An Improved Robust Regression Model for Response Surface Methodology

Efosa Edionwe[1,*], Mbegbu J.I.[2], Ekhosuehi N.[2] and Obiora-Ilouno H.O.[3]

[1]*Department of Mathematical Sciences, Edwin Clark University, Delta State, Nigeria.*
E-mail: ⟨trustnelson24@yahoo.com⟩

[2]*Department of Statistics, University of Benin, Benin City, Nigeria*
Email: ⟨{mbegbu, ekhosuehi}@uniben.edu⟩

[3]*Department of Statistics, Nnamdi Azikiwe University, Awka, Nigeria*
Email: ⟨obiorailounho@yahoo.com⟩

Abstract. In production, manufacturing and several other allied industries, it is important to apply appropriate tools for the analysis of data in order to enhance the opportunity for product and process optimization. A statistical tool that has successfully been used to achieve this goal is Response Surface Methodology (RSM). A recent trend in the modeling phase of RSM involves the use of semi-parametric regression models which are hybrids of the Ordinary Least Squares (OLS) and the Local Linear Regression (LLR) models. In this paper, we propose a modification in the current structure of the semi-parametric Model Robust Regression 2 (MRR2) with a view to improving its sensitivity to local trends and patterns in data. The proposed model is applied to two multiple response optimization problems from the literature. The results of goodness-of-fits and optimal solutions confirm that the proposed model performs better than the MRR2.

Keywords: desirability function, genetic algorithm, local linear regression, multiple response optimization problem, semi-parametric regression models

Received: August 1, 2018; accepted: November 11, 2018; available online: December 13, 2018

DOI: 10.17535/crorr.2018.0025

1. Introduction

Response Surface Methodology (RSM) is a collection of statistical techniques in the modeling and analysis of data in which a response is influenced by one or more explanatory variables [4, 6].

Three phases stand out in RSM, namely, the design of experiment phase, the Modeling phase, and the optimization phase [22, 31].

In the modeling phase, we assume that the relationship between the response variable, y and k explanatory variables $x_{1,}\ x_2, \dots, x_k$, takes the form:

$$\boldsymbol{y} = f\left(x_{1,}\ x_2, \dots, x_k\right) + \boldsymbol{\varepsilon}, \quad i = 1,2, \dots, n \tag{1}$$

[*] Corresponding author.

where $\boldsymbol{y}$ is a $1 \times n$ vector of responses, f represents the true but unknown relationship between the response variable and the k explanatory variables, $\boldsymbol{\varepsilon}$is a $1 \times n$ vector of random errors assumed to have a normal distribution with mean zero and constant variance, $n$is the sample size [23, 29].

Existing regression models applied in the estimation of the unknown function f in (1) include the parametric regression models (e.g. the OLS), the nonparametric regression models (e.g. LLR), and the semi-parametric regression models (e.g. the MRR1 and MRR2 models) [2, 19, 14, 25].

The aim of the optimization phase of RSM is to obtainthe setof values of the explanatory variables (optimal setting),$\boldsymbol{x} = [x_1, x_2, \dots, x_k]$that optimizes the fitted regression model based on the production requirements for the study [17, 21].

For studies that involve $m$multiple response,$m > 1$,it is essential that we get asetting of the explanatory variables that optimize a real valued cost function based on the production requirements of all the responses [15, 28, 30].A common real valued cost function that is applied in the optimization of multiple response is the desirability function [1, 7, 13]

Desirability function assigns $0 \leq d_p\left(\hat{y}_p(\boldsymbol{x})\right) \leq 1$ to each estimated response $\hat{y}_p(\boldsymbol{x})$. Depending on whether a particular response is to be assigned a target value, maximized, or minimized, different desirability function can be used:

For the nominal-the-better (NTB) response where the pth response acceptable value lies between an upper limit, U_p and a lower limit, L_p, $d_p\left(\hat{y}_p(\boldsymbol{x})\right)$ is given as:

$$d_p\left(\hat{y}_p(\boldsymbol{x})\right) = \begin{cases} \left\{\frac{\hat{y}_p(x)-L_p}{\emptyset_p-L_p}\right\}^{t_1}, & L_p \leq \hat{y}_p(\boldsymbol{x}) \leq \emptyset_p, \\ \left\{\frac{U_p-\hat{y}_p(x)}{U_p-\emptyset_p}\right\}^{t_2}, & \emptyset_p \leq \hat{y}_p(\boldsymbol{x}) \leq U_p, \\ 0\ , & otherwise \end{cases} \tag{2}$$

where $\emptyset_p$ is the target value of the p^{th}response.

For the larger-the-better (LTB) response, where the objective is to maximize the p^{th} response, $d_p\left(\hat{y}_p(\boldsymbol{x})\right)$ is given by a one-sided transformation as:

$$d_p\left(\hat{y}_p(\boldsymbol{x})\right) = \begin{cases} 0 & \hat{y}_p(\boldsymbol{x}) < L_p, \\ \left\{\frac{\hat{y}_p(x)-L_p}{\emptyset_p-L_p}\right\}^{t_1} & L_p \leq \hat{y}_p(\boldsymbol{x}) \leq \emptyset_p, \\ 1 & \hat{y}_p(\boldsymbol{x}) > \emptyset_P, \end{cases} \tag{3}$$

where $\emptyset_p$ is interpreted as a large enough value of the p^{th} response.

When the response is of the smaller-the-better (STB) type, the objective is to minimize the p^{th} response and $d_p\left(\hat{y}_p(\boldsymbol{x})\right)$ is given by a one-sided transformation as:

$$d_p\left(\hat{y}_p(\boldsymbol{x})\right) = \begin{cases} 1 & \hat{y}_p(\boldsymbol{x}) < \emptyset_p, \\ \left\{\frac{U_p-\hat{y}_p(x)}{U_p-\emptyset_p}\right\}^{t_2} & \emptyset_p \leq \hat{y}_p(\boldsymbol{x}) \leq U_p, \\ 0 & \hat{y}_p(\boldsymbol{x}) > U_p, \end{cases} \tag{4}$$

where $\emptyset_p$ denotes a small enough value of the pth response.

For RSM data, the values of t_1 and $t_2$are taken to be 1 [5, 15 28]. The m individual desirabilities are then combined using the geometric mean into a single scalar value given as:

$$D(\boldsymbol{x}) = maximize\left(\left(\prod_{p=1}^{m} d_p\left(\hat{y}_p(\boldsymbol{x})\right)\right)^{1/m}\right), \tag{5}$$

The goal of the desirability function reduces to maximizing $D(\boldsymbol{x})$ with respect to the $\boldsymbol{x}$.

The GA is an evolutionary optimization tool that can be applied to variety of optimization problems including those that are not well suited for standard optimization algorithms [16, 31]. Examples of such optimization problems include those in which the objective functions do not have closed form expressions [25, 29]. The nonparametric and semi-parametric regression models pose this kind of problem. Hence, the search for $\boldsymbol{x}$ that optimizes $D(\boldsymbol{x})$ in (5),b^*, N^* and C^* that optimize $PRESS^{**}$ in (12) and λ^*that optimizes $PRESS^{**}$in (15) are achieved using the GA optimization toolbox in Matlab.

The remainder of the paper is organized as follows: A review of existing regression models in RSM is presented in Section 2. In Section 3, the proposed modification to the current structure of the MRR2 is presented. Using two examples from the literature and a simulation study, comparisons of results from the proposed model and those from the MRR2 model are presented in Section 4.

The paper concludes in Section 5.

2. A Review of Regression Models Applied in RSM

Regression analysis is a statistical procedure for estimating a mean function of dependent (response) variable using either a bivariate or multivariate paired data. Such an estimated mean function is used for prediction [2, 24].Below is an overview of existing regression models applied in RSM

2.1. The Ordinary Least Squares(OLS) Model

The OLS is applied for estimating the unknown parameters in the model that the experimenter assumes adequate in the approximation of the unknown function f[22, 23].

The OLS estimate $\hat{y}_i^{(OLS)}$ response in the i^{th}data point is given as:

$$\hat{y}_i^{(OLS)} = \boldsymbol{x}_i(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{y} = \boldsymbol{h}_i^{(OLS)}\boldsymbol{y}, \tag{6}$$

where $\boldsymbol{y}$ is $n \times 1$ vector of response, $\mathbf{X}$ is$\mathrm{n} \times \mathrm{p}$ model matrix,p is the number of model parameters,$\boldsymbol{X}^T$is the transpose of the matrix$\mathbf{X}$, and $\boldsymbol{x}_i$ is the i^{th}row vector of $\boldsymbol{X}$[29].

Using matrix notation, the vector of OLS estimated response can be expressed as:

$$\hat{\boldsymbol{y}}^{(OLS)} = \boldsymbol{H}^{(OLS)}\boldsymbol{y} = \begin{bmatrix} \boldsymbol{h}_1^{(OLS)} \\ \boldsymbol{h}_2^{(OLS)} \\ \vdots \\ \boldsymbol{h}_n^{(OLS)} \end{bmatrix}\boldsymbol{y}, \tag{7}$$

where the $1 \times n$ vector $\boldsymbol{h}_i^{(OLS)}$ is the i^{th} row of the $n \times n$ OLS Hat matrix,

A limitation of the OLS model is that it performs poorly if the assumed model is misspecified or inadequate for the data [25, 29].

2.2. The Local Linear Regression (LLR)Model

The LLR model is a nonparametric regression version of the weighted least squares model. The weights utilized in the LLR model are derived from one of the several kernel functions such as the Gaussian kernel [2, 9, 10].

The LLR estimate, $\hat{y}_i^{(LLR)}$of y_i, is given as:

$$\hat{y}_i^{(LLR)} = \tilde{\boldsymbol{x}}_i(\tilde{\boldsymbol{X}}^T\boldsymbol{W}_i^{**}\tilde{\boldsymbol{X}})^{-1}\tilde{\boldsymbol{X}}^T\boldsymbol{W}_i^{**}\boldsymbol{y} = \boldsymbol{h}_i^{(LLR)}\boldsymbol{y}, \tag{8}$$

where $\tilde{\boldsymbol{x}}_i$is the i^{th} row of the LLR model matrix $\tilde{\boldsymbol{X}}$ given as:

$$\widetilde{\boldsymbol{X}} = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & & x_{nk} \end{bmatrix}_{n\times(k+1)},$$

where $x_{ij}, i = 1,2,\ldots,n, j = 1,2,\ldots,k,$ denotes the value of the j^{th} explanatory variable in the i^{th}data point, $\boldsymbol{W}_i^{**}$ is an $n \times n$ diagonal weight matrix used for estimation of the i^{th} response [28].

For instance, the first entry say w_1^{**} of W_i^{**}, $i = 1$,is obtained from the product kernel as:

$$w_1^{**} = \prod_{j=1}^{k} K\left(\frac{x_{ij}-x_{1j}}{b_i}\right) / \sum_{i=1}^{n} \prod_{j=1}^{k} K\left(\frac{x_{ij}-x_{1j}}{b_i}\right), i = 1,2,\ldots,n, j = 1,2,\ldots,k, \tag{9}$$

where $K\left(\frac{x_{ij}-x_{1j}}{b_i}\right) = e^{-\left(\frac{x_{ij}-x_{1j}}{b_i}\right)^2}$ is the simplified Gaussian kernel function andb_i, $0 < b_i \leq 1, i = 1,2,\ldots,n$, are the locally adaptive smoothing parameters or bandwidths [18, 25, 26].

Smoothing in nonparametric regression is said to be done using fixed or global smoothing parameter $b$if, in (9), we have $b_1 = b_2 = \cdots = b_n = b$, otherwise $b_i, i = 1,2,\ldots,n,$ are referred to as variable or locally adaptive smoothing parameters [3, 26]. Locally adaptive smoothing parameters have been found to perform better than their fixed counterpart because of their comparatively better sensitivity to local trends and patterns in the data[9, 32].

The locally adaptive smoothing parameter selector presented in [8] is given as:

$$b_i = \frac{b^* N(TC - y_i)}{T(Cn-1)}, \qquad i = 1,\ldots,n \tag{10}$$

where b^* is the fixed optimal smoothing parameter selected based on the minimization of the Penalized Prediction Error Sum of Squares ($PRESS^{**}$), $T = \sum_i^n y_i$, $C \geq 0,\ N > 0$ are data-driven tuning parameters.

In matrix notation, the LLR estimates of the response can be expressed as:

$$\hat{\boldsymbol{y}}^{(LLR)} = \boldsymbol{H}^{(LLR)}\boldsymbol{y} = \begin{bmatrix} \boldsymbol{h}_1^{(LLR)} \\ \boldsymbol{h}_2^{(LLR)} \\ \vdots \\ \boldsymbol{h}_n^{(LLR)} \end{bmatrix} \boldsymbol{y}, \tag{11}$$

where $\boldsymbol{H}^{(LLR)}$ is the $n \times n$LLR Hat matrix, and $\boldsymbol{h}_i^{(LLR)} = \widetilde{\boldsymbol{x}}_i(\widetilde{\boldsymbol{X}}^T\boldsymbol{W}_i^{**}\widetilde{\boldsymbol{X}})^{-1}\widetilde{\boldsymbol{X}}^T\boldsymbol{W}_i^{**}$is the i^{th} row vector of the LLR Hat matrix estimating y_i [20].

The smoothing parameter is the most crucial parameter in nonparametric regression procedure because the choice selected determines the shape of the fitted curve [12].

For RSM, the set of optimal smoothing parameters $\Phi = [b_1^*,\ b_2^*,\ldots,b_n^*]$ from locally adaptive smoothing parameters selector in (10) are derived from the optimal values C^* and N^* of C and N, respectively, based on the minimization of the $PRESS^{**}$ [8].

The form of the $PRESS^{**}$ criterion for selecting the smoothing parameters is given as:

$$PRESS^{**}(b_1, b_2, \ldots, b_n) = \frac{\sum_{i=1}^{n}\left(y_i - \hat{y}_{i,-i}^{(LLR)}\right)^2}{n - trace\left(H^{(LLR)}(\Phi)\right) + (n-k-1)\frac{SSE_{max} - SSE_{\Phi}}{SSE_{max}}}, \tag{12}$$

where SSE_{max} is the maximum Sum of Squared Errors obtained as$b_1, b_2, \ldots, b_n$ tend to infinity, SSE_{Φ} is the Sum of Squared Errors associated with a set of smoothing parameters$b_1, b_2, \ldots, b_n$, $tr\left(H^{(LLR)}\Phi\right)$ is the trace of the LLR Hat matrix and $\hat{y}_{i,-i}^{(LLR)}$ is the leave-one-out cross-validation estimate of y_i with the i^{th} observation left out [19, 25, 29].

The goal is to minimize $PRESS^{**}$ with respect to $(b_1, b_2, \ldots, b_n)$. $PRESS^{**}$ was developed as an alternative to Prediction ErrorSum of Squares (PRESS) which tends to overfit the data [20, 25, 29].

The LLR model is flexible and can capture local trends which may be overlooked by the OLS model. However, its performance is generally poor when applied in studies that involve$k > 1$ explanatory variables.This poor performance is referred to as'curse of dimensionality' in the nonparametric regression literature [11].

2.3. Semi-parametric Regression Models

Semi-parametric regression models are ideal in situations where a researcher has partial knowledge of the model that can be used to estimate$f$ in (1). However, this model is deficient in the capture of local trends in the entire range of the data points [20, 29].

An overview of the Model Robust Regression 1 (MRR1) and the Model Robust Regression 2 (MRR2) models is presented in the subsections that follow.

2.3.1. The MRR1 Model

The MRR1 model is a convex combination of the OLS and the LLR models given in (6) and (8), respectively, via a mixing parameter, λ, $0 \leq \lambda \leq 1$ [19, 25].

The MRR1 estimate, $\hat{y}_i^{(MRR1)}$ of y_i is given as:

$$\hat{y}_i^{(MRR1)} = (1-\lambda)\boldsymbol{x_i}(\boldsymbol{X^T X})^{-1}\boldsymbol{X^T y} + \boldsymbol{\lambda \tilde{x}_i}(\boldsymbol{\tilde{X}^T W_i^{**} \tilde{X}})^{-1}\boldsymbol{\tilde{X}^T W_i^{**} y}, \tag{13}$$

In matrix notation, the vector of MRR1 estimates of the responses may be expressed as:

$$\hat{\boldsymbol{y}}^{(MRR1)} = \boldsymbol{H}^{(\boldsymbol{MRR1})}\boldsymbol{y} = \begin{bmatrix} \boldsymbol{h}_1^{(\boldsymbol{MRR1})} \\ \boldsymbol{h}_2^{(\boldsymbol{MRR1})} \\ \vdots \\ \boldsymbol{h}_n^{(\boldsymbol{MRR1})} \end{bmatrix} \boldsymbol{y}, \tag{14}$$

where the$1 \times n$ vector, $\boldsymbol{h}_i^{(\boldsymbol{MRR1})} = (1-\lambda)\boldsymbol{x_i}(\boldsymbol{X^T X})^{-1}\boldsymbol{X^T} + \boldsymbol{\lambda \tilde{x}_i}(\boldsymbol{\tilde{X}^T W_i^{**} \tilde{X}})^{-1}\boldsymbol{\tilde{X}^T W_i^{**}}$is the i^{th} row of the MRR1 Hat matrix, $\boldsymbol{H}^{(MRR1)}$[25].

The optimal value $\lambda^*$of λ, may be selected based on the minimization of the form of the $PRESS^{**}$ criterion given as:

$$minimize\ PRESS^{**}(\lambda) = \frac{\sum_{i=1}^{n}\left(y_i - \hat{y}_{i,-i}^{(MRR1)}\right)^2}{n - tr\left(H^{(MRR1)}(\Phi,\lambda)\right) + (n-k-1)\frac{SSE_{max} - SSE_\Phi}{SSE_{max}}}, \tag{15}$$

where Φ=$[b_1^*, b_2^*, \ldots, b_n^*]$ denotes optimal smoothing parameters, SSE_Φ is the Sum of Squared Errors associated with the set of the optimal smoothing parameters,$[b_1^*, b_2^*, \ldots, b_n^*]$, $tr\left(H^{(MRR1)}(\Phi,\lambda)\right)$ is the trace of MRR1 Hat matrix, and $\hat{y}_{i,-i}^{(MRR1)}(\Phi,\lambda)$ is the leave-one-out cross-validation MRR1 estimate of y_i [19, 20, 25].

A computer program for a regression model coupled into the GA toolbox in Matlab obtains the optimal values of Φ in (10) (and, by extension,λ^*) based on the minimization of $PRESS^{**}$.

A drawback in the application of the MRR1 model is the problem associated with convex combination of two quantities or functions. Notice that for $0 \leq \lambda \leq 1$, min $[\hat{y}_i^{(OLS)}, \hat{y}_i^{(LLR)}] \leq \hat{y}_i^{(MRR1)} \leq$ max $[\hat{y}_i^{(OLS)}, \hat{y}_i^{(LLR)}]$. Hence, in data points where both the OLS and the LLR estimates of the response are either larger or smaller than the true value of the response from (1), the MRR1 estimates will be larger or smaller, respectively, than the true value of the response from (1) [20, 25].

2.3.2. MRR2 Model

The MRR2 model was proposed by [19]. It combines both the OLS estimates of the response and LLR estimates of the OLS residuals via a mixing parameter, λ, $0 \leq \lambda \leq 1$.

The MRR2 estimate, $\hat{y}_i^{(MRR2)}$of y_i, is given as:

$$\hat{y}_i^{(MRR2)} = \boldsymbol{x_i}(\boldsymbol{X^T X})^{-1}\boldsymbol{X^T y} + \lambda\tilde{\mathbf{x}}_i(\tilde{\boldsymbol{X}}^T\boldsymbol{W}_i^*\tilde{\boldsymbol{X}}^T)^{-1}\tilde{\boldsymbol{X}}^T\boldsymbol{W}_i^*(\boldsymbol{I} - \boldsymbol{X}(\boldsymbol{X^T X})^{-1}\boldsymbol{X^T})\boldsymbol{y},$$

$$= \boldsymbol{x_i}(\boldsymbol{X^T X})^{-1}\boldsymbol{X^T y} + \lambda\tilde{\mathbf{x}}_i\left(\tilde{\boldsymbol{X}}^T\boldsymbol{W}_i^*\tilde{\boldsymbol{X}}\right)^{-1}\tilde{\boldsymbol{X}}^T\boldsymbol{W}_i^*\boldsymbol{r}^{(\boldsymbol{OLS})}, \tag{16}$$

where $\boldsymbol{r}^{(\boldsymbol{OLS})}$ is the $n \times 1$ vector of the OLS residuals,$\boldsymbol{I}$ is an $n \times n$ identity matrix and $\boldsymbol{W}_i^*$is a $n \times n$ diagonal weights matrix for estimating the i^{th}OLS residual [29].

Equation (16) may be expressed in matrix form as:

$$\boldsymbol{y}^{(MRR2)} = \begin{bmatrix} \boldsymbol{h}_1^{(MRR2)} \\ \boldsymbol{h}_2^{(MRR2)} \\ \vdots \\ \boldsymbol{h}_n^{(MRR2)} \end{bmatrix}\boldsymbol{y} = \boldsymbol{H}^{(\boldsymbol{MRR2})}\boldsymbol{y}, \tag{17}$$

where $\boldsymbol{h}_i^{(\boldsymbol{MRR2})} = \boldsymbol{x_i}(\boldsymbol{X^T X})^{-1}\boldsymbol{X^T} + \lambda\tilde{\mathbf{x}}_i(\tilde{\boldsymbol{X}}^T\boldsymbol{W}_i^*\tilde{\boldsymbol{X}}^T)^{-1}\tilde{\boldsymbol{X}}^T\boldsymbol{W}_i^*(\boldsymbol{I} - \boldsymbol{X}(\boldsymbol{X^T X})^{-1}\boldsymbol{X^T})$ is the i^{th} row of the $n \times n$ MRR2 Hat matrix, $\boldsymbol{H}^{(MRR2)}$[20, 25, 29].The optimal values of Φ and$\lambda^*$are selected using (12) and (15), respectively, with MRR2 replacing LLR in (12) and MRR1 in (15).

The MRR2 is robust, flexible, and presently considered a better regression model than the OLS, LLR and the MRR1 for data emanating from response surface studies [25, 29]. However, it enjoys only a fractional advantage of the flexibility that the LLR model can offer since only the OLS residuals are considered by its LLR component.

3. Methodology (Modification of the MRR2 Model)

We propose a modification in the structure and configuration of the MRR2 model in order to accomplish a blend of the philosophy of both the MRR1 and the MRR2 models.

In order to achieve our purpose, the following objectives are considered:

(i) The inclusion of a LLR component that estimates the observed (raw) response just as in the case of the MRR1 model in (13) in order to fully utilize the flexibility of the LLR model. This provides better opportunity for capturing more local trends and pattern in the data.

(ii) The adoption of'smoothing of residuals' in the MRR2 model in (16) in order to retain the advantage of MRR2 over the MRR1 model. However, unlike the MRR2 model that uses only the residuals from the OLS estimates, the proposed model utilizes the combined residuals obtained from both the OLS and the LLR estimates of response.

The expression for the combined residuals is given as:

$$r_i^{(OLS+LLR)} = y_i - (\hat{y}_i^{(OLS)} + \hat{y}_i^{(LLR)}), \qquad i = 1,2,\dots,n, \tag{18}$$

For ease of reference, we designate the proposed model as the Modified MRR2 (MMRR2) model. Based on the objectives, the estimate, $\hat{y}_i^{(MMRR2)}$of$y_i, i = 1,2,\dots,n,$ is given as:

$$\hat{y}_i^{(MMRR2)} = \boldsymbol{x_i}(\boldsymbol{X^T X})^{-1}\boldsymbol{X^T y} + \tilde{\mathbf{x}}_i\left(\tilde{\boldsymbol{X}}^T\boldsymbol{W}_i^{**}\tilde{\boldsymbol{X}}\right)^{-1}\tilde{\boldsymbol{X}}^T\boldsymbol{W}_i^{**}\boldsymbol{y} +$$

$$+\lambda\tilde{\mathbf{x}}_i(\tilde{\boldsymbol{X}}^T\boldsymbol{W}_i\tilde{\boldsymbol{X}})^{-1}\tilde{\boldsymbol{X}}^T\boldsymbol{W}_i\left[\boldsymbol{I} - (\boldsymbol{X}(\boldsymbol{X^T X})^{-1}\boldsymbol{X^T} + \tilde{\boldsymbol{X}}\left(\tilde{\boldsymbol{X}}^T\boldsymbol{W}_i^{**}\tilde{\boldsymbol{X}}\right)^{-1}\tilde{\boldsymbol{X}}^T\boldsymbol{W}_i^{**})\right]\boldsymbol{y},$$

$$= \boldsymbol{x_i}(\boldsymbol{X^T X})^{-1}\boldsymbol{X^T y} + \tilde{\mathbf{x}}_i(\tilde{\boldsymbol{X}}^T\boldsymbol{W}_i^{**}\tilde{\boldsymbol{X}})^{-1}\tilde{\boldsymbol{X}}^T\boldsymbol{W}_i^{**}\mathbf{y} + \lambda\tilde{\mathbf{x}}_i\left(\tilde{\boldsymbol{X}}^T\boldsymbol{W}_i\tilde{\boldsymbol{X}}\right)^{-1}\tilde{\boldsymbol{X}}^T\boldsymbol{W}_i\boldsymbol{r}^{(\boldsymbol{OLS+LLR})},$$

$$= \boldsymbol{h}_i^{(OLS)}\boldsymbol{y} + \boldsymbol{h}_i^{(LLR)}\boldsymbol{y} + \boldsymbol{\lambda}\boldsymbol{h}_i^{(LLR)}\boldsymbol{r}^{(\boldsymbol{OLS+LLR})}, \tag{19}$$

where $\mathbf{I}$ is the $n \times n$ identity matrix, the $\boldsymbol{h}_i^{(OLS)}\boldsymbol{y}$ and $\boldsymbol{h}_i^{(LLR)}\boldsymbol{y}$ are OLS and LLR estimates of i^{th}response, as in (6) and (8), respectively, and the $\boldsymbol{h}_i^{(LLR)}\boldsymbol{r}^{(\boldsymbol{OLS+LLR})}$ is the LLR estimateof the i^{th} combined residual, $\boldsymbol{r}^{(OLS+LLR)}$, as in (18), $i = 1,2, \dots, n$.

Equation (19) may be expressed in matrix form as:

$$\widehat{\boldsymbol{y}}^{(MMRR2)} = \begin{bmatrix} \boldsymbol{h}_1^{(MMRR2)} \\ \boldsymbol{h}_2^{(MMRR2)} \\ \vdots \\ \boldsymbol{h}_n^{(MMRR2)} \end{bmatrix} \boldsymbol{y} = \boldsymbol{H}^{(MRR2)}\boldsymbol{y}, \tag{20}$$

where the $1 \times n$ vector, $\boldsymbol{h}_i^{(MMRR2)} = \boldsymbol{x_i}(\boldsymbol{X^T X})^{-1}\boldsymbol{X^T} + \widetilde{\boldsymbol{x}}_i\left(\widetilde{\boldsymbol{X}}^T\boldsymbol{W}_i^{**}\widetilde{\boldsymbol{X}}\right)^{-1}\widetilde{\boldsymbol{X}}^T\boldsymbol{W}_i^{**} + \lambda\widetilde{\boldsymbol{x}}_i(\widetilde{\boldsymbol{X}}^T\boldsymbol{W}_i\widetilde{\boldsymbol{X}})^{-1}\widetilde{\boldsymbol{X}}^T\boldsymbol{W}_i\left[\boldsymbol{I} - (\boldsymbol{X}(\boldsymbol{X^T X})^{-1}\boldsymbol{X^T} + \widetilde{\boldsymbol{X}}\left(\widetilde{\boldsymbol{X}}^T\boldsymbol{W}_i^{**}\widetilde{\boldsymbol{X}}\right)^{-1}\widetilde{\boldsymbol{X}}^T\boldsymbol{W}_i^{**})\right]$is the i^{th} row of the $n \times n$ MMRR2 Hat matrix, $\boldsymbol{H}^{(\boldsymbol{MMRR2})}$.

4. Application

Two multiple response problems are used in order to compare the performance of the proposed model with that of the MRR2 model. The goodness-of-fits used for comparison include the Sum of Squared Errors(SSE), the Coefficient of Determination, (R^2), the Prediction Error Sum of Squares, ($PRESS$), and two versions of the penalized $PRESS$ criteria, namely the $PRESS^{**}$ criterion given in (15) and the $PRESS^{*}$ criterion given as: $PRESS^{*} = \frac{PRESS}{n - tr(H^{(.)})}$, where $PRESS = \sum_{i=1}^{n}\left(y_i - \hat{y}_{i,-i}^{(.)}\right)^2$, $\hat{y}_{i,-i}^{(.)}$ is the leave-one-out estimated value of y_i, $tr(H^{(.)})$ is the trace of the Hat matrix, and $(.)$ refers to any of the regression models MRR2 or MMRR2.

4.1. The Chemical Process Data

This problem was analysed in [14, 15]. The aim of the study was to get the setting of the explanatory variables x_1 and x_2 (representing reaction time and temperature, respectively) that would simultaneously optimize three quality measures of a chemical solution y_1, y_2 and y_3 (representing yield, viscosity, and molecular weight, respectively). The process requirements for each response are as follows:
Maximize y_1 with lower limit $L = 78.5$, and a large enough value $\emptyset = 80$;
y_2 should take a value in the range $L = 62$ and $U = 68$ with $\emptyset$ =65;
Minimize $y_3$with upper limit $U = 3300$ and a small enough value $\emptyset$ = 3100.

i	x_1	x_2	y_1	y_2	y_3
1	0.1464	0.1464	76.5	62	2940
2	0.8536	0.1464	78.0	66	3680
3	0.1464	0.8536	77.0	60	3470
4	0.8536	0.8536	79.5	59	3890
5	0.0000	0.5000	75.6	71	3020
6	1.0000	0.5000	78.4	68	3360
7	0.5000	0.0000	77.0	57	3150
8	0.5000	1.0000	78.5	58	3630
9	0.5000	0.5000	79.9	72	3480
10	0.5000	0.5000	80.3	69	3200
11	0.5000	0.5000	80.0	68	3410
12	0.5000	0.5000	79.7	70	3290
13	0.5000	0.5000	79.8	71	3500

Table 1: *Chemical process data*

As it is the procedure when nonparametric regression is involved, the real values of the explanatory variables are coded to lie between 0 and 1.The data collected via a Central Composite Design is presented in Table 1. A full second-order polynomial model is specified for fitting each response using the OLS model [24, 28].

	Proposed model (MMRR2)							MRR2			
	$N^*(W^{**})$	$C^*(W^{**})$	$N^*(W)$	$C^*(W)$	$b^*(W^{**})$	$b^*(W)$	λ^*	N^*	C^*	b^*	λ^*
y_1	3.5579	0.0500	4.2015	0.0841	0.5985	0.2917	1.0000	12.9930	1.0000	0.2611	0.6973
y_2	6.5539	16.9999	3.1839	0.0950	0.5490	0.3893	1.0000	13.7602	0.0000	0.2670	1.0000
y_3	18.4685	0.1447	4.2078	0.0944	0.3833	0.2734	1.0000	4.4073	0.0937	0.2599	1.0000

Table 2: *Optimal values of the tuning parameters, fixed bandwidth and mixing parameter of the MMRR2 and the MRR2 models for the chemical process data*

The optimal values of the tuning parameters, fixed optimal smoothing parameter, and mixing parameter for the MMRR2 and MRR2 models for each response are presented in Table 2. Table 3 shows the comparison of goodness-of-fit for both models.

Response	Model	DF	$PRESS^{**}$	$PRESS^*$	$PRESS$	SSE	$R^2(\%)$
y_1	$MRR2$	5.3185	0.2695	0.4992	2.6548	0.2714	99.0558
	$MMRR2$	4.0000	**0.1191**	**0.4134**	**1.6535**	**0.2120**	**99.2624**
y_2	MRR2	4.6773	17.0310	42.1432	197.1171	11.2415	96.8880
	MMRR2	4.0000	**9.0119**	**30.8936**	**123.57143**	**10.0000**	**97.2317**
y_3	MRR2	4.0000	54640	148050	592210	**65720**	**92.3804**
	MMRR2	4.0000	**48416**	**135380**	**540530**	**65720**	**92.3804**

Table 3: *Goodness-of-fit of the MMRR2 and the MRR2 models for the chemical process data*

The results presented in Table 3 reveal that the MMRR2 model performs better in terms of all the statistics for y_1 and y_2.Fory_3, the MMRR2 model performs as well as the MRR2 model in terms of both the

SSE and the R^2 but provides significantly improved prediction statistics ($PRESS^{**}$, $PRESS^{*}$ and $PRESS$).

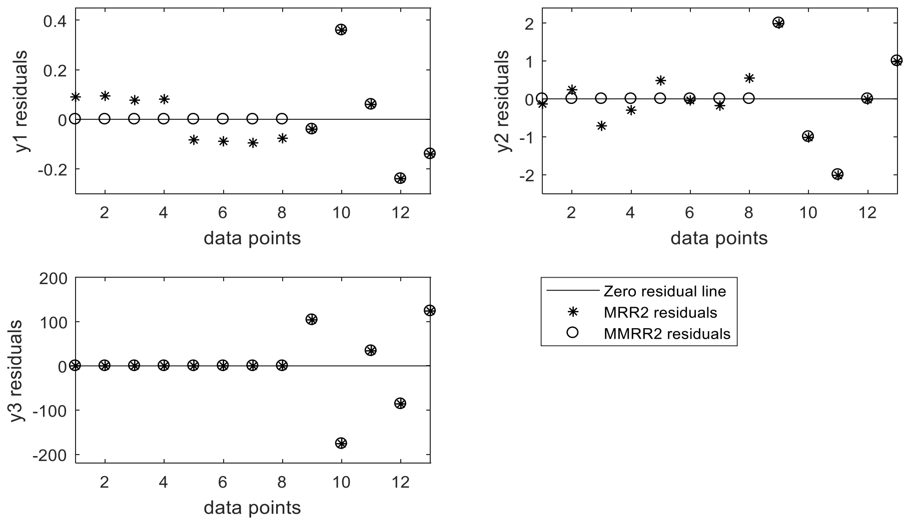

Figure 1: *Plots of residuals from the MRR2 and MMRR2 models for the chemical process data*

Figure 1 shows that the residuals of $y_3$from both regression models overlap while those from the MMRR2,for the most part, are closer to the zero residual line than those from the MRR2 for y_1 and y_2.Eight MMRR2 residuals of y_1 and y_2 are observed to lie on the zero residual line compared to zero and one, respectively, for MRR2. This suggests that the MMRR2 provide better fits of these responses.

Model	x_1	x_2	$\hat{y}_1$	$\hat{y}_2$	$\hat{y}_3$	d_1	d_2	d_3	$D(\%)$
$MRR2$	0.5180	0.2123	78.8603	66.1749	3157.7	0.2402	0.6084	0.7116	47.0216
$MMRR2$	0.5155	0.3467	79.0634	64.9985	3228.2	0.3756	0.9995	0.3589	**51.2641**

Table 4: *Comparison of the optimal results from the desirability function of the MMRR2 and MRR2 models for the chemical process data*

Table 4 presents the optimization results of both the MMRR2 and MRR2models obtained via the desirability function. The MMRR2 model is found to provide the settings of the explanatory variables that give the higher desirability measure.

4.2. The Minced Fish Quality Data

This example uses the Minced Fish Quality Data presented in [8, 28, 29]. The problem seeks to obtain a setting of three explanatory variables x_1 (washing temperature),x_2(washing time) andx_3 (washing ratio of water volume to sample weight) that would optimize four aspects of quality of minced fish including springiness (y_1), thiobarbituric acid number (y_2), cooking loss (y_3), and whiteness index (y_4). The data collected via a Central Composite Design is presented in Table 5.

i	x_1	x_2	x_3	y_1	y_2	y_3	y_4
1	0.2030	0.2030	0.2030	1.83	29.31	29.50	50.36
2	0.7970	0.2030	0.2030	1.73	39.32	19.40	48.16
3	0.2030	0.7970	0.2030	1.85	25.16	25.70	50.72
4	0.7970	0.7970	0.2030	1.67	40.18	27.10	49.69
5	0.2030	0.2030	0.7970	1.86	29.82	21.40	50.09
6	0.7970	0.2030	0.7970	1.77	32.20	24.00	50.61
7	0.2030	0.7970	0.7970	1.88	22.01	19.60	50.36
8	0.7970	0.7970	0.7970	1.66	40.02	25.10	50.42
9	0.0000	0.5000	0.5000	1.81	33.00	24.20	29.31
10	1.0000	0.5000	0.5000	1.37	51.59	30.60	50.67
11	0.5000	0.0000	0.5000	1.85	20.35	20.90	48.75
12	0.5000	1.0000	0.5000	1.92	20.53	18.90	52.79
13	0.5000	0.5000	0.0000	1.88	23.85	23.00	50.19
14	0.5000	0.5000	1.0000	1.90	20.16	21.20	50.86
15	0.5000	0.5000	0.5000	1.89	21.72	18.50	50.84
16	0.5000	0.5000	0.5000	1.88	21.21	18.60	50.93
17	0.5000	0.5000	0.5000	1.87	21.55	16.80	50.98

Table 5: *Minced fish quality data*

The process requirements for each response given in [29] are as follows:
Maximize y_1 with lower bound L=1.70, and large enough value $\emptyset$ = 1.92;
Minimize y_2 with small enough value $\emptyset$ =20.16 and upper bound U=21.00;
Minimize y_3 with small enough value $\emptyset$ =16.80, and upper bound U =20.00;
Maximize y_4 with lower bound L =45.00, and large enough value $\emptyset$ =50.98.

The parametric models specified for the OLS model for response variables $y_1$and y_4 include the intercept, $x_1$and x_1^2. The one specified for y_2 includes the intercept, x_1, x_2, x_1^2, and $x_1x_2$and for y_3 we have the intercept, x_1, x_2, x_3, x_1^2, x_1x_2, x_1x_3, x_3^2[28, 29].

The optimal values of the tuning parameters, fixed smoothing parameter and mixing parameter of the proposed model and the MRR2 models for each response are presented in Table 6. Table 7 shows the results of the goodness-of-fit.

	proposed model (MMRR2)							MRR2			
	$\boldsymbol{N^*(W^{**}}$	$\boldsymbol{C^*(W^{**})}$	$\boldsymbol{N^*(W)}$	$\boldsymbol{C^*(W)}$	$\boldsymbol{b^*(W^{**})}$	$\boldsymbol{b^*(W)}$	$\boldsymbol{\lambda^*}$	$\boldsymbol{N^*}$	$\boldsymbol{C^*}$	$\boldsymbol{b^*}$	$\boldsymbol{\lambda^*}$
y_1	2.1715	13.7116	8.4216	0.1026	0.5507	0.1665	1.0000	8.1582	1.4176	0.1665	1.0000
y_2	4.5552	0.1474	5.7697	0.1066	0.2435	0.2744	1.0000	5.8642	0.1094	0.2567	0.7149
y_3	3.0149	0.0395	4.2963	0.0874	0.4450	0.4013	1.0000	10.6022	0.0796	0.3646	0.8664
y_4	7.3073	16.5049	15.6914	24.6609	0.1201	0.0810	1.0000	11.0554	2.4078	0.1218	1.0000

Table 6: *Optimal values of the tuning parameters, fixed smoothing parameter and mixing parameter of the MMRR2 and the MRR2for the minced fish quality data*

Response	Model	DF	$PRESS^{**}$	$PRESS^{*}$	$PRESS$	SSE	R^2(%)
y_1	$MRR2$	12.0000	0.0021	0.0034	0.0407	**0.0123**	**95.7916**
	$MMRR2$	12.0000	**0.0018**	**0.0028**	**0.0333**	**0.0123**	**95.7916**
y_2	$MRR2$	9.1404	11.8142	**21.5425**	**196.9073**	42.0322	96.9414
	$MMRR2$	8.0000	**11.5081**	26.3707	210.9655	**37.7084**	**97.2560**
y_3	$MRR2$	3.1016	7.5469	**36.8943**	114.4304	2.9713	98.8486
	$MMRR2$	2.0000	**6.1772**	45.1033	**90.2066**	**2.0467**	**99.2069**
y_4	$MRR2$	12.0000	17.4484	37.9271	455.1257	**12.1387**	**97.1990**
	$MMRR2$	12.0000	**6.1794**	**13.4320**	**161.1843**	**12.1387**	**97.1990**

Table 7: *Comparison of goodness-of-fit of MMRR2 and MRR2 for the minced fish quality data*

Results in Table 7 shows that the proposed model provides better values for the $PRESS^{**}$, SSE and R^2 across the four responses. Overall, it is observed that the proposed model exclusively produces better results in thirteen cells, and jointly, in additional four, thus accounting for the better results in seventeen out of a total of twenty cells.

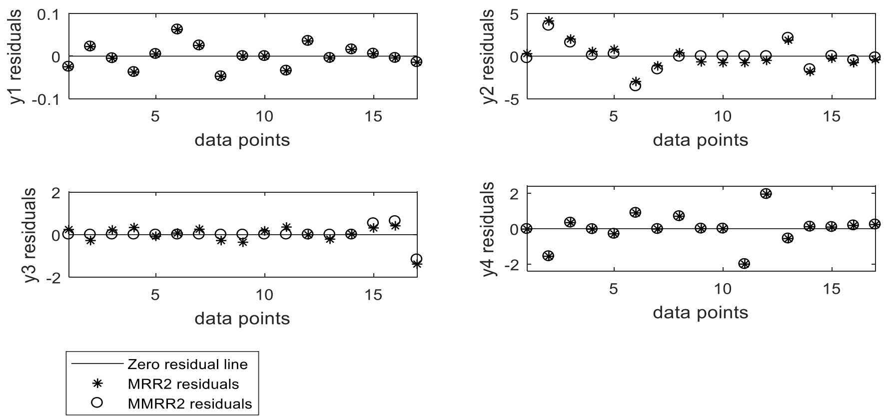

Figure 2: *Plots of residuals from MMRR2 and MRR2 model for the minced quality fish data*

Figure 2 shows that, for all the data points, y_1 and y_4 residuals from both models overlap while those from the MMRR2 are closer to the zero residual line than those from the MRR2 for y_2 and y_3.Seven and fourteen MMRR2 residuals of y_2 and y_3, respectively, are observed to lie squarely on the zero residual line compared to zero and two, respectively, for MRR2.These observations indicate that the MMRR2 provides more accurate fits.

Model	x_1	x_2	x_3	$\hat{y}_1$	$\hat{y}_2$	$\hat{y}_3$	$\hat{y}_4$	d_1	d_2	d_3	d_4	D(%)
$MRR2$	0.3491	1.0000	0.6506	1.8953	18.9230	18.1057	51.4171	0.8876	1.0	0.5920	1.0	85.1387
$MMRR2$	0.4690	0.9341	0.4957	1.8906	20.0080	16.0660	51.0311	0.8662	1.0	1.0000	1.0	**96.4736**

Table 8: *Comparison of the optimal results from the desirability function of the MMRR2 and MRR2 models for the minced fish quality data*

From Table 8, the MMRR2 model is found to give the setting of the explanatory variables that corresponds to a higher desirability measure.

4.3. Simulation Study

In the examples given in section 4.1 and 4.2, it was shown that the goodness of fits and the optimal solutions of fits of LLR_{PBS}were either better than or highly competitive when compared with the results from the MRR2. In this Subsection, we compare the performances of the respective regression models via simulated data. Each Monte Carlo simulation comprises 500 data sets based on the following underlying models:

Underlying Model 1:

$$y_i = 30 + 8x_{1i} + \gamma\{4\sin(3\pi x_{1i}) + 5\cos(3\pi x_{1i})\} + \varepsilon_i;$$

Underlying Model 2:

$$y_i = 42 + 15x_{1i} + 9x_{2i} + 17x_{1i}x_{2i} - 19x_{1i}^2 - 21x_{2i}^2 + \\ +\gamma\{4\sin(3\pi x_{1i}) - 3\cos(3\pi x_{2i}) + 3\sin(5\pi x_{1i}x_{2i})\} + \varepsilon_i,$$

Underlying Model 3:

$$y_i = 35 + 11x_{1i} + 6x_{2i} + 20x_{3i} - 3x_{1i}x_{2i} - 7x_{2i}x_{3i} - 2x_{1i}x_{3i} + 5x_{1i}^2 + 4x_{2i}^2 + 13x_{3i}^2 \\ + \gamma\{\sin(\pi x_{1i}) - \cos(\pi x_{2i}) - \cos(\pi x_{3i}) + \sin(\pi x_{1i}x_{2i}) + \cos(\pi x_{2i}x_{3i}) + \sin(\pi x_{1i}x_{3i})\} + \varepsilon_i,$$

where the x_{1i},x_{2i} and x_{3i} are the values of the explanatory variables, $\varepsilon_i, i = 1,2,\dots,n$, are the error terms which are normally distributed with mean zero and variance 1, and $\gamma$is a misspecification parameter which represents a departure of the model from one specified by the user. The values of the explanatory variables$x_1$for underlying model 1, x_1, $x_2$for underlying model 2 are obtained from Table 1 while those ofx_1, x_2 and $x_3$for underlying model 3 are obtained from Table 5.

Five degrees of model misspecifications, namelyγ =0.0, 0.25, 0.5, 0.75 and 1.0 are considered. For each of the degrees of γ, a full second-order polynomial is specified by the user and this perfectly approximates the underlying model only for the case where $\gamma = 0$ and the random error terms excluded. However, as the value of γ increases, the adequacy of the specified model for the underlying model deteriorates steadily.

The goal of the simulation study is to demonstrate the resolve of each of the regression models when applied to studies that consist of one, two or three explanatory variables, respectively. The regression models Average Sum of Squares of Error (AVESSE) for each degree of λ are presented in Table 9.

Underlying Model	γ	MRR2	MMRR2
1	0.00	8.3213	**8.3209**
	0.25	8.3606	**8.3414**
	0.50	8.4071	**8.4060**
	0.75	8.4083	**8.4071**
	1.00	8.4096	**8.4087**
2	0.00	6.0662	**5.3345**
	0.25	8.6776	**6.0033**
	0.50	14.2233	**11.1331**
	0.75	20.7073	**16.3665**
	1.00	29.7809	**22.2343**
3	0.00	6.9393	**6.2276**
	0.25	8.1300	**7.2754**
	0.50	9.4443	**8.2353**
	0.75	15.5143	**11.5050**
	1.00	18.1191	**13.4171**

Table 9: *Comparison of the AVESSE of each method for each model in the simulation studies*

Table 9 shows the AVESSE of $MRR2$ and that of $MMRR2$ for underlying model 1 is close across the five degrees of model misspecifications. However, $MMRR2$ gives better AVESSE for models 2 and 3 where the curse of dimensionality is more intense.

5. Conclusion

In this paper, we proposed a regression model that modifies the configuration of the MRR2 model. The model combines the philosophy, and hence, the advantages of both the MRR1 and the MRR2 models. Comparisons of the overall performances of the proposed model and the MRR2 model(in terms of goodness-of-fit, optimal solution based on the desirability function and plots of residuals) indicate that the proposed model leverages more from the hybridization of the OLS and the LLR models than the MRR2 model.

Acknowledgement

The authors would like to thank the Editor-in-Chief and the two anonymous reviewers for their very helpful comments and suggestions made on the previous draft of this manuscript.

References

[1] Adalarasan, R. and Santhanakumar, M. (2015). Response surface methodology and desirability analysis for optimizing μWEDM parameters for A16351/20% Al_2O_3 composite, International Journal of ChemTech Research, 7(6), 2625 – 2631.

[2] Anderson-Cook, C.M. and Prewitt, K. (2005). Some guidelines for using nonparametric models for modeling data from response surface designs. Journal of Modern Applied Statistical Models, 4, 106-119.

[3] Atkeson, C.G., Moore, A.W., and Schaal, S. (1997). Locally weighted learning, Artificial Intelligence Review, 11, 11-73.

[4] Box, G.E.P. and Wilson, B. (1951). On the experimental attainment of optimum conditions. Journal of the Royal Statistical Society, Series B, 13, 1- 45.

[5] Castillo, D.E. (2007). Process Optimization: A Statistical Method. New York: Springer International Series in Operations Research and Management Science.

[6] Chen, Y. and Ye, K. (2009). Bayesian hierarchical modelling on dual response surfaces in partially replicated designs. Journal of Quality Technology & Quantitative Management, 6, 371 – 389.

[7] Derringer, G. and Suich, R. (1980). Simultaneous optimization of several response variables, Journal of Quality Technology, 12(4), 214 – 219.

[8] Edionwe, E., Mbegbu, J.I., and Chinwe, R. (2016). A new function for generating local bandwidths for semi–parametric MRR2 model in response surface methodology, Journal of Quality Technology, 48(4), 388 – 404.

[9] Fan, J. and Gijbels, I. (1995). Data-driven bandwidth selection in local polynomial fitting: A variable bandwidth and spatial adaptation. Journal of the Royal Statistical Society, Series b, 57(2), 371 – 394.

[10] Fan, J. and Gijbels, I. (1996). Local Polynomial Modeling and its Applications, Chapman and Hall, London.

[11] Geenens, G. (2011). Curse of dimensionality and related issues in nonparametric functional regression. Statistics Surveys, 5, 30 – 43.

[12] Hardle, W., Muller, M., Sperlich, S., and Werwatz, A. (2005). Nonparametric and Semiparametric Models: An Introduction. Berlin: Springer-Verlag.

[13] Harrington, E.C. (1965). The desirability function, Industrial Quality Control, 21(10), 494 – 498.

[14] He, Z., Wang, J., Oh, J., and Park, S.H. (2009). Robust optimization for multiple responses using response surface methodology, Applied Stochastic Models in Business and Industry, 26, 157 – 171.

[15] He, Z., Zhu, P.F., and Park, S.H. (2012). A robust desirability function for multi-response surface optimization. European Journal of Operational Research, 221, 241-247.

[16] Holland, J. (1975). Adaptation in Natural and Artificial Systems. *University of Michigan Press, Ann Arbor.*

[17] Johnson, R.T. and Montgomery, D.C. (2009). Choice of second-order response surface designs for logistics and poisson regression models, Int. Journal of Experimental Design and Process Optimization, 1(1), pp. 2 – 23.

[18] Kohler, M., A. Schindler, and S. Sperlich (2014). A review and comparison of bandwidth selection methods for kernel regression. International Statistical Review, 82(2), 243–274.

[19] Mays, J.E., Birch, J.B., and Einsporn, R.L. (2000). An overview of model robust Regression. Journal of Statistical Computation and Simulation, 66, 79-100.

[20] Mays, J.E., Birch, J.B., and Starnes, B.A. (2001).Model robust regression: Combining parametric, nonparametric, and semi-parametric models. Journal of Nonparametric Statistics, 13, 245-277.

[21] Mondal, A. and Datta, A.K. (2011). Investigation of the process parameters using response surface methodology on the quality of crustless bread baked in a water-spraying oven, Journal of food process engineering, 34, 1819 – 1837.

[22] Myers, R.H. (1999). Response surface methodology - Current status and future directions. Journal of Quality Technology, 31, 30-44.

[23] Myers, R., Montgomery, D.C., and Anderson-Cook, C.M. (2009). Response Surface Methodology: Process and Product Optimization Using Designed Experiments, Wiley.

[24] Nadaraya, E.A. (1964). On estimating regression. Journal of Theory of Probability and its Application, 9, 141 – 142.

[25] Pickle, S.M., Robinson, T.J., Birch, J.B., and Anderson-Cook, C.M (2008). "A semi-parametric model to robust parameter design. Journal of Statistical Planning and Inference, 138, 114-131.

[26] Prewitt, K. and Lohr, S. (2006). Bandwidth selection in local polynomial regression using eigenvalues, Journal of Royal Statistical Society, Series B, 68(1), 135 – 154.

[27] Sestelo, M., Villanueva, N.M., Meira-Machado, L., and Roca-Pardinas, J. (2017).An R package for nonparametric estimation and inference in life sciences, Journal of Statistical Software, 82(12), 1- 27.

[28] Shah, K.H., Montgomery, D.C., and Carlyle, W.M. (2004). Response surface modeling and optimization in multi-response experiments using seemingly unrelated regressions. Quality Engineering, 16, 387-397.

[29] Wan, W. and Birch, J.B. (2011). A semi-parametric technique for multi-response optimization. Journal of Quality and Reliability Engineering International, 27, 47-59.

[30] Wu, C.F.J. and Hamada M.S. (2000). Experiments: Planning, Analysis and Parameter Design Optimization. New York. John Wiley & Sons, Inc.

[31] Yeniay, O. (2014). Comparative study of algorithm for response surface optimization. Journal of Mathematical and Computational Applications, 19, 93 -104.

[32] Zheng, Q., Gallagher, C. and Kulasekera, K.B. (2013). Adaptively weighted kernel regression. Journal of Nonparametric Statistics, 25(4), 855 – 872.