



## OPTIMIZATION OF AGRICULTURAL YIELD ON THE IMPACT OF OZONE AND SULPHUR DIOXIDE USING RESPONSE SURFACE METHODOLOGY

\*<sup>1</sup>Bodunwa, O. Kikelomo, <sup>1</sup>Aladeniyi, O. Bodunde, <sup>2</sup>Adewole, I. Ayoade

<sup>1</sup>Department of Statistics, Federal University of Technology, Akure,

<sup>2</sup>Department of Mathematical Sciences, Tai Solarin University of Education Ijebu-Ode

\*Corresponding authors' email: [okbodunwa@futa.edu.ng](mailto:okbodunwa@futa.edu.ng)

### ABSTRACT

In this study, Response Surface Methodology (RSM) was applied to the analysis of ozone and Sulphur dioxide on soybean. The data used was secondary data which came from research on the effects of air pollutants on crop yields conducted in North Carolina State University and USDA in 1981. From the analyses, the adequacy of the model used was validated with the F-value of 18.71 and 18.55 respectively suggests the models are significant. The Coefficient of Determination  $R^2$ , the Adjusted  $R^2$ , written as  $R_{adj}^2$  and Predicted  $R^2$  were also used to find the adequacy of the model used. Effects of ozone and Sulphur dioxide were looked into in the yields of soybean in two different locations. It was found that the 0.395 inlet of sulphur and 0.018 inlet of ozone gives optimum yield of 724.427 grams of soybeans in yield 1 and same quantity gives optimum yield of 248.282 grams of soybeans in yield 2 which made ozone and sulphurdioxide to have significant effect on the yield of soyabeans.

**Keywords:** Model Adequacy, optimization, response surface methodology, soybeans, yield

### INTRODUCTION

The soybean, soy bean or soya bean (*Glycine max*) (Multiscript, Retrieved Feb 16, 2012) is a species of legume native to East Asia, widely grown for its edible bean, which has numerous uses. Soybean meal is a significant and cheap source of protein for animal feeds and many packaged meals. It is one of the most important crops worldwide. This seed is a good source for both protein and vegetable oil. The crop is grown on an estimated 6% of the world's arable land and since 1970s, the area in soybean production has the highest percentage increase compared to any other major crop. Therefore, recent increases in the production of soybeans coincide with increases in demand for meal and oil (Hartman et.al., 2011).

Soybeans are a major crop worldwide and one of the leading cash crops in the world. Brazil is the most soybean production country with about 38% percent of world production with 138000 production (1000MT) followed by United State with 31% 114749 Mt (Marcela & Mohammad, 2016).

Ozone ( $O_3$ ) is a photochemically produced oxidant gas while Sulphur dioxide ( $SO_2$ ) is a product of burning of Sulphur containing fossil fuels and smelting of ores. Both gases occur widely in many regions and both are toxic to plants. Reviews of literature indicate that much information is available on plant responses to each (Treshow, 1984; Emberson, et al., 2018). As well, the combined effects of these two gases have been researched more widely than those of any other combination of pollutant gases (Ormrod, 1982).

The effects of daily ozone and the periodic Sulphur exposures on the yield of soybean was the primary aim of the study. The exposure regime was designed to simulate Sulphur exposures from a point source superimposed on a regional ozone stress (Kress et.al., 1986). Amundson and Kress (1990) determined if the exposure of ozone and Sulphur dioxide in combination produce greater- than-additive effects on yields of economically important crops. Monitoring of several physiological responses of the crops provided a means of assessing short term effects of these variables. Some crops like wheat rice, maize and soybeans in China were exposed to

ozone. The aggregated production and associated economic losses and major provinces were evaluated by combining annual production yields and crop market prices, this was done by (Shindell et.al., 2022).

Response Surface Methodology (RSM) has gained more interest in research methodology almost in every area of study in science and industry. It is a widely used mathematical and statistical method for modelling and analyzing a process in which the response of interest is affected by various variables (Braumah et.al., 2016). In recent literatures, (Yolmeh and Jafari, 2017) used RSM to get the optimization of different food processes such as extraction, drying, blanching, enzymatic hydrolysis and clarification, production of microbial metabolites and formulation. Also, (Sushanta et.al., 2018) developed the efficient technique for the production of clean coal by optimizing the operating parameters with the help of RSM. (Kumar et.al 2019) used RSM to optimize the temperature and time for maximum bio-oil yield

From all the literatures, the relationship that exist among the yield, ozone and Sulphur dioxide has not been addressed using RSM. Application of Response surface methodology was used in this study for optimizing operation condition and mitigating economic risk. Response surface methodology is a collection of statistical and mathematical methods that are useful for designing experiments, building models, evaluating the effect of factors, and searching for optimum conditions for desirable responses (Box and Wilson, 1951).

### MATERIAL AND METHODS

The data used in this study was a secondary data which came from research on the effects of air pollutants on crop yields conducted in North Carolina State University and USDA in 1981. The yield data and the observed seasonal averages of Ozone ( $X_1$ ) and Sulphur Dioxide ( $X_2$ ) for each experimental unit are given in Table1. The north and south halves of the experimental plots are recorded as the Yield ( $Y_1$ ) and Yield ( $Y_2$ ) respectively.

**Table 1: Data of yields of Soybean from the exposure of Ozone and Sulphur dioxide**

Block 1						Block 2					
Coded Input		Original Input		Response		Coded Input		Original Input		Response	
$X_1$	$X_2$	$X_1$	$X_2$	$Y_1$	$Y_2$	$X_1$	$X_2$	$X_1$	$X_2$	$Y_1$	$Y_2$
-0.979	-1	0.025	0	516.5	519.5	-0.979	-1	0.025	0	603	635
-0.357	-0.888	0.234	0.022	552	596	-0.988	-0.924	0.022	0.015	796	454.5
-0.970	-0.620	0.028	0.075	569	500.5	-1	-0.493	0.018	0.1	597.5	697
-0.967	0.969	0.029	0.389	419	358.5	-0.979	0.9240	0.025	0.38	458	365.5
-0.877	-1	0.059	0	503.5	449.5	-0.901	-1	0.051	0	652	496
-0.880	-0.918	0.058	0.016	411	484	-0.898	-0.858	0.052	0.028	590.5	292.5
-0.880	-0.645	0.058	0.07	502.5	477	-0.889	-0.534	0.055	0.092	440	427.5
-0.880	0.772	0.058	0.35	353	338.5	-0.901	0.7265	0.051	0.341	487	284
-0.851	-1	0.068	0	449.5	480.5	-0.854	-1	0.067	0	533.5	321.5
-0.836	-0.918	0.073	0.016	472.5	478	-0.857	-0.883	0.066	0.023	486	317
-0.839	-0.569	0.072	0.085	382.5	411.5	1	-0.473	0.69	0.104	420.5	456
-0.851	1	0.068	0.395	291	266.5	-0.851	0.9088	0.068	0.377	271	280.5
-0.803	-1	0.084	0	399	414.5	-0.788	-1	0.089	0	390.5	324.5
-0.851	-0.827	0.068	0.034	321.5	336.5	-0.794	-0.797	0.087	0.04	373	320.5
-0.809	-0.660	0.082	0.067	373	384.5	-0.800	-0.539	0.085	0.091	321	246
-0.785	0.772	0.09	0.35	269	303	-0.806	0.9189	0.083	0.379	246.5	274
-0.741	-1	0.105	0	438	345	-0.726	-1	0.11	0	307	281.5
-0.723	-0.908	0.111	0.018	346.5	347.5	-0.735	-0.762	0.107	0.047	387.5	329.5
-0.732	-0.574	0.108	0.084	297	316.5	-0.755	-0.503	0.1	0.098	270	246
-0.738	0.868	0.106	0.369	242.5	244	-0.755	0.8329	0.1	0.362	197.5	196
-0.696	-1	0.12	0	342.5	331.5	-0.693	-1	0.121	0	275	278.5
-0.663	-0.893	0.131	0.021	269	298.5	-0.681	-0.858	0.125	0.028	266	243.5
-0.678	-0.716	0.126	0.056	297.5	308.5	-0.675	-0.498	0.127	0.099	303	215.5
-0.687	0.7468	0.123	0.345	211	227	-0.690	0.7974	0.122	0.355	283.5	208

The independent variables Ozone and  $SO_2$  were coded between -1 and +1 as presented in the Table above for the response surface methodology used in this study.

The RSM technique can improve product yields and provide closer confirmation of the output response toward the nominal and target requirements. In recent years, RSM played an important role in oil fields, especially applications into enhanced oil recovery. In most RSM problems, the objective function of the response and independent variables is unknown. Thus, the first step is to find a suitable approximation for the true functional relationship between the response ( $Y$ ) and the set of independent variables ( $X_i$ ). If the response is well modelled by a linear function of the independent variables, then the approximation function is the first-order model. A model that incorporates curvature is usually required to approximate the response in the region close to optimum, and in most cases, a second order model is adequate (Montgomery, 2001)

In this study, a higher model, cubic is more adequate for these data compared to the second model.

Estimation of parameters of a multi-response model with random block effects

Let  $y_1, y_2, \dots, y_r$  be  $r$  response variables of interest that can be measured for each setting of a group of  $k$  input (control) variables denoted by  $x_1, x_2, \dots, x_k$ . Suppose that the  $n$

experimental runs for each response are divided into  $b$  blocks of size  $n_1, n_2, \dots, n_b$  such that

$$n = \sum_{j=1}^b n_j \quad (1)$$

in order to control an extraneous source of variation, hereafter referred to as a block effect. The response vector  $y_i$  can then be represented by the model

$$y_i = \beta_{0i} \mathbf{1}_n + X_i \beta_i + \mathbf{Z} \delta_i + \epsilon_i \quad (2)$$

Where  $\mathbf{1}_n$  is a vector of ones of order  $n \times 1$ ,  $\beta_{0i}$  and the elements of  $\beta_i$  are unknown constant parameters,  $X_i$  is an  $n \times p_i$  matrix.  $\delta_{ij}$  denotes the effect of the  $j$ th block for the  $i$ th response data ( $j = 1, 2, \dots, b$ ) and  $\mathbf{Z}$  is a block-diagonal matrix of the form  $\mathbf{Z} = \text{diag}(\mathbf{1}_{n_1}, \mathbf{1}_{n_2}, \dots, \mathbf{1}_{n_b})$  and the vectors  $\epsilon_i$  are the random errors assumed to have zero means with variance-covariance matrices.

For the purpose of this study, the model used is

$$y_i = \beta_{0i} \mathbf{1}_n + \mathbf{Z} \delta_i + \sum_{i=1}^k X_i \beta_i + \sum_{i=1}^k X_i^2 \beta_{ii} + \sum_{i=1}^k X_i^3 \beta_{iii} + \sum_{i=1}^k X_i X_j \beta_{ij} + \epsilon_i \quad (3)$$

The coefficient parameters of the models were estimated using a multiple linear regression analysis employing the Design-Expert Software (version 8.0.1.0, Stat-Ease, Inc., Minneapolis, USA). The Design-Expert was also used to demonstrate the 3D surface and 2D contour plots of the response models

## RESULTS AND DISCUSSION

## Model adequacy for the yields

Table 2: Regression coefficients of the predicted cubic polynomial model for the responses

		Intercept	A	B	AB	A <sup>2</sup>	B <sup>2</sup>	A <sup>2</sup> B	A B <sup>2</sup>	A <sup>3</sup>	B <sup>3</sup>
Yield 1	Estimate	700.19	792.81	-375.43	-1025.55	-672.52	-40.27	-667.15	-13.09	-1387.65	-108.27
	Standard Error	580.59	1178.3	786.68	1915.81	437.36	324.73	1144.84	370.51	358.39	87.91
	P value	<0.0001	0.5052	0.6360	0.5956	0.1326	0.9020	0.5636	0.9720	0.0004	0.2259
Yield 2	Estimate	311.82	306.28	-532.32	-1381.32	-564.87	489.15	-937.37	622.77	-1172.36	-27.44
	Standard Error	508.16	1031.3	688.54	1868.83	382.81	284.22	1002.03	324.29	313.69	76.94
	P value	<0.0001	0.7681	0.4442	0.4153	0.1485	0.0936	0.3556	0.0625	0.0006	0.7234

From (3), the polynomial model describing the relationship between the response Yield 1 and the variables can be written as

$$\text{Yield 1} = 700.19 + 792.81X_1 - 375.43X_2 - 1025.55X_1X_2 - 672.52X_1^2 - 40.27X_2^2 - 667.15X_1^2X_2 - 13.09X_1X_2^2 - 1387.65X_1^3 - 108.27X_2^3 \quad (4)$$

$$\text{Yield 2} = 311.82 + 306.28X_1 - 532.52X_2 - 1381.32X_1X_2 - 564.87X_1^2 - 489.15X_2^2 - 937.37X_1^2X_2 + 622.77X_1X_2^2 - 1172.36X_1^3 - 27.44X_2^3 \quad (5)$$

The Models in (4) and (5) with F-value of 18.71 and 18.55 respectively implies the models are significant. There is only

a 0.01% chance that an F-value this large could occur due to noise. P-values less than 0.0500 indicate model terms are significant. In this case A<sup>3</sup> is a significant model term. To prove the accuracy of the model (4), statistical analysis techniques were checked by the experimental error, the suitability of the model, the Goodness of Fit and the statistical significance of the terms in the model. The quality of the model is statistically measured by Akaike Information Criterion (AIC) and Bayesian Information Criteria (BIC) for Goodness of Fit. These values were compared to the other model and the model used here is having minimum values (561.56 and 548.31 compared to the quadratic model of 585.33 and 575.03 respectively)

Table 3: ANOVA for cubic model for Yield 1 and 2  
Response 1: Yield 1

Source	Sum of Squares	df	Mean Square	F-value	p-value
Block	11011.02	1	11011.02		
<b>Model</b>	6.346E+05	9	70511.75	18.71	< 0.0001 significant
A-OZONE	1706.05	1	1706.05	0.4528	0.5052
B-SO2	858.22	1	858.22	0.2278	0.6360
AB	1079.79	1	1079.79	0.2866	0.5956
A <sup>2</sup>	8909.39	1	8909.39	2.36	0.1326
B <sup>2</sup>	57.94	1	57.94	0.0154	0.9020
A <sup>2</sup> B	1279.62	1	1279.62	0.3396	0.5636
AB <sup>2</sup>	4.70	1	4.70	0.0012	0.9720
A <sup>3</sup>	56490.48	1	56490.48	14.99	0.0004
B <sup>3</sup>	5715.46	1	5715.46	1.52	0.2259
<b>Residual</b>	1.394E+05	37	3768.16		
<b>Cor Total</b>	7.850E+05	47			
<b>Std.Dev =</b>	61.39, Mean =39.67		R <sup>2</sup> = 0.8199	Adj R <sup>2</sup> = 0.7761	Ad Pre =15.80

## Response 2: Yield 2

Source	Sum of Squares	df	Mean Square	F-value	p-value
Block	21952.13	1	21952.13		
<b>Model</b>	4.820E+05	9	53555.08	18.55	< 0.0001 significant
A-OZONE	254.61	1	254.61	0.0882	0.7681
B-SO2	1726.65	1	1726.65	0.5981	0.4442
AB	1958.91	1	1958.91	0.6786	0.4153
A <sup>2</sup>	6285.61	1	6285.61	2.18	0.1485
B <sup>2</sup>	8550.25	1	8550.25	2.96	0.0936
A <sup>2</sup> B	2526.14	1	2526.14	0.8751	0.3556

AB <sup>2</sup>	10646.16	1	10646.16	3.69	0.0625
A <sup>3</sup>	40321.68	1	40321.68	13.97	0.0006
B <sup>3</sup>	367.05	1	367.05	0.1272	0.7234
<b>Residual</b>	1.068E+05	37	2886.70		
<b>Cor Total</b>	6.108E+05	47			
<b>Std.Dev =</b>	53.73, Mean =362.7	$R^2 = 0.8186$	$Adj R^2 = 0.7745$	Ad Pre =15.83	

The **Model F-value** of 18.71 and 18.55 for the two responses implies the models are significant. There is only a 0.01% chance that an F-value this large could occur due to noise. The values of Adequate Precision which are 15.80 and 15.83 respectively indicate an adequate signal, which implies that the model can be used to navigate the design space.

The Coefficient of Determination  $R^2$ , the Adjusted  $R^2$ , written as  $R_{adj}^2$  and Predicted  $R^2$  were also used to find the adequacy of the model used. Any model with values of any of  $R^2$  that is closed to 1 indicates an excellent quality in fitting the observed data. In this study, their values were

0.8199, 0.7761 and -1034.3492 respectively. Also, the Adequate Precision value is 15.8074

A negative Predicted  $R^2$  implies that the overall mean may be a better predictor of your response than the current model. Adequate Precision measures the signal to noise ratio. A ratio greater than 4 is desirable therefore this model can be used to navigate the design space. Once a model is constructed as a result of the above consequences, it can be used to predict reservoir performance and to optimize controllable variable. Another way of checking the adequacy of the model is the use of normal probability plot of the residual shown in the figure below.

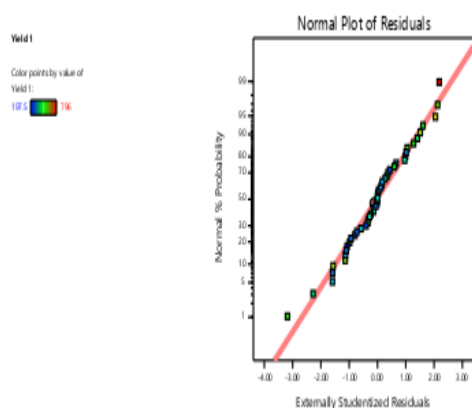


Figure 1 A

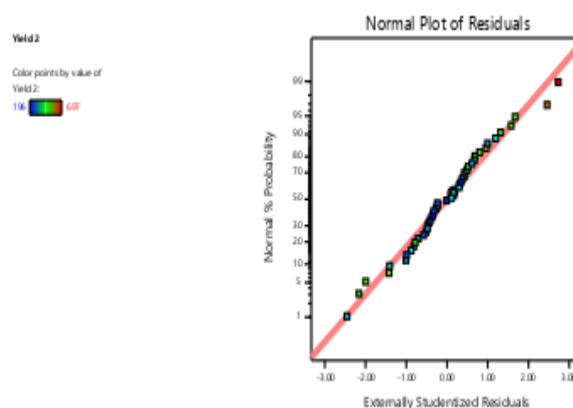


Figure 1 B

Figure 1 A and B above show that there is no severe indication of nonnormality, nor is there any evidence pointing to possible outliers from the plot of Yield 1 and 2 respectively.

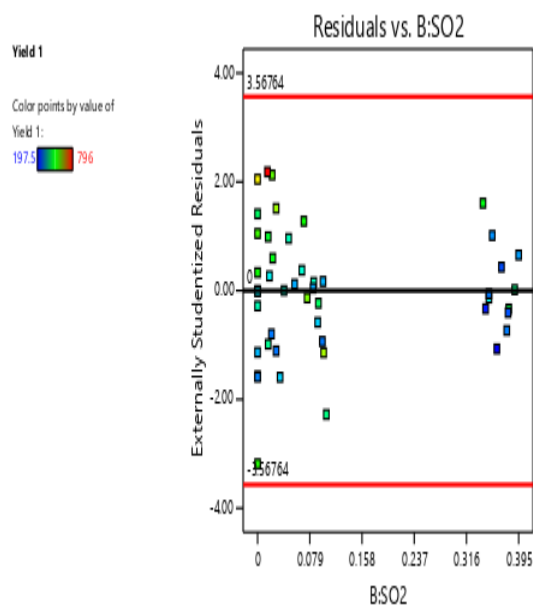


Figure 2A

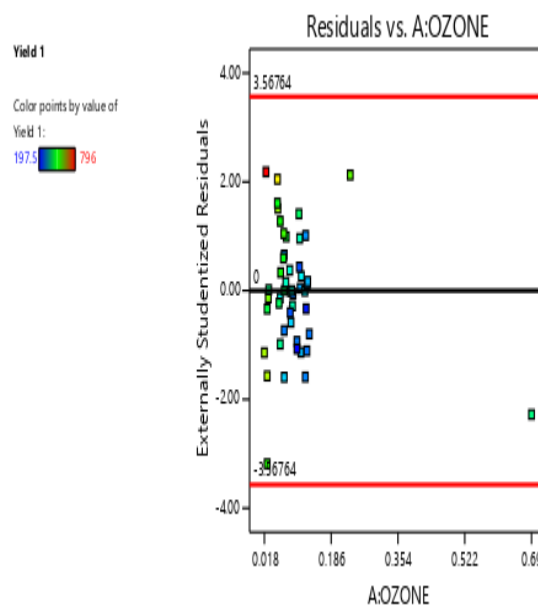


Figure 2B

Figure 2 A and B show plots of the residuals by ozone and by Sulphur dioxide respectively. These plots are potentially very informative. If there is more scatter in the residuals for a particular treatment, that could indicate that this treatment produces more erratic response readings than the others. More scatter in the residuals for a particular block could indicate that the block is not homogeneous. However, in this study,

Figure 2 gives no indication of inequality of variance by Sulphur dioxide but there is an indication that there is less variability in the yield for ozone. However, since all of the other residual plots are satisfactory, we will ignore the remaining plots

The predicted versus actual is also presented here

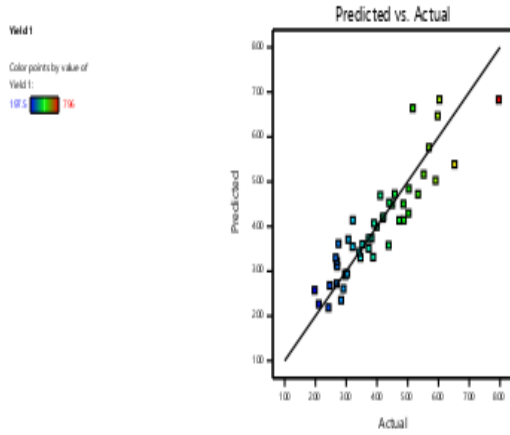


Figure 3A

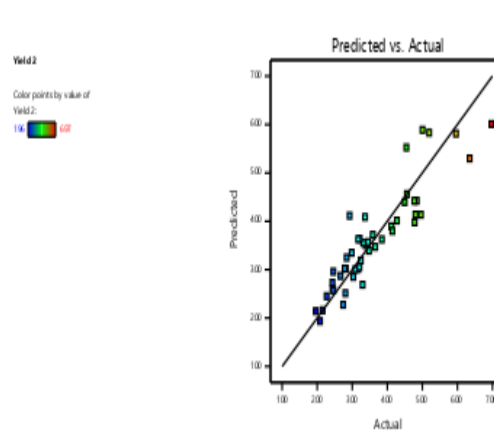


Figure 3B

Figure 3 show the plots of predicted and actual values of the yields. Since the observed and predicted are reasonably close which indicates that they are very similar, this confirm the successful run of the experiment.

**Optimization of the Ozone and Sulphur dioxide**

The graphical representation of yields Contour and 3D Response Surface Plots were shown in Figures 4 and 5 respectively

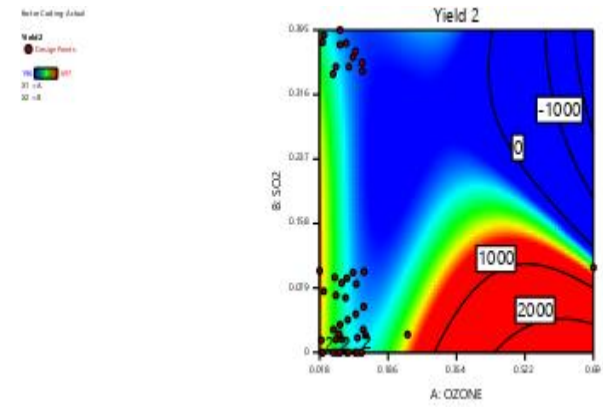
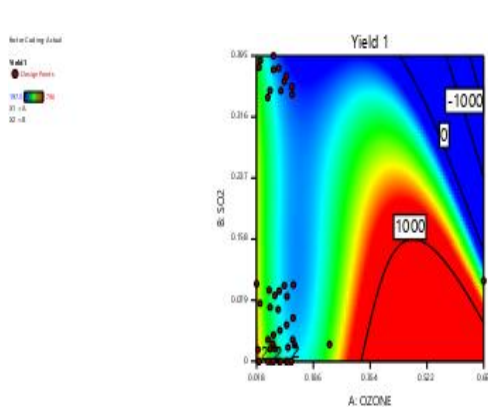


Figure 4: 2D Contour plots showing the effects of Ozone and Sulphur operations for both the Yields

In Fig. 4, presented the design points for the yields from ozone and sulphur dioxide. It was noted that at points 0.079 and 0.395, the effect of sulphurdioxide was much and at point 0.018ppm the effect of ozone was seen.

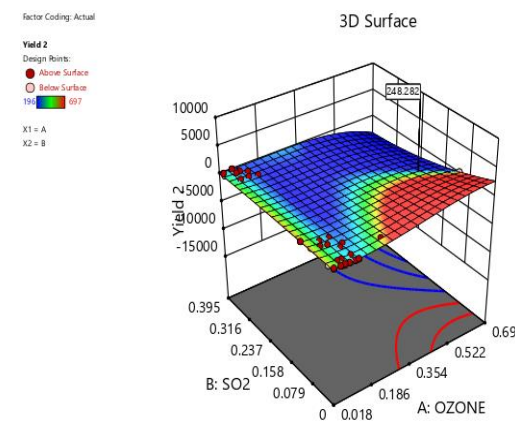
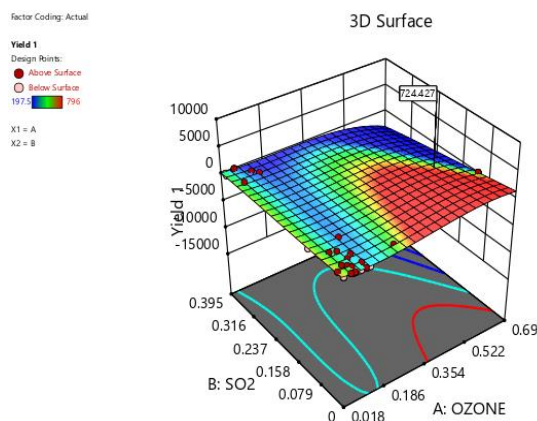


Figure 5: 3D Response Surface plots showing the effects of Ozone and Sulphur operations for both the yields

The 3D plots are the graphical representation of the regression equations in order to determine the optimum value of the variables within the design space (Khuri & Cornell, 1996). The value of predicted maximum on the surface is confined in the smallest ellipse in the contour diagram. It was also shown in Fig 5 that the optimal values of the process variables were found which is the combination of ozone and sulphur dioxide. The 0.395 inlet of sulphur and 0.018 inlet of ozone will give optimum yield of 724.427 grams of soybeans in yield 1 and same quantity will give optimum yield of 248.282 grams of soybeans in yield 2.

## CONCLUSIONS

The application of a mathematical model and the optimization on the basis of statistical design of experiments is proven to be a useful tool to predict and analyse the interaction effects between operating factors. In this study, the effect of ozone and Sulphur di oxide were looked into in the yields of soy bean in two different locations. It was found that the 0.395 inlet of sulphur and 0.018 inlet of ozone gives optimum yield of 724.427 grams of soybeans in yield 1 and same quantity gives optimum yield of 248.282 grams of soybeans in yield 2. The regression Models in (4) and (5) with F-value of 18.71 and 18.55 respectively implies the models are significant which made ozone and sulphurdioxide to have significant effect on the yield of soybeans..

## REFERENCES

- Amundson, R. G., & Kress, L. (1990). *Sulfur dioxide and ozone effects on crops*. United state: U. S. Department of Energy, Office of Scientific and Technical Information.
- Box, G. E., & Wilson, K. B. (1951). On the experimental attainment of optimum conditions. *J. Roy. Stat. Soc., Ser. A* 13, 1-45.
- Braimah, M. N., Anozie, A. N., & Odejebi, O. J. (2016). Utilization of response surface methodology (RSM) in the optimization of crude oil refinery. *Journal of Multidisciplinary Engineering Science and Technology (JMEST)*, 4361- 4369.
- Emberson, L. D., Pleijel, H., Ainsworth, E. A., Berg, M. V., Pandey, D., Dentener, F., & Buker, P. (2018). Ozone effects on crops and consideration in crop models. *European Journal of Agronomy*, 19 - 34.
- Hartman, G. L., West, E. D., & Herman, T. K. (2011). Crops that feed the World 2. Soybean-worldwide production, use and constraints caused by pathogens and pests. *Food Sec.* 3, 5-17.
- Khuri, A. I., & Cornell, J. A. (1996). *Response Surfaces: Designs and Analyses. 2nd Ed. Marcel Dekker Inc. New York.*
- Kress, L. W., Miller, J. E., Smith, H. J., & Rawlings, J. O. (1986). Impact of ozone and Sulphur dioxide on soybean yield. *Environmental Pollution Series A, Ecological and Biological*, 105 - 123.
- Kumar, M., Mishra, P. K., & Upadhyay, S. N. (2019). Pyrolysis of *Saccharum munia*: Optimization of process parameters using Response Surface Methodology and evaluation of kinetic parameters. *Bioresource Technology Reports*.
- Marcela, C. P., & Mohammad, M. (2016). The importance of Soybean production worldwide. *Academic Press*, 1-26.
- Montgomery, D. C. (2001). *Design and Analysis of Experiments, 5th ed.* New York, USA pp. 427-510: John Wiley & Sons.
- Multiscript, M. (Retrieved Feb 16, 2012). *Plant Name Database*.
- Ormrod, D. P. (1982). Air pollutant interactions in mixtures, Effects of gaseous air pollution in agriculture and horticulture. *Butterworths, London*, 532.
- Shindell, L. D., Ding, D., Zhang, L. X., & Zhang, Y. (2022). Surface ozone impacts on major crop production in China from 2010 to 2017. *Atmos. Chem. Phys*, 2625 - 2638.
- Sushanta, K. B., Himanshu, M., Sudipto, C., & Meikap, B. C. (2018). Application of Response Surface Methodology for optimization of leaching parameters for ash reduction from low-grade coal. *International Journal of Mining Science and Technology*, 621 -629.
- Treshow, M. (1984). Air pollution and plant life. *Wiley*, 486.
- Yolmeh, M., & Jafari, S. M. (2017). Applications of Response Surface Methodology in Food Industry Processes. *Food Bioprocess Technol*, 413 - 433.



©2023 This is an Open Access article distributed under the terms of the Creative Commons Attribution 4.0 International license viewed via <https://creativecommons.org/licenses/by/4.0/> which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is cited appropriately.