A COMPARATIVE STUDY OF SEMIPARAMETRIC REGRESSION MODELS: A NOVEL BLEND VIA LOCALLY ADAPTIVE BANDWITHS SELECTOR FOR RESPONSE SURFACE METHODOLOGY (RSM)

EGUASA, O.* AND OKANIGBUAN, P.N.
Department of Physical Sciences, Benson Idahosa University, Benin City, Nigeria

ABSTRACT

Response Surface Methodology (RSM) is a sequential statistical approach employed by engineers and industrial statistician for empirical model building where the processes and products are optimized. In RSM, the parametric regression models are frequently used but lack credibility due to model misspecification and as such affects the process mean and variance and ultimately, the estimated response is misscalulated. Although, the nonparametric regression models are flexible, but lack recognition in RSM due to the idiosyncracies of RSM data; such as dimensionality problem, sparseness of RSM data and small sample size. In the literature, semiparametric regression models are considered the most suitable methods in RSM because it combines attributes of the parametric and nonparametric regression models in a fashionable manner. In this paper, we give a comparative analysis of the OLS and three semiparametric regression models that utilizes two existing locally adaptive bandwidths from the literature were applied to obtain a novel blend of the semiparametric regression models used to smooth the two data for the application problems and the results tend to improve the goodness of fit statistics, with minimum residual plots for the responses and optimization of processes and products.

Keywords: Dimensionality problem, ordinary least squares (OLS), parametric regression model, nonparametric regression model, sequential statistical approach
***Correspondence:** oeguasa@biu.edu.ng**,** +2348028631580

INTRODUCTION

Response Surface Methodology (RSM) is a sequential statistical tool employed by engineers and statistician for empirical model building, such that the response variable is optimized [1]. RSM consists of three core stages namely, experimental design stage, modeling stage and the optimization stage of the fitted regression models where the settings of the explanatory variables optimize the response(s).

MATERIALS AND METHODS

The parametric regression models are superior if the user can specify a parametric form for the data, otherwise misspecified. The nonparametric regression model is not restricted to a user specified form as in the parametric counterpart. In spite of its flexibility, nonparametric regression models are challenged in a study such as RSM due to idiosyncracies of RSM data namely;

- Sparseness of RSM data
- Cost efficient design (small sample sizes)
- The study utilizes more than one explanatory variable (a term referred to as curse of dimensionality).

The parametric regression model (OLS)
Consider the parametric regression model:

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon}, \tag{1}$$

Where $\boldsymbol{y} = \begin{bmatrix} y_1 \\ y_2 \\ . \\ . \\ . \\ y_n \end{bmatrix}_{(n\times 1)}$ is the vector of response**,**

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

where $\boldsymbol{X} = \boldsymbol{X^{(OLS)}}$**,** $\boldsymbol{\beta}$ is the unknown parameter vector and $\boldsymbol{\varepsilon}$ is the vector of error term.

The estimated responses for the i^{th} data points are:

$$\widehat{\boldsymbol{\beta}}^{(OLS)} = \left(\boldsymbol{X'^{(OLS)}X^{(OLS)}}\right)^{-1}\boldsymbol{X'^{(OLS)}y},$$

$$\widehat{\boldsymbol{y}}_i^{(OLS)} = \boldsymbol{x_i'^{(OLS)}}\left(\boldsymbol{X'^{(OLS)}X^{(OLS)}}\right)^{-1}\boldsymbol{X'^{(OLS)}y}, \quad i = 1,2,\dots,n. \tag{2}$$

In matrix form, equation (2) is expressed as:

$$\widehat{\boldsymbol{y}}^{(OLS)} = \boldsymbol{H}^{(OLS)}\boldsymbol{y} = \begin{bmatrix} \boldsymbol{h_1}^{(OLS)} \\ \boldsymbol{h_2}^{(OLS)} \\ \vdots \\ \boldsymbol{h_n}^{(OLS)} \end{bmatrix}\boldsymbol{y}, \tag{3}$$

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

The disadvantage of the parametric regression model is that if misspecified, the estimates are usually biased [2, 3].

The Local Linear Regression model (LLR)

Using the weighted least squares theory [4], the LLR estimator $\hat{\boldsymbol{y}}_i^{(LLR)}$ is given as:

$$\hat{\boldsymbol{y}}_i^{(LLR)} = \boldsymbol{x}_i'^{(LLR)}\left(\boldsymbol{X}'^{(LLR)}\boldsymbol{W}_i\boldsymbol{X}^{(LLR)}\right)^{-1}\boldsymbol{X}'^{(LLR)}\boldsymbol{W}_i\boldsymbol{y},\ i = 1,2,\dots,n \quad (4)$$

where $\boldsymbol{X}^{(LLR)}$ is the LLR model matrix that depends solely on the number of explanatory variables utilized in the experiment, $\boldsymbol{W}_i = \boldsymbol{W}^{\boldsymbol{Raw}}$ is the diagonal matrix of kernel (Gaussian) weight used in the estimation of the i^{th} response and $\boldsymbol{x}_i'^{(LLR)}$ is the i^{th} row of the LLR model matrix.

In terms of locations, the LLR estimator is expressed as:

$$\hat{\boldsymbol{y}}_i^{(\boldsymbol{LLR})} = \boldsymbol{h}_i'^{(LLR)}\boldsymbol{y},\ i = 1,\ 2\ ,\ \dots,\ n \quad (5)$$

The shortcoming of the LLR model is that it suffers high bias in region where the data exhibit curvature [5, 6].

Bandwidths for nonparametric and semiparametric regression models

The choice of bandwidths for nonparametric and semiparametric regression models is a critical criterion and challenging in regression analysis [7, 8]. Bandwidth selection was designed to minimize bias and variance of the estimate [6].

A bandwidth b, is said to be fixed if it's value is constant for all the locations in a given regression method, else it is referred to as locally adaptive bandwidths [9].

Hence, the kernel function, $K(.)$ employed in RSM is the simplified Gaussian kernel given in [10] as:

$$K\left(\frac{x_i-x_0}{b}\right) = K\left(\frac{x_0-x_i}{b}\right) = e^{-\left(\frac{x_i-x_0}{b}\right)^2}, i = 1,\ 2,\ .\ .\ .\ ,\ n. \quad (6)$$

where the kernel weights w_{i0} in the kernel weight matrix is given as:

$$w_{i0} = \frac{K\left(\frac{x_i-x_0}{b}\right)}{\sum_{j=1}^{n} K\left(\frac{x_j-x_0}{b}\right)},\ \ i = 1,\ \ 2\ ,\ .\ .\ .\ .\ ,n\ . \quad (7)$$

According to [11], $K\left(\frac{x_i-x_0}{b}\right)$ in Equation (7) is referred to as kernel function which regulates the shape of the kernel weights (e.g. Gaussian kernel), x_0 is a dummy known as target point, b is the bandwidth.

A situation where more than one explanatory variable are used in the model matrix $\boldsymbol{X}^{(LLR)}$, the kernel weight w_{i0} is a product from simplified Gaussian kernel given as:

$$w_{i0} = \prod_{j=1}^{k} K\left(\frac{x_{ij}-x_0}{b}\right)\Big/\sum_{p=1}^{n}\left[\prod_{j=1}^{k} K\left(\frac{x_{pj}-x_0}{b}\right)\right],\ \ i = 1,\ \ .\ .\ .\ ,\ n\ . \quad (8)$$

[4, 12].

For data originated from RSM, the vector of optimal bandwidths $\Omega = [b_1^*,\ b_1^*, \dots, b_n^*]$ is obtained based on the minimization of the Penalized Prediction Error Sum of Squares ($PRESS^{**}$) [10].

The $PRESS^{**}$ criterion for selecting the bandwidths is given as:

$$PRESS^{**}(\Omega) = \frac{\sum_{i=1}^{n}\left(y_i-\hat{y}_{i,-i(.)}\right)^2}{n-trace\left(H^{(.)}(\Omega)\right)+(n-k-1)\frac{SSE_{max}-SSE_{\Omega}}{SSE_{max}}}, \quad (9)$$

where SSE_{max} is the maximum Sum of Squared Errors obtained as the $b_1, b_2, \dots, b_n$ approaches infinity, SSE_{Ω} is the sum of squared errors associated with a set of bandwidths $b_1, b_2, \dots, b_n$, $tr\left(H^{(.)}\Omega\right)$ is the trace of the Hat matrix and $\hat{y}_{i,-i}$ is the leave-one-out cross-validation estimated value of y_i with the i^{th} observation left out [10, 12].

Semiparametric regression models

The semiparametric regression models combines estimates of both parametric and nonparametric regression models either to the raw (observed) data or to the residuals via the mixing parameters, λ [7, 13]. Semiparametric regression models in recent times have received great attention due to its flexibility and wide range of applications to discipline such as Engineering, Pharmacology, Ecominomis and Industries [14].

Model Robust Regression 1 *(MRR1)*

An operative and simple model that discourses the drawbacks inherent in both parametric and nonparametric regression models is the use of one semiparametric regression model, model robust regression 1 (MRR1) that combines estimate of parametric and nonparametric regression models, both to the observed data in the most adequate proportion in a convex combination via the mixing parameter, λ[15].

The goal of MRR1 is to offer a smooth estimation of the mean response that can capture important patterns or trends in the data, in that way reducing bias while at the same time reducing variance [13].

The mathematical expression for the MRR1 as given in [16], and [17] as:

$$\hat{y}^{(MRR1)} = \lambda\hat{y}^{(LLR)} + (1-\lambda)\hat{y}^{(OLS)} \quad (10)$$

where the parameter λ is the mixing parameter with an interval [0, 1]. The purpose of λ is to combine estimates of parametric and nonparametric regression models in more efficient manner, such that as λ increases from zero to one, model misspecification also increases for the data. On the other-hand, when $\lambda = 0$, MRR1 is purely estimates of OLS, whereas, when $\lambda = 1$, MRR1 is estimates of LLR.

The optimal value λ^* of λ is selected similar to bandwidth, based on the minimization of the $PRESS^{**}$ criterion:

$$PRESS^{**}(\Omega, \lambda) = \frac{\sum_{i=1}^{n}\left(y_i - \hat{y}_{i,-i}^{(.)}(\Omega,\lambda)\right)^2}{n - trace\left(H^{(.)}(\Omega,\lambda)\right) + (n-k-1)\frac{SSE_{max} - SSE_{\Omega}}{SSE_{max}}}, \quad (11)$$

where $\Omega=[b_{1j}^*, b_{2j}^*, \dots, b_{nj}^*]$ is the vector of optimal bandwidths, $SSE_{\boldsymbol{\Omega}}$ is the Sum of Squared Errors associated with the set of the optimal bandwidths, $[b_{1j}^*, b_{2j}^*, \dots, b_{nj}^*]$, $trace\left(H^{(.)}(\Omega,\lambda)\right)$ is the trace of the Hat matrix, and $\hat{y}_{i,-i}^{(.)}(\Omega,\lambda)$ is the leave-one-out cross-validation estimate of y_i [10, 12].
The MRR1 suffers from the problem related with convex combination [4, 12].

Model Robust Regression 2 *(MRR2)*
MRR2 combines estimates of parametric regression model to the raw data, while the aspect of nonparametric regression model uses the LLR Hat matrix to fit the residuals from the estimates of parametric regression model via a mixing parameter, λ. In this study, the estimate of the parametric regression model is the use of OLS. Nevertheless, other parametric and nonparametric regression models could be exploited in a study such as RSM [4, 18].

The MRR2 was developed by [12] and is expressed as:

$$\hat{\boldsymbol{y}}^{(MRR2)} = \hat{\boldsymbol{y}}^{(OLS)} + \lambda\hat{\boldsymbol{r}}^{(LLR)},$$
$$\hat{\boldsymbol{r}}^{(LLR)} = \boldsymbol{H}_{\boldsymbol{r}}^{(LLR)}\boldsymbol{r} \quad (12)$$

$\boldsymbol{r} = \boldsymbol{y} - \boldsymbol{y}^{OLS}$ is the vector of residuals that represents the structure in the data not captured by the user specified parametric regression model. The vector of residuals, $\boldsymbol{r}$ is used to fit the LLR Hat matrix, $\boldsymbol{H}_{\boldsymbol{r}}^{(LLR)}\boldsymbol{r}$ resulting to vector of smoothed or fitted residuals, $\hat{\boldsymbol{r}}^{(LLR)}$. The role of λ is such that if the parametric fit is adequate, then $\lambda = 0$, otherwise estimates of nonparametric fit is added back to the parametric fit via λ in the interval (0 1] to improve the estimates of MRR2. Also, the choice of λ is selected through the minimization of $PRESS^{**}$ criterion [15].
The MRR2 does not fully profit from the flexibility LLR can offer, since only OLS residuals is utilized in its fitting process [18].

Adaptive bandwidths *(AB)*
According to [19] locally adaptive bandwidths:

$$b_i = \frac{b^* N\left(C[\sum_{j=1}^{n} y_j] - y_i\right)}{(Cn-1)\sum_{j=1}^{n} y_j}, \qquad i = 1,2,\dots,n \quad (13)$$

where b^* is a fixed optimal bandwidth, y_i, $i = 1,2,\dots,n$, could be taken as any statistics that mirrors the insufficiencies in the OLS estimates of the responses, $T = \sum_{j=1}^{n} y_j, N > 0$, and $C \geq 0$, are parameters introduced to address the problem of clustering within the interval [0, 1]. The optimal chosen tuning parameters of N and C are hereafter refers to as $N^*$and C^* respectively.

Locally adaptive bandwidths *(PAB)*
According to [20] presented data-driven locally adaptive bandwidths:

$$b_{ij} = T_{1j}\left(\frac{1}{2} - \frac{x_{ij}}{T_{2j}}\right)^2, \quad i = 1,2, \dots, n; j = 1,2,\dots,k \quad (14)$$

where, $0 < b_{ij} \leq 1,$, $T_{1j} > 0$, $T_{2j} > 0$.

The b^*_{ij} of the locally adaptive optimal bandwidths from Equation (14) is obtained at an optimally selected values of T_{1j}, T_{2j}, (hereafter referred to as T_{1j}^* and T_{2j}^*, respectively), $j = 1,2,\dots,k$, based on the minimization of the $PRESS^{**}$ criterion in Equation (9).
where $b_{ij} = b$, is called a fixed bandwidth, otherwise b_{ij}, $i = 1,2, \dots, n; j = 1,2,\dots,k$ are locally adaptive bandwidths.

The locally adaptive bandwidths of Equations (13) and (14) were applied to the two semiparametric regression models of Equations (11) and (12) to obtain a novel blend of semiparametric regression model used to smooth the two data for the application problems.

RESULTS AND DISCUSSION

In this section, we shall apply the OLS and the two semiparametric regression models utilizing the two locally adaptive bandwidths selector to adequately smooth the two RSM data and thereafter discusses the results and give their respective statistical significance.

Application I (single response chemical process data)
The problem of the study as given in [4, 21] was to relate chemical yield (y) to temperature (x_1) and time (x_2) with the intention to maximize the chemical yield. The data is obtained using the Central Composite Design is given in Table 1.

Table 1: Single Response Chemical Process Data generated from the CCD

i	x_1	x_2	y
1	-1	-1	88.55
2	1	-1	85.80
3	-1	1	86.29
4	1	1	80.44
5	-1.414	0	85.50
6	1.414	0	85.39
7	0	-1.414	86.22
8	0	1.414	85.70
9	0	0	90.21
10	0	0	90.85
11	0	0	91.31

Source: [4, 21]

Data transformation using central composite design (CCD)

Following nonparametric regression procedures in RSM, the values of the explanatory variables are coded between 0 and 1. The data collected via a CCD is transformed by a mathematical relation:

$$x_{NEW} = \frac{Min(x_{OLD}) - x_0}{(Min(x_{OLD}) - Max(x_{OLD}))} \quad (15)$$

where x_{NEW} is the transformed value, x_0 is the target value that needed to be transformed in the vector containing the old coded value, represented as x_{OLD}, Min (x_{OLD}) and $Max(x_{OLD})$ are the minimum and maximum values in the vector x_{OLD} respectively, [22]. The natural or coded variables in Table 1 can be transformed to explanatory variables in Table 2 using Equation (15)

Target points needed to be transformed for location 2 under the coded variables are given below:

Target points x_0: 1 , −1; $Min(x_{OLD})$: − 1.414 , −1.414; $Max(x_{OLD})$: 1.414, 1.414

$$x_{NEW} = \frac{Min(x_{OLD}) - x_0}{(Min(x_{OLD}) - Max(x_{OLD}))}$$

$$Explanatory\ variable\ x_1 : x_{21} = \frac{-1.414 - (1)}{((-1.414) - (1.414))} = 0.8536$$

$$Explanatory\ variable\ x_2 : x_{22} = \frac{-1.414 - (-1)}{((-1.414) - (1.414))} = 0.1464$$

Table 2: Single Response Chemical Process Data

i	x_1	x_2	y
1	0.1464	0.1464	88.55
2	0.8536	0.1464	85.80
3	0.1464	0.8536	86.29
4	0.8536	0.8536	80.44
5	0.0000	0.5000	85.50
6	1.0000	0.5000	85.39
7	0.5000	0.0000	86.22
8	0.5000	1.0000	85.70
9	0.5000	0.5000	90.21
10	0.5000	0.5000	90.85
11	0.5000	0.5000	91.31

A second-order model was specified for the parametric technique [4]. The R^2_{adj} from the OLS method using the full second-order model gives 67.77%. Here, the interest is to determine if the amount of variability not explained by the specified model can be reduced by the application of the LLR method.

Table 3: Locally Adaptive Optimal Bandwidths (PAB) for $MRR1_{PAB}$ in the Single Response Chemical Process Data

i	$MRR1_{PAB}$	
	x_1 $T_{11}^* = 1.3151000000000000$ $T_{21}^* = 2.9740000000000000$	x_2 $T_{12}^* = 1.4134000000000000$ $T_{22}^* = 1.0412000000000000$
	b_{i1}	b_{i2}
1	0.2672	0.1826
2	0.0597	0.1826
3	0.2672	0.1446
4	0.0597	0.1446

5	0.3288	0.0006
6	0.0353	0.0006
7	0.1448	0.3534
8	0.1448	0.2996
9	0.1448	0.0006
10	0.1448	0.0006
11	0.1448	0.0006

Table 3, represents varying tuning parameters for $MRR1_{PAB}$ via the locally adaptive bandwidths for the single response chemical process problem.

Table 4: Locally Adaptive Optimal Bandwidths (PAB) for $MRR2_{PAB}$ in the Single Response Chemical Process Data

i	$\boldsymbol{MRR2_{PAB}}$	
	x_1	x_2
	b_{i1} $T_{11}^* = 0.4386583586285935$ $T_{21}^* = 0.4989813147061812$	b_{i2} $T_{12}^* = 1.4145296569455679$ $T_{22}^* = 2.0399967966514610$
1	0.0187	0.2594
2	0.6430	0.2594
3	0.0187	0.0094
4	0.6430	0.0094
5	0.1097	0.0919
6	0.9924	0.0919
7	0.1106	0.3536
8	0.1106	0.0001
9	0.1106	0.0919
10	0.1106	0.0919
11	0.1106	0.0919

Table 4., represents varying tuning parameters for $MRR2_{PAB}$ via the locally adaptive bandwidths (PAB) for the single response chemical process problem.
The fixed mixing parameters for the models, $MRR1_{PAB}$, $MRR2_{PAB}$, and the $MRR2_{AB}$ as obtained via genetic algorithm tool in Matlab are presented in Table 5.

Table 5: Mixing Parameters of different models for Single Response Chemical Process Data

Response	**Model**	$\boldsymbol{\lambda}$
y	OLS	NOT APPLICABLE
	$MRR1_{PAB}$	0.9588225842457119
	$MRR2_{PAB}$	1.0000000000000000
	$MRR2_{AB}$	NOT FIXED (Edionwe *et al*. 2016)

Table 6: Comparison of the goodness-of-fit statistics of each method for the Chemical Process Data

METHOD	$\boldsymbol{b^*}$	$\boldsymbol{DF_{error}}$	$\boldsymbol{MSE}$	$\boldsymbol{SSE}$	$\boldsymbol{R^2}$	$\boldsymbol{R^2_{adj}}$	$\boldsymbol{PRESS}$	$\boldsymbol{PRESS^*}$	$\boldsymbol{PRESS^{**}}$
OLS	-	5.000	3.1600	15.8182	83.8800	67.7700	109.5179	21.9036	21.9036

$MRR1_{PAB}$	*	2.1751	0.3771	0.8203	99.1644	96.1584	45.6825	**21.0028**	4.5288
$MRR2_{PAB}$	*	2.0000	0.3053	**0.6107**	**99.3800**	96.8900	**43.2389**	21.6194	**4.4617**
$MRR2_{AB}$	**	2.4057	**0.2988**	0.7189	99.2783	**97.0000**	72.1981	30.0116	7.1896

In Table 6, $MRR1_{PAB}$ and $MRR2_{PAB}$ that uses PAB performed better in terms *PRESS*, $PRESS^{*}$, $PRESS^{**}$ R^2 and SSE statistics, whereas $MRR2_{AB}$ that uses AB performed better in respect to *MSE* and R^2_{adj} statistics. "*" and "**" represents PAB and existing AB respectively.

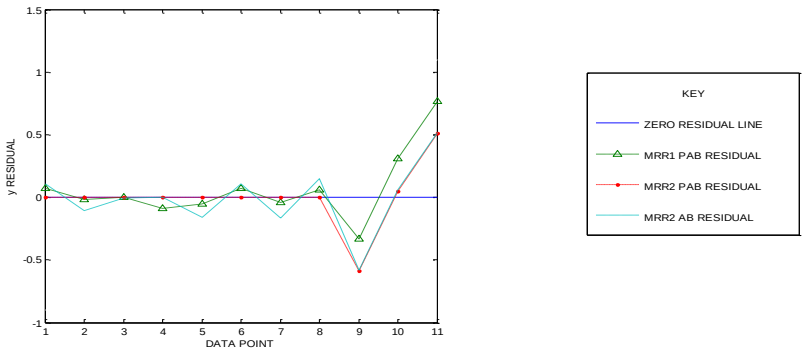

Figure 1: Residual plot for single response chemical process data

Figure 1, is the residual plots for the six models as specified in the KEY for single response chemical process data. Obviously, $MRR2_{PAB}$ that utilizes the PAB estimated the data better in the overall residual performance.

Table 7: Comparison of optimization results for the Chemical Process Data

Approach	x_1	x_2	$\hat{y}$
OLS	0.4393000000000000	0.4361000000000000	90.9783
$MRR1_{PAB}$	0.0006672519293997789	0.0017565336318757729	89.2913
$MRR2_{PAB}$	**0.26048040669568984**	**0.8008798602146169**	**91.5727**
$MRR2_{AB}$	0.3828000000000000	0.3921000000000000	91.4237

The model, $MRR2_{PAB}$ perform better than in terms of maximum chemical yield for single response chemical process data as given in Table 7. Obviously, $MRR2_{PAB}$ has a better experimental relationship between temperature (x_1) and time (x_2) as it relates to chemical yield.

Application II (the multiple response chemical process data)

This problem is analyzed in [22, 23]. The aim of the study is 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 target 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 target value $\emptyset = 3100$.

The problem as given in [22, 23] was to obtain the settings 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, with target value = 80; y_2 should take a value in the range L = 62 and U = 68 with target value = 65; minimize y_3 with upper limit U = 3300 and target value = 3100.

Based on the process requirements a Central Composite Design (CCD) was conducted to establish the design experiment and observed responses as presented in Table 8.

Table 8: Designed experiment and response values [22, 23]

i	***Experimental variables***		**Responses**		
	x_1	x_2	y_1	y_2	y_3
1	-1	-1	76.5	62	2940
2	1	-1	78.0	66	3680
3	-1	1	77.0	60	3470
4	1	1	79.5	59	3890
5	-1.414	0	75.6	71	3020
6	1.414	0	78.4	68	3360
7	0	-1.414	77.0	57	3150
8	0	1.414	78.5	58	3630
9	0	0	79.9	72	3480
10	0	0	80.3	69	3200
11	0	0	80.0	68	3410
12	0	0	79.7	70	3290
13	0	0	79.8	71	3500

Table 9, is the transformed data from CCD to RSM data using the mathematical relation in Equation (15) that needed to lie between zero and one inclusively.

Table 9: Multiple Response 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

In the multiple response chemical process data as given in section 3.3, we seek to show the performance of OLS, $MRR1_{PAB}$, $MRR2_{PAB}$ and $MRR2_{AB}$ based on the goodness-of-fit statistics and the process requirements.

In Table 10 through Table 15 represents different varying tuning parameters and locally adaptive bandwidths for $MRR1_{PAB}, MRR2_{PAB}$ and $MRR2_{AB}$ for a multi-response chemical process data.

Table 10: Optimal values of tuning parameters and Locally Adaptive Bandwidths for $y_1$using $MRR1_{PAB}$

i	$\boldsymbol{MRR1_{PAB}}$ ***Regression***

	x_1	x_2
	b_{i1} $T_{11}^{*} = 1.0031000000000000$ $T_{21}^{*} = 5.9998000000000000$	b_{i2} $T_{12}^{*} = 1.2694000000000000$ $T_{22}^{*} = 1.0541000000000000$
1	0.2269	0.1655
2	0.1284	0.1655
3	0.2269	0.1218
4	0.1284	0.1218
5	0.2508	0.0008
6	0.1115	0.0008
7	0.1741	0.3174
8	0.1741	0.2555
9	0.1741	0.0008
10	0.1741	0.0008
11	0.1741	0.0008
12	0.1741	0.0008
13	0.1741	0.0008

Table 11: Optimal values of tuning parameters and Locally Adaptive Bandwidths for $y_2$using $MRR1_{PAB}$

i	$\boldsymbol{MRR1_{PAB}}$ **Regression**	
	x_1	x_2
	b_{i1} $T_{11}^{*} = 1.5240000000000000$ $T_{21}^{*} = 4.4808000000000000$	b_{i2} $T_{12}^{*} = 0.4149000000000000$ $T_{22}^{*} = 7.6721000000000000$
1	0.3328	0.0960
2	0.1460	0.0960
3	0.3328	0.0627
4	0.1460	0.0627
5	0.3810	0.0784
6	0.1168	0.0784
7	0.2299	0.1037
8	0.2299	0.0567
9	0.2299	0.0784
10	0.2299	0.0784
11	0.2299	0.0784
12	0.2299	0.0784
13	0.2299	0.0784

Table 12: Optimal values of tuning parameters and Locally Adaptive Bandwidths for $y_3$using $MRR1_{PAB}$

i	$\boldsymbol{MRR1_{PAB}}$ **Regression**	
	x_1	x_2
	b_{i1} $T_{11}^{*} = 3.9999996701204728$ $T_{21}^{*} = 2.2512665995031690$	b_{i2} $T_{12}^{*} = 3.9999996360686363$ $T_{22}^{*} = 9.9999987155096690$
1	0.7568	0.9423
2	0.0584	0.9423
3	0.7568	0.6877
4	0.0584	0.6877
5	1.0000	0.8100
6	0.0125	0.8100
7	0.3089	1.0000
8	0.3089	0.6400
9	0.3089	0.8100

10	0.3089	0.8100
11	0.3089	0.8100
12	0.3089	0.8100
13	0.3089	0.8100

Table 13: Optimal values of tuning parameters and Locally Adaptive Bandwidths for $y_1$using $MRR2_{PAB}$

i	$\boldsymbol{MRR2_{PAB}}$ **Regression**	
	x_1	x_2
	b_{i1} $T_{11}^* = 0.5428430834699849$ $T_{21}^* = 4.0616785539017375$	b_{i2} $T_{12}^* = 1.517132347364459$ $T_{22}^* = 0.9631401273256596$
1	0.1168	0.1837
2	0.0456	0.1837
3	0.1168	0.2264
4	0.0456	0.2264
5	0.1357	0.0006
6	0.0350	0.0006
7	0.0771	0.3793
8	0.0771	0.4396
9	0.0771	0.0006
10	0.0771	0.0006
11	0.0771	0.0006
12	0.0771	0.0006
13	0.0771	0.0006

Table 14: Optimal values of tuning parameters and Locally Adaptive Bandwidths for $y_2$using $MRR2_{PAB}$

i	$\boldsymbol{MRR2_{PAB}}$ **Regression**	
	x_1	x_2
	b_{i1} $T_{11}^* = 3.9589536215421037$ $T_{21}^* = 1.5349728681657138$	b_{i2} $T_{12}^* = 2.251860516028199$ $T_{22}^* = 1.603928067610862$
1	0.6482	0.3762
2	0.0125	0.3762
3	0.6482	0.0023
4	0.0125	0.0023
5	0.9897	0.0798
6	0.0908	0.0798
7	0.1202	0.5630
8	0.1202	0.0343
9	0.1202	0.0798
10	0.1202	0.0798
11	0.1202	0.0798
12	0.1202	0.0798
13	0.1202	0.0798

T**able 15:** Optimal values of tuning parameters and Locally Adaptive Bandwidths for $y_3$using $MRR2_{PAB}$

i	$\boldsymbol{MRR2_{PAB}}$ **Regression**

	x_1	x_2
	b_{i1} T_{11}^* =0.47582949147420456 T_{21}^* =4.8815664001472030	b_{i2} T_{12}^* =1.4226028183585662 T_{22}^* =0.8797899080227980
1	0.1051	0.1583
2	0.0503	0.1583
3	0.1051	0.3146
4	0.0503	0.3146
5	0.1190	0.0066
6	0.0415	0.0066
7	0.0752	0.3557
8	0.0752	0.5766
9	0.0752	0.0066
10	0.0752	0.0066
11	0.0752	0.0066
12	0.0752	0.0066
13	0.0752	0.0066

The fixed mixing parameters for the models, $MRR1_{PAB}$, $MRR2_{PAB}$, and the $MRR2_{AB}$ as obtained via genetic algorithm tool in Matlab are presented in Table 16

Table 16: Mixing Parameters of different models for Multiple Chemical Process Data

Response	**Model**	$\boldsymbol{\lambda}$
y_1	OLS	NOT APPLICABLE
	$MRR1_{PAB}$	1.0000000000000000
	$MRR2_{PAB}$	1.0000000000000000
	$MRR2_{AB}$	NOT FIXED
y_2	OLS	NOT APPLICABLE
	$MRR1_{PAB}$	0.7085329578261933
	$MRR2_{PAB}$	1.0000000000000000
	$MRR2_{AB}$	NOT FIXED
y_3	OLS	NOT APPLICABLE
	$MRR1_{PAB}$	0.9319961759530458
	$MRR2_{PAB}$	1.0000000000000000
	$MRR2_{AB}$	NOT FIXED

Table 17: Model goodness-of-fits statistics for Multiple Chemical Process Data

Response	**Model**	$\boldsymbol{DF}$	$\boldsymbol{PRESS}^{**}$	$\boldsymbol{PRESS}$	$\boldsymbol{SSE}$	$\boldsymbol{MSE}$	$\boldsymbol{R^2}$**(%)**	$\boldsymbol{R^2_{Adj}}$**(%)**
y_1	OLS	7.0000	-	-	0.4962	0.4962	98.2733	97.0400
	$MRR1_{PAB}$	4.0144	**0.0481**	**0.6687**	0.2165	0.0539	99.2469	97.7489
	$MRR2_{PAB}$		0.0984	0.9548	**0.2131**	0.0533	**99.2600**	97.7800
	$MRR2_{AB}$	5.3254	0.2691	2.6547	0.2711	**0.0509**	99.0547	**97.8700**
y_2	OLS	7.0000	-	-	36.2242	5.1749	89.9725	82.8100
	$MRR1_{PAB}$	4.8751	**7.5752**	**107.9471**	12.2280	2.5083	96.6149	91.6676
	$MRR2_{PAB}$	4.0000	9.7470	109.5441	**10.0023**	2.5006	**97.2300**	91.6900
	$MRR2_{AB}$	5.3830	16.7814	197.1175	12.8545	**2.3880**	96.4427	**92.0700**
y_3	OLS	7.0000	-	-	207870	29696	75.8967	58.6800
	$MRR1_{PAB}$	6.5922	**26522**	**361000**	79164	**12009**	90.8216	**83.2923**
	$MRR2_{PAB}$	4.0000	50382	545270	66047	16512	**92.3400**	77.0300
	$MRR2_{AB}$	4.9155	52728	593130	68698	13976	92.0369	80.5600

In Table 17, $MRR1_{PAB}$, outperformed *OLS*, $MRR2_{PAB}$ and $MRR2_{AB}$ in terms of *PRESS* and *PRESS*** with respect to chemical yield (y_1). Whereas, $MRR2_{AB}$ performed better than other models considered in terms of *MSE* and R^2_{Adj} for chemical yield (y_1) and $MRR2_{PAB}$ performed better with respect to *SSE* and R^2 than *OLS*, $MRR1_{PAB}$ for chemical yield (y_1). For viscosity (y_2), $MRR1_{PAB}$ outperformed $MRR2_{PAB}$ and $MRR2_{AB}$ in terms of *PRESS* and *PRESS*** statistics whereas, and, $MRR2_{PAB}$ performed better over *OLS*, $MRR1_{PAB}$ and $MRR2_{AB}$in terms of SSE and R^2, and $MRR2_{AB}$ performed better over $MRR1_{PAB}$ and $MRR2_{PAB}$ in terms of MSE and R^2_{Adj}. In terms of molecular weight (y_3), $MRR1_{PAB}$ outperformed *OLS*, $MRR2_{PAB}$ and $MRR2_{AB}$ in terms of *PRESS*, *PRESS***, *MSE* and R^2_{Adj}. Whereas, in terms of R^2 statistics, $MRR2_{PAB}$ performed better than $MRR1_{PAB}$ and $MRR2_{AB}$. Therefore, in the overall goodness of fit statistics $MRR1_{PAB}$ performed better than *OLS*, $MRR2_{PAB}$and $MRR2_{AB}$.

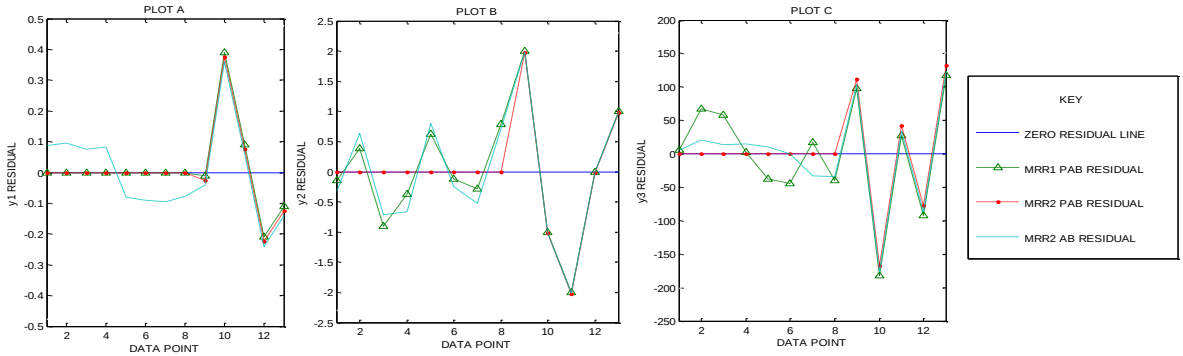

Figure 2: Plot A, Maximize Chemical yield; Plot B, a two sided transformation for viscosity; Plot C, minimize molecular weight

Figure 2 is basically the residual plots for different models as given in the KEY for which the model $MRR1_{PAB}$ performed better in terms of minimum residual points, meaning $MRR1_{PAB}$ estimated the data better than other models considered.

Table 18: Model optimal solution based on the Desirability function for multiple chemical Process Data

Model	x_1	x_2	$\hat{y}_1$	$\hat{y}_2$	$\hat{y}_3$	d_1	d_2	d_3	D(%)
OLS	0.444900000000 0000	0.222600000000 0000	78.76 16	66.48 27	3229.90 00	0.17 44	0.50 58	0.35 04	31.38 00
$MRR1_{PAB}$	**0.545836619093 2497**	**0.146400485050 63175**	79.35 68	64.79 11	3196.80 00	0.57 12	0.93 04	0.51 62	**64.97 58**
$MRR2_{PAB}$	0.536038290847 3450	0.229653107038 3355	78.78 80	66.42 83	3193.70 00	0.19 20	0.52 39	0.53 14	37.67 23
$MRR2_{AB}$	0.515900000000 0000	0.211000000000 0000	78.86 06	66.12 67	3158.20 00	0.24 04	0.62 44	0.70 90	47.39 00

In Table 18, the proposed model $MRR1_{PAB}$ satisfies the choice of process requirements for a multiple chemical process data. Hence, the overall desirability with the highest percentage gives the best production requirements.

CONCLUSION

In RSM, the stages are sequential, a new product is exposed to experimental design phase, modeling phase of the fitted regression model and the optimization phase with the aim to find setting of the explanatory variables that optimize responses as it relates to the quality of the new product. This sequential procedure is referred to as product qualification in the manufacturing industries, see [1, 24].

We have looked at four regression models such as OLS, $MRR1_{PAB}$, $MRR2_{PAB}$ and $MRR2_{AB}$for two RSM data. For the single response chemical problem, $MRR2_{PAB}$ in terms of goodness-of-fit statistics gives the overall best performance in the data analysis. Whereas, in the residual plot, , $MRR2_{PAB}$ residual gave the smallest residual line. In terms of the optimization results, $MRR2_{PAB}$ optimizes the response (chemical yield) better with a maximum chemical yield of **91.5727%** with respect to temperature and time.

In the multi-response chemical process data, $MRR1_{PAB}$ performed better in terms of goodness-of-fit statistics which by implication gives a better

explanation to the data. For the residual plot, $MRR1_{PAB}$ gives a residual line that is closer to the zero residual line. Lastly, the $MRR1_{PAB}$ satisfies the process requirement for the reaction time and temperature that optimizes three chemical solutions representing yield, viscosity and molecular weight respectively with overall desirability of **64.9758%**.

ACKNOWLEDGEMENTS

I sincerely want to thank my Ph.D. supervisor, Prof. J.I. Mbegbu for his tutelage and Dr. E. Edionwe for introducing me to my Ph.D. research area – Nonparametric Regression Models for Response Surface Methodology (RSM). Also, my sincere and cordial thanks to the reviewers.

REFERENCES

1. NAIR, A.T., MAKWANA, A.R. & AHAMMED, M.M. (2014). The use of Response Surface Methodology for modelling and analysis of water and waste – water treatment processes: *A Review. Water Science and Technology*, **69**(3): 464 – 478.
2. SWAMY, P.A.V.B., TAVLAS, G.S. & HALL, S.T. & HONDROYIAN, N.I.S. (2008). Estimation of Parameters in the Presence of Model Misspecification and Measurement Error. *University of Leicester, Working Paper No. 08/27*.
3. FATHI, H.K., MOGHADAM, M.B., AHMADI, N., FAYAZ, M. & NAVAEE, M. (2011). Application of Nonparametric Optimization Methods for Dyeing of Wool. *Middle-East Journal of Scientific Research*, **9**(2): 270 – 278.
4. PICKLE, S.M. (2006)."Semiparametric Techniques for Response Surface Methodology". Ph.D. Dissertation. *Department of Statistics, Virginia Polytechnic Institute & State University, Blacksburg, VA*.
5. HASTIE, T., TIBSHIRANI, R. & FRIEDMAN, T. A. (2009). The Elements of statistical learning, data mining, Inference and Prediction. *Springer Series in Statistics*. *Second Edition*.
6. RIVERS, D.L. (2009). A Graphical Analysis of Simultaneously Choosing the Bandwidth and Mixing Parameter for Semiparametric Regression Techniques, M. Sc. Thesis and Dissertations, *Virginia Commonwealth University.*
7. KAI, B., LI, R. & ZOU, H. (2010). Local Composite Quantile Regression Smoothing: an efficient and safe alternative to the local polynomial regression", *Journal of the Royal Statistical Society*, **72**: 49 – 69.
8. AYDIN, D., MEMMEDLI, M. & OMAY, R.E. (2013). Smoothing parameter selection for nonparametric regression using smoothing spline. *European Journal of Pure and Applied Mathematics*, **6**(2): 222 –238.
9. PREWITT, K. & LOHR, S. (2006). Bandwidth selection in local polynomial regression using eigenvalues. *Journal of Royal Statistical Society, Series B,* **68**(1): 135 – 154.
10. WAN, W. & BIRCH, J.B. (2011). A semi-parametric technique for multi-response optimization. *Journal of Quality and Reliability Engineering. International*, **27**(1): 47-59.
11. WAN, W. (2007). Semi-parametric techniques for multi-response optimization. Ph.D Dissertation submitted to the faculty of the *Virginia Polytechnic Institute and State University, Blacksburg, Virginia.*
12. MAYS, J., BIRCH, J.B. & STARNES, B.A. (2001). Model robust regression: combining parametric, nonparametric and semiparametric methods. *Journals of Nonparametric Statistics*, **13**: 245 – 277.
13. WATERMAN, M.J.T. (2002). Linear mixed model robust regression. Ph.D Dissertation submitted to the faculty of the *Virginia Polytechnic Institute and State University, Blacksburg, Virginia.*
14. FAN, Y.Q. and ULLAH, A. (1999). Asymptotic normality of a combined regression estimator. *Journal of Multivariate Analysis*, **71**(2): 191-240.
15. WALKER, E.L., STARNES, B.A., BIRCH, J.B. & MAYS, J.E., (2002). Model Robust Calibration: Methods and Application to Electrically-Scanned Pressure Transducers, *Technical Report, Dept. of Mathematics, Virginia Commonwealth University, Richmond, VA, USA.*
16. EINSPORN, R. (1987). HATLINK: A Link Between Least Squares Regression and Nonparametric Curve Estimation., Ph.D. dissertation, *Department of Statistics, Virginia Polytechnic Institute and State University, Blacksburg, VA*.
17. EINSPORN, R. (1993). Model robust regression: Using Nonparametric Regression to improve Parametric Regression Analysis. Technical Report 93-5, *Department of Statistics, Virginia Polytechnic Institute and State University, Blacksburg*.
18. PICKLE, S.M., ROBINSON, T.J., BIRCH, J.B. & ANDERSON-COOK, C.M. (2008). A semi-parametric method to robust parameter design. *Journal of Statistical Planning and Inference*, **138**: 114-131.
19. EDIONWE, E., MBEGBU, J.I. & 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.

20. EGUASA, O., EDIONWE, E. & 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

21. MYERS, R.H. & MONTGOMERY, D.C. (2002). Response Surface Methodology: Product and Process Optimization Using Designed Experiments. *Second Edition, John Wiley & Sons, New York*.

22. HE, Z., ZHU, P.F., & PARK, S.H. (2012). A robust desirability function for multi-response surface optimization. *European Journal of Operational Research*, **221**: 241-247.

23. HE, Z., WANG, J., OH, J. & PARK, S.H. (2009). Robust optimization for multiple responses using response surface methodology, *Applied Stochastic Models in Business and Industry*, **26:** 157 – 171.

24. NAJAFI, S., SALMASNIA, A. & KAZEMZADEH, R.B. (2011). Optimization of robust design for multiple response problem. Australian Journal of Basic and Applied Sciences, **5**: 1566 – 1577.