Explanation of datasets

Sheet 1: Effect of emulsifier concentration on the emulsion stability index (ESI)

Table 1 to **table 4** on sheet 1 shows the emulsion stability index (ESI) and the emulsions average droplet size at various emulsifier concentrations (1, 1.5, 2, and 2.5%) and oil ratios, the homogenization speed, the water salinity, and the time was kept constant at 10000rpm, 1000ppm of NaCl and 5 minutes. The experiment was repeated 3 times and the average ESI was calculated using equation 1.

The graph emulsions stability index vs. the emulsifier concentration, and that for the average droplet size vs. the emulsifier concentration for various oil ratios (15, 25, 35, and 45% diesel to water ratio) were potted as shown in **Figure 1** and **Figure 2** of sheet1. It was found that the ESI increased as the emulsifier concentration increased, and the average droplet size decrease as the emulsifier concentration increase. There is a direct proportionality between the ESI and the average droplet size of the emulsions. The highest ESI was found at a high emulsifier concentration (2%).

Sheet 2: Effect of homogenization speed on emulsions stability index (ESI)

Table 1 to **table 4** on sheet 2 represent the emulsions stability index and the average droplet size at various homogenization speeds (10000, 14000, 19000, and 24000 rpm) and oil ratios, the emulsifier concentration, the water salinity, and the homogenization time were kept constant at 2%, 1000ppm NaCl and 5 minutes. The experiments were repeated 3 times and the average ESI and average droplet size were plotted against the homogenization speed. **Figure 1** and **Figure 2** on sheet 2 show that the ESI increase as the homogenization speed increased for all oil rations (15, 25, 35, and 45% diesel to water ratio).

Sheet 3: Effect of Salinity on the emulsions stability index (ESI)

Table 1 to Table 4 on sheet 3 represents the emulsion stability index and the average droplet size (avg DSD) at various water salinity (1000, 2000, 3000, 4000ppm) and various oil ratios (15, 25, 35, and 45% diesel to water ratio). The emulsifier concentration, the homogenization speed, and the time were kept constant at 2%, 24000rpm, and 5 minutes. 3 experimental repeat was done and the average droplet size and ESI were plotted against the water salinity. **Figure 1** and **Figure 2** of sheet 3 show that the ESI decreases and the droplet size increase as the water salinity increases.

Sheet 4: Effect of oil content on the emulsions stability index (ESI)

Table 1 in sheet 4 shows the emulsion stability index and the average droplet size distribution at various oil ratios (15, 25, 35, and 45% diesel to water ratio). The emulsifier concentration, homogenization speed and water salinity were kept constant at 2%, 24000rpm, and 1000ppm NaCl.

Sheet 5: Effect of droplet size distribution on the emulsion stability index (ESI)

Table 1 to **Table** 3, **Figure** 1 to **Figure 3** in sheet 6 show the effect of the emulsifier concentration, homogenization speed, salinity and the oil content on the average droplets size distribution of the emulsions.

Sheet 6: Design of experiment (DoE) for the demulsification of crude oil-in-waster emulsions

Table: 1 and **Table: 2** show the amount of oil separated at different settling time (2 to 12h), different demulsifier concentration (350 to 930ppm) and different oil ration (15, 25, 35, and 45% diesel to water ratio) with cetyl-trimethylammonium bromide (CTAB), and trimethyl-tetradecyl ammonium chloride (TTAC).

ANOVA results for the demulfication of crude oil-in-water emulsions

ANOVA for Reduced Quadratic model

Response 1: Demul. Efficiency (CTAB)

Source	**Sum of Squares**	**df**	**Mean Square**	**F-value**	**p-value**	
Model	13484.67	17	793.22	22.83	< 0.0001	significant
A-Time	2858.66	1	2858.66	82.28	< 0.0001	
B-Concentration	7869.37	1	7869.37	226.51	< 0.0001	
C-Oil content	812.23	3	270.74	7.79	0.0004	
AB	112.82	1	112.82	3.25	0.0804	
AC	279.74	3	93.25	2.68	0.0621	
BC	601.81	3	200.60	5.77	0.0026	
A^2	427.48	1	427.48	12.30	0.0013	
B^2	39.08	1	39.08	1.13	0.2963	
B^2C	460.61	3	153.54	4.42	0.0100	
Residual	1181.21	34	34.74			
Lack of Fit	651.70	18	36.21	1.09	0.4316	not significant
Pure Error	529.51	16	33.09			
Cor Total	14665.88	51				

Response 2: Demul. Efficiency (TTAC)

Source	Sum of Squares	df	Mean Square	F-value	p-value	
Model	23082.61	14	1648.76	16.64	< 0.0001	significant
A-Time	3930.01	1	3930.01	39.66	< 0.0001	
B-D.Con	14298.41	1	14298.41	144.30	< 0.0001	
C-Oil content	2926.25	3	975.42	9.84	< 0.0001	
AB	672.91	1	672.91	6.79	0.0131	
AC	285.15	3	95.05	0.9592	0.4222	
BC	419.45	3	139.82	1.41	0.2549	
A^2	507.91	1	507.91	5.13	0.0295	
B^2	28.26	1	28.26	0.2852	0.5965	
Residual	3666.26	37	99.09			
Lack of Fit	2565.90	21	122.19	1.78	0.1222	not significant
Pure Error	1100.35	16	68.77			
Cor Total	26748.86	51				

The **Model F-value** of 22.83 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, C, BC, A^2, B^2C are significant model terms. Values greater than 0.1000 indicate the model terms are not significant. If there are many insignificant model terms (not counting those required to support hierarchy), model reduction may improve your model.

The **Lack of Fit F-value** of 1.09 implies the Lack of Fit is not significant relative to the pure error. There is a 43.16% chance that a Lack of Fit F-value this large could occur due to noise. Non-significant lack of fit is good -- we want the model to fit.

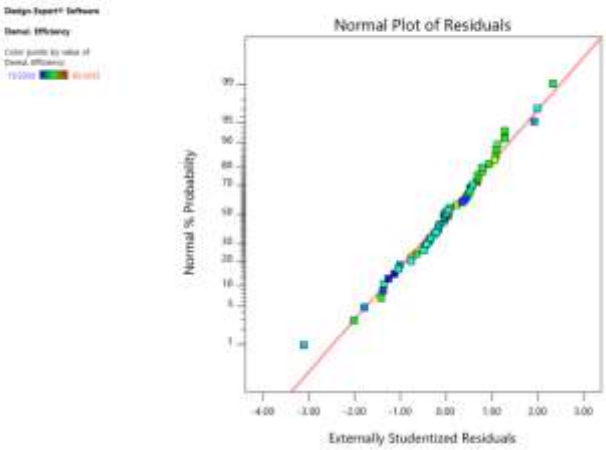
Fit Statistics (CTAB)

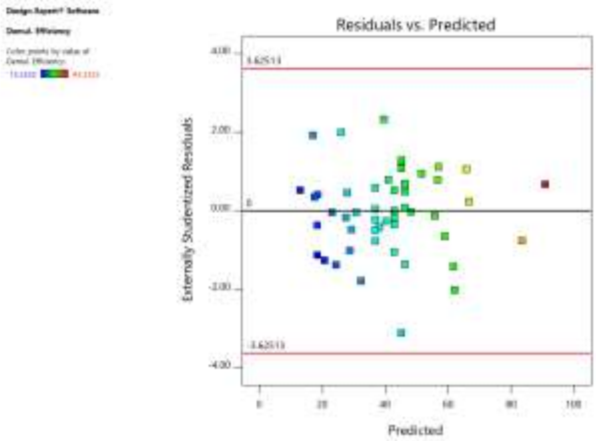
Std. Dev.	5.89	**R^2**	0.9195
Mean	41.23	**Adjusted R^2**	0.8792
C.V. %	14.30	**Predicted R^2**	0.7906
		Adeq Precision	22.4748

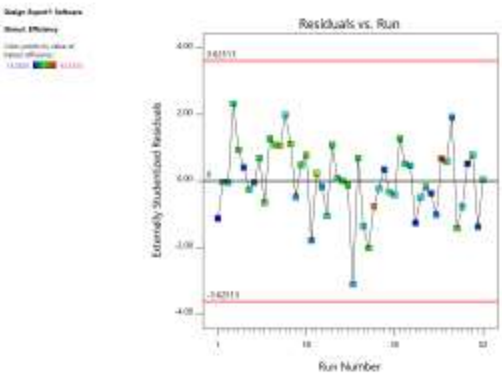
Std. Dev.	9.95	**R^2**	0.8629
Mean	47.34	**Adjusted R^2**	0.8111
C.V. %	21.03	**Predicted R^2**	0.6807
		Adeq Precision	15.9289

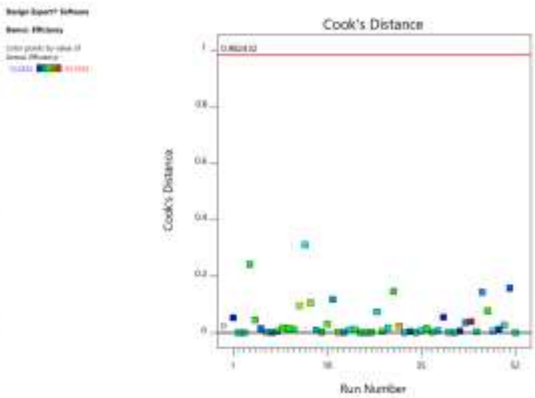
The **Predicted R^2** of 0.7906 is in reasonable agreement with the **Adjusted R^2** of 0.8792; i.e. the difference is less than 0.2.

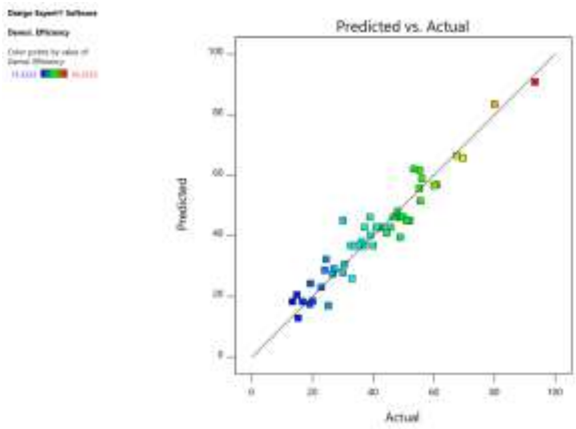
Adeq Precision measures the signal to noise ratio. A ratio greater than 4 is desirable. Your ratio of 22.475 indicates an adequate signal. This model can be used to navigate the design space.

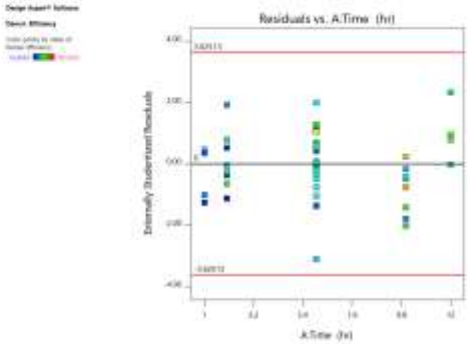


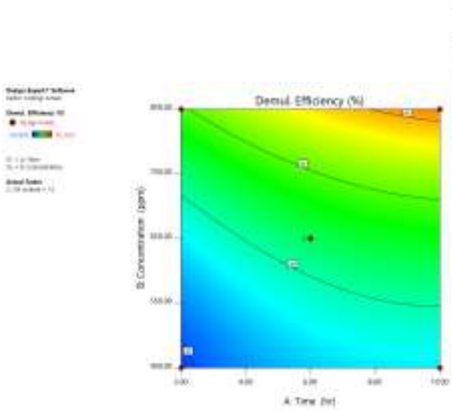


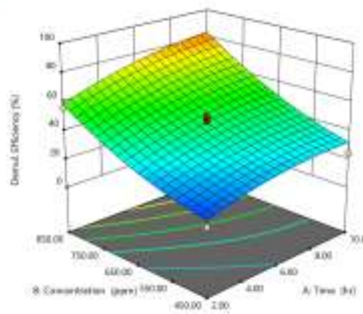


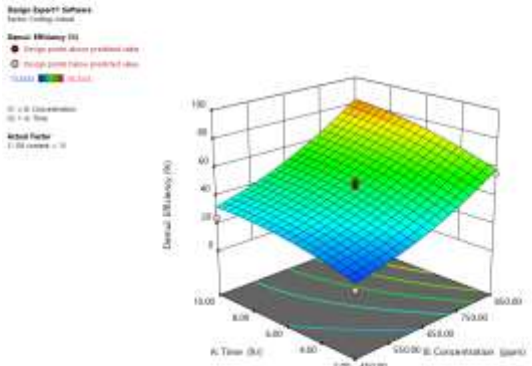


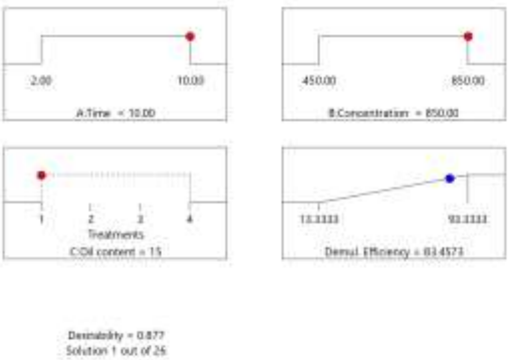

Design-Expert® Software
Predicted vs. Actual
Predicted
Actual
0
20
40
60
80
100

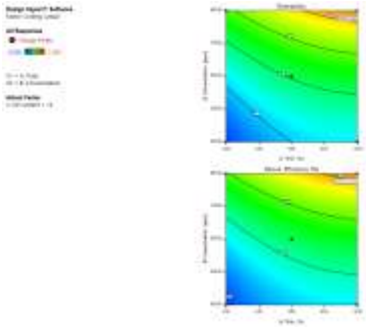

Residuals vs. A:Time (hr)
Externally Studentized Residuals
A:Time (hr)
4.00
2.00
0.00
-2.00
-4.00

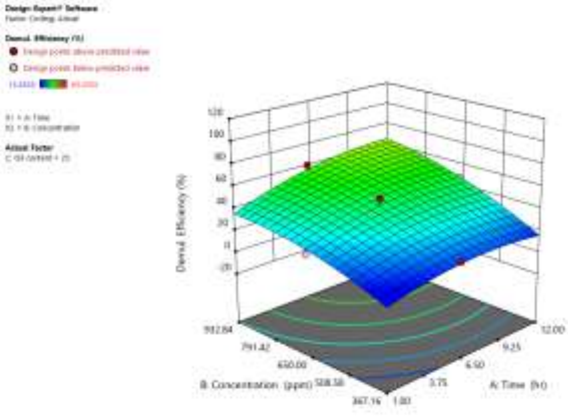

Demul. Efficiency (%)
B: Concentration (ppm)
A: Time (hr)

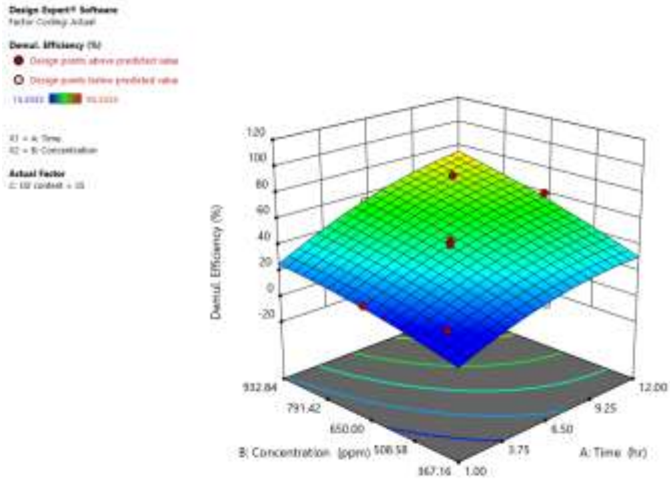

Design-Expert® Software
Demul. Efficiency (%)
850.00
750.00
650.00
550.00
450.00
B: Concentration (ppm)
10.00
8.00
6.00
4.00
2.00
A: Time (hr)

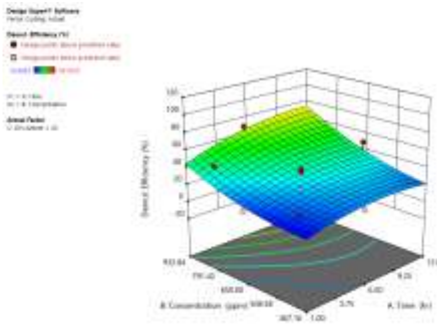

Design-Expert® Software
Demul. Efficiency (%)
10.00
8.00
6.00
4.00
2.00
A: Time (hr)
850.00
750.00
650.00
550.00
450.00
B: Concentration (ppm)


2.00
10.00
A:Time = 10.00
450.00
850.00
B:Concentration = 850.00
1
2
3
4
Treatments
C:Oil content = 15
13.3333
93.3333
Demul. Efficiency = 83.4573
Desirability = 0.877
Solution 1 out of 26




Design-Expert® Software
Factor Coding: Actual
Demul. Efficiency (%)
Design points above predicted value
Design points below predicted value
X1 = A: Time
X2 = B: Concentration
Actual Factor
Demul. Efficiency (%)
B: Concentration (ppm)
A: Time (hr)


Design-Expert® Software
Factor Coding: Actual
Demul. Efficiency (%)
Design points above predicted value
Design points below predicted value
X1 = A: Time
X2 = B: Concentration
Actual Factor
120
100
80
60
40
20
0
-20
Demul. Efficiency (%)
932.84
791.42
650.00
508.58
367.16
B: Concentration (ppm)
12.00
9.25
6.50
3.75
1.00
A: Time (hr)


Design-Expert® Software
Demul. Efficiency (%)
B: Concentration (ppm)
A: Time (hr)