19	23.65	48534.09
$\hat{Y}_{\sigma}$	0.45	0.16	43.82	28.95	93075.31

From Table 9, according to the performance metrics, it can be said that the $\hat{Y}_{MED}$ and $\hat{Y}_{MAD}$ give better model performance values than the $\hat{Y}_{\mu}$ and $\hat{Y}_{\sigma}$ even the *RMSE* and the *MAE* values of $\hat{Y}_{\mu}$ is slightly smaller than the $\hat{Y}_{MED}$. The PRESS of $\hat{Y}_{MED}$ slightly similar with PRESS of $\hat{Y}_{\mu}$. However, the PRESS value of $\hat{Y}_{MAD}$ is quite smaller than PRESS value of $\hat{Y}_{\sigma}$. It can be easily said from the performance metric results of $\hat{Y}_{MAD}$ that modeling with considering the distribution of errors gives more realistic results in accordance with the nature of the data. After obtaining predicted robust dual response, it was aimed to optimize dual response problem as MOO problem given in Equation (15). Here, the target value of the *MED* response and the search domain of input variables are considered as T =500 and $\mathbf{X} \in [-1, 1]$, respectively. In order to obtain Pareto solution set, the NSGA-II is applied by using tunable parameters as in Table 10.

Table 10. Tunable parameters values of the NSGA-II

Algorithm Parameters	**Values**
Number of variable (v)	3
n_{pop}	100
Selection operator	Tournament
Crossover operator	SBX
Mutation operator	Polynomial
$\Pr_{cross}$	0.90
$\Pr_{mut}$	0.01
n_{gen}	100

The obtained Pareto solution set is presented in Figure 7. From Figure 7, it can be easily said that all the non-dominated solutions, sized 100, have an importance in the Pareto set. However, a researcher may want to get a compromise solution among many non-dominated solutions. For this purpose, in this study, the TOPSIS method was applied to Pareto set with assuming that the objectives have equal importance. The obtained compromise solution is presented in Figure 8 with the value of objective function vector and input vector, $\mathbf{f} = [120.9 \ 14.19]$ and $\mathbf{X} = [0.0858 \ 0.0185 \ 0.1288]$, respectively.

It should be noted here that the relative closeness coefficient, R, is obtained as equal to 0.5778 during the TOPSIS application. This compromise solution is low risk value among the Pareto solutions according to the decision rules given in Table 3.

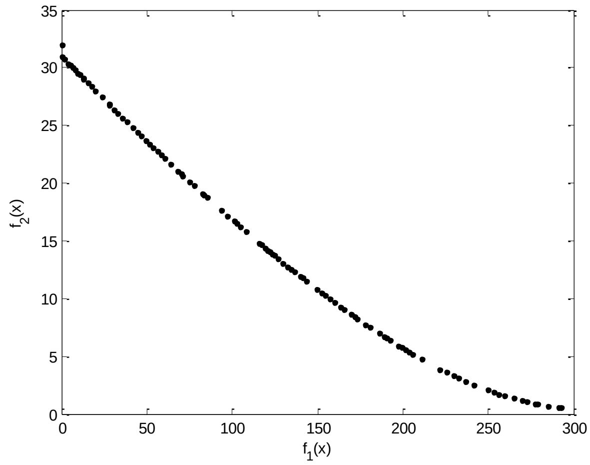

Figure 7. Pareto solution set for robust dual responses

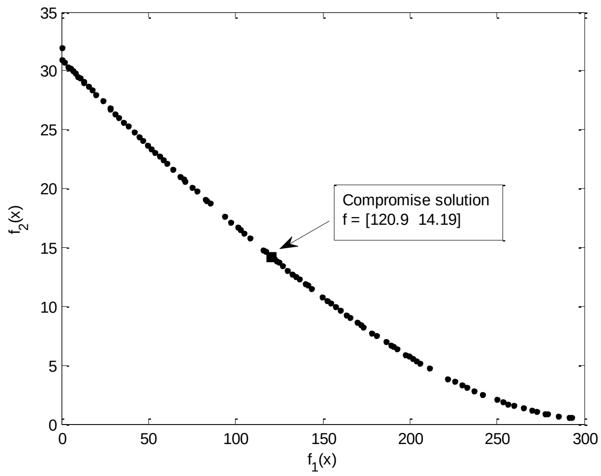

Figure 8. Pareto set with a compromise solution for robust dual responses

For the purpose of comparison, if $\left|\hat{Y}_{\mu}(\mathbf{X})-500\right|$ and $\hat{Y}_{\sigma}(\mathbf{X})$ are considered as $f_1(\mathbf{X})$ and $f_2(\mathbf{X})$, respectively, the Pareto solution set can be obtained as in Figure 9 through the NSGA-II by using the tunable parameters as given in Table 10. The obtained compromise solution is presented in Figure 10 with the value of objective function vector and input vector, $\mathbf{f}=[210.1\ 21.94]$ and $\mathbf{X}=[0.8836\ -1\ -0.1809]$, respectively. It can be said from Figure 8 and Figure 10 that the compromise solution of robust dual response model is more preferable than the compromise solution of mean-standard deviation dual response model.

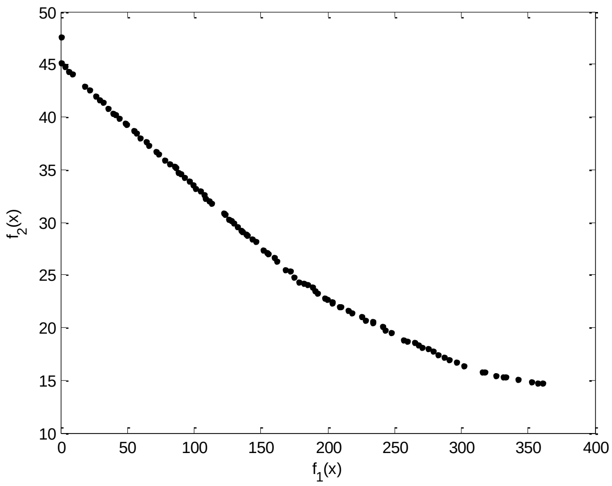

Figure 9. Pareto solution set for mean-standard deviation dual response

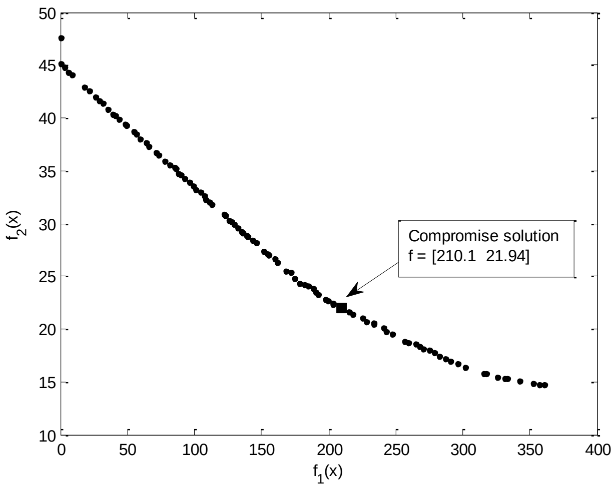

Figure 10. Pareto set with a compromise solution for mean-standard deviation dual response

5. CONCLUSION

In this study, modeling and optimization stages of replicated response measured data set were presented in robust and multi-objective perspectives, respectively. In modeling stage, replicated response measures were presented as dual response by using robust statistics, the *MED* and the *MAD*. The fitting performance of the predicted second order polynomial robust dual responses were analyzed according to the several model performance metrics, e.g. *RMSE*, *MAE*, *PRESS*. To optimize robust dual response, a multi-objective optimization algorithm, called NSGA-II, was applied and many alternative input settings were obtained as Pareto front. A compromise solution was chosen among many non-dominated input settings by using TOPSIS method. A case example from literature, called printing ink data set, was carried out to present the performance of the proposed robust dual response model which is encouraging with the most satisfactory input settings.

REFERENCES

[1] Myers RH, Carter WH. Response surface techniques for dual response systems. Technometrics 1973; 15(2): 301-317.

[2] Vining GG, Myers RH. Combining Taguchi and response surface philosophies: a dual response approach. Journal of Quality Technology 1990; 22: 38-45.

[3] Lin DKJ, Tu W. Dual Response Surface Optimization. Journal of Quality Technology 1995; 27(1): 34-39.

[4] Copeland KAF, Nelson PR. Dual Response Optimization via Direct Function Minimization. Journal of Quality Technology 1996; 28(3): 331-336.

[5] Del Castillo E. Multiresponse Process Optimization via Constrained Confidence Regions. Journal of Quality Technology 1996; 28(1): 61-70.

[6] Kim KJ, Lin DKJ. Dual Response Surface Optimization: A Fuzzy Modeling Approach. Journal of Quality Technology 1998; 30(1): 1-10.

[7] Khoo LP, Chen CH. Integration of response surface methodology with genetic algorithms. The International Journal of Advanced Manufacturing Technology 2001; 18(7): 483-489.

[8] Ding R, Lin DK, Wei D. Dual-response surface optimization: A weighted MSE approach. Quality engineering 2004; 16(3): 377-385.

[9] Jeong IJ, Kim KJ, Chang SY. Optimal weighting of bias and variance in dual response surface optimization. Journal of Quality Technology 2005; 37(3): 236-247.

[10] Shaibu AB, Cho BR. Another view of dual response surface modeling and optimization in robust parameter design. The International Journal of Advanced Manufacturing Technology 2009; 41: 631-641.

[11] Lee DH, Jeong IJ, Kim KJ. A posterior preference articulation approach to dual-response-surface optimization. IIE Transactions 2010; 42: 161–171.

[12] Baba I, Midi H, Rana S, Ibragimov G. An Alternative Approach of Dual Response Surface Optimization Based on Penalty Function Method. Mathematical Problems in Engineering 2015; 1-6.

[13] Le TH, Shin S. A literature review on RSM-based robust parameter design (RPD): Experimental design, estimation modeling, and optimization methods. Journal of the Korean Society for Quality Management 2018; 46(1): 39-74.

[14] Park C, Cho B. Development of Robust Design Under Contamined and Non-normal Data. Quality engineering 2003; 15(3): 463-469.

[15] Lee SB, Park C, Cho BR. Development of a highly efficient and resistant robust design. International Journal of Production Research 2007; 45(1): 157-167.

[16] Deb K, Pratap A, Agarwal S, Meyarivan T. A Fast and Elitist Multiobjective Genetic Algorithm: NSGA-II. IEEE Transactions on Evolutionary Computation 2002; 6(2): 182-197.

[17] Unune DR, Nirala CK, Mali HS. ANN-NSGA-II dual approach for modeling and optimization in abrasive mixed electro discharge diamond grinding of Monel K-500. Engineering Science and Technology, an International Journal 2018; 21: 322-329.

[18] Han H, Yu R, Li B, Zhang Y. Multi-objective optimization of corrugated tube inserted with multi-channel twisted tape using RSM and NSGA-II. Applied Thermal Engineering 2019; 159: 113731.

[19] Chen SJ, Hwang CL. Fuzzy multiple attribute decision making methods. In: Fuzzy multiple attribute decision making. Berlin, Heidelberg: Springer, 1992. pp. 289-486.

[20] Hwang CL, Yoon K. Multiple Attribute Decision Making. In: Lecture Notes in Economics and Mathematical Systems. Berlin: Springer-Verlag, 1981.

[21] Khuri AI, Cornell JA. Response surfaces: designs and analyses. CRC press 1996; 152.

[22] Ames AE, Mattucci N, Macdonald S, Szonyi G, Hawkins DM. Quality loss functions for optimization across multiple response surfaces. Journal of Quality Technology 1997; 29(3): 339-346.

[23] Del Castillo E, Montgomery DC. A Nonlinear Programming Solution to the Dual Response Problem. Journal of Quality Technology 1993; 25(3): 199-204.

[24] Cho BR, Philips MD, Kapur KC. Quality improvement by RSM modeling for robust design. In: The 5th Industrial Engineering Research Conference; 1996; Minneapolis. pp. 650-655.

[25] Köksoy O, Doğanaksoy N. Joint Optimization of Mean and Standard Deviation Using Response Surface Methods. Journal of Quality Technology 2003; 35(3): 239-252.

[26] Shin S, Cho BR. Studies on a biobjective robust design optimization problem. IIE Transactions 2009; 41(11): 957-968.

[27] Truong NKV, Shin S. Development of a new robust design methodology based on Bayesian perspectives. International Journal of Quality Engineering and Technology 2012; 3(1): 50-78.

[28] Nha VT, Shin S, Jeong SH. Lexicographical dynamic goal programming approach to a robust design optimization within the pharmaceutical environment. European Journal of Operational Research 2013; 229(2): 505-517.

[29] Jeong IJ, Lee DH. Generating evenly distributed nondominated solutions in dual response surface optimization. Quality Technology & Quantitative Management 2017; 1-17.

[30] Lee DH, Kim BR, Yang JK, Oh SH. Dual Response Surface Optimization using Multiple Objective Genetic Algorithms. Journal of the Korean Institute of Industrial Engineers 2017; 43(3): 164-175.

[31] Park C, Leeds M. A highly efficient robust design under data contamination. Computers & Industrial Engineering 2016; 93: 131-142.

[32] Chen CT, Lin CT, Huang SF. A fuzzy approach for supplier evaluation and selection in supply chain management. International Journal of Production Economics 2006; 102: 289-301.

[33] Box GEP, Draper NR. Wiley series in probability and mathematical statistics. Empirical model-building and response surfaces. Oxford: John Wiley & Sons, 1987.

[34] Türkşen Ö, Tez M. An application of nelder-mead heuristic-based hybrid algorithms: estimation of compartment model parameters. International Journal of Artificial Intelligence 2016; 14(1): 112-129.