ISSN No:-2456-2165

Response Surface Methodology (RSM): An Application to Cocoa Yield for Microclimate Regulation

[1]David Umolo
Department of Statistics
Federal Polytechnic Orogun, Delta State, Nigeria

[2]George. O. Edeki
Survey and Design of Investigation Unit, Department of Statistics University of Ibadan, Ibadan, Nigeria

Abstract:- This work investigates the impact of temperature and rainfall on the yield of cocoa using an experimental dataset from the Cocoa Research Institute of Nigeria, Ibadan (n=213) with a blocking factor having 4 levels.

The data were analyzed using exploratory data analysis and the response surface methodology. The exploratory data analysis relationship/distributional plot shows that there exists significant negative relationship between the yield of cocoa and the predictors (temperature and rainfall). The estimated boxplot with respect to blocking factor indicates that there is presence of outlier in the yield of cocoa with majority of the yields measured below 500kg over the period of study.

Results from the response surface models without blocking indicate that all the estimated models were statistically significant with all the lack of fit test estimated to be insignificant (an indication of good fit). On the basis of incorporated blocking factor to the experiment, we observed that all models which range from first order to second order outperformed those without blocking factor by considering the estimated adjusted R^2. The blocking factors incorporated into the experiment were found to be statistically significant with all contour plots on the basis of the Eigen analysis suggesting insignificant lack of fit. This implies that incorporating blocking factor helped minimize the sum of squared error and in turn improved the precision.This study recommends that CRIN and other cocoa farmers should learn to adopt newly developed techniques that could militate against the impact of weather change being experienced.

Keywords:- *Response surface model, cocoa yield, rainfall, temperature, Eigen-analysis, contour, lack of fit.*

I. INTRODUCTION

Cocoa bean being the seed of the Cacao tree (also known as Theobroma Cacao) is a tropical indigenous to the equatorial region of the Americas. Among the extracts from cocoa bean is the fluid paste or liquor wherewith chocolate and cocoa powder are generated.

Billions of people across the globe gain benefits from utilizing cocoa and chocolate products. The growing of cocoa could be observed at 300m above average sea level and in that vein requires a minimum of 90 to 100mm rainfall per month with an annual rainfall of 1500 to 2000mm.

It is of great interest to note that cocoa need equitable climate with well distributed rainfall and if dry periods are prolonged, irrigation scheduling will become highly recommendable. Thus, in an ideal setting, the most considerable temperature range is 15° to 39° whereby 25° is identified as optimum preferred ideal temperature(Fisal Ahmad et al,2013)

Cocoa is an essential agricultural farm produce for economic advancement. Its depletion might be greatly attributed to its exposure to temperature and some other environmental parameters. High temperature decreases the life span of leaves, thereby increasing the speed of seed ripening and in turns, cocoa butter is hardened.

No doubt, pests that thrive under the new temperature condition may then pose an additional threat to the yield. The farmers are then expected to pay much attention to resilience, and the agroforestry production system in order to mitigate the negative direct and indirect effects of temperature and rainfall changes so as to regulate microclimate via any plausible standard approach to tackling and handling such situation.

Studies on the utilization of response surface methods for investigating the impacts of rainfall and temperature on cocoa yield are very limited and thus much attention will be drawn to the methodological reviews rather than the conceptual.

Study on the extraction of phenolic compounds from cocoa shell using RSM in conjunction with Artificial Neural Network was carried out in 2021 to model and optimize a green extraction method via a strategy to revolutionize the by-product, thereby obtaining novel high-value product (Miguel Reboillo-Hernanzet et al., 2021).

The optimization of nitric acid mediated extraction of pectin from cocoa pod husks was ascertained using response surface methodology. It is worthwhile to have discovered from the study variables which influenced the nitric acid extraction and hence the optimization was done such that the yield was optimized by increasing temperature among others. (Lucia Cristina Vriesmann et al., 2021).

The development of free and encapsulated Arjuna her extract added vanilla chocolate diary by using response surface methodology. This was achieved by optimizing both drinks containing free herbs and drinks containing encapsulated herb by incorporating alcoholic extract of

Terminalia arjuna i.e. free and encapsulated forms, cocoa powder and sugar into milk with the help of Central Composite Rotatable Design (CCRD) of the response surface methodology (Pravin Diagambarsawale et al., 2020).

The joint impacts of temperature, pH, and water activity on the growth of Trichoderma asperellum was carried out using response surface methodology and finding shows that water activity is more critical in determining the growth as compared to its counterparts but low water activity could be responsible for hindering growth in some conditions.

II. DATA AND METHODS

The data for the study are secondary data sourced from the Cocoa Research Institute of Nigeria (CRIN) Ibadan, Oyo State. Nigeria. The dataset is made up a response variable: Cocoa yield (Cy) and two stimuli namely: Temperature (Tmp) and Rainfall (Rnf) with 4 blocking variables termed A_1, A_2, A_3 and A_4 representing the plots in the field of experiment in the area of study for the period of 10 years which span 2011 to 2021.

The methodology adopted for the study is response surface methodology having the capacity to produce result in good fit of the model, reliable estimates of the parameters including pure error parameter estimates. It is a technique that is insensitive to the presence of outliers in dataset and thus promote bocking strategy in experimental design. This study is thereby carried out using this technique to select most plausible model design that minimizes the heterogeneity of the model parameter (Montgomery, 2013).

The first order model with interaction in RSM is considered the first approach to ascertain the set of factors that impacts the response values substantially. It is sometimes called the screening designs. Since two stimuli exists in this study the first order model with interaction is expressed as

$$Cy_i = \beta_0 + \beta_1 \delta Tmp_{i1} + \beta_2 \delta Rnf_{i2} + \beta_{12} \delta Tmp.\delta Rfn$$

where Cy is the actual value of cocoa yield, both δTmp and δRfn are the revised or coded designs unit of Temperature and Rainfall such that $\delta Tmp, \delta Rfn \in [-1, +1]$. $\delta Tmp.\delta Rfn$ denotes the interaction term. Assuming the interaction term does not exist, the model thus may be referred to as the main effects model.

The second order model with interaction is presented as

$$Cy_i = \beta_0 + \beta_1 \delta Tmp_{i1} + \beta_2 \delta Rnf_{i2} + \beta_{11} \delta Tmp^2 + \beta_{22} \delta Rfn^2 + \beta_{12} \delta Tmp.\delta Rfn$$

Via the incorporation of blocking factor, it of interest to note that the models above become more power and reliable. The RSM model for first order with interaction is given as

$$Cy_i = \beta_0 + \beta_1 \delta Tmp + \beta_2 \delta Rfn + \sum_{i=2}^{b} A_i$$

where$b$ is the number of blocks available, A_1 implies the reference while. In this manner the additional term is incorporated into other forms of the RSM explored to account for blocking.

$$Cy_i = \beta_0 + \beta_1 \delta Tmp_{i1} + \beta_2 \delta Rnf_{i2} + \beta_{11} \delta Tmp^2 + \beta_{22} \delta Rfn^2 + \beta_{12} \delta Tmp.\delta Rfn + \sum_{i=2}^{b} A_i$$

A. Research Design Flowchart

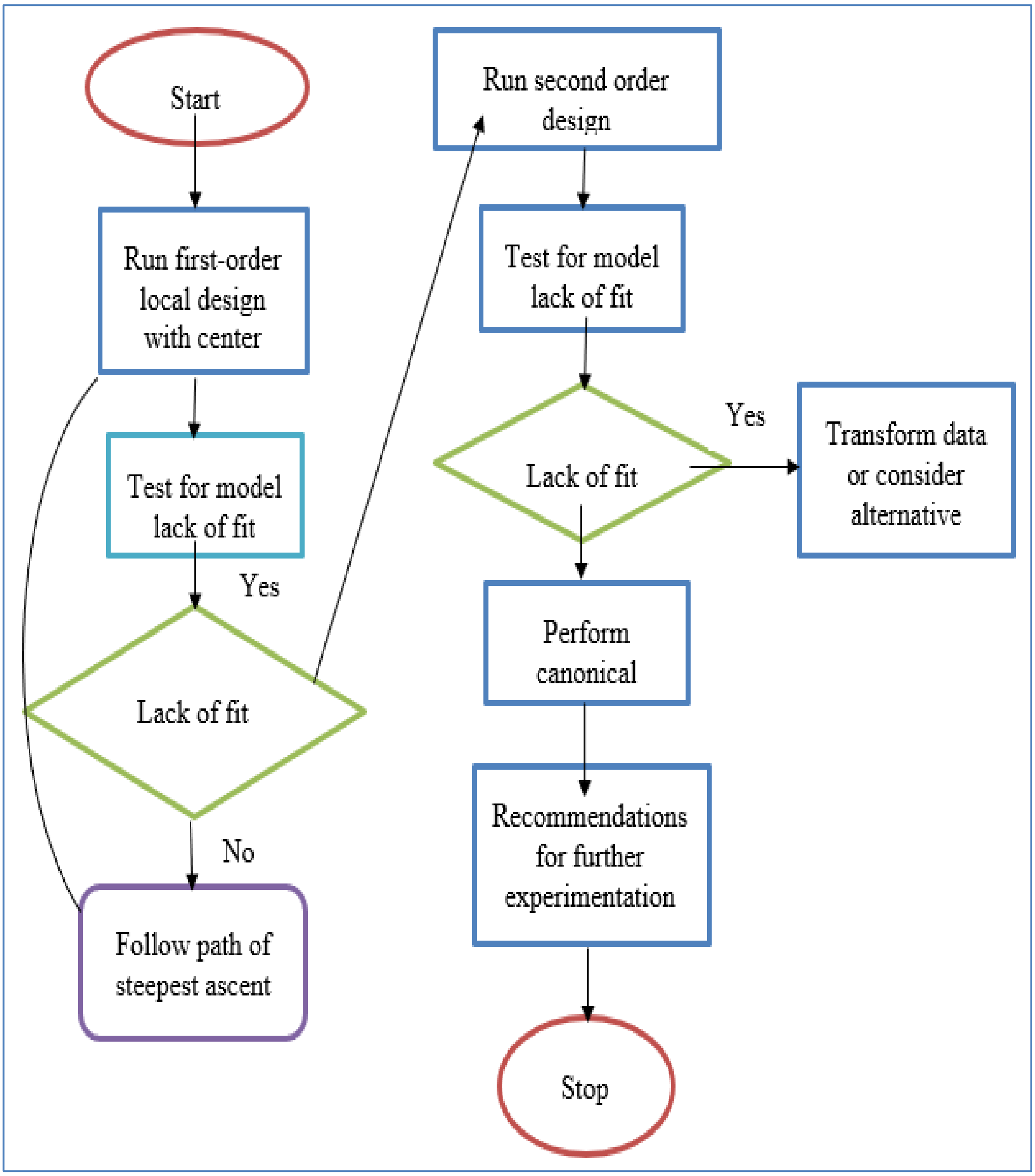


Fig. 1: Research Design Flowchart

B. The Model Structure

The model structure observed in the study includes those of first order response surface method (FO), second order surface method (SO), Pure quadratic surface method (PQ) and Two-Way interaction method (TWI).

This Techniques are explored via the coding structure

$$\delta Tmp = \frac{Max(Tmp) - \frac{[Max(Tmp)+Min(Tmp)]}{2}}{\frac{[Max(Tmp)-Min(Tmp)]}{2}}$$

and

$$\delta Rfn = \frac{Max(Rfn) - \frac{[Max(Rfn)+Min(Rfn)]}{2}}{\frac{[Max(Rfn)-Min(Rfn)]}{2}}$$

for $-1 < \delta Tmp, \delta Rfn < +1$ (Kowalski, Cornell, and Montgomery (2005).

III. RESULTS AND DISCUSSION

A. Exploratory Data Analysis

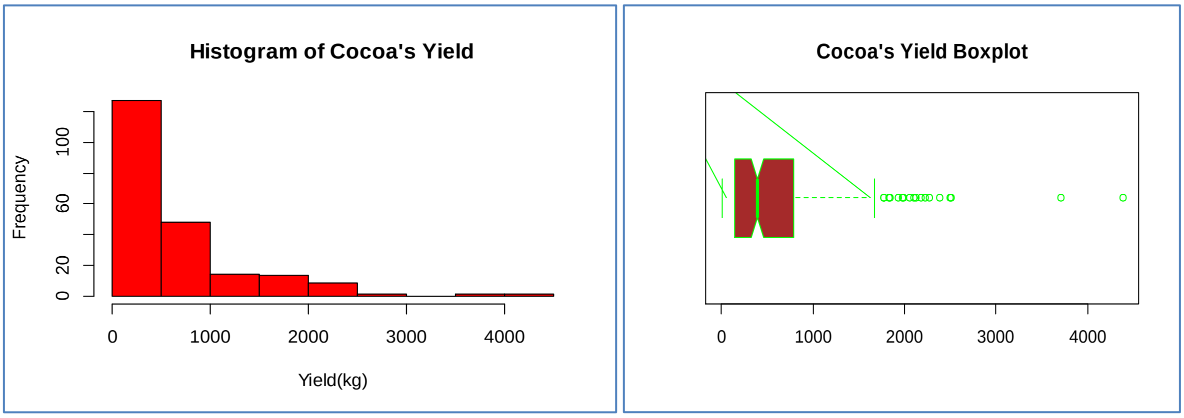


Fig. 1: Distribution of Cocoa yield using histogram and boxplot

The histogram of the cocoa yield indicates that more of the kg of cocoa produced over the period of study is less than 500kg, followed by yield between 50kg and 1000kg. This suggests a left skewed distribution for cocoa yield. The boxplot helps us understand that there is presence of outliers in the yield of cocoa though they are not extreme outliers. The five statistics-summary from the plot were estimated as (3.0, 145.4, 388.7, 789.9, 1668.6) corresponding to (minimum yield, first quartile, median, third quarter, maximum yield).

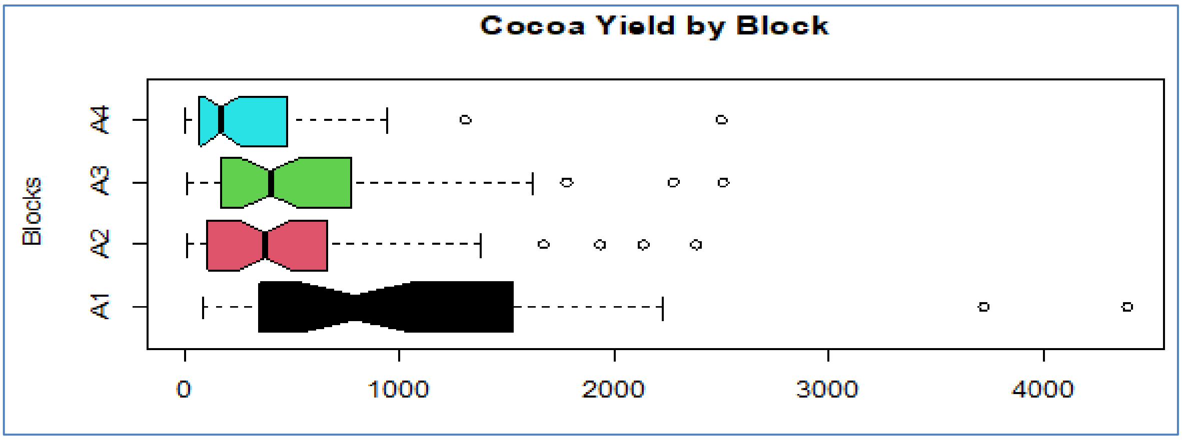


Fig. 2: Distribution of Cocoa Yield by Block

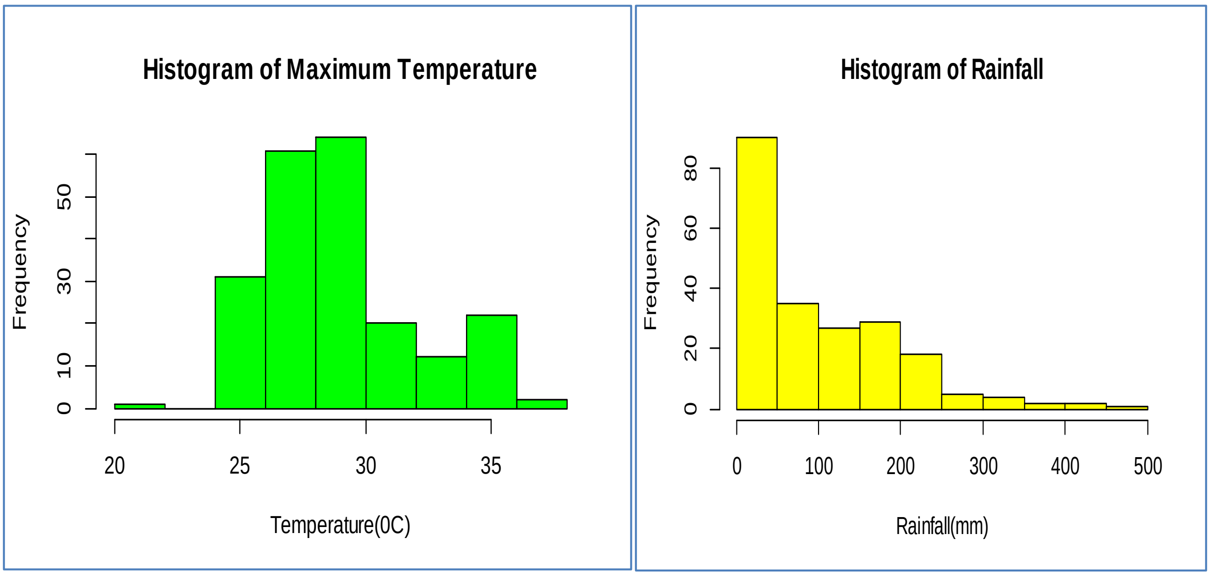


Fig. 3: Distribution of maximum temperature and rainfall

Cocoa yield is found to be highly generated from block A_1 followed by A_3 then A_2. Yield of cocoa is less for block A_4 (see figure 2).The distribution of the maximum temperature onmonthly basis is found to be higher than $37°C$ and minimax temperature is around $20°C$. The highest observed rainfall is $500mm$ while the minimum is slightly above $0mm$. In the sample under study, in comparison with an idea situation, majority of the monthly rainfall account for a measurement below $90mm$ whereas the temperature is higher than $25°C$ ideally in the study area in most of the month. This will of no doubt leave the yield of cocoa to a state of poorness and malnourishment(see figure 3). Nevertheless, the distributions are skewed, it is not a challenge in the case of response surface methodology since it is sophisticated method that overwrites the impact of outliers.

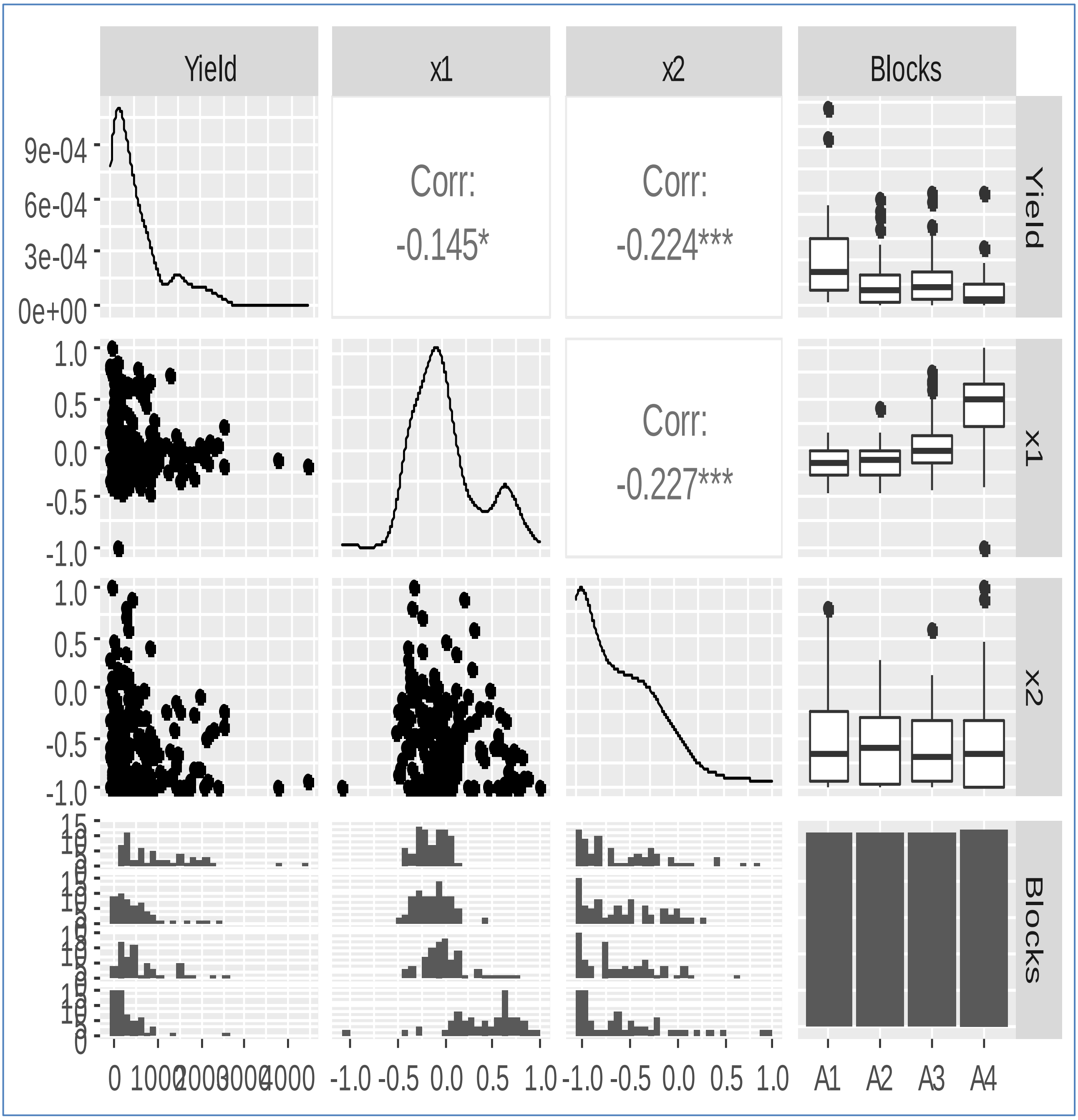


Fig. 4: Distributional and correlation plot of the variables

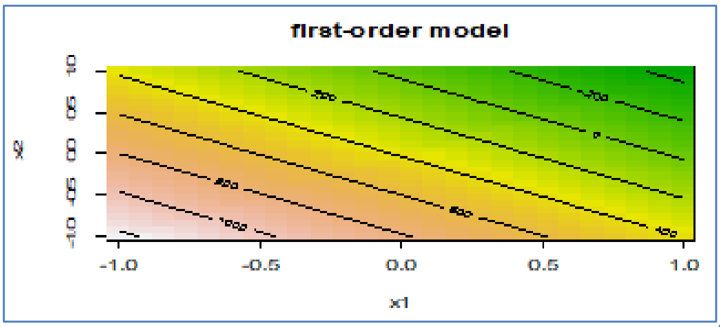
Considering figure 4 the paired relationship plot shows that there exists negative relationship between the yield of cocoa and temperature, also between rainfall and they are both significant at 0.05 level of significance. The relationship between temperature and rainfall were discovered to be significant at 0.05 though not very high indicating no presence of multicollinearity. Cocoa yield in block A_1 is more dispersed as compared to those of other blocks. The average temperature is highest for block A_4 while the distribution of rainfall is approximately evenly likely between the blocks.

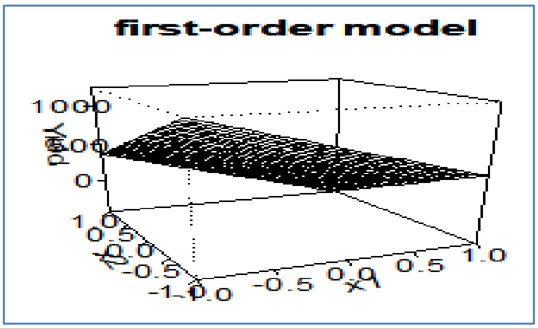
Table 1: Parameter Estimates for Models without Blocking

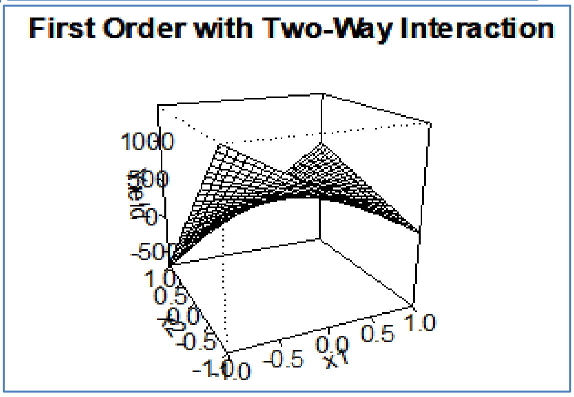
Model	Coefficient	Estimate	Std. Error	t-value	P-value	Decision	AP-v	LoFP-v
FO	**Intercept**	387.173	74.554	5.1932	4.94e-07	***	$4.6e^{-05}$	0.2863
	Tmp (X_1)	-416.209	135.999	-3.0604	0.002499	**		
	Rfn (X_2)	-428.164	106.738	-4.0114	8.4e-05	***		
TWI	**Intercept**	643.30	46.67	13.784	<2e-16	***		
	TWI	525.04	172.30	3.047	0.0026	**		
FO-TWI	**Intercept**	423.343	76.501	5.5338	9.324e-08	***	$FO(4.1e^{-05})$ $TWI(0.0588)$	0.29375
	X_1	59.712	284.611	0.2098	0.8340270			
	X_2	-397.254	107.321	-3.7016	0.0002741	***		
	$X_1:X_2$	674.767	355.118	1.9001	0.0587937	.		
PQ	**Intercept**	493.17	74.78	6.595	3.39e-10	***		
	PQ(X_1,X_2)X_1^2	-1053.65	249.08	-4.230	3.49e-05	***		
	PQ(X_1,X_2)X_2^2	477.98	117.91	4.054	7.10e-05	***		
SO	**(Intercept)**	450.943	80.453	5.6051	6.597e-08	***	$FO\ (3.0e^{-05})$ $TWI(0.055)$ $PQ(0.0154)$	0.3129
	X_1	166.135	283.286	0.5865	0.55821			
	X_2	-246.904	163.408	-1.5110	0.13232			
	$X_1:X_2$	482.698	369.631	1.3059	0.19304			
	X_1^2	-812.760	315.252	-2.5781	0.01063	*		
	X_2^2	269.462	180.726	1.4910	0.13749			

AP-v: ANOVA P-value, LoFP-v: Lack of Fit P-value, "***", "**", "*", "." Indicate significant at 0.1%, 1%, 5%, and 10% respectively. FO: First order model; TWI: Two model with interaction; First order model with interaction; PQ: Pure quadratic model; Second order model with interaction (full model).

Findings as reported in table 1 shows that across all the variant of RSM without blocking factor, it is observed that the first order terms are statistically significant. This shows that both temperature and rainfall are paramount to cocoa yield under normal microclimatic condition. Also, it is discovered that the joint effect of both across the SRM forms is not significant while accounting for individual factor effect within the same model. On the other hand, the Two-Way Interaction model showed that both rainfall and temperature jointly have significant impact on the yield of cocoa during the study period assuming no separate factor effect is considered. The lack of fit (LoFP-v) of the model suggest that the underlined SRM models fit the data accordingly. There is no lack of fit. Considering the quadratic terms, temperature impose a significant negative impact on the yield of cocoa in the studied area.







ISSN No:-2456-2165

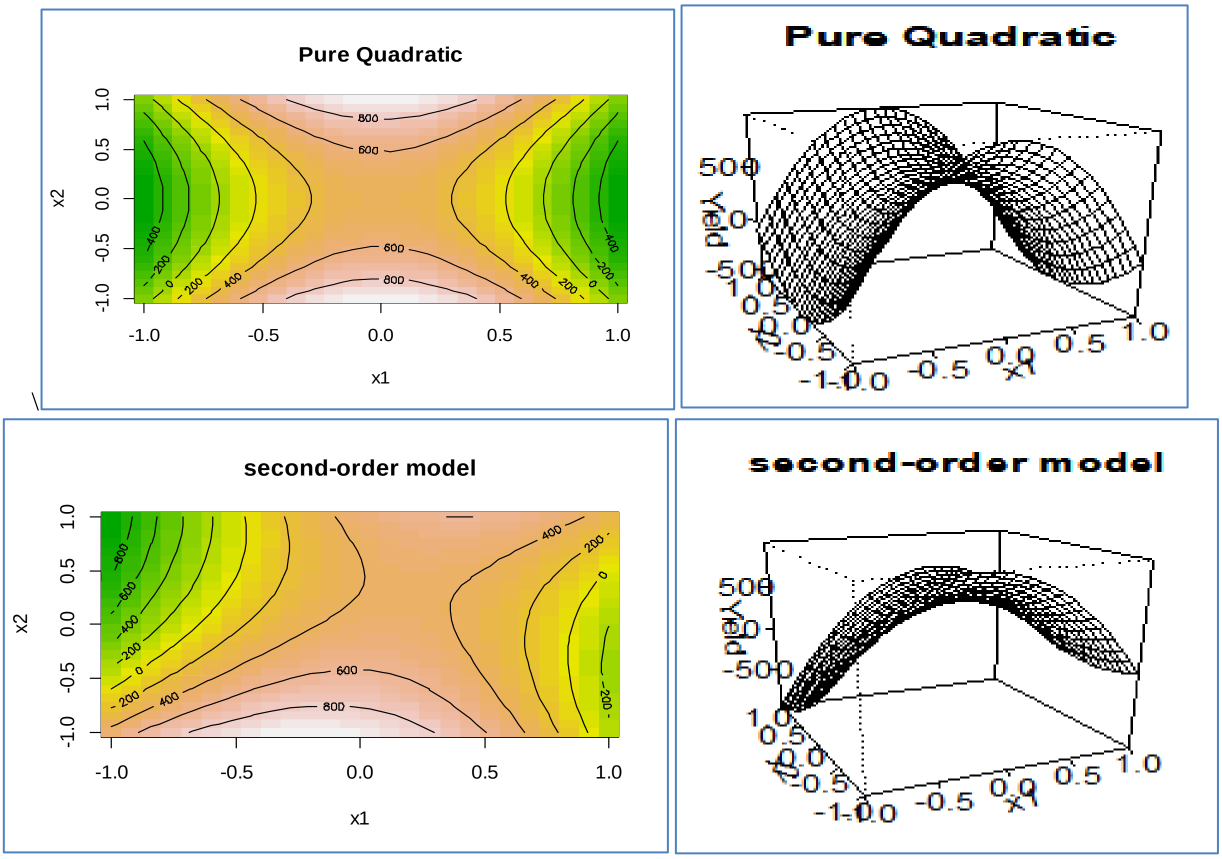


Fig. 5: Surface and contour plot for models without blocking

The surface and contour plots from figure 5 suggest that there is no significant dependency between temperature and rainfall. While the stationary points of the response surface for the first order with interaction are 0.5887 and -0.0885 corresponding to eigenvalue and eigenvectors (337.3837 [-0.2082, -0.9781]), (-37.3837[-0.782,0.2082]) respectively while for the second order with interaction, the stationary points are 0.1882, 0.2896 corresponding to eigenvalue and eigenvectors (320.8457[-0.2082,-0.9781]) and (-864.1441[-0.9782,0.2082]) respectively. The plots shows the possibility of negative impact on the yield of cocoa.

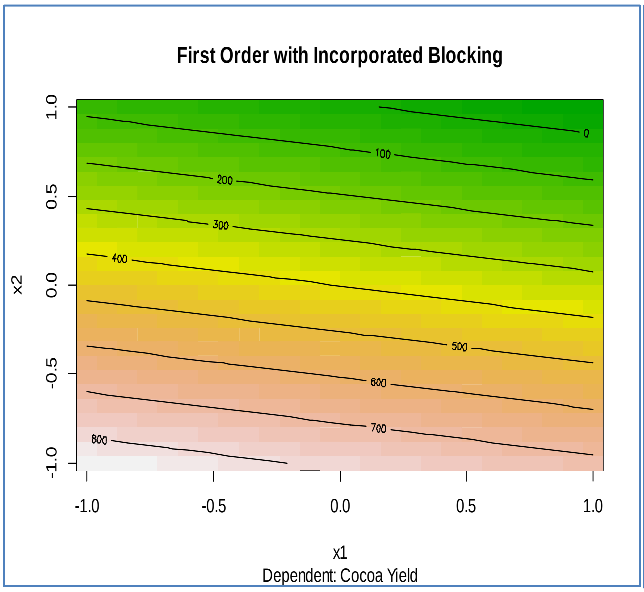
Table 2: Parameter Estimates for Models with Incorporated Blocking

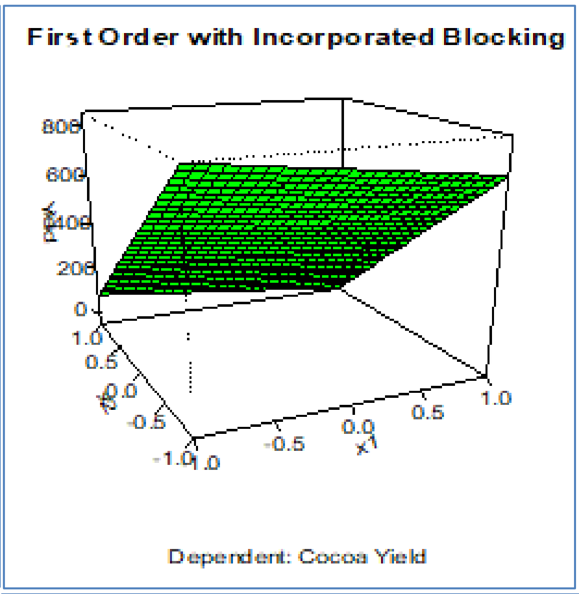
Model	Coefficient	Estimate	Std. Error	t-Value	P-value	Decision	AP-v	LoFP-v
FO-B	Intercept	795.690	108.357	7.3432	4.699e-12	***	0.00083	0.0545
	BlocksA2	-516.345	120.037	-4.3016	2.613e-05	***		
	BlocksA3	-409.595	123.132	-3.3265	0.0010406	**		
	BlocksA4	-664.593	153.246	-4.3368	2.258e-05	***		
	X_1	-68.462	175.527	-0.3900	0.6969089			
	X_2	-387.660	102.994	-3.7639	0.0002179	***		
SO-WOI-B	Intercept	802.951	109.685	7.3205	5.517e-12	***	$FO(0.0007)$ $PQ(0.028)$ $B(2.6e^{-07})$	0.0573
	BlocksA2	-507.443	118.876	-4.2687	3.005e-05	***		
	BlocksA3	-406.124	121.771	-3.3352	0.0010119	**		
	BlocksA4	-564.970	158.893	-3.5557	0.0004681	***		
	X_1	75.714	181.977	0.4161	0.6777990			
	X_2	-278.985	157.092	-1.7759	0.0772274	.		
	X_1^2	-756.430	308.379	-2.4529	0.0150054	*		
	X_2^2	212.108	172.886	1.2269	0.2212794			
TWI-B	Intercept	847.63	109.89	7.7132	5.166e-13	***	$FO(0.0007)$ $TWI(0.028)$ $B(2.8e^{-07})$	0.0565
	BlocksA2	-516.67	118.93	-4.3445	2.191e-05	***		
	BlocksA3	-438.84	122.71	-3.5763	0.0004343	***		
	BlocksA4	-680.85	152.00	-4.4791	1.241e-05	***		
	X_1	471.13	299.68	1.5721	0.1174637			
	X_2	-351.90	103.31	-3.4061	0.0007923	***		
	$X_1:X_2$	747.55	338.13	2.2109	0.0281460	*		
SO-B	Intercept	833.26	110.47	7.5426	1.491e-12	***	$FO(0.00064)$	0.0584

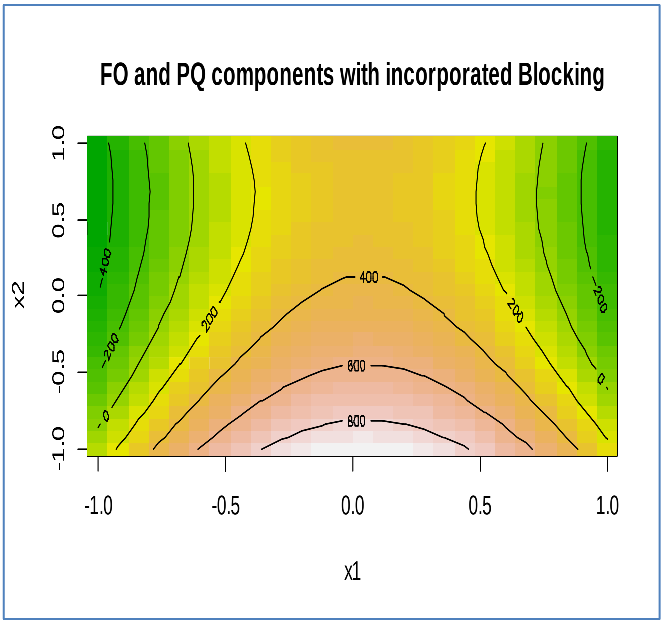
	BlocksA2	-504.71	118.28	-4.2670	3.032e-05	***	$TWI(0.0269)$ $PQ(0.0670)$ $B(2.21E^{-07})$	
	BlocksA3	-427.98	121.78	-3.5143	0.0005432	***		
	BlocksA4	-609.72	160.11	-3.8081	0.0001852	***		
	X_1	502.07	302.21	1.6613	0.0981827	.		
	X_2	-206.38	161.63	-1.2769	0.2030973			
	$X_1:X_2$	630.76	357.98	1.7620	0.0795679	.		
	X_1^2	-577.77	323.13	-1.7880	0.0752557	.		
	X_2^2	260.60	174.19	1.4960	0.1361934			

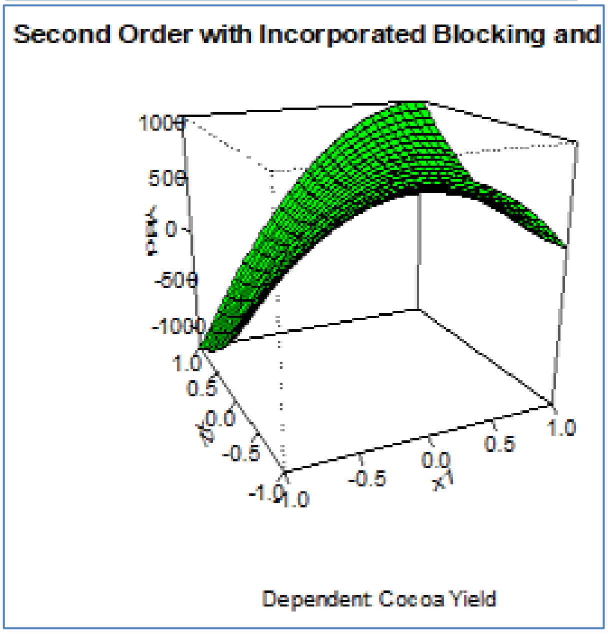
AP-v: ANOVA P-value, LoFP-v: Lack of Fit P-value, "***", "**", "*", "." Indicate significant at 0.1%, 1%, 5%, and 10% respectively. FO-B: First order model with blocking; SO-WOI-B: Second order model with blocking without Interaction; Two model with interaction with blocking; Second order model with blocking (Full model).

Table 2 accounts for the impact of temperature and rainfall having controlled for the impact plots being the blocking factor in the experimental unit. The first order model with incorporated blocking shows that plots play an essential role in determining the yield of cocoa in the study area. This result is ascertained through out all the different models observed. It duly noted that while plot is being controlled for, temperature impacts cocoa yield insignificantly across all model but rainfall plays a significant role in determine the yield of cocoa. Considering the estimates of the model parameter, it worthwhile to note that the yield of cocoa decreases as rainfall increases. Hence, rainfall poorly impact cocoa yield in the area under study.









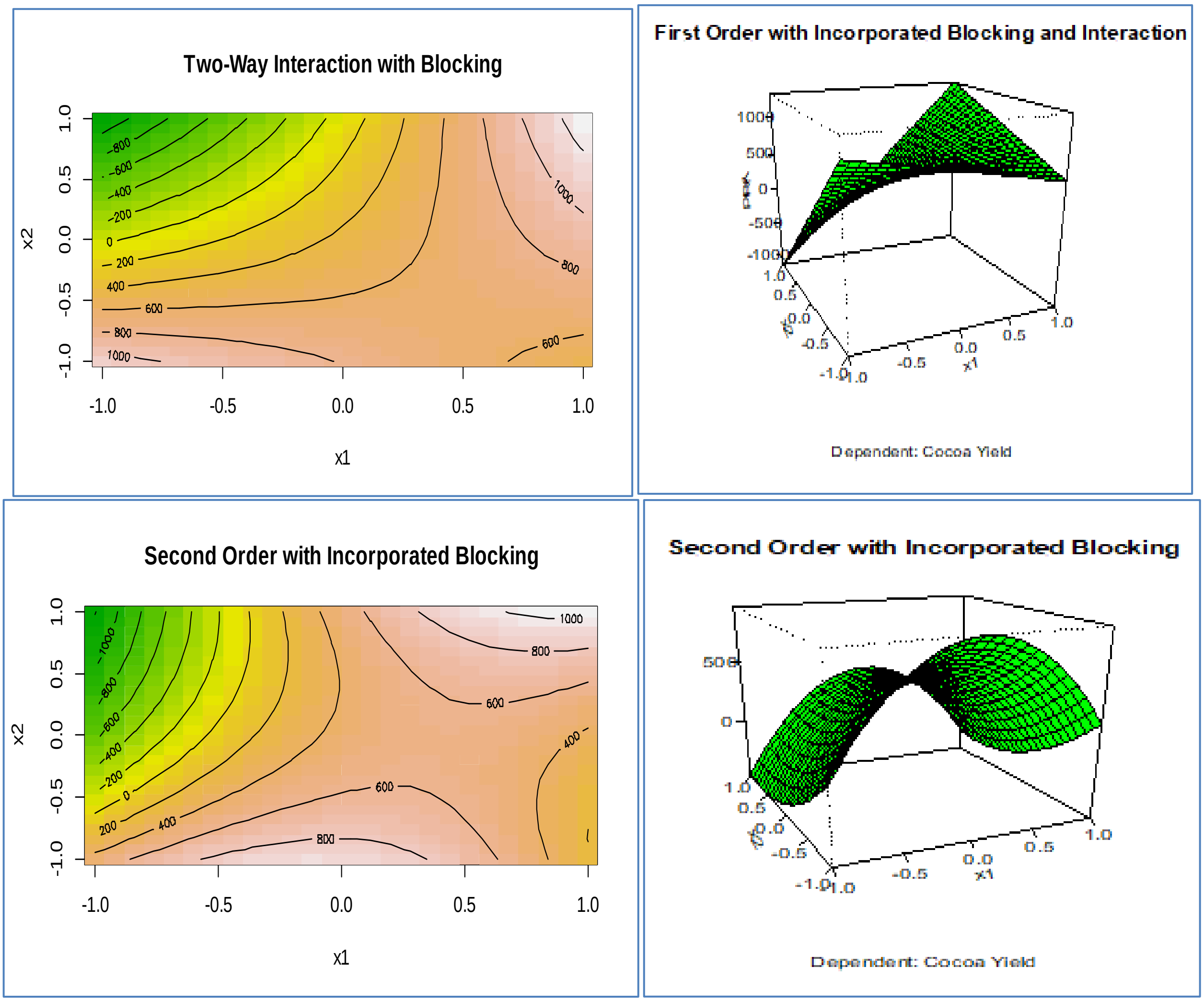


Fig. 6: Surface and contour plot for models with incorporated blocking

Considering the estimated contour plot, it is observed that the insignificance of the lack of fit of first order RSM results to a contour with stationary points of response surface as (0.0500, 0.6576) for (x_1, x_2) which resulted to set of Eigen-values (x_1=212.1085 and x_2=-756.4302) corresponding to set of Eigen-vectors ([0,1], [0,-1]). From the estimated contour plot for the second order RSM. it is discovered to have stationary points as (0.3918, -0.0782) corresponding to (x_1 and x_2) respectively. The estimated Eigen-values for (x1, x2) are (365.9881, -683.1568) associated with set of Eigen-vectors ([-0.3169, -0.9484], [-0.948, 0.3169])

Table 3: Models comparison for response surface models without blocking

S/N	Model	R^2	$Adj\ R^2$	P-value
1	Linear Regression	0.0908	0.08212	4.576×10^{-05}
2	First Order	0.0908	0.08212	4.576×10^{-05}
3	Pure Quadratic	0.1185	0.1101	1.78×10^{-06}
4	Two Way Interaction	0.04215	0.03761	0.002604
5	Second Order	0.1415	0.1208	6.46×10^{-06}

Table 3 shows the models performance or adequacy given that no blocking was incorporated into the experiment. Findings show that second order model outperformed all estimated models since its coefficient of determination is the highest even in terms of the robust measure adjusted R^2.

Table 4: Models comparison for response surface models with Incorporated Blocking

S/N	Model	R^2	$Adj\ R^2$	P-value
1	First Order	0.1947	0.1753	1.354×10^{-08}
2	Second order Model without Interaction	0.2223	0.1957	5.448×10^{-09}
3	Two Way Interaction	0.2134	0.1905	4.894×10^{-09}
4	Second Order	0.234	0.2039	4.132×10^{-09}

Findings from table 4 indicates that second order model outperforms every other model with incorporated blocking factor since it has the highest value of adjustedR^2. In relation to Table 4.22 all models with blocking factor perform better than those without blocking. Hence, with blocking there is every tendency for high precision to be ascertained.

IV. CONCLUSION

Cocoa has been identified as one of the major agricultural producein the world with diverse benefits as decreased inflammation, balanced heart and brain health, weight control and healthy teeth and skin. Increased temperature, decreasesin rainfall negatively impact cocoa crop health and its yields over time. Hence cocoa will be negatively affected as climate changes progress. In the presence of blocking factors such as soil type and PH scale, there is tendency for the yield of cocoa to improve significantly (Jijakhi, 2007). Following from the findings of this research, the study recommends that Cocoa Research Institute of Nigeria and other cultivators of cocoa should adopt newly developed cocoa varieties, methods of planting and cultivation which are well adapted to beat the vagaries of weather changes being experienced for improved income and livelihood.

REFERENCES

[1]. Fisal Ahmad, Wan Mahd Ashri Wan Daud, Mohd Azmier Ahmad, and RosinahRadzi(2013),"The effect of acid leaching on porosity and surface functional groups of cocoa (Theobroma cacao)-shell based activated carbon, using response methodology,"Journal of Chemical Engineering Research and Design 91 (6) 1028-1038,2013.

[2]. Jijakli M. H., Begoude B. A. D, Lahlahi R. Friel D. and Tondje P. R..(2007),"Response Surface Methodology study of the combined effects of temperature , pH, and a_w, on the growth rate of Trichoderma asperellum," Journal of Applied Microbiology ISSN 1364-5072.

[3]. Kowalski,S.M, Cornell,J.A., and Montgnomery,D.C.(2005),"A Modified Path of Steepest Ascent for Response Surface Methodology ," Journal of Quality Technology,37,75-83.

[4]. Lucia C. V., Reinaldo F., T. and Carmen L. O. P(2011),"Estimated optimization ofnitric acid-mediated extraction of pectin from cacao pod husks(Theobroma cacao L)."using response urface methodology" ,ScienceDirect Vol 84, Issue 4,pp 1230-1236.

[5]. Miguel R. H. Silvia C., Diego T., Angela S. et al,(2021),"Extraction of Phenolic compounds from cocoa shell ;Modeling using response surface methodology and artificial neural network." ScienceDirect Vol 270, 1 September 2021, 118779.

[6]. Montgnomery, D.C.,(2013), Design and Analysis of Experiments,7th edition ,John Wiley and Sons, New York.

[7]. Pravin D.S. (2020),"Development of free and encapsulated Arjuna herb extract added vanilla chocolate dairy drink by using response surface methodology . " ScienceDirect Vol 2, December 2020,100020.