

Chapter 6

The 2^k Factorial Design

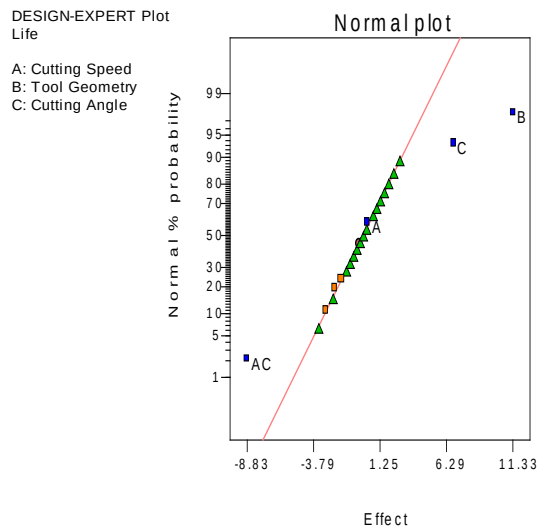
Solutions

6.1. An engineer is interested in the effects of cutting speed (A), tool geometry (B), and cutting angle on the life (in hours) of a machine tool. Two levels of each factor are chosen, and three replicates of a 2^3 factorial design are run. The results are as follows:

A	B	C	Treatment Combination	Replicate		
				I	II	III
-	-	-	(1)	22	31	25
+	-	-	a	32	43	29
-	+	-	b	35	34	50
+	+	-	ab	55	47	46
-	-	+	c	44	45	38
+	-	+	ac	40	37	36
-	+	+	bc	60	50	54
+	+	+	abc	39	41	47

(a) Estimate the factor effects. Which effects appear to be large?

From the normal probability plot of effects below, factors B , C , and the AC interaction appear to be significant.



(b) Use the analysis of variance to confirm your conclusions for part (a).

The analysis of variance confirms the significance of factors B , C , and the AC interaction.

Design Expert Output

Response: Life in hours					
ANOVA for Selected Factorial Model					
Analysis of variance table [Partial sum of squares]					
Source	Sum of Squares	DF	Mean Square	F Value	Prob > F
Model	1612.67	7	230.38	7.64	0.0004
A	0.67	1	0.67	0.022	0.8837
B	770.67	1	770.67	25.55	0.0001
C	280.17	1	280.17	9.29	0.0077
AB	16.67	1	16.67	0.55	0.4681
AC	468.17	1	468.17	15.52	0.0012
BC	48.17	1	48.17	1.60	0.2245
ABC	28.17	1	28.17	0.93	0.3483
Pure Error	482.67	16	30.17		
Cor Total	2095.33	23			

The Model F-value of 7.64 implies the model is significant. There is only a 0.04% chance that a "Model F-Value" this large could occur due to noise.

The reduced model ANOVA is shown below. Factor A was included to maintain hierarchy.

Design Expert Output

Response: Life in hours					
ANOVA for Selected Factorial Model					
Analysis of variance table [Partial sum of squares]					
Source	Sum of Squares	DF	Mean Square	F Value	Prob > F
Model	1519.67	4	379.92	12.54	< 0.0001
A	0.67	1	0.67	0.022	0.8836
B	770.67	1	770.67	25.44	< 0.0001
C	280.17	1	280.17	9.25	0.0067
AC	468.17	1	468.17	15.45	0.0009
Residual	575.67	19	30.30		
Lack of Fit	93.00	3	31.00	1.03	0.4067
Pure Error	482.67	16	30.17		
Cor Total	2095.33	23			

The Model F-value of 12.54 implies the model is significant. There is only a 0.01% chance that a "Model F-Value" this large could occur due to noise.

Effects B, C and AC are significant at 1%.

(c) Write down a regression model for predicting tool life (in hours) based on the results of this experiment.

$$y_{ijk} = 40.8333 + 0.1667x_A + 5.6667x_B + 3.4167x_C + 4.4167x_Ax_C$$

Design Expert Output

Factor	Coefficient Estimate	DF	Standard Error	95% CI Low	95% CI High	VIF
Intercept	40.83	1	1.12	38.48	43.19	
A-Cutting Speed	0.17	1	1.12	-2.19	2.52	1.00
B-Tool Geometry	5.67	1	1.12	3.31	8.02	1.00
C-Cutting Angle	3.42	1	1.12	1.06	5.77	1.00
AC	-4.42	1	1.12	-6.77	-2.06	1.00

Final Equation in Terms of Coded Factors:

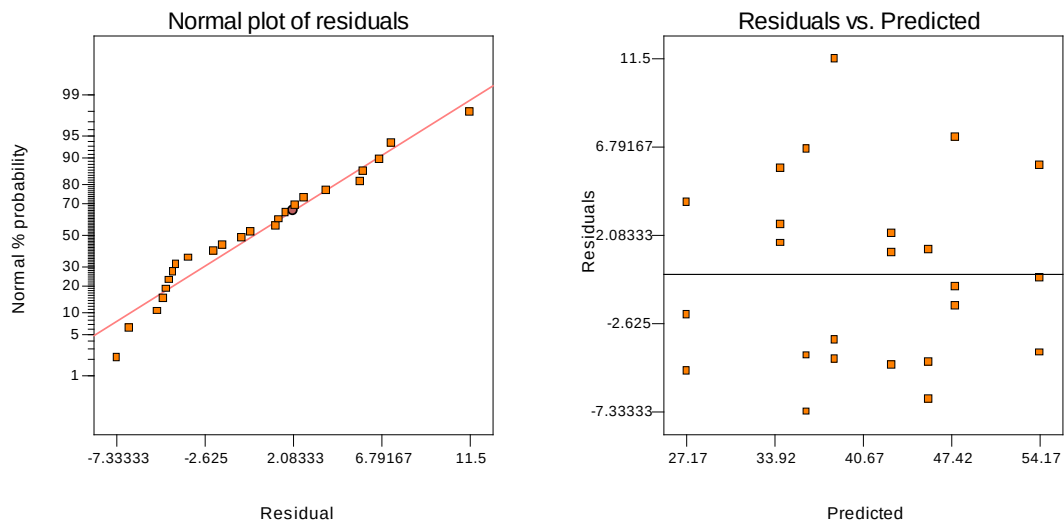
Life =
+40.83
+0.17 * A
+5.67 * B
+3.42 * C
-4.42 * A * C

Final Equation in Terms of Actual Factors:

Life	=
+40.83333	
+0.16667	* Cutting Speed
+5.66667	* Tool Geometry
+3.41667	* Cutting Angle
-4.41667	* Cutting Speed * Cutting Angle

The equation in part (c) and in the given in the computer output form a “hierarchical” model, that is, if an interaction is included in the model, then all of the main effects referenced in the interaction are also included in the model.

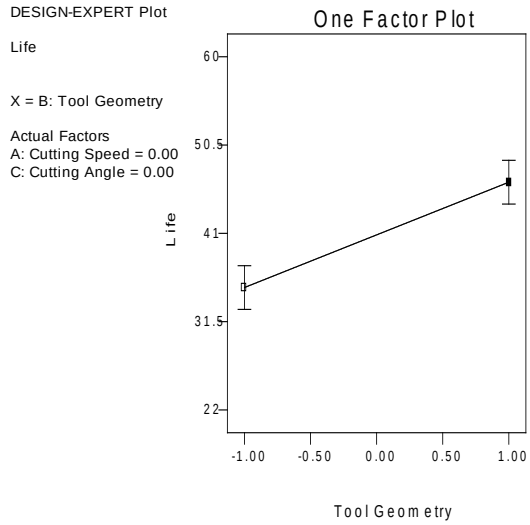
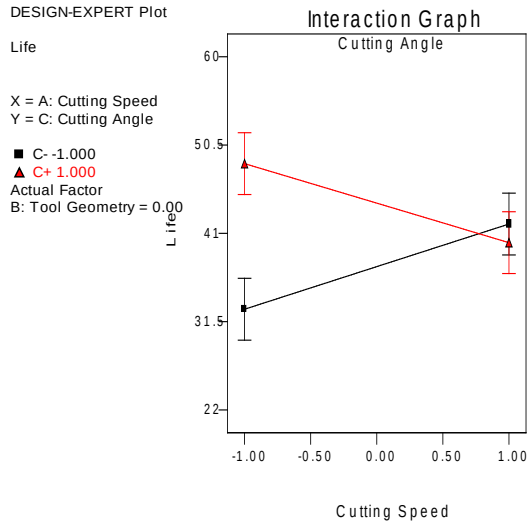
(d) Analyze the residuals. Are there any obvious problems?



There is nothing unusual about the residual plots.

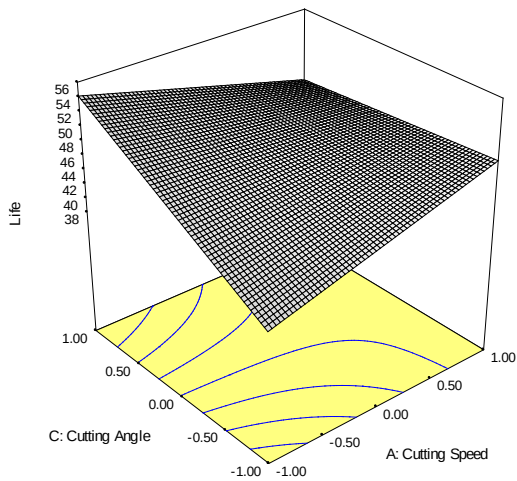
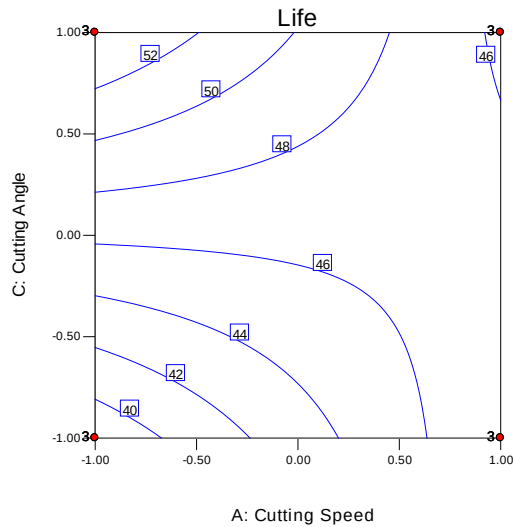
(e) Based on the analysis of main effects and interaction plots, what levels of *A*, *B*, and *C* would you recommend using?

Since *B* has a positive effect, set *B* at the high level to increase life. The *AC* interaction plot reveals that life would be maximized with *C* at the high level and *A* at the low level.



6.2. Reconsider part (c) of Problem 6.1. Use the regression model to generate response surface and contour plots of the tool life response. Interpret these plots. Do they provide insight regarding the desirable operating conditions for this process?

The response surface plot and the contour plot in terms of factors *A* and *C* with *B* at the high level are shown below. They show the curvature due to the *AC* interaction. These plots make it easy to see the region of greatest tool life.



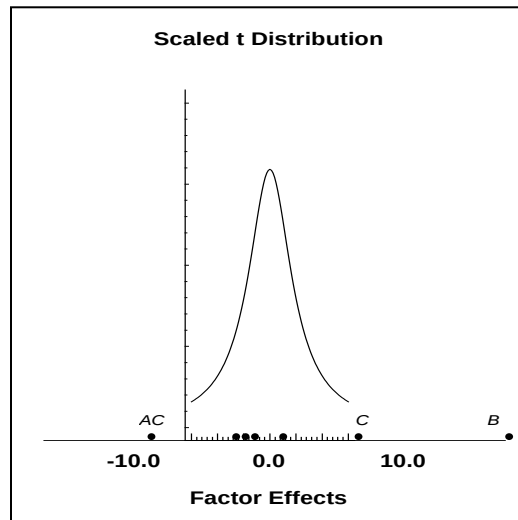
6.3. Find the standard error of the factor effects and approximate 95 percent confidence limits for the factor effects in Problem 6.1. Do the results of this analysis agree with the conclusions from the analysis of variance?

$$SE_{(effect)} = \sqrt{\frac{1}{n2^{k-2}} S^2} = \sqrt{\frac{1}{(3)2^{3-2}} 30.17} = 2.24$$

Variable	Effect	
A	0.333	
B	11.333	*
AB	-1.667	
C	6.833	*
AC	-8.833	*
BC	-2.833	
ABC	-2.167	

The 95% confidence intervals for factors B, C and AC do not contain zero. This agrees with the analysis of variance approach.

6.4. Plot the factor effects from Problem 6.1 on a graph relative to an appropriately scaled t distribution. Does this graphical display adequately identify the important factors? Compare the conclusions from this plot with the results from the analysis of variance. $S = \sqrt{\frac{MS_E}{n}} = \sqrt{\frac{30.17}{3}} = 3.17$



This method identifies the same factors as the analysis of variance.

6.5. A router is used to cut locating notches on a printed circuit board. The vibration level at the surface of the board as it is cut is considered to be a major source of dimensional variation in the notches. Two factors are thought to influence vibration: bit size (A) and cutting speed (B). Two bit sizes (1/16 and 1/8 inch) and two speeds (40 and 90 rpm) are selected, and four boards are cut at each set of conditions shown below. The response variable is vibration measured as a resultant vector of three accelerometers (x, y, and z) on each test circuit board.

A	B	Treatment Combination	Replicate			
			I	II	III	IV
-	-	(1)	18.2	18.9	12.9	14.4
+	-	a	27.2	24.0	22.4	22.5
-	+	b	15.9	14.5	15.1	14.2
+	+	ab	41.0	43.9	36.3	39.9

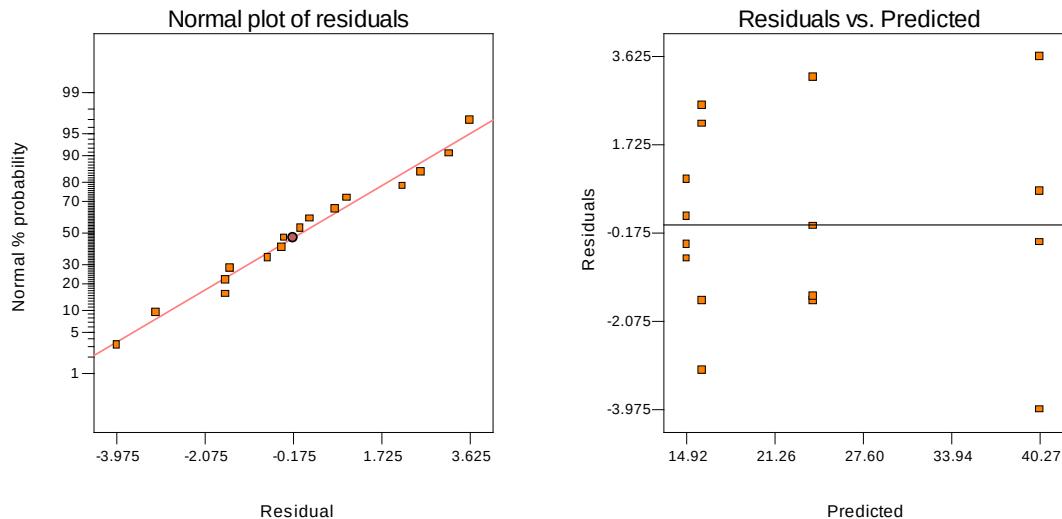
(a) Analyze the data from this experiment.

Design Expert Output

Response: Vibration						
ANOVA for Selected Factorial Model						
Analysis of variance table [Partial sum of squares]						
Source	Sum of Squares	DF	Mean Square	F Value	Prob > F	
Model	1638.11	3	546.04	91.36	< 0.0001	significant
A	1107.23	1	1107.23	185.25	< 0.0001	
B	227.26	1	227.26	38.02	< 0.0001	
AB	303.63	1	303.63	50.80	< 0.0001	
Residual	71.72	12	5.98			
Lack of Fit	0.000	0				
Pure Error	71.72	12	5.98			
Cor Total	1709.83	15				

The Model F-value of 91.36 implies the model is significant. There is only a 0.01% chance that a "Model F-Value" this large could occur due to noise.

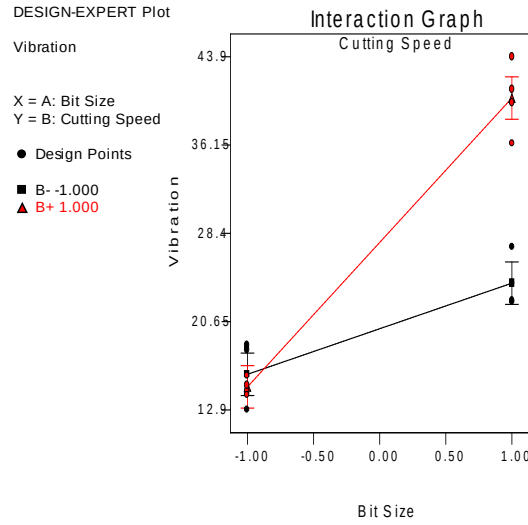
(b) Construct a normal probability plot of the residuals, and plot the residuals versus the predicted vibration level. Interpret these plots.



There is nothing unusual about the residual plots.

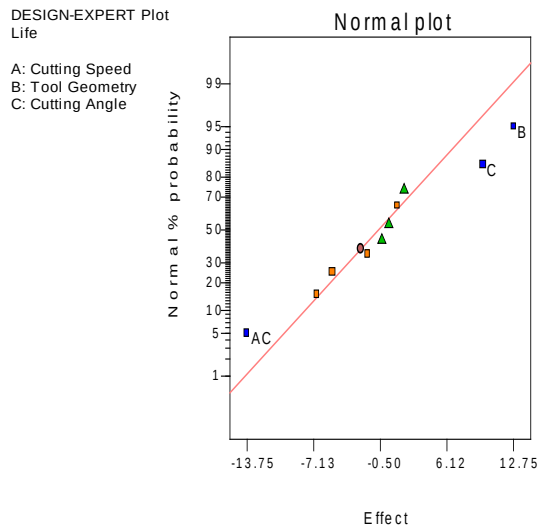
(c) Draw the AB interaction plot. Interpret this plot. What levels of bit size and speed would you recommend for routine operation?

To reduce the vibration, use the smaller bit. Once the small bit is specified, either speed will work equally well, because the slope of the curve relating vibration to speed for the small tip is approximately zero. The process is robust to speed changes if the small bit is used.



6.6. Reconsider the experiment described in Problem 6.1. Suppose that the experimenter only performed the eight trials from replicate I. In addition, he ran four center points and obtained the following response values: 36, 40, 43, 45.

(a) Estimate the factor effects. Which effects are large?



Effects B, C, and AC appear to be large.

(b) Perform an analysis of variance, including a check for pure quadratic curvature. What are your

conclusions?
$$SS_{PureQuadratic} = \frac{n_F n_C (\bar{y}_F - \bar{y}_C)^2}{n_F + n_C} = \frac{(8)(4)(40.875 - 41.000)^2}{8 + 4} = 0.0417$$

Design Expert Output

Response: Life in hours						
ANOVA for Selected Factorial Model						
Analysis of variance table [Partial sum of squares]						
Source	Sum of Squares	DF	Mean Square	F Value	Prob > F	
Model	1048.88	7	149.84	9.77	0.0439	significant
A	3.13	1	3.13	0.20	0.6823	
B	325.13	1	325.13	21.20	0.0193	
C	190.12	1	190.12	12.40	0.0389	
AB	6.13	1	6.13	0.40	0.5722	
AC	378.12	1	378.12	24.66	0.0157	
BC	55.12	1	55.12	3.60	0.1542	
ABC	91.12	1	91.12	5.94	0.0927	
Curvature	0.042	1	0.042	2.717E-003	0.9617	not significant
Pure Error	46.00	3	15.33			
Cor Total	1094.92	11				

The Model F-value of 9.77 implies the model is significant. There is only a 4.39% chance that a "Model F-Value" this large could occur due to noise.

The "Curvature F-value" of 0.00 implies the curvature (as measured by difference between the average of the center points and the average of the factorial points) in the design space is not significant relative to the noise. There is a 96.17% chance that a "Curvature F-value" this large could occur due to noise.

Design Expert Output

Response: Life in hours						
ANOVA for Selected Factorial Model						
Analysis of variance table [Partial sum of squares]						
Source	Sum of Squares	DF	Mean Square	F Value	Prob > F	
Model	896.50	4	224.13	7.91	0.0098	significant
A	3.13	1	3.13	0.11	0.7496	
B	325.12	1	325.12	11.47	0.0117	
C	190.12	1	190.12	6.71	0.0360	
AC	378.12	1	378.12	13.34	0.0082	
Residual	198.42	7	28.35			
Lack of Fit	152.42	4	38.10	2.49	0.2402	not significant
Pure Error	46.00	3	15.33			
Cor Total	1094.92	11				

The Model F-value of 7.91 implies the model is significant. There is only a 0.98% chance that a "Model F-Value" this large could occur due to noise.

Effects B, C and AC are significant at 5%. There is no effect of curvature.

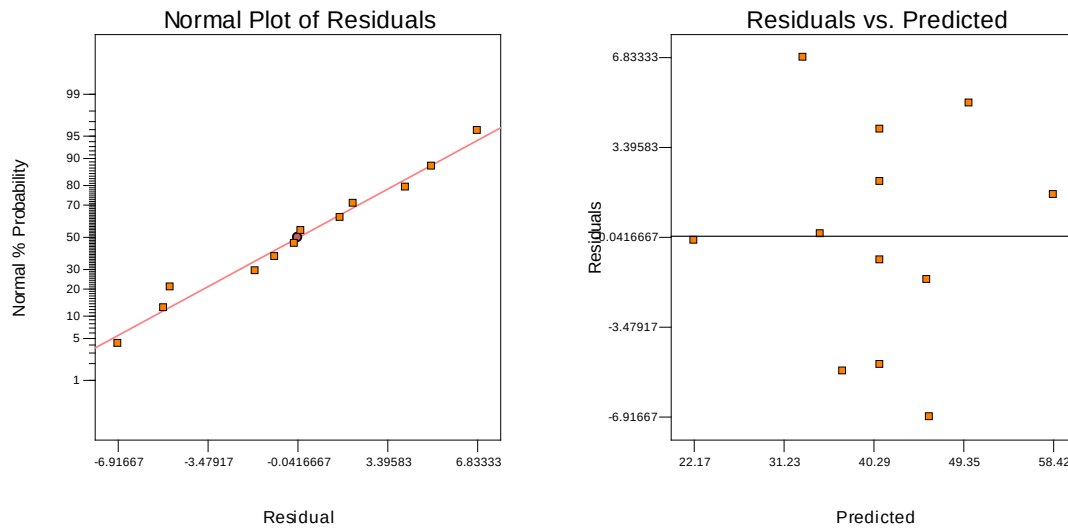
- (c) Write down an appropriate model for predicting tool life, based on the results of this experiment. Does this model differ in any substantial way from the model in Problem 6.1, part (c)?

The model shown in the *Design Expert* output below does not differ substantially from the model in Problem 6.1, part (c).

Design Expert Output

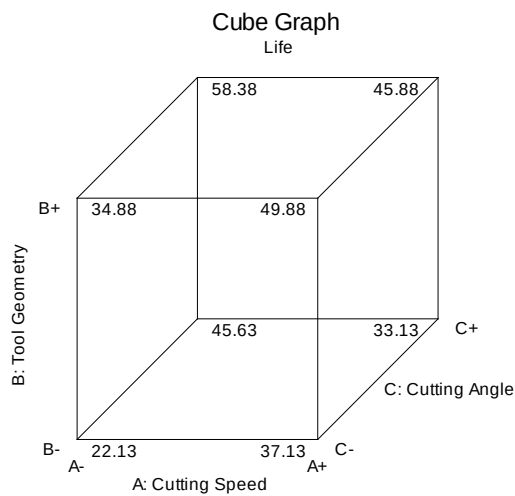
Final Equation in Terms of Coded Factors:		
Life	=	
+40.88		
+0.62	* A	
+6.37	* B	
+4.87	* C	
-6.88	* A * C	

- (d) Analyze the residuals.



(e) What conclusions would you draw about the appropriate operating conditions for this process?

To maximize life run with *B* at the high level, *A* at the low level and *C* at the high level



6.7. An experiment was performed to improve the yield of a chemical process. Four factors were selected, and two replicates of a completely randomized experiment were run. The results are shown in the following table:

Treatment Combination	Replicate I	Replicate II	Treatment Combination	Replicate I	Replicate II
(1)	90	93	<i>d</i>	98	95
<i>a</i>	74	78	<i>ad</i>	72	76
<i>b</i>	81	85	<i>bd</i>	87	83

<i>ab</i>	83	80	<i>abd</i>	85	86
<i>c</i>	77	78	<i>cd</i>	99	90
<i>ac</i>	81	80	<i>acd</i>	79	75
<i>bc</i>	88	82	<i>bcd</i>	87	84
<i>abc</i>	73	70	<i>abcd</i>	80	80

(a) Estimate the factor effects.

Design Expert Output

Model	Term	Effect	SumSqr	% Contribtn
Error	Intercept			
Error	A	-9.0625	657.031	40.3714
Error	B	-1.3125	13.7812	0.84679
Error	C	-2.6875	57.7813	3.55038
Error	D	3.9375	124.031	7.62111
Error	AB	4.0625	132.031	8.11267
Error	AC	0.6875	3.78125	0.232339
Error	AD	-2.1875	38.2813	2.3522
Error	BC	-0.5625	2.53125	0.155533
Error	BD	-0.1875	0.28125	0.0172814
Error	CD	1.6875	22.7812	1.3998
Error	ABC	-5.1875	215.281	13.228
Error	ABD	4.6875	175.781	10.8009
Error	ACD	-0.9375	7.03125	0.432036
Error	BCD	-0.9375	7.03125	0.432036
Error	ABCD	2.4375	47.5313	2.92056

(b) Prepare an analysis of variance table, and determine which factors are important in explaining yield.

Design Expert Output

Response: yield						
ANOVA for Selected Factorial Model						
Analysis of variance table [Partial sum of squares]						
Source	Sum of Squares	DF	Mean Square	F Value	Prob > F	
Model	1504.97	15	100.33	13.10	< 0.0001	significant
A	657.03	1	657.03	85.82	< 0.0001	
B	13.78	1	13.78	1.80	0.1984	
C	57.78	1	57.78	7.55	0.0143	
D	124.03	1	124.03	16.20	0.0010	
AB	132.03	1	132.03	17.24	0.0007	
AC	3.78	1	3.78	0.49	0.4923	
AD	38.28	1	38.28	5.00	0.0399	
BC	2.53	1	2.53	0.33	0.5733	
BD	0.28	1	0.28	0.037	0.8504	
CD	22.78	1	22.78	2.98	0.1038	
ABC	215.28	1	215.28	28.12	< 0.0001	
ABD	175.78	1	175.78	22.96	0.0002	
ACD	7.03	1	7.03	0.92	0.3522	
BCD	7.03	1	7.03	0.92	0.3522	
ABCD	47.53	1	47.53	6.21	0.0241	
Residual	122.50	16	7.66			
Lack of Fit	0.000	0				
Pure Error	122.50	16	7.66			
Cor Total	1627.47	31				

The Model F-value of 13.10 implies the model is significant. There is only a 0.01% chance that a "Model F-Value" this large could occur due to noise.

Values of "Prob > F" less than 0.0500 indicate model terms are significant.
In this case A, C, D, AB, AD, ABC, ABD, ABCD are significant model terms.

$F_{0.01,1,16} = 8.53$, and $F_{0.025,1,16} = 6.12$ therefore, factors *A* and *D* and interactions *AB*, *ABC*, and *ABD* are significant at 1%. Factor *C* and interactions *AD* and *ABCD* are significant at 5%.

- (b) Write down a regression model for predicting yield, assuming that all four factors were varied over the range from -1 to +1 (in coded units).

Model with hierarchy maintained:

Design Expert Output

Final Equation in Terms of Coded Factors:

```
yield =
+82.78
-4.53 * A
-0.66 * B
-1.34 * C
+1.97 * D
+2.03 * A * B
+0.34 * A * C
-1.09 * A * D
-0.28 * B * C
-0.094 * B * D
+0.84 * C * D
-2.59 * A * B * C
+2.34 * A * B * D
-0.47 * A * C * D
-0.47 * B * C * D
+1.22 * A * B * C * D
```

Model without hierarchy terms:

Design Expert Output

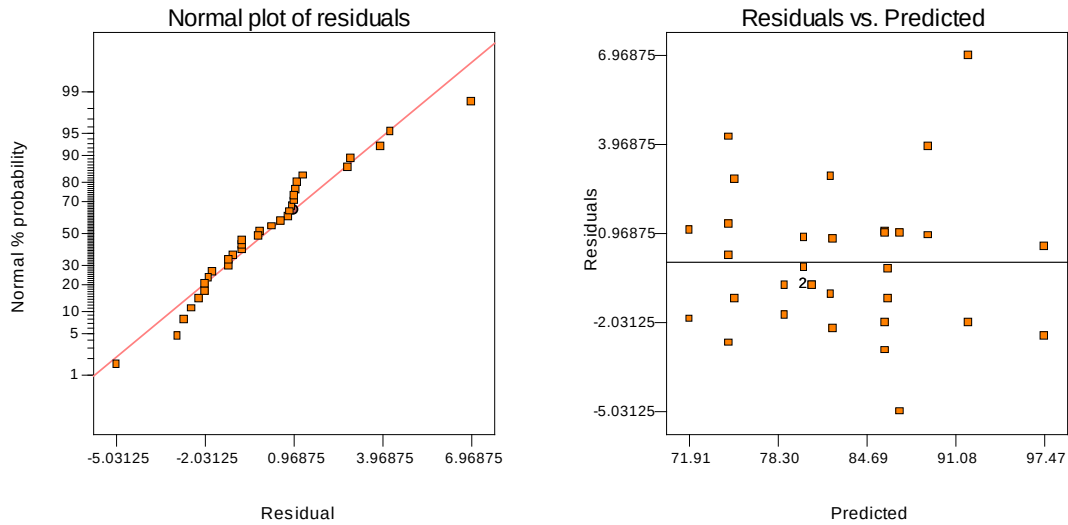
Final Equation in Terms of Coded Factors:

```
yield =
+82.78
-4.53 * A
-1.34 * C
+1.97 * D
+2.03 * A * B
-1.09 * A * D
-2.59 * A * B * C
+2.34 * A * B * D
+1.22 * A * B * C * D
```

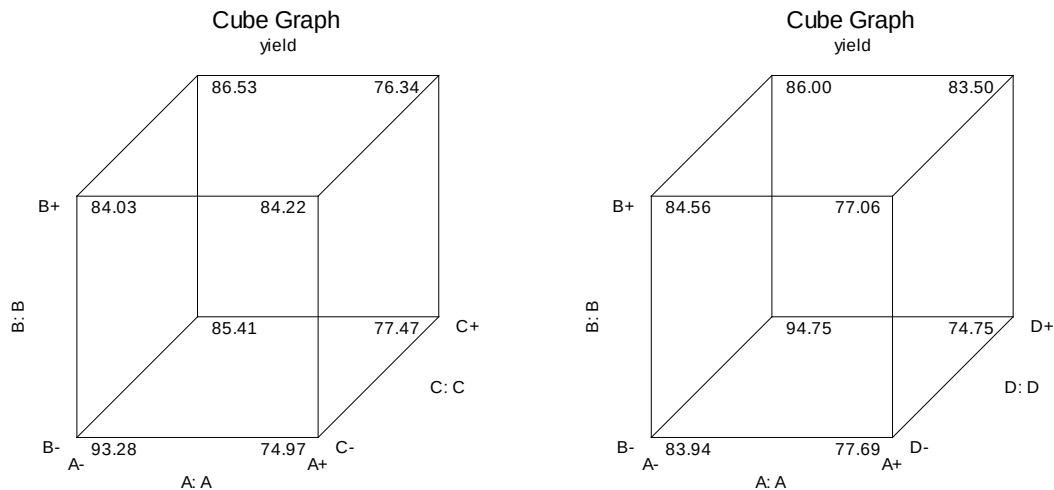
Confirmation runs might be run to see if the simpler model without hierarchy is satisfactory.

- (d) Plot the residuals versus the predicted yield and on a normal probability scale. Does the residual analysis appear satisfactory?

There appears to be one large residual both in the normal probability plot and in the plot of residuals versus predicted.



- (e) Two three-factor interactions, ABC and ABD , apparently have large effects. Draw a cube plot in the factors A , B , and C with the average yields shown at each corner. Repeat using the factors A , B , and D . Do these two plots aid in data interpretation? Where would you recommend that the process be run with respect to the four variables?



Run the process at A low B low, C low and D high.

6.8. A bacteriologist is interested in the effects of two different culture media and two different times on the growth of a particular virus. She performs six replicates of a 2^2 design, making the runs in random order. Analyze the bacterial growth data that follow and draw appropriate conclusions. Analyze the residuals and comment on the model's adequacy.

Time	Culture Medium			
	1	2	3	4
12 hr	21	22	25	26
	23	28	24	25
	20	26	29	27
	37	39	31	34
18 hr	38	38	29	33
	35	36	30	35

Design Expert Output

Response: Virus growth

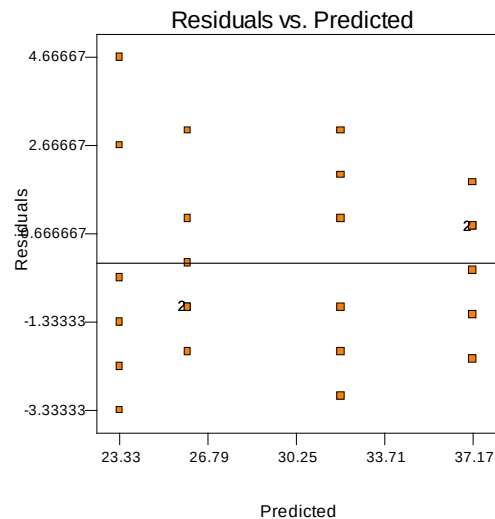
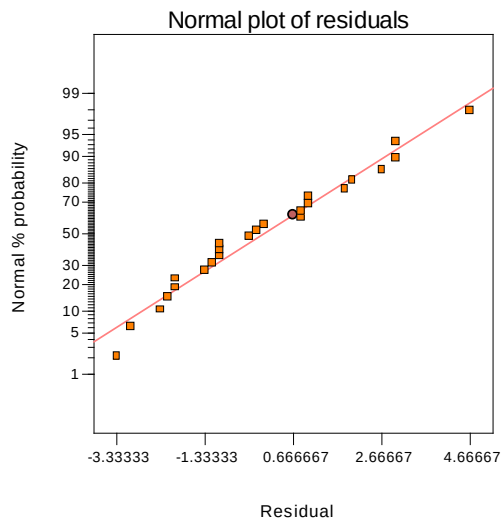
ANOVA for Selected Factorial Model

Analysis of variance table [Partial sum of squares]

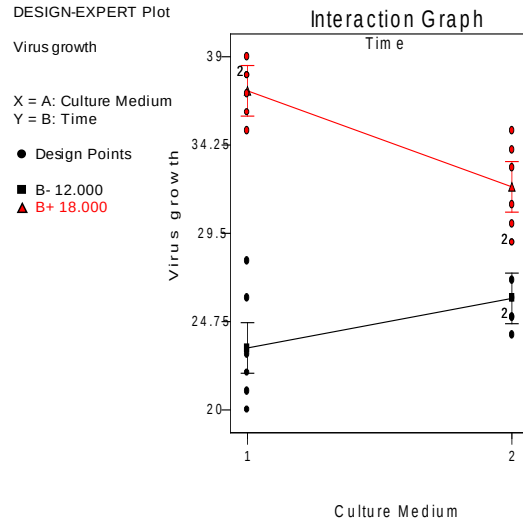
Source	Sum of Squares	DF	Mean Square	F Value	Prob > F	
Model	691.46	3	230.49	45.12	< 0.0001	significant
A	9.38	1	9.38	1.84	0.1906	
B	590.04	1	590.04	115.51	< 0.0001	
AB	92.04	1	92.04	18.02	0.0004	
Residual	102.17	20	5.11			
Lack of Fit	0.000	0				
Pure Error	102.17	20	5.11			
Cor Total	793.63	23				

The Model F-value of 45.12 implies the model is significant. There is only a 0.01% chance that a "Model F-Value" this large could occur due to noise.

Values of "Prob > F" less than 0.0500 indicate model terms are significant. In this case B, AB are significant model terms.



Growth rate is affected by factor *B* (Time) and the *AB* interaction (Culture medium and Time). There is some very slight indication of inequality of variance shown by the small decreasing funnel shape in the plot of residuals versus predicted.



6.9. An industrial engineer employed by a beverage bottler is interested in the effects of two different types of 32-ounce bottles on the time to deliver 12-bottle cases of the product. The two bottle types are glass and plastic. Two workers are used to perform a task consisting of moving 40 cases of the product 50 feet on a standard type of hand truck and stacking the cases in a display. Four replicates of a 2^2 factorial design are performed, and the times observed are listed in the following table. Analyze the data and draw the appropriate conclusions. Analyze the residuals and comment on the model's adequacy.

Bottle Type	Worker			
	1	1	2	2
Glass	5.12	4.89	6.65	6.24
	4.98	5.00	5.49	5.55
Plastic	4.95	4.43	5.28	4.91
	4.27	4.25	4.75	4.71

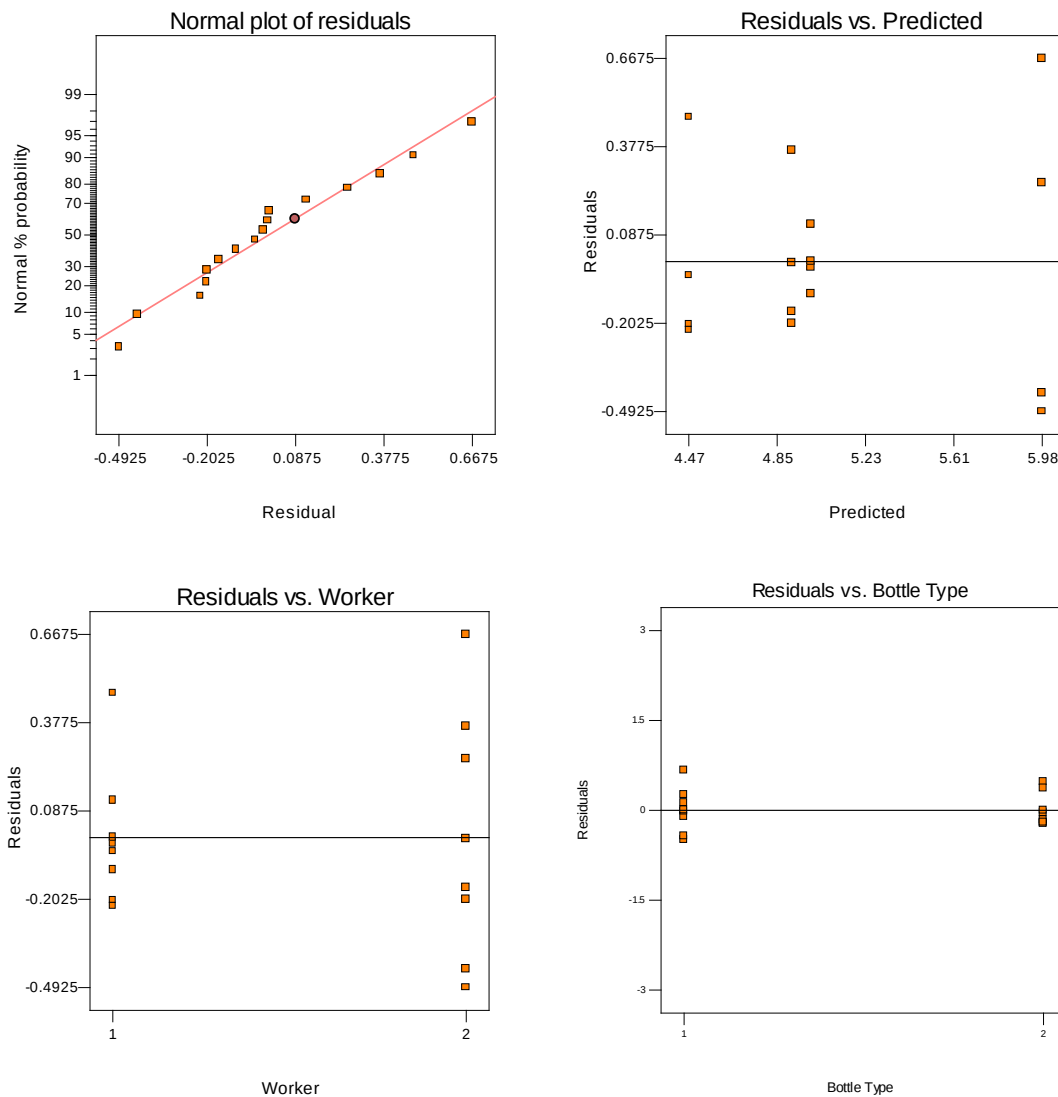
Design Expert Output

Response: Times						
ANOVA for Selected Factorial Model						
Analysis of variance table [Partial sum of squares]						
	Source	Sum of Squares	DF	Mean Square	F Value	Prob > F
Model	4.86	3	1.62	13.04	0.0004	significant
A	2.02	1	2.02	16.28	0.0017	
B	2.54	1	2.54	20.41	0.0007	
AB	0.30	1	0.30	2.41	0.1463	
Residual	1.49	12	0.12			
Lack of Fit	0.000	0				
Pure Error	1.49	12	0.12			
Cor Total	6.35	15				

The Model F-value of 13.04 implies the model is significant. There is only a 0.04% chance that a "Model F-Value" this large could occur due to noise.

Values of "Prob > F" less than 0.0500 indicate model terms are significant. In this case A, B are significant model terms.

There is some indication of non-constant variance in this experiment.



6.10. In problem 6.9, the engineer was also interested in potential fatigue differences resulting from the two types of bottles. As a measure of the amount of effort required, he measured the elevation of heart rate (pulse) induced by the task. The results follow. Analyze the data and draw conclusions. Analyze the residuals and comment on the model's adequacy.

Bottle Type	Worker			
	1	1	2	2
Glass	39	45	20	13
	58	35	16	11
Plastic	44	35	13	10
	42	21	16	15

Design Expert Output

Response: Pulse

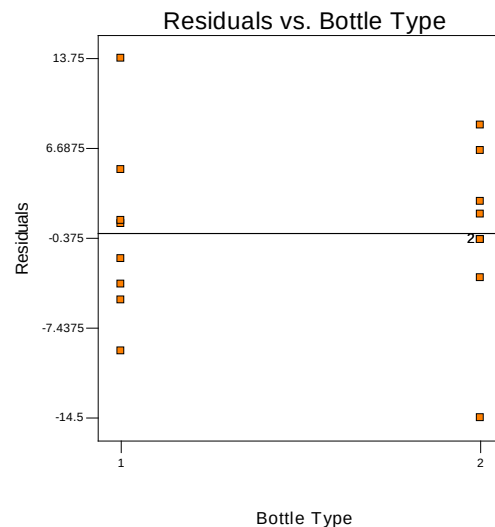
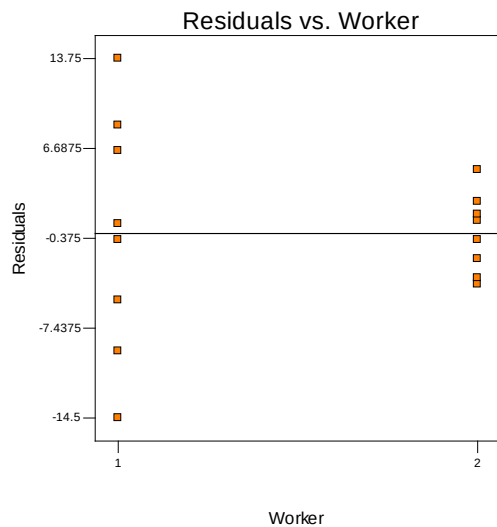
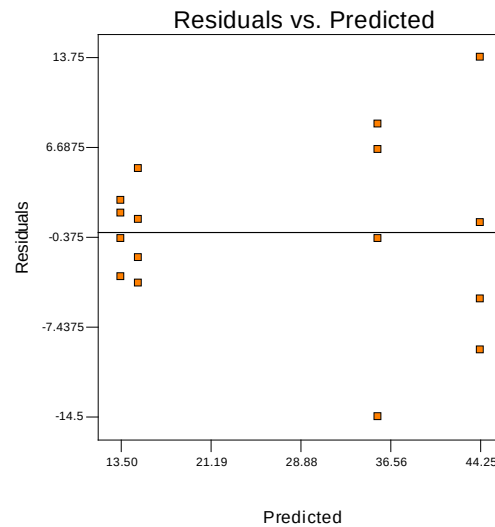
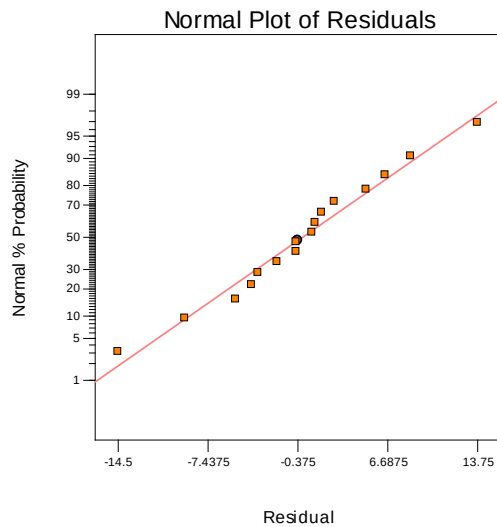
ANOVA for Selected Factorial Model

Analysis of variance table [Partial sum of squares]

Source	Sum of Squares	DF	Mean Square	F Value	Prob > F	
Model	2784.19	3	928.06	16.03	0.0002	significant
A	2626.56	1	2626.56	45.37	< 0.0001	
B	105.06	1	105.06	1.81	0.2028	
AB	52.56	1	52.56	0.91	0.3595	
Residual	694.75	12	57.90			
Lack of Fit	0.000	0				
Pure Error	694.75	12	57.90			
Cor Total	3478.94	15				

The Model F-value of 16.03 implies the model is significant. There is only a 0.02% chance that a "Model F-Value" this large could occur due to noise.

Values of "Prob > F" less than 0.0500 indicate model terms are significant. In this case A are significant model terms.



There is an indication that one worker exhibits greater variability than the other.

6.11. Calculate approximate 95 percent confidence limits for the factor effects in Problem 6.10. Do the results of this analysis agree with the analysis of variance performed in Problem 6.10?

$$SE_{(effect)} = \sqrt{\frac{1}{n2^{k-2}} S^2} = \sqrt{\frac{1}{(4)2^{2-2}} 57.90} = 3.80$$

Variable	Effect	C.I.
A	-25.625	$\pm 3.80(1.96) = \pm 7.448$
B	-5.125	$\pm 3.80(1.96) = \pm 7.448$
AB	3.625	$\pm 3.80(1.96) = \pm 7.448$

The 95% confidence interval for factor A does not contain zero. This agrees with the analysis of variance approach.

6.12. An article in the *AT&T Technical Journal* (March/April 1986, Vol. 65, pp. 39-50) describes the application of two-level factorial designs to integrated circuit manufacturing. A basic processing step is to grow an epitaxial layer on polished silicon wafers. The wafers mounted on a susceptor are positioned inside a bell jar, and chemical vapors are introduced. The susceptor is rotated and heat is applied until the epitaxial layer is thick enough. An experiment was run using two factors: arsenic flow rate (A) and deposition time (B). Four replicates were run, and the epitaxial layer thickness was measured (in mm). The data are shown below:

A	B	Replicate				Factor	Levels
		I	II	III	IV		
-	-	14.037	16.165	13.972	13.907	A	55%
+	-	13.880	13.860	14.032	13.914		
-	+	14.821	14.757	14.843	14.878	B	Short
+	+	14.888	14.921	14.415	14.932		

(a) Estimate the factor effects.

Design Expert Output

	Term	Effect	SumSqr	% Contribtn
Model	Intercept			
Error	A	-0.31725	0.40259	6.79865
Error	B	0.586	1.37358	23.1961
Error	AB	0.2815	0.316969	5.35274
Error	Lack Of Fit		0	0
Error	Pure Error		3.82848	64.6525

(b) Conduct an analysis of variance. Which factors are important?

From the analysis of variance shown below, no factors appear to be important. Factor B is only marginally interesting with an *F*-value of 4.31.

Design Expert Output

Response: Thickness						
ANOVA for Selected Factorial Model						
Analysis of variance table [Partial sum of squares]						
Source	Sum of Squares	DF	Mean Square	F Value	Prob > F	
Model	2.09	3	0.70	2.19	0.1425	not significant

A	0.40	1	0.40	1.26	0.2833
B	1.37	1	1.37	4.31	0.0602
AB	0.32	1	0.32	0.99	0.3386
Residual	3.83	12	0.32		
Lack of Fit	0.000	0			
Pure Error	3.83	12	0.32		
Cor Total	5.92	15			

The "Model F-value" of 2.19 implies the model is not significant relative to the noise. There is a 14.25 % chance that a "Model F-value" this large could occur due to noise.

Values of "Prob > F" less than 0.0500 indicate model terms are significant.
In this case there are no significant model terms.

- (c) Write down a regression equation that could be used to predict epitaxial layer thickness over the region of arsenic flow rate and deposition time used in this experiment.

Design Expert Output

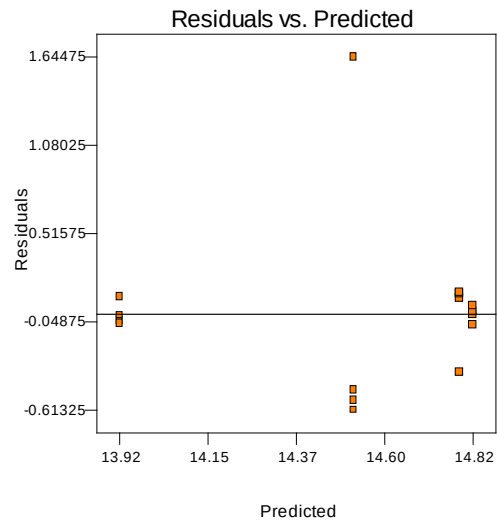
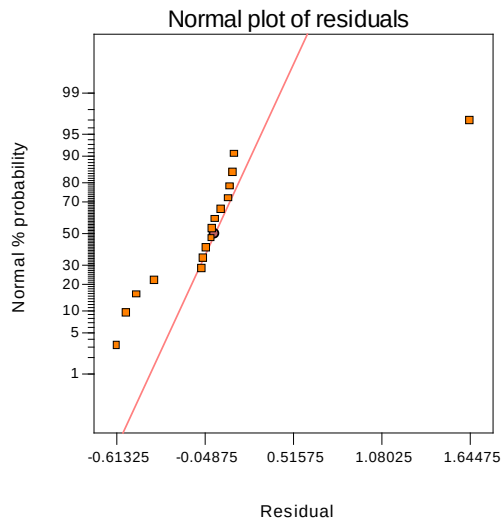
Final Equation in Terms of Coded Factors:

$$\begin{aligned} \text{Thickness} = & \\ & +14.51 \\ & -0.16 \quad * A \\ & +0.29 \quad * B \\ & +0.14 \quad * A * B \end{aligned}$$

Final Equation in Terms of Actual Factors:

$$\begin{aligned} \text{Thickness} = & \\ & +37.62656 \\ & -0.43119 \quad * \text{Flow Rate} \\ & -1.48735 \quad * \text{Dep Time} \\ & +0.028150 \quad * \text{Flow Rate} * \text{Dep Time} \end{aligned}$$

- (d) Analyze the residuals. Are there any residuals that should cause concern? Observation #2 falls outside the groupings in the normal probability plot and the plot of residual versus predicted.



- (e) Discuss how you might deal with the potential outlier found in part (d).

One approach would be to replace the observation with the average of the observations from that experimental cell. Another approach would be to identify if there was a recording issue in the original data.

The first analysis below replaces the data point with the average of the other three. The second analysis assumes that the reading was incorrectly recorded and should have been 14.165.

The analysis with the run associated with standard order 2 replaced with the average of the remaining three runs in the cell, 13.972, is shown below.

Design Expert Output

Response: Thickness						
ANOVA for Selected Factorial Model						
Analysis of variance table [Partial sum of squares]						
Source	Sum of Squares	DF	Mean Square	F Value	Prob > F	
Model	2.97	3	0.99	53.57	< 0.0001	significant
A	7.439E-003	1	7.439E-003	0.40	0.5375	
B	2.96	1	2.96	160.29	< 0.0001	
AB	2.176E-004	1	2.176E-004	0.012	0.9153	
Pure Error	0.22	12	0.018			
Cor Total	3.19	15				

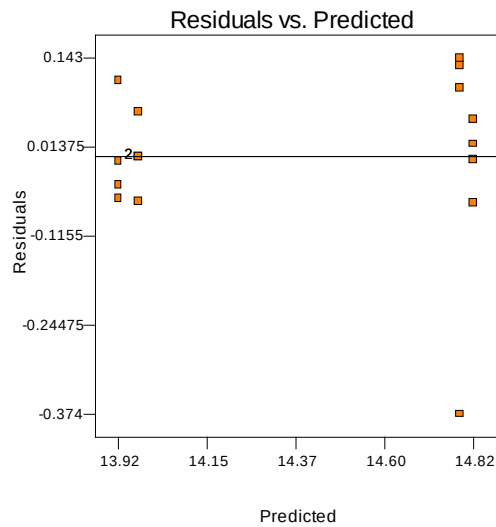
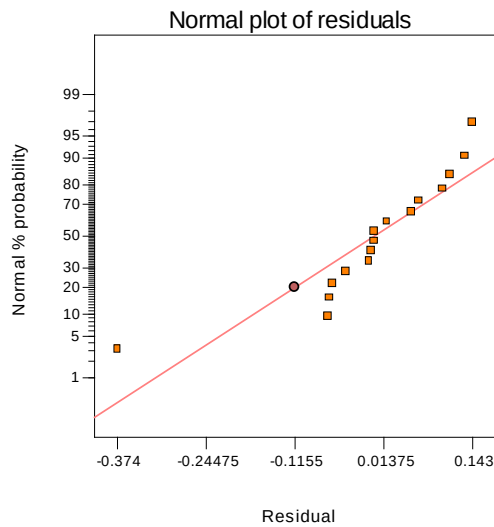
The Model F-value of 53.57 implies the model is significant. There is only a 0.01% chance that a "Model F-Value" this large could occur due to noise.

Values of "Prob > F" less than 0.0500 indicate model terms are significant. In this case B are significant model terms.

Final Equation in Terms of Coded Factors:

$$\text{Thickness} = +14.38 - 0.022 * A + 0.43 * B + 3.688E-003 * A * B$$

Final Equation in Terms of Actual Factors:

$$\text{Thickness} = +13.36650 - 0.020000 * \text{Flow Rate} + 0.12999 * \text{Dep Time} + 7.37500E-004 * \text{Flow Rate} * \text{Dep Time}$$


A new outlier is present and should be investigated.

Analysis with the run associated with standard order 2 replaced with the value 14.165:

Design Expert Output

Response: Thickness						
ANOVA for Selected Factorial Model						
Analysis of variance table [Partial sum of squares]						
Source	Sum of Squares	DF	Mean Square	F Value	Prob > F	
Model	2.82	3	0.94	45.18	< 0.0001	significant
A	0.018	1	0.018	0.87	0.3693	
B	2.80	1	2.80	134.47	< 0.0001	
AB	3.969E-003	1	3.969E-003	0.19	0.6699	
Pure Error	0.25	12	0.021			
Cor Total	3.07	15				

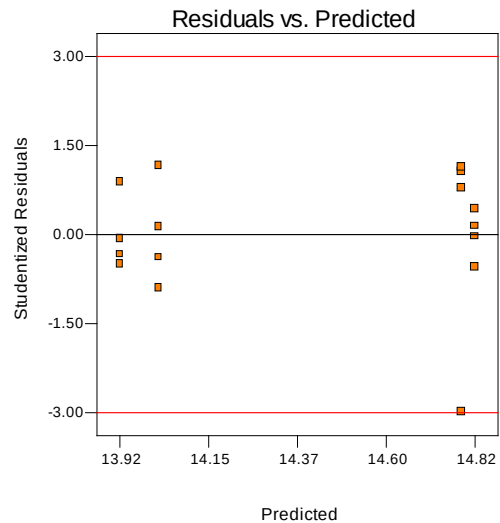
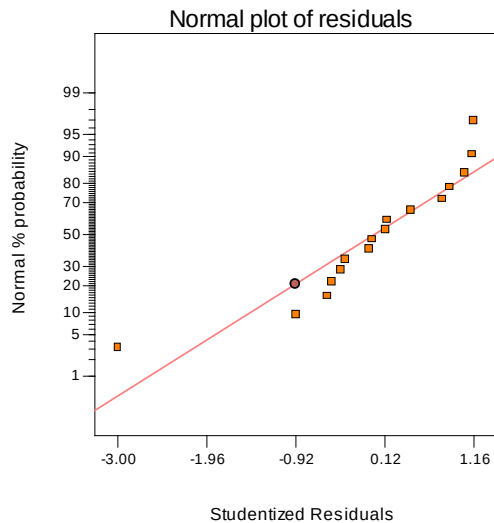
The Model F-value of 45.18 implies the model is significant. There is only a 0.01% chance that a "Model F-Value" this large could occur due to noise.

Values of "Prob > F" less than 0.0500 indicate model terms are significant. In this case B are significant model terms.

Final Equation in Terms of Coded Factors:

$$\text{Thickness} = +14.39 - 0.034 * A + 0.42 * B + 0.016 * A * B$$

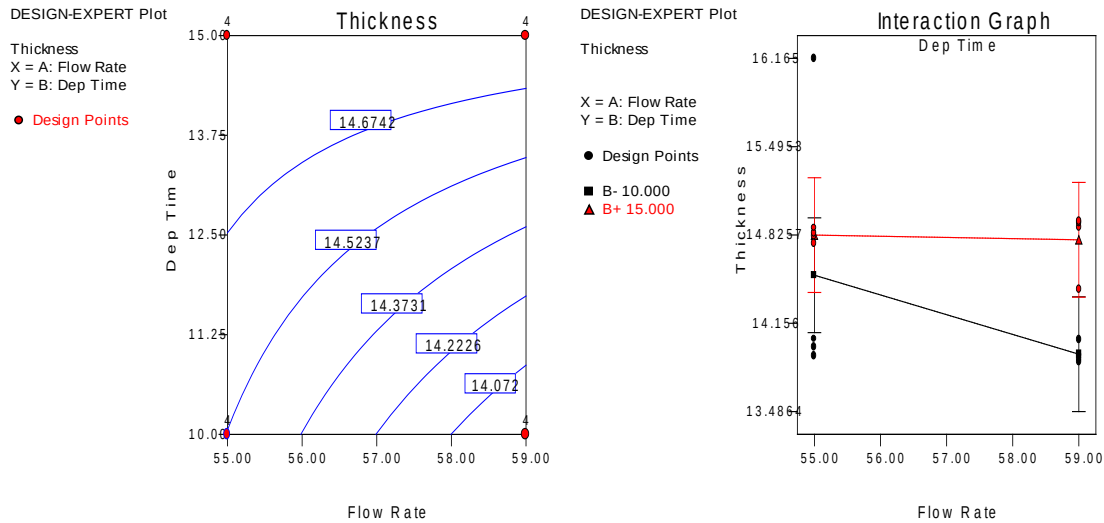
Final Equation in Terms of Actual Factors:

$$\text{Thickness} = +15.50156 - 0.056188 * \text{Flow Rate} - 0.012350 * \text{Dep Time} + 3.15000\text{E-}003 * \text{Flow Rate} * \text{Dep Time}$$


Another outlier is present and should be investigated.

6.13. Continuation of Problem 6.12. Use the regression model in part (c) of Problem 6.12 to generate a response surface contour plot for epitaxial layer thickness. Suppose it is critically important to obtain layer thickness of 14.5 mm. What settings of arsenic flow rate and deposition time would you recommend?

Arsenic flow rate may be set at any of the experimental levels, while the deposition time should be set at 12.4 minutes.



6.14. Continuation of Problem 6.13. How would your answer to Problem 6.13 change if arsenic flow rate was more difficult to control in the process than the deposition time?

Running the process at a high level of Deposition Time there is no change in thickness as flow rate changes.

6.15. A nickel-titanium alloy is used to make components for jet turbine aircraft engines. Cracking is a potentially serious problem in the final part, as it can lead to non-recoverable failure. A test is run at the parts producer to determine the effects of four factors on cracks. The four factors are pouring temperature (A), titanium content (B), heat treatment method (C), and the amount of grain refiner used (D). Two replicated of a 2^4 design are run, and the length of crack (in μm) induced in a sample coupon subjected to a standard test is measured. The data are shown below:

A	B	C	D	Treatment Combination	Replicate I	Replicate II
-	-	-	-	(1)	7.037	6.376
+	-	-	-	a	14.707	15.219
-	+	-	-	b	11.635	12.089
+	+	-	-	ab	17.273	17.815
-	-	+	-	c	10.403	10.151
+	-	+	-	ac	4.368	4.098
-	+	+	-	bc	9.360	9.253
+	+	+	-	abc	13.440	12.923
-	-	-	+	d	8.561	8.951

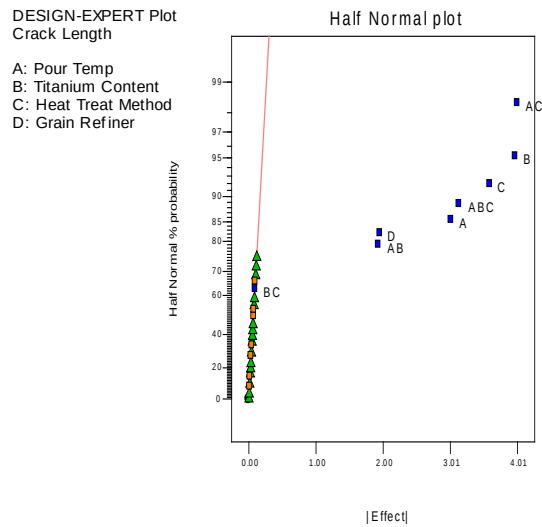
+	-	-	+	<i>ad</i>	16.867	17.052
-	+	-	+	<i>bd</i>	13.876	13.658
+	+	-	+	<i>abd</i>	19.824	19.639
-	-	+	+	<i>cd</i>	11.846	12.337
+	-	+	+	<i>acd</i>	6.125	5.904
-	+	+	+	<i>bcd</i>	11.190	10.935
+	+	+	+	<i>abcd</i>	15.653	15.053

(a) Estimate the factor effects. Which factors appear to be large?

From the half normal plot of effects shown below, factors *A*, *B*, *C*, *D*, *AB*, *AC*, and *ABC* appear to be large.

Design Expert Output

	Term	Effect	SumSqr	% Contribtn
Model	Intercept			
Model	A	3.01888	72.9089	12.7408
Model	B	3.97588	126.461	22.099
Model	C	-3.59625	103.464	18.0804
Model	D	1.95775	30.6623	5.35823
Model	AB	1.93412	29.9267	5.22969
Model	AC	-4.00775	128.496	22.4548
Error	AD	0.0765	0.046818	0.00818145
Error	BC	0.096	0.073728	0.012884
Error	BD	0.04725	0.0178605	0.00312112
Error	CD	-0.076875	0.0472781	0.00826185
Model	ABC	3.1375	78.7512	13.7618
Error	ABD	0.098	0.076832	0.0134264
Error	ACD	0.019125	0.00292613	0.00051134
Error	BCD	0.035625	0.0101531	0.00177426
Error	ABCD	0.014125	0.00159613	0.000278923



(b) Conduct an analysis of variance. Do any of the factors affect cracking? Use $\alpha=0.05$.

The Design Expert output below identifies factors A, B, C, D, AB, AC, and ABC as significant.

Design Expert Output

Response: Crack Length in mm x 10 ⁻²						
ANOVA for Selected Factorial Model						
Analysis of variance table [Partial sum of squares]						
Source	Sum of Squares	DF	Mean Square	F Value	Prob > F	
Model	570.95	15	38.06	468.99	< 0.0001	significant
A	72.91	1	72.91	898.34	< 0.0001	
B	126.46	1	126.46	1558.17	< 0.0001	
C	103.46	1	103.46	1274.82	< 0.0001	
D	30.66	1	30.66	377.80	< 0.0001	
AB	29.93	1	29.93	368.74	< 0.0001	
AC	128.50	1	128.50	1583.26	< 0.0001	
AD	0.047	1	0.047	0.58	0.4586	
BC	0.074	1	0.074	0.91	0.3547	
BD	0.018	1	0.018	0.22	0.6453	
CD	0.047	1	0.047	0.58	0.4564	
ABC	78.75	1	78.75	970.33	< 0.0001	
ABD	0.077	1	0.077	0.95	0.3450	
ACD	2.926E-003	1	2.926E-003	0.036	0.8518	
BCD	0.010	1	0.010	0.13	0.7282	
ABCD	1.596E-003	1	1.596E-003	0.020	0.8902	
Residual	1.30	16	0.081			
Lack of Fit	0.000	0				
Pure Error	1.30	16	0.081			
Cor Total	572.25	31				

The Model F-value of 468.99 implies the model is significant. There is only a 0.01% chance that a "Model F-Value" this large could occur due to noise.

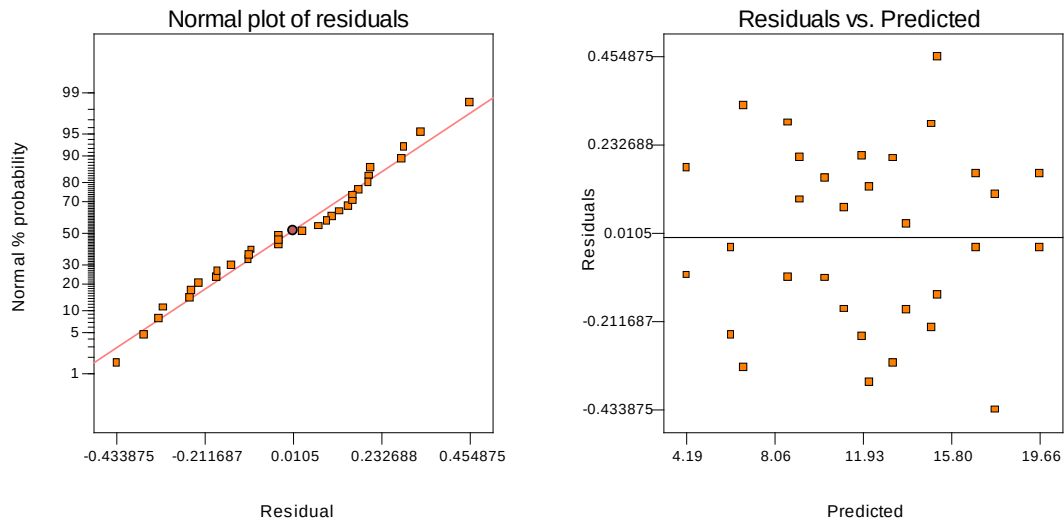
Values of "Prob > F" less than 0.0500 indicate model terms are significant. In this case A, B, C, D, AB, AC, ABC are significant model terms.

(c) Write down a regression model that can be used to predict crack length as a function of the significant main effects and interactions you have identified in part (b).

Design Expert Output

Final Equation in Terms of Coded Factors:	
Crack Length=	
+11.99	
+1.51 *A	
+1.99 *B	
-1.80 *C	
+0.98 *D	
+0.97 *A*B	
-2.00 *A*C	
+1.57 *A*B*C	

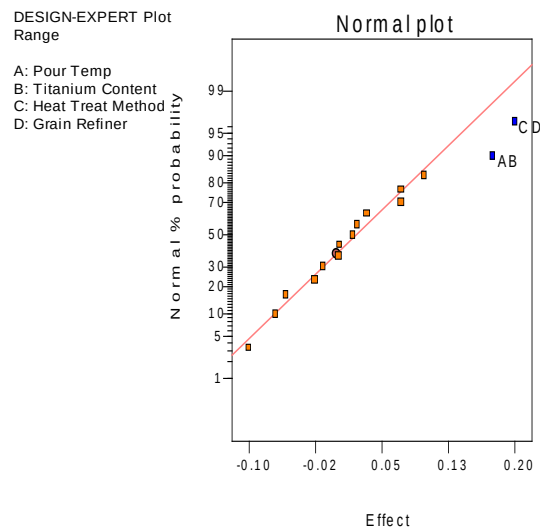
(d) Analyze the residuals from this experiment.



There is nothing unusual about the residuals.

(e) Is there an indication that any of the factors affect the variability in cracking?

By calculating the range of the two readings in each cell, we can also evaluate the effects of the factors on variation. The following is the normal probability plot of effects:



It appears that the *AB* and *CD* interactions could be significant. The following is the ANOVA for the range data:

Design Expert Output

Response:					
Range					
ANOVA for Selected Factorial Model					
Analysis of variance table [Partial sum of squares]					
Source	Sum of Squares	DF	Mean Square	F Value	Prob > F
Model	0.29	2	0.14	11.46	0.0014
AB	0.13	1	0.13	9.98	0.0075
CD	0.16	1	0.16	12.94	0.0032
Residual	0.16	13	0.013		
Cor Total	0.45	15			

The Model F-value of 11.46 implies the model is significant. There is only a 0.14% chance that a "Model F-Value" this large could occur due to noise.

Values of "Prob > F" less than 0.0500 indicate model terms are significant. In this case AB, CD are significant model terms.

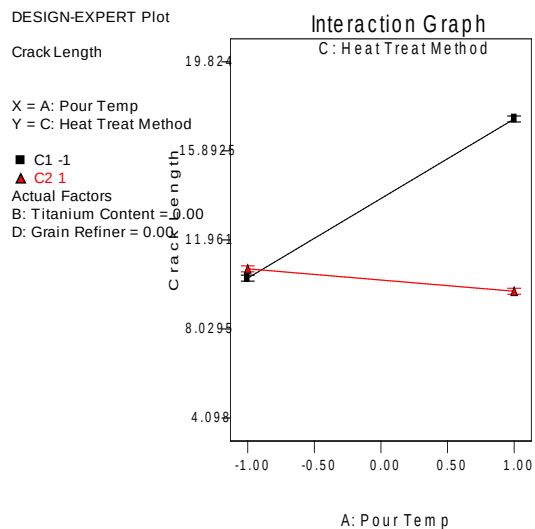
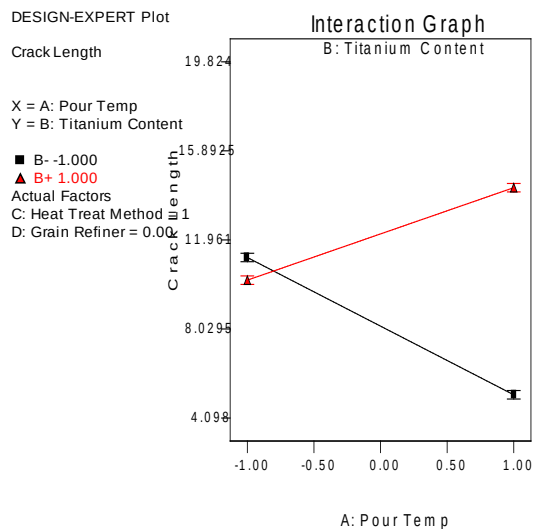
Final Equation in Terms of Coded Factors:

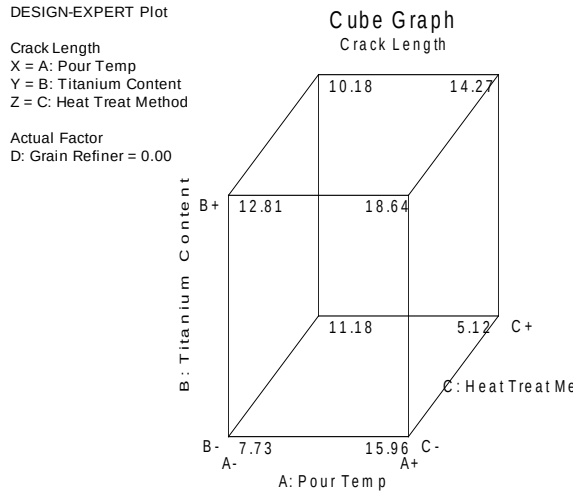
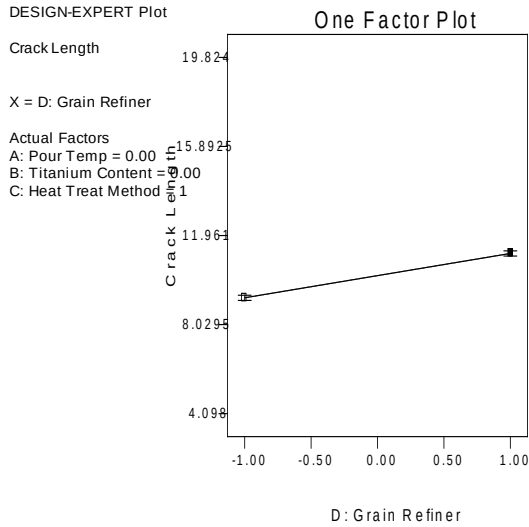
$$\text{Range} = +0.37 + 0.089 * A * B + 0.10 * C * D$$

- (f) What recommendations would you make regarding process operations? Use interaction and/or main effect plots to assist in drawing conclusions.

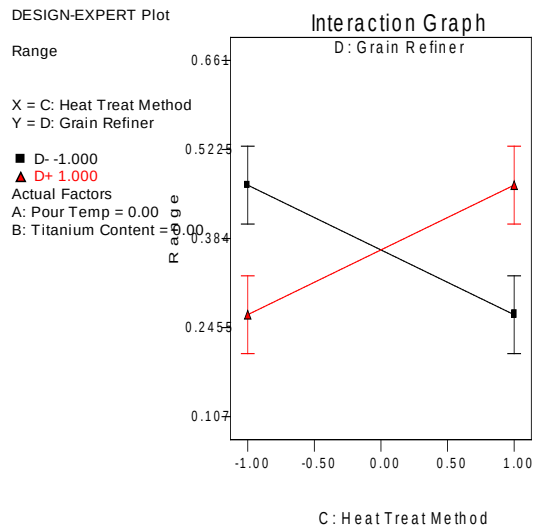
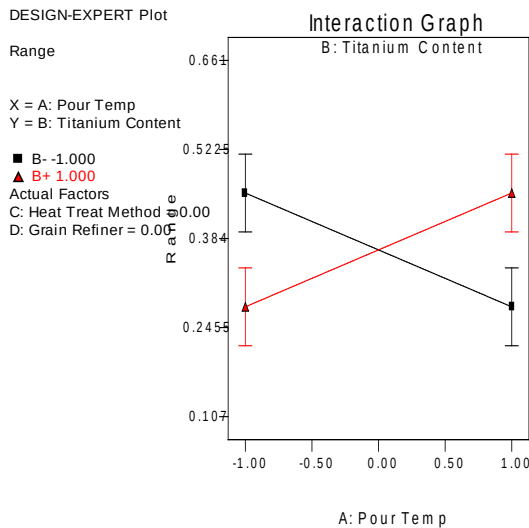
From the interaction plots, choose *A* at the high level and *B* at the low level. In each of these plots, *D* can be at either level. From the main effects plot of *C*, choose *C* at the high level. Based on the range analysis, with *C* at the high level, *D* should be set at the low level.

From the analysis of the crack length data:





From the analysis of the ranges:



6.16. Continuation of Problem 6.15. One of the variables in the experiment described in Problem 6.15, heat treatment method (c), is a categorical variable. Assume that the remaining factors are continuous.

- (a) Write two regression models for predicting crack length, one for each level of the heat treatment method variable. What differences, if any, do you notice in these two equations?

Design Expert Output

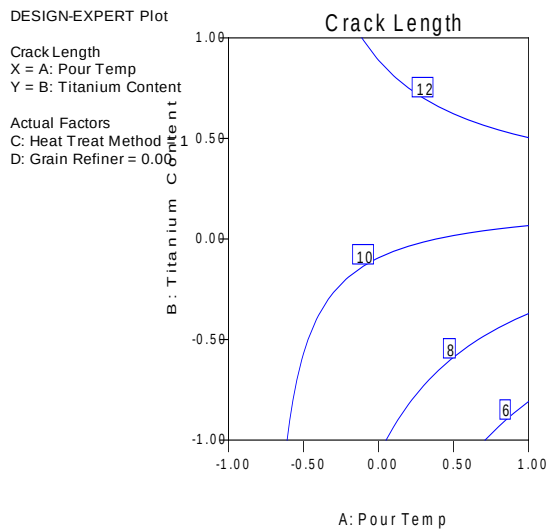
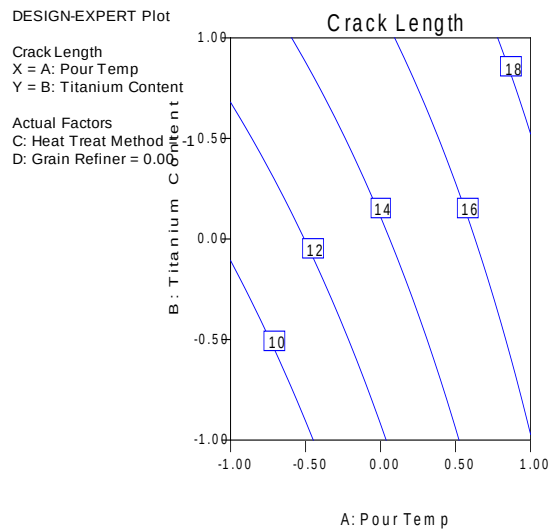
Final Equation in Terms of Actual Factors

Heat Treat Method -1
Crack Length =
+13.78619
+3.51331 * Pour Temp
+1.93994 * Titanium Content
+0.97888 * Grain Refiner

-0.60169 * Pour Temp * Titanium Content

Heat Treat Method 1
 Crack Length =
 +10.18994
 -0.49444 * Pour Temp
 +2.03594 * Titanium Content
 +0.97888 * Grain Refiner
 +2.53581 * Pour Temp * Titanium Content

(b) Generate appropriate response surface contour plots for the two regression models in part (a).



(c) What set of conditions would you recommend for the factors A , B and D if you use heat treatment method $C=+$?

High level of A , low level of B , and low level of D .

(d) Repeat part (c) assuming that you wish to use heat treatment method $C=-$.

Low level of A , low level of B , and low level of D .

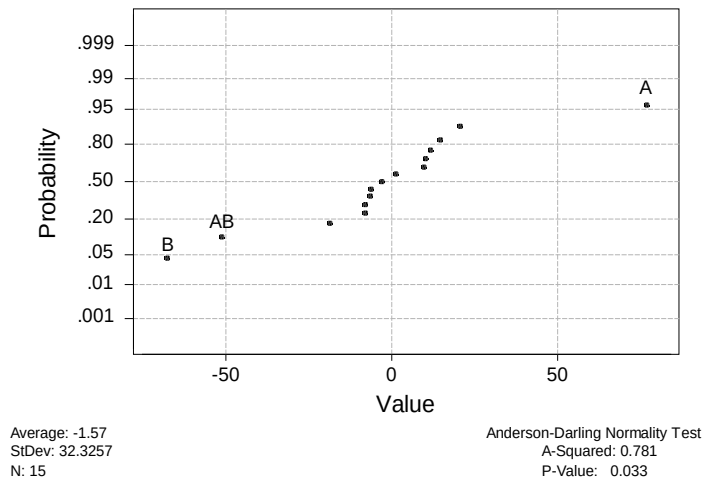
6.17. An experimenter has run a single replicate of a 2^4 design. The following effect estimates have been calculated:

$A = 76.95$	$AB = -51.32$	$ABC = -2.82$
$B = -67.52$	$AC = 11.69$	$ABD = -6.50$
$C = -7.84$	$AD = 9.78$	$ACD = 10.20$
$D = -18.73$	$BC = 20.78$	$BCD = -7.98$
	$BD = 14.74$	$ABCD = -6.25$
	$CD = 1.27$	

(a) Construct a normal probability plot of these effects.

The plot from Minitab follows.

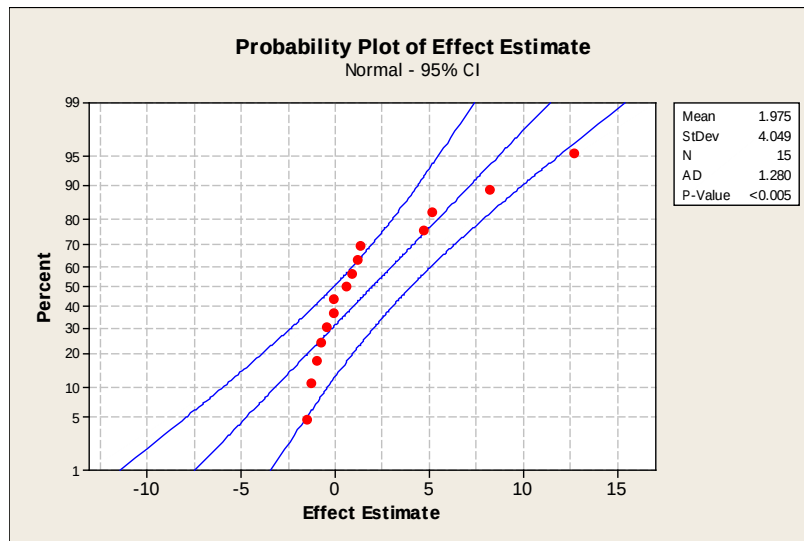
Normal Probability Plot



(b) Identify a tentative model, based on the plot of the effects in part (a).

$$\hat{\mu} = \text{Intercept} + 38.475x_A - 33.76x_B - 25.66x_Ax_B$$

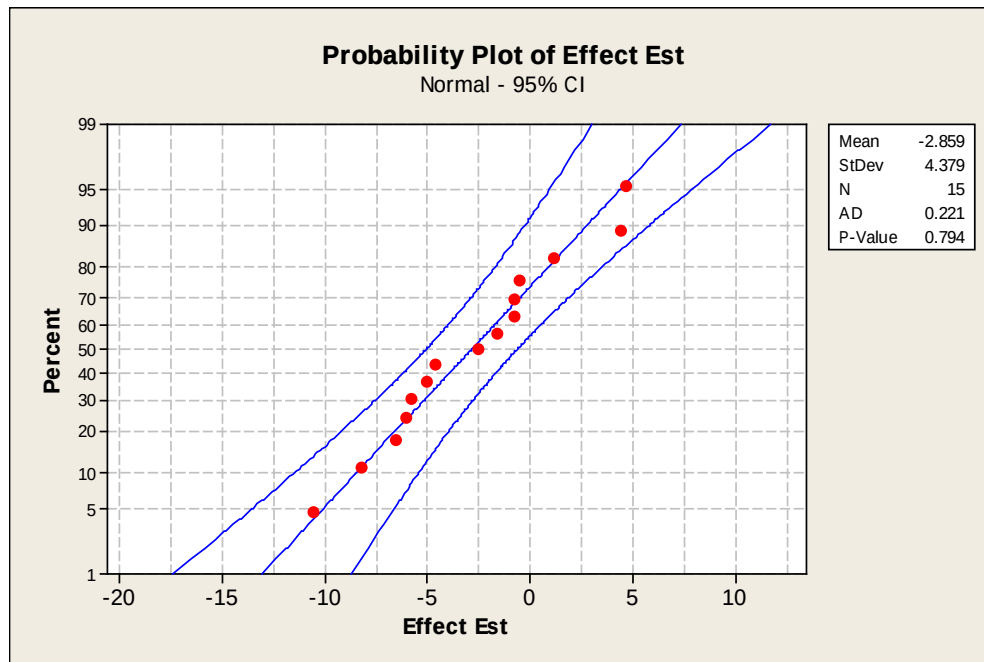
6.18. The effect estimates from a 2^4 factorial design are as follows: ABCD = -1.5138, ABC = -1.2661, ABD = -0.9852, ACD = -0.7566, BCD = -0.4842, CD = -0.0795, BD = -0.0793, AD = 0.5988, BC = 0.9216, AC = 1.1616, AB = 1.3266, D = 4.6744, C = 5.1458, B = 8.2469, and A = 12.7151. Are you comfortable with the conclusions that all main effects are active?



The upper right 4 dots are the four main effects. Since they do not follow the rest of the data, the normal probability plot shows that they are active.

6.19. The effect estimates from a 2^4 factorial experiment are listed here. Are any of the effects significant?

ABCD = -2.5251	AD = -1.6564
BCD = 4.4054	AC = 1.1109
ACD = -0.4932	AB = -10.5229
ABD = -5.0842	D = -6.0275
ABC = -5.7696	C = -8.2045
CD = 4.6707	B = -6.5304
BD = -4.6620	A = -0.7914
BC = -0.7982	



No effects appear to be significant.

6.20. Consider a variation of the bottle filling experiment from Example 5.3. Suppose that only two levels of carbonation are used so that the experiment is a 2^3 factorial design with two replicates. The data are shown below.

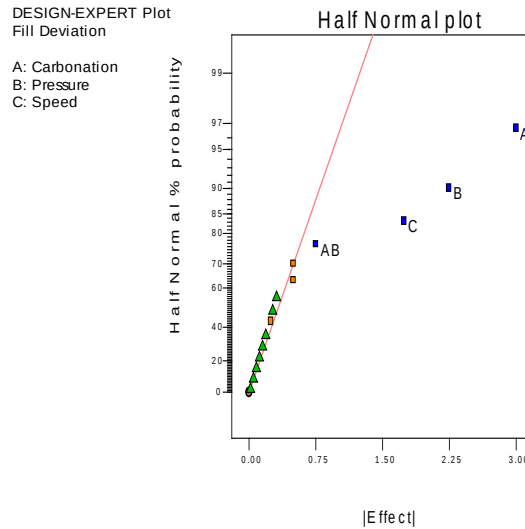
Run	Coded Factors			Fill Height Deviation	
	A	B	C	Replicate 1	Replicate 2
1	-	-	-	-3	-1
2	+	-	-	0	1
3	-	+	-	-1	0
4	+	+	-	2	3
5	-	-	+	-1	0
6	+	-	+	2	1
7	-	+	+	1	1
8	+	+	+	6	5

A (%)	Factor Levels	
	Low (-1)	High (+1)
	10	12

<i>B</i> (psi)	25	30
<i>C</i> (b/m)	200	250

(a) Analyze the data from this experiment. Which factors significantly affect fill height deviation?

The half normal probability plot of effects shown below identifies the factors *A*, *B*, and *C* as being significant and the *AB* interaction as being marginally significant. The analysis of variance in the Design Expert output below confirms that factors *A*, *B*, and *C* are significant and the *AB* interaction is marginally significant.

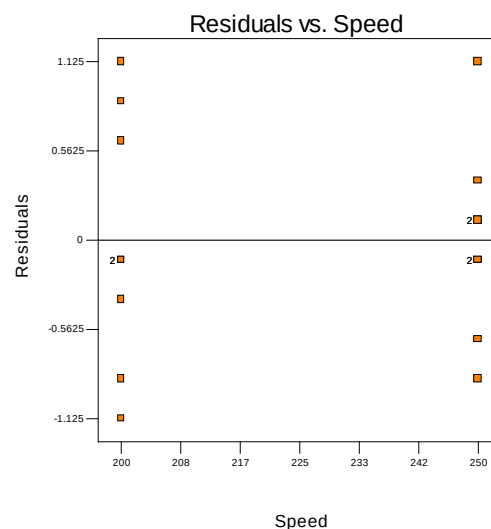
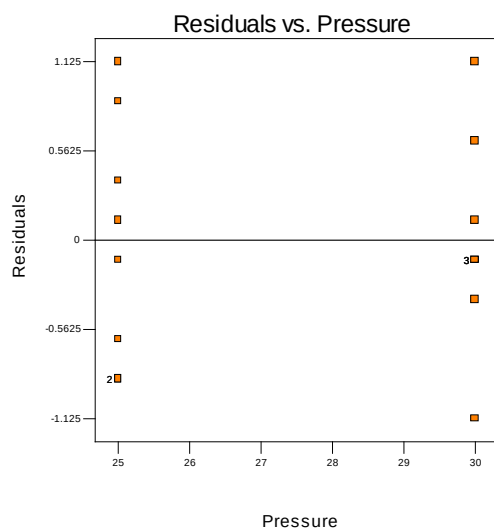
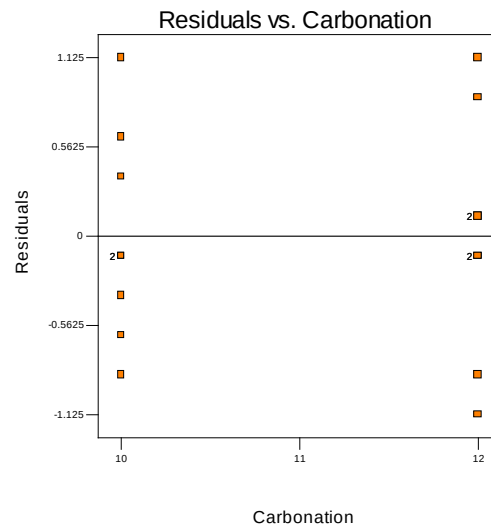
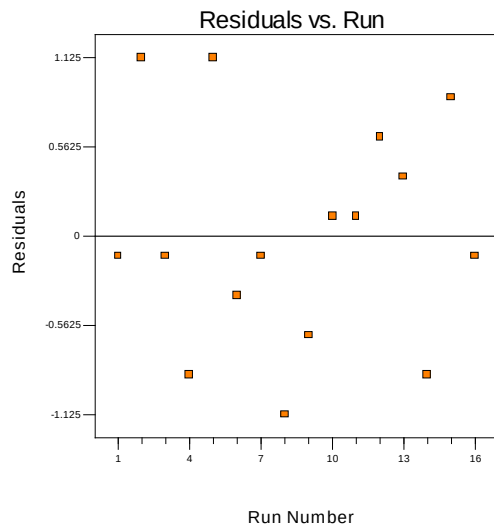
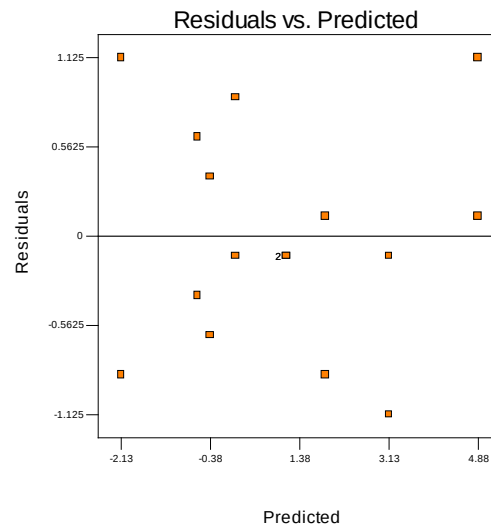
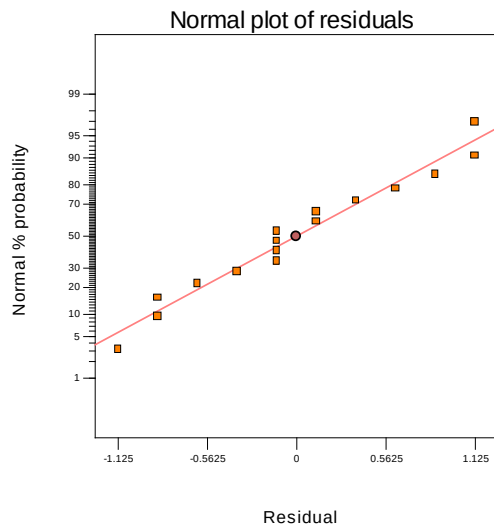


Design Expert Output

Response: Fill Deviation						
ANOVA for Selected Factorial Model						
Analysis of variance table [Partial sum of squares]						
Source	Sum of Squares	DF	Mean Square	F Value	Prob > F	
Model	70.75	4	17.69	26.84	< 0.0001	significant
<i>A</i>	36.00	1	36.00	54.62	< 0.0001	
<i>B</i>	20.25	1	20.25	30.72	0.0002	
<i>C</i>	12.25	1	12.25	18.59	0.0012	
<i>AB</i>	2.25	1	2.25	3.41	0.0917	
Residual	7.25	11	0.66			
<i>Lack of Fit</i>	2.25	3	0.75	1.20	0.3700	not significant
<i>Pure Error</i>	5.00	8	0.63			
Cor Total	78.00	15				
The Model F-value of 26.84 implies the model is significant. There is only a 0.01% chance that a "Model F-Value" this large could occur due to noise.						
Values of "Prob > F" less than 0.0500 indicate model terms are significant. In this case <i>A</i> , <i>B</i> , <i>C</i> are significant model terms.						
Std. Dev.	0.81	R-Squared	0.9071			
Mean	1.00	Adj R-Squared	0.8733			
C.V.	81.18	Pred R-Squared	0.8033			
PRESS	15.34	Adeq Precision	15.424			

(b) Analyze the residuals from this experiment. Are there any indications of model inadequacy?

The residual plots below do not identify any violations to the assumptions.



- (c) Obtain a model for predicting fill height deviation in terms of the important process variables. Use this model to construct contour plots to assist in interpreting the results of the experiment.

The model in both coded and actual factors are shown below.

Design Expert Output

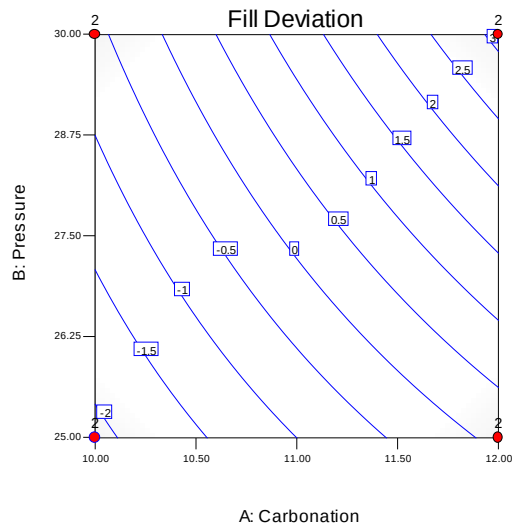
Final Equation in Terms of Coded Factors

$$\begin{aligned} \text{Fill Deviation} = & \\ & +1.00 \\ & +1.50 * A \\ & +1.13 * B \\ & +0.88 * C \\ & +0.38 * A * B \end{aligned}$$

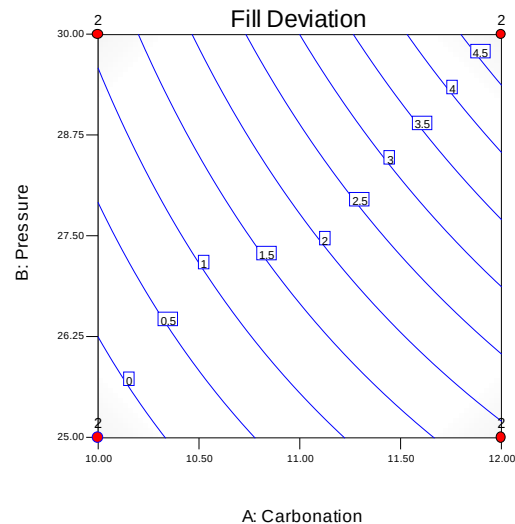
Final Equation in Terms of Actual Factors

$$\begin{aligned} \text{Fill Deviation} = & \\ & +9.62500 \\ & -2.62500 * \text{Carbonation} \\ & -1.20000 * \text{Pressure} \\ & +0.035000 * \text{Speed} \\ & +0.15000 * \text{Carbonation} * \text{Pressure} \end{aligned}$$

The following contour plots identify the fill deviation with respect to carbonation and pressure. The plot on the left sets the speed at 200 b/m while the plot on the right sets the speed at 250 b/m. Assuming a faster bottle speed is better, settings in pressure and carbonation that produce a fill deviation near zero can be found in the lower left hand corner of the contour plot on the right.



Speed set at 200 b/m



Speed set at 250 b/m

- (d) In part (a), you probably noticed that there was an interaction term that was borderline significant. If you did not include the interaction term in your model, include it now and repeat the analysis. What difference did this make? If you elected to include the interaction term in part (a), remove it and repeat the analysis. What difference does this make?

The following analysis of variance, residual plots, and contour plots represent the model without the interaction. As in the original analysis, the residual plots do not identify any concerns with the assumptions. The contour plots did not change significantly either. The interaction effect is small relative to the main effects.

Design Expert Output

Response: Fill Deviation

ANOVA for Selected Factorial Model

Analysis of variance table [Partial sum of squares]

Source	Sum of Squares	DF	Mean Square	F Value	Prob > F	
Model	68.50	3	22.83	28.84	< 0.0001	significant
A	36.00	1	36.00	45.47	< 0.0001	
B	20.25	1	20.25	25.58	0.0003	
C	12.25	1	12.25	15.47	0.0020	
Residual	9.50	12	0.79			
Lack of Fit	4.50	4	1.13	1.80	0.2221	not significant
Pure Error	5.00	8	0.63			
Cor Total	78.00	15				

The Model F-value of 28.84 implies the model is significant. There is only a 0.01% chance that a "Model F-Value" this large could occur due to noise.

Values of "Prob > F" less than 0.0500 indicate model terms are significant. In this case A, B, C are significant model terms.

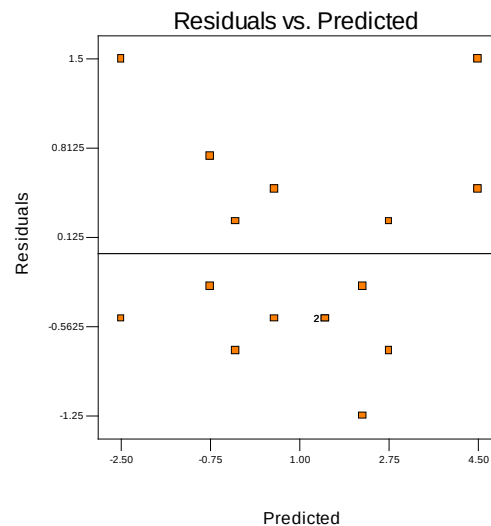
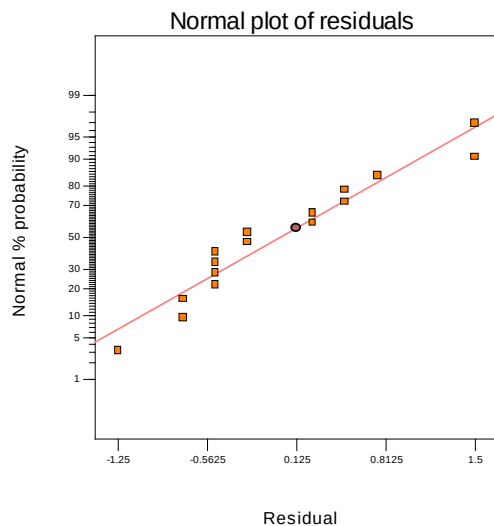
Std. Dev.	0.89	R-Squared	0.8782
Mean	1.00	Adj R-Squared	0.8478
C.V.	88.98	Pred R-Squared	0.7835
PRESS	16.89	Adeq Precision	15.735

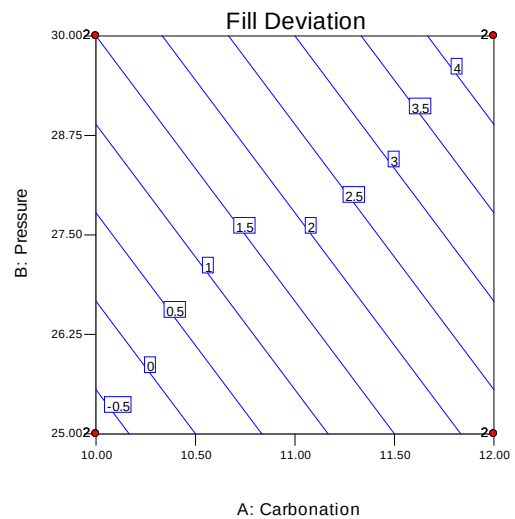
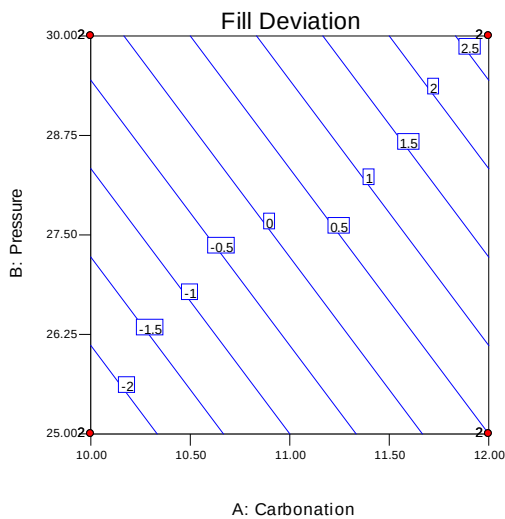
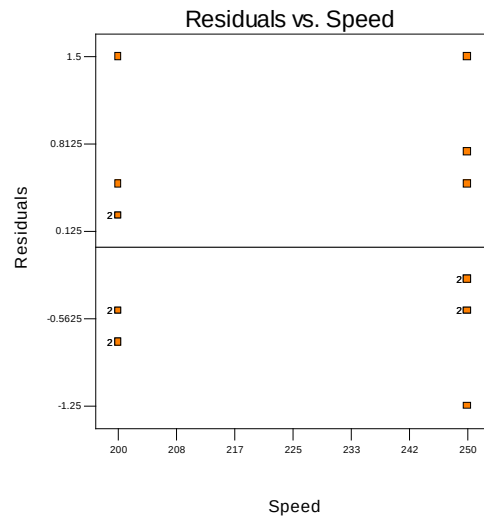
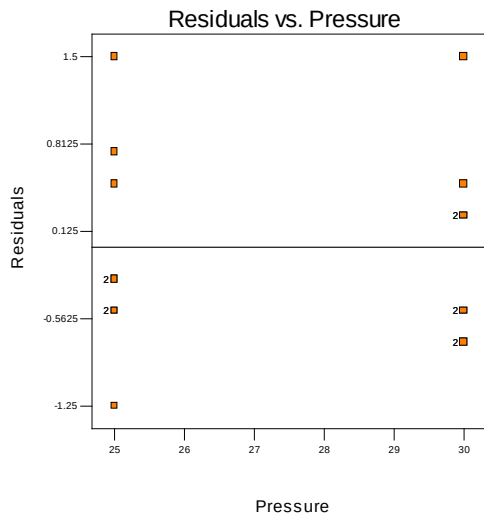
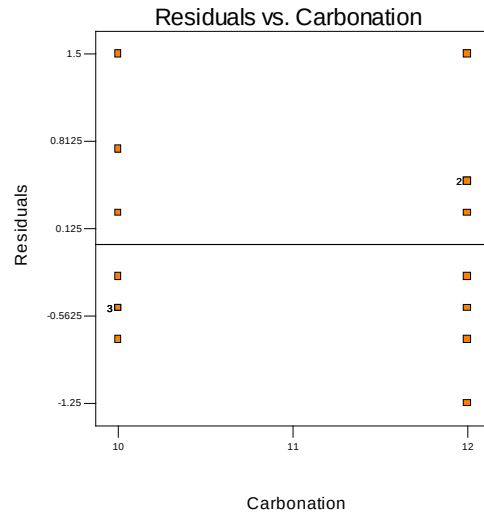
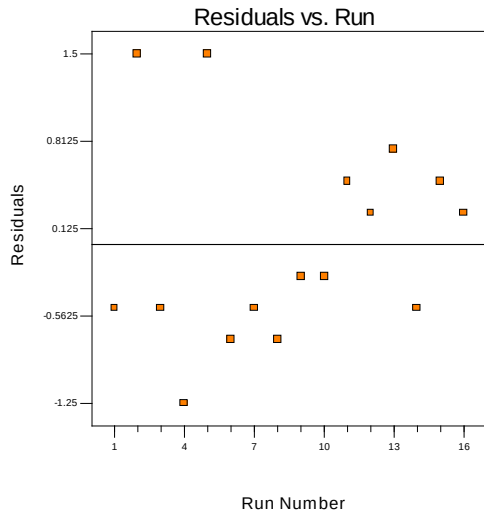
Final Equation in Terms of Coded Factors

$$\begin{aligned} \text{Fill Deviation} = & \\ & +1.00 \\ & +1.50 * A \\ & +1.13 * B \\ & +0.88 * C \end{aligned}$$

Final Equation in Terms of Actual Factors

$$\begin{aligned} \text{Fill Deviation} = & \\ & -35.75000 \\ & +1.50000 * \text{Carbonation} \\ & +0.45000 * \text{Pressure} \\ & +0.035000 * \text{Speed} \end{aligned}$$



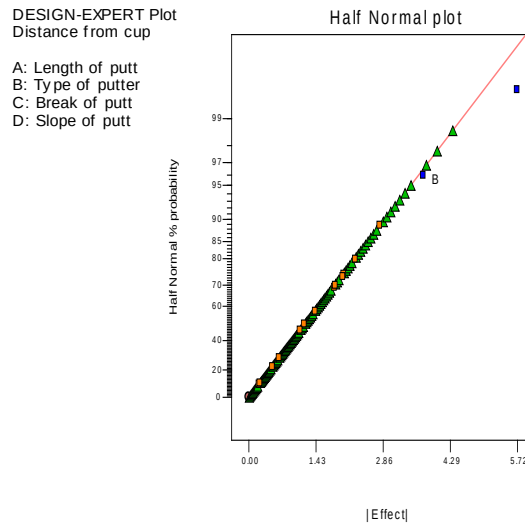


6.21. I am always interested in improving my golf scores. Since a typical golfer uses the putter for about 35-45% of his or her strokes, it seems logical that in improving one's putting score is a logical and perhaps simple way to improve a golf score ("The man who can putt is a match for any man." – Willie Parks, 1864-1925, two-time winner of the British Open). An experiment was conducted to study the effects of four factors on putting accuracy. The design factors are length of putt, type of putter, breaking putt vs. straight putt, and level versus downhill putt. The response variable is distance from the ball to the center of the cup after the ball comes to rest. One golfer performs the experiment, a 2^4 factorial design with seven replicates was used, and all putts were made in random order. The results are as follows.

Design Factors				Distance from cup (replicates)						
Length of putt (ft)	Type of putter	Break of putt	Slope of putt	1	2	3	4	5	6	7
10	Mallet	Straight	Level	10.0	18.0	14.0	12.5	19.0	16.0	18.5
30	Mallet	Straight	Level	0.0	16.5	4.5	17.5	20.5	17.5	33.0
10	Cavity-back	Straight	Level	4.0	6.0	1.0	14.5	12.0	14.0	5.0
30	Cavity-back	Straight	Level	0.0	10.0	34.0	11.0	25.5	21.5	0.0
10	Mallet	Breaking	Level	0.0	0.0	18.5	19.5	16.0	15.0	11.0
30	Mallet	Breaking	Level	5.0	20.5	18.0	20.0	29.5	19.0	10.0
10	Cavity-back	Breaking	Level	6.5	18.5	7.5	6.0	0.0	10.0	0.0
30	Cavity-back	Breaking	Level	16.5	4.5	0.0	23.5	8.0	8.0	8.0
10	Mallet	Straight	Downhill	4.5	18.0	14.5	10.0	0.0	17.5	6.0
30	Mallet	Straight	Downhill	19.5	18.0	16.0	5.5	10.0	7.0	36.0
10	Cavity-back	Straight	Downhill	15.0	16.0	8.5	0.0	0.5	9.0	3.0
30	Cavity-back	Straight	Downhill	41.5	39.0	6.5	3.5	7.0	8.5	36.0
10	Mallet	Breaking	Downhill	8.0	4.5	6.5	10.0	13.0	41.0	14.0
30	Mallet	Breaking	Downhill	21.5	10.5	6.5	0.0	15.5	24.0	16.0
10	Cavity-back	Breaking	Downhill	0.0	0.0	0.0	4.5	1.0	4.0	6.5
30	Cavity-back	Breaking	Downhill	18.0	5.0	7.0	10.0	32.5	18.5	8.0

(a) Analyze the data from this experiment. Which factors significantly affect putting performance?

The half normal probability plot of effects identifies only factors *A* and *B*, length of putt and type of putter, as having a potentially significant affect on putting performance. The analysis of variance with only these significant factors is presented as well and confirms significance.

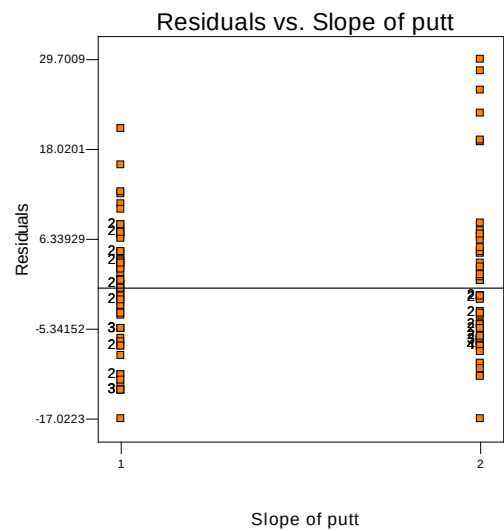
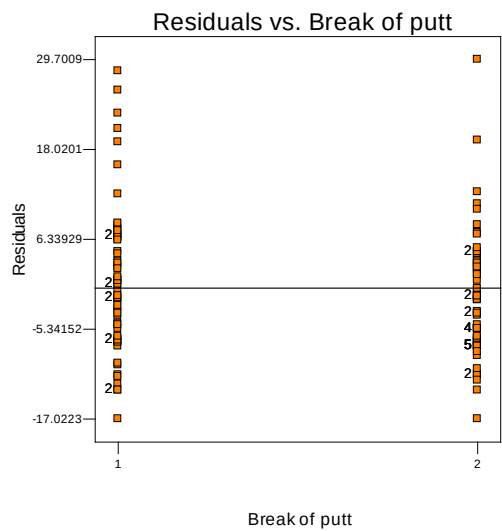
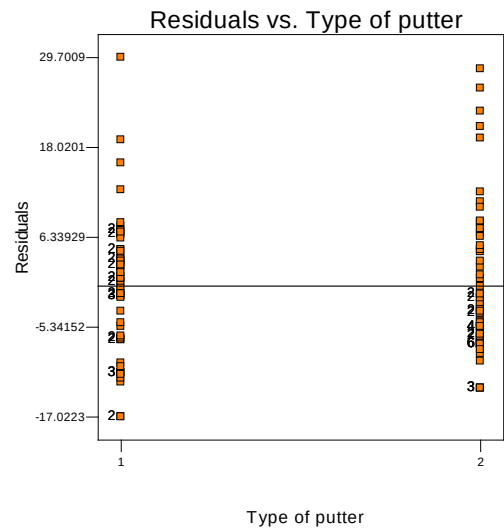
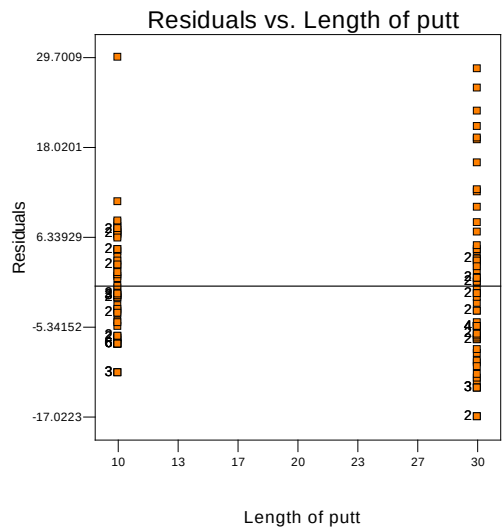
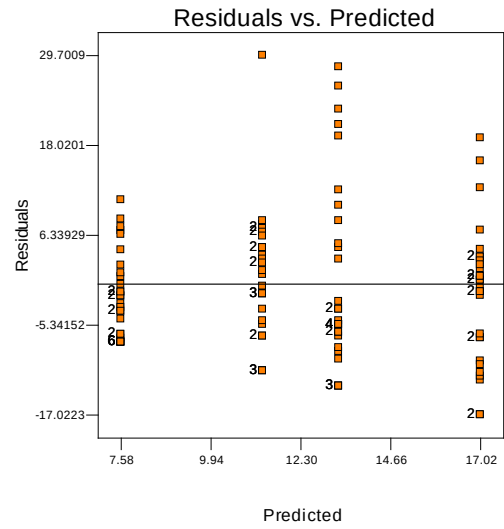
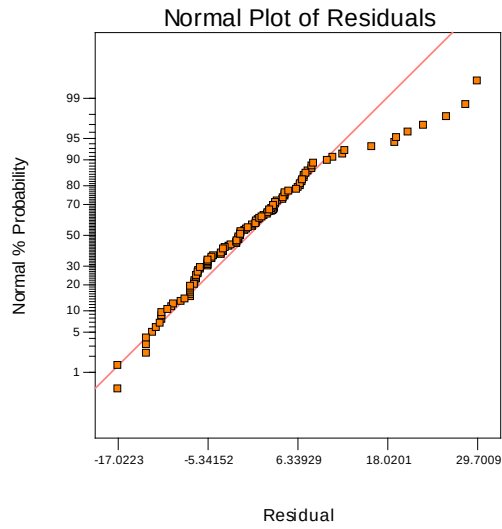


Design Expert Output with Only Factors A and B

Response: Distance from cup						
ANOVA for Selected Factorial Model						
Analysis of variance table [Terms added sequentially (first to last)]						
Source	Sum of Squares	DF	Mean Square	F Value	Prob > F	
Model	1305.29	2	652.65	7.69	0.0008	significant
A	917.15	1	917.15	10.81	0.0014	
B	388.15	1	388.15	4.57	0.0347	
Residual	9248.94	109	84.85			
Lack of Fit	933.15	13	71.78	0.83	0.6290	not significant
Pure Error	8315.79	96	86.62			
Cor Total	10554.23	111				
The Model F-value of 7.69 implies the model is significant. There is only a 0.08% chance that a "Model F-Value" this large could occur due to noise.						
Values of "Prob > F" less than 0.0500 indicate model terms are significant. In this case A, B, are significant model terms.						
Std. Dev.	9.21	R-Squared	0.1237			
Mean	12.30	Adj R-Squared	0.1076			
C.V.	74.90	Pred R-Squared	0.0748			
PRESS	9765.06	Adeq Precision	6.266			

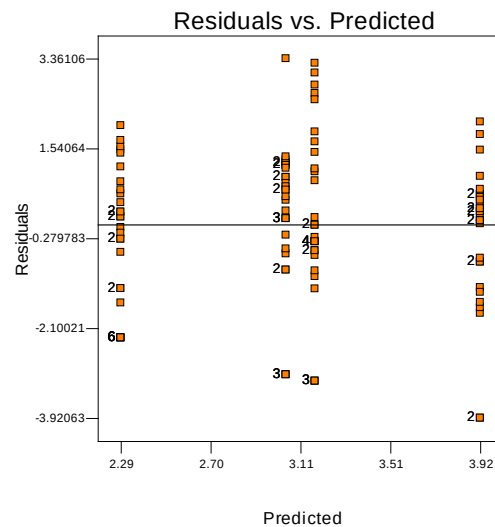
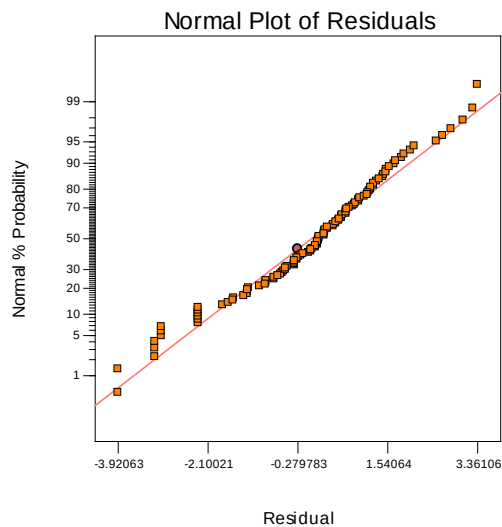
(b) Analyze the residuals from this experiment. Are there any indications of model inadequacy?

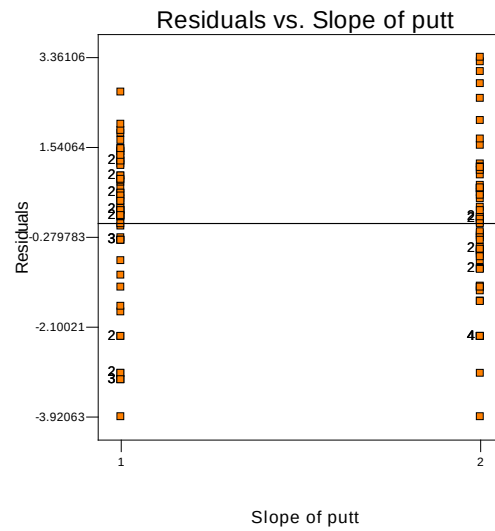
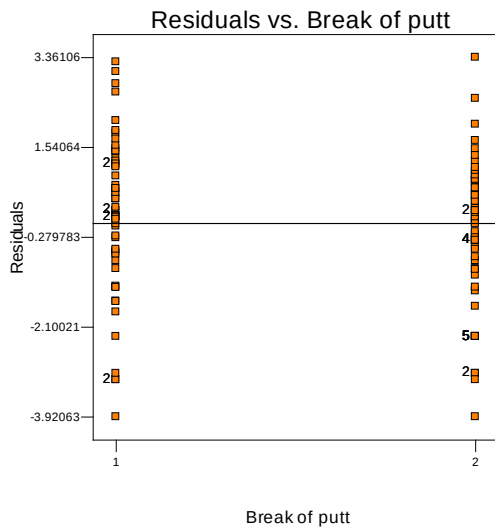
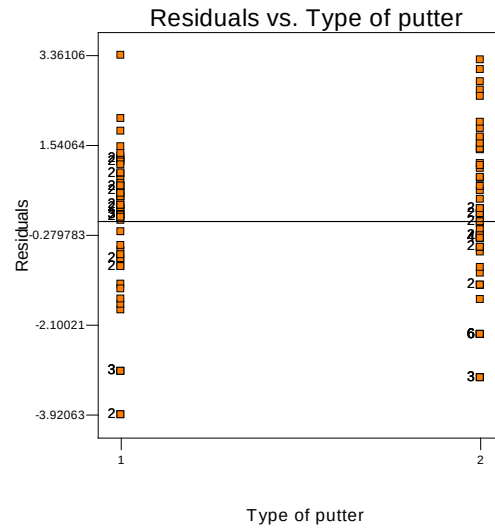
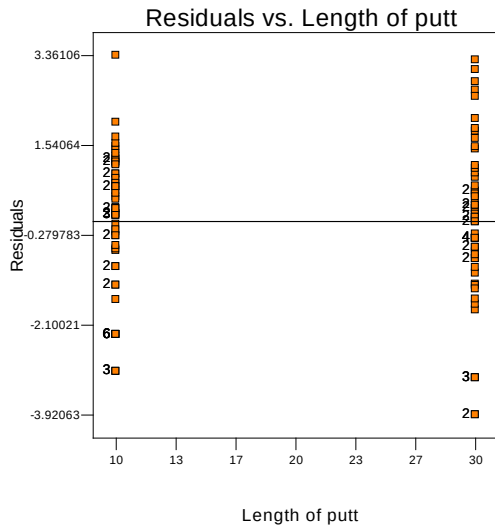
The residual plots for the model containing only the significant factors A and B are shown below. The normality assumption appears to be violated. Also, as a golfer might expect, there is a slight inequality of variance with regards to the length of putt. A square root transformation is applied which corrects the violations. The analysis of variance and corrected residual plots are also presented. Finally, an effects plot identifies a 10 foot putt and the cavity-back putter reduce the mean distance from the cup.



Design Expert Output with Only Factors A and B and a Square Root Transformation

Response:		Distance from cup		Transform: Square root		Constant:		0	
ANOVA for Selected Factorial Model									
Analysis of variance table [Terms added sequentially (first to last)]									
Source	Sum of Squares	DF	Mean Square	F Value	Prob > F				
Model	37.26	2	18.63	7.85	0.0007	significant			
A	21.61	1	21.61	9.11	0.0032				
B	15.64	1	15.64	6.59	0.0116				
Residual	258.63	109	2.37						
Lack of Fit	30.19	13	2.32	0.98	0.4807	not significant			
Pure Error	228.45	96	2.38						
Cor Total	295.89	111							
The Model F-value of 7.85 implies the model is significant. There is only a 0.07% chance that a "Model F-Value" this large could occur due to noise.									
Values of "Prob > F" less than 0.0500 indicate model terms are significant. In this case A, B, are significant model terms.									
Std. Dev.	1.54	R-Squared	0.1259						
Mean	3.11	Adj R-Squared	0.1099						
C.V.	49.57	Pred R-Squared	0.0771						
PRESS	273.06	Adeq Precision	6.450						





DESIGN-EXPERT Plot

Sqrt(Distance from cup)

X = A: Length of putt

Y = B: Type of putter

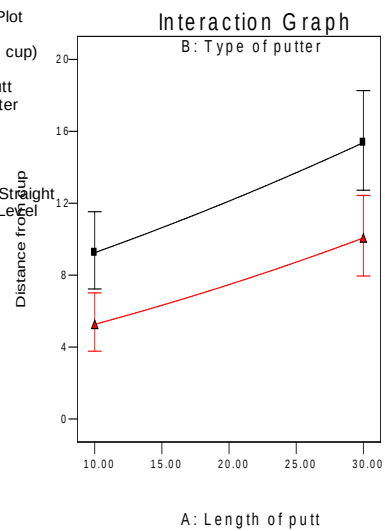
■ B1 Mallet

▲ B2 Cavity-back

Actual Factors

C: Break of putt = Straight

D: Slope of putt = Level

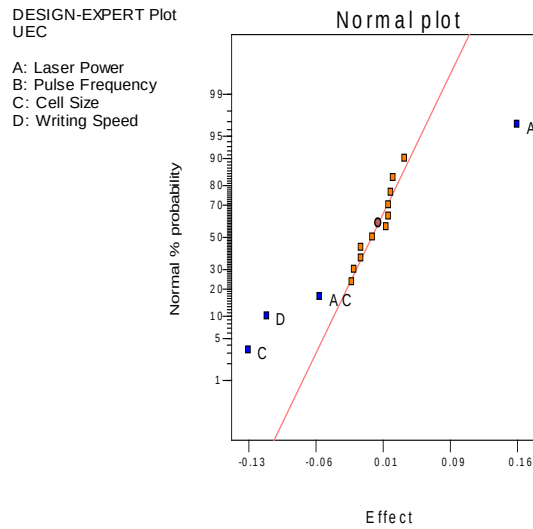


6.22. Semiconductor manufacturing processes have long and complex assembly flows, so matrix marks and automated 2d-matrix readers are used at several process steps throughout factories. Unreadable matrix marks negatively effect factory run rates, because manual entry of part data is required before manufacturing can resume. A 2^4 factorial experiment was conducted to develop a 2d-matrix laser mark on a metal cover that protects a substrate mounted die. The design factors are A = laser power (9W, 13W), B = laser pulse frequency (4000 Hz, 12000 Hz), C = matrix cell size (0.07 in, 0.12 in), and D = writing speed (10 in/sec, 20 in/sec), and the response variable is the unused error correction (UEC). This is a measure of the unused portion of the redundant information embedded in the 2d matrix. A UEC of 0 represents the lowest reading that still results in a decodable matrix while a value of 1 is the highest reading. A DMX Verifier was used to measure UEC. The data from this experiment are shown below.

Standard Order	Run Order	Laser Power	Pulse Frequency	Cell Size	Writing Speed	UEC
8	1	1	1	1	-1	0.80
10	2	1	-1	-1	1	0.81
12	3	1	1	-1	1	0.79
9	4	-1	-1	-1	1	0.60
7	5	-1	1	1	-1	0.65
15	6	-1	1	1	1	0.55
2	7	1	-1	-1	-1	0.98
6	8	1	-1	1	-1	0.67
16	9	1	1	1	1	0.69
13	10	-1	-1	1	1	0.56
5	11	-1	-1	1	-1	0.63
14	12	1	-1	1	1	0.65
1	13	-1	-1	-1	-1	0.75
3	14	-1	1	-1	-1	0.72
4	15	1	1	-1	-1	0.98
11	16	-1	1	-1	1	0.63

(a) Analyze the data from this experiment. Which factors significantly affect UEC?

The normal probability plot of effects identifies A , C , D , and the AC interaction as significant. The Design Expert output including the analysis of variance confirms the significance and identifies the corresponding model. Contour plots identify factors A and C with B held constant at zero and D toggled from -1 to +1.



Design Expert Output

Response: UEC

ANOVA for Selected Factorial Model

Analysis of variance table [Terms added sequentially (first to last)]

Source	Sum of Squares	DF	Mean Square	F Value	Prob > F	
Model	0.24	4	0.059	35.51	< 0.0001	significant
A	0.10	1	0.10	61.81	< 0.0001	
C	0.070	1	0.070	42.39	< 0.0001	
D	0.051	1	0.051	30.56	0.0002	
AC	0.012	1	0.012	7.30	0.0206	
Residual	0.018	11	1.657E-003			
Cor Total	0.25	15				

The Model F-value of 35.51 implies the model is significant. There is only a 0.01% chance that a "Model F-Value" this large could occur due to noise.

Values of "Prob > F" less than 0.0500 indicate model terms are significant. In this case A, C, D, AC are significant model terms.

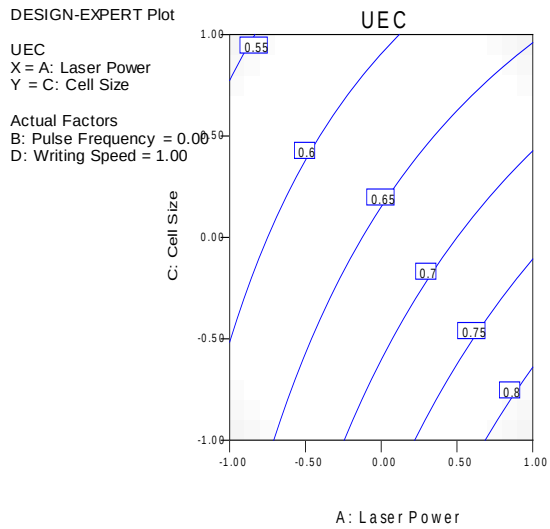
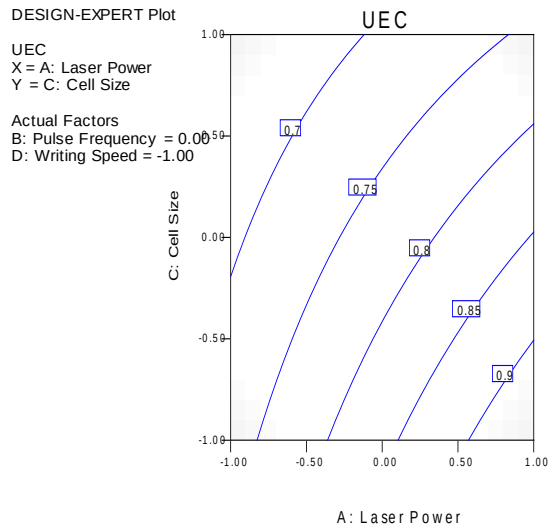
Std. Dev.	0.041	R-Squared	0.9281
Mean	0.72	Adj R-Squared	0.9020
C.V.	5.68	Pred R-Squared	0.8479
PRESS	0.039	Adeq Precision	17.799

Final Equation in Terms of Coded Factors

$$\begin{aligned} \text{UEC} = & \\ & +0.72 \\ & +0.080 * A \\ & -0.066 * C \\ & -0.056 * D \\ & -0.027 * A * C \end{aligned}$$

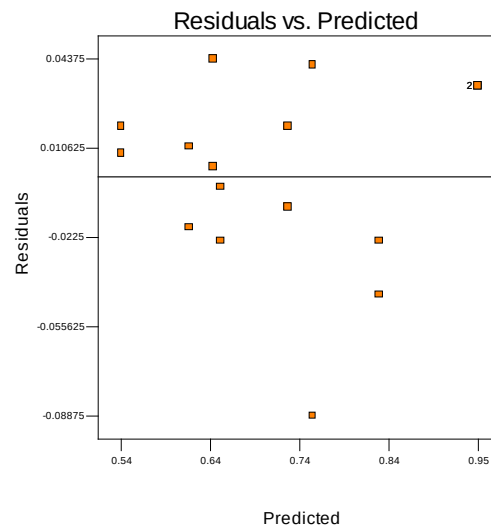
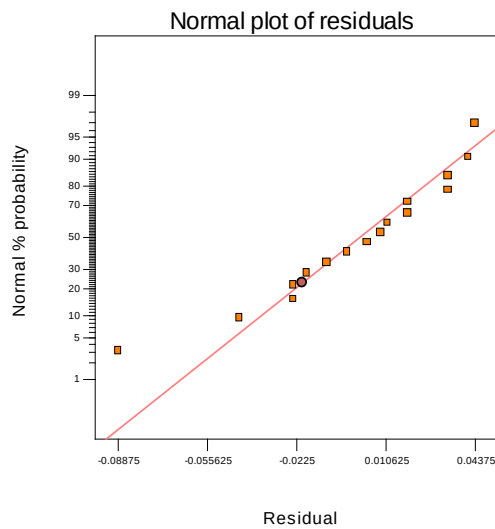
Final Equation in Terms of Actual Factors

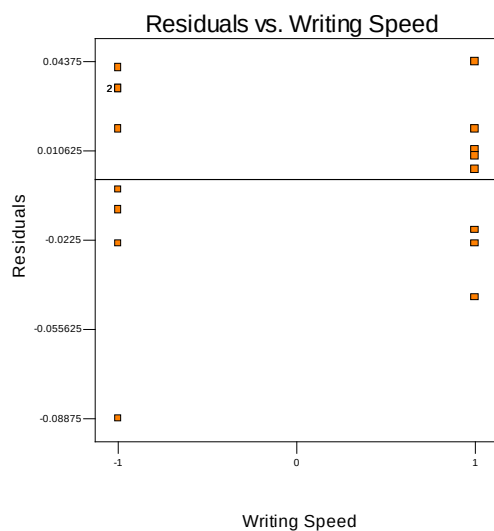
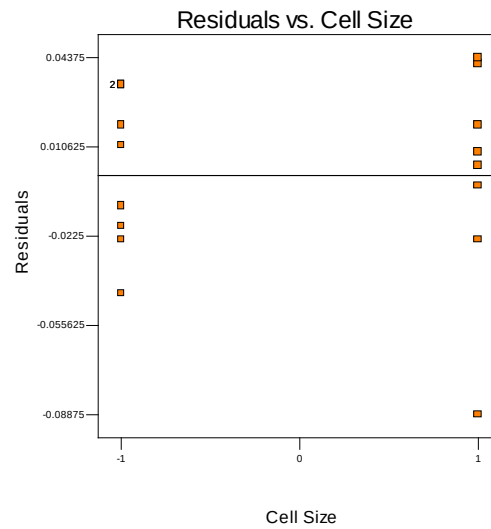
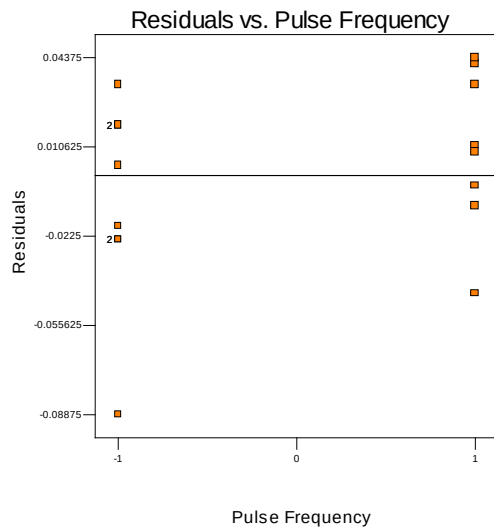
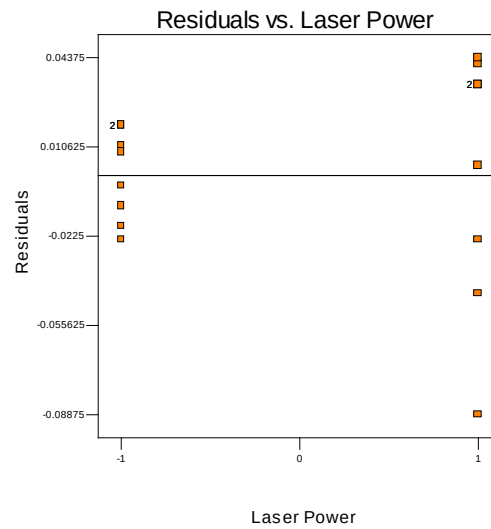
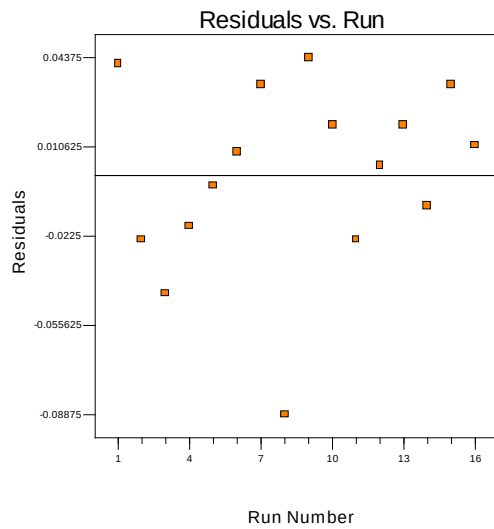
$$\begin{aligned} \text{UEC} = & \\ & +0.71625 \\ & +0.080000 * \text{Laser Power} \\ & -0.066250 * \text{Cell Size} \\ & -0.056250 * \text{Writing Speed} \\ & -0.027500 * \text{Laser Power} * \text{Cell Size} \end{aligned}$$



(b) Analyze the residuals from this experiment. Are there any indications of model inadequacy?

The residual plots appear acceptable with the exception of run 8, standard order 6. This value should be verified by the engineer.





6.23. Reconsider the experiment in Problem 6.22. Suppose that four center points are available, and the UEC response at these four runs is 0.98, 0.95, 0.93 and 0.96, respectively. Reanalyze the experiment incorporating a test for curvature into the analysis. What conclusions can you draw? What recommendations would you make to the experimenters?

As with the results of problem 6.20, factors *A*, *C*, *D*, and the *AC* interaction remain significant. However, the *CD* interaction and curvature are significant as well. The curvature is the strongest effect; unfortunately, we are not able to determine which factor(s) have a quadratic term. We recommend that the engineer augment the experiment with additional experimental runs, such as axial points and a couple of extra center points for blocking purposes. These extra runs will determine the pure quadratic effects and allow us to fit a second order model.

Design Expert Output

Response: UEC						
ANOVA for Selected Factorial Model						
Analysis of variance table [Terms added sequentially (first to last)]						
Source	Sum of Squares	DF	Mean Square	F Value	Prob > F	
Model	0.24	5	0.048	45.07	< 0.0001	significant
A	0.10	1	0.10	95.77	< 0.0001	
C	0.070	1	0.070	65.68	< 0.0001	
D	0.051	1	0.051	47.35	< 0.0001	
AC	0.012	1	0.012	11.32	0.0051	
CD	5.625E-003	1	5.625E-003	5.26	0.0391	
Curvature	0.18	1	0.18	170.59	< 0.0001	significant
Residual	0.014	13	1.069E-003			
Lack of Fit	0.013	10	1.260E-003	2.91	0.2057	not significant
Pure Error	1.300E-003	3	4.333E-004			
Cor Total	0.44	19				
The Model F-value of 45.07 implies the model is significant. There is only a 0.01% chance that a "Model F-Value" this large could occur due to noise.						
Values of "Prob > F" less than 0.0500 indicate model terms are significant. In this case A, C, D, AC, CD are significant model terms.						
Std. Dev.	0.033	R-Squared	0.9455			
Mean	0.76	Adj R-Squared	0.9245			
C.V.	4.28	Pred R-Squared	0.9209			
PRESS	0.035	Adeq Precision	20.936			
Coefficient Factor	Standard Estimate	95% CI DF	95% CI Error	Low	High	VIF
Intercept	0.72	1	8.175E-003	0.70	0.73	
A-Laser Power	0.080	1	8.175E-003	0.062	0.098	1.00
C-Cell Size	-0.066	1	8.175E-003	-0.084	-0.049	1.00
D-Writing Speed	-0.056	1	8.175E-003	-0.074	-0.039	1.00
AC	-0.028	1	8.175E-003	-0.045	-9.839E-003	1.00
CD	0.019	1	8.175E-003	1.089E-003	0.036	1.00
Center Point	0.24	1	0.018	0.20	0.28	1.00

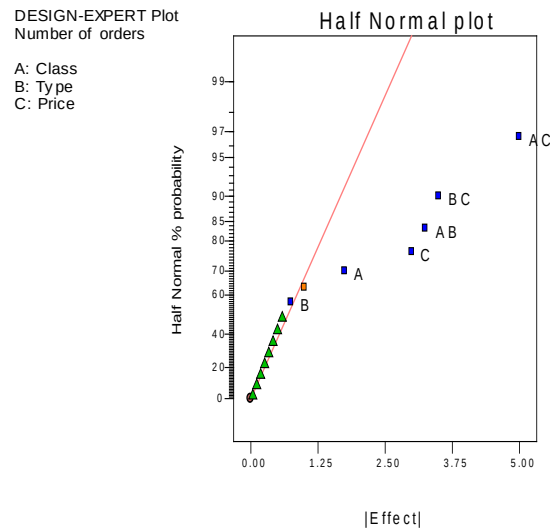
6.24. A company markets its products by direct mail. An experiment was conducted to study the effects of three factors on the customer response rate for a particular product. The three factors are *A* = type of mail used (3rd class, 1st class), *B* = type of descriptive brochure (color, black-and-white), and *C* = offer price (\$19.95, \$24.95). The mailings are made to two groups of 8,000 randomly selected customers, with 1,000 customers in each group receiving each treatment combination. Each group of customers is considered as a replicate. The response variable is the number of orders placed. The experimental data is shown below.

Run	Coded Factors			Number of Orders	
	A	B	C	Replicate 1	Replicate 2
1	-	-	-	50	54
2	+	-	-	44	42
3	-	+	-	46	48
4	+	+	-	42	43
5	-	-	+	49	46
6	+	-	+	48	45
7	-	+	+	47	48
8	+	+	+	56	54

	Factor Levels	
	Low (-1)	High (+1)
A (class)	3rd	1st
B (type)	BW	Color
C (\$)	\$19.95	\$24.95

(a) Analyze the data from this experiment. Which factors significantly affect the customer response rate?

The half normal probability plot of effects identifies the two factor interactions, AB , AC , BC , and factors A and C as significant. Factor B is not significant; however, remains in the model to satisfy the hierarchical principle. The analysis of variance confirms the significance of two factor interactions and factor C . Factor A is marginally significant.



Design Expert Output

Response: Number of orders**ANOVA for Selected Factorial Model****Analysis of variance table [Partial sum of squares]**

Source	Sum of Squares	DF	Mean Square	F Value	Prob > F	
Model	241.75	6	40.29	12.95	0.0006	significant
A	12.25	1	12.25	3.94	0.0785	
B	2.25	1	2.25	0.72	0.4171	
C	36.00	1	36.00	11.57	0.0079	
AB	42.25	1	42.25	13.58	0.0050	
AC	100.00	1	100.00	32.14	0.0003	
BC	49.00	1	49.00	15.75	0.0033	
Residual	28.00	9	3.11			
Lack of Fit	4.00	1	4.00	1.33	0.2815	not significant
Pure Error	24.00	8	3.00			
Cor Total	269.75	15				

The Model F-value of 12.95 implies the model is significant. There is only a 0.06% chance that a "Model F-Value" this large could occur due to noise.

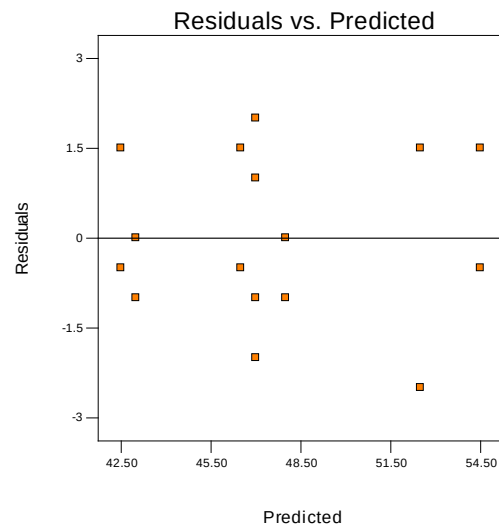
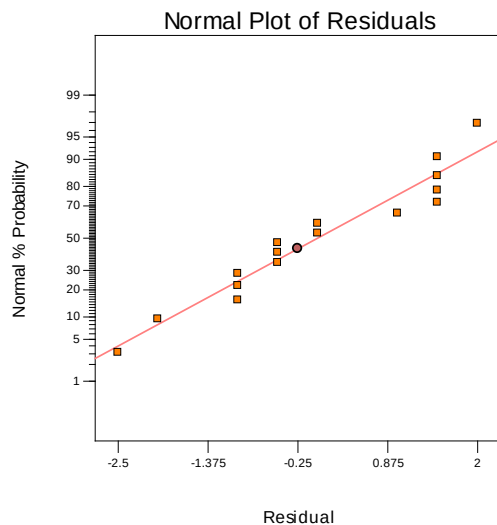
Values of "Prob > F" less than 0.0500 indicate model terms are significant. In this case C, AB, AC, BC are significant model terms.

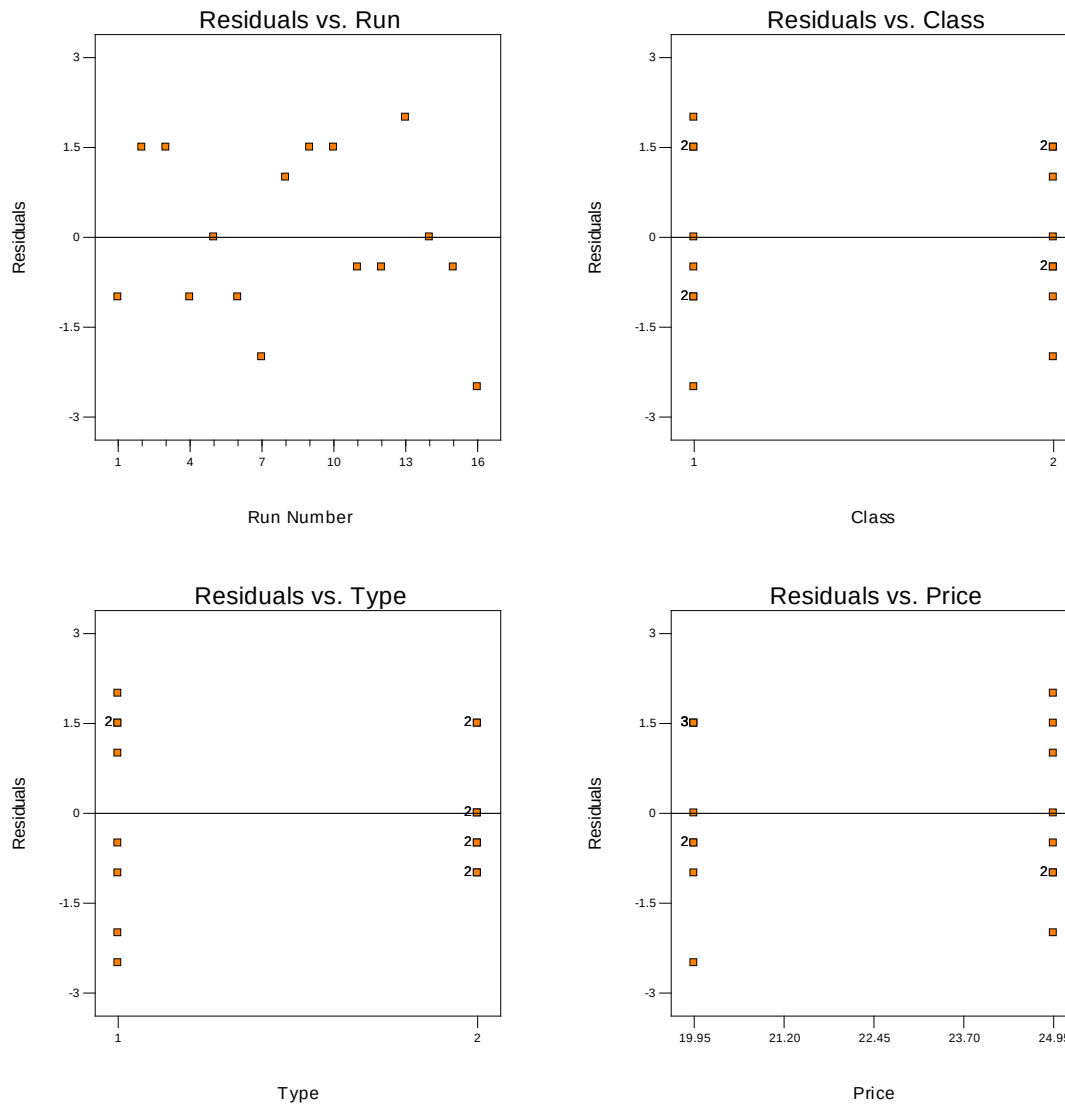
Std. Dev.	1.76	R-Squared	0.8962
Mean	47.63	Adj R-Squared	0.8270
C.V.	3.70	Pred R-Squared	0.6719
PRESS	88.49	Adeq Precision	10.286

Coefficient Factor	Standard Estimate	95% CI DF	95% CI Error	Low	High	VIF
Intercept	47.63	1	0.44	46.63	48.62	
A-Class	-0.88	1	0.44	-1.87	0.12	1.00
B-Type	0.37	1	0.44	-0.62	1.37	1.00
C-Price	1.50	1	0.44	0.50	2.50	1.00
AB	1.63	1	0.44	0.63	2.62	1.00
AC	2.50	1	0.44	1.50	3.50	1.00
BC	1.75	1	0.44	0.75	2.75	1.00

(b) Analyze the residuals from this experiment. Are there any indications of model inadequacy?

The residual plots below do not identify model inadequacy.





(c) What would you recommend to the company?

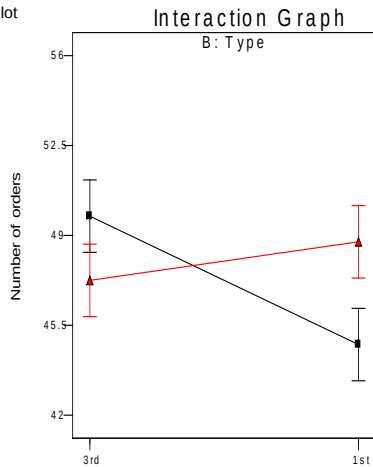
Based on the interaction plots below, we recommend 3rd class mail, black-and-white brochures, and an offered price of \$19.95 would achieve the greatest number of orders. If the offered price must be \$24.95, then the 1st class mail with color brochures is recommended.

DESIGN-EXPERT Plot

Number of orders

X = A: Class
Y = B: Type

■ B1 BW
▲ B2 Color
Actual Factor
C: Price = 22.45



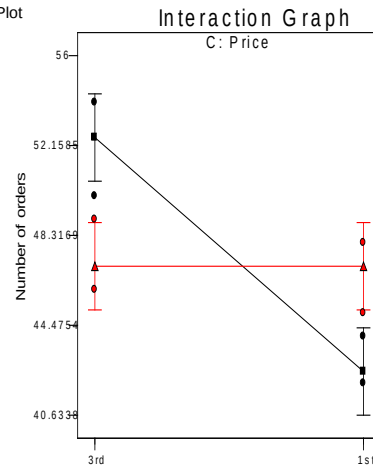
A: Class

DESIGN-EXPERT Plot

Number of orders

X = A: Class
Y = C: Price

● Design Points
■ C- 19.950
▲ C+ 24.950
Actual Factor
B: Type = BW



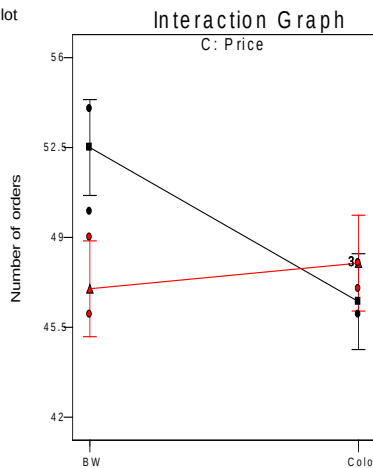
A: Class

DESIGN-EXPERT Plot

Number of orders

X = B: Type
Y = C: Price

● Design Points
■ C- 19.950
▲ C+ 24.950
Actual Factor
A: Class = 3rd



B: Type

6.25. Consider the single replicate of the 2^4 design in Example 6.2. Suppose we had arbitrarily decided to analyze the data assuming that all three- and four-factor interactions were negligible. Conduct this analysis and compare your results with those obtained in the example. Do you think that it is a good idea to arbitrarily assume interactions to be negligible even if they are relatively high-order ones?

Design Expert Output

Response: Filtration Rate in A/min						
ANOVA for Selected Factorial Model						
Analysis of variance table [Partial sum of squares]						
Source	Sum of Squares	DF	Mean Square	F Value	Prob > F	
Model	5603.13	10	560.31	21.92	0.0016	significant
A	1870.56	1	1870.56	73.18	0.0004	
B	39.06	1	39.06	1.53	0.2713	
C	390.06	1	390.06	15.26	0.0113	
D	855.56	1	855.56	33.47	0.0022	
AB	0.063	1	0.063	2.445E-003	0.9625	
AC	1314.06	1	1314.06	51.41	0.0008	
AD	1105.56	1	1105.56	43.25	0.0012	

BC	22.56	1	22.56	0.88	0.3906
BD	0.56	1	0.56	0.022	0.8879
CD	5.06	1	5.06	0.20	0.6749
Residual	127.81	5	25.56		
Cor Total	5730.94	15			

The Model F-value of 21.92 implies the model is significant. There is only a 0.16% chance that a "Model F-Value" this large could occur due to noise.

Values of "Prob > F" less than 0.0500 indicate model terms are significant. In this case A, C, D, AC, AD are significant model terms.

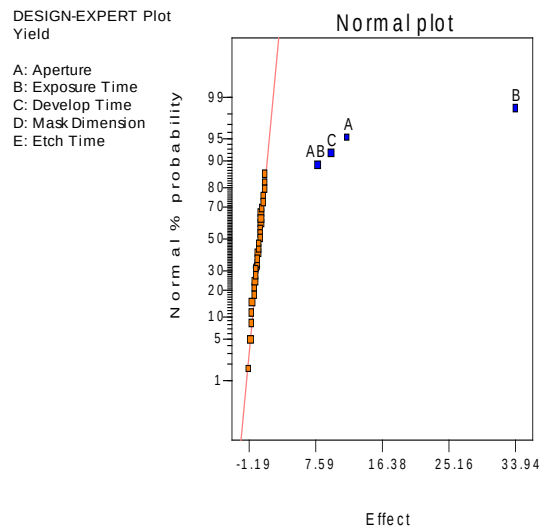
This analysis of variance identifies the same effects as the normal probability plot of effects approach used in Example 6.2. In general, it is not a good idea to arbitrarily pool interactions. Use the normal probability plot of effect estimates as a guide in the choice of which effects to tentatively include in the model.

6.26. An experiment was run in a semiconductor fabrication plant in an effort to increase yield. Five factors, each at two levels, were studied. The factors (and levels) were *A* = aperture setting (small, large), *B* = exposure time (20% below nominal, 20% above nominal), *C* = development time (30 s, 45 s), *D* = mask dimension (small, large), and *E* = etch time (14.5 min, 15.5 min). The unreplicated 2^5 design shown below was run.

(1) = 7	<i>d</i> = 8	<i>e</i> = 8	<i>de</i> = 6
<i>a</i> = 9	<i>ad</i> = 10	<i>ae</i> = 12	<i>ade</i> = 10
<i>b</i> = 34	<i>bd</i> = 32	<i>be</i> = 35	<i>bde</i> = 30
<i>ab</i> = 55	<i>abd</i> = 50	<i>abe</i> = 52	<i>abde</i> = 53
<i>c</i> = 16	<i>cd</i> = 18	<i>ce</i> = 15	<i>cde</i> = 15
<i>ac</i> = 20	<i>acd</i> = 21	<i>ace</i> = 22	<i>acde</i> = 20
<i>bc</i> = 40	<i>bcd</i> = 44	<i>bce</i> = 45	<i>bcde</i> = 41
<i>abc</i> = 60	<i>abcd</i> = 61	<i>abce</i> = 65	<i>abcde</i> = 63

(a) Construct a normal probability plot of the effect estimates. Which effects appear to be large?

From the normal probability plot of effects shown below, effects *A*, *B*, *C*, and the *AB* interaction appear to be large.



(b) Conduct an analysis of variance to confirm your findings for part (a).

Design Expert Output

Response: Yield					
ANOVA for Selected Factorial Model					
Analysis of variance table [Partial sum of squares]					
Source	Sum of Squares	DF	Mean Square	F Value	Prob > F
Model	11585.13	4	2896.28	991.83	< 0.0001
A	1116.28	1	1116.28	382.27	< 0.0001
B	9214.03	1	9214.03	3155.34	< 0.0001
C	750.78	1	750.78	257.10	< 0.0001
AB	504.03	1	504.03	172.61	< 0.0001
Residual	78.84	27	2.92		
Cor Total	11663.97	31			

The Model F-value of 991.83 implies the model is significant. There is only a 0.01% chance that a "Model F-Value" this large could occur due to noise.

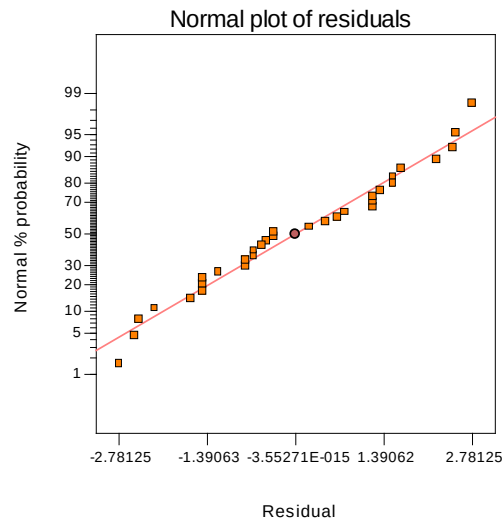
Values of "Prob > F" less than 0.0500 indicate model terms are significant. In this case A, B, C, AB are significant model terms.

(c) Write down the regression model relating yield to the significant process variables.

Design Expert Output

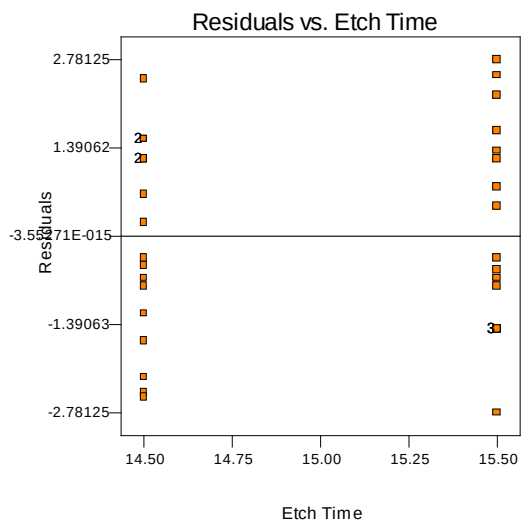
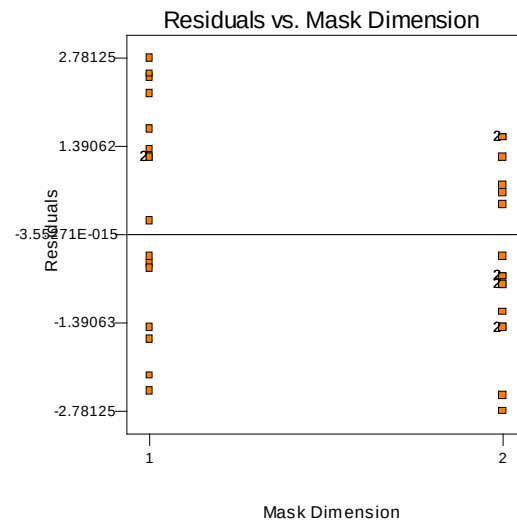
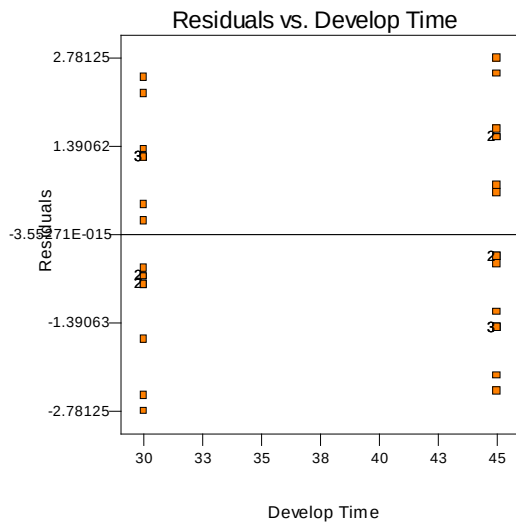
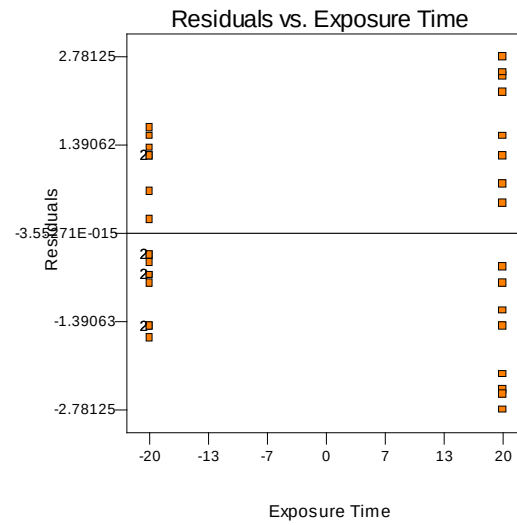
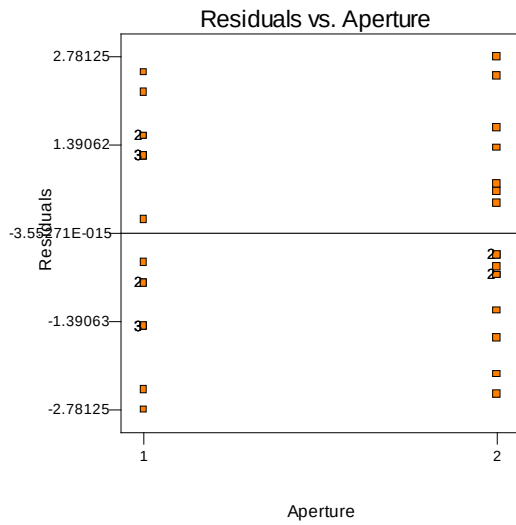
Final Equation in Terms of Actual Factors:			
Aperture	small	Yield	=
			+0.40625
			+0.65000 * Exposure Time
			+0.64583 * Develop Time
Aperture	large	Yield	=
			+12.21875
			+1.04688 * Exposure Time
			+0.64583 * Develop Time

(d) Plot the residuals on normal probability paper. Is the plot satisfactory?



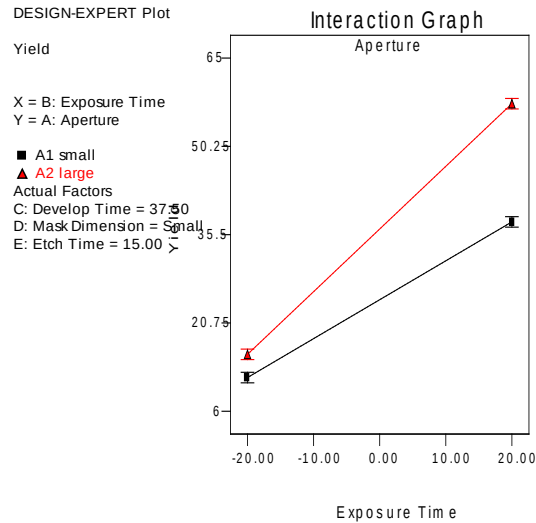
There is nothing unusual about this plot.

(e) Plot the residuals versus the predicted yields and versus each of the five factors. Comment on the plots.



The plot of residual versus exposure time shows some very slight inequality of variance. There is no strong evidence of a potential problem.

(f) Interpret any significant interactions.

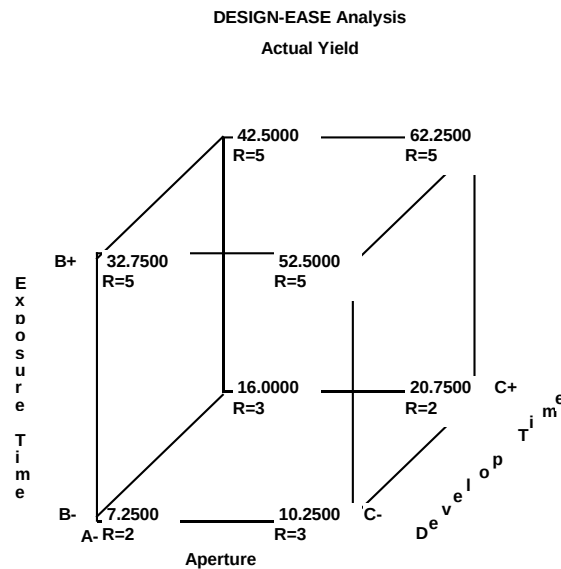


Factor A does not have as large an effect when B is at its low level as it does when B is at its high level.

(g) What are your recommendations regarding process operating conditions?

To achieve the highest yield, run B at the high level, A at the high level, and C at the high level.

(h) Project the 2^5 design in this problem into a 2^k design in the important factors. Sketch the design and show the average and range of yields at each run. Does this sketch aid in interpreting the results of this experiment?



This cube plot aids in interpretation. The strong AB interaction and the large positive effect of C are clearly evident.

6.27. Continuation of Problem 6.26. Suppose that the experimenter had run four runs at the center points in addition to the 32 trials in the original experiment. The yields obtained at the center point runs were 68, 74, 76, and 70.

(a) Reanalyze the experiment, including a test for pure quadratic curvature.

Because aperture and mask dimension are not continuous variables, the four center points were split amongst these two factors as follows.

Aperture	Mask Dimension	Yield
Small	Small	68
Large	Small	74
Small	Large	76
Large	Large	70

The sum of squares for the curvature can be estimated with the following equation and is confirmed with the analysis of variance shown in the Design Expert output.

$$SS_{PureQuadratic} = \frac{n_F n_C (\bar{y}_F - \bar{y}_C)^2}{n_F + n_C} = \frac{(32)(4)(30.53125 - 72)^2}{32 + 4} = 6114.337$$

Design Expert Output

Response: Yield

ANOVA for Selected Factorial Model

Analysis of variance table [Partial sum of squares]

Source	Sum of Squares	DF	Mean Square	F Value	Prob > F	
Model	11461.09	4	2865.27	353.92	< 0.0001	significant
A	992.25	1	992.25	122.56	< 0.0001	
B	9214.03	1	9214.03	1138.12	< 0.0001	
C	750.78	1	750.78	92.74	< 0.0001	
AB	504.03	1	504.03	62.26	< 0.0001	
Curvature	6114.34	1	6114.34	755.24	< 0.0001	significant
Residual	242.88	30	8.10			
Cor Total	17818.31	35				

The Model F-value of 353.92 implies the model is significant. There is only a 0.01% chance that a "Model F-Value" this large could occur due to noise.

Values of "Prob > F" less than 0.0500 indicate model terms are significant.
In this case A, B, C, AB are significant model terms.

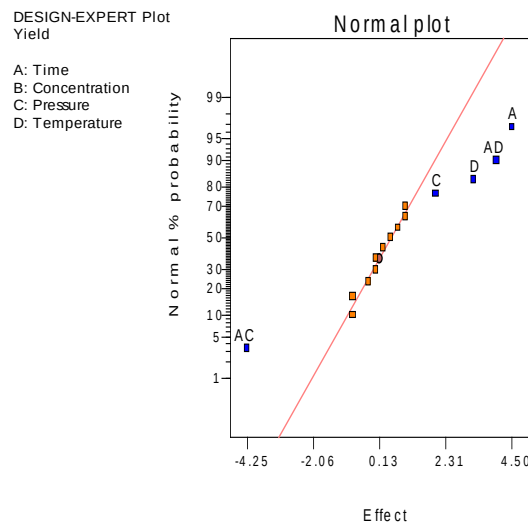
Discuss what your next step would be.

Add axial points for factors *B* and *C* along with four more center points to fit a second-order model and satisfy blocking concerns.

6.28. In a process development study on yield, four factors were studied, each at two levels: time (*A*), concentration (*B*), pressure (*C*), and temperature (*D*). A single replicate of a 2^4 design was run, and the resulting data are shown in the following table:

Run Number	Actual Run Order	A	B	C	D	Yield (lbs)		Factor	Levels
								Low (-)	High (+)
1	5	-	-	-	-	12	A (h)	2.5	3.0
2	9	+	-	-	-	18	B (%)	14	18
3	8	-	+	-	-	13	C (psi)	60	80
4	13	+	+	-	-	16	D (°C)	225	250
5	3	-	-	+	-	17			
6	7	+	-	+	-	15			
7	14	-	+	+	-	20			
8	1	+	+	+	-	15			
9	6	-	-	-	+	10			
10	11	+	-	-	+	25			
11	2	-	+	-	+	13			
12	15	+	+	-	+	24			
13	4	-	-	+	+	19			
14	16	+	-	+	+	21			
15	10	-	+	+	+	17			
16	12	+	+	+	+	23			

(a) Construct a normal probability plot of the effect estimates. Which factors appear to have large effects?



A, C, D and the AC and AD interactions appear to have large effects.

(b) Conduct an analysis of variance using the normal probability plot in part (a) for guidance in forming an error term. What are your conclusions?

Design Expert Output

Response: Yield					
ANOVA for Selected Factorial Model					
Analysis of variance table [Partial sum of squares]					
Source	Sum of Squares	DF	Mean Square	F Value	Prob > F

Model	275.50	5	55.10	33.91	< 0.0001	significant
A	81.00	1	81.00	49.85	< 0.0001	
C	16.00	1	16.00	9.85	0.0105	
D	42.25	1	42.25	26.00	0.0005	
AC	72.25	1	72.25	44.46	< 0.0001	
AD	64.00	1	64.00	39.38	< 0.0001	
Residual	16.25	10	1.62			
Cor Total	291.75	15				

The Model F-value of 33.91 implies the model is significant. There is only a 0.01% chance that a "Model F-Value" this large could occur due to noise.

Values of "Prob > F" less than 0.0500 indicate model terms are significant. In this case A, C, D, AC, AD are significant model terms.

(c) Write down a regression model relating yield to the important process variables.

Design Expert Output

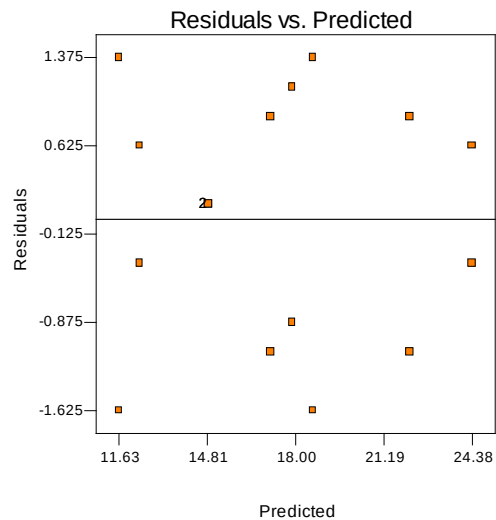
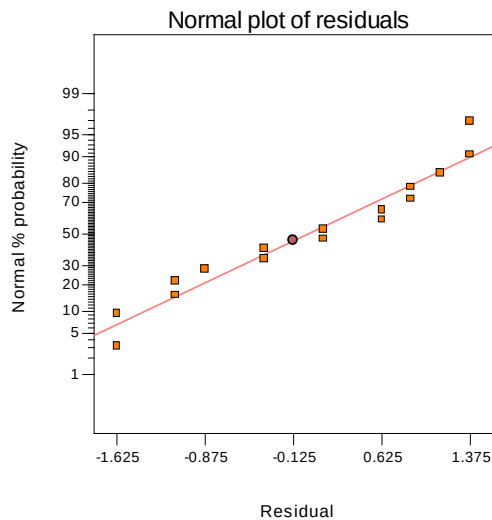
Final Equation in Terms of Coded Factors:

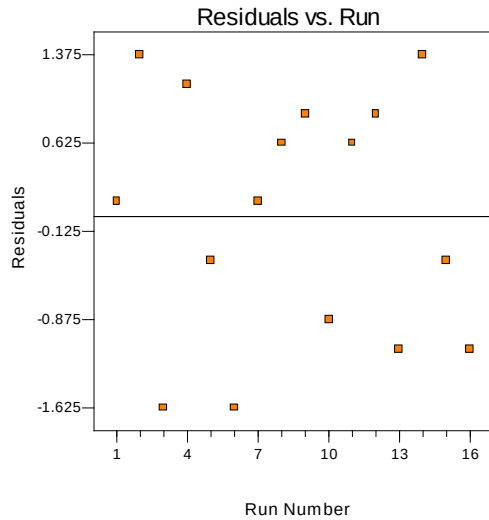
$$\begin{aligned} \text{Yield} = & \\ & +17.38 \\ & +2.25 *A \\ & +1.00 *C \\ & +1.63 *D \\ & -2.13 *A*C \\ & +2.00 *A*D \end{aligned}$$

Final Equation in Terms of Actual Factors:

$$\begin{aligned} \text{Yield} = & \\ & +209.12500 \\ & -83.50000 * \text{Time} \\ & +2.43750 * \text{Pressure} \\ & -1.63000 * \text{Temperature} \\ & -0.85000 * \text{Time} * \text{Pressure} \\ & +0.64000 * \text{Time} * \text{Temperature} \end{aligned}$$

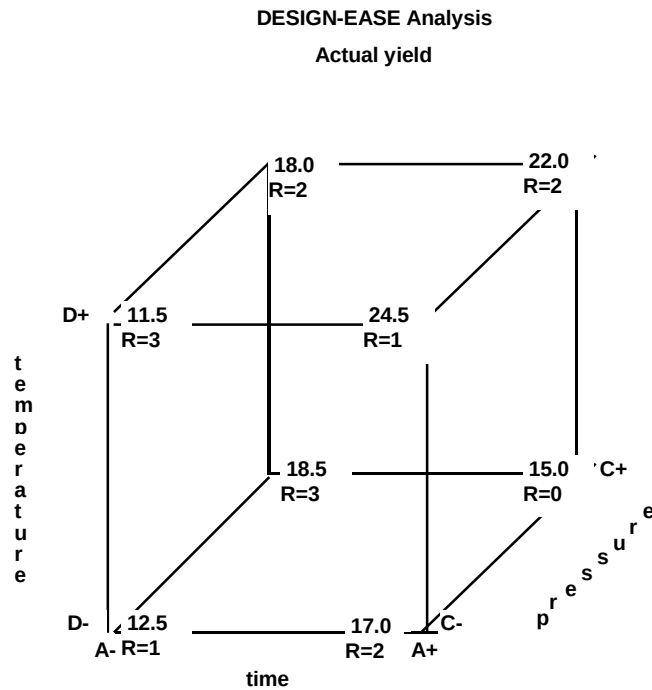
(d) Analyze the residuals from this experiment. Does your analysis indicate any potential problems?





There is nothing unusual about the residual plots.

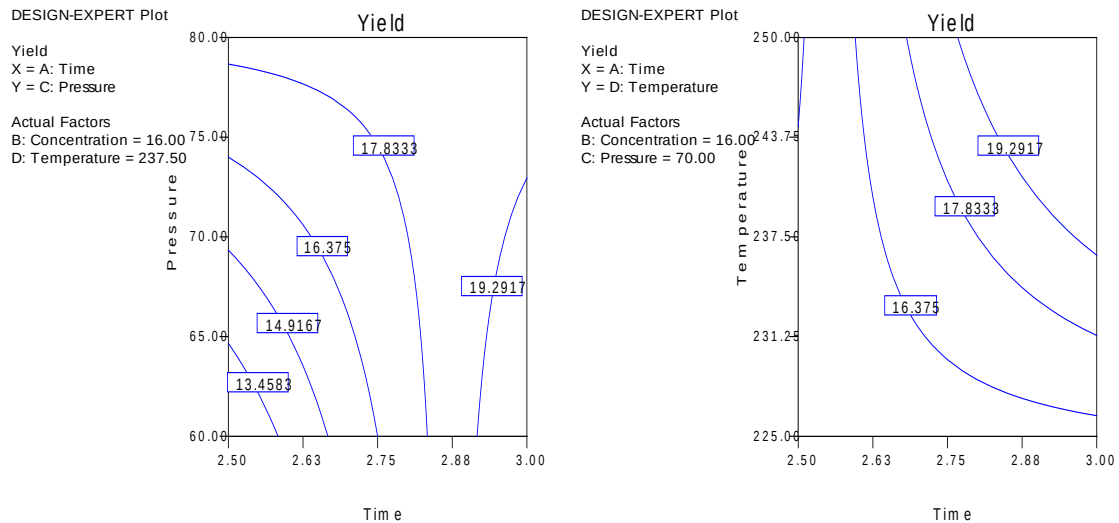
- (e) Can this design be collapsed into a 2^3 design with two replicates? If so, sketch the design with the average and range of yield shown at each point in the cube. Interpret the results.



6.29. Continuation of Problem 6.28. Use the regression model in part (c) of Problem 6.23 to generate a response surface contour plot of yield. Discuss the practical purpose of this response surface plot.

The response surface contour plot shows the adjustments in the process variables that lead to an increasing or decreasing response. It also displays the curvature of the response in the design region, possibly

indicating where robust operating conditions can be found. Two response surface contour plots for this process are shown below. These were formed from the model written in terms of the original design variables.



6.30. The scrumptious brownie experiment. The author is an engineer by training and a firm believer in learning by doing. I have taught experimental design for many years to a wide variety of audiences and have always assigned the planning, conduct, and analysis of an actual experiment to the class participants. The participants seem to enjoy this practical experience and always learn a great deal from it. This problem uses the results of an experiment performed by Gretchen Krueger at Arizona State University.

There are many different ways to bake brownies. The purpose of this experiment was to determine how the pan material, the brand of brownie mix, and the stirring method affect the scrumptiousness of brownies. The factor levels were

Factor	Low (-)	High (+)
A = pan material	Glass	Aluminum
B = stirring method	Spoon	Mixer
C = brand of mix	Expensive	Cheap

The response variable was scrumptiousness, a subjective measure derived from a questionnaire given to the subjects who sampled each batch of brownies. (The questionnaire dealt with such issues as taste, appearance, consistency, aroma, and so forth.) An eight-person test panel sampled each batch and filled out the questionnaire. The design matrix and the response data are shown below:

Brownie Batch	A	B	C	1	2	3	4	5	6	7	8
1	-	-	-	11	9	10	10	11	10	8	9
2	+	-	-	15	10	16	14	12	9	6	15
3	-	+	-	9	12	11	11	11	11	11	12
4	+	+	-	16	17	15	12	13	13	11	11
5	-	-	+	10	11	15	8	6	8	9	14
6	+	-	+	12	13	14	13	9	13	14	9
7	-	+	+	10	12	13	10	7	7	17	13
8	+	+	+	15	12	15	13	12	12	9	14

- (a) Analyze the data from this experiment as if there were eight replicates of a 2^3 design. Comment on the results.

Design Expert Output

ANOVA for Selected Factorial Model

Analysis of variance table [Partial sum of squares]

Source	Sum of Squares	DF	Mean Square	F Value	Prob > F	
Model	93.25	7	13.32	2.20	0.0475	significant
A	72.25	1	72.25	11.95	0.0010	
B	18.06	1	18.06	2.99	0.0894	
C	0.063	1	0.063	0.010	0.9194	
AB	0.062	1	0.062	0.010	0.9194	
AC	1.56	1	1.56	0.26	0.6132	
BC	1.00	1	1.00	0.17	0.6858	
ABC	0.25	1	0.25	0.041	0.8396	
Residual	338.50	56	6.04			
Lack of Fit	0.000	0				
Pure Error	338.50	56	6.04			
Cor Total	431.75	63				

The Model F-value of 2.20 implies the model is significant. There is only a 4.75% chance that a "Model F-Value" this large could occur due to noise.

Values of "Prob > F" less than 0.0500 indicate model terms are significant. In this case A are significant model terms.

In this analysis, A, the pan material and B, the stirring method, appear to be significant. There are 56 degrees of freedom for the error, yet only eight batches of brownies were cooked, one for each recipe.

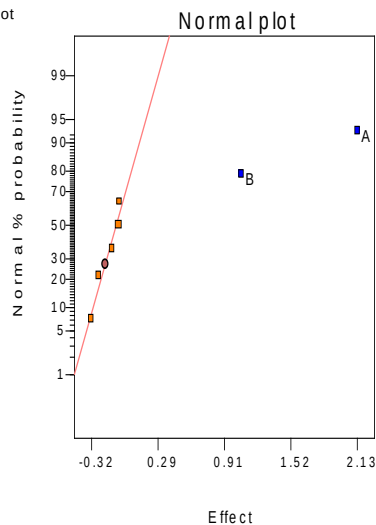
- (b) Is the analysis in part (a) the correct approach? There are only eight batches; do we really have eight replicates of a 2^3 factorial design?

The different rankings by the taste-test panel are not replicates, but repeat observations by different testers on the same batch of brownies. It is not a good idea to use the analysis in part (a) because the estimate of error may not reflect the batch-to-batch variation.

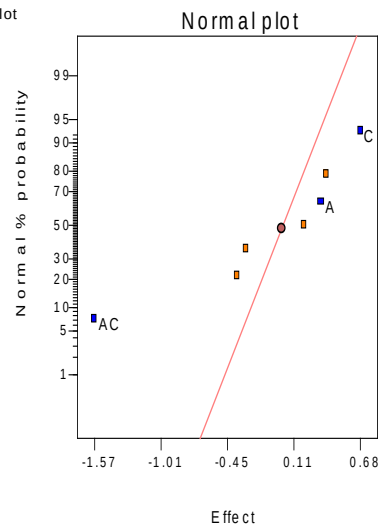
- (c) Analyze the average and standard deviation of the scrumptiousness ratings. Comment on the results. Is this analysis more appropriate than the one in part (a)? Why or why not?

DESIGN-EXPERT Plot
Average

A: Pan Material
B: Stirring Method
C: Mix Brand

DESIGN-EXPERT Plot
Stdev

A: Pan Material
B: Stirring Method
C: Mix Brand



Design Expert Output

Response: Average						
ANOVA for Selected Factorial Model						
Analysis of variance table [Partial sum of squares]						
Source	Sum of Squares	DF	Mean Square	F Value	Prob > F	
Model	11.28	2	5.64	76.13	0.0002	significant
A	9.03	1	9.03	121.93	0.0001	
B	2.25	1	2.25	30.34	0.0027	
Residual	0.37	5	0.074			
Cor Total	11.65	7				
The Model F-value of 76.13 implies the model is significant. There is only a 0.02% chance that a "Model F-Value" this large could occur due to noise.						
Values of "Prob > F" less than 0.0500 indicate model terms are significant. In this case A, B are significant model terms.						

Design Expert Output

Response: Stdev						
ANOVA for Selected Factorial Model						
Analysis of variance table [Partial sum of squares]						
Source	Sum of Squares	DF	Mean Square	F Value	Prob > F	
Model	6.05	3	2.02	9.77	0.0259	significant
A	0.24	1	0.24	1.15	0.3432	
C	0.91	1	0.91	4.42	0.1034	
AC	4.90	1	4.90	23.75	0.0082	
Residual	0.82	4	0.21			
Cor Total	6.87	7				
The Model F-value of 9.77 implies the model is significant. There is only a 2.59% chance that a "Model F-Value" this large could occur due to noise.						
Values of "Prob > F" less than 0.0500 indicate model terms are significant. In this case AC are significant model terms.						

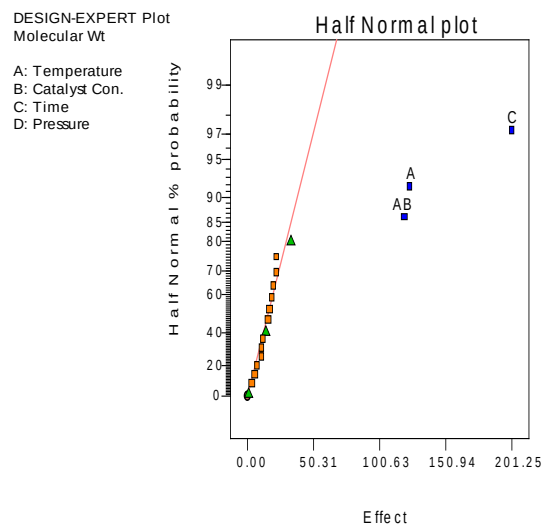
Variables *A* and *B* affect the mean rank of the brownies. Note that the *AC* interaction affects the standard deviation of the ranks. This is an indication that both factors *A* and *C* have some effect on the variability in the ranks. It may also indicate that there is some inconsistency in the taste test panel members. For the analysis of both the average of the ranks and the standard deviation of the ranks, the mean square error is now determined by pooling apparently unimportant effects. This is a more accurate estimate of error than obtained assuming that all observations were replicates.

6.31. An experiment was conducted on a chemical process that produces a polymer. The four factors studied were temperature (*A*), catalyst concentration (*B*), time (*C*), and pressure (*D*). Two responses, molecular weight and viscosity, were observed. The design matrix and response data are shown below:

Run Number	Actual Run				Molecular		Factor	Levels
	Order	A	B	C	D	Weight	Viscosity	
1	18	-	-	-	-	2400	1400	A (°C)
2	9	+	-	-	-	2410	1500	B (%)
3	13	-	+	-	-	2315	1520	C (min)
4	8	+	+	-	-	2510	1630	D (psi)
5	3	-	-	+	-	2615	1380	
6	11	+	-	+	-	2625	1525	
7	14	-	+	+	-	2400	1500	

8	17	+	+	+	-	2750	1620
9	6	-	-	-	+	2400	1400
10	7	+	-	-	+	2390	1525
11	2	-	+	-	+	2300	1500
12	10	+	+	-	+	2520	1500
13	4	-	-	+	+	2625	1420
14	19	+	-	+	+	2630	1490
15	15	-	+	+	+	2500	1500
16	20	+	+	+	+	2710	1600
17	1	0	0	0	0	2515	1500
18	5	0	0	0	0	2500	1460
19	16	0	0	0	0	2400	1525
20	12	0	0	0	0	2475	1500

- (a) Consider only the molecular weight response. Plot the effect estimates on a normal probability scale. What effects appear important?



A, C and the AB interaction appear to be important.

- (b) Use an analysis of variance to confirm the results from part (a). Is there an indication of curvature? A, C and the AB interaction are significant. While the main effect of B is not significant, it could be included to preserve hierarchy in the model. There is no indication of quadratic curvature.

Design Expert Output

Response: Molecular Wt						
ANOVA for Selected Factorial Model						
Analysis of variance table [Partial sum of squares]						
Source	Sum of Squares	DF	Mean Square	F Value	Prob > F	
Model	2.809E+005	3	93620.83	73.00	< 0.0001	significant
A	61256.25	1	61256.25	47.76	< 0.0001	
C	1.620E+005	1	1.620E+005	126.32	< 0.0001	
AB	57600.00	1	57600.00	44.91	< 0.0001	
Curvature	3645.00	1	3645.00	2.84	0.1125	not significant
Residual	19237.50	15	1282.50			
Lack of Fit	11412.50	12	951.04	0.36	0.9106	not significant

Pure Error	7825.00	3	2608.33
Cor Total	3.037E+005	19	

The Model F-value of 73.00 implies the model is significant. There is only a 0.01% chance that a "Model F-Value" this large could occur due to noise.

Values of "Prob > F" less than 0.0500 indicate model terms are significant. In this case A, C, AB are significant model terms.

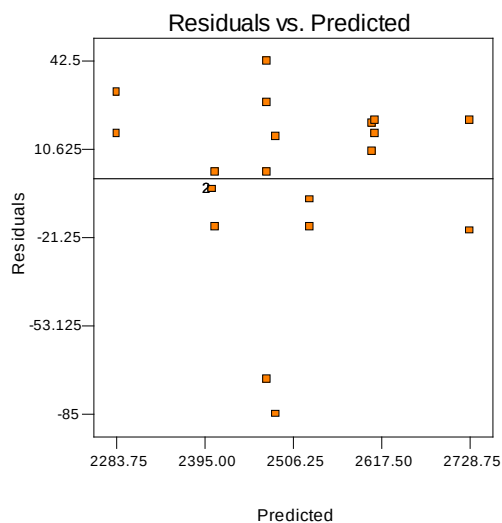
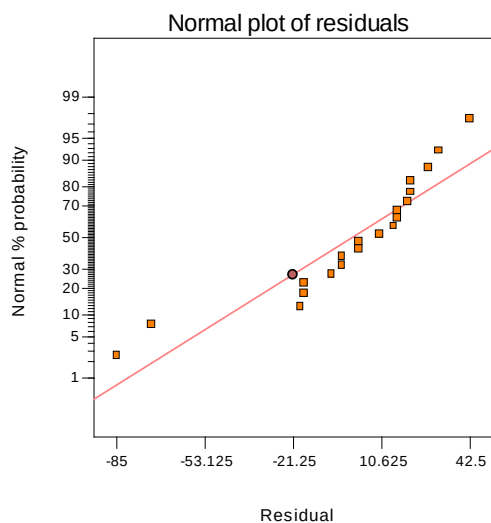
(c) Write down a regression model to predict molecular weight as a function of the important variables.

Design Expert Output

Final Equation in Terms of Coded Factors:

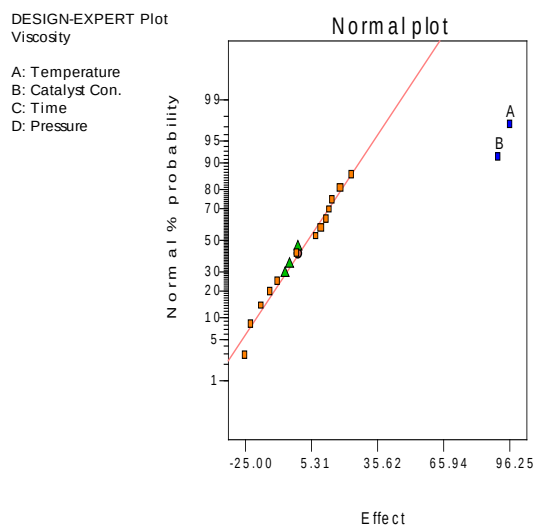
Molecular Wt =
 +2506.25
 +61.87 * A
 +100.63 * C
 +60.00 * A * B

(d) Analyze the residuals and comment on model adequacy.



There are two residuals that appear to be large and should be investigated.

(e) Repeat parts (a) - (d) using the viscosity response.



Design Expert Output

Response: Viscosity

ANOVA for Selected Factorial Model

Analysis of variance table [Partial sum of squares]

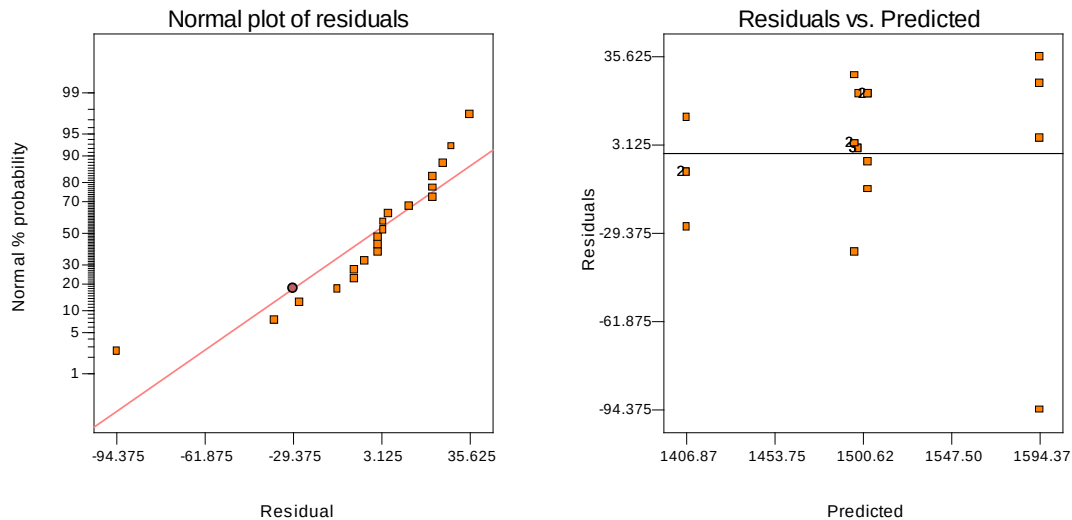
Source	Sum of Squares	DF	Mean Square	F Value	Prob > F	
Model	70362.50	2	35181.25	35.97	< 0.0001	significant
A	37056.25	1	37056.25	37.88	< 0.0001	
B	33306.25	1	33306.25	34.05	< 0.0001	
Curvature	61.25	1	61.25	0.063	0.8056	not significant
Residual	15650.00	16	978.13			
Lack of Fit	13481.25	13	1037.02	1.43	0.4298	not significant
Pure Error	2168.75	3	722.92			
Cor Total	86073.75	19				

The Model F-value of 35.97 implies the model is significant. There is only a 0.01% chance that a "Model F-Value" this large could occur due to noise.

Values of "Prob > F" less than 0.0500 indicate model terms are significant. In this case A, B are significant model terms.

Final Equation in Terms of Coded Factors:

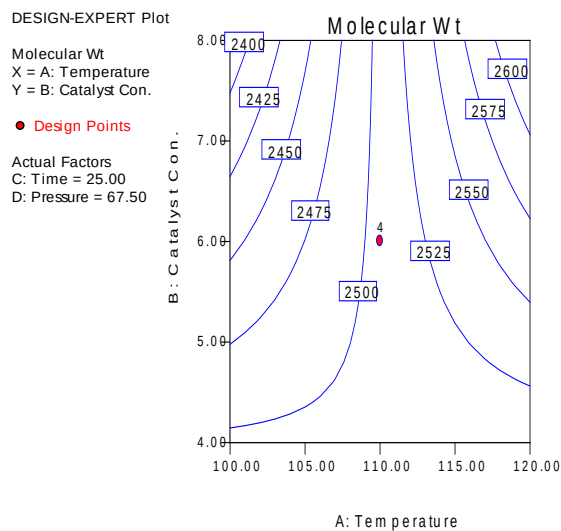
$$\begin{aligned} \text{Viscosity} = & \\ & +1500.62 \\ & +48.13 \quad * A \\ & +45.63 \quad * B \end{aligned}$$



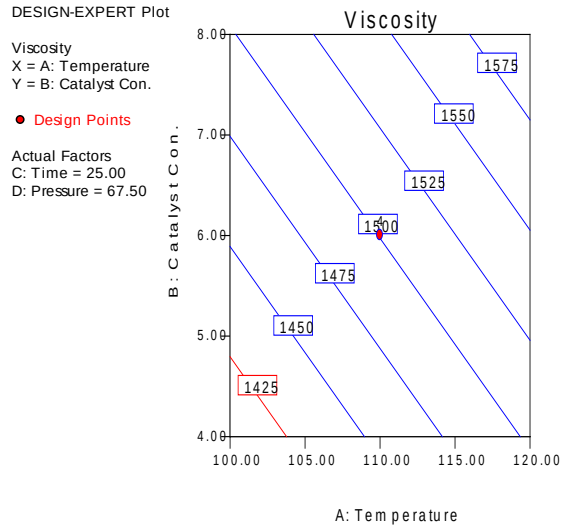
There is one large residual that should be investigated.

6.32. Continuation of Problem 6.31. Use the regression models for molecular weight and viscosity to answer the following questions.

- (a) Construct a response surface contour plot for molecular weight. In what direction would you adjust the process variables to increase molecular weight? Increase temperature, catalyst and time.

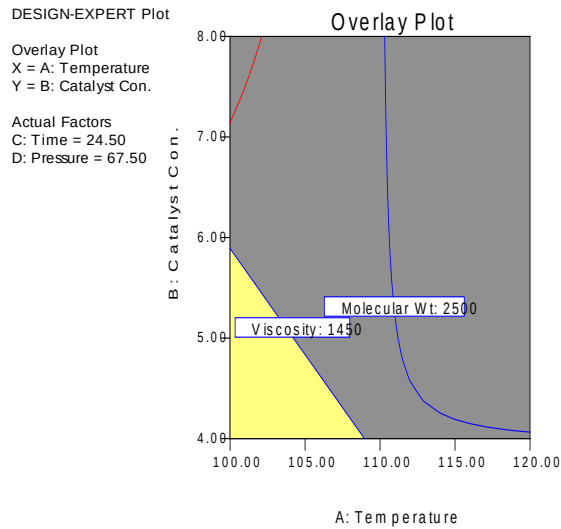


- (b) Construct a response surface contour plot for viscosity. In what direction would you adjust the process variables to decrease viscosity?



Decrease temperature and catalyst.

- (c) What operating conditions would you recommend if it was necessary to produce a product with a molecular weight between 2400 and 2500, and the lowest possible viscosity?



Set the temperature between 100 and 105, the catalyst between 4 and 5%, and the time at 24.5 minutes. The pressure was not significant and can be set at conditions that may improve other results of the process such as cost.

6.33. Consider the single replicate of the 2^4 design in Example 6.2. Suppose that we ran five points at the center (0, 0, 0, 0) and observed the following responses: 73, 75, 71, 69, and 76. Test for curvature in this experiment. Interpret the results.

Design Expert Output

Response: Filtration Rate						
ANOVA for Selected Factorial Model						
Analysis of variance table [Partial sum of squares]						
Source	Sum of Squares	DF	Mean Square	F Value	Prob > F	
Model	5535.81	5	1107.16	68.01	< 0.0001	significant
A	1870.56	1	1870.56	114.90	< 0.0001	
C	390.06	1	390.06	23.96	0.0002	
D	855.56	1	855.56	52.55	< 0.0001	
AC	1314.06	1	1314.06	80.71	< 0.0001	
AD	1105.56	1	1105.56	67.91	< 0.0001	
Curvature	28.55	1	28.55	1.75	0.2066	not significant
Residual	227.93	14	16.28			
Lack of Fit	195.13	10	19.51	2.38	0.2093	not significant
Pure Error	32.80	4	8.20			
Cor Total	5792.29	20				

The Model F-value of 68.01 implies the model is significant. There is only a 0.01% chance that a "Model F-Value" this large could occur due to noise.

Values of "Prob > F" less than 0.0500 indicate model terms are significant. In this case A, C, D, AC, AD are significant model terms.

The "Curvature F-value" of 1.75 implies the curvature (as measured by difference between the average of the center points and the average of the factorial points) in the design space is not significant relative to the noise. There is a 20.66% chance that a "Curvature F-value" this large could occur due to noise.

There is no indication of curvature.

6.34. A missing value in a 2^k factorial. It is not unusual to find that one of the observations in a 2^k design is missing due to faulty measuring equipment, a spoiled test, or some other reason. If the design is replicated n times ($n > 1$) some of the techniques discussed in Chapter 5 can be employed. However, for an unreplicated factorial ($n=1$) some other method must be used. One logical approach is to estimate the missing value with a number that makes the highest-order interaction contrast zero. Apply this technique to the experiment in Example 6.2 assuming that run ab is missing. Compare the results with the results of Example 6.2.

Treatment Combination	Response	Response *					
		<i>ABCD</i>	<i>ABCD</i>	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>
(1)	45	45	1	-1	-1	-1	-1
<i>a</i>	71	-71	-1	1	-1	-1	-1
<i>b</i>	48	-48	-1	-1	1	-1	-1
<i>ab</i>	missing	missing * 1	1	1	1	-1	-1
<i>c</i>	68	-68	-1	-1	-1	1	-1
<i>ac</i>	60	60	1	1	-1	1	-1
<i>bc</i>	80	80	1	-1	1	1	-1
<i>abc</i>	65	-65	-1	1	1	1	-1
<i>d</i>	43	-43	-1	-1	-1	-1	1
<i>ad</i>	100	100	1	1	-1	-1	1
<i>bd</i>	45	45	1	-1	1	-1	1
<i>abd</i>	104	-104	-1	1	1	-1	1
<i>cd</i>	75	75	1	-1	-1	1	1
<i>acd</i>	86	-86	-1	1	-1	1	1
<i>bcd</i>	70	-70	-1	-1	1	1	1
<i>abcd</i>	96	96	1	1	1	1	1

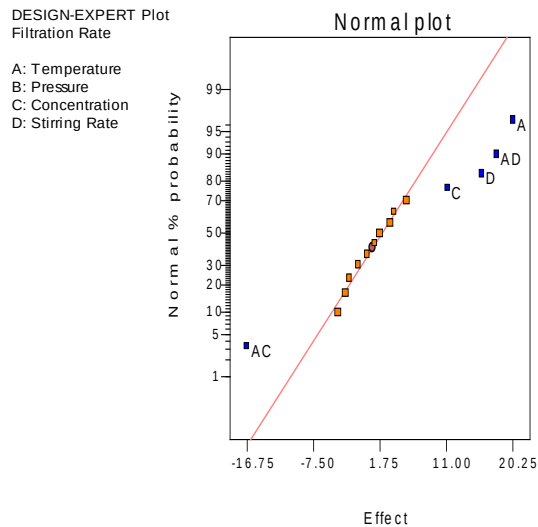
$$\text{Contrast (ABCD)} = \text{Missing} - 54 = 0$$

$$\text{Missing} = 54$$

Substitute the value 54 for the missing run at *ab*. From the effects list and half normal plot shown below, factors A, C, D, AC, and AD appear to be large; the same result as found in Example 6.2.

Design Expert Output

	Term	Effect	SumSqr	% Contribtn
Model	Intercept			
Model	A	20.25	1640.25	27.5406
Model	B	1.75	12.25	0.205684
Model	C	11.25	506.25	8.50019
Model	D	16	1024	17.1935
Model	AB	-1.25	6.25	0.104941
Model	AC	-16.75	1122.25	18.8431
Model	AD	18	1296	21.7605
Model	BC	3.75	56.25	0.944465
Model	BD	1	4	0.067162
Model	CD	-2.5	25	0.419762
Model	ABC	3.25	42.25	0.709398
Model	ABD	5.5	121	2.03165
Model	ACD	-3	36	0.604458
Model	BCD	-4	64	1.07459
Model	ABCD	0	0	0
	Lenth's ME	11.5676		
	Lenth's SME	23.4839		

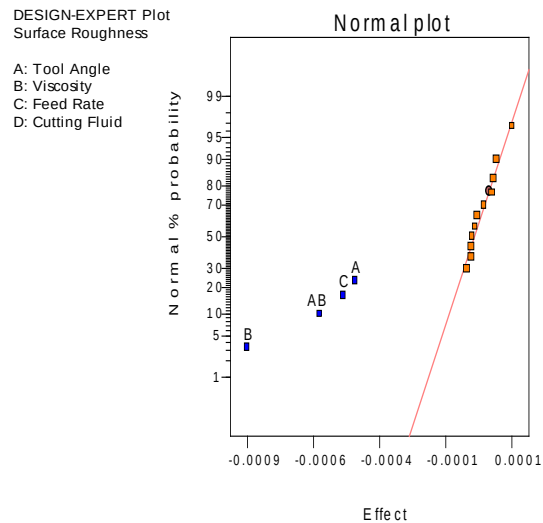


6.35. An engineer has performed an experiment to study the effect of four factors on the surface roughness of a machined part. The factors (and their levels) are *A* = tool angle (12 degrees, 15 degrees), *B* = cutting fluid viscosity (300, 400), *C* = feed rate (10 in/min, 15 in/min), and *D* = cutting fluid cooler used (no, yes). The data from this experiment (with the factors coded to the usual -1, +1 levels) are shown below.

Run	A	B	C	D	Surface Roughness
1	-	-	-	-	0.00340
2	+	-	-	-	0.00362
3	-	+	-	-	0.00301
4	+	+	-	-	0.00182

5	-	-	+	-	0.00280
6	+	-	+	-	0.00290
7	-	+	+	-	0.00252
8	+	+	+	-	0.00160
9	-	-	-	+	0.00336
10	+	-	-	+	0.00344
11	-	+	-	+	0.00308
12	+	+	-	+	0.00184
13	-	-	+	+	0.00269
14	+	-	+	+	0.00284
15	-	+	+	+	0.00253
16	+	+	+	+	0.00163

- (a) Estimate the factor effects. Plot the effect estimates on a normal probability plot and select a tentative model.



- (b) Fit the model identified in part (a) and analyze the residuals. Is there any indication of model inadequacy?

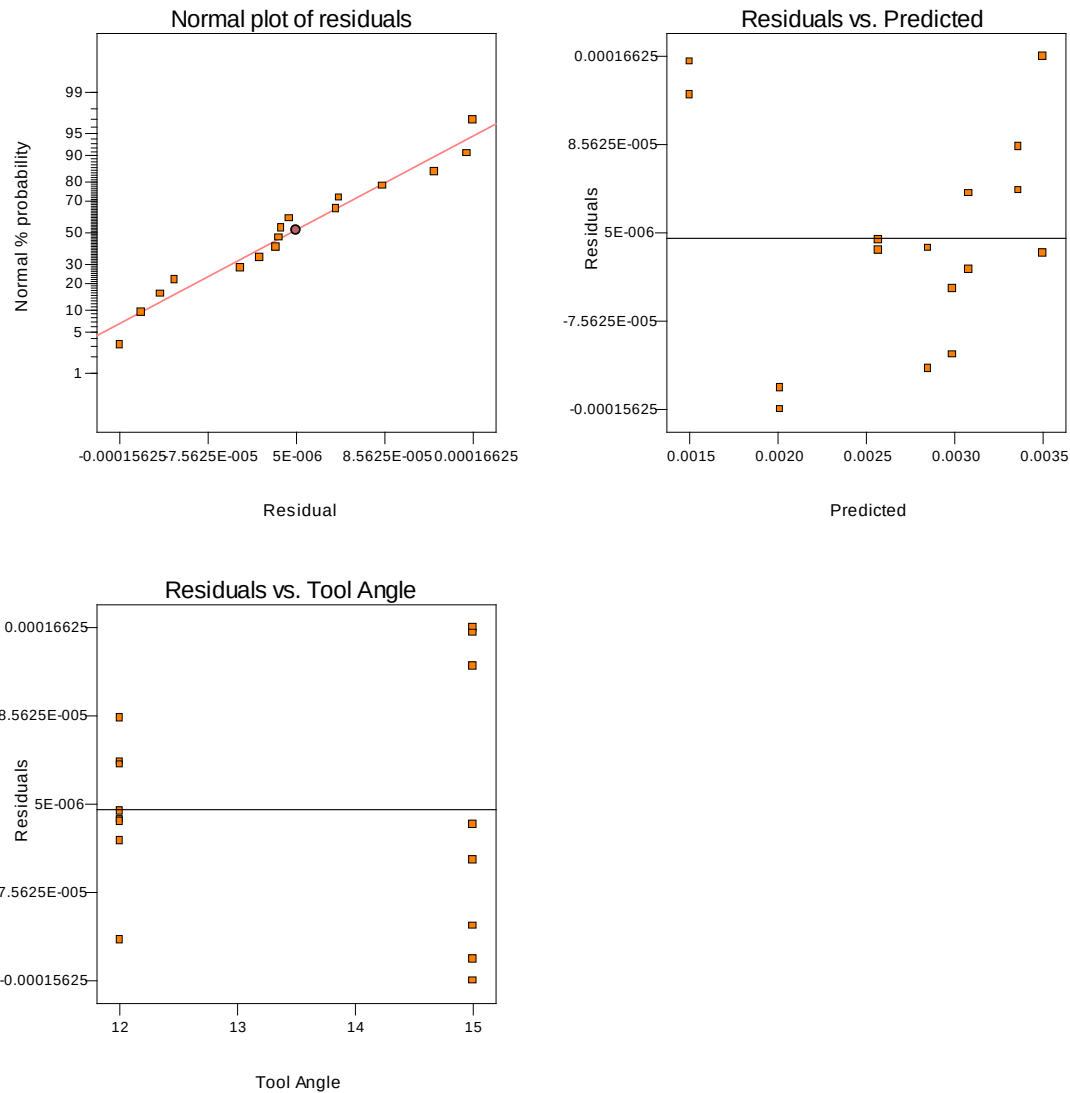
Design Expert Output

Response: Surface Roughness						
ANOVA for Selected Factorial Model						
Analysis of variance table [Partial sum of squares]						
Source	Sum of Squares	DF	Mean Square	F Value	Prob > F	
Model	6.406E-006	4	1.601E-006	114.97	< 0.0001	significant
A	8.556E-007	1	8.556E-007	61.43	< 0.0001	
B	3.080E-006	1	3.080E-006	221.11	< 0.0001	
C	1.030E-006	1	1.030E-006	73.96	< 0.0001	
AB	1.440E-006	1	1.440E-006	103.38	< 0.0001	
Residual	1.532E-007	11	1.393E-008			
Cor Total	6.559E-006	15				

The Model F-value of 114.97 implies the model is significant. There is only a 0.01% chance that a "Model F-Value" this large could occur due to noise.

Values of "Prob > F" less than 0.0500 indicate model terms are significant.

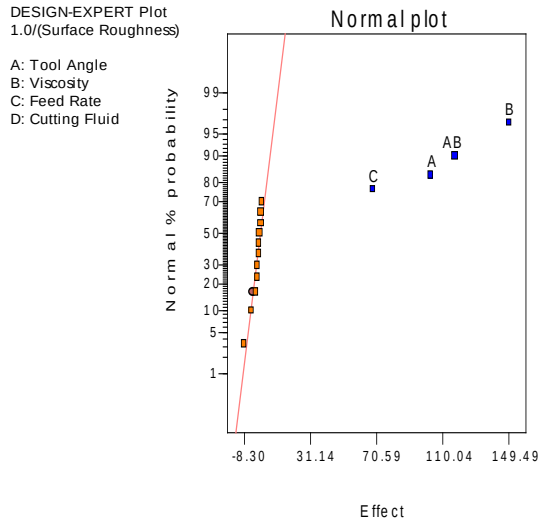
In this case A, B, C, AB are significant model terms.



The plot of residuals versus predicted shows a slight “u-shaped” appearance in the residuals, and the plot of residuals versus tool angle shows an outward-opening funnel.

- (c) Repeat the analysis from parts (a) and (b) using $1/y$ as the response variable. Is there an indication that the transformation has been useful?

The plots of the residuals are more representative of a model that does not violate the constant variance assumption.



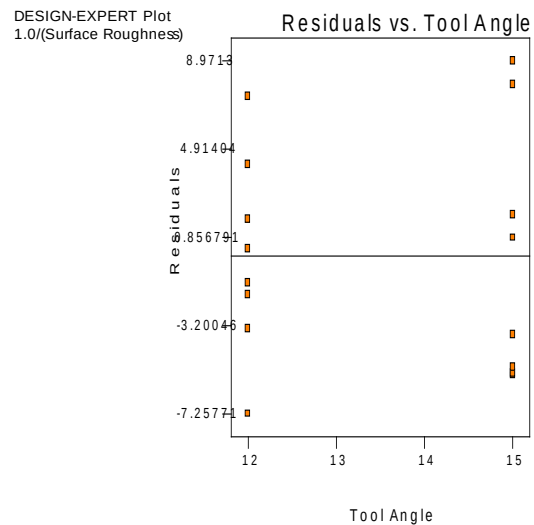
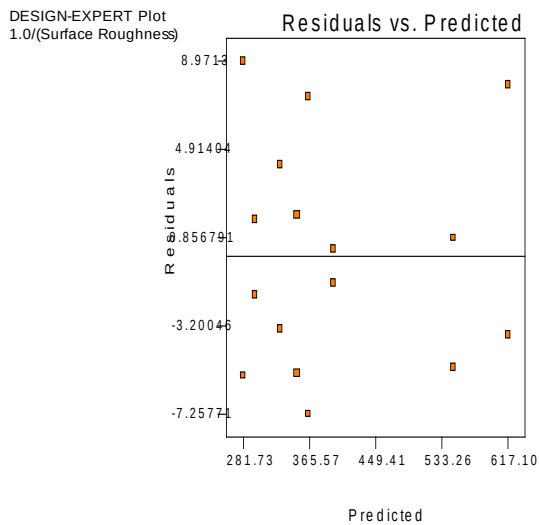
Design Expert Output

Response: Surface RoughnessTransform:Inverse
ANOVA for Selected Factorial Model
Analysis of variance table [Partial sum of squares]

Source	Sum of Squares	DF	Mean Square	F Value	Prob > F	
Model	2.059E+005	4	51472.28	1455.72	< 0.0001	significant
A	42610.92	1	42610.92	1205.11	< 0.0001	
B	89386.27	1	89386.27	2527.99	< 0.0001	
C	18762.29	1	18762.29	530.63	< 0.0001	
AB	55129.62	1	55129.62	1559.16	< 0.0001	
Residual	388.94	11	35.36			
Cor Total	2.063E+005	15				

The Model F-value of 1455.72 implies the model is significant. There is only a 0.01% chance that a "Model F-Value" this large could occur due to noise.

Values of "Prob > F" less than 0.0500 indicate model terms are significant.
In this case A, B, C, AB are significant model terms.



- (d) Fit a model in terms of the coded variables that can be used to predict the surface roughness. Convert this prediction equation into a model in the natural variables.

Design Expert Output

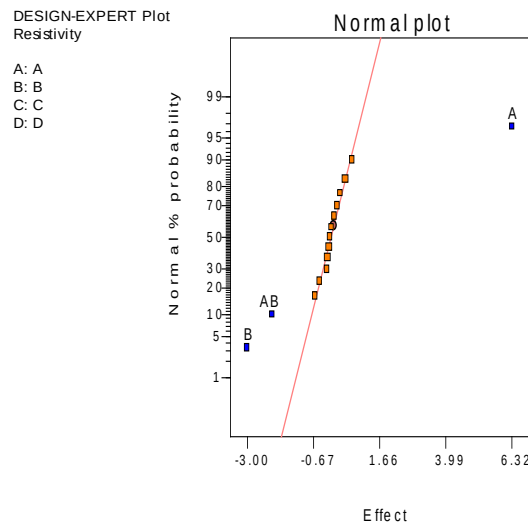
Final Equation in Terms of Coded Factors:

$$\begin{aligned}
 &1.0/(\text{Surface Roughness}) = \\
 &+397.81 \\
 &+51.61 \quad * A \\
 &+74.74 \quad * B \\
 &+34.24 \quad * C \\
 &+58.70 \quad * A * B
 \end{aligned}$$

6.36. Resistivity on a silicon wafer is influenced by several factors. The results of a 2^4 factorial experiment performed during a critical process step is shown below.

Run	A	B	C	D	Resistivity
1	-	-	-	-	1.92
2	+	-	-	-	11.28
3	-	+	-	-	1.09
4	+	+	-	-	5.75
5	-	-	+	-	2.13
6	+	-	+	-	9.53
7	-	+	+	-	1.03
8	+	+	+	-	5.35
9	-	-	-	+	1.60
10	+	-	-	+	11.73
11	-	+	-	+	1.16
12	+	+	-	+	4.68
13	-	-	+	+	2.16
14	+	-	+	+	9.11
15	-	+	+	+	1.07
16	+	+	+	+	5.30

- (a) Estimate the factor effects. Plot the effect estimates on a normal probability plot and select a tentative model.



- (b) Fit the model identified in part (a) and analyze the residuals. Is there any indication of model inadequacy?

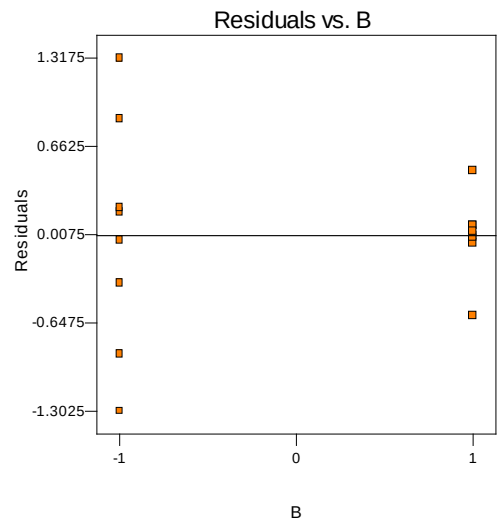
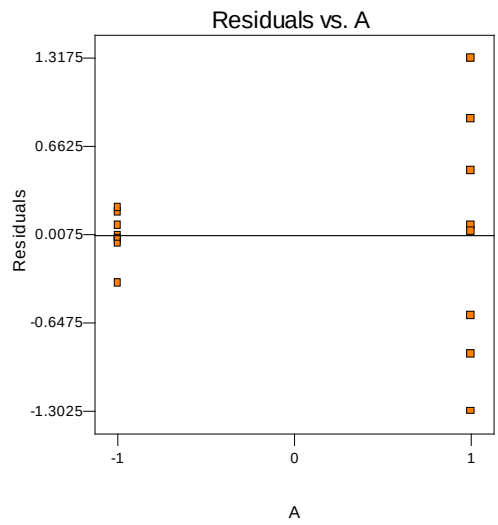
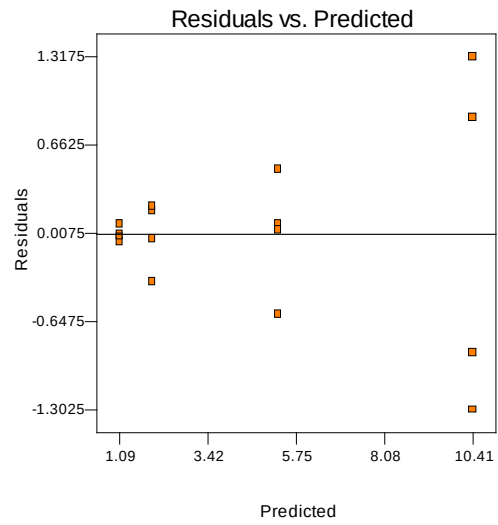
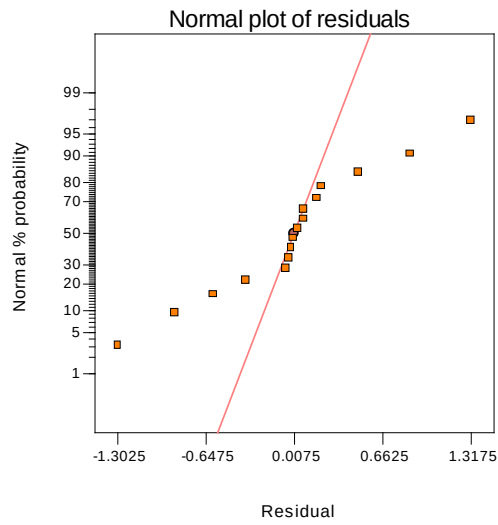
The normal probability plot of residuals is not satisfactory. The plots of residual versus predicted, residual versus factor A, and the residual versus factor B are funnel shaped indicating non-constant variance.

Design Expert Output

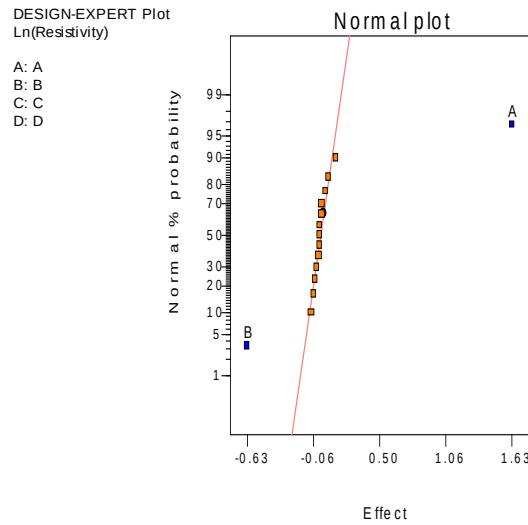
Response: Resistivity					
ANOVA for Selected Factorial Model					
Analysis of variance table [Partial sum of squares]					
Source	Sum of Squares	DF	Mean Square	F Value	Prob > F
Model	214.22	3	71.41	148.81	< 0.0001
A	159.83	1	159.83	333.09	< 0.0001
B	36.09	1	36.09	75.21	< 0.0001
AB	18.30	1	18.30	38.13	< 0.0001
Residual	5.76	12	0.48		
Cor Total	219.98	15			

The Model F-value of 148.81 implies the model is significant. There is only a 0.01% chance that a "Model F-Value" this large could occur due to noise.

Values of "Prob > F" less than 0.0500 indicate model terms are significant.
In this case A, B, AB are significant model terms.



- (c) Repeat the analysis from parts (a) and (b) using $\ln(y)$ as the response variable. Is there any indication that the transformation has been useful?



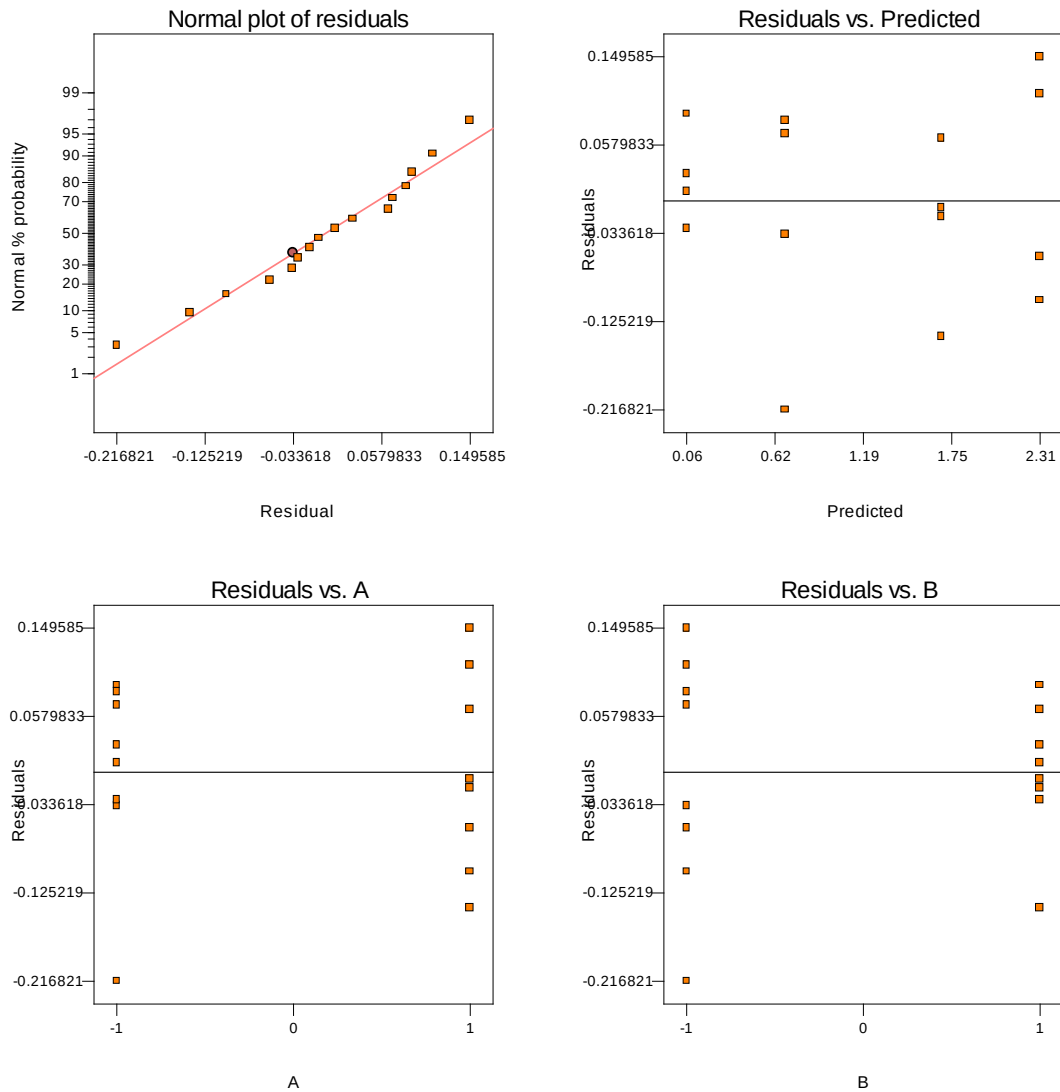
Design Expert Output

Response:	Resistivity Transform: Natural log			Constant:	0.000	
ANOVA for Selected Factorial Model						
Analysis of variance table [Partial sum of squares]						
	Sum of		Mean	F		
Source	Squares	DF	Square	Value	Prob > F	
Model	12.15	2	6.08	553.44	< 0.0001	significant
A	10.57	1	10.57	962.95	< 0.0001	
B	1.58	1	1.58	143.94	< 0.0001	
Residual	0.14	13	0.011			
Cor Total	12.30	15				

The Model F-value of 553.44 implies the model is significant. There is only a 0.01% chance that a "Model F-Value" this large could occur due to noise.

Values of "Prob > F" less than 0.0500 indicate model terms are significant. In this case A, B are significant model terms.

The transformed data no longer indicates that the *AB* interaction is significant. A simpler model has resulted from the log transformation.



The residual plots are much improved.

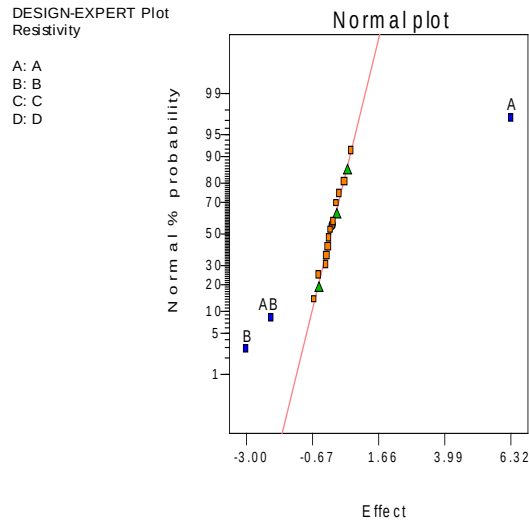
(d) Fit a model in terms of the coded variables that can be used to predict the resistivity.

Design Expert Output

Final Equation in Terms of Coded Factors:

$$\begin{aligned} \text{Ln(Resistivity)} &= \\ &+1.19 \\ &+0.81 \quad * A \\ &-0.31 \quad * B \end{aligned}$$

6.37. Continuation of Problem 6.36. Suppose that the experiment had also run four center points along with the 16 runs in Problem 6.36. The resistivity measurements at the center points are: 8.15, 7.63, 8.95, 6.48. Analyze the experiment again incorporating the center points. What conclusions can you draw now?



Design Expert Output

Response: Resistivity

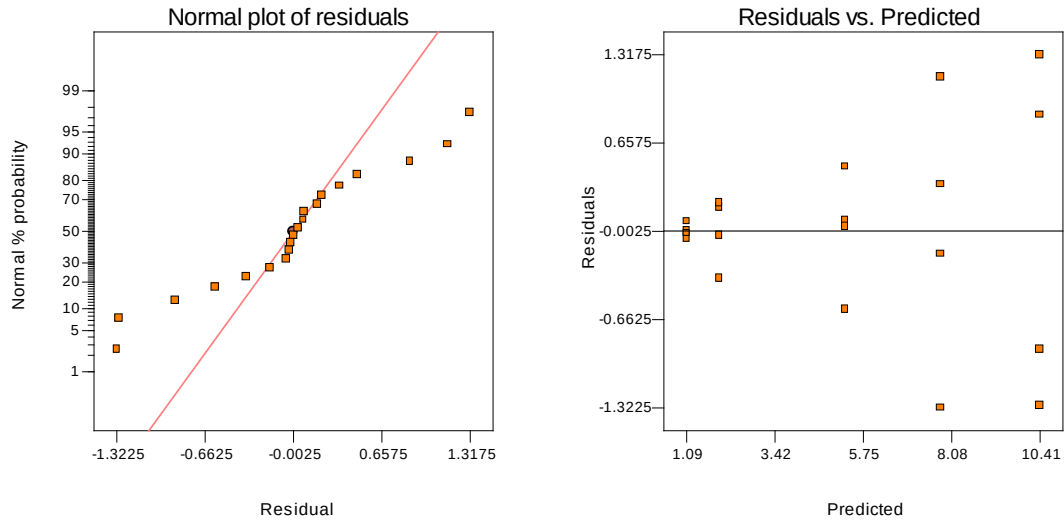
ANOVA for Selected Factorial Model

Analysis of variance table [Partial sum of squares]

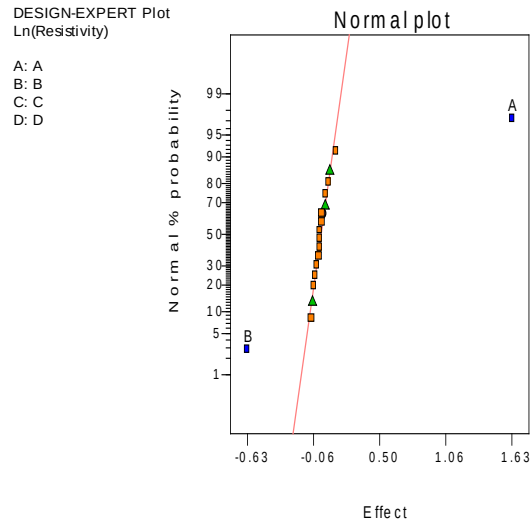
Source	Sum of Squares	DF	Mean Square	F Value	Prob > F	
Model	214.22	3	71.41	119.35	< 0.0001	significant
A	159.83	1	159.83	267.14	< 0.0001	
B	36.09	1	36.09	60.32	< 0.0001	
AB	18.30	1	18.30	30.58	< 0.0001	
Curvature	31.19	1	31.19	52.13	< 0.0001	significant
Residual	8.97	15	0.60			
Lack of Fit	5.76	12	0.48	0.45	0.8632	not significant
Pure Error	3.22	3	1.07			
Cor Total	254.38	19				

The Model F-value of 119.35 implies the model is significant. There is only a 0.01% chance that a "Model F-Value" this large could occur due to noise.

Values of "Prob > F" less than 0.0500 indicate model terms are significant. In this case A, B, AB are significant model terms.



Because of the funnel shaped residual versus predicted plot, the analysis was repeated with the natural log transformation.



Design Expert Output

Response:			Resistivity Transform: Natural log		Constant: 0.000	
ANOVA for Selected Factorial Model						
Analysis of variance table [Partial sum of squares]						
Source	Sum of Squares	DF	Mean Square	F Value	Prob > F	
Model	12.15	2	6.08	490.37	< 0.0001	significant
A	10.57	1	10.57	853.20	< 0.0001	
B	1.58	1	1.58	127.54	< 0.0001	
Curvature	2.38	1	2.38	191.98	< 0.0001	significant
Residual	0.20	16	0.012			
Lack of Fit	0.14	13	0.011	0.59	0.7811	not significant
Pure Error	0.056	3	0.019			
Cor Total	14.73	19				
The Model F-value of 490.37 implies the model is significant. There is only a 0.01% chance that a "Model F-Value" this large could occur due to noise.						

Values of "Prob > F" less than 0.0500 indicate model terms are significant.
In this case A, B are significant model terms.

The "Curvature F-value" of 191.98 implies there is significant curvature (as measured by difference between the average of the center points and the average of the factorial points) in the design space. There is only a 0.01% chance that a "Curvature F-value" this large could occur due to noise.

The curvature test indicates that the model has significant pure quadratic curvature.

6.38. The book by Davies (*Design and Analysis of Industrial Experiments*) describes an experiment to study the yield of isatin. The factors studied and their levels are as follows:

Factor	Low (-)	High (+)
A: Acid strength (%)	87	93
B: Reaction time (min)	15	30
C: Amount of acid (ml)	35	45
D: Reaction temperature (°C)	60	70

The data from the 2^4 factorial is shown in Table P6.11.

Table P6.11					
A	B	C	D	Yield	
-1	-1	-1	-1	6.08	
1	-1	-1	-1	6.04	
-1	1	-1	-1	6.53	
1	1	-1	-1	6.43	
-1	-1	1	-1	6.31	
1	-1	1	-1	6.09	
-1	1	1	-1	6.12	
1	1	1	-1	6.36	
-1	-1	-1	1	6.79	
1	-1	-1	1	6.68	
-1	1	-1	1	6.73	
1	1	-1	1	6.08	
-1	-1	1	1	6.77	
1	-1	1	1	6.38	
-1	1	1	1	6.49	
1	1	1	1	6.23	

- (a) Fit a main-effects-only model to the data from this experiment. Are any of the main effects significant?

Temperature appears to be significant.

Design Expert Output

Response	1	Yield				
ANOVA for selected factorial model						
Analysis of variance table [Partial sum of squares - Type III]						
Source	Sum of Squares	df	Mean Square	F Value	p-value Prob > F	
Model	0.47	4	0.12	2.26	0.1287	not significant
A-Acid strength	0.15	1	0.15	2.81	0.1221	
B-Reaction time	1.806E-003	1	1.806E-003	0.035	0.8558	

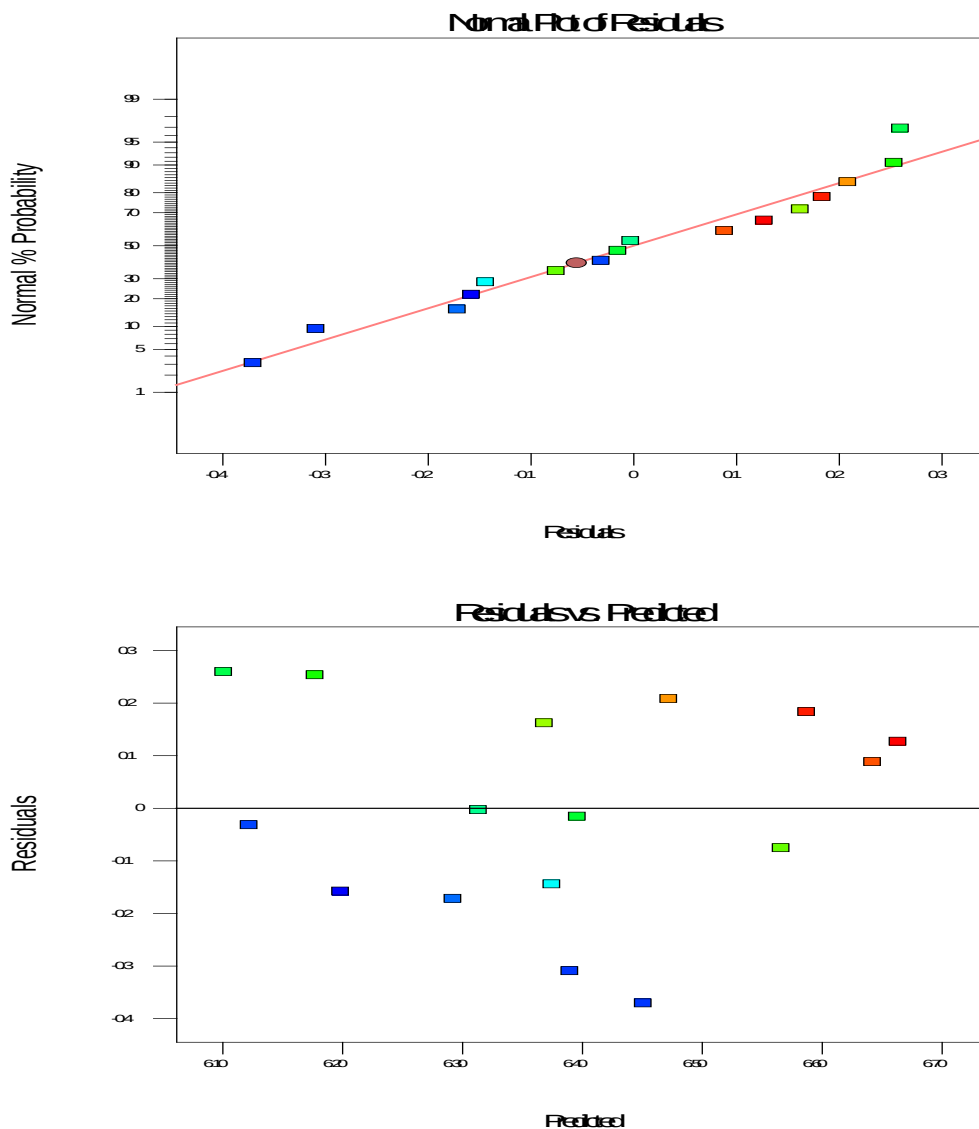
C-Amount of acid	0.023	1	0.023	0.45	0.5181
D-Temperature	0.30	1	0.30	5.75	0.0354
Residual	0.57	11	0.052		
Cor Total	1.04	15			

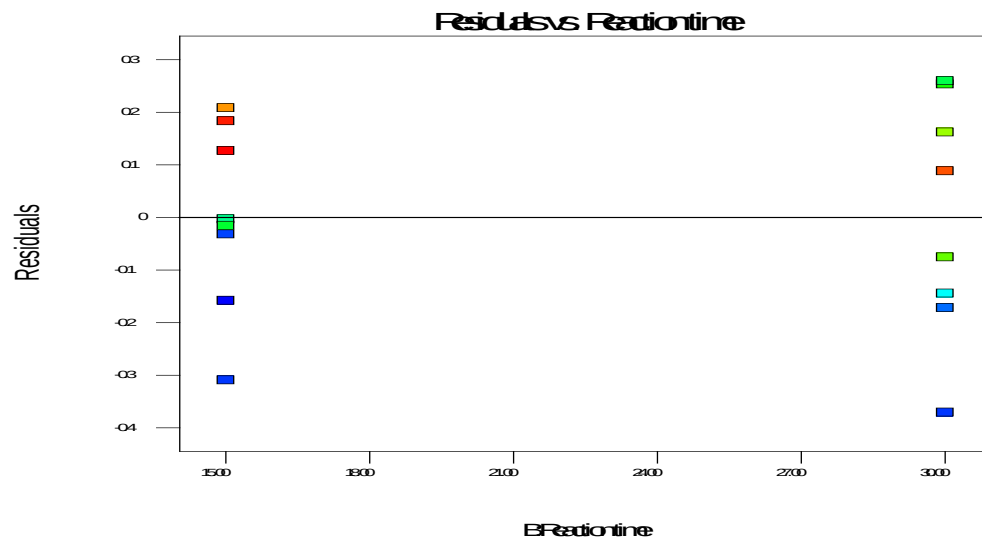
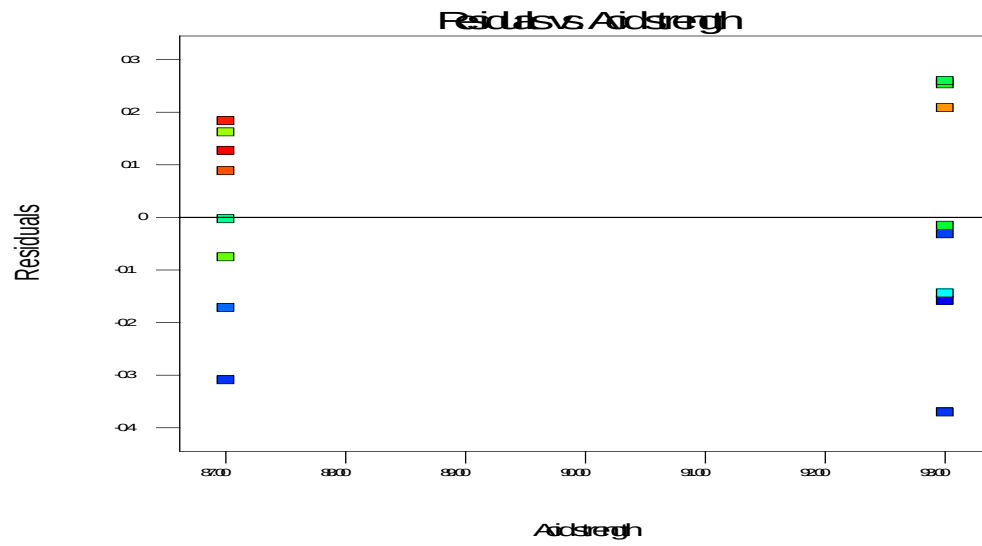
The "Model F-value" of 2.26 implies the model is not significant relative to the noise. There is a 12.87 % chance that a "Model F-value" this large could occur due to noise.

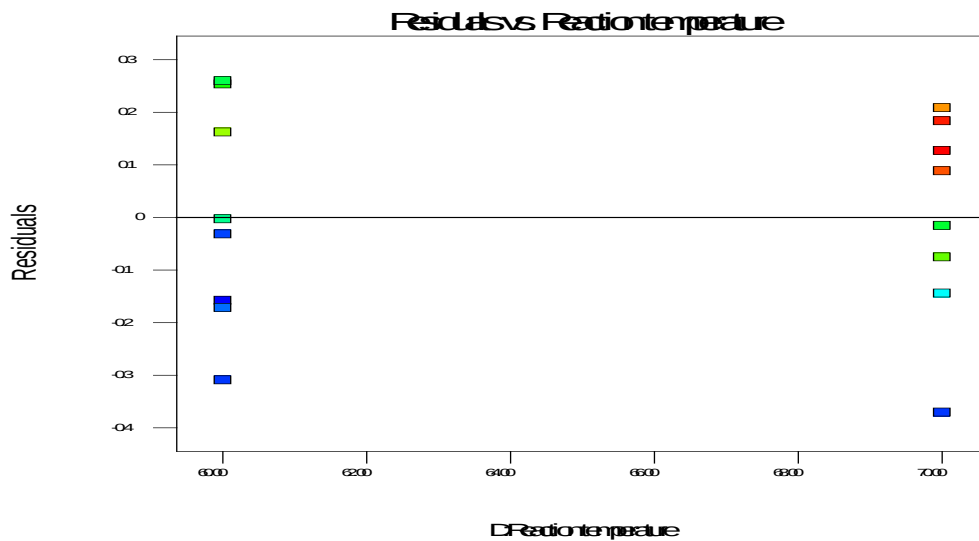
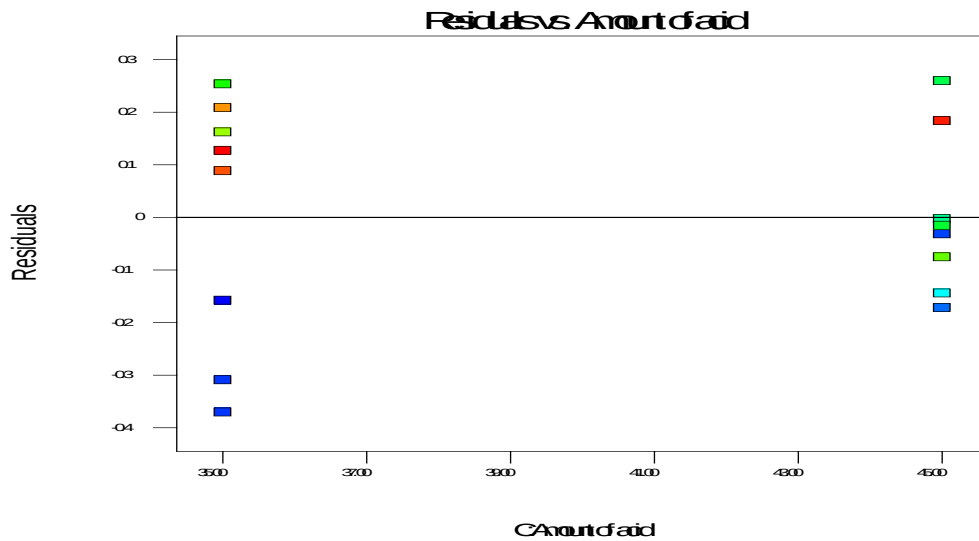
Values of "Prob > F" less than 0.0500 indicate model terms are significant.
In this case D are significant model terms.

- (b) Analyze the residuals. Are there any indications of model inadequacy or violation of the assumptions?

The residual plots appear to be acceptable.







- (c) Find an equation for predicting the yield of isatin over the design space. Express the equation in both coded and engineering units.

Design Expert Output

Final Equation in Terms of Coded Factors:

$$\begin{aligned} \text{Yield} = & \\ & +6.38 \\ & -0.096 * A \\ & -0.011 * B \\ & -0.038 * C \\ & +0.14 * D \end{aligned}$$

Design Expert Output

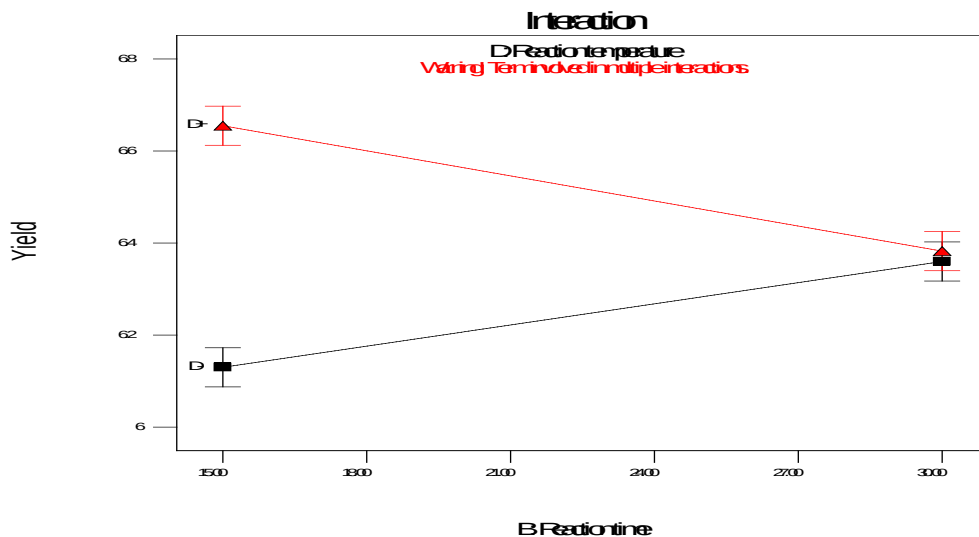
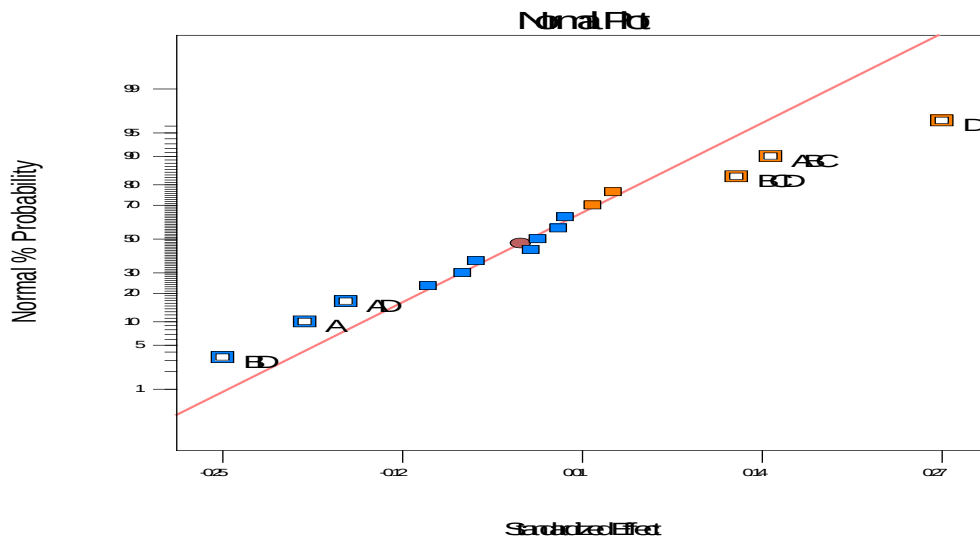
Final Equation in Terms of Coded Factors:

$$\begin{aligned} \text{Yield} = & \\ & +7.80813 \end{aligned}$$

-0.031875	* Acid strength
-1.41667E-003	* Reaction time
-7.62500E-003	* Amount of acid
+0.027375	* Reaction temperature

- (d) Is there any indication that adding interactions to the model would improve the results that you have obtained?

The normal probability of effect shown below suggests that the *BD* and other interactions may improve the results. The *BD* interaction plot shows why the interaction is strong, yet the main effect *B* is not significant.

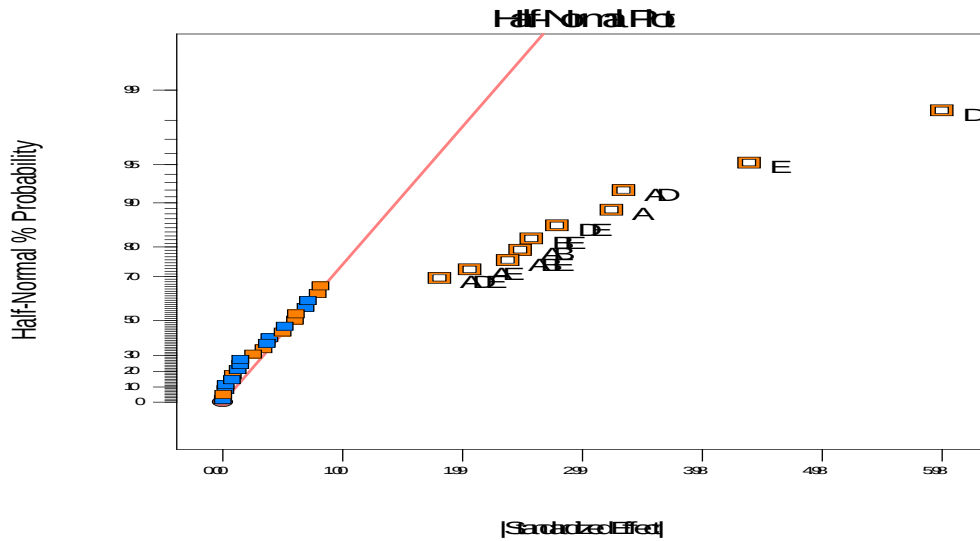


6.39. An article in *Quality and Reliability Engineering International* (2010, Vol. 26, pp. 223-233) presents a 2^5 factorial design. The experiment is shown in the following table:

A	B	C	D	E	y
-1	-1	-1	-1	-1	8.11
1	-1	-1	-1	-1	5.56
-1	1	-1	-1	-1	5.77
1	1	-1	-1	-1	5.82
-1	-1	1	-1	-1	9.17
1	-1	1	-1	-1	7.8
-1	1	1	-1	-1	3.23
1	1	1	-1	-1	5.69
-1	-1	-1	1	-1	8.82
1	-1	-1	1	-1	14.23
-1	1	-1	1	-1	9.2
1	1	-1	1	-1	8.94
-1	-1	1	1	-1	8.68
1	-1	1	1	-1	11.49
-1	1	1	1	-1	6.25
1	1	1	1	-1	9.12
-1	-1	-1	-1	1	7.93
1	-1	-1	-1	1	5
-1	1	-1	-1	1	7.47
1	1	-1	-1	1	12
-1	-1	1	-1	1	9.86
1	-1	1	-1	1	3.65
-1	1	1	-1	1	6.4
1	1	1	-1	1	11.61
-1	-1	-1	1	1	12.43
1	-1	-1	1	1	17.55
-1	1	-1	1	1	8.87
1	1	-1	1	1	25.38
-1	-1	1	1	1	13.06
1	-1	1	1	1	18.85
-1	1	1	1	1	11.78
1	1	1	1	1	26.05

(a) Analyze the data from this experiment. Identify the significant factors and interactions.

The half normal plot of effects below identifies the significant factors and interactions.



Design Expert Output

Response 1 y
ANOVA for selected factorial model

Analysis of variance table [Partial sum of squares - Type III]

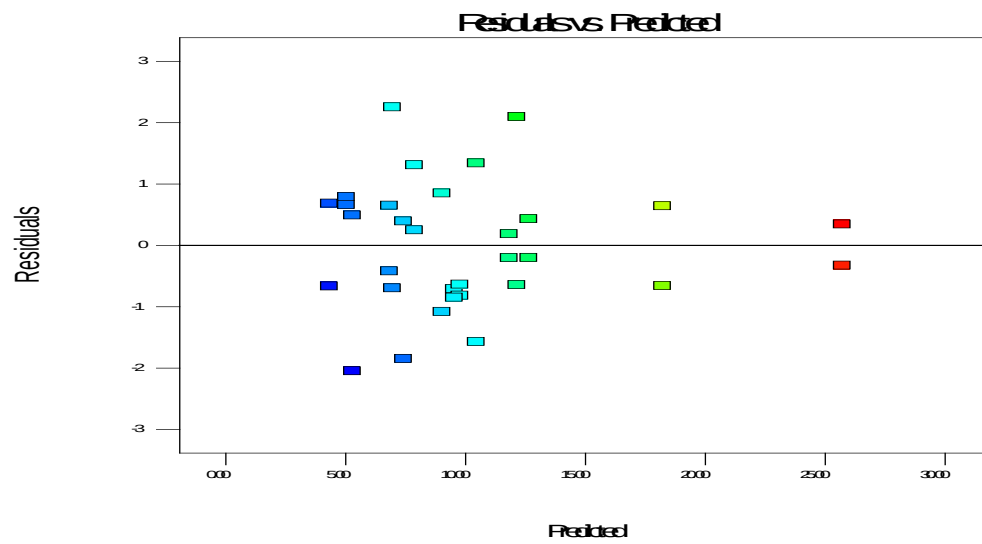
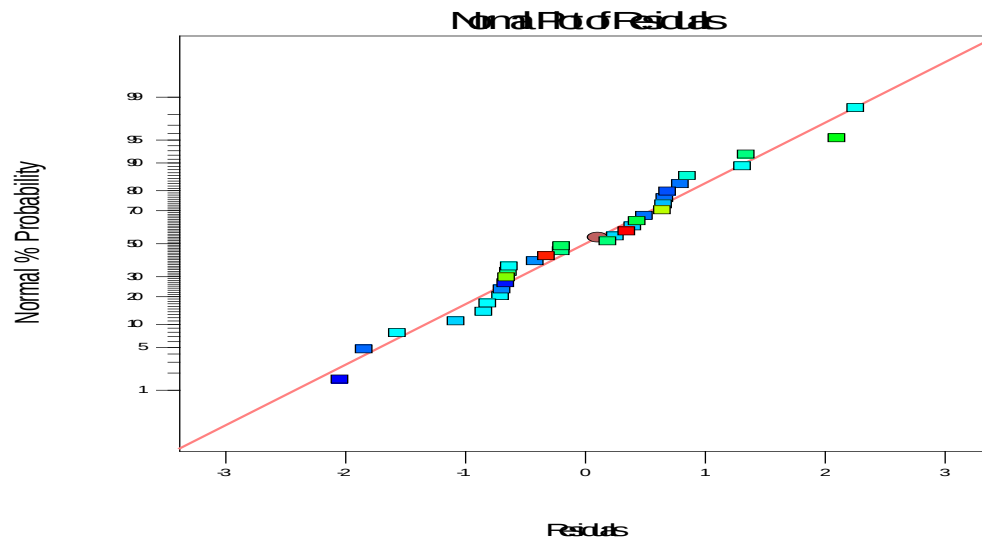
Source	Sum of Squares	df	Mean Square	F Value	p-value Prob > F	significant
Model	879.62	11	79.97	49.15	< 0.0001	
A-A	83.56	1	83.56	51.36	< 0.0001	
B-B	0.060	1	0.060	0.037	0.8492	
D-D	285.78	1	285.78	175.66	< 0.0001	
E-E	153.17	1	153.17	94.15	< 0.0001	
AB	48.93	1	48.93	30.08	< 0.0001	
AD	88.88	1	88.88	54.63	< 0.0001	
AE	33.76	1	33.76	20.75	0.0002	
BE	52.71	1	52.71	32.40	< 0.0001	
DE	61.80	1	61.80	37.99	< 0.0001	
ABE	44.96	1	44.96	27.64	< 0.0001	
ADE	26.01	1	26.01	15.99	0.0007	
Residual	32.54	20	1.63			
Cor Total	912.16	31				

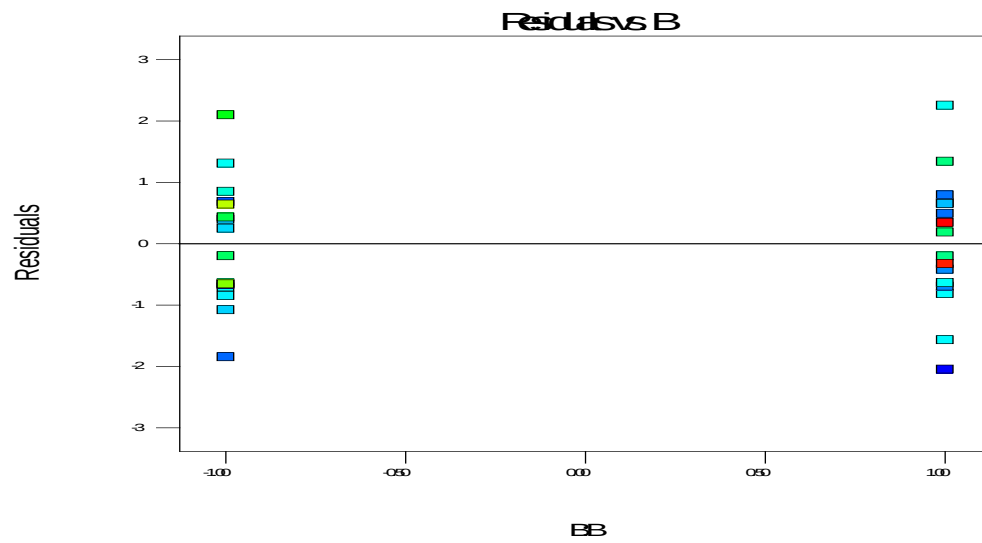
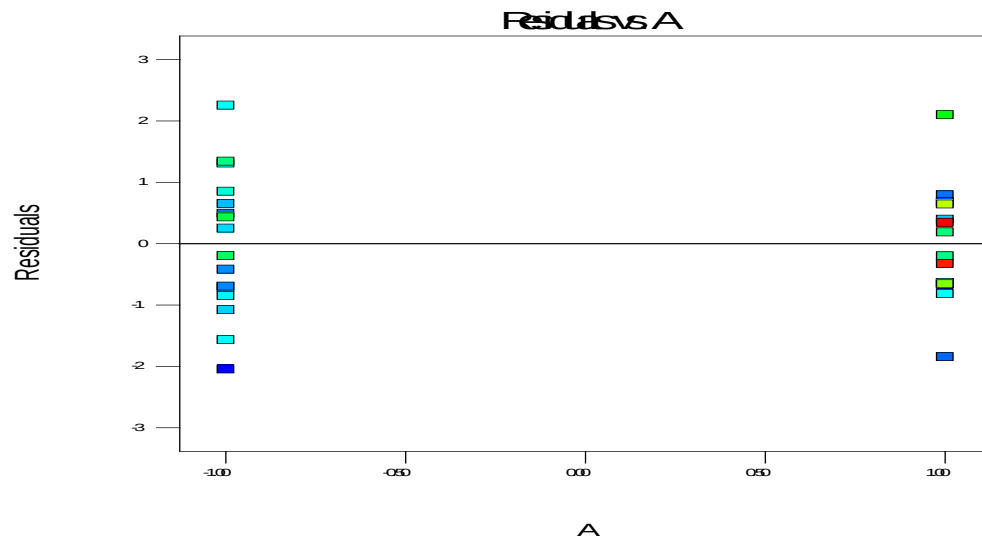
The Model F-value of 49.15 implies the model is significant. There is only a 0.01% chance that a "Model F-Value" this large could occur due to noise.

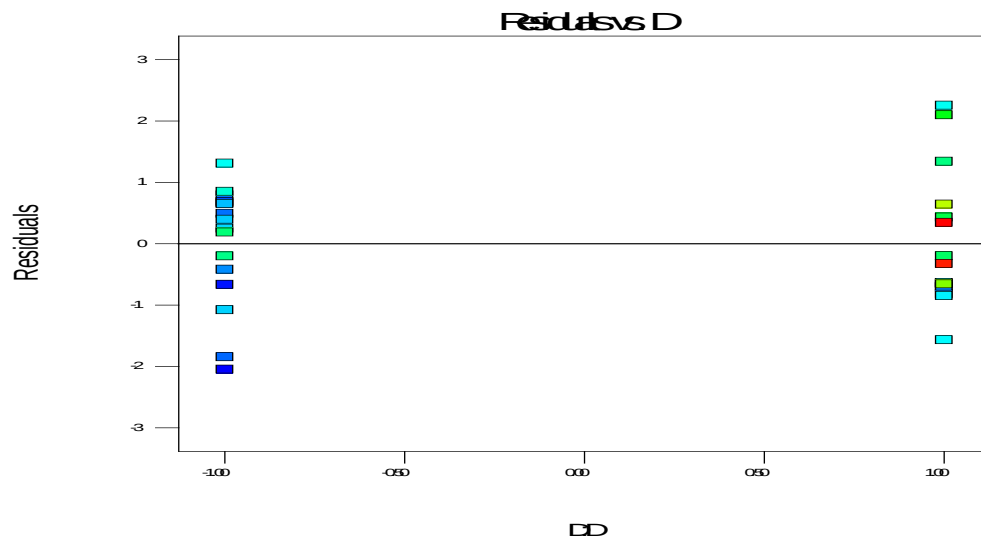
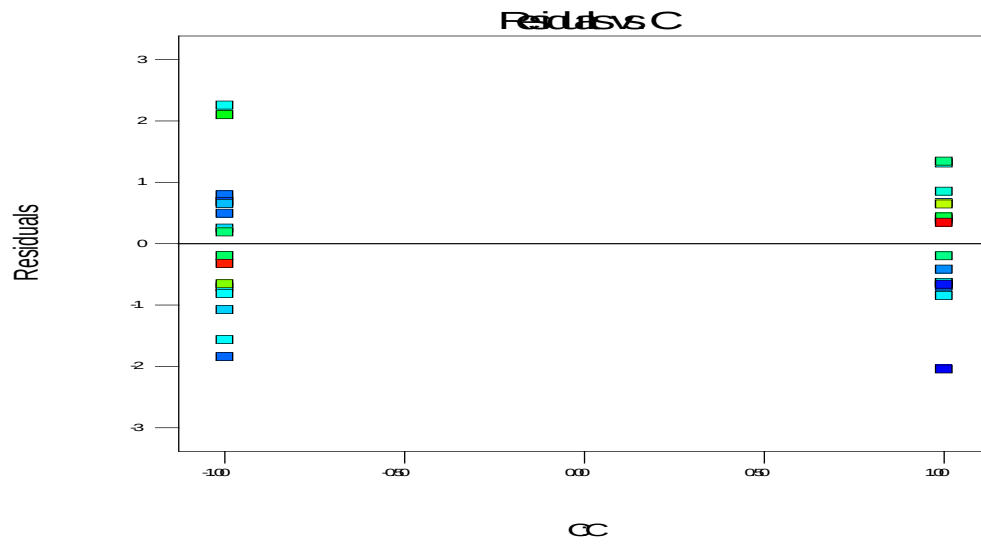
Values of "Prob > F" less than 0.0500 indicate model terms are significant.
 In this case A, D, E, AB, AD, AE, BE, DE, ABE, ADE are significant model terms.

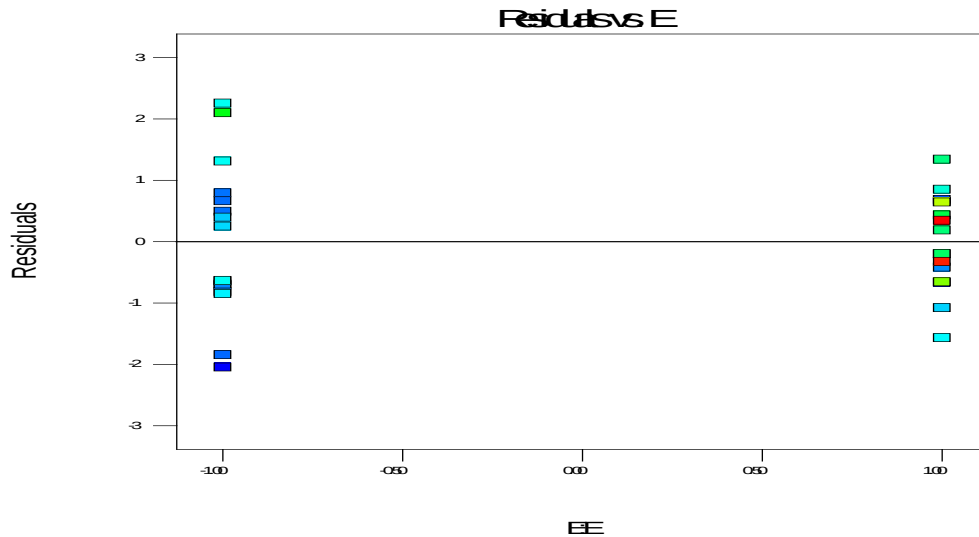
- (b) Analyze the residuals from this experiment. Are there any indications of model inadequacy or violations of the assumptions?

The residual plots below do not identify any concerns with model adequacy or the violations of the assumptions.









- (c) One of the factors from this experiment does not seem to be important. If you drop this factor, what type of design remains? Analyze the data using the full factorial model for only the four active factors. Compare your results with those obtained in part (a).

The resulting experimental design is a replicated 2^4 full factorial design. The ANOVA is shown below. The factor names in the output below were modified to match the factor names in the original problem. The same factors are significant below as were significant in the original analysis.

Design Expert Output

Response 1 y					
ANOVA for selected factorial model					
Analysis of variance table [Partial sum of squares - Type III]					
Source	Sum of Squares	df	Mean Square	F Value	p-value Prob > F
Model	888.80	15	59.25	40.58	< 0.0001
A-A	83.56	1	83.56	57.23	< 0.0001
B-B	0.060	1	0.060	0.041	0.8414
D-D	285.78	1	285.78	195.74	< 0.0001
E-E	153.17	1	153.17	104.91	< 0.0001
AB	48.93	1	48.93	33.51	< 0.0001
AD	88.88	1	88.88	60.88	< 0.0001
AE	33.76	1	33.76	23.13	0.0002
BD	5.778E-003	1	5.778E-003	3.958E-003	0.9506
BE	52.71	1	52.71	36.10	< 0.0001
DE	61.80	1	61.80	42.33	< 0.0001
ABD	3.82	1	3.82	2.61	0.1255
ABE	44.96	1	44.96	30.79	< 0.0001
ADE	26.01	1	26.01	17.82	0.0006
BDE	0.050	1	0.050	0.035	0.8549
ABDE	5.31	1	5.31	3.63	0.0747
Pure Error	23.36	16	1.46		
Cor Total	912.16	31			

The Model F-value of 40.58 implies the model is significant. There is only a 0.01% chance that a "Model F-Value" this large could occur due to noise.

Values of "Prob > F" less than 0.0500 indicate model terms are significant. In this case A, C, D, AB, AC, AD, BD, CD, ABD, ACD are significant model terms.

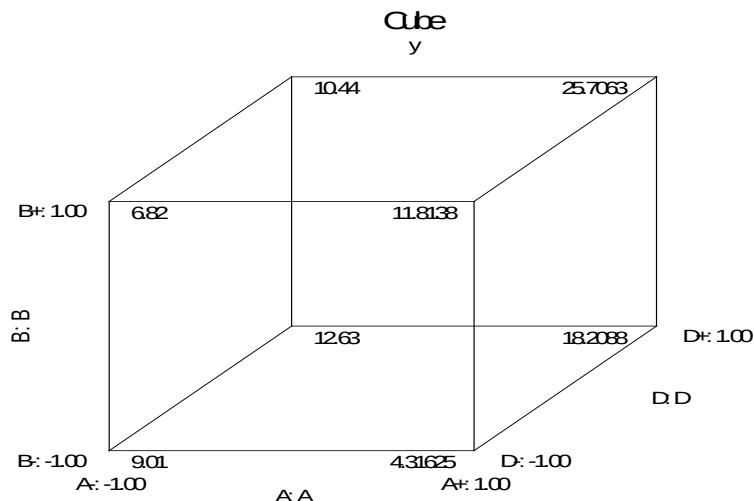
- (d) Find the settings of the active factors that maximize the predicted response.

The cube plot below, with factors E set at +1 and C set at 0, identifies the maximum predicted response with the remaining factors, A , B , and D all set at +1.

Design-Expert® Software
Factor Coding: Actual

y
X1 = A: A
X2 = B: B
X3 = D: D

Actual Factors
C: C = 0.00
E: E = 1.00

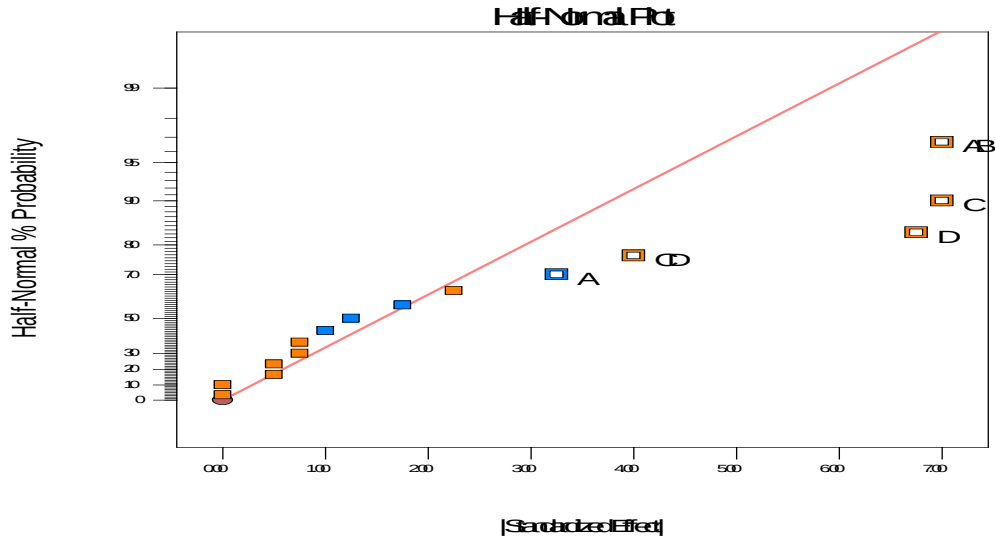


6.40. A paper in the *Journal of Chemical Technology and Biotechnology* ("Response Surface Optimization of the Critical Media Components for the Production of Surfactin," 1997, Vol. 68, pp. 263-270) describes the use of a designed experiment to maximize surfactin production. A portion of the data from this experiment is shown in the table below. Surfactin was assayed by an indirect method, which involves measurement of surface tension of the diluted broth samples. Relative surfactin concentrations were determined by serially diluting the broth until the critical micelle concentration (CMC) was reached. The dilution at which the surface tension starts rising abruptly was denoted by CMC^{-1} and was considered proportional to the amount of surfactant present in the original sample.

Run	Glucose (g dm^{-3})	NH_4NO_3 (g dm^{-3})	FeSO_4 ($\text{g dm}^{-3} \times 10^{-4}$)	MnSO_4 ($\text{g dm}^{-3} \times 10^{-2}$)	y (CMC^{-1})
1	20.00	2.00	6.00	4.00	23
2	60.00	2.00	6.00	4.00	15
3	20.00	6.00	6.00	4.00	16
4	60.00	6.00	6.00	4.00	18
5	20.00	2.00	30.00	4.00	25
6	60.00	2.00	30.00	4.00	16
7	20.00	6.00	30.00	4.00	17
8	60.00	6.00	30.00	4.00	26
9	20.00	2.00	6.00	20.00	28
10	60.00	2.00	6.00	20.00	16
11	20.00	6.00	6.00	20.00	18
12	60.00	6.00	6.00	20.00	21
13	20.00	2.00	30.00	20.00	36
14	60.00	2.00	30.00	20.00	24
15	20.00	6.00	30.00	20.00	33
16	60.00	6.00	30.00	20.00	34

- (a) Analyze the data from this experiment. Identify the significant factors and interactions.

The half normal probability plot of effects, followed by the ANOVA, identify the significant factors and interactions. Although factor *B* is not significant, the *AB* interaction is.



Design Expert Output

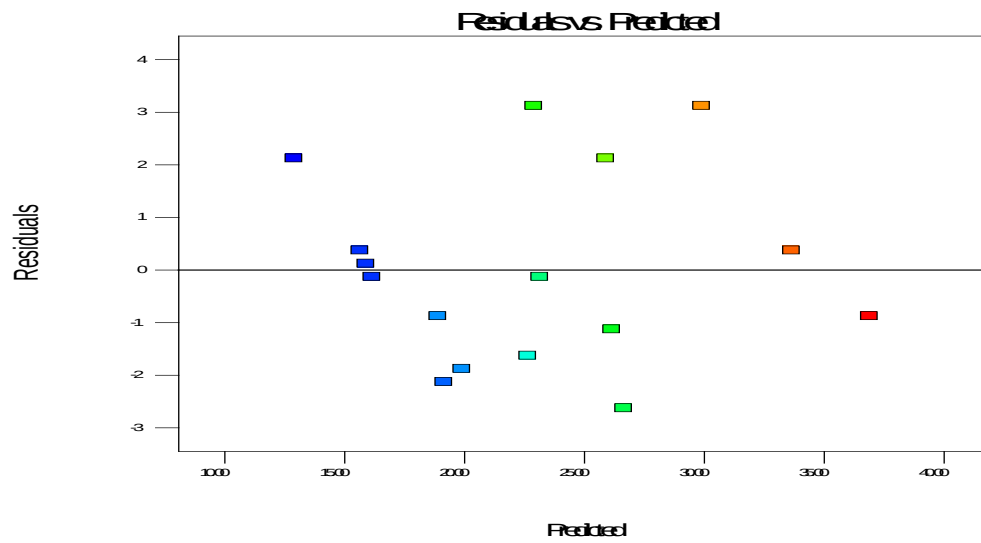
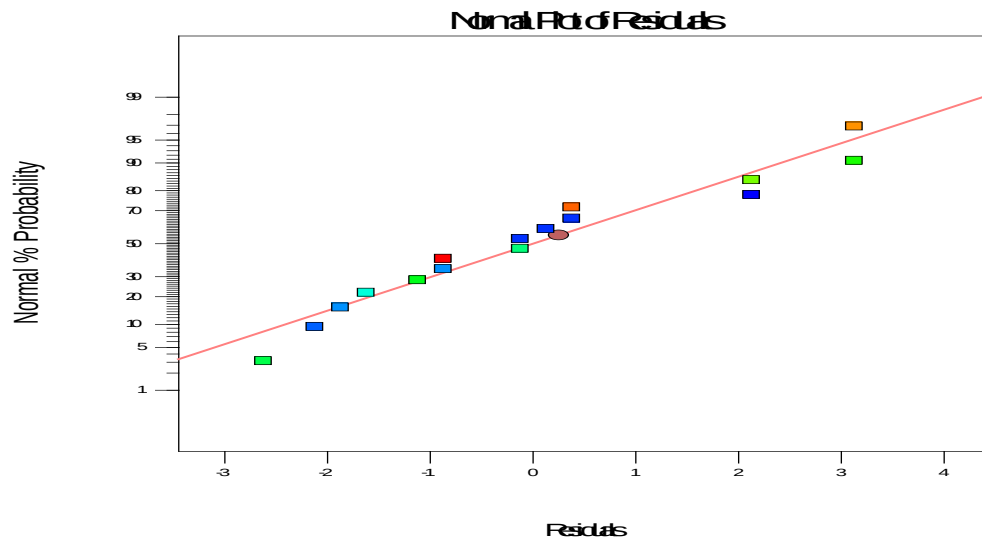
Response 1 y					
ANOVA for selected factorial model					
Analysis of variance table [Partial sum of squares - Type III]					
Source	Sum of Squares	df	Mean Square	F Value	p-value Prob > F
Model	680.50	6	113.42	20.73	< 0.0001
A-Glucose	42.25	1	42.25	7.72	0.0214
B-NH ₄ NO ₃	0.000	1	0.000	0.000	1.0000
C-FeSO ₄	196.00	1	196.00	35.82	0.0002
D-MnSO ₄	182.25	1	182.25	33.30	0.0003
AB	196.00	1	196.00	35.82	0.0002
CD	64.00	1	64.00	11.70	0.0076
Residual	49.25	9	5.47		
Cor Total	729.75	15			

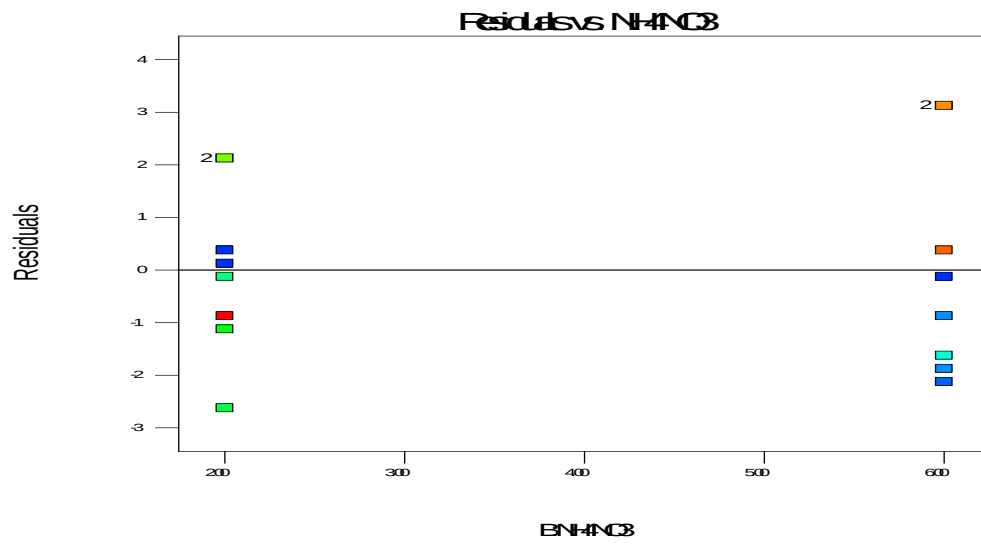
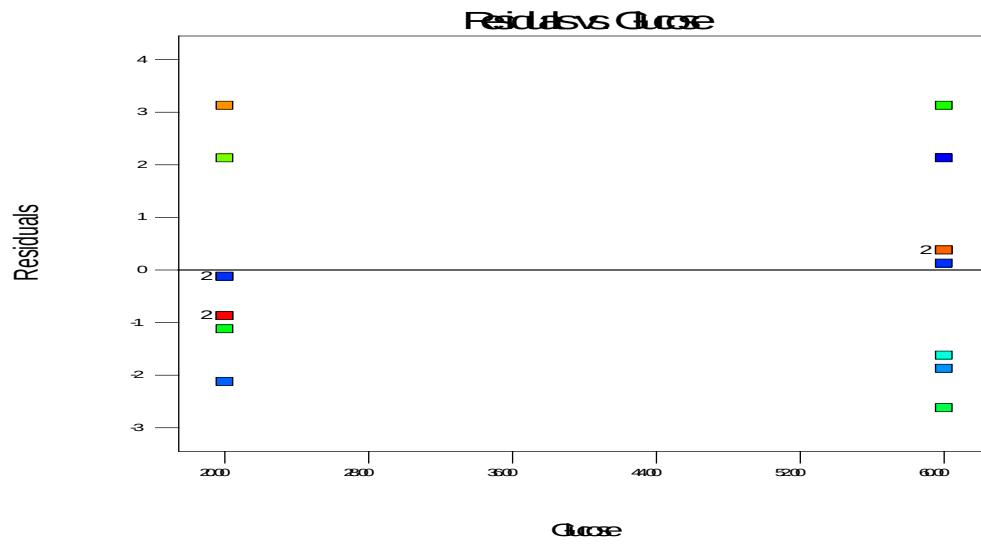
The Model F-value of 20.73 implies the model is significant. There is only a 0.01% chance that a "Model F-Value" this large could occur due to noise.

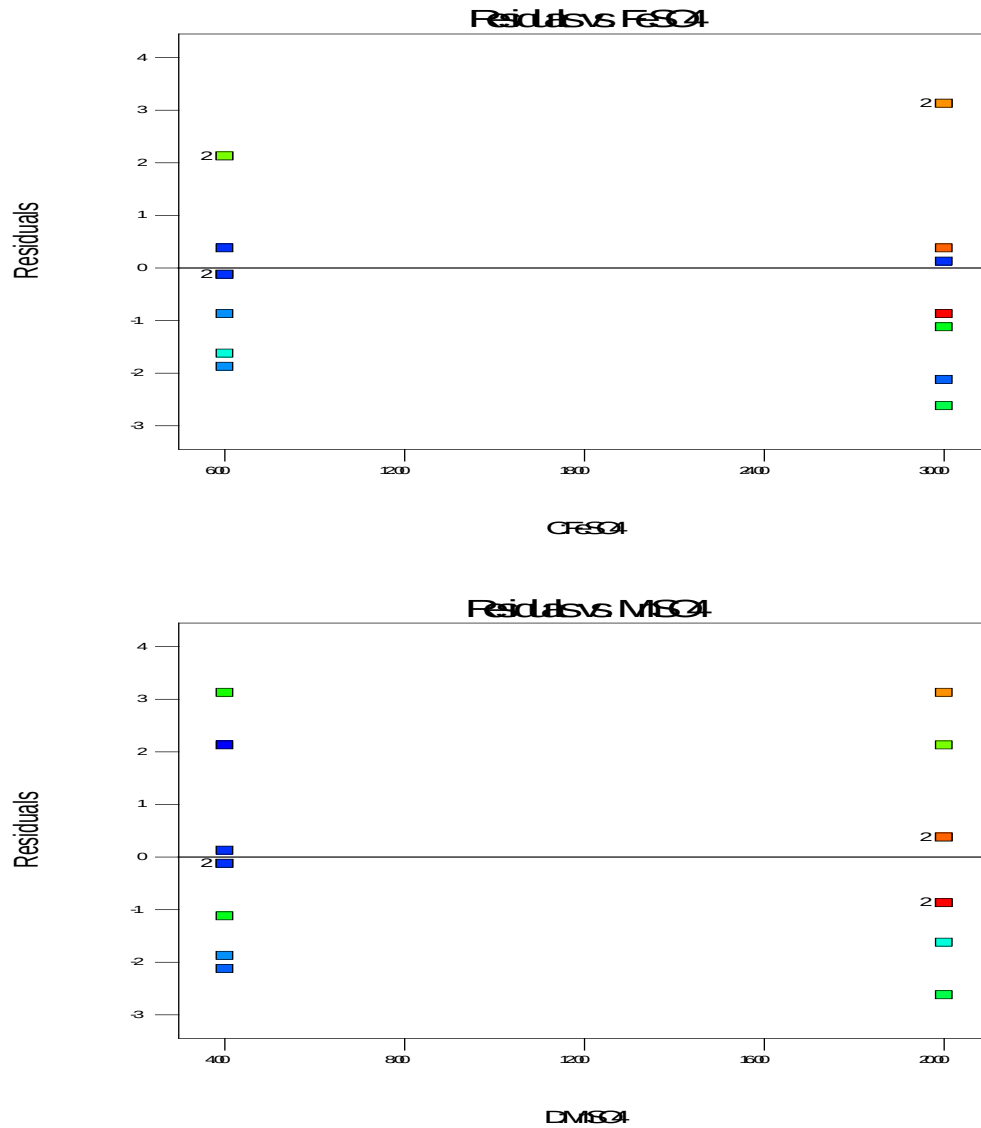
Values of "Prob > F" less than 0.0500 indicate model terms are significant. In this case A, C, D, AB, CD are significant model terms.

- (b) Analyze the residuals from this experiment. Are there any indications of model inadequacy or violations of the assumptions?

The residual plots below do not identify any concerns with model adequacy or the violations of the assumptions.



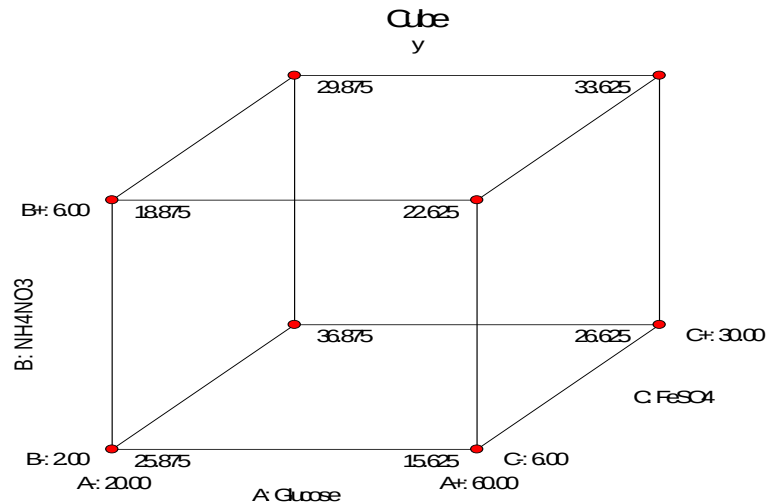




(c) What conditions would optimize the surfactin production?

The response, y , is maximized when factor A is 20, B is 2, C is 30, and D is 20.

Design-Expert® Software
 Factor Coding: Actual
 y
 X1 = A: Glucose
 X2 = B: NH₄NO₃
 X3 = C: FeSO₄
 Actual Factor
 D: MnSO₄ = 20.00



6.41. Continuation of Problem 6.40. The experiment in Problem 6.40 actually included six center points. The responses at these conditions were 35, 35, 35, 36, 36, and 34. Is there any indication of curvature in the response function? Are additional experiments necessary? What would you recommend doing now?

Curvature appears to be very significant with a p value less than 0.0001. Axial runs, along with additional center point runs to identify blocking effects, should be run.

Design Expert Output

Response	1	y				
ANOVA for selected factorial model						
Analysis of variance table [Partial sum of squares - Type III]						
Source	Sum of Squares	df	Mean Square	F Value	p-value Prob > F	
Model	713.00	8	89.12	54.61	< 0.0001	significant
A-Glucose	42.25	1	42.25	25.89	0.0003	
B-NH4NO3	0.000	1	0.000	0.000	1.0000	
C-FeSO4	196.00	1	196.00	120.10	< 0.0001	
D-MnSO4	182.25	1	182.25	111.68	< 0.0001	
AB	196.00	1	196.00	120.10	< 0.0001	
AD	12.25	1	12.25	7.51	0.0179	
BC	20.25	1	20.25	12.41	0.0042	
CD	64.00	1	64.00	39.22	< 0.0001	
Curvature	659.28	1	659.28	403.98	< 0.0001	significant
Residual	19.58	12	1.63			
Lack of Fit	16.75	7	2.39	4.22	0.0660	not significant
Pure Error	2.83	5	0.57			
Cor Total	1391.86	21				

6.42. An article in the *Journal of Hazardous Materials* ("Feasibility of Using Natural Fishbone Apatite as a Substitute for Hydroxyapatite in Remediating Aqueous Heavy Metals," Vol. 69, Issue 2, 1999, pp. 187-197) describes an experiment to study the suitability of fishbone, a natural, apatite rich substance, as a substitute for hydroxyapatite in the sequestering of aqueous divalent heavy metal ions. Direct comparison of hydroxyapatite and fishbone apatite was performed using a three-factor two-level full factorial design. Apatite (30 or 60 mg) was added to 100mL deionized water and gently agitated overnight in a shaker. The pH was then adjusted to 5 or 7 using nitric acid. Sufficient concentration of lead nitrate solution was added

to each flask to result in a final volume of 200 mL and a lead concentration of 0.483 or 2.41 mM, respectively. The experiment was a 2^3 replicated twice and it was performed for both fishbone and synthetic apatite. Results are shown below:

Apatite	pH	Pb	Fishbone		Hydroxyapatite	
			Pb,mM	pH	Pb,mM	pH
+	+	+	1.82	5.22	0.11	3.49
+	+	+	1.81	5.12	0.12	3.46
+	+	-	0.01	6.84	0.00	5.84
+	+	-	0.00	6.61	0.00	5.90
+	-	+	1.11	3.35	0.80	2.70
+	-	+	1.04	3.34	0.76	2.74
+	-	-	0.00	5.77	0.03	3.36
+	-	-	0.01	6.25	0.05	3.24
-	+	+	2.11	5.29	1.03	3.22
-	+	+	2.18	5.06	1.05	3.22
-	+	-	0.03	5.93	0.00	5.53
-	+	-	0.05	6.02	0.00	5.43
-	-	+	1.70	3.39	1.34	2.82
-	-	+	1.69	3.34	1.26	2.79
-	-	-	0.05	4.50	0.06	3.28
-	-	-	0.05	4.74	0.07	3.28

(a) Analyze the lead response for fishbone apatite. What factors are important?

As shown below, all main effects and interactions are significant.

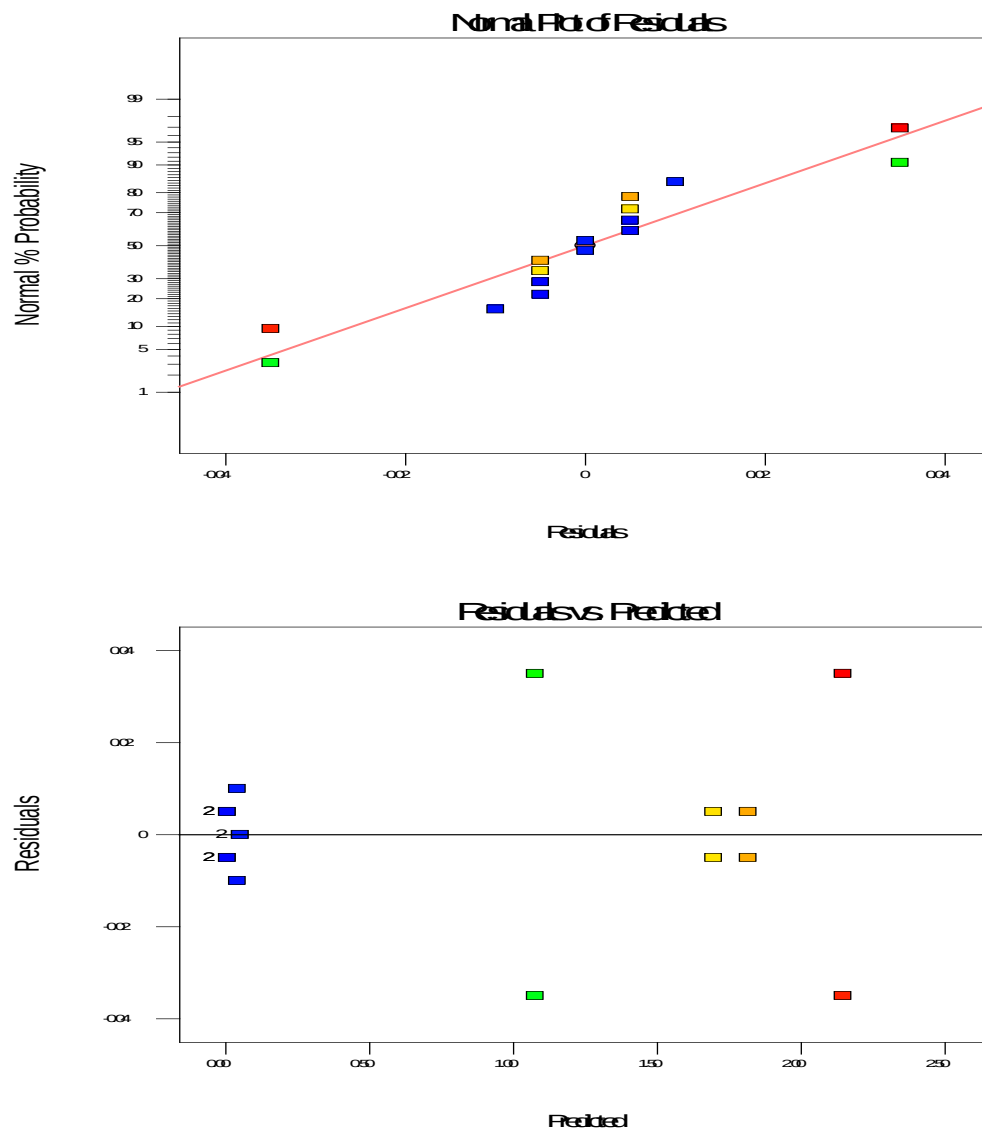
Design Expert Output

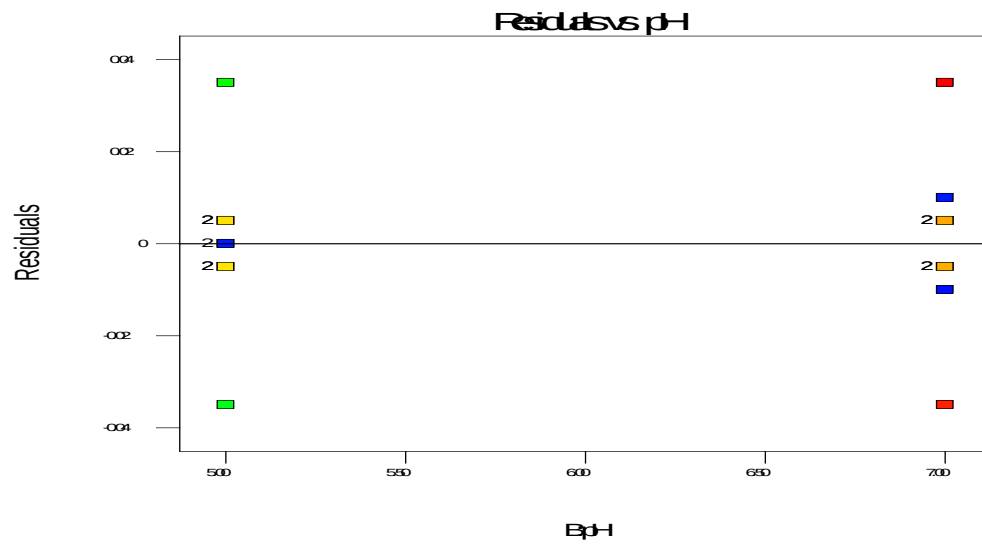
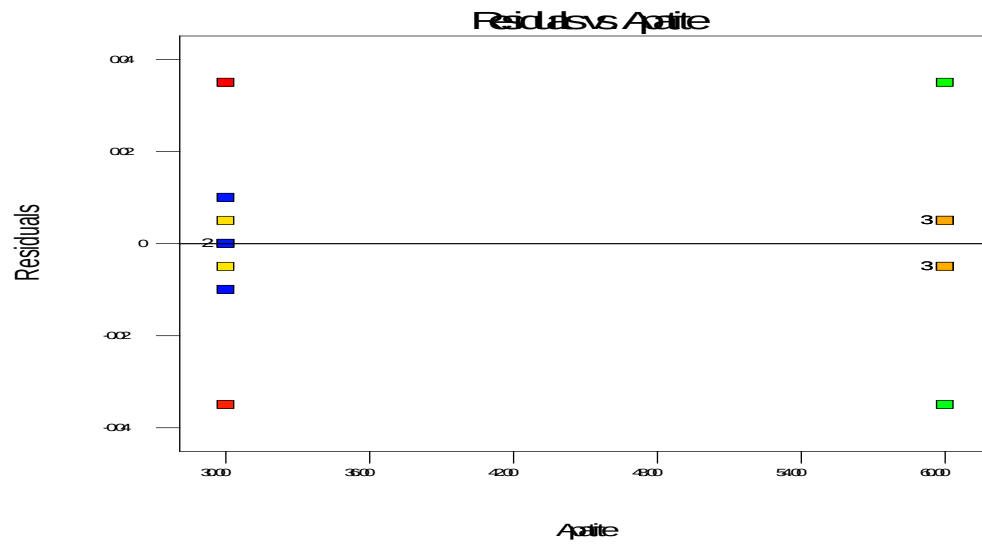
Response 1 Fishbone Pb						
ANOVA for selected factorial model						
Analysis of variance table [Partial sum of squares - Type III]						
Source	Sum of Squares	df	Mean Square	F Value	p-value Prob > F	
Model	12.19	7	1.74	2629.41	< 0.0001	significant
A-Apatite	0.27	1	0.27	400.34	< 0.0001	
B-pH	0.35	1	0.35	525.43	< 0.0001	
C-Pb	10.99	1	10.99	16587.51	< 0.0001	
AB	0.023	1	0.023	33.96	0.0004	
AC	0.19	1	0.19	285.62	< 0.0001	
BC	0.36	1	0.36	543.40	< 0.0001	
ABC	0.020	1	0.020	29.58	0.0006	
Pure Error	5.300E-003	8	6.625E-004			
Cor Total	12.20	15				

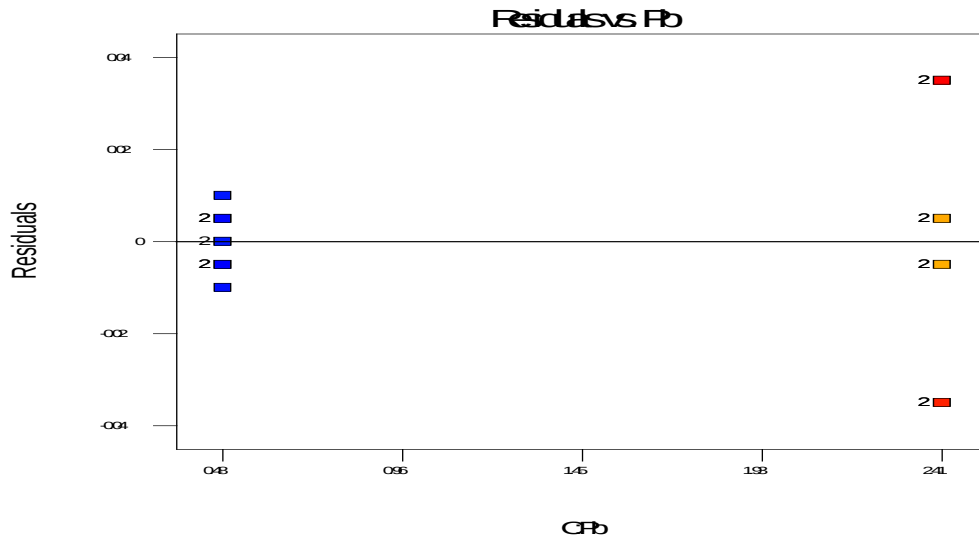
Factor	Coefficient Estimate	df	Standard Error	95% CI Low	95% CI High	VIF
Intercept	0.85	1	6.435E-003	0.84	0.87	
A-Apatite	-0.13	1	6.435E-003	-0.14	-0.11	1.00
B-pH	0.15	1	6.435E-003	0.13	0.16	1.00
C-Pb	0.83	1	6.435E-003	0.81	0.84	1.00
AB	0.038	1	6.435E-003	0.023	0.052	1.00
AC	-0.11	1	6.435E-003	-0.12	-0.094	1.00
BC	0.15	1	6.435E-003	0.14	0.16	1.00
ABC	0.035	1	6.435E-003	0.020	0.050	1.00

(b) Analyze the residuals from this response and comment on model adequacy.

The normal plot identifies slightly thicker tails in the distribution of the residuals. The plots of residuals vs. predicted and residuals vs. the Pb effect identify nonconstant variance.







(c) Analyze the pH response for fishbone apatite. What factors are important?

The *AB* and *ABC* interactions are only moderately significant; all other main effects and interactions are significant.

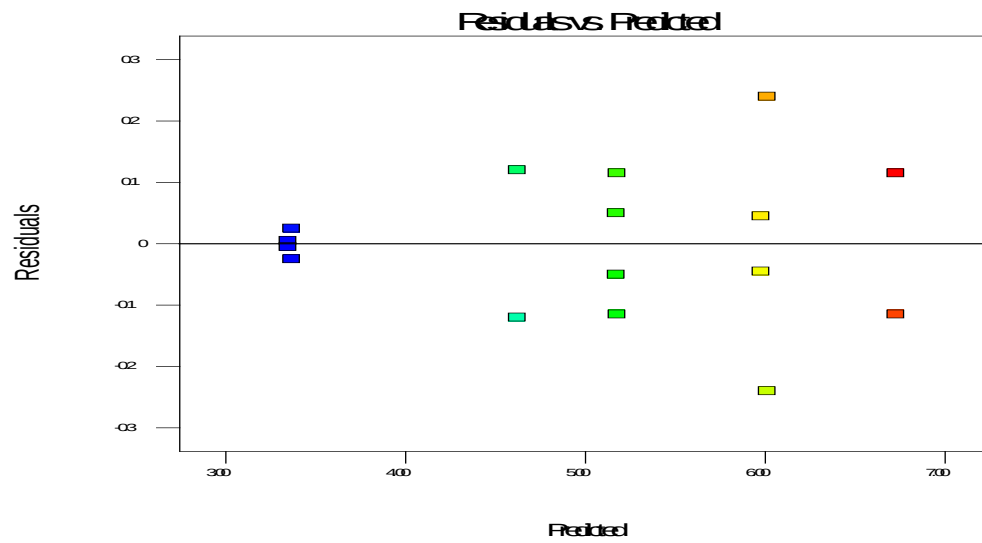
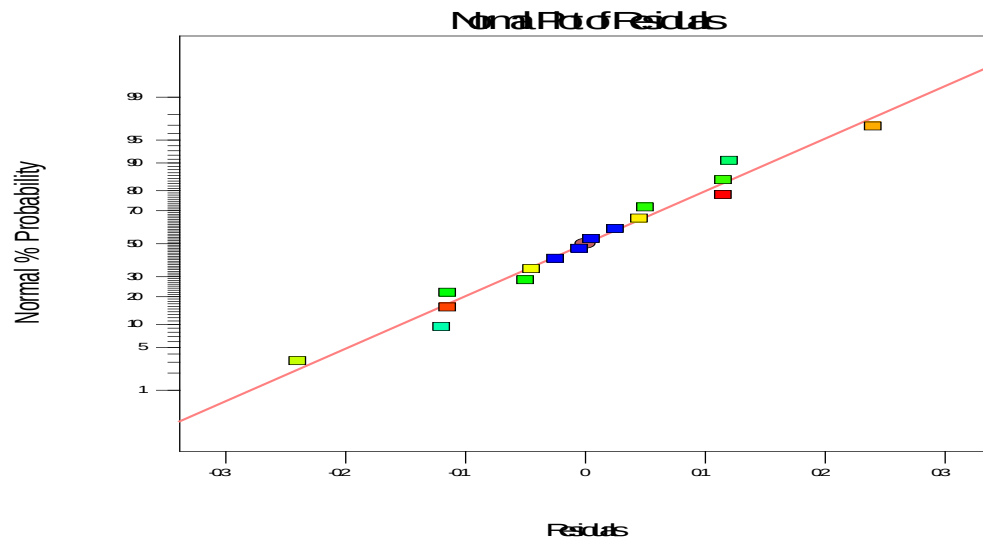
Design Expert Output

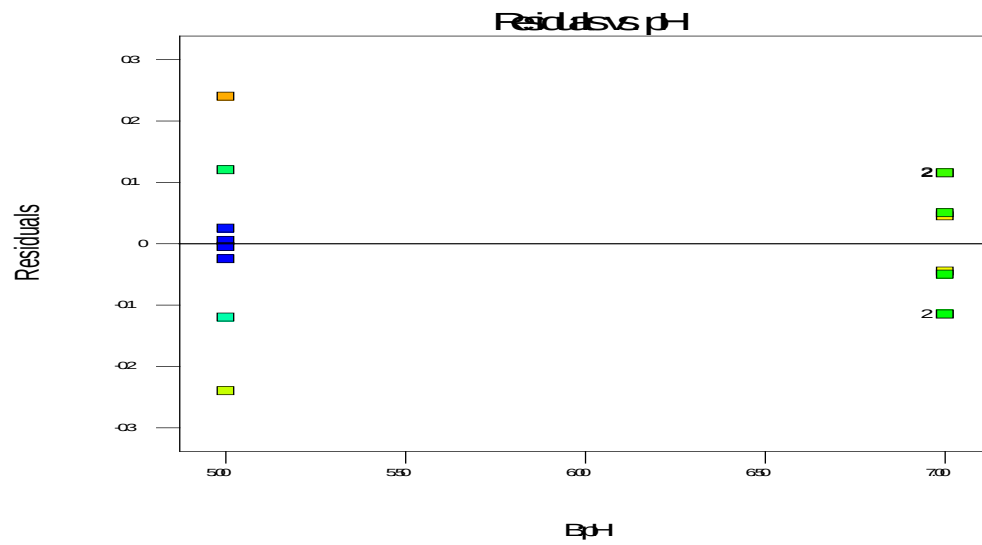
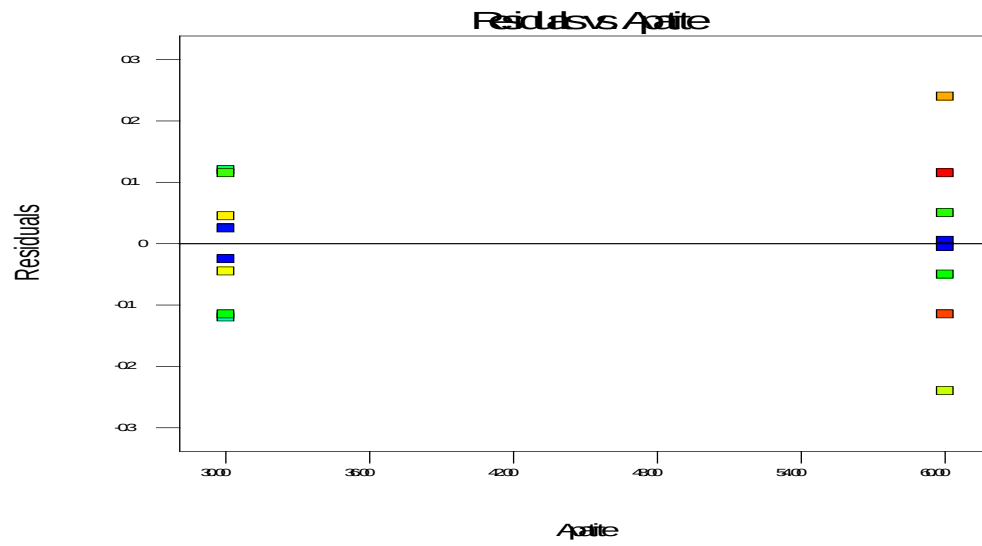
Response 1 Fishbone pH						
ANOVA for selected factorial model						
Analysis of variance table [Partial sum of squares - Type III]						
Source	Sum of Squares	df	Mean Square	F Value	p-value Prob > F	
Model	21.09	7	3.01	116.29	< 0.0001	significant
A-Apatite	1.12	1	1.12	43.17	0.0002	
B-pH	8.14	1	8.14	314.08	< 0.0001	
C-Pb	9.84	1	9.84	379.98	< 0.0001	
AB	0.098	1	0.098	3.77	0.0881	
AC	1.17	1	1.17	45.23	0.0001	
BC	0.61	1	0.61	23.64	0.0013	
ABC	0.11	1	0.11	4.14	0.0763	
Pure Error	0.21	8	0.026			
Cor Total	21.30	15				

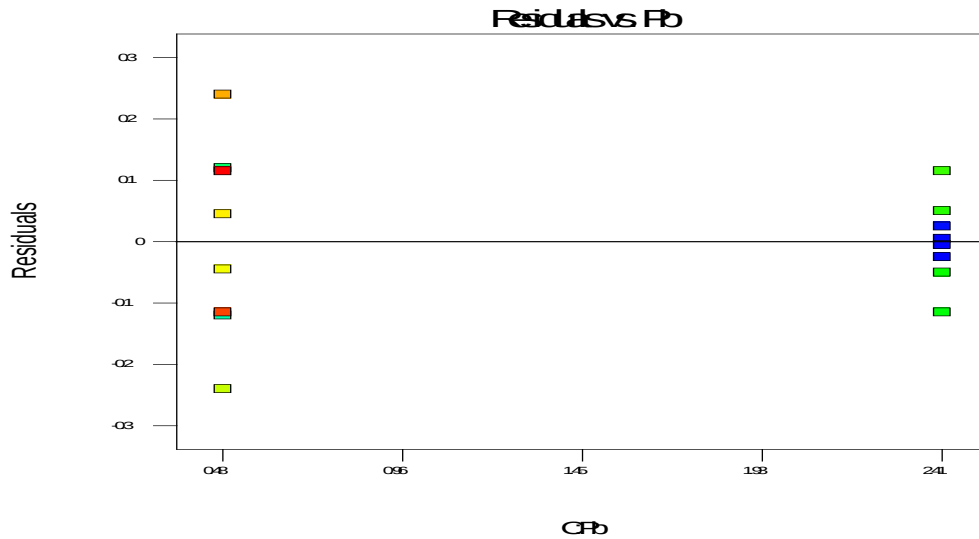
Factor	Coefficient Estimate	df	Standard Error	95% CI Low	95% CI High	VIF
Intercept	5.05	1	0.040	4.96	5.14	
A-Apatite	0.26	1	0.040	0.17	0.36	1.00
B-pH	0.71	1	0.040	0.62	0.81	1.00
C-Pb	-0.78	1	0.040	-0.88	-0.69	1.00
AB	-0.078	1	0.040	-0.17	0.015	1.00
AC	-0.27	1	0.040	-0.36	-0.18	1.00
BC	0.20	1	0.040	0.10	0.29	1.00
ABC	0.082	1	0.040	-0.011	0.17	1.00

(d) Analyze the residuals from this response and comment on model adequacy.

Although the normal probability plot is acceptable, the plots of residuals vs. predicted identifies nonconstant variance.







(e) Analyze the lead response for hydroxyapatite apatite. What factors are important?

As shown below, all main effects and interactions are significant.

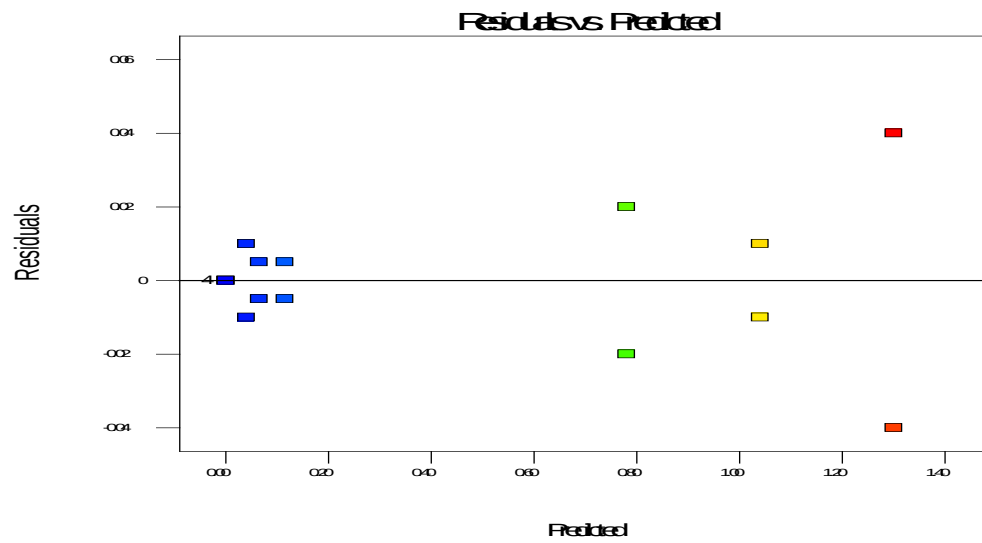
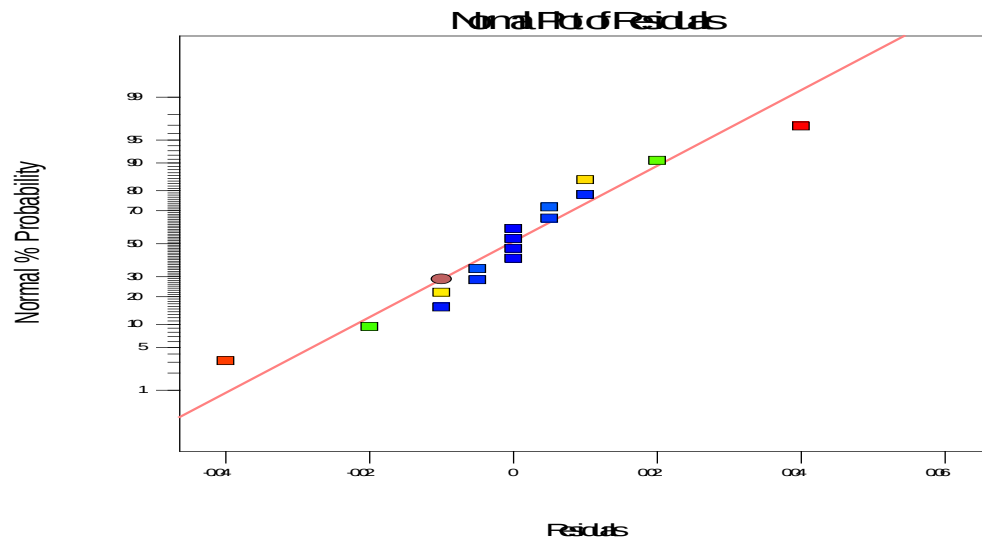
Design Expert Output

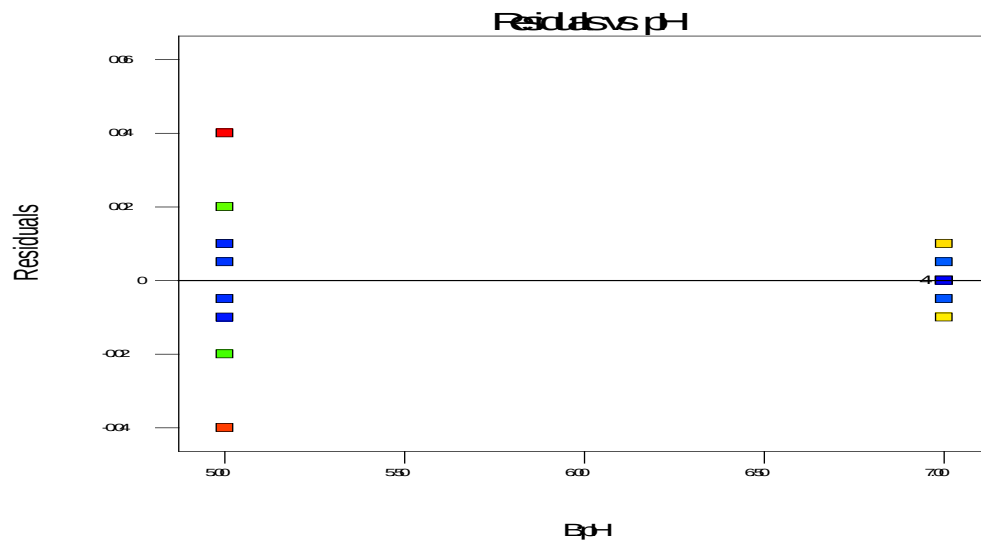
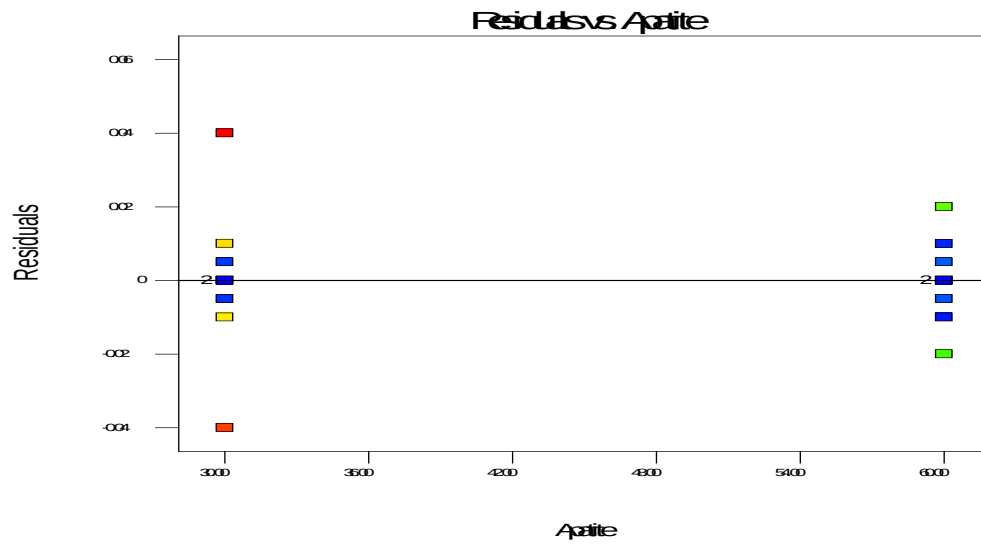
Response 1 Hydroxyapatite Pb						
ANOVA for selected factorial model						
Analysis of variance table [Partial sum of squares - Type III]						
Source	Sum of Squares	df	Mean Square	F Value	p-value Prob > F	
Model	4.01	7	0.57	1018.21	< 0.0001	significant
A-Apatite	0.54	1	0.54	960.40	< 0.0001	
B-pH	0.27	1	0.27	471.51	< 0.0001	
C-Pb	2.45	1	2.45	4354.18	< 0.0001	
AB	0.036	1	0.036	64.18	< 0.0001	
AC	0.50	1	0.50	896.18	< 0.0001	
BC	0.17	1	0.17	298.84	< 0.0001	
ABC	0.046	1	0.046	82.18	< 0.0001	
Pure Error	4.500E-003	8	5.625E-004			
Cor Total	4.01	15				

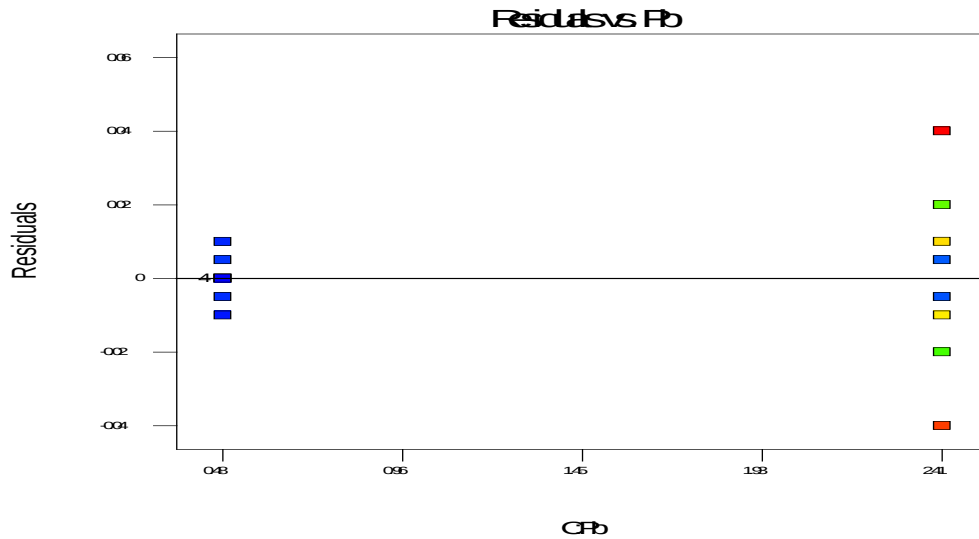
Factor	Coefficient Estimate	df	Standard Error	95% CI Low	95% CI High	VIF
Intercept	0.42	1	5.929E-003	0.40	0.43	
A-Apatite	-0.18	1	5.929E-003	-0.20	-0.17	1.00
B-pH	-0.13	1	5.929E-003	-0.14	-0.12	1.00
C-Pb	0.39	1	5.929E-003	0.38	0.40	1.00
AB	-0.048	1	5.929E-003	-0.061	-0.034	1.00
AC	-0.18	1	5.929E-003	-0.19	-0.16	1.00
BC	-0.10	1	5.929E-003	-0.12	-0.089	1.00
ABC	-0.054	1	5.929E-003	-0.067	-0.040	1.00

(f) Analyze the residuals from this response and comment on model adequacy.

The normal plot identifies slightly thicker tails in the distribution of the residuals. The plots of residuals vs. predicted and residuals vs. the effects identifies nonconstant variance.







(g) Analyze the pH response for hydroxyapatite apatite. What factors are important?

The ABC interaction is not significant; all of the main effects and two factor interactions are significant.

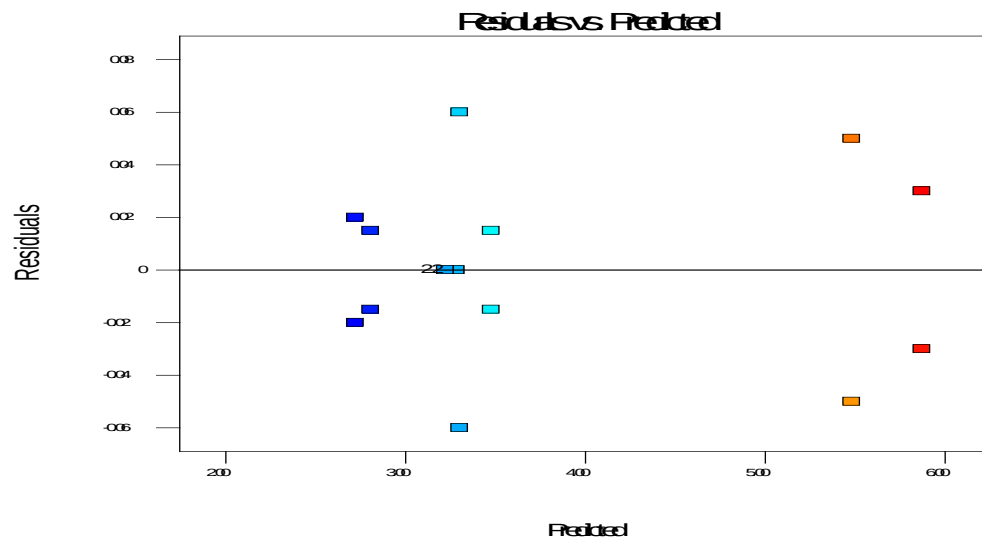
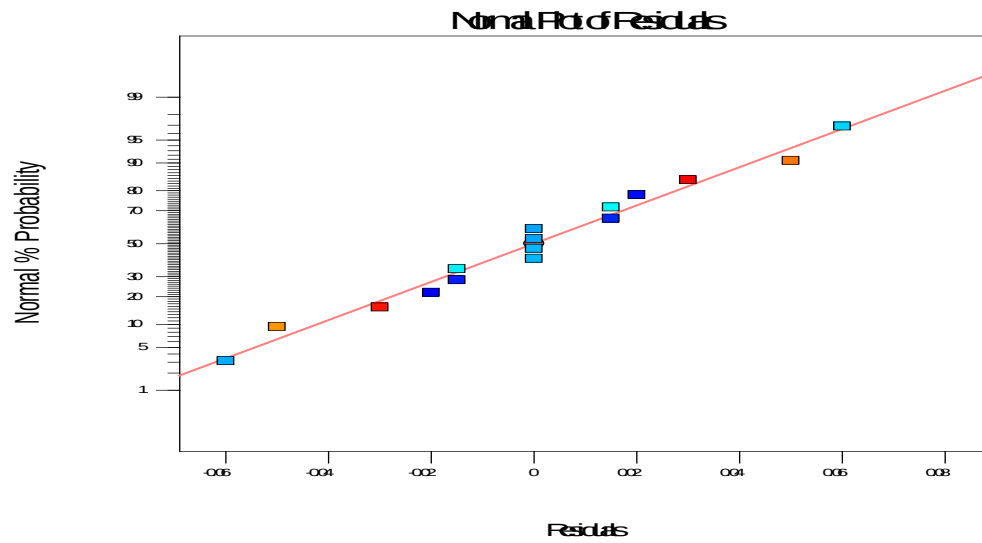
Design Expert Output

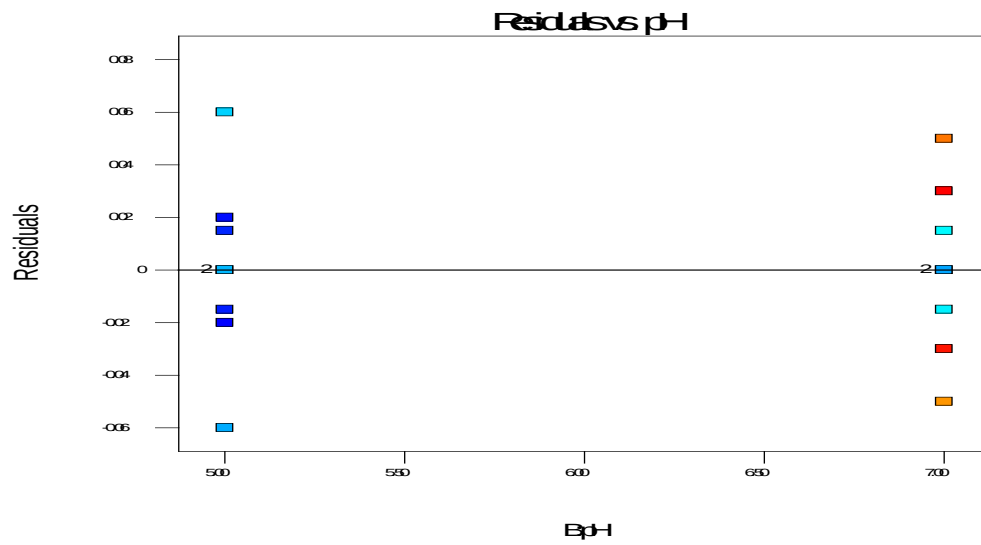
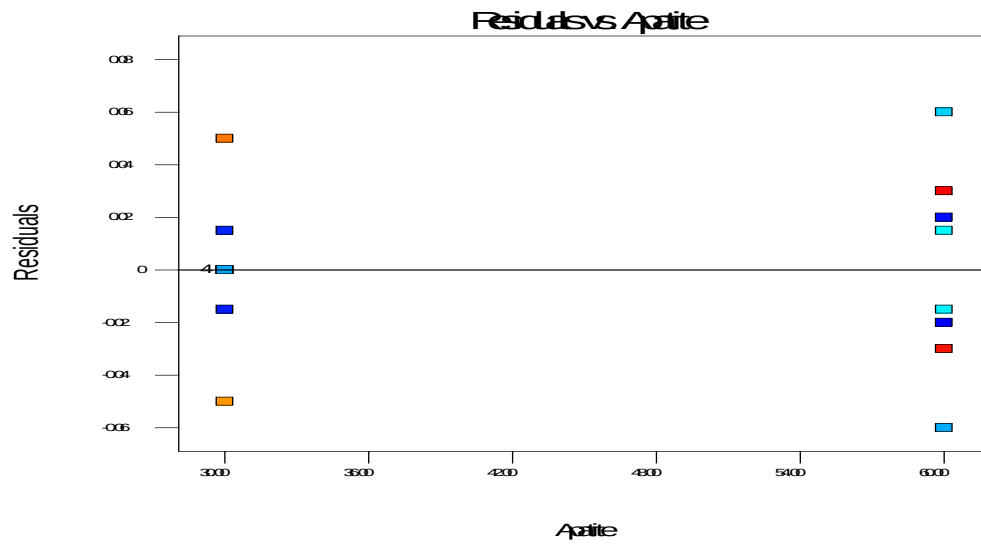
Response 1 Hydroxyapatite pH						
ANOVA for selected factorial model						
Analysis of variance table [Partial sum of squares - Type III]						
Source	Sum of Squares	df	Mean Square	F Value	p-value Prob > F	
Model	20.44	7	2.92	1487.66	< 0.0001	significant
A-Apatite	0.084	1	0.084	42.85	0.0002	
B-pH	8.82	1	8.82	4494.73	< 0.0001	
C-Pb	8.15	1	8.15	4153.39	< 0.0001	
AB	0.13	1	0.13	64.22	< 0.0001	
AC	0.014	1	0.014	7.34	0.0267	
BC	3.24	1	3.24	1650.96	< 0.0001	
ABC	2.250E-004	1	2.250E-004	0.11	0.7436	
Pure Error	0.016	8	1.963E-003			
Cor Total	20.45	15				

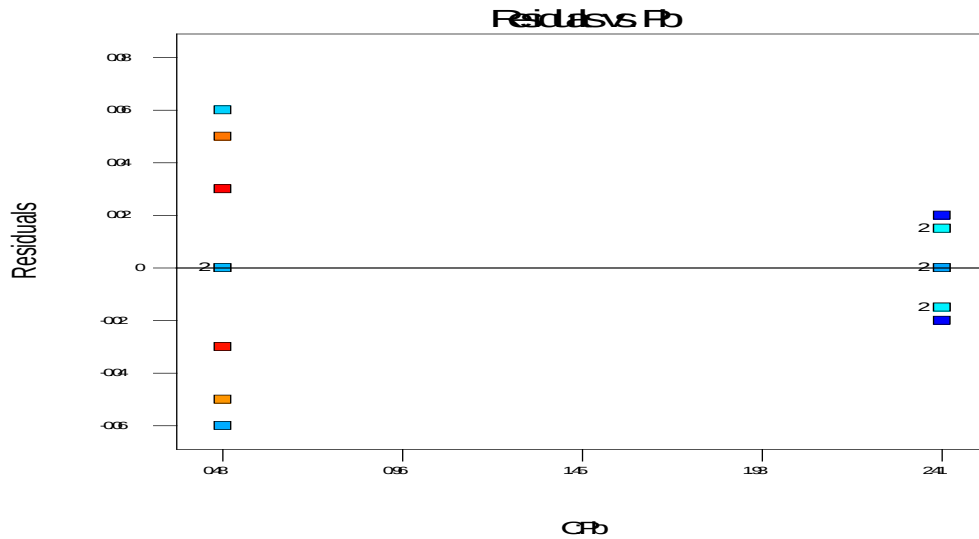
Factor	Coefficient Estimate	df	Standard Error	95% CI Low	95% CI High	VIF
Intercept	3.77	1	0.011	3.74	3.79	
A-Apatite	0.072	1	0.011	0.047	0.098	1.00
B-pH	0.74	1	0.011	0.72	0.77	1.00
C-Pb	-0.71	1	0.011	-0.74	-0.69	1.00
AB	0.089	1	0.011	0.063	0.11	1.00
AC	-0.030	1	0.011	-0.056	-4.461E-003	1.00
BC	-0.45	1	0.011	-0.48	-0.42	1.00
ABC	-3.750E-003	1	0.011	-0.029	0.022	1.00

(h) Analyze the residuals from this response and comment on model adequacy.

The only potential concern with the residual plots is the nonconstant variance shown in the plot of residuals vs. pH.







- (i) What differences do you see between fishbone and hydroxyapatite apatite? The authors of this paper concluded that fishbone apatite was comparable to hydroxyapatite apatite. Because the fishbone apatite is cheaper, it was recommended for adoption. Do you agree with these conclusions?

The authors of the journal article did not show their analysis for this experiment. When comparing the Fishbone and Hydroxyapatite models main effects and interactions for the Pb and pH responses, we might disagree with the authors.

A more effective approach to understand the differences between Fishbone and Hydroxyapatite would be to include this as a factor in the experimental design. The modified table is shown below followed by the analysis.

Apatite	pH	Pb	Type	Pb,mM	pH
+	+	+	Fishbone	1.82	5.22
+	+	+	Fishbone	1.81	5.12
+	+	-	Fishbone	0.01	6.84
+	+	-	Fishbone	0.00	6.61
+	-	+	Fishbone	1.11	3.35
+	-	+	Fishbone	1.04	3.34
+	-	-	Fishbone	0.00	5.77
+	-	-	Fishbone	0.01	6.25
-	+	+	Fishbone	2.11	5.29
-	+	+	Fishbone	2.18	5.06
-	+	-	Fishbone	0.03	5.93
-	+	-	Fishbone	0.05	6.02
-	-	+	Fishbone	1.70	3.39
-	-	+	Fishbone	1.69	3.34
-	-	-	Fishbone	0.05	4.50
-	-	-	Fishbone	0.05	4.74
+	+	+	Hydroxyapatite	0.11	3.49
+	+	+	Hydroxyapatite	0.12	3.46

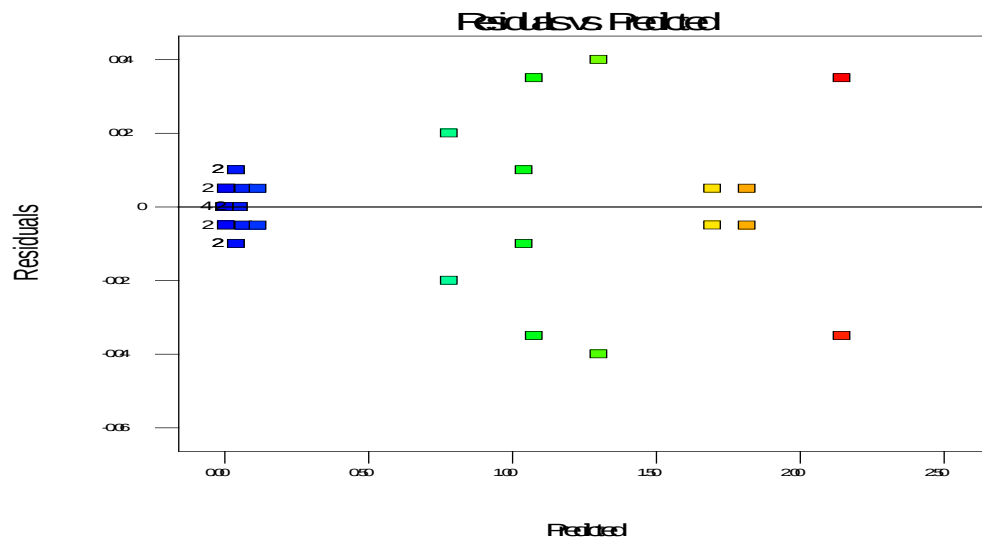
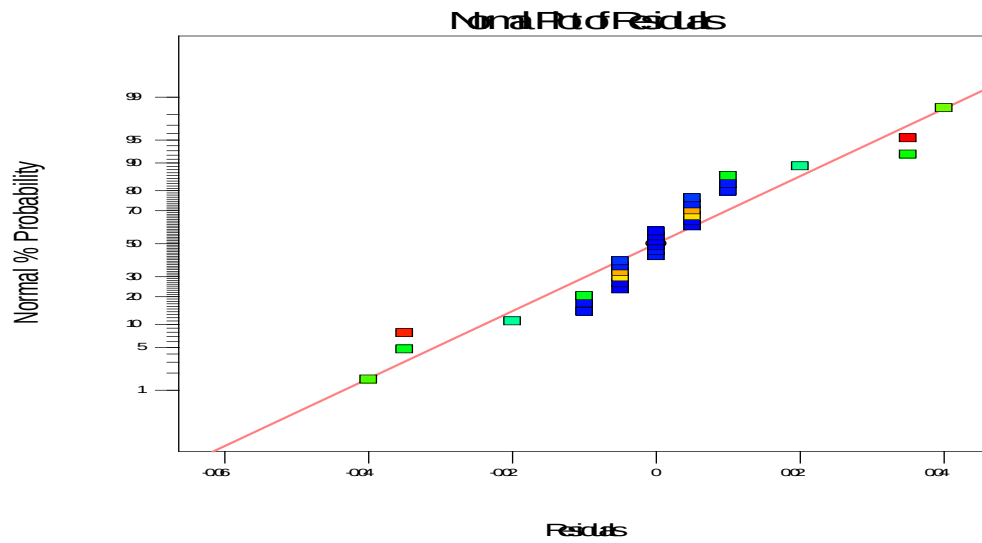
+	+	-	Hydroxyapatite	0.00	5.84
+	+	-	Hydroxyapatite	0.00	5.90
+	-	+	Hydroxyapatite	0.80	2.70
+	-	+	Hydroxyapatite	0.76	2.74
+	-	-	Hydroxyapatite	0.03	3.36
+	-	-	Hydroxyapatite	0.05	3.24
-	+	+	Hydroxyapatite	1.03	3.22
-	+	+	Hydroxyapatite	1.05	3.22
-	+	-	Hydroxyapatite	0.00	5.53
-	+	-	Hydroxyapatite	0.00	5.43
-	-	+	Hydroxyapatite	1.34	2.82
-	-	+	Hydroxyapatite	1.26	2.79
-	-	-	Hydroxyapatite	0.06	3.28
-	-	-	Hydroxyapatite	0.07	3.28

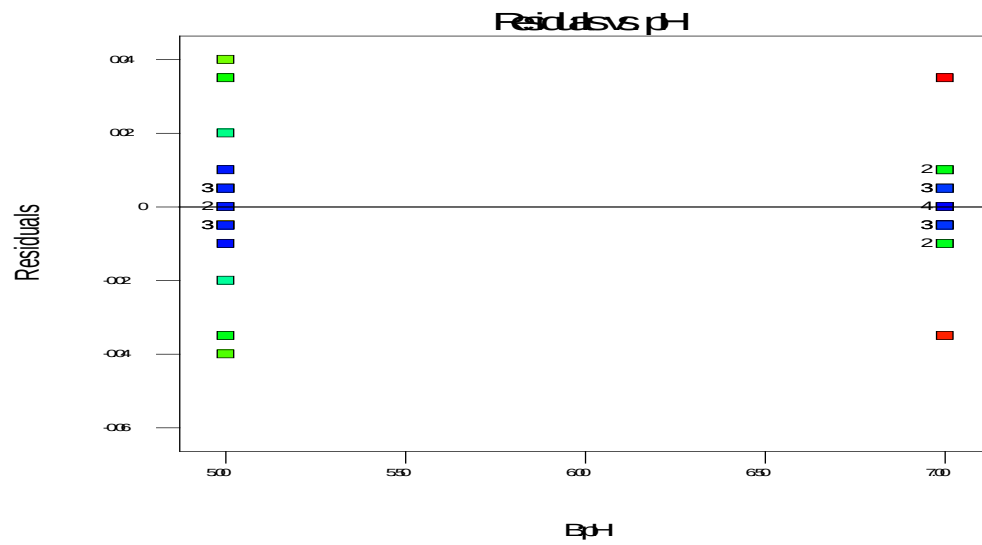
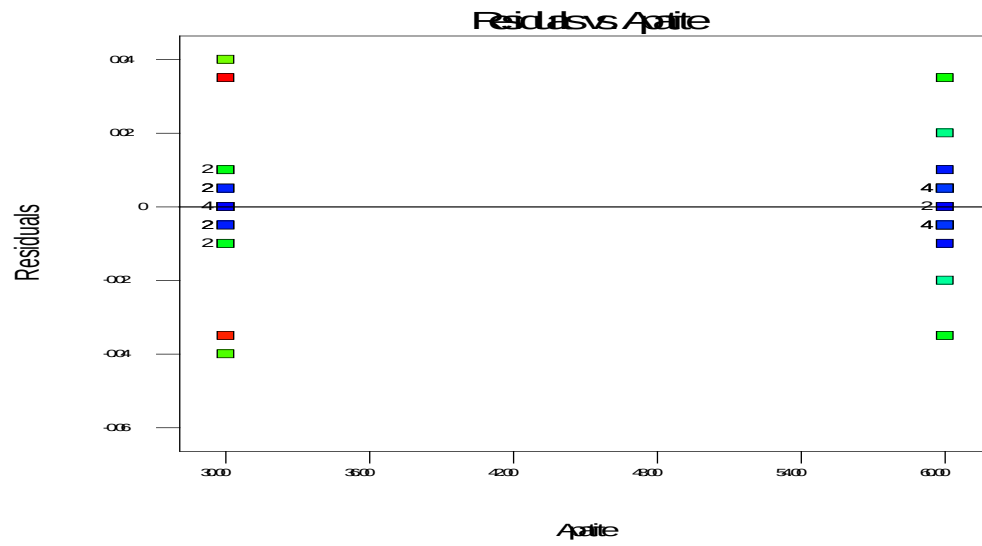
The ANOVA below identifies factor *D*, the type of apatite, as being very significant.

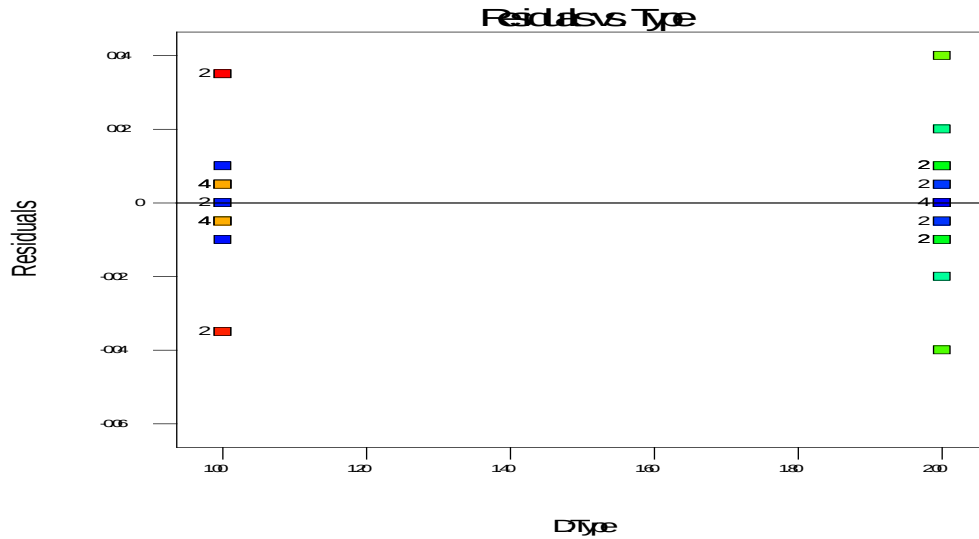
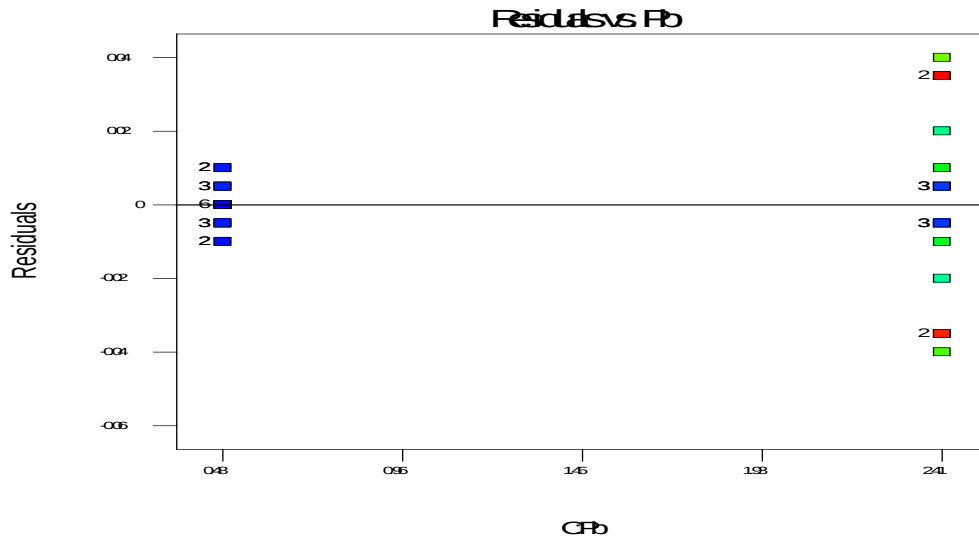
Design Expert Output

Response	1	Pb Response				
ANOVA for selected factorial model						
Analysis of variance table [Partial sum of squares - Type III]						
Source	Sum of Squares	df	Mean Square	F Value	p-value Prob > F	
Model	17.73	15	1.18	1929.32	< 0.0001	significant
A-Apatite	0.78	1	0.78	1275.51	< 0.0001	
B-pH	2.813E-003	1	2.813E-003	4.59	0.0478	
C-Pb	11.91	1	11.91	19440.33	< 0.0001	
D-Type	1.52	1	1.52	2485.73	< 0.0001	
AB	8.000E-004	1	8.000E-004	1.31	0.2699	
AC	0.66	1	0.66	1070.22	< 0.0001	
AD	0.024	1	0.024	39.51	< 0.0001	
BC	0.018	1	0.018	29.47	< 0.0001	
BD	0.61	1	0.61	996.76	< 0.0001	
CD	1.53	1	1.53	2500.00	< 0.0001	
ABC	2.813E-003	1	2.813E-003	4.59	0.0478	
ABD	0.058	1	0.058	94.37	< 0.0001	
ACD	0.038	1	0.038	61.73	< 0.0001	
BCD	0.51	1	0.51	832.73	< 0.0001	
ABCD	0.063	1	0.063	102.88	< 0.0001	
Pure Error	9.800E-003	16	6.125E-004			
Cor Total	17.74	31				

The residual plots below identify concerns, so a power transformation with lambda of 0.7 was applied to the Pb response.





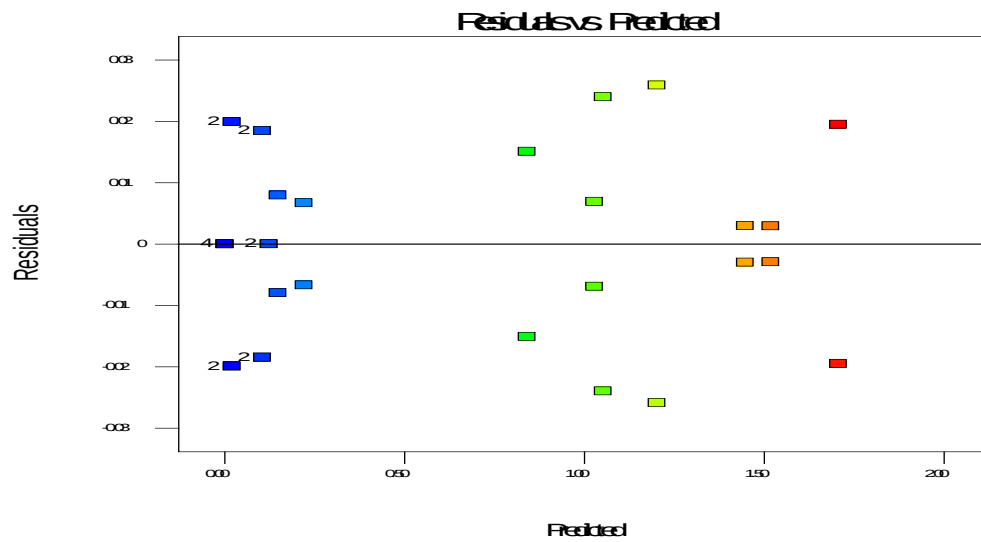
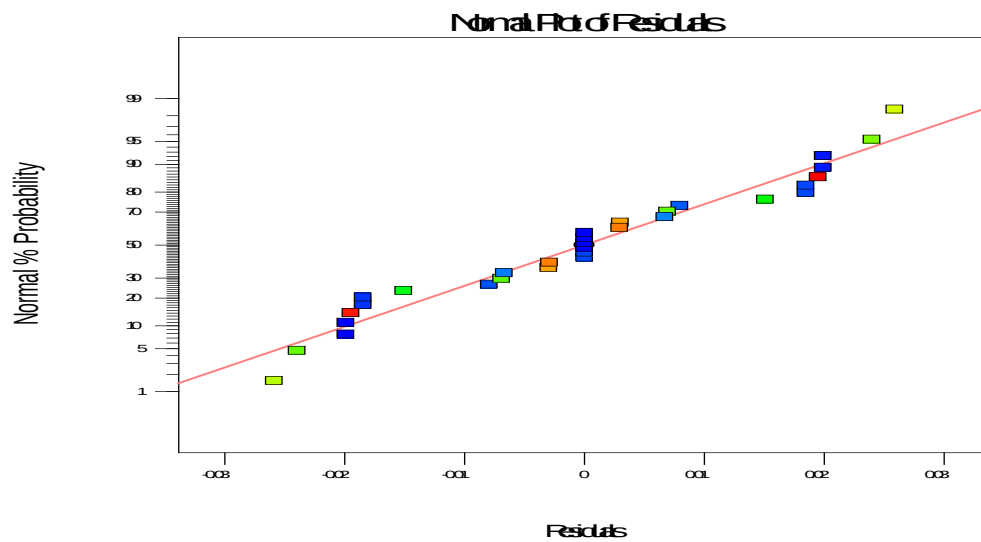


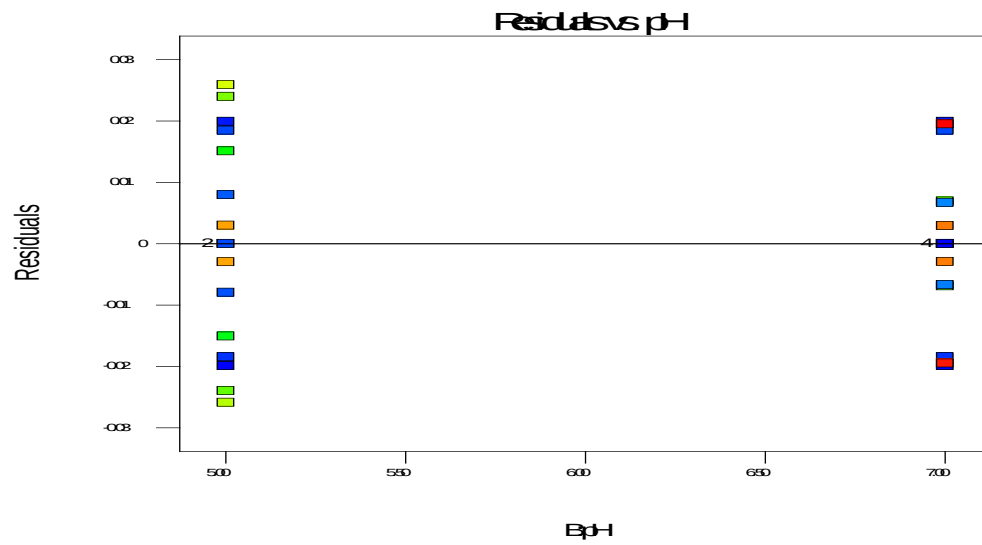
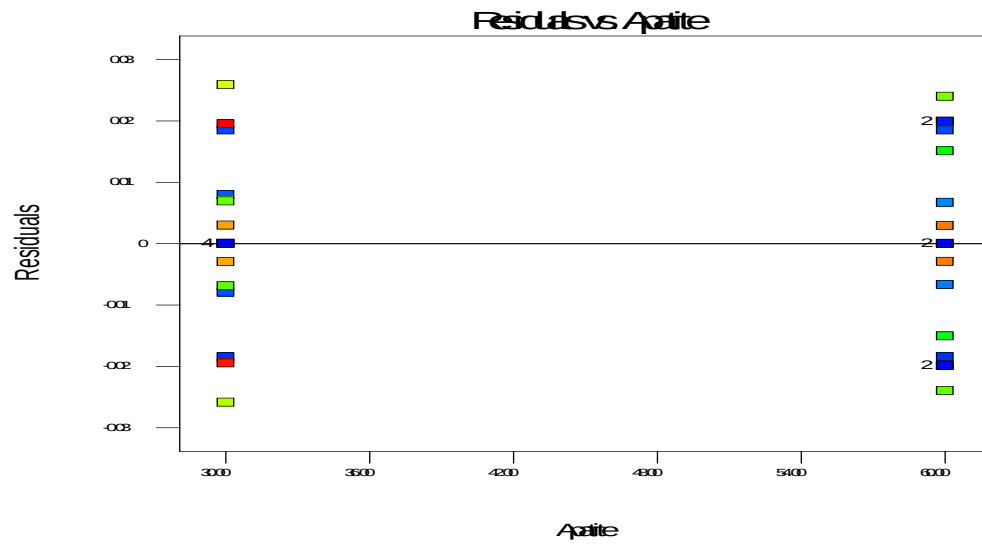
Design Expert Output

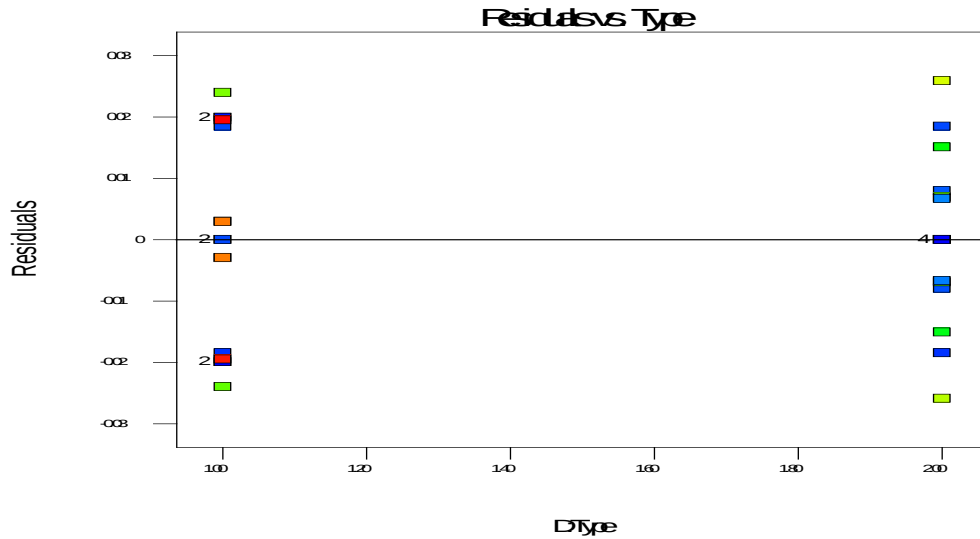
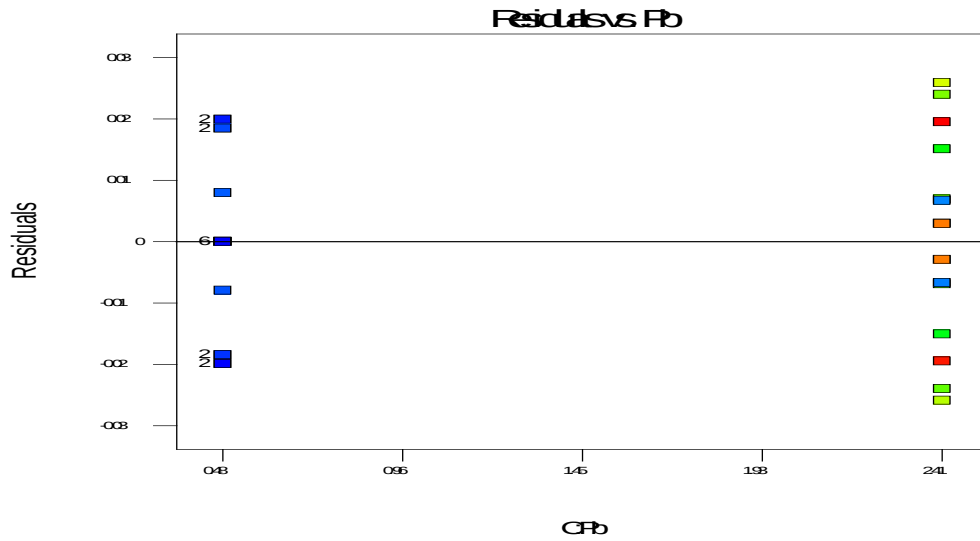
Response	1	Pb Response				
Transform:	Power	Lambda:	0.7	Constant:	0	
ANOVA for selected factorial model						
Analysis of variance table [Partial sum of squares - Type III]						
Source	Sum of Squares	df	Mean Square	F Value	p-value Prob > F	
Model	12.10	15	0.81	1844.15	< 0.0001	significant
A-Apatite	0.49	1	0.49	1123.40	< 0.0001	
B-pH	0.014	1	0.014	32.86	< 0.0001	
C-Pb	9.02	1	9.02	20612.51	< 0.0001	
D-Type	0.75	1	0.75	1712.43	< 0.0001	
AB	3.973E-003	1	3.973E-003	9.08	0.0082	
AC	0.29	1	0.29	661.73	< 0.0001	
AD	0.024	1	0.024	55.70	< 0.0001	
BC	5.077E-003	1	5.077E-003	11.61	0.0036	
BD	0.38	1	0.38	877.65	< 0.0001	
CD	0.73	1	0.73	1670.33	< 0.0001	
ABC	0.011	1	0.011	25.97	0.0001	

ABD	0.049	1	0.049	112.97	< 0.0001
ACD	0.067	1	0.067	152.27	< 0.0001
BCD	0.21	1	0.21	472.40	< 0.0001
ABCD	0.057	1	0.057	131.40	< 0.0001
Pure Error	6.999E-003	16	4.374E-004		
Cor Total	12.11	31			

There are no concerns with the residual plots shown below.







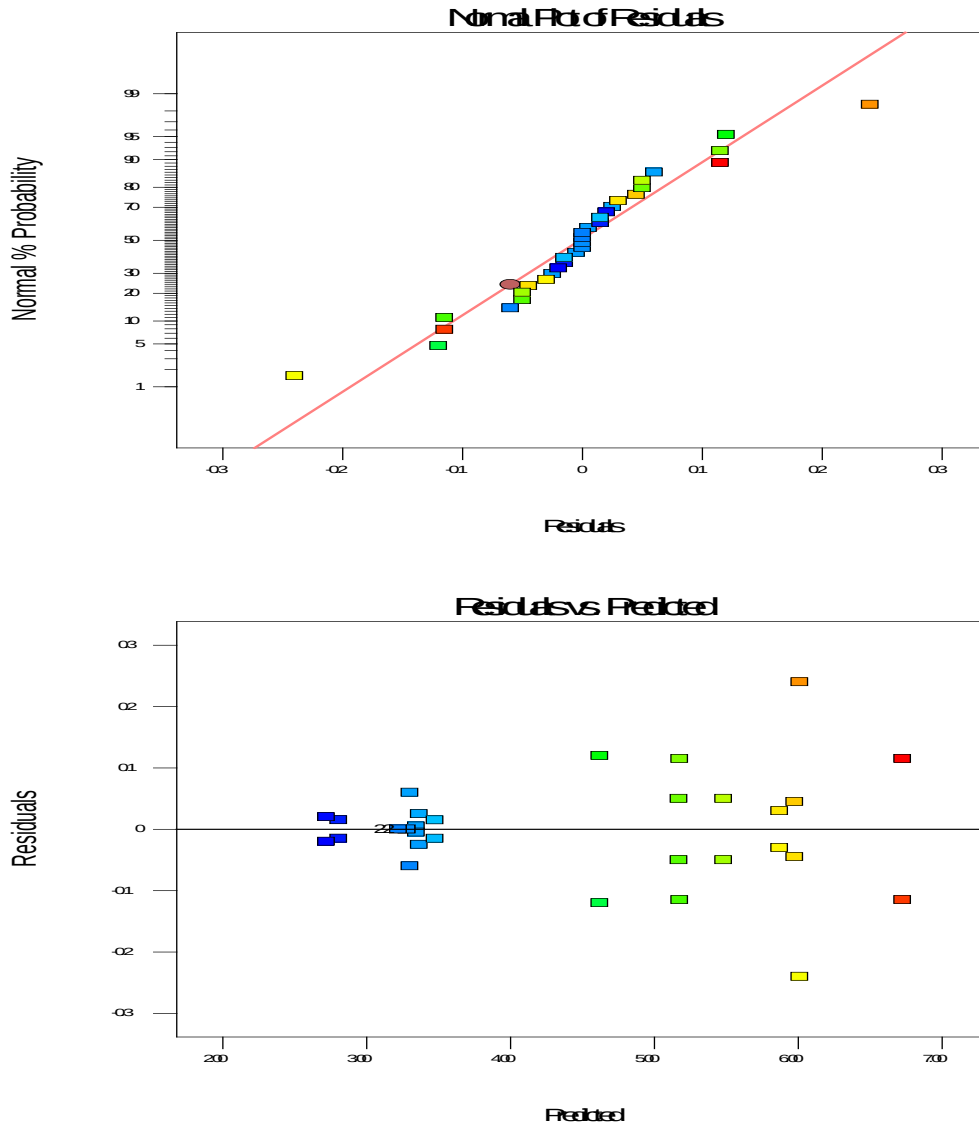
The analysis for the pH response is shown below.

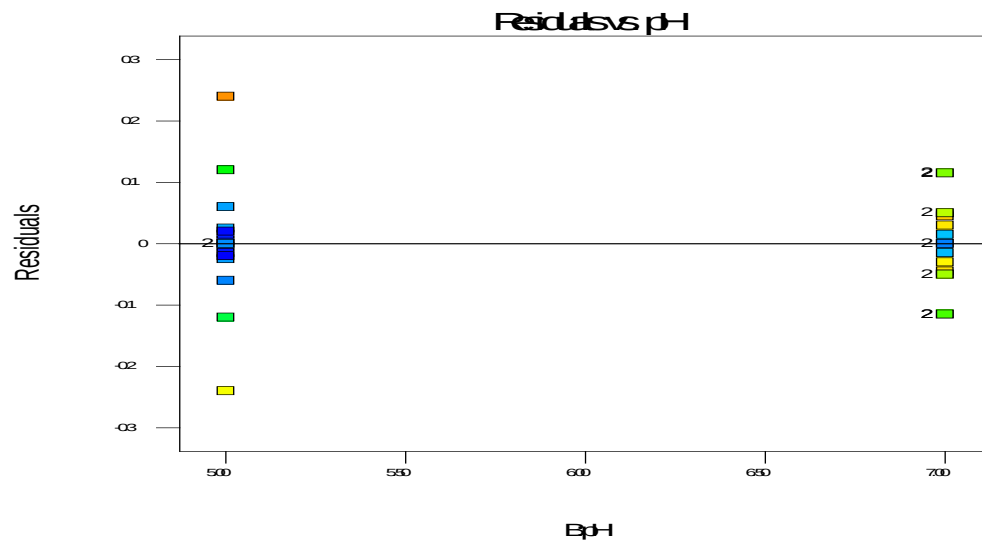
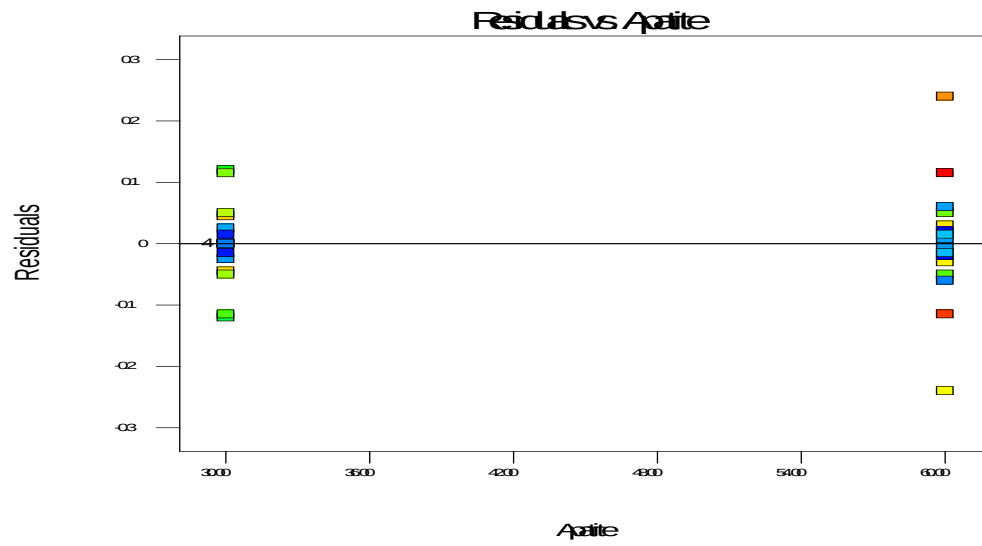
Design Expert Output

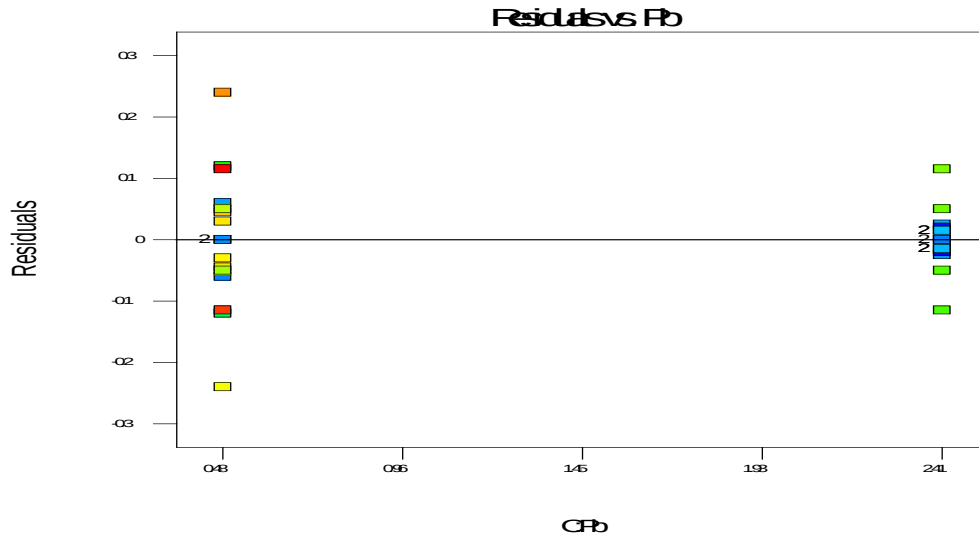
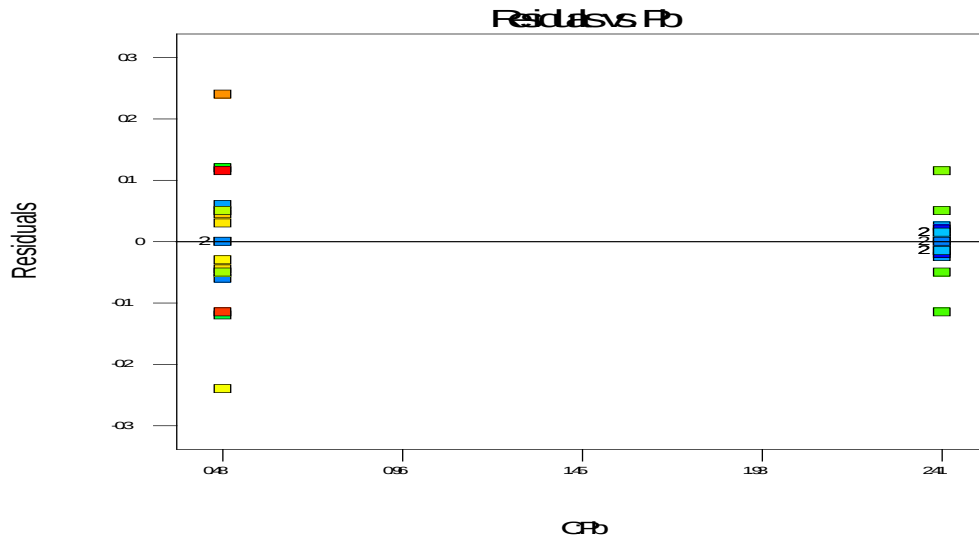
Response 2 pH Response						
ANOVA for selected factorial model						
Analysis of variance table [Partial sum of squares - Type III]						
Source	Sum of Squares	df	Mean Square	F Value	p-value Prob > F	
Model	54.62	15	3.64	261.32	< 0.0001	significant
A-Apatite	0.91	1	0.91	65.15	< 0.0001	
B-pH	16.95	1	16.95	1216.47	< 0.0001	
C-Pb	17.96	1	17.96	1288.54	< 0.0001	
D-Type	13.09	1	13.09	939.72	< 0.0001	
AB	9.031E-004	1	9.031E-004	0.065	0.8023	
AC	0.72	1	0.72	51.89	< 0.0001	
AD	0.29	1	0.29	21.14	0.0003	
BC	0.52	1	0.52	37.15	< 0.0001	
BD	6.903E-003	1	6.903E-003	0.50	0.4916	
CD	0.040	1	0.040	2.86	0.1100	

ABC	0.049	1	0.049	3.50	0.0796
ABD	0.22	1	0.22	15.99	0.0010
ACD	0.46	1	0.46	33.24	< 0.0001
BCD	3.33	1	3.33	239.31	< 0.0001
ABCD	0.059	1	0.059	4.21	0.0569
Pure Error	0.22	16	0.014		
Cor Total	54.84	31			

Nonconstant variance is identified in the residual plots below, so an inverse square root transformation was applied.







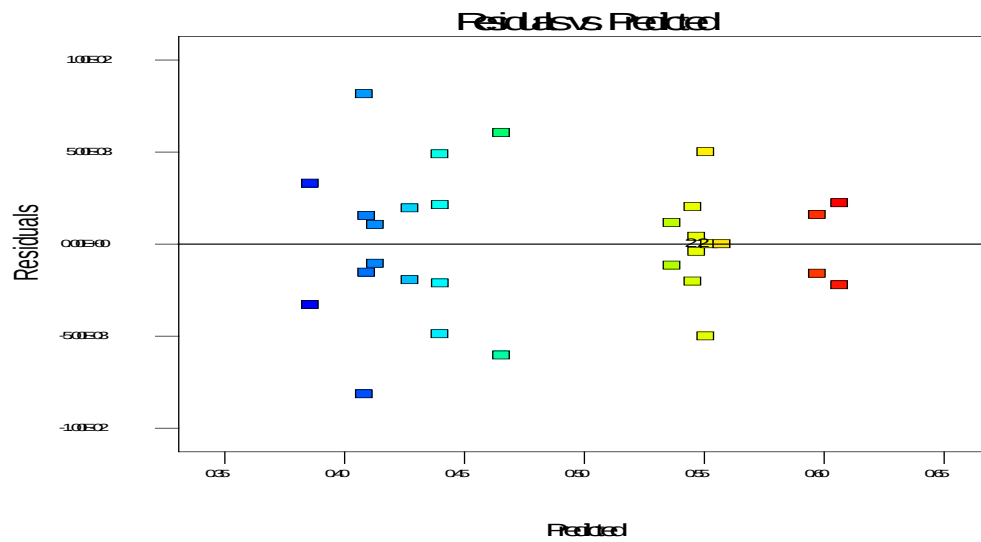
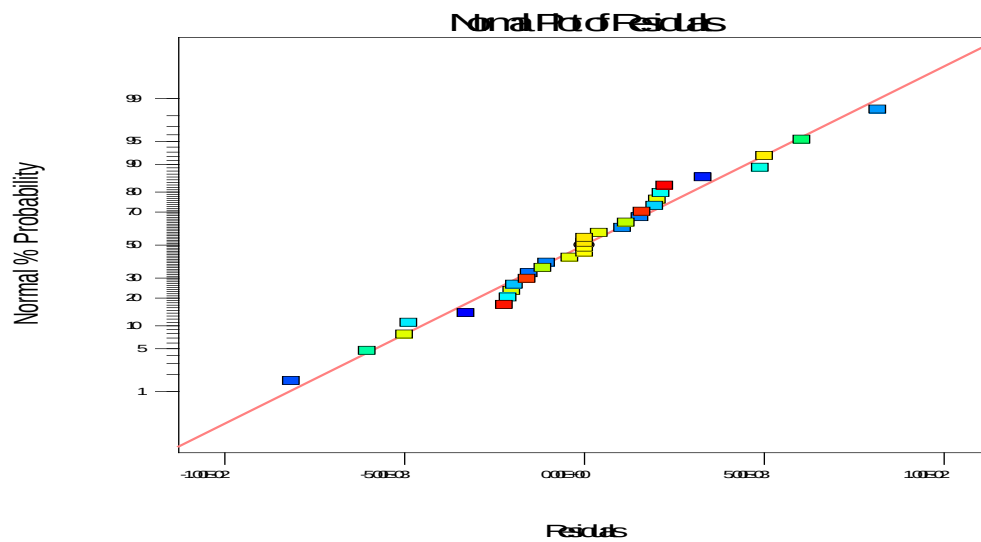
Design Expert Output

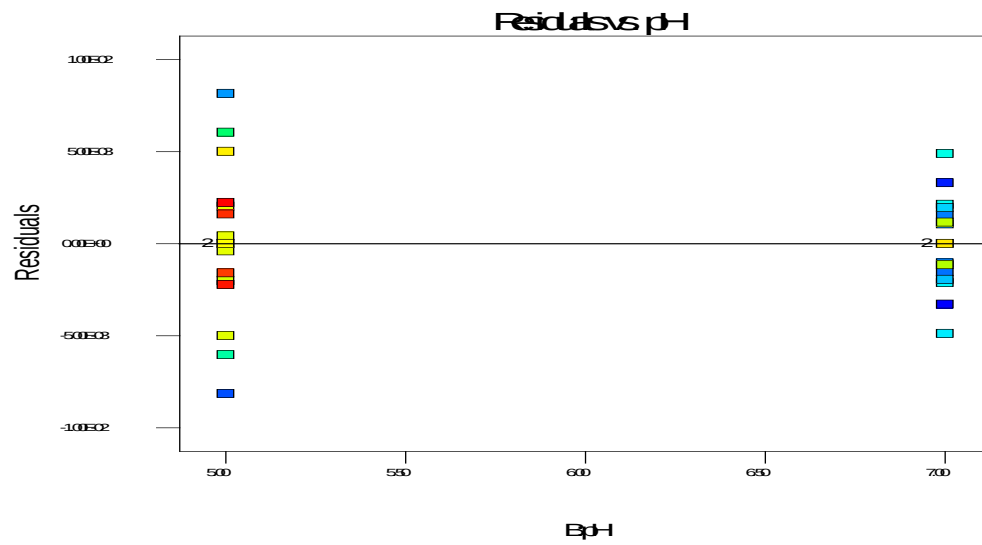
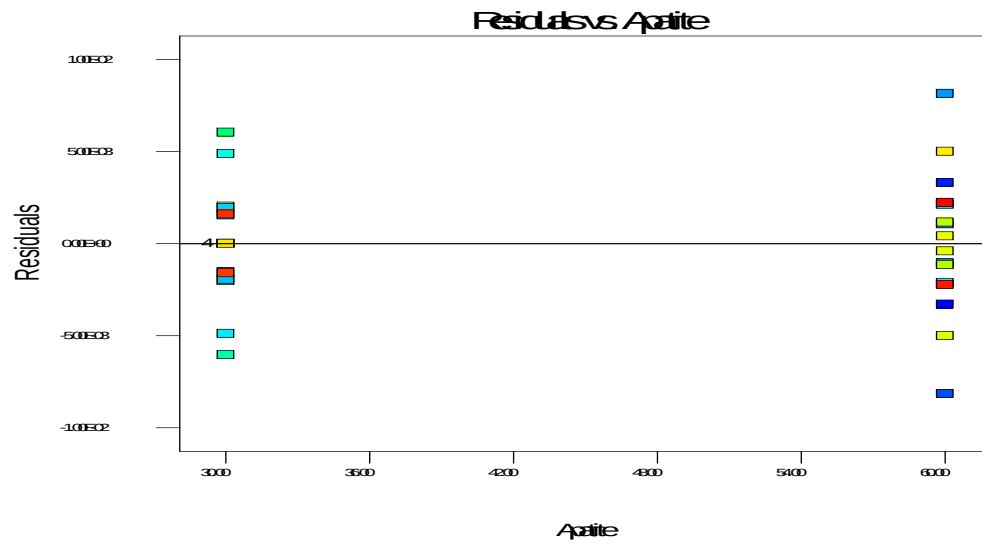
Response	2 pH Response					
Transform:	Inverse Sqrt			Constant:	0	
ANOVA for selected factorial model						
Analysis of variance table [Partial sum of squares - Type III]						
Source	Sum of Squares	df	Mean Square	F Value	p-value Prob > F	
Model	0.17	15	0.011	483.43	< 0.0001	significant
A-Apatite	1.418E-003	1	1.418E-003	60.44	< 0.0001	
B-pH	0.055	1	0.055	2346.06	< 0.0001	
C-Pb	0.054	1	0.054	2303.65	< 0.0001	
D-Type	0.045	1	0.045	1918.39	< 0.0001	
AB	1.417E-005	1	1.417E-005	0.60	0.4484	
AC	9.441E-004	1	9.441E-004	40.23	< 0.0001	
AD	3.287E-004	1	3.287E-004	14.00	0.0018	
BC	4.687E-005	1	4.687E-005	2.00	0.1768	
BD	8.263E-004	1	8.263E-004	35.21	< 0.0001	
CD	3.302E-004	1	3.302E-004	14.07	0.0017	
ABC	3.443E-004	1	3.443E-004	14.67	0.0015	

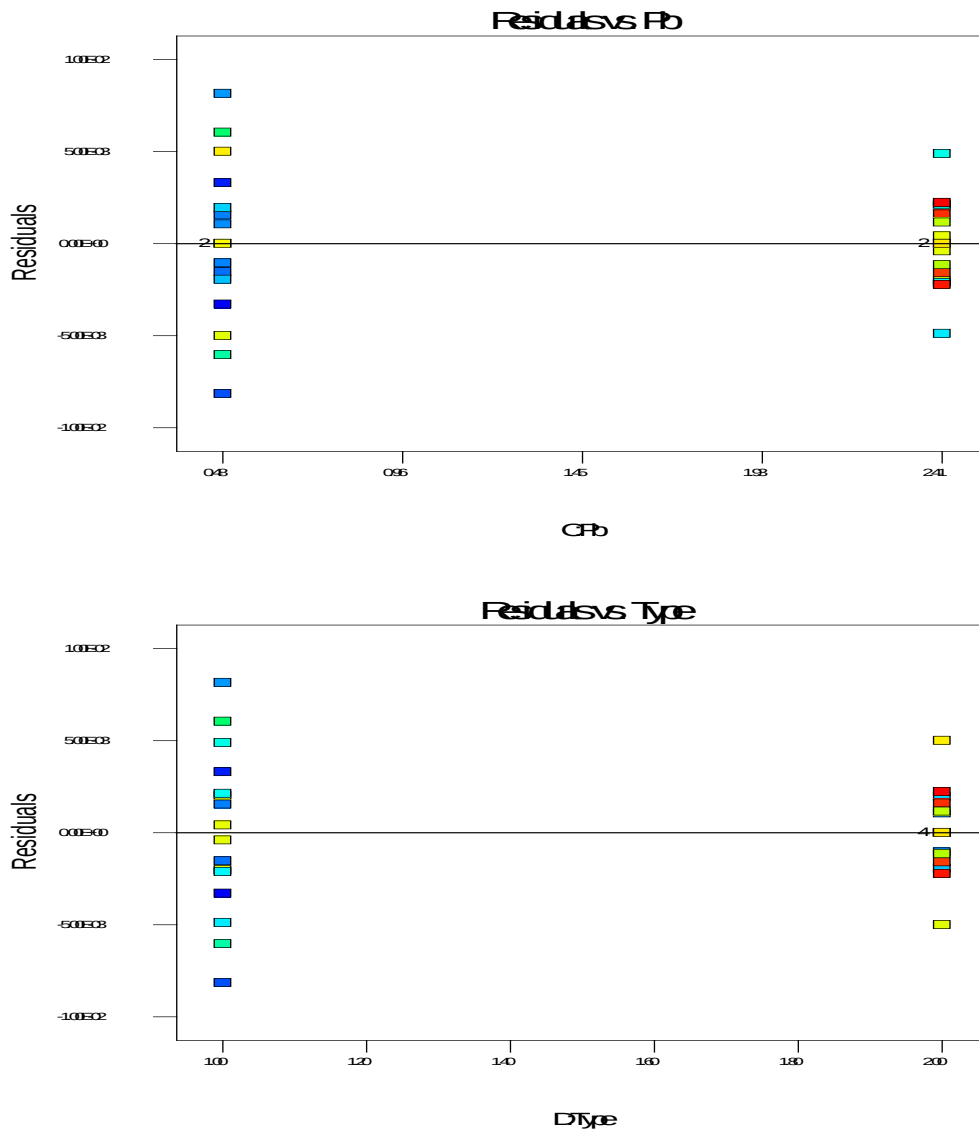
significant

ABD	7.072E-004	1	7.072E-004	30.14	< 0.0001
ACD	7.593E-004	1	7.593E-004	32.36	< 0.0001
BCD	0.010	1	0.010	437.95	< 0.0001
ABCD	4.034E-005	1	4.034E-005	1.72	0.2083
Pure Error	3.755E-004	16	2.347E-005		
Cor Total	0.17	31			

There are no concerns with the residual plots shown below.







In summary, there is a difference between the Fishbone Apatite and the synthetic Hydroxyapatite.

6.43. Often the fitted regression model from a 2^k factorial design is used to make predictions at points of interest in the design space.

- (a) Find the variance of the predicted response $\hat{y}$ at the point $x_1, x_2, \dots, x_k$ in the design space. Hint: Remember that the x 's are coded variables, and assume a 2^k design with an equal number of replicates n at each design point so that the variance of a regression coefficient $\hat{\beta}$ is $\frac{\sigma^2}{n2^k}$ and that the covariance between any pair of regression coefficients is zero.

Let's assume that the model can be written as follows:

$$\hat{y}(\mathbf{x}) = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \dots + \hat{\beta}_p x_p$$

where $\mathbf{x}' = [x_1, x_2, \dots, x_k]$ are the values of the original variables in the design at the point of interest where a prediction is required, and the variables in the model $x_1, x_2, \dots, x_p$ potentially include interaction terms among the original k variables. Now the variance of the predicted response is

$$\begin{aligned} V[\hat{y}(\mathbf{x})] &= V(\hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \dots + \hat{\beta}_p x_p) \\ &= V(\hat{\beta}_0) + V(\hat{\beta}_1 x_1) + V(\hat{\beta}_2 x_2) + \dots + V(\hat{\beta}_p x_p) \\ &= \frac{\sigma^2}{n2^k} \left(1 + \sum_{i=1}^p x_i^2 \right) \end{aligned}$$

This result follows because the design is orthogonal and all model parameter estimates have the same variance. Remember that some of the x 's involved in this equation are potentially interaction terms.

- (b) Use the result of part (a) to find an equation for a $100(1-\alpha)\%$ confidence interval on the true mean response at the point $x_1, x_2, \dots, x_k$ in the design space.

The confidence interval is

$$\hat{y}(\mathbf{x}) - t_{\alpha/2, df_E} \sqrt{V[\hat{y}(\mathbf{x})]} \leq y(\mathbf{x}) \leq \hat{y}(\mathbf{x}) + t_{\alpha/2, df_E} \sqrt{V[\hat{y}(\mathbf{x})]}$$

where df_E is the number of degrees of freedom used to estimate σ^2 and the estimate of σ^2 has been used in computing the variance of the predicted value of the response at the point of interest.

6.44. Hierarchical Models. Several times we have utilized the hierarchy principal in selecting a model; that is, we have included non-significant terms in a model because they were factors involved in significant higher-order terms. Hierarchy is certainly not an absolute principle that must be followed in all cases. To illustrate, consider the model resulting from Problem 6.1, which required that a non-significant main effect be included to achieve hierarchy. Using the data from Problem 6.1:

- (a) Fit both the hierarchical model and the non-hierarchical model.

Design Expert Output for Hierarchical Model

Response: Life in hours						
ANOVA for Selected Factorial Model						
Analysis of variance table [Partial sum of squares]						
Source	Sum of Squares	DF	Mean Square	F Value	Prob > F	
Model	1519.67	4	379.92	12.54	< 0.0001	significant
A	0.67	1	0.67	0.022	0.8836	
B	770.67	1	770.67	25.44	< 0.0001	
C	280.17	1	280.17	9.25	0.0067	
AC	468.17	1	468.17	15.45	0.0009	
Residual	575.67	19	30.30			not significant
Lack of Fit	93.00	3	31.00	1.03	0.4067	
Pure Error	482.67	16	30.17			
Cor Total	2095.33	23				
The Model F-value of 12.54 implies the model is significant. There is only a 0.01% chance that a "Model F-Value" this large could occur due to noise.						
Values of "Prob > F" less than 0.0500 indicate model terms are significant. In this case B, C, AC are significant model terms.						
Std. Dev.	5.50		R-Squared	0.7253		
Mean	40.83		Adj R-Squared	0.6674		

C.V.	13.48	Pred R-Squared	0.5616
PRESS	918.52	Adeq Precision	10.747

The "Pred R-Squared" of 0.5616 is in reasonable agreement with the "Adj R-Squared" of 0.6674. A difference greater than 0.20 between the "Pred R-Squared" and the "Adj R-Squared" indicates a possible problem with your model and/or data.

Design Expert Output for Non-Hierarchical Model

Response: Life in hours**ANOVA for Selected Factorial Model****Analysis of variance table [Partial sum of squares]**

	Sum of Squares	DF	Mean Square	F Value	Prob > F	
Model	1519.00	3	506.33	17.57	< 0.0001	significant
B	770.67	1	770.67	26.74	< 0.0001	
C	280.17	1	280.17	9.72	0.0054	
AC	468.17	1	468.17	16.25	0.0007	
Residual	576.33	20	28.82			
Lack of Fit	93.67	4	23.42	0.78	0.5566	not significant
Pure Error	482.67	16	30.17			
Cor Total	2095.33	23				

The Model F-value of 17.57 implies the model is significant. There is only a 0.01% chance that a "Model F-Value" this large could occur due to noise.

Values of "Prob > F" less than 0.0500 indicate model terms are significant. In this case B, C, AC are significant model terms.

The "Lack of Fit F-value" of 0.78 implies the Lack of Fit is not significant relative to the pure error. There is a 55.66% 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.

Std. Dev.	5.37	R-Squared	0.7249
Mean	40.83	Adj R-Squared	0.6837
C.V.	13.15	Pred R-Squared	0.6039
PRESS	829.92	Adeq Precision	12.320

The "Pred R-Squared" of 0.6039 is in reasonable agreement with the "Adj R-Squared" of 0.6837. A difference greater than 0.20 between the "Pred R-Squared" and the "Adj R-Squared" indicates a possible problem with your model and/or data.

- (b) Calculate the PRESS statistic, the adjusted R^2 and the mean square error for both models.

The PRESS and R^2 are in the Design Expert Output above. The PRESS is smaller for the non-hierarchical model than the hierarchical model suggesting that the non-hierarchical model is a better predictor.

- (c) Find a 95 percent confidence interval on the estimate of the mean response at a cube corner ($x_1 = x_2 = x_3 = \pm 1$). Hint: Use the result of Problem 6.36.

Design Expert Output

	Prediction	SE Mean	95% CI low	95% CI high	SE Pred	95% PI low	95% PI high
Life	27.45	2.18	22.91	31.99	5.79	15.37	39.54
Life	36.17	2.19	31.60	40.74	5.80	24.07	48.26
Life	38.67	2.19	34.10	43.24	5.80	26.57	50.76
Life	47.50	2.19	42.93	52.07	5.80	35.41	59.59
Life	43.00	2.19	38.43	47.57	5.80	30.91	55.09
Life	34.17	2.19	29.60	38.74	5.80	22.07	46.26
Life	54.33	2.19	49.76	58.90	5.80	42.24	66.43
Life	45.50	2.19	40.93	50.07	5.80	33.41	57.59

- (d) Based on the analyses you have conducted, which model would you prefer?

Notice that PRESS is smaller and the adjusted R^2 is larger for the non-hierarchical model. This is an indication that strict adherence to the hierarchy principle isn't always necessary. Note also that the confidence interval is shorter for the non-hierarchical model.

6.45. Suppose that you want to run a 2^3 factorial design. The variance of an individual observation is expected to be about 4. Suppose that you want the length of a 95% confidence interval on any effect to be less than or equal to 1.5. How many replicates of the design do you need to run?

With the equations for the $se(\text{Effect})$ and $100(1-\alpha)$ percent confidence interval on the effects shown below, we can iteratively estimate the number of replicates. From the table of iterations, 14 replicates are required.

$$se(\text{Effect}) = \frac{2S}{\sqrt{n2^k}}$$

$$\text{Effect} \pm t_{(\alpha/2, N-p)} se(\text{Effect})$$

n	se(Effect)	t(0.025,N-p)	95% CI Length
12	0.408	1.987	1.623
13	0.392	1.985	1.557
14	0.378	1.983	1.499
15	0.365	1.981	1.447