

## TABLE CONTENTS

<b>CHAPTER 1 .....</b>	<b>2</b>
<b>GENERATION OF DESIGN FOR BIODIESEL PRODUCTION USING DESIGN EXPERT SOFTWARE.....</b>	<b>2</b>
<b>1.1 Experimental Design Matrix .....</b>	<b>2</b>
<b>1.2 ANOVA Analysis.....</b>	<b>3</b>
<b>1.3 Effects of Variables on Biodiesel Yield .....</b>	<b>4</b>
<b>1.4 Optimization .....</b>	<b>7</b>
<b>CHAPTER 2 .....</b>	<b>10</b>
<b>Case Studies of Different Response Surface Methodology Design .....</b>	<b>10</b>
<b>2.1 Taguchi Orthogonal Array (OA) Design – (Jee Pei Qi) .....</b>	<b>10</b>
<b>2.2 Central Composite Design to Optimize Coagulation- Kinetic Model, Thermodynamics, and Chemical Composition of Pomelo Essential Oil Extraction Using Steam Distillation (Choong Yi Jie).....</b>	<b>14</b>
<b>2.3 3 Level Factorial Design – Thermal cracking of petroleum residue oil using three level factorial design (Khadija Nadim) .....</b>	<b>18</b>
<b>2.4 Box-Behnken Design – Parametric optimization by Box-Behnken design for synthesis of magnetic chitosan-benzil/ZnO/Fe3O4 nanocomposite and textile dye removal (Reghiou et al., 2021) (Wong Pei Yu).....</b>	<b>25</b>
<b>2.5 Optimal Design (D-optimal) of Response Surface Methodology - Application on Optimizing Drug Release from Formulation of Acrylamide-based Hydrogels (Tham Ting Woon) .....</b>	<b>30</b>
<b>REFERENCES.....</b>	<b>37</b>

## CHAPTER 1

### GENERATION OF DESIGN FOR BIODIESEL PRODUCTION USING DESIGN EXPERT SOFTWARE

#### 1.1 Experimental Design Matrix

Company ABC decide to produce biodiesel by using enzymatic process with waste cooking oil. Firstly, in Design Expert software, 5 levels Response Surface Methodology (RSM) couple with Central Composite Design (CCD) is selected. The numeric factors were selected as 4 because there are 4 important variables for the biodiesel production with a certain experimental range which are enzyme loading (g), oil to methanol molar ratio, temperature ( $^{\circ}\text{C}$ ) and reaction time (hour) shown in Table Q1a from the question. The value inserted into the column of low and high instead of -alpha and +alpha column because the information was provided in range of values as shown in figure 1.1.

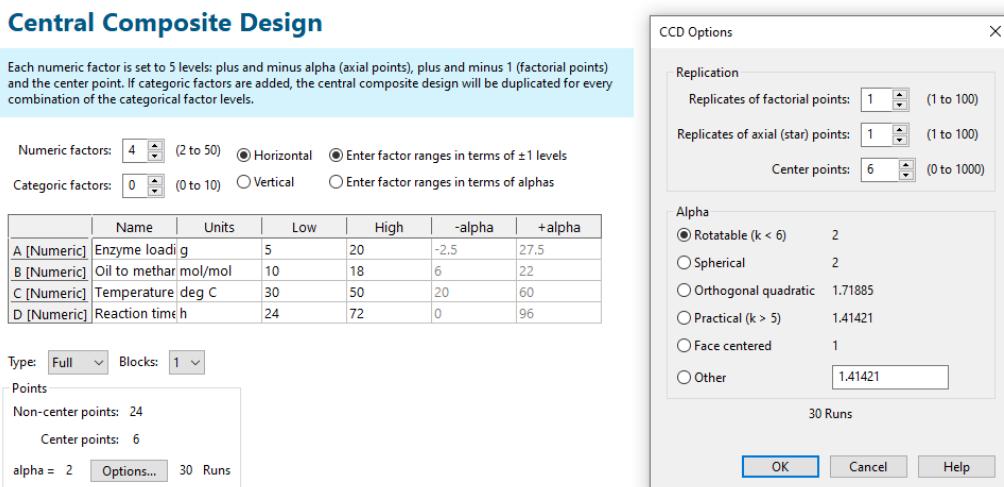


Figure 1.1: Data Input in RSM CCD

Then, the responses is selected as 1 and the response name is the biodiesel yield with unit wt.% and the all the experimental results are tabulated by inserted the value from Table Q1b. There are total 30 runs in the design and figure 1.2 shows the results of the design matrix of the CCD.

Std	Run	Factor 1 A:Enzyme loading g	Factor 2 B:Oil to methanol mol/mol	Factor 3 C:Temperature deg C	Factor 4 D:Reaction time h	Response 1 Biodiesel Yield wt.%
30	1	12.5	14	40	48	79.38
27	2	12.5	14	40	48	84.81
21	3	12.5	14	20	48	93.61
17	4	-2.5	14	40	48	68.13
14	5	20	10	50	72	68.98
19	6	12.5	6	40	48	80.95
8	7	20	18	50	24	70.01
18	8	27.5	14	40	48	77.54
11	9	5	18	30	72	62.93
10	10	20	10	30	72	93.61
12	11	20	18	30	72	93.61
26	12	12.5	14	40	48	66.66
23	13	12.5	14	40	0	74.9
24	14	12.5	14	40	96	82.67
4	15	20	18	30	24	69.98
7	16	5	18	50	24	78.29
3	17	5	18	30	24	53.32
1	18	5	10	30	24	72.95
22	19	12.5	14	60	48	93.61
6	20	20	10	50	24	67.21
29	21	12.5	14	40	48	63.32
15	22	5	18	50	72	58.21
28	23	12.5	14	40	48	50.44
9	24	5	10	30	72	60.24
25	25	12.5	14	40	48	93.61
5	26	5	10	50	24	60.22
13	27	5	10	50	72	52.21
20	28	12.5	22	40	48	93.61
2	29	20	10	30	24	72.47
16	30	20	18	50	72	87.28

Figure 1.2: Results of Design Matrix of CCD

By clicking R1: Biodiesel Yield under the analysis [+] tab, and then click fit summary, the model suggested by Design Expert shows is Linear.

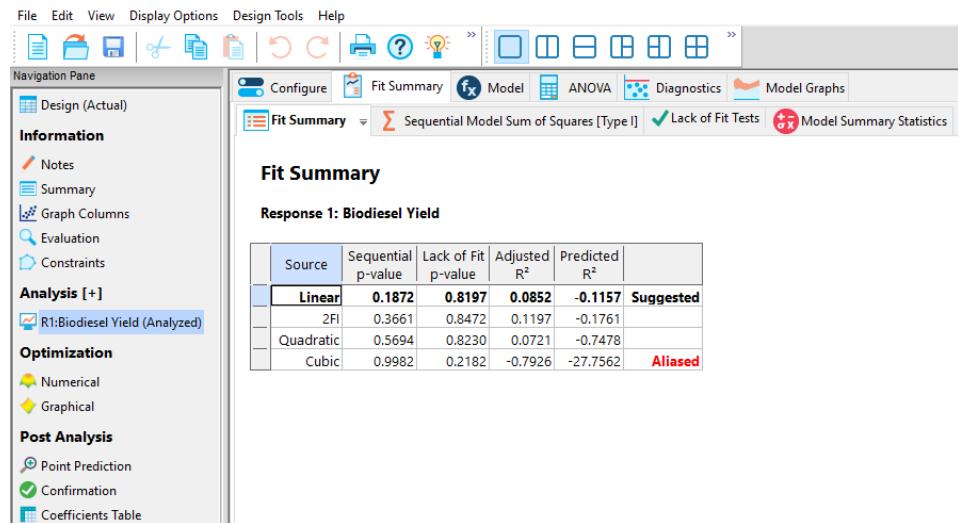


Figure 1.3: Model Suggested by Design Expert

## 1.2 ANOVA Analysis

Anova is a statistical tool used to validate the significance of the model among the different models. It is important to choose the correct and appropriate model to obtain reliable ANOVA generated data which interpret the relationships between the 4 factors and the response variable. Hence, it is crucial to examine the fit summary before choosing a model.

Based on the obtained sequential p-value, it suggested that model with  $p < 0.05$  is statistical significance. All models are not in the range suggested. However, Linaer model has the lowest sequential p-value among all the models. Hence, it is suggested to be used. In addition, Linear Model has the lowest difference between the adjusted R<sup>2</sup> and Predicted R<sup>2</sup>, indicating it is the best fit model. The obtained negative predicted R<sup>2</sup> values indicate overfitting which should be avoided as it restrains the usage of the model to determine the new observed data entry to the model (IBM Education, 2021). The Linear model has the lowest prediction sum of squares (PRESS) indicates its best predictive ability. All models have insignificant lack of fit p-values represent they are fitting the data. The Adeq Precision value represent the signal-to-noise ratio and Adeq Precision>4 indicates an adequate signal. The Linear model has the highest Adeq Precision, represent that it is not affected by unwanted noise to provide more accurate data. Therefore, Linear model is chosen for analysis and optimization tasks.

**Table 1.1: Fit Summary Obtained from Design Expert**

Source	Sequential p-value	Lack of Fit p-value	R <sup>2</sup>			PRESS	Adeq Precision
			Adjusted	Predicted	Differences		
Linear	0.1872*	0.8197*	0.0852*	0.1157*	0.0305*	5915.05*	4.5331*
2FI	0.3661	0.8472	0.1197	-0.1761	0.2958	6235.63	-
Quadratic	0.5694	0.8230	0.0721	-0.7478	0.8199	9266.41	-
Cubic	0.9982	0.2182	-0.7926	-27.7562	26.9636	152500	-

\* indicates the desired condition for each statistical parameter

The ANOVA table for the Linear model for the response on biodiesel yield in weight basis (%) is presented in Table 1.2. The table shows the distribution of the observed data and the relationship between each of the factor quantitatively.

Firstly, the Linear model has a high F-value of 1.68, with the p-value of 0.1872 represent it has only 18.72% to occur due to noise. However, the factors “B-Oil to Methanol Molar Ratio”, “C-Temperature” and “D-Reaction Time” are found to be insignificant with the p-values > 0.05. On the other hand, the factor “A-Enzyme Loading” has significant effect on the response. The ANOVA analysis suggested that the enzymatic process for biodiesel production is enzyme loading sensitive in the range of [5g, 20g].

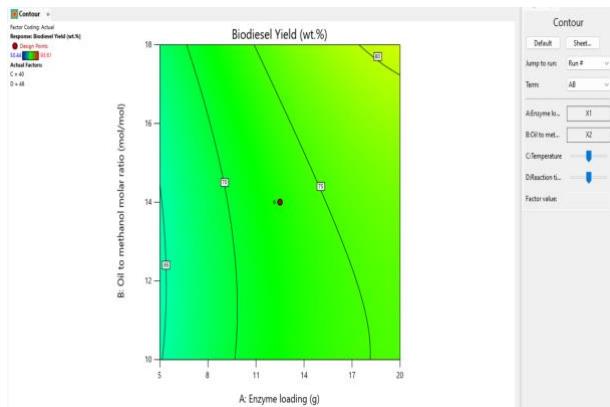
Design Expert suggested that p-value > 0.100 indicates an insignificant term and model reduction may be able to improve this model. Nevertheless, in bioprocess engineering, oil to methanol molar ratio, temperature and reaction time are crucial parameters to be considered in the batch mode of operation. Therefore, certain arrangements and selections of investigated range shall be implemented. Firstly, the design of experiment can consider widening the range of investigation for oil to methanol molar ratio to study the effect of molar ratio of oil to methanol on biodiesel yield. From the perspective of temperature, 30°C to 50°C has almost no effects on the biodiesel production. The analysis indicates that the higher range of temperature in the batch production of biodiesel should be implemented to study its effects throughout the process. From the perspective of reaction time, 24 hours to 72 hours has almost no effects on the biodiesel production. The analysis indicates that the shorter reaction time should be implemented to study its effects on the biodiesel yield.

**Table 1.2: ANOVA Table for *Linear Model***

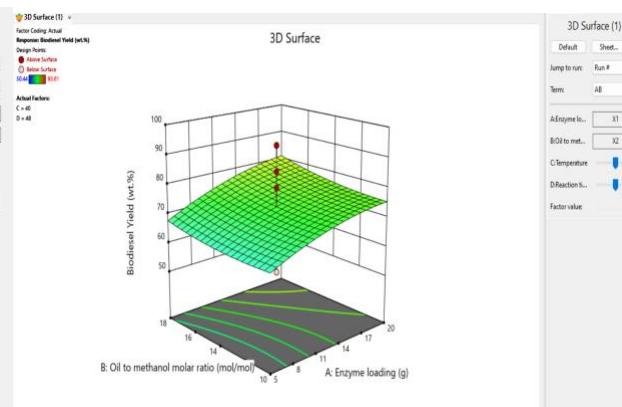
Source	Sum of Squares	df	Mean Square	F-value	p-value	
<b>Model</b>	1120.60	4	280.15	1.68	0.1872	Not Significant
A-Enzyme Loading	859.21	1	859.21	5.14	0.0323	Significant
B-Oil to Methanol Molar Ratio	108.63	1	108.63	0.6495	0.4279	Not Significant
C-Temperature	56.12	1	56.12	0.3356	0.5676	Not Significant
D-Reaction time	96.64	1	96.64	0.5778	0.4543	Not Significant
<b>Residual</b>	4181.19	25	167.25			
Lack of Fit	2933.40	20	146.67	0.5877	0.887	Not Significant
Pure Error	1247.80	5	249.56			
<b>Cor Total</b>	5301.79	29				

### 1.3 Effects of Variables on Biodiesel Yield

The perturbation or disturbance plot in response surface designs shows how the response changes as an element moves away from the designated reference point while all the other variables stay unchanged. Furthermore, a contour graph was utilized to compare the components of a mixture and a two-dimensional (2D) depiction of the response. It may show the relationship between numerical inputs, responses, and mixture elements. There are a total of 4 variables involved in this case which is the enzyme loading (g), oil to methanol molar ratio, temperature (°C), and reaction time (h). However, there is only one response that needed to be optimized which is the biodiesel yield (%). A 3-dimensional (3D) surface can aid to illustrate a contour plot as it is a graphical technique that plots constant z slices which is the contour in a 2D format.

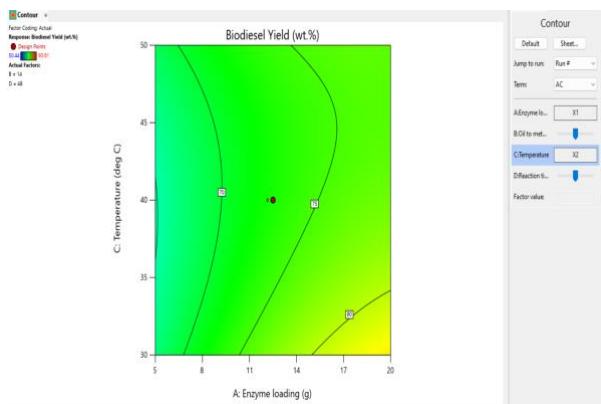


**Figure x: Contour graph of Enzyme Loading Oil to Methanol Molar Ratio**

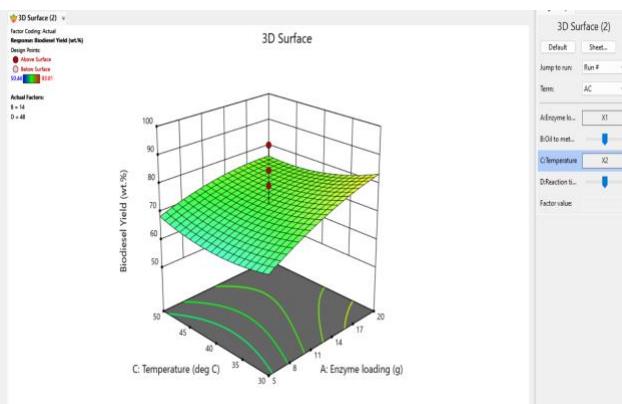


**Figure x: 3D graph of Enzyme and Oil to Methanol Molar Ratio**

Figures x and x depict the connection between enzyme loading (A) and oil-to-methanol molar ratio (B). These two figures show the connection between enzyme loading and the oil-to-methanol molar ratio while the reaction duration temperature was kept constant at 40°C and was held constant at 48 hours. The lowest biodiesel yield obtained is where the enzyme loading of around 12.5 g and oil to methanol molar ratio of 14, which get the lowest biodiesel yield of 70%. From the observation, it proves that the interaction of oil to methanol molar ratio and temperature (°C) was insignificant, which agrees with the results of the p-value (probability of obtaining an F-ratio as large or larger than the observed ones) from the AVONA table obtained for BC which is 0.8286 which is relatively large.

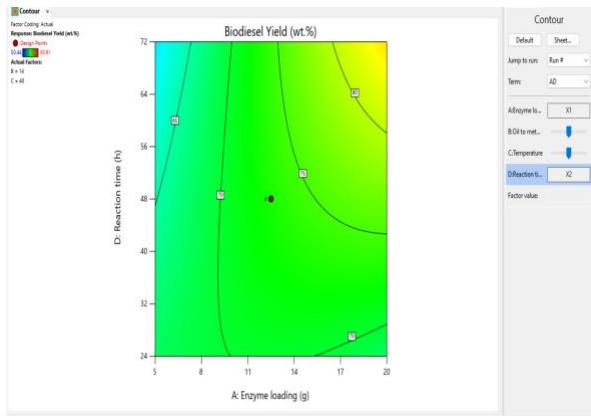


**Figure x: Contour graph of Enzyme loading and temperature**

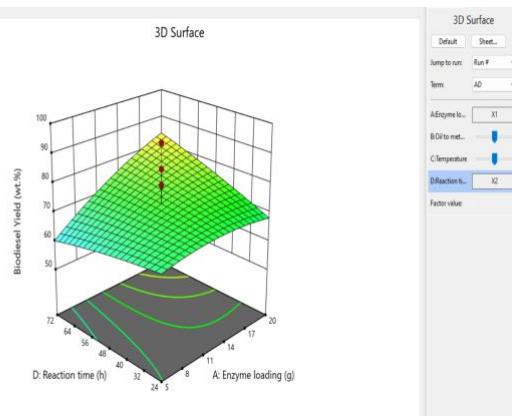


**Figure x: 3D graph of Enzyme loading and temperature**

Figures x and x depict the connection between enzyme loading (A) and temperature (C). These two graphs illustrate the relationship between enzyme loading and temperature while maintaining a constant oil-to-methanol molar ratio of 14 and a 48-hour reaction duration. Where the enzyme dosage is around 12.5 g and the temperature is 40 °C, the lowest biodiesel production of 70% is attained. From the observation, it proves that the interaction of oil to methanol molar ratio and temperature (°C) was insignificant, which agrees with the results of the p-value (probability of obtaining an F-ratio as large or larger than the observed ones) from the AVONA table obtained for BC which is 0.5039 which is relatively large.

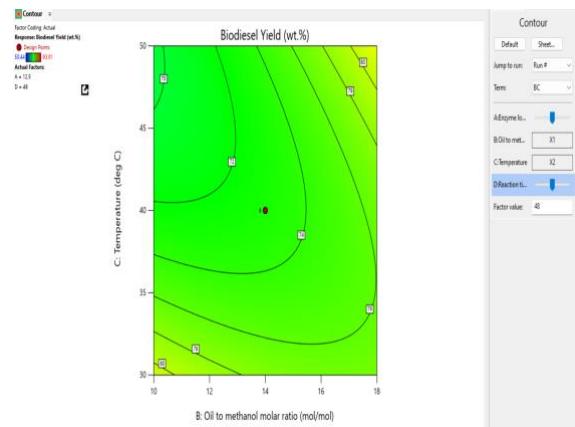


**Figure x: Contour graph of Enzyme loading versus reaction time**

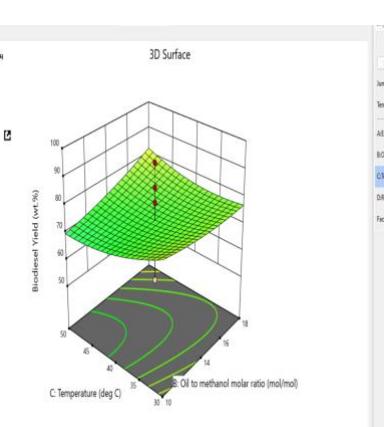


**Figure x: 3D graph of Enzyme loading versus reaction time**

Figures x and x depict the connection between enzyme loading (A) and reaction time (D). These two figures illustrate the link between enzyme loading and reaction time while maintaining a constant oil to methanol ratio of 14 and a 40-hour reaction duration. When the temperature is 48 °C and the oil-to-methanol molar ratio is approximately 12.5, the lowest biodiesel yield is produced, which is 60%. From the observation, it proves that the interaction of oil to methanol molar ratio and temperature (°C) was significant, which agrees with the results of the p-value (probability of obtaining an F-ratio as large or larger than the observed ones) from the AVONA table obtained for BC which is 0.0882 which is relatively small.

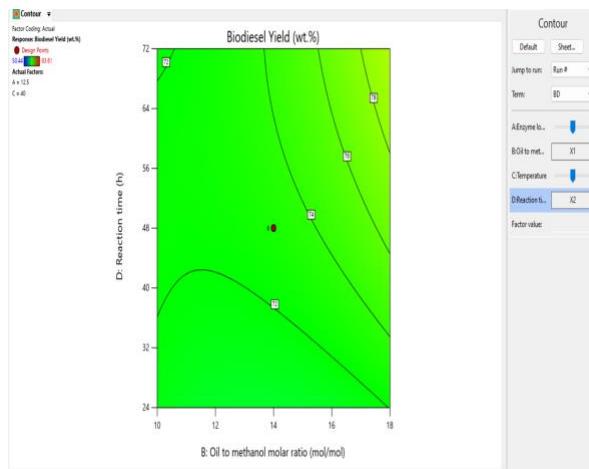


**Figure x: Contour graph of oil to methanol molar ratio and temperature**

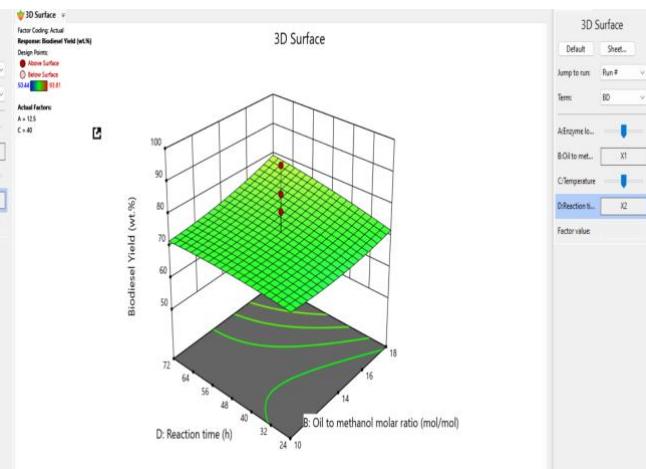


**Figure x: 3D graph of oil to methanol molar ratio and temperature**

Figures x and x depict the connection between the oil-to-methanol molar ratio (B) and the temperature (C). These 2 figures show the relationship between oil to methanol molar ratio and temperature while the enzyme loading was kept constant at 12.5 g and the reaction time was also kept constant at 48 h. The oil-to-methanol molar ratio of around 14 and the 40-hour reaction duration produce the lowest biodiesel output of 58%, respectively. From the observation, it proves that the interaction of oil to methanol molar ratio and temperature (°C) was insignificant, which agrees with the results of the p-value from the AVONA table obtained for BC which is 0.2340 which is relatively large.

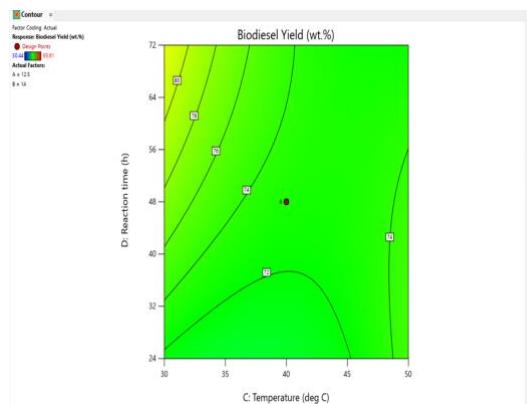


**Figure x: Contour graph of oil to methanol molar ratio and reaction time**

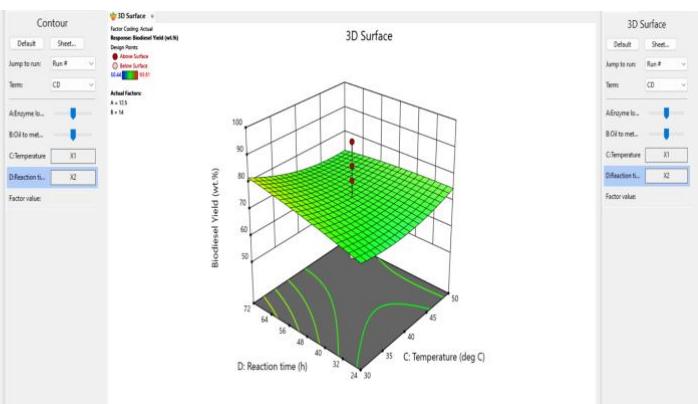


**Figure x: 3D graph of oil to methanol molar ratio and reaction time**

Figures x and x depict the connection between the oil-to-methanol molar ratio (B) and the reaction time (D). These two figures demonstrate the relationship between the oil-to-methanol molar ratio and the reaction time while maintaining a constant enzyme loading of 12.5 g and a temperature of 40 °C. The oil-to-methanol molar ratio of around 14 and the 40-hour reaction duration produce the lowest biodiesel output of 58%, respectively. From the observation, it proves that the interaction of oil to methanol molar ratio and temperature (°C) was insignificant, which agrees with the results of the p-value (probability of obtaining an F-ratio as large or larger than the observed ones) from the AVONA table obtained for BC which is 0.5957 which is relatively large.



**Figure x: Contour graph of temperature and reaction time**



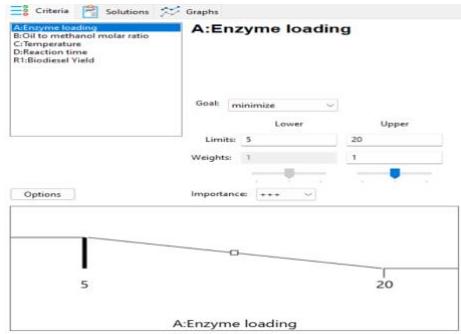
**Figure x: 3D graph of temperature and reaction time**

The relationship between the temperature (C) and reaction time (D) was shown in Figure x and Figure x. These 2 figures show the relationship between temperature and reaction time while the enzyme loading was kept constant at 12.5 g and the oil to methanol molar ratio was also kept constant at 14. Figure x demonstrates that when temperature and reaction time rise, the biodiesel output rises as well. The lowest biodiesel yield obtained is where the temperature of around 40 °C and reaction time of 40 hours, which get the lowest biodiesel yield of 57 %. This contour-shaped graph shows that the interaction of reaction time (h) and temperature (°C) was insignificant, which agrees with the results of the p-value from the AVONA table obtained for the CD which is 0.3457 which is relatively large.

## 1.4 Optimization

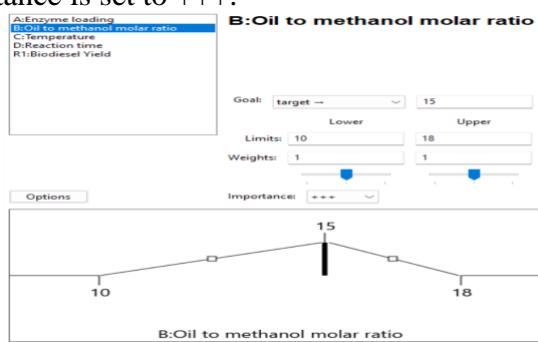
The criteria for the 4 independent variables such as enzyme loading (A), oil to methanol molar ratio (B), temperature (C) and reaction time (D) and response variable such as biodiesel yield should be set accordingly to obtain the desirable results. Numerical optimization of Design Expert can be maximized, minimize or in target. Importance is a tool to change the relative priorities to achieve the goals establish for some or all variables. The higher the importance we set, it means to emphasize one over the rest.

There are 5 levels of importance from single plus (+) to 5 plus (++++). According to this scenario, we set the importance to +++, which is a medium setting. By leaving all importance criteria at their defaults, no goals are favored over others.



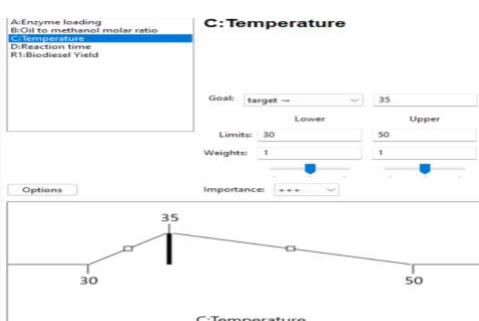
**Figure x: Set the Goal of Enzyme Loading to Minimise**

In Figure x, the goal of enzyme loading is set to be minimise. The lower limit and upper limit of A is set to be 5 g and 20 g. The importance is set to +++.



**Figure x: Set the Goal of Oil to Methanol Molar Ratio to Target**

In figure x, the goal of oil to methanol molar ratio is set to be target. As mention in question, the oil and methanol ratio are fixed at 1:15. The upper and lower limit is set as 10 and 18 respectively. The importance is set as +++.



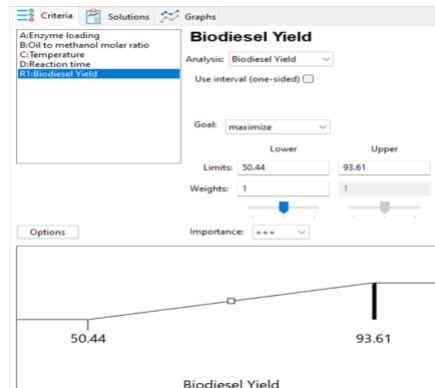
**Figure x: Set the Goal of Temperature to Target**

In figure x above, the goal of temperature is set as target. Accoding to the question, the temperature is fixed at  $35^{\circ}\text{C}$  with +++. The lower and upper limit is set as 30 and 50 respectively.



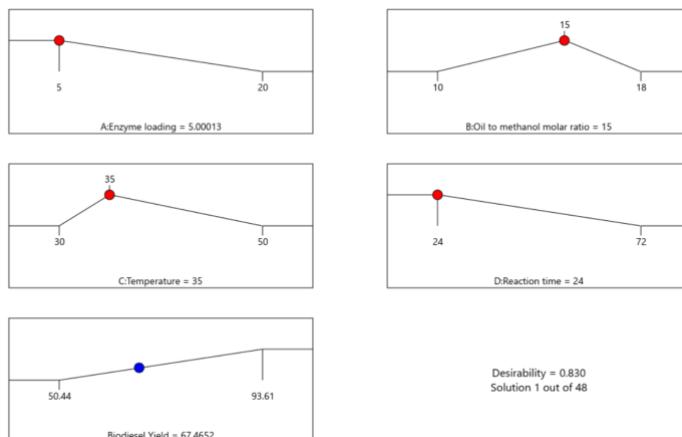
### Figure x : Set the Goal of Reaction Time to Minimise

In figure x, the goal of reaction time is set to be minimise. The lower and upper limit is set as 24 hours and 72 hours respectively with +++ under the selection of importance.



### Figure x: Set the Goal of Biodiesel Yield to Maximise

In figure x, the goal of biodiesel yield is set to maximise to achieve the highest biodiesel yield. 50.44 % is set as the lower limit while 93.61 % is set as the upper limit. The importance is set to +++.



### Figure x: Solution Generated for A,B,C,D and Biodiesel Yield

From figure x, there are 48 solutions generated by Design Expert Software. The solution 1 out of 48 was chosen to be the best answer because it has a highest desirability among all the solutions. The value of A, B, C, D and biodiesel yield in solution 1 are 5 g, 15, 35 °C, 24 hours and 67.4652 %. The solution obtained has a desirability of 0.83 which is acceptable because it is almost equal to 1. This design can be referred to achieve the optimum biodiesel yield. Under the “Factor Tool”, other alternatives can be checked to achieve the optimum biodiesel yield with different values for each factors.

### Constraints

	Name	Goal	Lower Limit	Upper Limit	Lower Weight	Upper Weight	Importance
	A:Enzyme loading	minimize	5	20	1	1	3
	B:Oil to methanol molar ratio	is target = 15	10	18	1	1	3
	C:Temperature	is target = 35	30	50	1	1	3
	D:Reaction time	minimize	24	72	1	1	3
	Biodiesel Yield	maximize	50.44	93.61	1	1	3

### Figure x: Overall Numerical Optimization

## **CHAPTER 2**

### **Case Studies of Different Response Surface Methodology Design**

#### **2.1 Taguchi Orthogonal Array (OA) Design – A case study to optimize the chemical process by using Taguchi Method of Experimental Design (Jee Pei Qi)**

##### **a) A brief background of the example**

A business that manufactures chemical products is based in Mumbai, India. Product X which is made from rubber, polymers, plastics, adhesives, textiles, paper, and medicines all use product X extensively. When added to polyvinyl chloride (PVC), zinc borate retards flames and serves as an epoxy resin and smoke suppressor. Wide-ranging uses for Zinc Borate exist in numerous fields. Due to the significant Loss-on-Ignition (LoI) and inadequate chemical Borate composition in Zinc Borate, this business streak has declined tremendously due to the poor output of zinc borate drew the management's notice, and they emphasized the proportion of the chemical borate and the larger LoI concerns. Furthermore, an optimization method should be applied to raise the chemical Borate and decrease the LoI. In addition, the temperature, the amount of borex, the pH, and the length of time the reaction takes to complete will all have an impact on the product's final shape. Since they will also have an impact on the end result in this situation, the aforementioned elements will also need to be optimized.

##### **b) The objective function to be optimized and the relevant variables, conditions and assumptions**

This case study's main goals are to increase chemical Borate concentration ( $B_2O_3$ ) and decrease Loss-on-Ignition (LoI) %. However, this case study's focus will simply be reducing the Loss-on-Ignition (LoI). The temperature, amount of borex, pH, and reaction time are other important factors in this case study since they all have an impact on the final result. Additionally, as the signal-to-noise (S/N) ratio is used to create equations in this case study, the lower the value of Loss-on-Ignition (LoI) is deemed to be better.

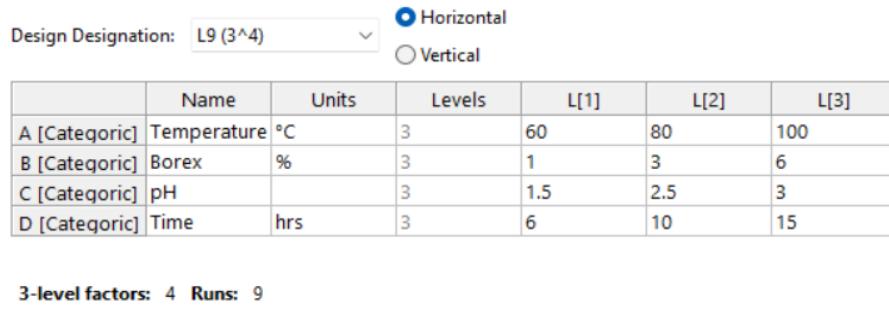
$$\frac{S}{N} \text{ Ratio} = -10 \times \log \frac{1}{n} (\sum y^2)$$

Three replicates of the LoI in each run are considered while creating the conditions for this case study, resulting in three replies for each LoI. The reaction is predicated on the premise that Zinc Borate is the sole chemical created throughout the whole process.

##### **c) Generation of data by using software simulation or other equivalent methods**

The Design Expert programme will be employed, and Taguchi design will be the method of optimisation. A design for LoI will be made to minimize LoI. Under the Miscellaneous in Factorial function, the Taguchi Orthogonal Array (OA) design will be used in the designs. In this case study, the product will be impacted by a total of four elements (Temperature, Borex, pH, and Time), each of which will have three levels. As a result, the design designation employed will be the L9 (34) as shown

in Figure x. L9(3<sup>4</sup>) denotes that there will be a total of nine levels, divided into four categories, each of which has three levels.



**Figure x: Four factors involved in this case study**

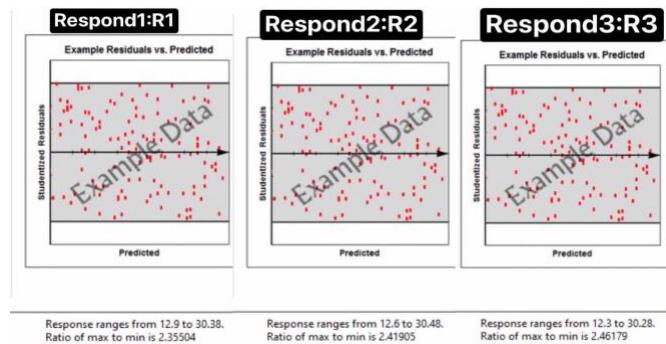
Next, each run will include three copies of each LoI, resulting in three responses for each LoI. The three replies' input values will need to be entered in the next step; these values were obtained directly from the article. Prior to entering the answer data, the standard must sort in ascending order. The value of replies for LoI is depicted in Figure x below.

Std	Run	Factor 1 A:Temperature °C	Factor 2 B:Borex %	Factor 3 C:pH	Factor 4 D:Time hrs	Response 1 R1	Response 2 R2	Response 3 R3
1	5	60	1	1.5	6	25.55	25.66	25.77
2	3	60	3	2.5	10	30.38	30.48	30.28
3	6	60	6	3	15	25.76	25.36	25.56
4	7	80	1	2.5	15	19.9	19.95	20
5	9	80	3	3	6	20.58	20.08	20.33
6	2	80	6	1.5	10	19.22	19.02	19.42
7	4	100	1	3	10	15.34	15.24	15.14
8	1	100	3	1.5	15	14.03	14.18	14.33
9	8	100	6	2.5	6	12.9	12.6	12.3

**Figure x: Value of three responses for LoI**

After entering all the response values, the four components must convert their numerical values into discrete numbers so that the response data can be analyzed and the f and p values can be obtained in the ANOVA table.

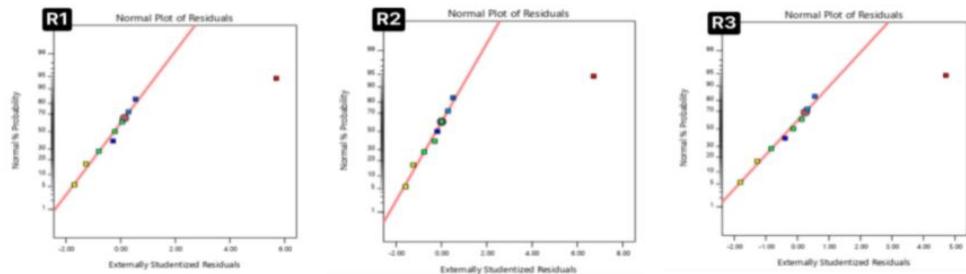
#### d) Analysis of the data and responses obtained



**Figure x: Transformation of the model**

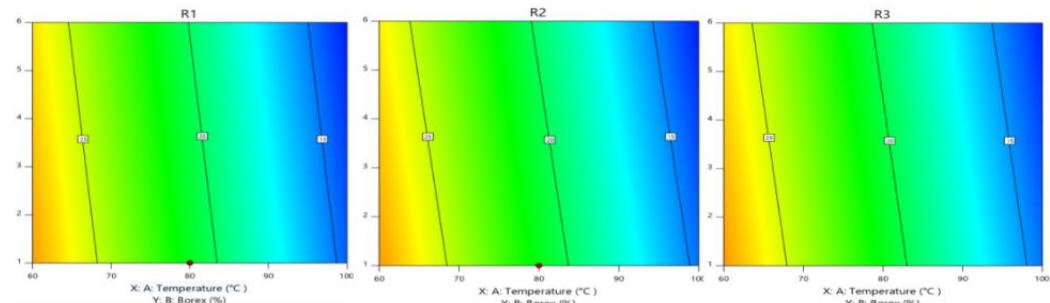


space when the ratio is greater than 4. For this case study, all three responses have a ratio that is bigger than 4 which mean that it is desirable. of the replies for this case study have a ratio greater than 4, indicating the desired result.



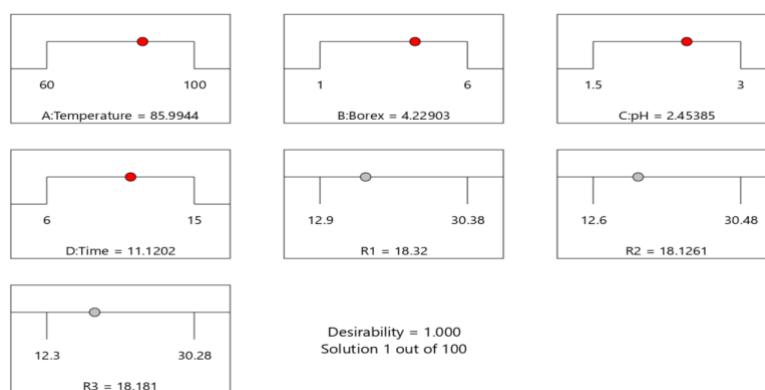
**Figure x: Normal Plot of Distribution**

To ascertain whether or not the residuals follow a normal distribution, the normal plot of the residuals is utilized. The residuals are considered to have a normal distribution if the points follow a straight line. Figure x demonstrates that virtually all of the data are fitted onto the line, with just one point not being linearly fitted. The linear line's plot demonstrates that it adheres to the normal distribution, producing a successful outcome.



**Figure x: Contour graph between temperature and percentage of borex**

Figure x above is the contour graph between temperature and percentage of borex for three responses. In this case study, the time taken, and pH will be constant. Moreover, figure x shows that the temperature will be 80°C and the borex will be one which indicates the influence of temperature and percentage of borex toward the product formed.



### **Figure x: Ramp solution for the optimization process**

The red point in Figure x will represent the ideal factor setting. There are 100 possible solutions in this scenario, with solution 1 being the term A (temperature), which may be fixed at 86°C, term B (percentage of borex), which is 4.23%, term c (pH), which is 2.45385, which is acidic, and term d (duration), which will be 11.1202 hours. The LoI value in the product, which has been reduced to roughly 18 for each of the three replies, is what is meant by the R-value. In reality, the value of LoI may be reduced by setting all the variables to the desired value.

#### **e) Justify the selection of the experimental design compared to others**

The case study's goal is to reduce the value of LoI, and there are a total of four components that will have an impact on the final output. As a result, the Taguchi design technique is used in this case study to optimize the problem of the relationship between design elements and performance. The Taguchi approach, which has four elements and may modify many variables simultaneously, is appropriate for this case study. In addition, it offers a quick and effective way to obtain additional quantitative data while concurrently optimizing aspects. Additionally, there are two types of elements that might have an impact on a product: noisy factors and controlled factors. Both controlled parameters and noise factors are established and maintained at a certain level during the experimental phase. While the noise factor is used to evaluate the degree of their influence on the process, controllable elements are utilized to alter it. This is another justification for using the Taguchi approach over alternative designs. This is so that the controllable factor and the noise factor may be employed in the Taguchi method's two orthogonal arrays, the inner array and the outer array. The two orthogonal arrays were referred to as a cross-array experiment because it can identify the best design parameters with the least amount of noise effect on the final result. In reality, compared to other design methods, the Taguchi technique will be better appropriate for usage in this case study.

## **2.2 Central Composite Design to Optimize Coagulation- Kinetic Model, Thermodynamics, and Chemical Composition of Pomelo Essential Oil Extraction Using Steam Distillation (Choong Yi Jie)**

### **Brief Background**

Pomelo is a fruit tree usually found in Vietnam. It has a thick exocarp that contributes to a significant portion of its total weight. As a result, small amount of the fruit pulp is consumed. Thus, researchers wish to put more attention on studying other parts of the pomelo such as leaves, seeds and peel to show their potential benefits. For example, the essential oil extracted from the pomelo peel can help to release stress or be utilized in hair care therapy.

There are various methods used to extract essential oil from the pomelo peel. For example, solvent-free microwave extraction (SFME), hydrodistillation (HD) and supercritical CO<sub>2</sub> extraction. The distillation process consists of two phases involves rapid distillation which extracts the volatile oil

components on or near to the surface and followed by the slow distillation, which involves the distillation of plant tissues and followed by distillation on the outer surface of the raw material.

The distillation process is affected by 3 factors, including the mass of the material, steam flow rate and the extraction time. To optimize the extraction process, response surface methodology, specifically central composite design (CCD) is utilized. It can determine the optimal conditions for the extraction process by analyzing the relationship between the input variables and the response variables.

### **Objective Function, Relevant Variables, Conditions and Assumptions**

In this study, the main objective is to optimize the process parameters to obtain maximize yield of essential oil. The Table 2.3 shows the variables and levels.

Table 2.3: Range and levels of independent variable.

Variables	Range and levels				
	- $\alpha$	-1	0	1	$\alpha$
A. Mass of materials (g)	163	300	500	700	836
B. Steam Flow Rate (mL/min)	1.6	1.8	2.1	2.4	2.6
C. Extraction Time (min)	40	60	90	120	140

By using the central composite design (CCD), a quadratic model was selected to represent the relationship between the three factors with five levels, as shown in Table 2.3. The equation below shows how the essential oil (Y) is being affected by the variables (A, B and C).

$$Y = +6.12 - 0.1929A + 0.2315B + 0.1865C + 0.0992AB - 0.0073AC + 0.0936BC - 0.2252A^2 - 0.5781B^2 - 0.1988C^2$$

### **Generation of data by using the Design Expert**

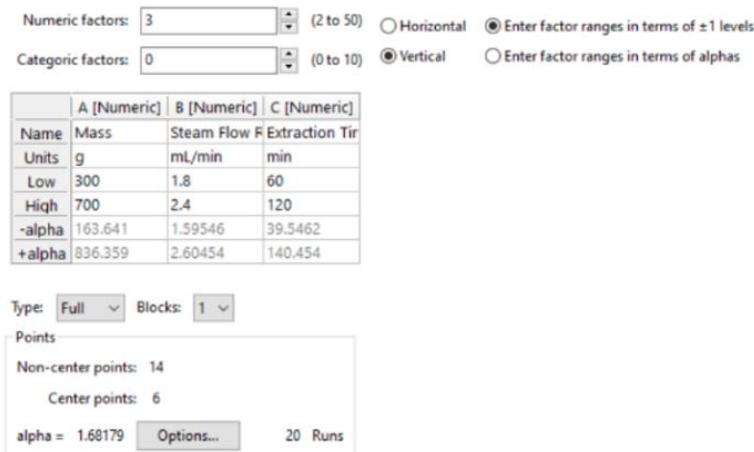


Figure 2.3.1: Central Composite Design inserted with Numeric Factor

Design Expert is chosen as the software to simulate the data for this study. All the data variables are filled into prepared table. Figure 2.3.1 shows the data obtained from Table 2.3. The ‘low’ column in Figure 2.3.1 is filled with the value in the level of -1 from Figure 2.3, whereas the ‘high’ column in Figure 2.3.1 is filled with the values in the level of +1 from Figure 2.3. As the value of alpha ( $\alpha$ ) is not specified, so the default rotatable value with  $k < 6$ , which is 1.68179 is used. Although it has a slight difference from the value given in Table 2.3, it does not significantly affect the result. The study involves 20 runs to improve the accuracy of the results.

## Central Composite Design

Responses:  (1 to 999)

Horizontal  
 Vertical

	Name	Units
	Yield (Actual)	%, mL/100g
	Yield (Residu)	%, mL/100g

Figure 2.3.2: Responses of the study

After the variables have been inserted, the responses of the study required to be filled. There are 2 responses in this study, the actual yield and residual yield of essential oil.

Std	Run	Factor 1 A:Mass g	Factor 2 B:Steam Flow Rate mL/min	Factor 3 C:Extraction Time min	Response 1 Yield (Actual) %, mL/100g	Response 2 Yield (Residual) %, mL/100g
17	1	500	2.1	90	6.2757	-0.0748
13	2	500	2.1	39.5462	5.2734	0.096
8	3	700	2.4	120	5.6235	0.0006
11	4	500	1.59546	90	4.1334	-0.1422
6	5	700	1.8	120	4.6539	0.0571
20	6	500	2.1	90	6.1761	-0.1422
19	7	500	2.1	90	5.977	-0.0138
1	8	300	1.8	60	4.964	0.1567
3	9	300	2.4	60	5.1625	0.0386
15	10	500	2.1	90	6.25	0.0996
9	11	163.641	2.1	90	5.9062	-0.0645
5	12	300	1.8	120	5.2784	-0.0645
10	13	836.359	2.1	90	5.0829	-0.0252
16	14	500	2.1	90	6.0546	-0.0054
12	15	500	2.60454	90	4.8595	0.1308
4	16	700	2.4	60	4.964	-0.0181
14	17	500	2.1	140.454	5.8649	-0.1135
18	18	500	2.1	90	5.977	0.0077
7	19	300	2.4	120	5.653	-0.0765
2	20	700	1.8	60	4.5669	0.0302

Figure 2.3.3: Design table of CCD in Design-Expert

After the responses have been filled, a set of data will be generated for the factors with random sequence (which will be same sequence for standardized but different for the run). The data for responses are then filled according to the study and the data for the optimization will be generated.

### Analysis of data and responses

#### Fit Summary

##### Response 1: Yield (Actual)

Source	Sequential p-value	Lack of Fit p-value	Adjusted R <sup>2</sup>	Predicted R <sup>2</sup>	
Linear	0.2308	0.0008	0.0853	-0.1316	
2FI	0.9495	0.0004	-0.0966	-0.8110	
Quadratic	< 0.0001	0.7265	0.9643	0.9271	Suggested
Cubic	0.5624	0.7813	0.9614	0.9374	Aliased

Figure 2.3.4: Fit summary of actual yield of essential oil

## Fit Summary

### Response 2: Yield (Residual)

Source	Sequential p-value	Lack of Fit p-value	Adjusted R <sup>2</sup>	Predicted R <sup>2</sup>	
Linear	0.1991	0.5441	0.1046	-0.2112	Suggested
2FI	0.4313	0.5284	0.1016	-0.9523	
Quadratic	0.9129	0.3435	-0.1106	-1.9957	
Cubic	0.2474	0.4756	0.1588	-5.5619	Aliased

Figure 2.3.5: Fit summary of residual yield of essential oil

According to Figure 2.3.4, Expert Design suggested that a quadratic model is selected for the actual yield of essential oil, while a linear model is suggested for the residual yield of essential oil.

### ANOVA for Quadratic model

### Response 1: Yield (Actual)

Source	Sum of Squares	df	Mean Square	F-value	p-value	
Model	7.32	9	0.8137	58.10	< 0.0001	significant
A-Mass	0.5081	1	0.5081	36.28	0.0001	
B-Steam Flow Rate	0.7316	1	0.7316	52.24	< 0.0001	
C-Extraction Time	0.4747	1	0.4747	33.90	0.0002	
AB	0.0787	1	0.0787	5.62	0.0392	
AC	0.0004	1	0.0004	0.0304	0.8650	
BC	0.0701	1	0.0701	5.00	0.0493	
A <sup>2</sup>	0.7308	1	0.7308	52.18	< 0.0001	
B <sup>2</sup>	4.82	1	4.82	343.87	< 0.0001	
C <sup>2</sup>	0.5696	1	0.5696	40.68	< 0.0001	
Residual	0.1400	10	0.0140			
Lack of Fit	0.0506	5	0.0101	0.5657	0.7265	not significant
Pure Error	0.0894	5	0.0179			
Cor Total	7.46	19				

Figure 2.3.6: Anova Analysis of Actual Yield of Essential Oil

Based on the result given in Figure 2.3.6, it shows that the actual yield model is significant, indicating that the quadratic model is suitable for the analysis. For the factor analysis, the p-value of AB, BC, A<sup>2</sup>, B<sup>2</sup> and C<sup>2</sup> are < 0.05, indicating that they are significant factors in this model. On the other hand, the p-value of AC > 0.1, indicating that it is insignificant factor. Insignificant term may be removed from the model. However, AC will not be removed from this model as A<sup>2</sup> and C<sup>2</sup> are significant terms that may be ignored if AC is removed.

The lack of fit for this model has a F-value of 0.5657 and P-value of 0.7265, indicating that it is insignificant relative to pure error. Thus, the chance for the lack of fit F-value to occur due to noise is around 73%. The insignificant of lack of fit indicates that the model is suitable for analysis.

## Fit Statistics

<b>Std. Dev.</b>	0.1183	<b>R<sup>2</sup></b>	0.9812
<b>Mean</b>	5.43	<b>Adjusted R<sup>2</sup></b>	0.9643
<b>C.V. %</b>	2.18	<b>Predicted R<sup>2</sup></b>	0.9271
		<b>Adeq Precision</b>	24.1908

Figure 2.3.7: Fit statistic for Actual Yield of Essential Oil

Figure 2.3.7 shows that  $R^2$  value of 0.9812 indicates that high reliable of the model. The predicted  $R^2$  value and adjusted  $R^2$  value are close enough with a slight difference of 0.04 ( $< 0.2$ ). At the same time, they are quite close to the  $R^2$  value. These values suggested that quadratic model is a better fit. In addition, the Adeq Precision with value of 24 indicates that the adequate signal can measure the design space accurately.

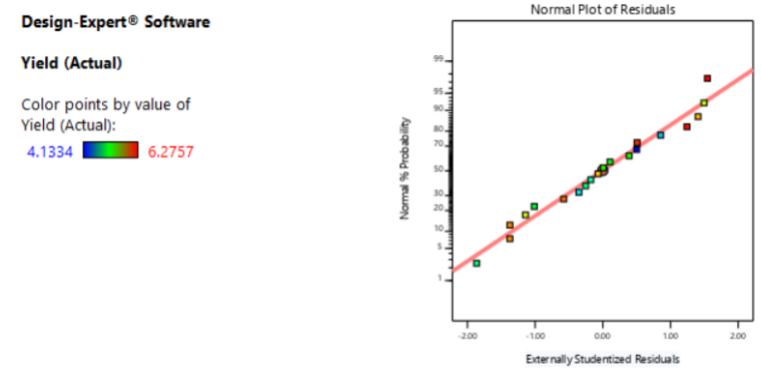


Figure 2.3.8: Graph indicates a normal plot of residuals

Finally, the result with the value  $> 4$  indicates that it is considered suitable for utilization.

### Numerical optimization

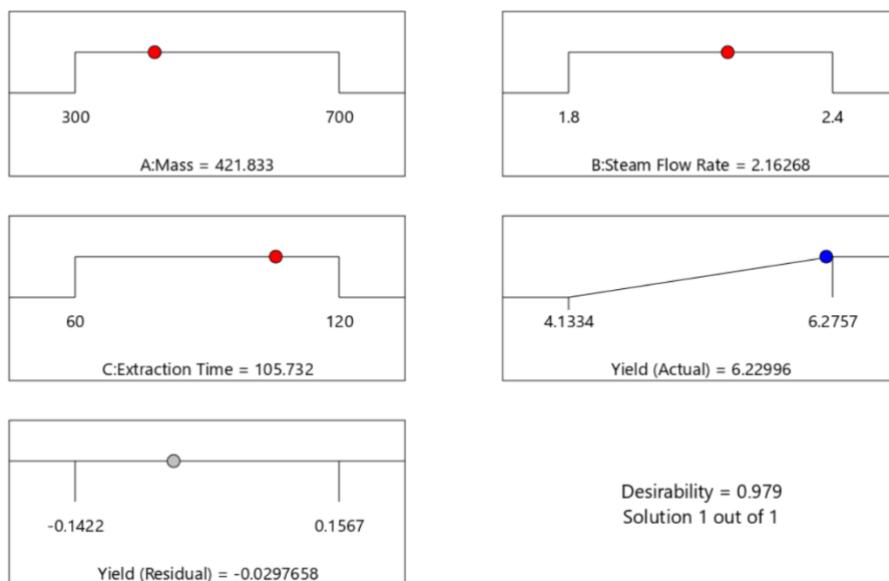


Figure 2.3.9: Numerical optimization of this case

To maximize the actual yield of essential oil, only a solution is provided. Hence, the optimal conditions to achieve actual yield of 6.23%, including (A) the mass of the material at 421.833g, (B) the steam flow rate at 2.163 mL/min and (C) the extraction time of 105.732 min. The desirability of the solution is 0.979.

### Justify the selection of the experimental design

The response surface design (RSM) is suggested to be used for this study as all the variables are numerical variables, indicating there are limited options. The Design Expert Software provides 3 RSM designs, including optimal design, Box-Behnken and central composite design CCD. In this study, CCD is used as the alpha values are take into consideration for the design. The alpha values can only be filled into the CCD section in the software.

## **2.3 3 Level Factorial Design – Thermal cracking of petroleum residue oil using three level factorial design (Khadija Nadim)**

**Background:**

As the demand for lighter fuels such as gasoline, diesel and kerosene has increased, heavy oils are subjected to thermal cracking process to break down the larger particles to smaller ones. The crude oils undergo high pressure and intense heat, thus making the process more costly. Thus, the use of adequate experimental design is particularly important. The experiment consists of nitrogen and hydrogen gases being supplied to the batch reactor vessel at constant pressure. The reactor is then heated to the temperatures between 120 – 180 kPa. The reaction is carried out for a fixed period of time in the range 40 – 100 min. After which the reactor is cooled down to room temperature.

### **Objective function and Variables:**

The operating conditions that are examined under these reactions is temperature, pressure, and reaction time. The temperature ranges from 400 – 480°C, reaction time 40 – 100 min, and pressure 120–180 kPa. Based on three-level factorial design model, the value of total conversion obtained, and the yields of total distillate fuels, gasoline, kerosene, and diesel are found. The most influential factor on each experimental design response is identified and the predicted values are also compared with the experimental values.

Factor	Code	Unit	Low level (-1)	Medium level (0)	High level (+1)
Temperature	A	°C	400	440	480
Time	B	Min	40	70	100
Pressure	C	kPa	120	150	180

Source: (Alsobaai, 2011)

As the design aims to increase the yield, a quadratic model is used to correlate the thermal cracking factors to the total conversion.

$$yi = \beta_0 + \sum_{j=1}^n \beta_i X_i + \sum_{j=1}^n \beta_{ii} X_i^2 + \sum_{i=1}^{n-1} \sum_{i < j=2}^n \beta_{ij} X_i X_j$$

where  $yi$  is the yield of the product,  $\beta_0$  is the mean of the observations,  $\beta_i$ ,  $\beta_j$  and  $\beta_{ij}$  are unknown constants and  $n$  is the number of observations.

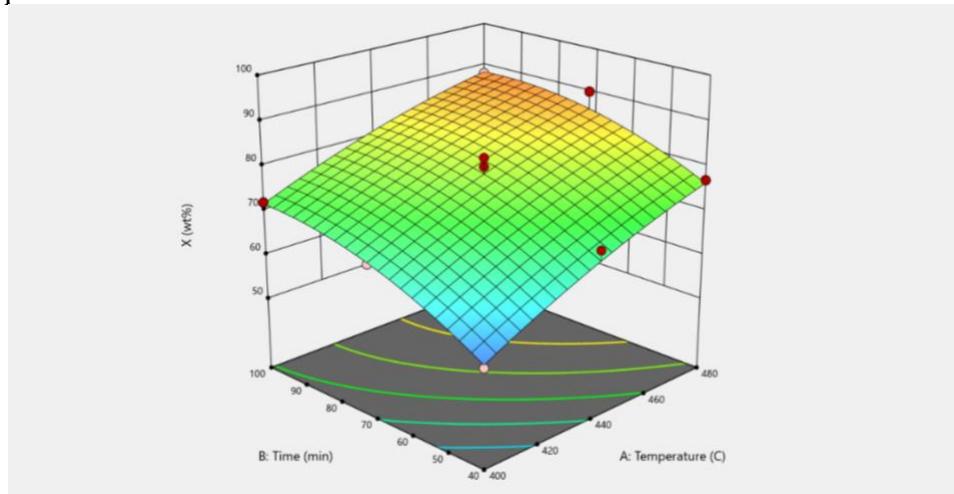
### **Data obtained:**

Runs =  $3^3 + 5$  (centre points) = 32



$$\text{Conversion, } X = -238.97909 + 1.01929A + 1.38121B - 0.25388C - 4.09200E-003B^2$$

3D Model Graph:



Analysis of variance (ANOVA) for total distillate fuels, Y:

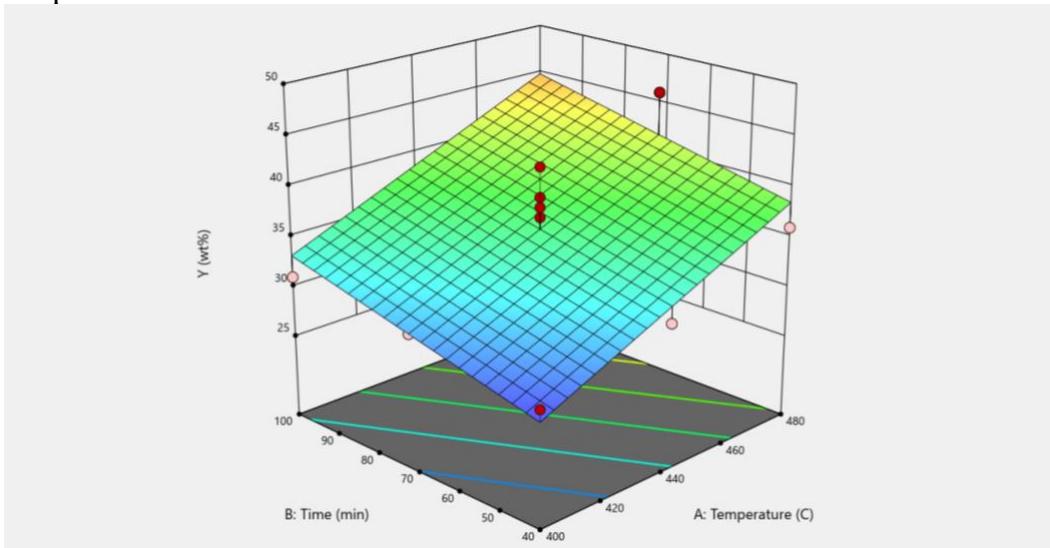
	Source	Sum of Squares	df	Mean Square	F-value	p-value	
	<b>Model</b>	798.56	3	266.19	36.24	< 0.0001	significant
	A-Temperature	612.50	1	612.50	83.39	< 0.0001	
	B-Time	180.50	1	180.50	24.57	< 0.0001	
	C-Pressure	5.56	1	5.56	0.7564	0.3919	
	<b>Residual</b>	205.66	28	7.35			
	Lack of Fit	164.16	23	7.14	0.8599	0.6429	not significant
	Pure Error	41.50	5	8.30			
	<b>Cor Total</b>	1004.22	31				

The **Model F-value** of 36.24 implies the model is 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, B are significant model terms. Thus, the rest are removed to improve the model.

The equation obtained becomes,

$$\text{Total distilled fuels, } Y = -38.48958 + 0.14583A + 0.10556B + 0.018519C$$

3D Model Graph:



### Analysis of variance (ANOVA) for gasoline yield:

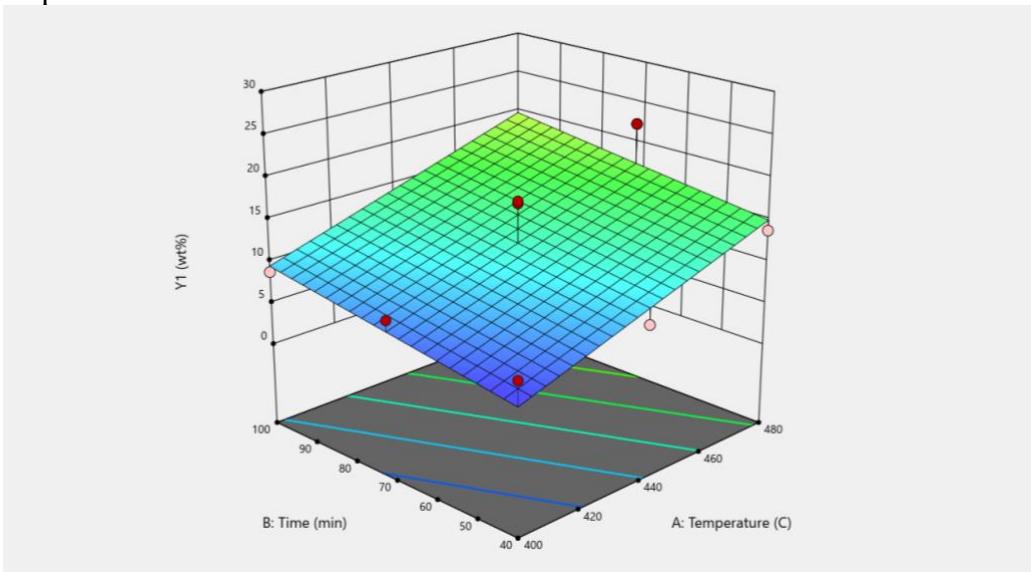
Source	Sum of Squares	df	Mean Square	F-value	p-value	
<b>Model</b>	591.57	3	197.19	16.15	< 0.0001	significant
A-Temperature	472.27	1	472.27	38.67	< 0.0001	
B-Time	99.88	1	99.88	8.18	0.0079	
C-Pressure	19.43	1	19.43	1.59	0.2176	
<b>Residual</b>	341.93	28	12.21			
Lack of Fit	269.21	23	11.70	0.8049	0.6781	not significant
Pure Error	72.71	5	14.54			
<b>Cor Total</b>	933.50	31				

The Model F-value of 16.15 implies the model is 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, B are significant model terms. Thus, the rest are removed to improve the model.

The equation obtained becomes,

$$\text{Gasoline yield, } Y_1 = -54.74144 + 0.12806A + 0.078519B + 0.034630C$$

3D Model Graph:



### Analysis of variance (ANOVA) for kerosene yield:

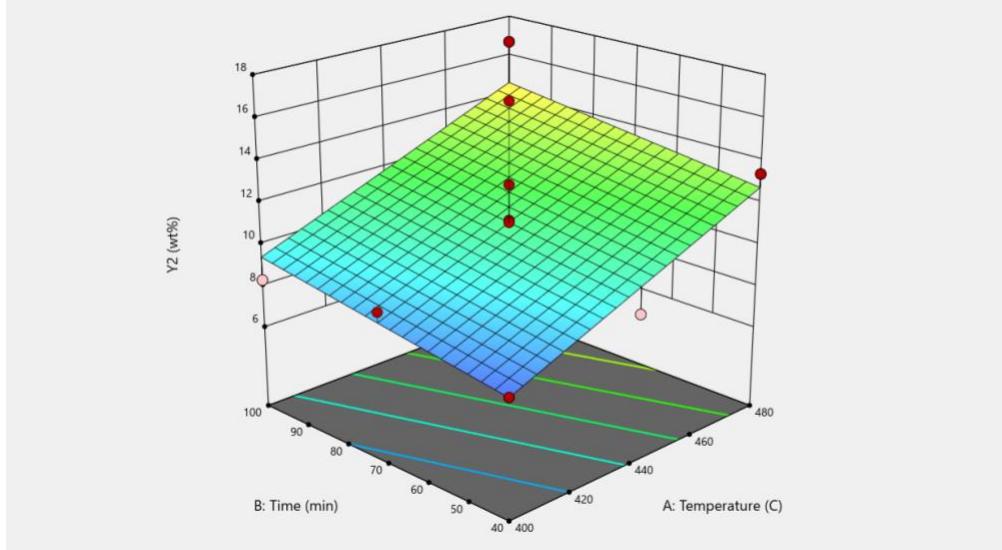
Source	Sum of Squares	df	Mean Square	F-value	p-value	
<b>Model</b>	139.88	3	46.63	13.91	< 0.0001	significant
A-Temperature	120.64	1	120.64	35.98	< 0.0001	
B-Time	15.12	1	15.12	4.51	0.0426	
C-Pressure	4.11	1	4.11	1.23	0.2777	
<b>Residual</b>	93.88	28	3.35			
Lack of Fit	67.25	23	2.92	0.5489	0.8518	not significant
Pure Error	26.63	5	5.33			
<b>Cor Total</b>	233.76	31				

The Model F-value of 13.91 implies the model is 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, B are significant model terms. Thus, the rest are removed to improve the model.

The equation obtained becomes,

$$\text{Kerosene Yield, } Y_2 = -17.13403 + 0.064722A + 0.030556B + 0.015926C$$

3D Model Graph:



Analysis of variance (ANOVA) for diesel yield:

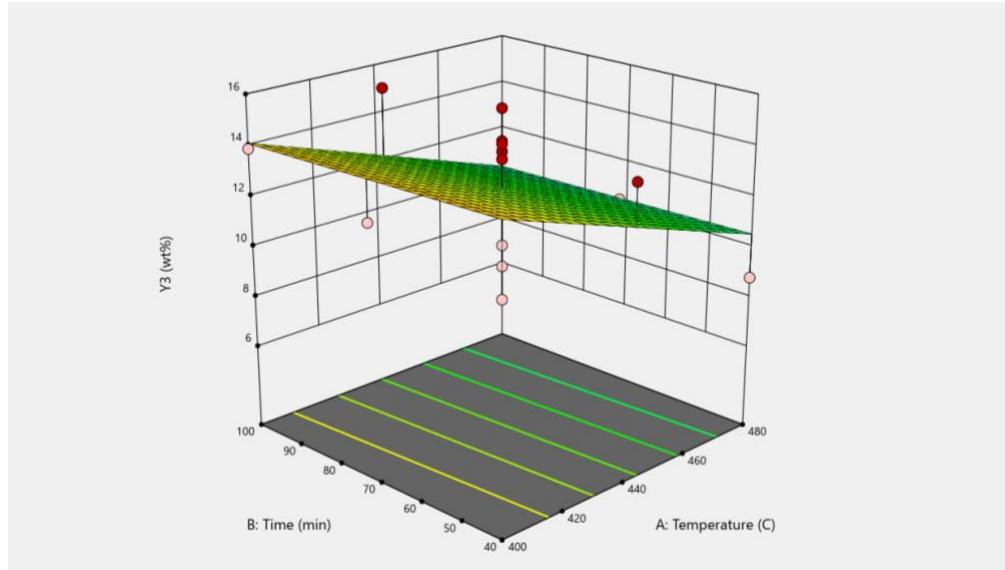
	Source	Sum of Squares	df	Mean Square	F-value	p-value	
	<b>Model</b>	64.05	3	21.35	3.83	0.0205	significant
	A-Temperature	63.84	1	63.84	11.45	0.0021	
	B-Time	0.2006	1	0.2006	0.0360	0.8510	
	C-Pressure	0.0006	1	0.0006	0.0001	0.9921	
	<b>Residual</b>	156.15	28	5.58			
	Lack of Fit	121.41	23	5.28	0.7599	0.7078	not significant
	Pure Error	34.73	5	6.95			
	<b>Cor Total</b>	220.19	31				

The Model F-value of 3.83 implies the model is significant. There is only a 2.05% 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 is a significant model term. Thus, the rest are removed to improve the model.

The equation obtained becomes,

$$\text{Diesel Yield, } Y_3 = 33.32512 - 0.047083A - 3.51852E-003B - 1.85185E-004C$$

3D Model Graph:



## Results

Predicted vs Actual Value

Conversion:	Yield of total distillate fuels:
Yield of gasoline:	Yield of kerosene:
Yield of diesel:	



It is seen from table above that the predicted values are close to the experimental values. Therefore, this model is accurate enough to predict the conversion and yields of total distillate fuels, gasoline, kerosene, and diesel within the range of temperature, time and pressure ranges.

There are 36 solutions obtained by maximizing the temperature, contact time and pressure. This is done because the conversion and yields are found to increase with increasing operating conditions. The optimum operating conditions for thermal cracking of petroleum residue oil with a desirability of 98% are,

	Temperature	Time	Pressure
	480	100	180

## Confirmation

Two-sided Confidence = 95%

	Solution 1 of 36 Response	Predicted Mean	Predicted Median	Observed	Std Dev	n	SE Pred	95% PI low	Data Mean	95% PI high
	X (wt%)	89.8781	89.8781		3.3359	1	4.08196	81.4127		98.3436
	Y (wt%)	45.3993	45.3993		2.71019	1	2.96628	39.3232		51.4755
	Y1 (wt%)	20.8104	20.8104		3.49452	1	3.82473	12.9758		28.645
	Y2 (wt%)	14.1215	14.1215		1.83111	1	2.00413	10.0162		18.2268
	Y3 (wt%)	10.3399	10.3399		2.36149	1	2.58464	5.04554		15.6343

## Justification

This model uses a 3-level factorial to find the optimal solution. This is because of the presence of a quadratic relationship between the factors and the conversion, X, which would result in curvature in the conversion. Additionally, this method works by checking the effect of change in one factor to the response in all levels. Therefore, it is much more applicable for this study as the changes in temperature, pressure and reaction time can be studies separately and simultaneously on the individual responses.

## 2.4 Box-Behnken Design – Parametric optimization by Box-Behnken design for synthesis of magnetic chitosan-benzil/ZnO/Fe<sub>3</sub>O<sub>4</sub> nanocomposite and textile dye removal (Reghioua et al., 2021) (Wong Pei Yu)

### Brief Background:

In this study, chitosan was modified with inorganic fillers and a crosslinking agent to form Cs-Bz/ZnO/Fe<sub>3</sub>O<sub>4</sub>, a novel material. Remazol Brilliant Blue R dye is aimed to remove from wastewater by this material through adsorption and removal rate and adsorption capacity of material is aimed to optimize. Adsorption is use due to its effective technique for removing pollutants due to its high selectivity, regeneration, low-cost, and designability. The Freundlich isotherm and a pseudo-second-order kinetic model were used to model the substance. Via a variety of physical and chemical interactions, including electrostatic and hydrogen bonding, the dye was eliminated. The material is reusable and environmentally acceptable, and it may be used to treat actual wastewater, get rid of heavy metal ions, and lower the demand for chemical oxygen. The study sought to address the requirement to

preserve ecosystems and eliminate pollutants from water and waste. Pollutant removal can be accomplished through adsorption utilising modified chitosan, and the material Cs-Bz/ZnO/Fe<sub>3</sub>O<sub>4</sub> nanocomposite is a promising choice for these uses.

### **Objective Function, Relevant Variables, Conditions, And Assumptions:**

#### **Objective Function:**

The decolorization of RBBR was analyzed and predicted using a polynomial model of second order, which is presented in the following equation.

$$\text{Maximize } Y = \beta_0 + \sum \beta_i X_i + \sum \beta_{ii} X_i^2 + \sum \sum \beta_{ij} X_i X_j$$

where Y and ( $X_i, X_j$ ) represent the response of RBBR dye removal (%) and coded independent factors,  $\beta_0$ : constant,  $\beta_i$ ,  $\beta_{ii}$  and  $\beta_{ij}$  is the linear, quadratic, and interactive coefficients of input independent variables respectively

#### **Variables:**

Adsorption inputs such as ZnO loading (%), adsorbent dose (g), solution pH, contact time (minutes), and temperature (°C) are the independent factors that affect the response of uptaking of RBBR dye (output).

#### **Conditions:**

Table 1: Experimental Levels of Independent Factors and Their Codes In BBD

Codes	Variables	Level 1 (-1)	Level 2 (0)	Level 3 (+1)
A	ZnO loading (%)	0	25	50
B	Adsorbent Dose (g)	0.02	0.04	0.06
C	pH	4	7	10
D	Temperature (°C)	30	45	60
E	Time (min)	5	10	15

### **Generation of Data Using Design Expert Simulation:**

Firstly, the information from table 1 is inserted into Response Surface BBD as shown in figure 1. The numeric factors are 5 and center points per block is 6.

#### **Box-Behnken Design**

Each numeric factor is set to 3 levels. If categoric factors are added, the Box-Behnken design will be duplicated for every combination of the categoric factor levels. These designs have fewer runs than 3-Level Factorials.

Numeric factors:  (3 to 21)  Horizontal  
 Categoric factors:  (0 to 10)  Vertical

	Name	Units	Low	High
A [Numeric]	ZnO Loading	%	0	50
B [Numeric]	Adsorbent dose	g	0.02	0.06
C [Numeric]	pH		4	10
D [Numeric]	Temperature	deg C	30	60
E [Numeric]	Time	min	5	15

Blocks:

Center points per block:  (0 to 1000) 46 Runs

Figure 1: Data Input in RSM BBD





than 0.20 is preferred. Since the difference in this model is only 0.0618, which is very small, it shows that the experimental data is fit with the model.

Figure 6a and 6b below shows the normal plot of residuals and predicted vs actual plot. From figure 6a, the points lie along a straight line and assume it is normally distributed. The normal distributions of the residuals signify the accuracy of the assumptions, and the independence of the residuals. From figure 6b it can be deduced that the predicted and actual points were mostly close to each other. This observation signifies that the experimental results for this research are highly acceptable.

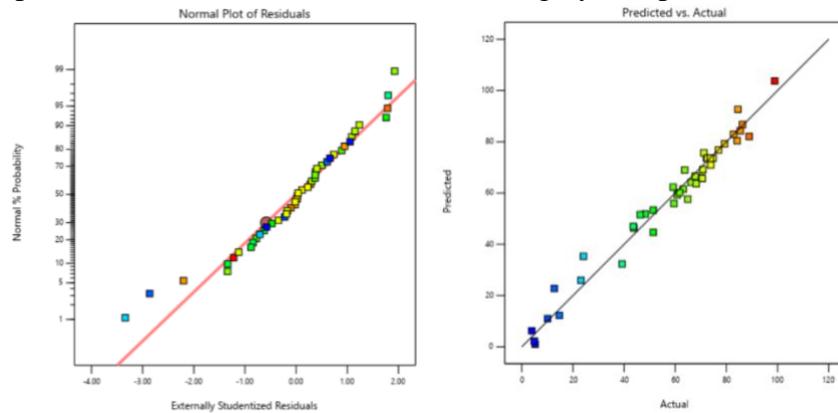


Figure 6a: Normal Plot of Residuals    Figure 6b: Predicted vs Actual

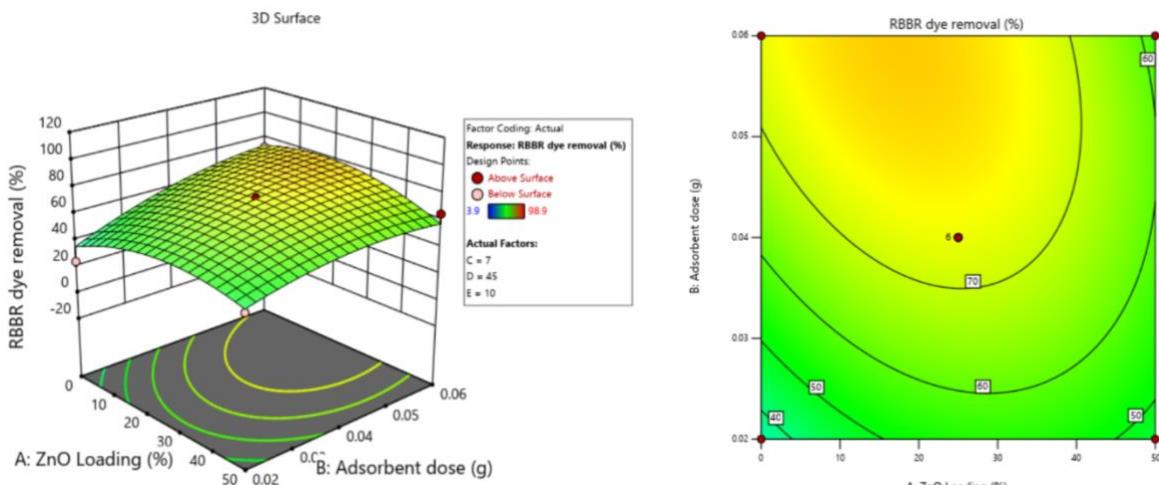


Figure 7a: 3D plot between AB on RBBR removal. Figure 7b: 2D plot between AB on RBBR removal

Surface and contour plots were generated to determine the importance of combined interactions among the selected factors and to optimize the adsorption behavior of RBBR dye.

The 3D response surface plots and 2D contour plot for the statistically significant interaction between ZnO loading (A) and adsorbent dose (B) are presented in Figures 7a and 7b, respectively. The p value of 0.0245 indicates their significant function. The parameters pH of 7, temperature of 45°C, and time of 10 minutes were kept constant. Increasing both ZnO loading and adsorbent dose improved the removal effectiveness of RBBR. This may be attributed to the homogeneous distribution of ZnO nanoparticles in the Cs-Bz/Fe<sub>3</sub>O<sub>4</sub> polymeric matrix, resulting in an increased surface area. A higher Cs-Bz/ZnO-25/ Fe<sub>3</sub>O<sub>4</sub> adsorbent dose will provide more available adsorption sites, thus capturing more RBBR dye molecules from the bulk solution.

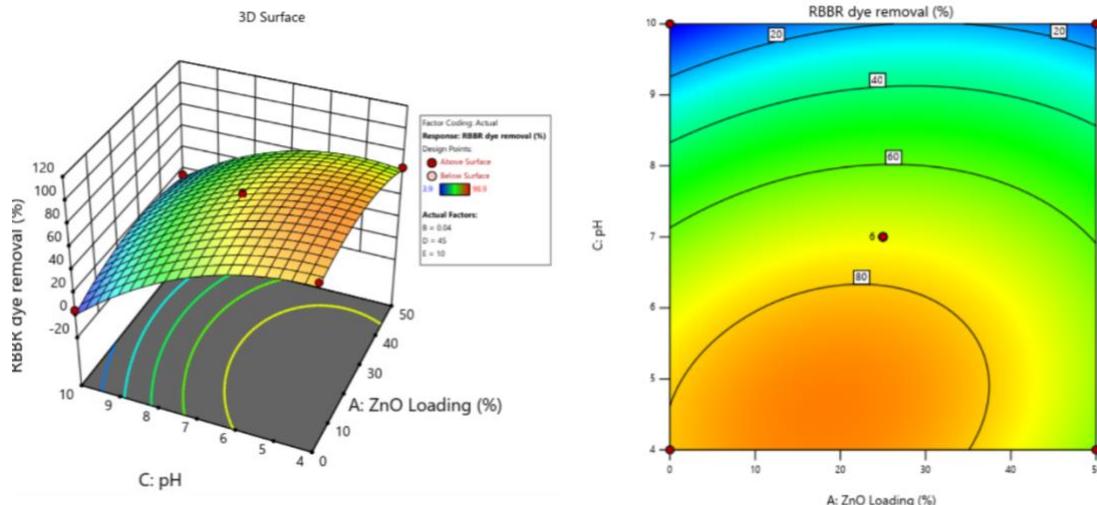


Figure 8a: 3D plot between AC on RBBR removal. Figure 8b: 2D plot between AC on RBBR removal

Figure 8a and Figure 8b illustrate the response surface plots (3D and 2D contour, respectively) for the parametric interaction between loading of ZnO nanoparticles (A) and working solution pH (C) as a statistically significant function ( $p$  value = 0.0222). The adsorbent dose of 0.04, temperature of 45 °C, and time of 10 minutes were constant. Adjusting the solution pH from pH 10 to pH 4 improved the RBBR removal efficiency from 3.9% to 98.9%.

#### **Justification of Selection of Box-Behnken Compared to Others:**

Compared to the Central Composite Design (CCD), the Box-Behnken Design (BBD) does not include extreme levels for all factors, making it suitable for experiments that aim to avoid operating under extreme conditions and finding the optimal operating conditions within a limited range of variables. This statistical experimental design method requires fewer experimental runs than other designs like CCD or 3-factorial design while achieving the same level of precision. Moreover, BBD allows for the evaluation of interactions between different variables, which provides valuable information for the optimization process. Additionally, it is an efficient response surface design that can optimize multiple variables simultaneously, making it a suitable method for both synthesizing the nanocomposite and removing textile dye

#### **2.5 Optimal Design (D-optimal) of Response Surface Methodology - Application on Optimizing Drug Release from Formulation of Acrylamide-based Hydrogels (Tham Ting Woon)**

##### **Background**

Acrylamide-based hydrogels are a class of hydrogels that widely used in biomedical and chemical industries. Hydrogels are formed from a cross-linked acrylamide polymer which are hydrophilic and can absorb and retain water. Hydrogels have potential application in drug delivery system because its unique properties such as high biocompatibility, tunable physical and chemical characteristics and high-water content (Sabbagh et al., 2018). These properties enable it to maintain its structure, making it suitable for continuous drug release. The formulation of acrylamide-based hydrogels process involves the selection of monomers, cross-linked agents, and initiators (Sabbagh et al., 2018). It can be synthesized through free radical polymerization and thermal polymerization.

##### **Objective Function, Variables, Conditions and Assumptions**

After the optimization of acrylamide-based hydrogels formulation, response surface methodology (RSM) had been used to further optimize the drug release by adjusting the concentration of raw materials, temperature and pH. The variables should be adjusted carefully to develop a safe and effective drug delivery system.

The objective function is to study release the drug in suitable condition through response surface methodology (RSM)-Optimal Design. The variables are MgO (A), NaCMC(B), temperature (C), pH (D) (Sabbagha et al., 2018). The ranges and level are listed in Table x below.

**Table x: Ranges and Condition of Variables**

Variables	Factors	Low Level	High Level
A	MgO	0.01	0.02
B	NaCMC	0.1	0.2
C	Temperature	37	38
D	pH	3.50	4.7

The assumption equation generated in this scenario is  $91 - 17*A + 3*B - 6*C + 5*D + 5*B*D + 4*C*D + 1*D^2$ . This equation is generated by Design Expert Software.

### Generation of Data by using Software Simulation

#### Optimal (Custom) Design

A flexible design structure to accommodate custom models, categoric factors, and irregular (constrained) regions. Runs are determined by a selection criterion chosen during the build.

Numeric factors: 4 (1 to 30)  Horizontal

Categoric factors: 0 (0 to 10)  Vertical

	Name	Units	Type	Levels	L[1]	L[2]
A [Numeric]	MgO	g	Continuous	N/A	0.01	0.02
B [Numeric]	NaCMC	g	Continuous	N/A	0.1	0.2
C [Numeric]	Temperature	degree C	Continuous	N/A	37	38
D [Numeric]	pH	pH	Continuous	N/A	3.5	4.7

[Edit constraints...](#)

#### Figure x: The Details of Factors in Design Expert

There are 4 factors in this case which are the concentration of MgO and NaCMC, temperature and pH. The mixing of MgO (g) and NaCMC (g) to prepare the optimum state of hydrogel polymer. The range of MgO is between 0.01 to 0.02 while NaCMC is from 0.1 to 0.2. The temperature is in between 37 °C which is a safe condition to 38 °C which is a disease condition (Sabbagha et al., 2018). The pH is in between 3.5 to 4.7.

#### Optimal (Custom) Design

Search: Both Exchanges	Optimality: I
<a href="#">Edit model...</a>	Quadratic
Blocks: 1 (1 to 1000)	Runs
	Required model points: 15
	Additional model points: 0
	Lack-of-fit points: 5
	Replicate points: 5
	Additional center points: 0
	Total runs: 25

Both Exchanges will try both Point Exchange and Coordinate Exchange searches of the design space. This could result in some unusual combinations of factors. If you require certain candidates or combinations of factors, switch to Point Exchange.

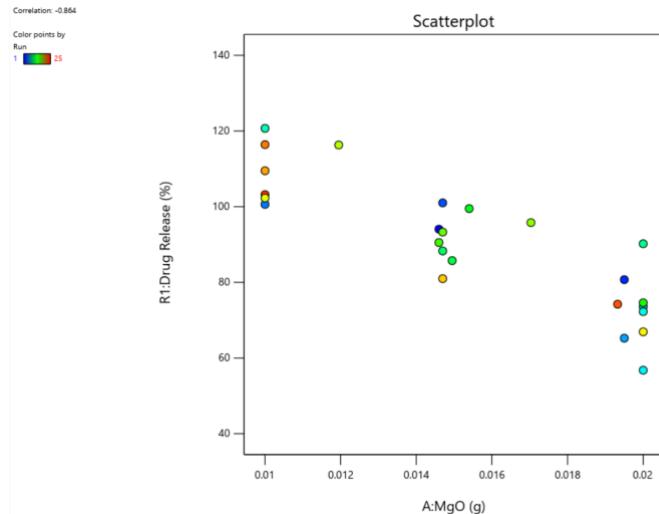
I-optimal designs (also called IV or Integrated Variance) provide lower average prediction variance across your region of experimentation. I-optimality is desirable for response surface methods (RSM) where prediction is important. The algorithm picks points that minimize the integral of the prediction variance across the design space.

#### Figure x: Run is Set by Default

The mathematic model used in this case is quadratic mathematical mode. This model shows the relationship between factor A,B,C,D and Response. There are total of 25 runs generated by Design Expert.

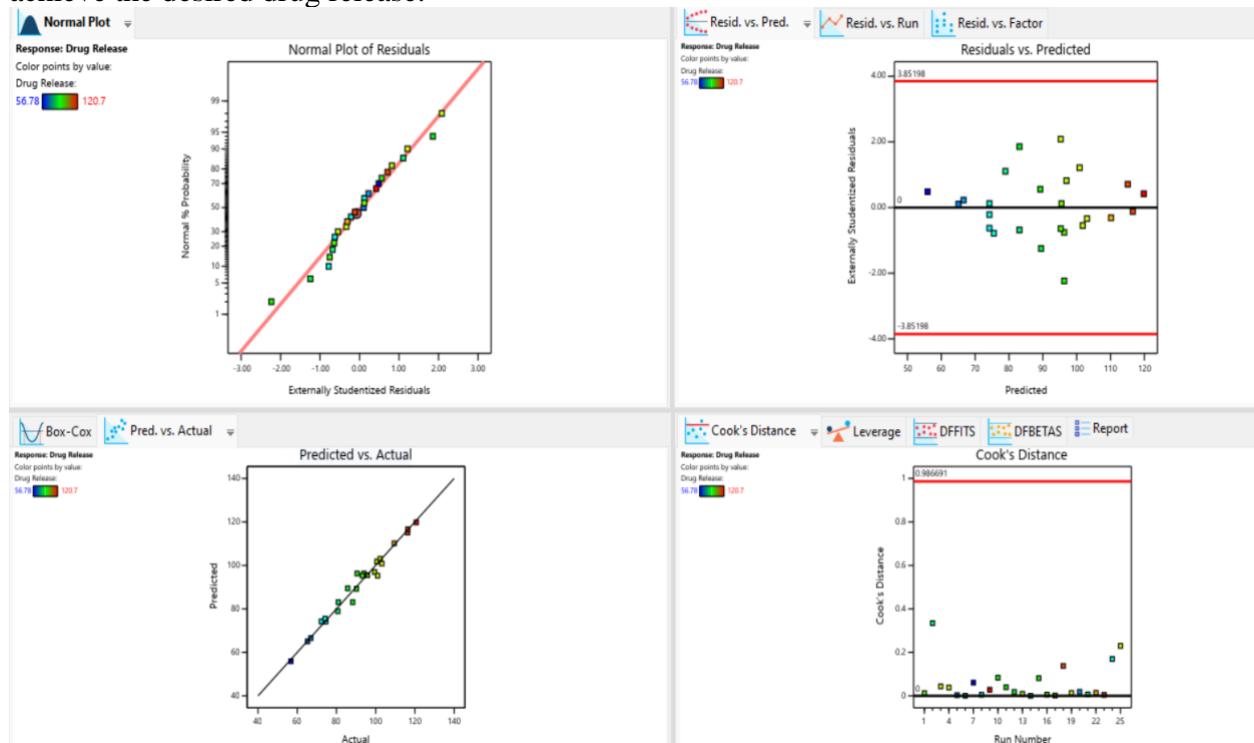






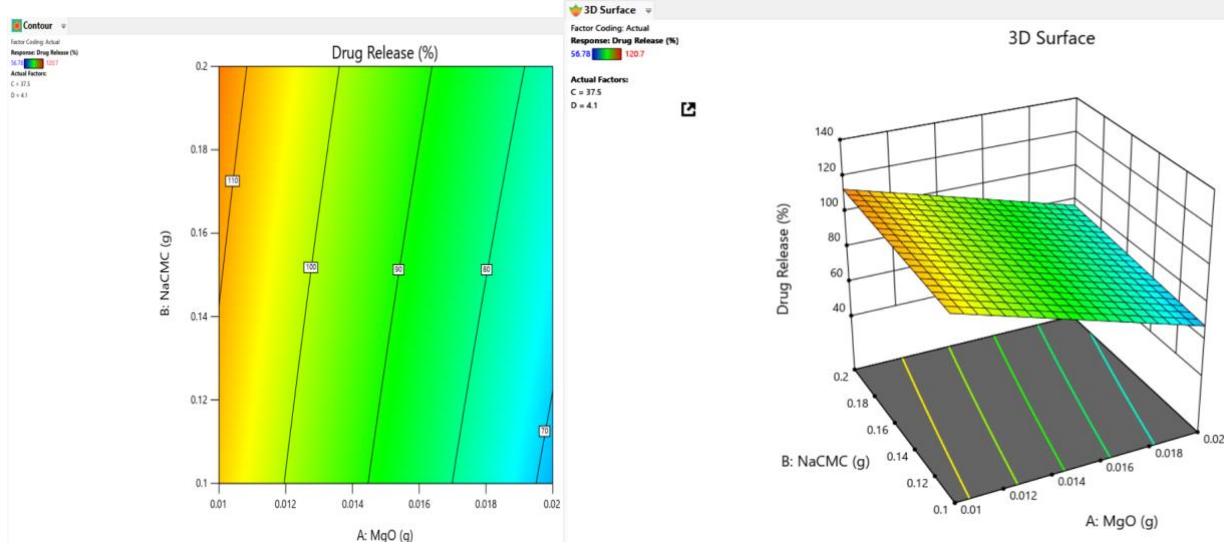
**Figure x: Scatterplot graph of Drug Release against MgO (A)**

From Figure x, the scatterplot show the relationship between drug release and concentration of MgO. Drug release is inversely proportional to the concentration of MgO. In this case, drug release need to be maximise. In order to maximise the drug release, the concentration of MgO should be minimise to achieve the desired drug release.



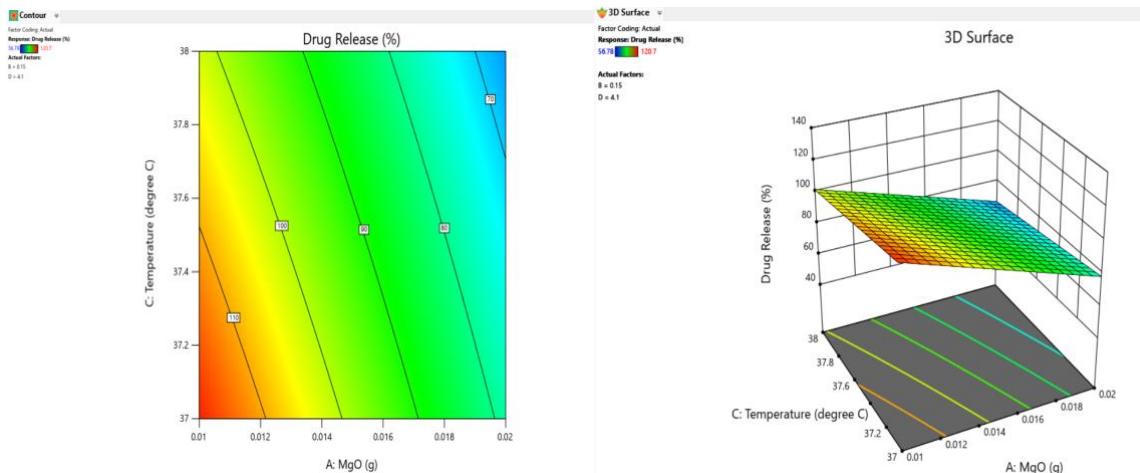
**Figure x: Details under Diagnostics**

Diagnostic plot was observed in order to ensure the assumption is fit with the ANOVA table. The suitability of the model was checked through different diagnostic plots such as normal plot of residuals, Residuals vs Predicted plot, Predicted vs Actual plot and Cook's Distance plot. In the normal plot of residuals graph, most of the points are lie along the line. This indicates that there is no large deviation of variance occurred. In Residuals vs Predicted plot, all the points are located inside the red line. This model is adequate as there is no outlier points. In the Predicted vs Actual plot, all the points are lie within the line. This indicated that there is no much different between the actual results from experiment and the model equation predicted by Design Expert.



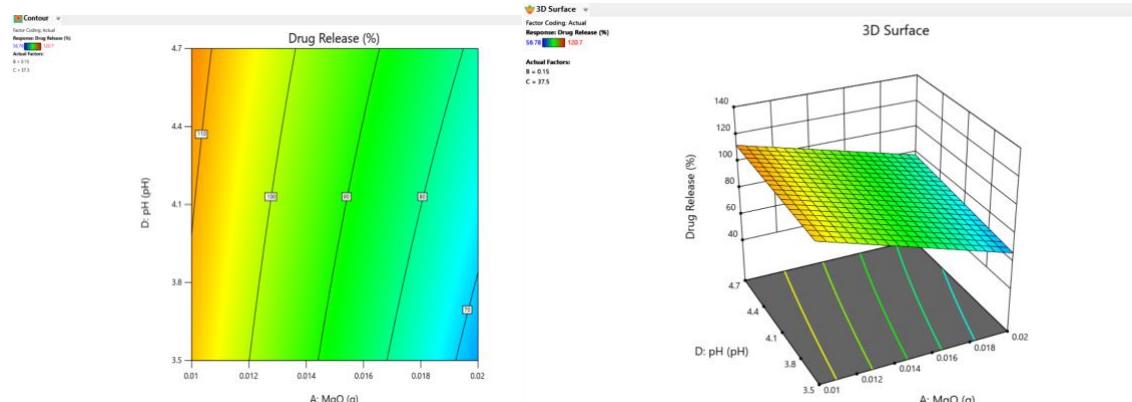
**Figure x: Contour Graph and 3D Surface Graph of MgO(A) and NaCMC(B)**

From Figure x above, this show the relationship of MgO and NaCMC to optimize drug release during the formulation of acrylamide-based hydrogel in the condition of fixed temperature at 37.5 °C and pH at 4.1.



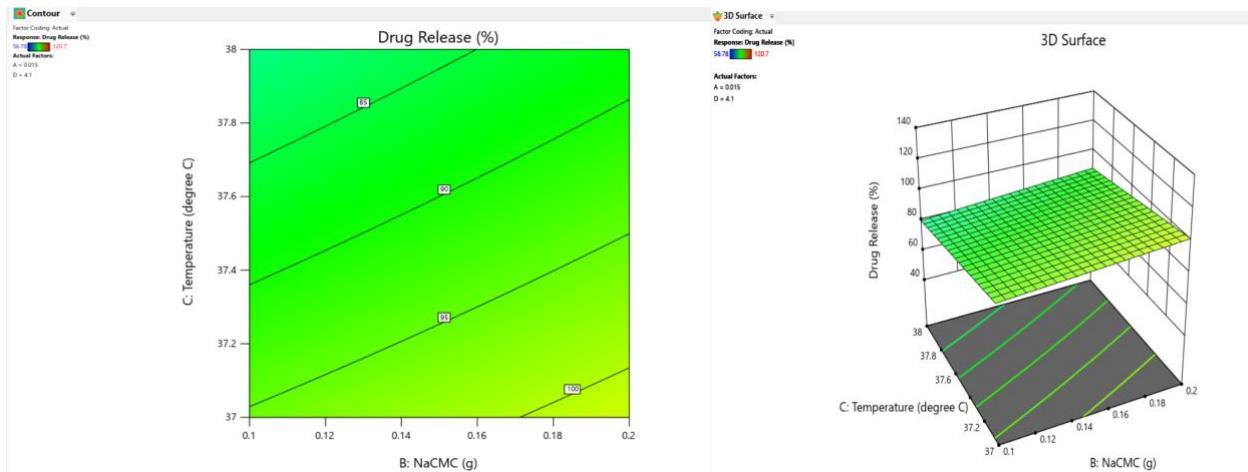
**Figure x: Contour Graph and 3D Surface Graph of MgO(A) and Temperature (C)**

From Figure x above, this show the relationship of MgO and temperature to optimize drug release during the formulation of acrylamide-based hydrogel in the condition of NaCMC at 0.15 g and pH at 4.1.



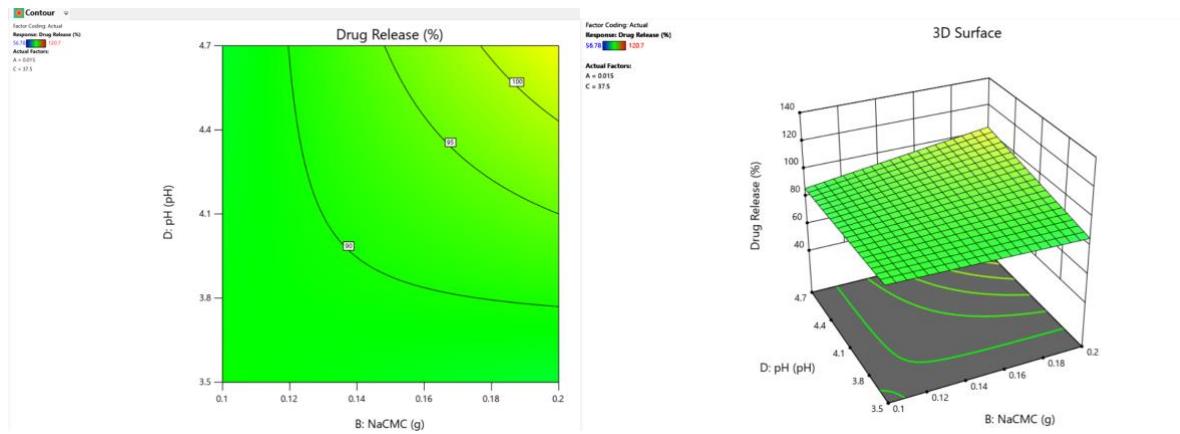
**Figure x: Contour Graph and 3D Surface Graph of MgO(A) and pH(D)**

From Figure x above, this show the relationship of MgO and pH to optimize drug release during the formulation of acrylamide-based hydrogel in the condition of NaCMC at 0.15 g and temperature at 37.5 °C.



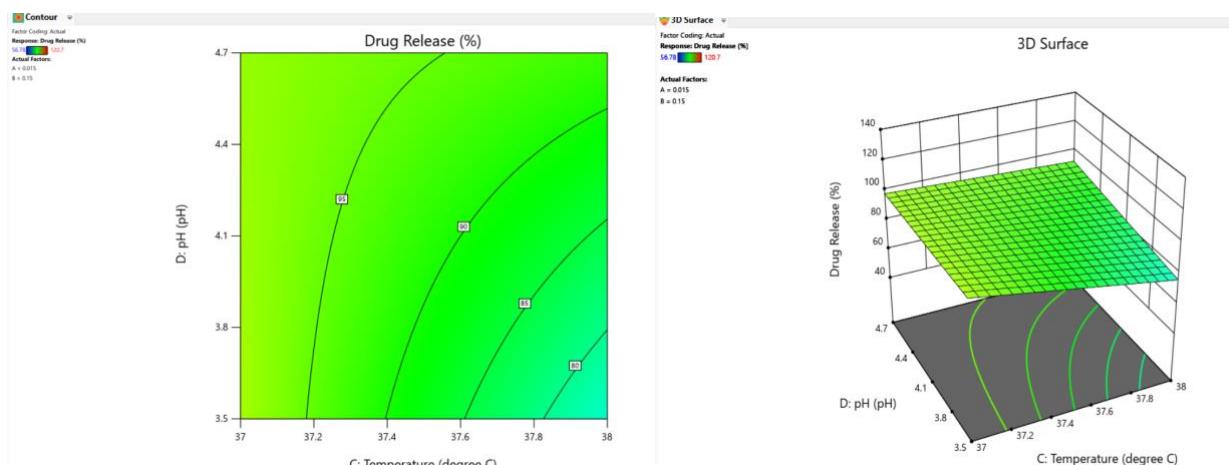
**Figure x: Contour Graph and 3D Surface Graph of NaCMC (B) and Temperature(C)**

From Figure x above, this show the relationship of NaCMC and temperature to optimize drug release during the formulation of acrylamide-based hydrogel in the condition of MgO at 0.015 g and pH at 4.1.



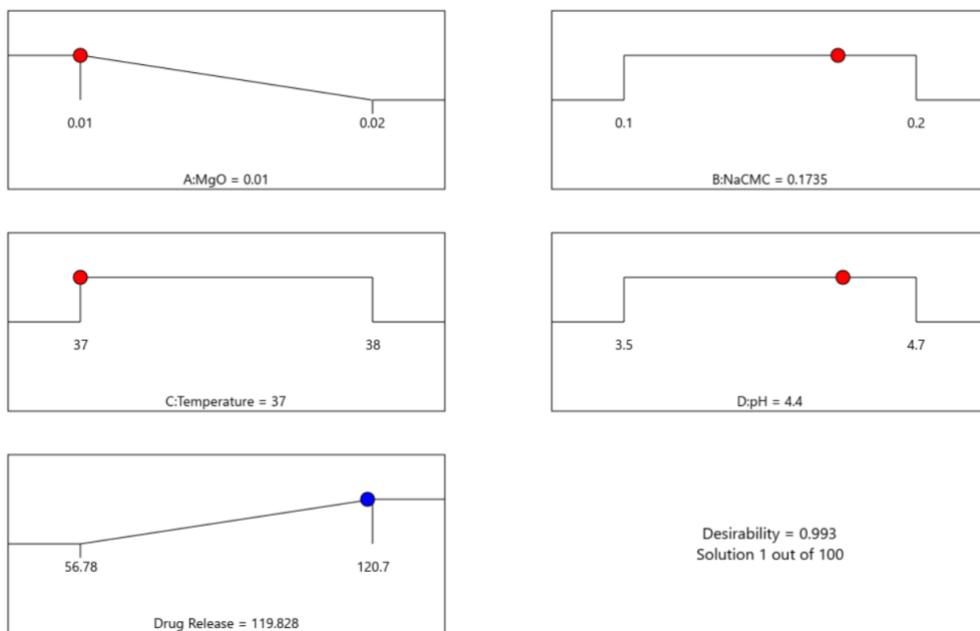
**Figure x: Contour Graph and 3D Surface Graph of NaCMC (B) and pH(D)**

From Figure x above, this show the relationship of NaCMC and pH to optimize drug release during the formulation of acrylamide-based hydrogel in the condition of MgO at 0.015 g and temperature 37.5 °C.



**Figure x: Contour Graph and 3D Surface Graph of Temperature (C) and pH(D)**

From Figure x above, this show the relationship of temperature and pH in the formulation of acrylamide-based hydrogel in the condition of MgO at 0.015 g and NaCMC at 0.15 g.



**Figure x: Optimum Solution in Minimizing the Concentration of MgO and Maximizing the Drug Release.**

From Figure x, the goal of the concentration of MgO is set to be minimize. The goal for concentration of NaCMC, temperature and time is set in range while the goal for drug release is set to be maximise. All the importance are set into ++++. According to Figure x above, the scatterplot of MgO and drug release show that the relationship between them are inversely proportional. Besides, minimizing the concentration of MgO can help in reducing the cost of the process. There are total of 100 solutions generated by Design Expert. Solution 1 is chosen as the best because it has a highest desirability which are 0.993 which close to 1. This design is consider valid. From the above solution, we can obtain the concentration of MgO is 0.01 g, concentration of NaCMC is 0.1735 g, the temperature is 37 °C, pH is at 4.4 and the drug release is 119.828 %.

#### **Selection of Optimal Design compared to others**

Optimal design is a systematic and efficient design to identify the condition of a system. The advantages of optimal design are it can save time and can accommodate custom models compared to other types of design. This can make optimal design more appropriate for complex system with multiple variables. The number of factors and responses can be adjusted freely according to the condition. Besides, it only require few experiment runs compared to other designs to achieve the same level of precision which can save time. Greater efficiency can be achieved in optimal design. This is because optimal design allow a more thorough exploration of the design space and lead to more accurate result. The disadvantages of using optimal design are time and resources constraints. This is because it is significant to collect and analyse data especially complex system with multiple variables while optimal design only requires few experiment run. All the data and condition need to decide properly to achieve a high accuracy of result. However, this can be solved by doing more researches to come out with the best solution. Optimal design can use to solve complex system with multiple variables, so it is important to have professional knowledge and skills to optimise. Overall, optimal design is selected because it has a greater efficiency, more flexible on design, save time and suitable in complex system with multiple variables.

#### **REFERENCES**

- Alsobaa, A.M. (2011) *Thermal cracking of petroleum residue oil using three level factorial design*, *Science Direct* . Journal of King Saud University - Engineering Sciences. Available at: <https://www.sciencedirect.com/science/article/pii/S1018363911000602> (Accessed: April 1, 2023).

Dao, T. et al. 2021. Central Composite Design, Kinetic Model, Thermodynamics, and Chemical Composition of Pomelo (*Citrus Maxima* (Burm.) Merr.) Essential Oil Extraction by Steam Distillation. *Processes*, 9(11), p.2075.

Reghioua, A., Barkat, D., Jawad, A.H., Abdulhameed, A.S., Al-Kahtani, A.A. and ALOthman, Z.A., 2021. Parametric optimization by Box–Behnken design for synthesis of magnetic chitosan-benzil/ZnO/Fe<sub>3</sub>O<sub>4</sub> nanocomposite and textile dye removal. *Journal of Environmental Chemical Engineering*, [online] 9(3), p.105166. Available at: <https://www.sciencedirect.com/science/article/pii/S2213343721001433> [Accessed 9 Apr. 2023].

Sabbagh, F. et al. (2018) *From formulation of acrylamide-based hydrogels to their optimization for drug release using response surface methodology*, C. Elsevier. Available at: <https://www.sciencedirect.com/science/article/pii/S0928493117340717> (Accessed: April 5, 2023).