BIU Journal of Basic and Applied Sciences 8(1): 81 – 90, 2023.

ISSN: 2563-6424

IMPROVING THE PERFORMANCE OF SEMIPARAMETRIC MODEL ROBUST REGRESSION 2 VIA THE INCLUSION OF THE STATISTICALLY SIGNIFICANT INTERACTION TERMS IN THE MODEL MATRIX

***EDIONWE, E.[1] AND OSEMWENKHA, S. O.[2]**
[1]Department of Mathematical Sciences, Edwin Clark University, Kiagbodo, Delta State, Nigeria
[2]Department of Statistics, University of Benin, Benin City, Nigeria
*Corresponding Author: trustnelson24@yahoo.com

ABSTRACT

Response Surface Methodology (RSM) comes handy when a researcher wants to determine the value of each of the explanatory variables that simultaneously optimize the response variables. In the modelling phase of RSM, a suitable regression model is fitted using the data generated from the experimental design phase. The fitted model is subsequently subjected to an appropriate optimization routine in order to obtain the optimal solution of the study. Currently, the semiparametric model robust regression 2 (MRR2) model is considered the best regression model for handling data emanating from response surface studies. MRR2 is a hybrid model, combining estimates of the response (output) from both the local linear regression (LLR) and the ordinary least squares (OLS) via mixing parameters. When MRR2 is applied in response surface studies, the current philosophy entails the exclusion of interaction terms in the model matrix of LLR component of MRR2 irrespective of the statistical significance of the interaction terms in the OLS model matrix. In this paper, we present results for a problem from the literature in which the significant interaction terms in the OLS model matrix were duly included in LLR model matrix. A multiple response problem from the literature was used to justify the inclusion of the interaction terms in the LLR model matrix. It is found that the MRR2 applied with the interaction terms included outperforms its counterpart both in terms of the goodness-of-fit statistics and the desirability-based optimal solutions. Specifically, the MRR2 with the proposed model matrix gives better prediction errors in the three responses as well as a desirability value of approximately 77.3% as against the 47.4% for the MRR2 which disregards the significant interaction terms in its model matrix.

KEYWORDS: Local bandwidths, Local linear regression, Model matrix, Model robust regression 2, Response surface methodology, Semi-parametric regression

INTRODUCTION

Response Surface Methodology (RSM) is a collection of statistical tools that provide a means of establishing an empirical relationship between a k explanatory variables, $x_1, x_2, \dots, x_k$ and a response variable (y) of a system or process using data from designed experiments (Montgomery, 2005; Matys *et al*., 2022; Joudi-Sarighayeh *et al*., 2023). In the modeling phase of RSM, researchers assume that the relationship between the response variable, y and the k explanatory variables $x_{1,}\ x_2, \dots, x_k$, takes the form:

$$y_i = f\left(x_{i1,}\ x_{i2}, \dots, x_{ik}\right) + \varepsilon_i, \quad i = 1,2,\dots,n \qquad (1)$$

where the mean function f denotes the true but unknown relationship between the response variable and the k explanatory variables, $x_{ij,}\ i = 1,\dots,n, j = 1,\dots,k,$ are the values of the j^{th} explanatory variable at the i^{th} data point, $\varepsilon_i,\ i = 1,2,\dots,n,$ are random error terms with the assumption that $\boldsymbol{\varepsilon} \sim N(0, \sigma^2)$, and n is the sample size (Castillo, 2007; Myers et al., 2009; Karlovic et al., 2023).

Existing classes of regression models applied in the estimation of the unknown function f in (1) include the parametric regression model, the nonparametric regression model and the semi-parametric regression model (Anderson-Cook and Prewitt, 2005; Pickle et al., 2008). However, for small-sample setting which is typical of RSM, the semi-parametric Model Robust Regression 2 (MRR2) is currently considered the best regression model (Wan and Birch, 2011; Eguasa et al., 2022).

Mathematically, the MRR2 estimate, $\hat{y}_i^{(MRR2)}$ of y_i, $i = 1,2,\dots,n,$ is given by:

$$\hat{y}_i^{(MRR2)} = \boldsymbol{x_i}(\boldsymbol{X^T X})^{-1}\boldsymbol{X^T y} + \lambda_i \tilde{\mathbf{x}}_{\mathbf{i}}(\boldsymbol{\tilde{X}^T W_i \tilde{X}^T})^{-1}\boldsymbol{\tilde{X}^T W_i}(\boldsymbol{I} - (\boldsymbol{X^T X})^{-1}\boldsymbol{X^T})\boldsymbol{y}, \quad (2)$$

$$= \boldsymbol{x_i}(\boldsymbol{X^T X})^{-1}\boldsymbol{X^T y} + \lambda_i \tilde{\mathbf{x}}_{\mathbf{i}}\left(\boldsymbol{\tilde{X}^T W_i \tilde{X}}\right)^{-1}\boldsymbol{\tilde{X}^T W_i r^{(OLS)}}, \quad (3)$$

where λ_i, $0 \le \lambda_i \le 1,\ i = 1,2,\dots,n,$ is the local mixing parameter for combining the OLS and the LLR estimates at the i^{th} data point, $\boldsymbol{x_i}(\boldsymbol{X^T X})^{-1}\boldsymbol{X^T y}$ is the OLS component which is the parametric regression component, $\boldsymbol{x_i}$ is the i^{th} row vector of the $n \times p$ OLS model matrix $\boldsymbol{X}$, where p is the number of model parameters, $\boldsymbol{y}$ is $n \times 1$ vector of response, $\tilde{\mathbf{x}}_{\mathbf{i}}(\boldsymbol{\tilde{X}^T W_i \tilde{X}^T})^{-1}\boldsymbol{\tilde{X}^T W_i}(\boldsymbol{I} - (\boldsymbol{X^T X})^{-1}\boldsymbol{X^T})\boldsymbol{y}$ is the LLR component, $\tilde{\mathbf{x}}_{\mathbf{i}}$ is the i^{th} row of the LLR model matrix $\boldsymbol{\tilde{X}}$**,** whose description and that of the one for OLS are both given in Section 2, $\boldsymbol{I}$ is an $n \times n$ identity matrix, $\boldsymbol{W_i}$ is a $n \times n$ diagonal weights matrix for estimating the i^{th} OLS residual, $\boldsymbol{r^{(OLS)}}$ is the $n \times 1$ vector of the OLS residuals, $\boldsymbol{X^T}$ and $\boldsymbol{\tilde{X}^T}$ are the transposed model matrices of the OLS and LLR, respectively (Mays et al., 2001; Pickle et al., 2008).

The rth-entry say w_r of weight matrix, $\boldsymbol{W_i}$ is obtained from the Gaussian product kernel as:

$$w_r^{**} = \prod_{j=1}^{k} K\left(\frac{x_{ij} - x_{rj}}{b_i}\right) \Big/ \sum_{i=1}^{n} \prod_{j=1}^{k} K\left(\frac{x_{ij} - x_{rj}}{b_i}\right), \qquad i = 1,2,\dots,n, j = 1,2,\dots,k, \qquad (4)$$

where $K\left(\frac{x_{ij} - x_{rj}}{b_i}\right) = e^{-\left(\frac{x_{ij} - x_{rj}}{b_i}\right)^2}$ is the simplified Gaussian kernel and b_i, $0 < b_i \le 1, i = 1,2,\dots,n,$ are the locally adaptive or simply the local bandwidths (Mays *et al*., 2002).

For data emanating from response surface studies, a model for selecting locally adaptive bandwidths presented in Edionwe *et al*. (2016) and Edionwe *et al*. (2018) is given by:

$$b_i = \frac{b^*\{N(TC-y_i)\}}{T(Cn-1)}, \quad i = 1, \dots, n \tag{5}$$

where b* (which may be taken to be equal to 1) is the fixed optimal bandwidth, $T = \sum_i^n y_i$, $C \geq 0$, $N > 0$ are data-driven tuning parameters whose optimal values C^*, N^*, respectively, generate a vector, Φ of n locally adaptive optimal bandwidths, $[b_1^*, b_2^*, \dots, b_n^*]$ based on the minimization of the penalized Prediction Error Sum of Squares (PRESS**) given by:

$$PRESS^{**}(\Phi) = \frac{\sum_{i=1}^n \left(y_i - \hat{y}_{i,-i}^{(MRR2)}(\Phi)\right)^2}{n - trace\left(H^{(MRR2)}(\Phi)\right) + (n-k-1)\frac{SSE_{max} - SSE_{\Phi}}{SSE_{max}}}, \tag{6}$$

where $\sum_{i=1}^n \left(y_i - \hat{y}_{i,-i}^{(MRR2)}(\Phi)\right)^2$ is the Prediction Error sum of Squares (PRESS), SSE_{max} is the maximum Sum of Squared Errors obtained as $b_1, b_2, \dots, b_n$ tend to infinity, SSE_{Φ} is the Sum of Squared Errors (SSE) for a specific vector of bandwidths $\Phi = [b_1, b_2, \dots, b_n]$, $tr\left(H^{(MRR2)}\right)$ is the trace of the MRR2 hat matrix and $\hat{y}_{i,-i}^{(MRR2)}(\Phi)$ is the leave-one-out cross-validation estimate of y_i with the i^{th} observation left out (Mays et al., 2001; Pickle et al., 2008).

Similarly, a model for selecting the local mixing parameters proposed by Edionwe *et al.* (2017) is given by:

$$\lambda_i(e_i) = \frac{N_0(T_0C_0 + e_i)}{T_0(C_0n+1)}, \quad i = 1,2, \dots, n, \tag{7}$$

where $T_o = \sum_i^n e_i$, $e_i = |y_i - y_i^{(OLS)}|$, $C_o \geq 0$, $N_o \geq 0$, are data-driven tuning parameters. The optimal values of C_o and N_o are herein designated as C_o^*, N_o^*, respectively. The optimal vector of mixing parameters generated from (7) is based on the minimization of a version of the PRESS** given by:

$$PRESS^{**}([\lambda_1, \lambda_2, \dots, \lambda_n]) = \frac{\sum_{i=1}^n \left(y_i - \hat{y}_{i,-i}^{(MRR2)}(\Phi^*, [\lambda_1, \lambda_2, \dots, \lambda_n])\right)^2}{n - tr\left(H^{(MRR2)}(\Phi^*, [\lambda_1, \lambda_2, \dots, \lambda_n])\right) + (n-k-1)\frac{SSE_{max} - SSE_{\Phi^*[\lambda_1, \lambda_2, \dots, \lambda_n]}}{SSE_{max}}}, \tag{8}$$

where Φ^*=$[b_1^*, b_2^*, \dots, b_n^*]$ denotes locally optimal bandwidths that minimizes (7) above, $SSE_{\Phi^*[\lambda_1, \lambda_2, \dots, \lambda_n]}$ is the SSE for optimal bandwidths associated with a given vector of mixing parameters, $[\lambda_1, \lambda_2, \dots, \lambda_n]$, $tr\left(H^{(MRR2)}(\Phi^*, [\lambda_1, \lambda_2, \dots, \lambda_n])\right)$ is the trace of MRR2 hat matrix given Φ^* and $[\lambda_1, \lambda_2, \dots, \lambda_n]$, $DF = n - tr\left(H^{(MRR2)}(\Phi^*, [\lambda_1, \lambda_2, \dots, \lambda_n])\right)$, and $\hat{y}_{i,-i}^{(MRR2)}(\Phi^*, [\lambda_1, \lambda_2, \dots, \lambda_n])$ is the leave-one-out cross-validation of MRR2 estimate of y_i given the optimal bandwidths and a vector of mixing parameters, $[\lambda_1, \lambda_2, \dots, \lambda_n]$.

The MRR2 as presented in equation (2) may be expressed in matrix form as:

$$y^{(MRR2)} = \begin{bmatrix} h_1^{(MRR2)} \\ h_2^{(MRR2)} \\ \vdots \\ h_n^{(MRR2)} \end{bmatrix} y \qquad = H^{(MRR2)}y, \tag{9}$$

where $h_i^{(MRR2)} = x_i(X^TX)^{-1}X^T + \lambda_i\tilde{x}_i(\tilde{X}^TW_i^*\tilde{X}^T)^{-1}\tilde{X}^TW_i^*(I - X(X^TX)^{-1}X^T)$ is the i^{th} row of the $n \times n$ MRR2 Hat matrix, $H^{(MRR2)}$.

Once the mean function, f, in (1) has been modelled, the resulting fitted curve is subjected to some optimization routine in order to obtain the values of the explanatory variables that optimize the response based on the prescribed production requirements. For multiple response studies involving $m > 1$ response variables, an efficient optimization criterion for achieving this task is the desirability function (Harrington, 1975; Derringer and Suich, 1980; He et al., 2012). Given the production requirement of a response, the desirability function transforms the estimated response, $\hat{y}(\boldsymbol{x})$ into a response-type dependent scalar measure, $d(\hat{y}(\boldsymbol{x}))$.

If the response is of nominal-the-better (NTB) type where the response acceptable value lies between an upper limit, U and a lower limit, L, $d(\hat{y}(\boldsymbol{x}))$ is given by:

$$d(\hat{y}(x)) = \begin{cases} 0 & \hat{y}(\boldsymbol{x}) < L \\ \left\{\frac{\hat{y}(x)-L}{\emptyset-L}\right\} & L \le \hat{y}(\boldsymbol{x}) < \emptyset, \\ \left\{\frac{U-\hat{y}(x)}{U-\emptyset}\right\} & \emptyset \le \hat{y}(\boldsymbol{x}) \le U, \\ 0 & \hat{y}(\boldsymbol{x}) > U, \end{cases} , \qquad (10)$$

where $\emptyset$ is the target value of the given response.

If the response is of larger-the-better (LTB) type, then the objective is to maximize the response and so $d(\hat{y}(\boldsymbol{x}))$ is given by a one-sided transformation as:

$$d(\hat{y}(\boldsymbol{x})) = \begin{cases} 0 & \hat{y}(\boldsymbol{x}) < L, \\ \left\{\frac{\hat{y}(x)-L}{\emptyset-L}\right\} & L \le \hat{y}(\boldsymbol{x}) \le \emptyset, \\ 1 & \hat{y}(\boldsymbol{x}) > \emptyset, \end{cases} \qquad (11)$$

where $\emptyset$ is interpreted as large enough value of the given response.

The $d(\hat{y}(\boldsymbol{x}))$ of a smaller-the-better (STB) response is given by a one-sided transformation as:

$$d(\hat{y}(\boldsymbol{x})) = \begin{cases} 1 & \hat{y}(\boldsymbol{x}) < \emptyset, \\ \left\{\frac{U-\hat{y}(x)}{U-\emptyset}\right\} & \emptyset \le \hat{y}(\boldsymbol{x}) \le U, \\ 0 & \hat{y}(\boldsymbol{x}) > U, \end{cases} \qquad (12)$$

where $\emptyset$ is a small enough value of the given response.

The overall objective of the desirability criterion is getting the values of the explanatory variables that maximize the geometric mean, D, of all the individual desirability measures given as:

$$D = maximize\{[d_1 \times d_2 \times \ldots \times d_m]^{1/m}\}, \qquad (13)$$

Model Matrices for MRR2 in RSM

In this Section, we give an overview of the current make-up of the model matrices for OLS and LLR. The Section concludes with a presentation of a general form of the model matrix for LLR when the statistically significant interaction terms in the OLS model matrix is duly incorporated

OLS Model Matrix in RSM

Suppose that the entire terms in a second-order polynomial assumed for estimating the unknown function f in (1) are all significant, then the model matrix $\boldsymbol{X}$ of the OLS fitted to a data of sample size n and consisting of k explanatory variables takes the form:

$$\boldsymbol{X} = \begin{bmatrix} 1 & x_{11} & \cdots & x_{1k} & x_{11}^2 & \cdots & x_{1k}^2 & \cdots & x_{11}x_{12} & \cdots & x_{1(k-1)}x_{1k} \\ 1 & x_{21} & \cdots & x_{2k} & x_{21}^2 & \cdots & x_{2k}^2 & \cdots & x_{21}x_{22} & \cdots & x_{2(k-1)}x_{2k} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_{n1} & \cdots & x_{n2} & x_{n1}^2 & \cdots & x_{nk}^2 & \cdots & x_{n1}x_{n2} & \cdots & x_{n(k-1)}x_{nk} \end{bmatrix}$$

where the column of ones, $[1, 1, \dots, 1]^T$, represents the intercept, x_{ij}, $i = 1,2,\dots,n, j = 1,2,\dots,k,$ denotes the value of the j^{th} explanatory variable at the i^{th} data point, x_{ij}^2 are the quadrature terms, and the interaction terms form the columns:

$$\begin{bmatrix} x_{11}x_{12} & x_{11}x_{13} & \cdots & x_{1(k-1)}x_{1k} \\ x_{21}x_{22} & x_{21}x_{23} & \cdots & x_{2(k-1)}x_{2k} \\ \vdots & \vdots & \vdots & \vdots \\ x_{n1}x_{n2} & x_{11}x_{12} & \cdots & x_{n(k-1)}x_{nk} \end{bmatrix}.$$

It is sometimes the case that not all the terms in the above OLS model matrix are statistically significant or have significant effect on the response variable. In order to be sure of which terms to include in the OLS model matrix for a particular study, statistical tests such as ANOVA is conducted and the p-values for each term is read off the ANOVA table. The terms whose p-values are greater than 0.05 are considered statistically insignificant and, thus, discarded from the OLS model matrix. Other statistical tests carried out to determine significance of regression terms include the F-test (for simultaneously testing the significance of all the terms), and the t-test for testing the significance of the individual term (Agarwal, 2015).

LLR Model Matrix for RSM

LLR is one of the nonparametric polynomial regression models developed and applied originally in the setting of scatterplot smoothing where no consideration is ascribed to interaction terms between the independent variables (Fan and Gijbels, 1992; Fan and Gijbels, 1996; Loader, 1999; Hardle et al., 2005). Hence, when LLR was imported to response surface studies, whether as a stand-alone regression model or a hybrid combination with OLS, the traditional practice of leaving out the interaction terms was preserved (Pickle *et al*, 2008; Wan and Birch, 2011; Eguasa *et al*., 2022).

Therefore, presently, the general form of the LLR model matrix $\widetilde{\boldsymbol{X}}$ for a data that comprises k explanatory variables is given by:

$$\widetilde{\boldsymbol{X}} = \begin{bmatrix} 1 & \tilde{x}_{11} & \cdots & \tilde{x}_{1k} \\ 1 & \tilde{x}_{21} & \cdots & \tilde{x}_{2k} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & \tilde{x}_{n1} & \cdots & \tilde{x}_{nk} \end{bmatrix}$$

where $\tilde{x}_{ij}$, $i = 1,2,\dots,n, j = 1,2,\dots,k,$ denotes the value of the j^{th} explanatory variable at the i^{th} data point.

METHODOLOGY

Essentially, an interaction term that is correctly declared statistically significant by a statistical test certainly would contain an important information regarding the degree of the quantitative relationship between the response and the explanatory variables and impact it in some appreciable ways. Hence, we proposed the inclusion of such significant interaction terms in the LLR model matrix for improved performance of the MRR2 in RSM. Therefore, the current paper proposes a new LLR model matrix which takes the general form given by:

$$\widetilde{\boldsymbol{X}} = \begin{bmatrix} 1 & \tilde{x}_{11} & \cdots & \tilde{x}_{1k} & \tilde{x}_{11}\tilde{x}_{12} & \cdots & \tilde{x}_{1(k-1)}\tilde{x}_{1k} \\ 1 & \tilde{x}_{21} & \cdots & \tilde{x}_{2k} & \tilde{x}_{21}\tilde{x}_{22} & \cdots & \tilde{x}_{2(k-1)}\tilde{x}_{2k} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & \tilde{x}_{n1} & \cdots & \tilde{x}_{nk} & \tilde{x}_{n1}\tilde{x}_{n2} & \cdots & \tilde{x}_{n(k-1)}\tilde{x}_{nk} \end{bmatrix}$$

where all the $\tilde{x}_{ij}, i = 1,2, \dots, n, j = 1,2, \dots, k,$ retain their previous definitions.

Application

This section presents the analysis of a multi-response problem from the literature. For easy reference, MRR2 results obtained using existing LLR model matrix, and the proposed LLR model matrix are designated MRR2, MRR2*, respectively. Tables that shows the goodness-of-fit such as SSE, Coefficient of Determination (R^2), PRESS as well as the optimization results based on the desirability function are presented. The best goodness-of-fit as well as the desirability value are shown in bold.

The Multiple Response Chemical Data

This problem originates from Montgomery (2005). It involves three response variables, namely the y_1 (yield), y_2 (viscosity), and y_3 (molecular weight). Two inputs (factors) were found to influence these responses: reaction time (x_1) and temperature (x_2). A full second-order polynomial were found to be adequate for each of the three response variables, meaning that the interaction term $x_{i1}x_{i2}, i = 1,2, \dots, 13,$ which represents the collective effects of reaction time and temperature on chemical yield, are statistically significant. Hence, its inclusion in the LLR model matrices for y_1, y_2 and y_3 for MRR2*.

The process specifications for each of the responses are as follows:

Maximize y_1 with lower limit $L = 78.5$, and target value, $\emptyset = 80$;

y_2 is to take a value in the range of $L = 62$ and $U = 68$ with target value, $\emptyset = 65$;

Minimize y_3 with upper limit $U = 3300$ and target value, $\emptyset = 3100$.

The data, collected via a Central Composite Design (CCD), is presented in Table 1. The optimal tuning parameters from (5), the corresponding local bandwidths based on (6), and the optimal tuning parameters from (7), the corresponding local mixing parameters based on (8) for MRR2 and MRR2* are shown in Tables 2 and 3, respectively. The goodness-of-fit and optimization results for each of the regression models are presented in Tables 4 and 5, respectively.

Table 1: Chemical process data

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 2: Optimal local bandwidths and mixing parameters for MRR2

	Optimal Local Bandwidths			Optimal Local Mixing Parameters		
i	y_1 $N = 3.4392$ $C = 0.0268$	y_2 $N = 3.6743$ $C = 0.0001$	y_3 $N = 1.2120$ $C = 0.0939$	y_1 $N_o = 8.8755$ $C_o = 0.0883$	y_2 $N_o = 8.6854$ $C_o = 0.1052$	y_3 $N_o = 9.0054$ $C_o = 0.1135$
1	0.2543	0.2677	0.1489	0.7142	0.4811	0.4392
2	0.2621	0.2850	0.0566	0.7317	0.6516	0.8870
3	0.2569	0.2590	0.0828	0.6606	1.0000	0.4795
4	0.2698	0.2547	0.0304	0.6718	0.6391	0.9273
5	0.2497	0.3066	0.1389	0.6839	0.8154	0.4592
6	0.2642	0.2936	0.0965	0.7086	0.4669	1.0000
7	0.2569	0.2461	0.1227	0.7342	0.5786	0.6548
8	0.2647	0.2504	0.0628	0.6583	0.9271	0.7118
9	0.2719	0.3109	0.0815	0.4354	0.7845	0.6873
10	0.2740	0.2979	0.1165	1.0000	0.5853	0.8771
11	0.2724	0.2936	0.0903	0.4708	0.7846	0.5026
12	0.2709	0.3023	0.1052	0.7882	0.3860	0.6397
13	0.2714	0.3066	0.0790	0.6118	0.5852	0.7400

Table 3: Optimal local bandwidths and mixing parameters for MRR2*

	Optimal Local Bandwidths			Optimal Local Mixing Parameters		
	y_1 $N = 1.2973$ $C = 0.6570$	y_2 $N = 4.8995$ $C = 0.0510$	y_3 $N = 1.2780$ $C = 0.0922$	y_1 $N_o = 7.0444$ $C_o = 0.0139$	y_2 $N_o = 9.9584$ $C_o = 0.2193$	y_3 $N_o = 8.9373$ $C_o = 0.1089$
1	0.1001	0.3177	0.1763	0.5873	0.6342	0.4296
2	0.0999	0.3861	0.0597	0.6126	0.7544	0.8850
3	0.1000	0.2836	0.0928	0.5099	1.0000	0.4706
4	0.0996	0.2665	0.0266	0.5352	0.7456	0.9260
5	0.1003	0.4715	0.1637	0.5435	0.8698	0.4499
6	0.0998	0.4203	0.1101	0.5792	0.6242	1.0000
7	0.1000	0.2323	0.1432	0.6161	0.7030	0.6489
8	0.0998	0.2494	0.0676	0.5066	0.9486	0.7069
9	0.0996	0.4886	0.0912	0.1847	0.8481	0.6820
10	0.0995	0.4373	0.1353	1.0000	0.7076	0.8750
11	0.0995	0.4203	0.1022	0.2359	0.8481	0.4942
12	0.0996	0.4544	0.1212	0.6941	0.5671	0.6335
13	0.0996	0.4715	0.0881	0.4394	0.7076	0.7356

Table 4: Comparison of goodness-of-fit for statistics for MRR2 and MRR2*

Response	Model	DF	SSE	R^2	PRESS	PRESS**
y_1	MRR2	5.3254	0.2711	99.0567	2.6539	0.2691
	MRR2*	5.3161	0.2655	99.0765	2.4785	0.2608
y_2	MRR2	5.3830	12.8545	96.4415	198.5980	16.7814
	MRR2*	5.0996	12.1397	96.6393	173.5069	15.6555
y_3	MRR2	4.9155	68698	92.0351	612200	52728
	MRR2*	4.9311	68800	92.0232	503110	45936

MRR2* gives the best PRESS and PRESS** across the three responses. In addition, MRR2* produces better SSE and R^2 across two of the three responses. In summary, it is seen that MRR2* accounts for the better values in ten cells out of a total of twelve.

Table 5: Optimization results based on desirability function in (13)

Model	x_1	x_2	max($\hat{y}_1$)	Ø($\hat{y}_2$)	min ($\hat{y}_3$)	d($\hat{y}_1$)	d($\hat{y}_2$)	d($\hat{y}_3$)	D(%)
MRR2	0.5159	0.2110	78.8604	66.1260	3158.2	0.2403	0.6247	0.7090	47.3887
MRR2*	0.9461	0.6828	79.2138	64.9084	2023.4	0.4759	0.9695	1.0000	77.2687

The relatively better results of PRESS and PRESS** for MRR2* in Table 4 is utilized to get a relatively better setting ($x_1 = 0.9461, x_2 = 0.6828$) of the explanatory variables that simultaneously optimizes the three responses ($y_1 = 79.2138, y_2 = 64.9084$, $y_3 = 2023.4$) based on their individual process specifications. Table 4 shows that MRR2* gives a desirability value of approximately 77.3% (from comparatively better individual desirability of $d(\hat{y}_1) = 0.4759$, $d(\hat{y}_2) = 0.9695$, $d(\hat{y}_3) = 1.0000$) in comparison to its counterpart which only manages a paltry value of approximately 47.4%.

Conclusion

In this paper, the inclusion of the statistically significant interaction terms to the model matrix for the LLR component of the MRR2 was proposed and the MRR2 resulting from such incorporation of such significant interaction terms was designated MRR2* in the current paper.

MRR2* was found to perform better than MRR2. Specifically, MRR2* was found to give smaller prediction errors in the three responses considered in the paper than the existing MRR2. Furthermore, MRR2* produced a desirability of 77.3% outperforming the MRR2 that produced 47.4%, and thus providing an improvement of $100\left(\frac{77.3-47.4}{47.4}\right)\% = 63.1\%$ in the capacity to meet prespecified product requirements. The practical relevance of the desirability values is that with MRR2*, the researcher is able to apply optimal values of both the reaction time and reaction temperature in order to produce a product that meets 77.3% of the product requirements.

REFERENCES

Agarwal, B. L. (2015). Basic Statistics. New Age International Publishers: New Delhi.

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.

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

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

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.

Edionwe, E., Mbegbu, J. I., Ekhosuehi, N. and Obiora-Iluouno, H. O. (2018). An improved robust regression model for response surface methodology. *Croatian Operational Research Review*, 9: 317 – 330.

Edionwe, E., Mbegbu, J. I. and Iguodala, W. A. (2017). Improving the performance of model-robust regression 2 (MRR2) method using new adaptive mixing parameters and a modified penalized error sum of squares, *Journal of the Nigerian Association of Mathematical Physics*, 41: 229 – 240.

Eguasa, O., Edionwe, E. and Mbegbu, J. I. (2022). Local linear regression and the problem of dimensionality: A remedial strategy via a new locally adaptive bandwidths selector, *Journal of Applied Statistics*, DOI: 10.1080/02664763.2022.2026895

Fan, J. and Gijbels, I. (1992). Variable bandwidth and local linear regression Smoothers. *The Annals of Statistics*, 20(4): 2008-2036.

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

Hardle, W., Muller, M., Sperlich, S. and Werwatz, A. (2005).

Nonparametric and Semiparametric Models: An Introduction. Berlin: Springer-Verlag

Harrington, E. C. (1965). The desirability function. *Industrial Quality Control*, 21(10): 494 – 498.

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.

Joudi-Sarighayeh, F., Abbaspour-Gilandeh, Y., Kaveh, M., Szymanek, M. and Kulig, R. (2023). Response Surface Methodology Approach for Predicting Convective/Infrared Drying, Quality, Bioactive and Vitamin C Characteristics of Pumpkin Slices. *Foods*, 12: 1114.

Karlovic, S., Dujmíc, F., Rimac, B. S., Badanjak, S. M., Nincevic, G. A., Škegro, M., Šimic, M. A. and Brncic, M. (2023). Mathematical modeling and optimization of ultrasonic pre-treatment for drying of pumpkin (*Cucurbita moschata*). *Processes*, 11: 469.

Loader, C. R. (1999). Bandwidth selection: Classical or plug-in? *The Annals of Statistics*, 2: 415-438.

Mays, J. E. and Birch, J. B. (2002). Smoothing for small samples with model misspecification: Nonparametric and Semi-parametric concerns. *Journal of Applied Statistics*, 29(7): 1023-1045.

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.

Matys, A., Dadan, M., Witrowa-Rajchert, D., Parniakov, O. and Wiktor, A. (2022). Response surface methodology as a tool for optimization of pulsed electric field pretreatment and microwave-convective drying of apple. *Appl. Sci*., 12: 3392.

Montgomery, D. C. (2005). Design and Analysis of Experiments, sixth ed., Wiley, New York.

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

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.

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.